The Astrophysical Journal, 938:110 (24pp), 2022 October 20 https://doi.org/10.3847/1538-4357/ac8e04

The Pantheon+ Analysis: Cosmological Constraints

Dillon Brout[1,32], Dan Scolnic[2], Brodie Popovic[2], Adam G. Riess[3,4], Anthony Carr[5], Joe Zuntz[6], Rick Kessler[7,8], Tamara M. Davis[5], Samuel Hinton[5], David Jones[9], W. D'Arcy Kenworthy[4], Erik R. Peterson[2], Khaled Said[5], Georgie Taylor[10], Noor Ali[11], Patrick Armstrong[12], Pranav Charvu[2], Arianna Dwomoh[2], Cole Meldorf[8], Antonella Palmese[13], Helen Qu[14], Benjamin M. Rose[2], Bruno Sanchez[2], Christopher W. Stubbs[1,15], Maria Vincenzi[2], Charlotte M. Wood[16], Peter J. Brown[17,18], Rebecca Chen[2], Ken Chambers[19], David A. Coulter[9], Mi Dai[4], Georgios Dimitriadis[20], Alexei V. Filippenko[21], Ryan J. Foley[9], Saurabh W. Jha[22], Lisa Kelsey[23], Robert P. Kirshner[1,24], Anais Möller[25,26], Jessie Muir[27], Seshadri Nadathur[28], Yen-Chen Pan[29], Armin Rest[3], Cesar Rojas-Bravo[9], Masao Sako[14], Matthew R. Siebert[9], Mat Smith[30], Benjamin E. Stahl[21], and Phil Wiseman[31]

[1] Center for Astrophysics, Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA; dillon.brout@cfa.harvard.edu
[2] Department of Physics, Duke University, Durham, NC, 27708, USA
[3] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[4] Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218 USA
[5] School of Mathematics and Physics, University of Queensland, Brisbane, QLD 4072, Australia
[6] Institute for Astronomy, University of Edinburgh, Edinburgh EH9 3HJ, UK
[7] Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
[8] Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA
[9] Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 92064, USA
[10] Research School of Astronomy and Astrophysics, Australian National University, Canberra, Australia
[11] Ume University, SE-901 87, Ume, Sweden
[12] Mt. Stromlo Observatory, The Research School of Astronomy and Astrophysics, Australian National University, ACT 2601, Australia
[13] Department of Physics, University of California, Berkeley, CA 94720-7300, USA
[14] Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA
[15] Department of Physics, Harvard University, 17 Oxford Street, Cambridge, MA 02138, USA
[16] Department of Physics and Astronomy, University of Notre Dame, Notre Dame, IN 46556, USA
[17] Department of Physics and Astronomy, Texas A&M University, 4242 TAMU, College Station, TX 77843, USA
[18] George P. and Cynthia Woods Mitchell Institute for Fundamental Physics & Astronomy, College Station, TX 77843, USA
[19] Institute of Astronomy, University of Hawaii, 2680 Woodlawn Drive, Honolulu, HI 96822, USA
[20] School of Physics, Trinity College Dublin, The University of Dublin, Dublin 2, Ireland
[21] Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA
[22] Department of Physics and Astronomy, Rutgers, the State University of New Jersey, Piscataway, NJ 08854, USA
[23] Institute of Cosmology and Gravitation, University of Portsmouth, Portsmouth, PO1 3FX, UK
[24] Gordon and Betty Moore Foundation, Palo Alto, CA 94304, USA
[25] Centre for Astrophysics & Supercomputing, Swinburne University of Technology, Victoria 3122, Australia
[26] LPC, Université Clermont Auvergne, CNRS/IN2P3, F-63000 Clermont-Ferrand, France
[27] Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON N2L 2Y5, Canada
[28] Department of Physics & Astronomy, University College London, Gower Street, London, WC1E 6BT, UK
[29] Graduate Institute of Astronomy, National Central University, 32001 Jhongli, Taiwan
[30] Université de Lyon, Université Claude Bernard Lyon 1, CNRS/IN2P3, IP2I Lyon, F-69622, Villeurbanne, France
[31] School of Physics and Astronomy, University of Southampton, Southampton SO17 1BJ, UK
Received 2022 February 8; revised 2022 August 1; accepted 2022 August 24; published 2022 October 19

Abstract

We present constraints on cosmological parameters from the Pantheon+ analysis of 1701 light curves of 1550 distinct Type Ia supernovae (SNe Ia) ranging in redshift from $z = 0.001$ to 2.26. This work features an increased sample size from the addition of multiple cross-calibrated photometric systems of SNe covering an increased redshift span, and improved treatments of systematic uncertainties in comparison to the original Pantheon analysis, which together result in a factor of 2 improvement in cosmological constraining power. For a flat ΛCDM model, we find $\Omega_M = 0.334 \pm 0.018$ from SNe Ia alone. For a flat w_0CDM model, we measure $w_0 = -0.90 \pm 0.14$ from SNe Ia alone, $H_0 = 73.5 \pm 1.1\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when including the Cepheid host distances and covariance (SH0ES), and $w_0 = -0.978^{+0.024}_{-0.031}$ when combining the SN likelihood with Planck constraints from the cosmic microwave background (CMB) and baryon acoustic oscillations (BAO); both w_0 values are consistent with a cosmological constant. We also present the most precise measurements to date on the evolution of dark energy in a flat w_0w_aCDM universe, and measure $w_a = -0.1^{+0.9}_{-2.0}$ from Pantheon+ SNe Ia alone, $H_0 = 73.3 \pm 1.1\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ when including SH0ES Cepheid distances, and $w_a = -0.65^{+0.28}_{-0.32}$ when combining Pantheon+ SNe Ia with CMB and BAO data. Finally, we find that systematic uncertainties in the use of SNe Ia along the distance ladder comprise less than one-third of the total uncertainty in the measurement of H_0

[32] NASA Einstein Fellow.

and cannot explain the present "Hubble tension" between local measurements and early universe predictions from the cosmological model.

Unified Astronomy Thesaurus concepts: Cosmology (343); Dark energy (351); Dark matter (353); Type Ia supernovae (1728); Cosmological models (337); Expanding universe (502)

Supporting material: machine-readable table

1. Introduction

Type Ia supernovae (SNe Ia) anchor the standard model of cosmology with their unmatched ability to map the past 10 billion years of expansion history. SNe Ia provided the first evidence of the accelerating expansion of the universe (Riess et al. 1998; Perlmutter et al. 1999), and they remain invaluable because they are (1) bright enough to be seen at large cosmic distances, (2) common enough to be found in large numbers, and (3) can be standardized to ∼0.1 mag precision in brightness or ∼5% in distance per object.

Statistical leverage from large samples of SNe Ia has grown rapidly over the last three decades, and well-calibrated and standardized compilations of these samples have facilitated measurements of the *relative* expansion history across the redshift range $0 < z < 1$ characterized by the equation-of-state parameter of dark energy ($w = P/(\rho c^2)$), and the measurement of the Hubble constant H_0, the current expansion rate determined from *absolute* distances. Measurements of w are constrained from the comparison of standardized SN Ia magnitudes over a wide range of redshifts obtained from different surveys with different observing-depth strategies. Measurements of H_0 require very nearby (<50 Mpc, ∼1 discovered per year) SNe Ia found by multiple surveys in galaxies that host calibrated primary distance indicators (e.g., Cepheids, tip of the red giant branch, TRGB), which are then compared to SNe in the Hubble flow, often from the same surveys.

However, simply combining several subsamples into a large sample of SNe Ia does not provide meaningful gains without rigorous cross-calibration, self-consistent analysis of their light curves and redshifts, and characterization of their numerous sources of related uncertainties or covariance. As samples and compilations grow, ever greater attention must be paid to the control of systematic uncertainties, which would otherwise dominate sample uncertainties.

This analysis, Pantheon+, is the successor to the original Pantheon analysis (Scolnic et al. 2018b) and builds on the analysis framework of the original Pantheon to combine an even larger number of SN Ia samples and include those that are in galaxies with measured Cepheid distances in order to be able to simultaneously constrain parameters describing the full expansion history (e.g., Ω_M, w_0, and w_a) with the local expansion rate (H_0). The original Pantheon compilation of 1048 SNe Ia was used to measure a value (from SNe Ia alone) of $w = -1.090 \pm 0.220$. Riess et al. (2016), in their measurement of the local expansion rate H_0, used a prerelease version of Pantheon based on Scolnic et al. (2015) and further augmented the sample as Pantheon did not extend to reach the low redshifts of the primary distance indicators at $z < 0.01$.

Although there was significant overlap in data and analysis between the Pantheon measurement of w and the Riess et al. (2016) measurement of H_0, the Riess et al. (2016) measurement included several Cepheid-calibrator SNe Ia that were not included in Pantheon, and the fitting for H_0 and parameters describing the expansion history were done independently rather than simultaneously. Dhawan et al. (2020) later established a framework for considering the covariance between SNe in primary distance indicator hosts and SNe in the Hubble flow. We build on that framework, which was developed originally for a redshift-binned Hubble diagram, and in this paper we create the first unbinned sample with covariance extending down to $z = 0.001$ that can be used to propagate correlated systematics for simultaneous measurements of H_0, Ω_M, w_0, and w_a. We (i) analyze the largest set of cosmologically viable SN Ia light curves to date, (ii) include low-redshift samples to extend the lower bound in redshift to 0.001, which contains the primary distance indicators (SNe in SH0ES Cepheid host galaxies), (iii) propagate systematic uncertainties for both primary distance indicators and higher-redshift SNe simultaneously, and (iv) leverage the large strides made in the field of SN Ia cosmology since the original Pantheon.

This paper is the culmination of a series of papers that comprise the Pantheon+ analysis. A graphic of an overview of the numerous Pantheon+ supporting analyses, on which this paper heavily relies, is shown in Figure 1. Details of each paper pertinent to this analysis are described in Section 3. One of these papers, Scolnic et al. (2022, hereafter S22), describes the sample of 1701 cosmologically viable SN Ia light curves of 1550 distinct SNe, which we will refer to as "the Pantheon+ sample." The redshifts and peculiar velocities of the SNe used here are given by Carr et al. (2021), and a comprehensive analysis of peculiar velocities is presented by Peterson et al. (2022). The cross-calibration of the different photometric systems used in this analysis can be found in Brout et al. (2022, hereafter Fragilistic), and calibration-related systematic uncertainty limits are determined by Brownsberger et al. (2021). The underlying SN Ia populations describing the data set are given by Popovic et al. (2021b). The model for intrinsic brightness variations was developed by Brout & Scolnic (2021) and then improved and evaluated by Popovic et al. (2021a). The novel systematic framework for simultaneous measurement of H_0 and cosmology was developed by Dhawan et al. (2020), and an improved methodology for systematic uncertainties is described by Brout et al. (2021).

In this work we discuss briefly the aforementioned papers in the context of their use in this analysis, evaluate several additional systematic uncertainties not addressed in these works, measure cosmological parameters, examine additional signals in the Hubble diagram, and compile systematic uncertainty budgets on cosmological parameters. A companion paper by the SH0ES Team (Riess et al. 2022, hereafter R22) combines from this work 277 Hubble flow ($0.023 < z < 0.15$) SNe Ia and 42 SNe Ia in Cepheid-calibrator hosts, their relative distances, and their covariance, with the absolute distances of primary distance anchors (Cepheids and TRGB) from R22 in order to measure H_0 under the assumption of flat ΛCDM. Similarly, in this work we utilize the full Pantheon+ sample of 1550 SNe Ia in combination with the R22 Cepheid host distances to show the impact of cosmological models with

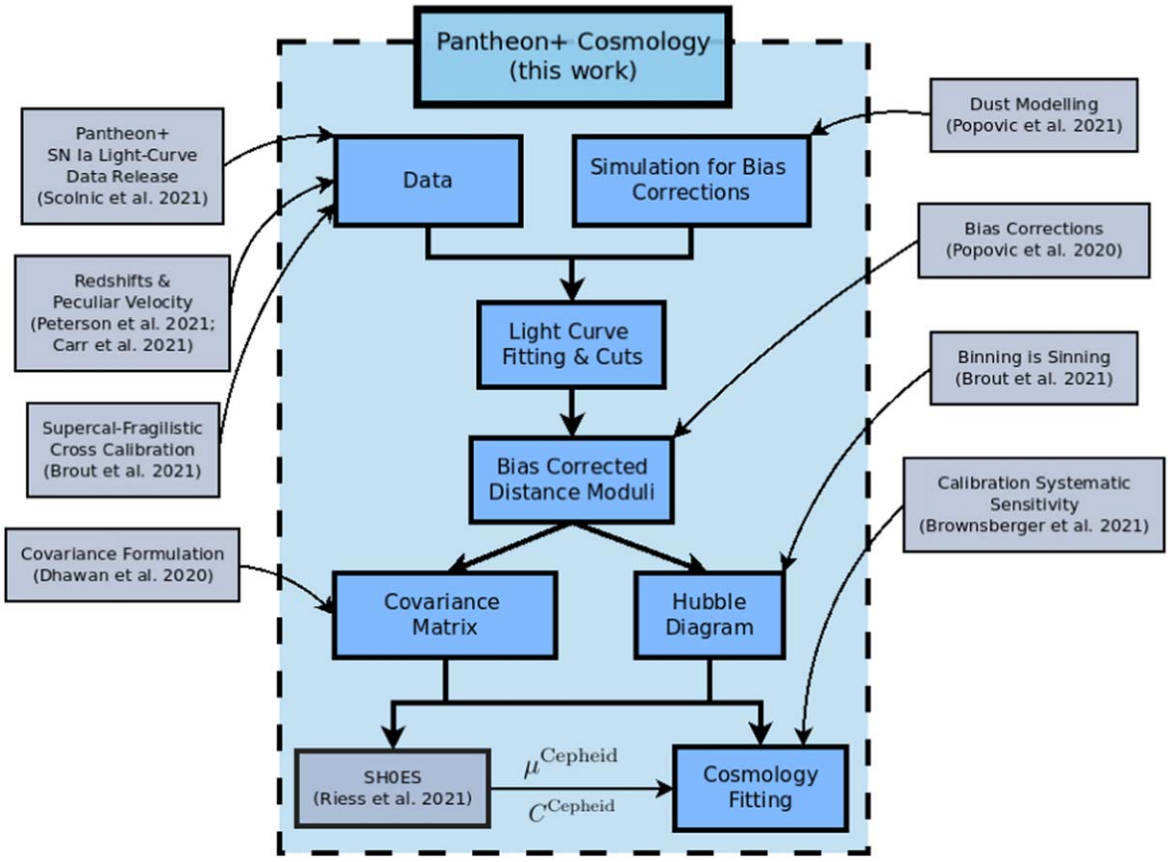

Figure 1. Analysis roadmap of this work and supporting/complementary Pantheon+ and SH0ES papers. Components of the analysis here are shown in blue. The companion paper (Riess et al. 2022, hereafter R22), which provides a constraint on H_0, requires the Hubble diagram and covariance computed in this work. Likewise, measurements of H_0 in this work require the R22 Cepheid distance and covariance. Supporting papers are shown in gray boxes.

more freedom than those used in R22 as well as the impact of SN-related systematic uncertainties on inferences of H_0.

An important aspect of this work is the public release of the data and simulations used here that allow for the reproduction of multiple different stages of this analysis. In Appendix C, we present the numerous products that will be made available, including SN distances, redshifts, uncertainties, covariance, and extensive SNANA simulations (Kessler et al. 2009) of the data that model astrophysical effects, cosmological effects, and the observation/telescope effects of each survey down to the level of cadence, weather history, etc. We encourage the community to validate alternate analyses of the publicly released Pantheon+ sample on these simulations.

The structure of the paper is as follows. In Section 2, we describe the methodology from fitting SN light curves to constraining cosmological parameters. Section 3 summarizes all of the inputs to the analysis including the data sample, calibration, and redshifts. In Section 4, we describe the cosmological results. Sections 5 and 6 are our discussions and conclusions, respectively.

2. Methodology of Constraining Cosmological Parameters with SNe Ia

2.1. Measuring Distances to SNe Ia

To standardize the SN Ia brightnesses, we fit light curves using SNANA with the SALT2 model as originally developed by Guy et al. (2010) and updated in Brout et al. (2022), hereafter SALT2-B22. For each SN, the SALT2 light-curve fit returns four parameters: the light-curve amplitude x_0 where $m_B \equiv -2.5\log_{10}(x_0)$; x_1, the stretch parameter corresponding to light-curve width; c, the light-curve color, which includes contributions from both intrinsic color and dust; and t_0, the time of peak brightness. Extinction due to Milky Way dust is accounted for in the SALT2 light-curve fitting. From the parameters m_B, x_1, and c, we standardize the SN brightnesses and infer distance moduli (μ), used in the Hubble diagram, with a modified version of the Tripp (1998) distance estimator. Following Kessler & Scolnic (2017, hereafter BBC), the distance modulus is defined as

$$\mu = m_B + \alpha x_1 - \beta c - M - \delta_{\rm bias} + \delta_{\rm host}, \quad (1)$$

where α and β are global nuisance parameters relating stretch and color, respectively, to luminosity. M is the fiducial magnitude of an SN Ia, which can be calibrated by setting an absolute distance scale with primary distance anchors such as Cepheids. $\delta_{\rm bias}$ is a correction term[33] to account for selection biases that is determined from simulations following Popovic et al. (2021b), described in detail in Appendix A. $\delta_{\rm host}$ is the luminosity correction (step) for residual correlations between the standardized brightness of an SN Ia and the host-galaxy

[33] Past analyses have the opposite sign $+\delta_{\rm bias}$; however, since the values of $\delta_{\rm bias}$ in the public release are meant to be subtracted, we change the sign compared to previous works.

mass,

$$\delta_{\rm host} = \gamma \times (1 + e^{(M_\star - S)/\tau_{M\star}})^{-1} - \frac{\gamma}{2}, \quad (2)$$

where γ is the magnitude of the SN Ia luminosity differences between SNe in high ($M_\star > 10^{10} M_\odot$) and low ($M_\star < 10^{10} M_\odot$) stellar-mass galaxies and where "hostless" SNe have been assumed to reside in galaxies with low stellar mass. $M_\star$ is the inferred stellar mass measured in units of solar mass ($M_\odot$) from spectral energy distribution fitting to the photometry of each host galaxy, S is the step location (nominal analysis assumes $S = 10^{10} M_\odot$), and $\tau_{M\star}$ describes the width of the step.

The total distance-modulus error, σ_μ, for SN i is described as

$$\sigma^2_{\mu,i} = f(z_i, c_i, M_{\star,i})\sigma^2_{{\rm meas},i} + \sigma^2_{\rm floor}(z_i, c_i, M_{\star,i}) + \sigma^2_{{\rm lens},i} + \sigma^2_{z,i} + \sigma^2_{{\rm vpec},i}, \quad (3)$$

where $\sigma_{\rm meas}$ is the measurement uncertainty of SALT2 light-curve fit parameters and their associated covariances (see Equation (3) of Kessler & Scolnic 2017) resulting from photometric uncertainties. The measurement uncertainty is scaled by $f(z_i, c_i, M_{\star,i})$ specific to each survey in order to account for selection effects that can reduce the observed scatter at the limits of each sample. The uncertainty contribution from gravitational lensing as given by Jönsson et al. (2010) is $\sigma_{\rm lens} = 0.055\, z$. We note that, as discussed by Kessler et al. (2019a), the correct lensing distribution is utilized in simulations. The nominal distance-modulus uncertainty contribution due to the combination of redshift-measurement uncertainty (σ_z) and peculiar-velocity uncertainty ($\sigma_{\rm vpec}$) have both been converted to distance-modulus uncertainty under the assumption of a cosmological model. Chen et al. (2022) noted that the optimal way to characterize redshift-measurement uncertainty at high redshifts (e.g., the Dark Energy Survey, DES, sample, $z > 0.3$) is to float the redshift and use the uncertainty in redshift as a prior in the light-curve fit. However, following previous analyses, we fix the redshift and include the associated distance-modulus uncertainty σ_z in Equation (3), which is a correct estimate at low redshifts ($z < 0.1$). Lastly, $\sigma_{\rm floor}$ represents the floor in standardizability owing to intrinsic unmodeled variations in SNe Ia such that

$$\sigma^2_{\rm floor}(z_i, c_i, M_{\star,i}) = \sigma^2_{\rm scat}(z_i, c_i, M_{\star,i}) + \sigma^2_{\rm gray}, \quad (4)$$

where $\sigma^2_{\rm scat}(z_i, c_i, M_{\star,i})$ is determined from a model that describes intrinsic brightness fluctuations, and $\sigma^2_{\rm gray}$ is a single number representing a gray (color-independent) floor in standardizability for all SNe Ia; $\sigma^2_{\rm gray}$ is determined after the BBC fitting process in order to bring the Hubble diagram reduced χ^2 to unity. The details of $\sigma^2_{\rm scat}(z_i, c_i, M_{\star,i})$, its model dependence, and its contribution to systematic uncertainties are discussed in further detail in Section 3.3.2 and Appendix A.

To determine the distance-modulus values of all of the SNe, we follow the BBC fitting process with updates to increase the dimensionality of bias corrections in Popovic et al. (2021b). The likelihood (as given in Equation (6) of Kessler & Scolnic 2017) results in a cosmology-independent minimization of the free parameters (α, β, γ, and $\sigma_{\rm gray}$) that minimize the scatter in the Hubble diagram. While the BBC process was designed for utility for photometric cosmology analyses and uses SN Ia classification probabilities, the data analyzed here are a spectroscopically confirmed SN Ia sample, and therefore we set the non-Ia SN probabilities to zero for the whole sample.

2.2. The Covariance Matrix

Following Conley et al. (2011), we compute covariance matrices $C_{\rm stat}$ and $C_{\rm syst}$ to account for statistical and systematic uncertainties and expected correlations between the SN Ia light curves in the sample when analyzing cosmological models. BBC produced both a redshift-binned and an unbinned Hubble diagram, enabling both binned and unbinned covariance matrices. For the original Pantheon (Scolnic et al. 2018b), Joint Light-Curve Analysis (JLA; Betoule et al. 2014), and DES3YR (Brout et al. 2019b), $C_{\rm stat}$ and $C_{\rm syst}$ were redshift-binned matrices (or smoothed as a function of redshift) citing computational limitations. Following Brout et al. (2021), in this work we utilize the unbinned Hubble diagrams to create unbinned covariance matrices. The Pantheon+ sample (Scolnic et al. 2022) also includes "duplicate SNe Ia," SNe Ia that have been observed simultaneously by numerous different surveys, so that statistical covariance $C_{\rm stat}$ is computed as

$$C_{\rm stat}(i,j) = \begin{cases} \sigma^2_\mu & i = j \\ \sigma^2_{\rm floor} + \sigma^2_{\rm lens} + & \\ \sigma^2_z + \sigma^2_{\rm vpec} & i \neq j \text{ and } {\rm SN}_i = {\rm SN}_j \end{cases}, \quad (5)$$

where each row of the matrix corresponds to an SN light curve, the diagonal of $C_{\rm stat}$ is the full distance error (σ^2_μ) of the ith light curve, and where measurement noise from components other than the light curve itself is included as off-diagonal covariance between entries corresponding to light curves of the same SN (${\rm SN}_i = {\rm SN}_j$) observed by two different surveys.

Systematic uncertainties can manifest in three key places in the analysis: (1) from changing aspects affecting the light-curve fitting (e.g., survey photometry, calibration, SALT2 model), (2) from changing redshifts that propagate to changes in distance moduli relative to a cosmological model, and (3) from changes in the astrophysical or survey-dependent assumptions in the simulations used for bias corrections. For each of these categories, we examine all of the known significant sources of systematic uncertainty (ψ) with sizes S_ψ, which result in residuals in the Hubble diagram relative to our baseline analysis ($\mu_{\rm BASE}$). In order to compute the effect of systematics, we first define

$$\Delta\mu^i_\psi \equiv \mu^i_\psi - \mu^i_{\rm BASE} - (\mu_{\rm ref}(z_\psi) - \mu_{\rm ref}(z_{\rm BASE})), \quad (6)$$

where μ^i_ψ is the set of distances for systematic ψ. For systematics that affect redshift, we have included a new methodology in Equation (6) that utilizes a reference cosmological model distance $\mu_{\rm ref}(z)$ corresponding to flat ΛCDM ($\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$). The $\mu_{\rm mod}(z_\psi)$ and $\mu_{\rm mod}(z_{\rm BASE})$ are the cosmological model distances corresponding to redshifts z_ψ and $z_{\rm BASE}$. In order to propagate redshift effects into a distance $\times$ distance covariance matrix, the additional component $\mu_{\rm mod}(z_\psi) - \mu_{\rm mod}(z_{\rm BASE})$ accounts for the difference in inferred model distance.

Assuming linearity between $\Delta\mu_\psi$ and ψ, we compute the derivative for each ψ in order to build a 1701×1701

systematic covariance matrix as

$$C_{\rm syst}^{ij} = \sum_{\psi} \frac{\partial \Delta\mu_{\psi}^{i}}{\partial S_{\psi}} \frac{\partial \Delta\mu_{\psi}^{j}}{\partial S_{\psi}} \sigma_{\psi}^{2}, \quad (7)$$

which denotes the covariance between the ith and jth light-curve fit summed over the different sources of systematic uncertainty (ψ) with uncertainty σ_{ψ} (see Section 3 for details). As shown by Brout et al. (2021), the σ_{ψ} serve as priors on the known size of systematic uncertainties, but the data itself can constrain the impact of each systematic under the condition that information has not been collapsed by binning/smoothing (as was done for the original Pantheon, JLA, and DES3YR).

Fluctuations of the sample of light curves that pass the sample quality cuts (Table 2 of S22) for different systematics result in an ill-defined covariance matrix. To have a well-defined unbinned covariance matrix requires a subtle treatment in order to ensure that the sample is consistent in both the light-curve fitting and BBC stages across all systematics in the analysis. Quality cuts at the light-curve stage are only applied to the set of SNe based on their values found in the baseline analysis, and this SN sample is used for all systematic tests. We perform the BBC process twice—the first iteration to identify the subset of <1% of SNe for which bias corrections are unable to be computed, and a second iteration using only the common set of SNe that have valid bias corrections in all systematic variants. The final cosmology sample of 1701 light curves that satisfy all criteria is described in detail in S22 (see the "Systematics" row in Table 2 of S22).

Finally, the statistical and systematic covariance matrices are combined and used to constrain cosmological models:

$$C_{\rm stat+syst} = C_{\rm stat} + C_{\rm syst}. \quad (8)$$

2.3. Cosmology

Constraining cosmological models with SN data using χ^2 has been used in previous SN Ia cosmology analyses (e.g., Riess et al. 1998; Astier et al. 2006) and first included systematic covariance in Conley et al. (2011). Here we follow closely the formalism of Conley et al. (2011) where cosmological parameters are constrained by minimizing a χ^2 likelihood:

$$-2\ln(\mathcal{L}) = \chi^2 = \Delta \boldsymbol{D}^T \, C_{\rm stat+syst}^{-1} \, \Delta \boldsymbol{D}, \quad (9)$$

where $\boldsymbol{D}$ is the vector of 1701 SN distance-modulus residuals computed as

$$\Delta D_i = \mu_i - \mu_{\rm model}(z_i), \quad (10)$$

and each SN distance (μ_i) is compared to the predicted model distance given the measured SN/host redshift ($\mu_{\rm model}(z_i)$). The model distances are defined as

$$\mu_{\rm model}(z_i) = 5\log(d_L(z_i)/10\ {\rm pc}), \quad (11)$$

where d_L is the model-based luminosity distance that includes the parameters describing the expansion history $H(z)$. For a flat cosmology ($\Omega_k = 0$), the luminosity distance is described by

$$d_L(z) = (1+z)c \int_0^z \frac{dz'}{H(z')}, \quad (12)$$

where $d_L(z)$ is calculated at each step of the cosmological fitting process, and the parameterization of the expansion history (used in Equation (12) and therefore in the likelihood Equation (9)) in this work is defined as

$$H(z) = H_0 \sqrt{\Omega_M (1+z)^3 + \Omega_\Lambda (1+z)^{3(1+w)}}. \quad (13)$$

See Hogg (1999) for the forms of the expansion history $H(z)$ used in the case that the assumption of flatness is relaxed.

The parameters M (Equation (1)) and H_0 (Equation (13)) are degenerate when analyzing SNe alone. However, we also present constraints that include the recently released SH0ES Cepheid host distance anchors (R22) in the likelihood that facilitates constraints on both M and H_0.

When utilizing SH0ES Cepheid host distances, the SN distance residuals are modified as follows:

$$\Delta D'_i = \begin{cases} \mu_i - \mu_i^{\rm Cepheid} & i \in \text{Cepheid hosts} \\ \mu_i - \mu_{\rm model}(z_i) & \text{otherwise}, \end{cases} \quad (14)$$

where $\mu_i^{\rm Cepheid}$ is the Cepheid calibrated host-galaxy distance provided by SH0ES, and where $\mu_i - \mu_i^{\rm Cepheid}$ is sensitive to the parameters M and H_0 and is largely insensitive to Ω_M or w. We also include the SH0ES Cepheid host-distance covariance matrix ($C_{\rm stat+syst}^{\rm Cepheid}$) presented by R22 such that the likelihood becomes

$$-2\ln(\mathcal{L}') = \Delta \boldsymbol{D}'^T \, (C_{\rm stat+syst}^{\rm SN} + C_{\rm stat+syst}^{\rm Cepheid})^{-1} \, \Delta \boldsymbol{D}', \quad (15)$$

where $C_{\rm stat+syst}^{\rm SN}$ denotes the SN covariance.

We evaluate the likelihoods with the PolyChord (Handley et al. 2015) sampler in the CosmoSIS package (Zuntz et al. 2015) using 250 live points, 30 repeats, and an evidence tolerance requirement of 0.1. This resulted in converged chains containing 1000–3000 independent samples. We verified the SN-only results with CosmoMC (Lewis & Bridle 2002) and with the fast cosmology grid-search program in `SNANA`. The likelihood for Pantheon+ and R22 Cepheid host distance samples will be made available in the public version of CosmoSIS. In this work we also utilize the additional public likelihoods in CosmoSIS in order to combine with and assess agreement with external cosmological probes: Planck (Planck Collaboration et al. 2020) and baryon acoustic oscillations (BAOs; likelihoods discussed in Section 4).

In this work we investigate four cosmological parameterizations:

1. Flat ΛCDM: Ω_M is floated, and we fix $w = -1$ and $\Omega_M + \Omega_\Lambda = 1$.
2. ΛCDM: Ω_M and Ω_Λ are floated, and we fix $w = -1$.
3. Flat wCDM: w and Ω_M are floated, and we fix $\Omega_M + \Omega_\Lambda = 1$.
4. Flat $w_0 w_a$CDM: $w = w_0 + w_a(1+z)$, Ω_M, w_0, w_a are floated, and we fix $\Omega_M + \Omega_\Lambda = 1$.

We blind our analysis in two ways simultaneously. First, we blind the binned distance residuals output by the BBC fit, as cosmological parameters could be inferred visually from simply looking at the Hubble diagram. Second, in order to prevent accidental viewing of the cosmological parameters themselves, the CosmoSIS chains were shifted by unknown values following the formalism of Hinton (2016).

3. Data and Analysis Inputs

Here we review each component of the data set and analysis. We discuss the fundamental *purpose*, the *baseline* treatment in

this analysis, and the *systematic* uncertainties associated with each aspect (if applicable). The impact of systematics in both distance and cosmological inference is shown in Section 4. We provide a brief overview of this section here.

Data

Section 3.1.1: SN Ia Light Curves
Section 3.1.2: Redshifts
Section 3.1.3: Peculiar Velocities
Section 3.1.4: Host-galaxy Properties

Calibration and Light-curve Fitting

Section 3.2.1: Calibration
Section 3.2.2: SALT2 Model
Section 3.2.3: Milky Way Extinction

Simulations

Section 3.3.1: Survey Modeling
Section 3.3.2: Intrinsic Scatter Models
Section 3.3.3: Uncertainty Modeling
Section 3.3.4: Validation

3.1. Data

3.1.1. SN Ia Light Curves

Purpose: The flux-calibrated light-curve photometry is fit to determine the SALT2 parameters used in standardization (Equation (1)).

Baseline: The light-curve data is described in detail by S22 and references therein. The full set of spectroscopically classified photometric light curves is compiled from 18 different publicly available and privately released samples. In total, 2077 SN light-curve fits converged using SALT2; after quality cuts are applied (Table 2 of S22), this results in 1701 SN light curves of 1550 unique SNe Ia usable for cosmological constraints. The sample includes a 3.5σ Hubble residual outlier cut to remove five potential contaminants that are likely nonnormal Type Ia or misidentified redshifts. The sample of cosmologically viable light curves includes 81 light curves of 42 SNe used to calibrate Cepheid brightnesses as utilized by R22. The survey SN photometry compiled in Scolnic et al. (2022) and analyzed here is from DES[34] (Brout et al. 2019a; Smith et al. 2020a), Foundation[1] (Foley et al. 2018), Pan-STARRS (PS1; Scolnic et al. 2018b), Supernova Legacy Survey (SNLS; Betoule et al. 2014), Sloan Digital Sky Survey (SDSS; Sako et al. 2011), Hubble Space Telescope (HST; Gilliland et al. 1999; Riess et al. 2001, 2004, 2007; Suzuki et al. 2012; Riess et al. 2018), and Low-z (grouped together as LOSS_1[1]; Ganeshalingam et al. 2010; LOSS_2[1]; Stahl et al. 2019; SOUSA[135]; Brown et al. 2014; CNIa0.02[1]; Chen et al. 2020; CSP; Krisciunas et al. 2017b; CfA1; Riess et al. 1999; CfA2; Jha et al. 2006; CfA3; Hicken et al. 2009; CfA4; Hicken et al. 2012; and numerous smaller low-redshift samples[1] of one to two SNe given by Milne et al. 2010; Stritzinger et al. 2010; Tsvetkov & Elenin 2010; Zhang et al. 2010; Krisciunas et al. 2017a; Burns et al. 2018; Gall et al. 2018; Burns et al. 2020; Kawabata et al. 2020.)

Systematics: See Calibration (Section 3.2.1).

[34] Not included in Pantheon 2018.
[35] https://pbrown801.github.io/SOUSA/

3.1.2. Redshifts

Purpose: The peculiar-velocity corrected cosmic microwave background (CMB) frame redshift of each SN/host is required to compare the inferred distance to a distance predicted by a cosmological model, as given in Equation (10). Additionally, heliocentric redshifts are required in the SALT2 light-curve fits in order to shift the model spectrum to match the data.

Baseline: The redshifts for all of the SNe (and their host galaxies, depending on what is available) are provided by Carr et al. (2021), who performed a comprehensive review of redshifts for the Pantheon+ samples and made numerous corrections. Carr et al. (2021) reported the heliocentric redshifts for each SN and converted the redshift into the CMB frame. The redshifts of the Pantheon+ sample cover a range of $0.001 < z < 2.3$. While redshifts of the 42 Cepheid host calibrator SNe are included, they are not used in the comparison of SN Ia magnitudes to the Cepheid distance scale and are only provided for reference and for SALT2 fitting.

Systematics: Following Carr et al. (2021), we apply a coherent shift to each redshift of $+4 \times 10^{-5}$. This was conservatively stated by Calcino & Davis (2017) for the potential size of a local void bias and by Davis et al. (2019) as a potential measurement bias.

3.1.3. Peculiar Velocities

Purpose: Peculiar motions of galaxies arise from coherent flows, motion of halos, inflow into clusters or superclusters, and intragroup motion. Corrections are applied to the observed redshifts (after light-curve fitting) based on peculiar-velocity maps derived from independent large spectroscopic galaxy surveys.

Baseline: The nominal peculiar velocities used for this analysis were determined by Peterson et al. (2022) from a comparison of multiple treatments of peculiar-velocity maps and group catalogs. Corrections were applied by Carr et al. (2021) for the Pantheon+ sample. The baseline corrections are based on 2M++ (Carrick et al. 2015) with gl obal parameters found in Said et al. (2020) and combined with group velocities estimated from Tully (2015) group assignments. The $\sigma_{\rm vpec}$ in Equation (3) is found using $240\,{\rm km\,s^{-1}}$ after accounting for uncertainties propagated into the covariance matrix described below. This $\sigma_{\rm vpec}$ floor is in agreement with what was used in Peterson et al. (2022), and for the SNe between $0.001 < z < 0.02$, it is likely a conservative estimate, as Kenworthy et al. (2022) found a floor of $155 \pm 25\,{\rm km\,s^{-1}}$ for the most nearby SN calibrators. This apparent reduction at the lowest redshifts may be due to the peculiar-velocity maps having higher fidelity at these redshifts and because Pantheon+ has relatively better virial-group information at these redshifts.

Systematics: Peterson et al. (2022) discuss multiple viable alternatives for the treatment of peculiar velocities. The first approach is to use the 2M++ corrections (Carrick et al. 2015) integrating over the line-of-sight relation (iLOS) between distance and the measured redshift. We take this variation as the first systematic with $\sigma_\psi^2 = 0.5$. The second approach is to use the Two-Micron All-Sky Redshift Survey (2MRS; Lilow & Nusser 2021) peculiar-velocity map; however, differences between 2MRS and 2M++ at very low redshift ($z < 0.01$) cause numerical stability issues for off-diagonal $C_{\rm syst}$ elements. We incorporate only the diagonal differences between 2MRS and 2M ++ into $C_{\rm syst}$ with $\sigma_\psi^2 = 0.5$. As a numerically stable estimate of

the
off-diagonal terms, we use the 2M++ velocities transformed by the slope and offset difference between the 2M++ and 2MRS maps found in Peterson et al. (2022). The two approaches added in quadrature result in an effective $\sigma_\psi^2 = 1.0$.

3.1.4. Host-galaxy Properties

Purpose: The observed host-galaxy mass versus SN luminosity relation is used to standardize the SN Ia brightnesses in two ways. First, simulations of the data set include correlations between SN color and SN stretch and host properties such as dust as a function of host mass, following Popovic et al. (2021a). Second, a further residual correction is applied in the Tripp Equation (1) where the "mass step" γ is fit in the BBC stage.

Baseline: The host-galaxy stellar masses are presented by S22 and references therein. Masses are determined for all host galaxies, and star formation rates and morphologies are also included the low-z sample. In the baseline analysis, we apply the mass step at $10^{10} M_\odot$ following Pantheon and DES3YR.

Systematics: Several independent analyses (Sullivan et al. 2010; Childress et al. 2013a; Kelsey et al. 2021) have suggested that the optimal location of the mass step could range between $10^{9.8} M_\odot$ and $10^{10.2} M_\odot$. We therefore include a systematic uncertainty where the mass step occurs at $10^{10.2} M_\odot$.

3.2. Calibration and Light-curve Fitting

3.2.1. Calibration

Purpose: Photometric calibration of each passband in each survey is needed to fit light curves and facilitate comparison of the brightnesses of SNe across different telescopes/instruments/filters. Photometric calibration is also important to homogenize spectrophotometric data sets used in the SALT2 model training.

Baseline: The calibration of all 25 photometric systems used in this work is discussed in Brout et al. (2022). The outputs of Fragilistic are a best-fit calibration solution for each of the 105 passbands and a joint 105×105 covariance matrix that describes the covariance between the zero-point calibrations of all passbands that arise from using a single common stellar catalog to tie all surveys together (PS1).

Systematics: The systematics due to calibration and their impact are discussed in detail in Fragilistic. We estimate the impact of the correlated filter zero-point and central wavelength uncertainties by refitting SALT2 light curves (with retrained SALT2 models; see Section 3.2.2) using nine realizations of the 105 zero-points. For each of the nine realizations, a value of $\sigma_\psi^2 = 1/9$ is adopted such that they add in quadrature to $\sim$1. The uncertainty in modeling the spectrum of the HST primary standard star C26202 has been tripled to account for the recent update in Bohlin et al. (2020); it is now set to 15 mmag over 7000 Å ($\sigma_\psi = 3$ for a systematic of 5 mmag over 7000 Å). Lastly, an additional conservative systematic is included only for the CSP SNe to account for the 2% recalibration in CSP tertiary stellar magnitudes from Stritzinger et al. (2010) to Krisciunas et al. (2017b; $\sigma_\psi = 1$).

3.2.2. SALT2 Model

Purpose: The trained SALT2 model is required to fit light curves and determine the light-curve parameters (m_b, c, and x_1) for each SN used in Equation (1).

Baseline: We use the Fragilistic calibration solution and newly trained SALT2-B22 model,[36] which was developed following the formalism of Guy et al. (2010) and Taylor et al. (2021). The SALT2 model includes a component of training statistical uncertainty, which is incorporated in the fitted light-curve parameters.

Systematics: For each of the nine correlated realizations of Fragilistic filter zero-points and central wavelengths discussed above (for calibration), we simultaneously retrain the SALT2 model. Additionally, to conservatively account for a possible systematic from the redevelopment of the SALT2 model-training process itself, we adopt an additional systematic by fitting the data set with the SALT2 model trained by Betoule et al. (2014) and applying a scaling of $\sigma_\psi = 1/3$ (see Section 5 and Figure 15 for impact).

3.2.3. Milky Way Extinction

Purpose: Values of the Milky Way (MW) dust extinction, $E(B-V)_{\rm MW}$, are applied to the SALT2 model spectra during both the model-training process and during the data light-curve fitting process. The "extinction curve" describes the relation between the amount of reddening and extinction as a function of wavelength.

Baseline: We account for MW extinction using maps from Schlegel et al. (1998), with a scale of 0.86 following Schlafly et al. (2010). We assume the MW extinction curve from Fitzpatrick (1999) with $R_V = 3.1$.

Systematics: Similarly to Pantheon, we adopt a global 5% uncertainty scaling of $E(B-V)_{\rm MW}$ based on the fact that Schlafly & Finkbeiner (2011), in a reanalysis of Schlafly et al. (2010), derive smaller values of reddening by 4%, despite using a very similar SDSS footprint ($\sigma_\psi = 1$). While Schlafly & Finkbeiner (2011) found that their results prefer the Fitzpatrick (1999) extinction curve, we conservatively include an additional systematic uncertainty in the MW extinction curve and analyze the data (training and light-curve fit) using the extinction curve from Cardelli et al. (1989) and apply a systematic scaling of $\sigma_\psi = 1/3$, as this reflects the preference of Fitzpatrick (1999) over Cardelli et al. (1989).

3.3. Simulations

3.3.1. Survey Modeling

Purpose: We utilize catalog-level simulations of large samples of SN Ia (>1,000,000 per survey) light curves. SNANA simulations specific to each survey in our analysis are prescribed by each aspect of acquiring an SN Ia sample. As detailed in Figure 1 of Kessler et al. (2019a), the simulations require three main sets of inputs:

A source model for generating SNe with realistic astrophysical properties and applying cosmological effects such as redshifting, dimming, lensing, peculiar velocities, and MW extinction.

A noise model, unique to each survey, for applying instrumental and atmospheric noise to determine a detection efficiency ("DETEFF").

A trigger model, unique to each survey, that includes the observing cadence and describes an efficiency as a function of B-band peak magnitude for detecting SNe and obtaining a spectroscopic confirmation ("SPECEFF").

[36] Released publicly at pantheonplussh0es.github.io.

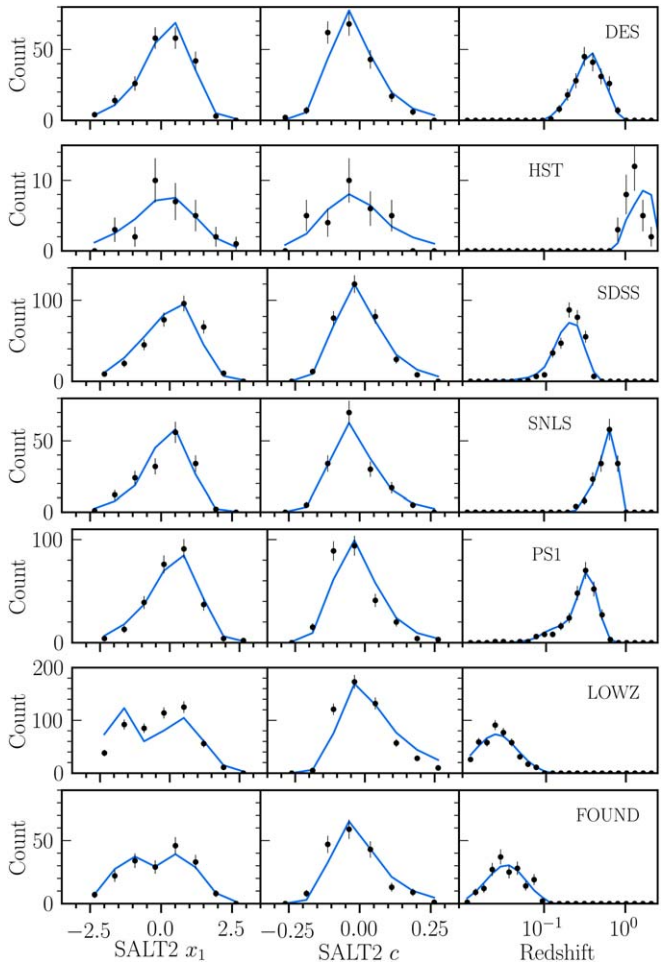

Figure 2. Comparison between observed data (black points) and simulations (blue lines) for the largest subsamples in this analysis: DES, HST, SDSS, SNLS, PS1, LOW-z, and Foundation (FOUND). We compare three key distributions: the SALT2 light-curve-fit parameters x_1 and c are shown as well as the measured redshift.

These simulations for each survey are combined and used to forward model the underlying populations of the SN properties (Popovic et al. 2021a, 2021b) and to determine the expected biases in measured SN distances that follow from the known selection effects. These biases are corrected in the $\delta_{\rm bias}$ term of Equation (1).

Baseline: Depicted in Figure 2 are the distributions of the key observables (z, x_1, and c) for both data and simulations of each survey used in this analysis. We find good agreement between the data and simulations, as described in detail by Popovic et al. (2021a) and Popovic et al. (2021b). We note that the agreement in the redshift dimension is achieved despite not explicitly tuning the redshift distribution of surveys.

We simulate SNe in LOW-z and Foundation down to $z = 0.001$. Novel for this work specifically are the simulations of primary distance indicator hosts of SNe in the range $0.001 < z < 0.01$, which are assumed to have the same color and stretch populations as those of their respective surveys (LOW-z and FOUND), and specifically over this redshift range, they are assumed to be complete with flat spectroscopic selection efficiency. These simulations facilitate bias corrections to the Cepheid-calibrator SNe and thus the propagation of modeling systematics to the SNe used in the companion SH0ES analysis (R22).

The simulation inputs for survey cadence, DETEFF, and SPECEFF functions have been evaluated in many analyses over the past decade. Table 1 shows a summary of where we obtain these inputs for each survey. Survey metadata is used to model the cadence and instrumental properties, if available, such as for FOUND, SDSS, PS1, DES, and SNLS. LOW-z data do not provide such metadata, and thus the cadence and noise properties are extracted from the data as described in Kessler et al. (2019a) following the procedure developed by Scolnic et al. (2018b), which assumes that the LOW-z subset of SNe is magnitude-limited. These are simulations of the CfA and CSP samples, but not of the newer samples included in this work (LOSS, SOUSA, and CNIa0.02), thereby implicitly assuming that the CfA and CSP samples have similar selection effects and therefore distance biases as the newer additions. To simulate SN-host correlations, a catalog of host-galaxy properties and specifically their stellar-mass distributions is taken from Popovic et al. (2021b). The simulations used for bias corrections for all surveys are performed in ΛCDM ($w = -1.0$, $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$) with the SALT2-B22 model.

Systematics: We increase the S/N of each simulation by 20%, resulting in all survey simulated distributions changing by more than 1σ, as a single conservative systematic in the determination of the selection biases. Kessler & Scolnic (2017) showed that the sensitivity of the bias corrections to the input cosmology is relatively weak; this was confirmed by Brout et al. (2019b) and found to be a negligible contribution to SN Ia uncertainty budgets. We therefore do not include this as an additional systematic.

3.3.2. Intrinsic Scatter Models

Purpose: A model of the intrinsic SN brightness variations, called "intrinsic scatter," is needed to account for the observed residual variation in SN Ia standardized luminosities that exceeds expectations from measurement uncertainties alone. In addition, models of the true ("parent") populations of SN Ia SALT2 parameters c and x_1 are required for the source model in `SNANA`. The intrinsic scatter model is utilized in the bias-correction simulations.

Baseline: We utilize the Brout & Scolnic (2021, hereafter BS21) model that prescribes SN Ia scatter into two color-dependent components: (i) a standard cosmological color law specific to SNe Ia and (ii) additional dust-based color laws and dust extinctions that vary with each galaxy/SN. This approach is preferred because of its novel replication of the observed relationships between SN color and residual Hubble diagram scatter as well as its ability to replicate the "mass step" as a function of SN Ia color. We use the scatter model parameters from BS21 with improvements from Popovic et al. (2021a) in our baseline bias-correction simulations; because the BS21 model includes within it the parent c population, we also utilize the separate parent population for x_1 derived by Popovic et al. (2021b). Improving upon Scolnic & Kessler (2016), Popovic et al. (2021b) fit for parent populations in bins of mass to account for host–SN Ia relationships. Popovic et al. (2021b) split their populations into high- and low-redshift groups, and notably for low-redshift surveys, the x_1 populations are fitted with a two-Gaussian model to recreate the observed double peak in the x_1 distribution.

Systematics: We include two categories of systematics for the intrinsic scatter model and parent populations: (1) different models of intrinsic scatter, and (2) determination

Table 1
References for Inputs to SNANA Simulations Used for This Analysis

Survey	Cadence	DETEFF	SPECEFF
LOW-z	Scolnic et al. (2018b)	Kessler et al. (2019b)	Scolnic et al. (2018b)
FOUND	Jones et al. (2019)	N/A	Jones et al. (2019)
SDSS	Kessler et al. (2013)	Kessler et al. (2009)	Popovic et al. (2021b)
PS1	Jones et al. (2018a)	Jones et al. (2018b)	Scolnic et al. (2018b)
DES	Smith et al. (2020b)	Kessler et al. (2015)	Abbott et al. (2019)
SNLS	Kessler et al. (2013)	N/A	Popovic et al. (2021b)
HST	Scolnic et al. (2018b)	N/A	N/A

Note. We give references for the "Cadence," which describes the observing history; the "DETEFF," which describes the detection efficiency based on the signal-to-noise ratio (S/N); and the "SPECEFF," which describes the spectroscopic selection efficiency as a function of SN magnitude.

of parameters for the BS21 model. For the former, we use two additional scatter models from Kessler et al. (2013) that have been used in previous cosmology analyses (JLA, Pantheon, and DES3YR). These are (1) the "G10" model based on Guy et al. (2010), which describes ∼70% of the excess Hubble scatter from "gray" variations and the remaining scatter from wavelength-dependent variations, and (2) the "C11" model based on Chotard (2011), which describes ∼30% of the excess Hubble scatter from coherent variations, and the remaining scatter from wavelength-dependent variations. For the G10 and C11 scatter models, bias corrections are performed in seven dimensions as given by Popovic et al. (2021b). For the systematic uncertainty in the determination of the BS21 model parameters, we adopt three different viable sets of dust and intrinsic SN populations from Popovic et al. (2021a). These populations are the best-fit (maximum likelihood) parameters (hereafter P21), the mean posterior set of parameters, and a set that represents a 1σ fluctuation in the uncertainty. Lastly, while the BS21 and P21 models impact the simulated bias corrections, the SALT2 training and light-curve fitting has not been altered. The choice of scatter model is propagated through the simulations used for the bias corrections applied in Equation (1) and for the uncertainty modeling in $\sigma_{\rm scat}$ of Equation (4).

3.3.3. Distance-modulus Uncertainty Modeling

Purpose: To match the reported SN distance-modulus uncertainties (Equation (3)) to the scatter in distance that is observed in the data.

Baseline: The BS21 model parameters have been fit to the observed scatter in the data set. We can utilize large BS21 simulations to determine $\sigma_{\rm scat}(z,\ c,\ M_\star)$ after accounting for selection effects. The efficacy of this method is shown in Figure 3, which demonstrates good agreement between the observed rms of the Hubble residuals and the uncertainties of the distance-modulus values.

Systematics: To conservatively account for how SN cosmology was done in the past (JLA and Pantheon), in Equation (3) we set $\sigma_{\rm scat}(z,\ c,\ M_\star) = 0$ and allow only a single $\sigma_{\rm gray}$ parameter to replicate the methodology used with historic intrinsic scatter models (G10 and C11). However, we note that for G10 and C11, the trends in rms seen for the data in Figure 3 do not match the reported uncertainties.

3.3.4. Validation

Purpose: To verify that our analysis can recover input values in data-sized simulated samples and does not produce biases. Such tests are sensitive to the light-curve fitting and BBC technique (as well as implementation and coding errors); however, they are not sensitive to certain aspects of the analysis such as the assumption of the SALT2 model or photometric calibration.

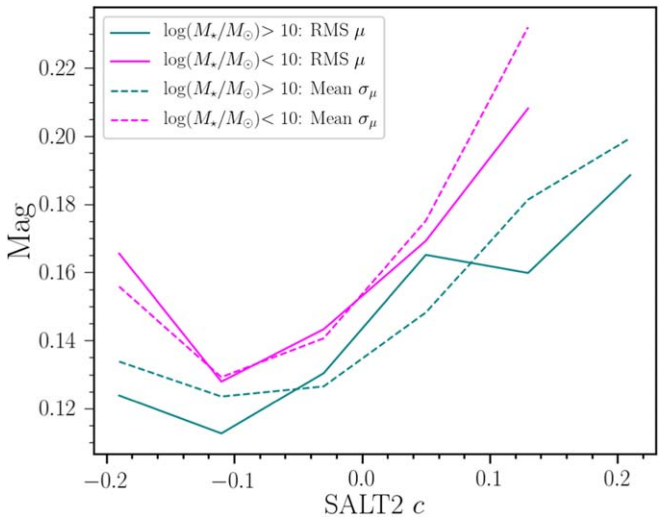

Figure 3. Pantheon+ distance-modulus uncertainties (shown as the dashed lines with mean σ_μ and split on host mass) in comparison to the observed rms of the distance-modulus residuals (shown as solid lines as rms μ split on host mass), as a function of color. This shows that the distance errors are adequately modeled (Equation (4)) as a function of SN color and host stellar mass. In previous analyses, the uncertainties were roughly flat as a function of color.

Baseline: We perform an end-to-end test of our baseline analysis pipeline from survey photometry catalog-level simulations. We create 20 realizations of each survey in an arbitrary cosmological model ($w = -1$): 10 with the BS21 scatter model and 10 with the G10 scatter model. We perform light-curve fitting, apply bias corrections, compile into 10 Hubble diagrams, and maximize the cosmological likelihoods (Equation (9)) using a fast cosmology grid-search program in SNANA (Kessler et al. 2009), with approximate priors from CMB measurements (Planck Collaboration et al. 2020) to obtain best-fit cosmological parameters and uncertainties. For the BS21 model simulations, we recover a mean best-fit $w = -1.012 \pm 0.011$, and for the G10 model simulations, we recover a mean best-fit $w = -0.983 \pm 0.015$; both are within $\sim 1\sigma$ of the input cosmology. The 20 realizations are made available publicly[37] along with bias-correction simulations.

[37] Will be made available after publication at pantheonplussh0es.github.io.

Table 2
Standardization Parameters and Results

Model	BBC Fit				CosmoSIS Fit	
	α	β	$\sigma_{\rm gray}$	γ	rms	$ln(\mathcal{L})$
BS21	0.148(4)	3.09(4)	0.00	−0.003(7)	0.171	−1635
P21	0.145(5)	3.00(5)	0.00	0.019(10)	0.171	−1674
G10	0.153(4)	2.98(5)	0.10	0.054(7)	0.173	−1676
C11	0.153(4)	3.44(6)	0.12	0.053(8)	0.173	−1681

Note. The nuisance parameters, as defined in Equations (1) and (3) are given here for different assumptions about the intrinsic scatter model, as described in Section 3 (intrinsic scatter model). The fact that $\sigma_{\rm int} \sim 0$ and $\gamma \sim 0$ for the BS21 and P21 models is due to modeling the scatter and mass step as part of the BBC process, which is discussed in further detail in Appendix A. The BS21 is the baseline choice for intrinsic scatter. The rms is given in units of magnitudes. The Hubble diagram likelihood values for each model ($\mathcal{L}$) include an uncertainty normalization term.

4. Results

4.1. Standardization Parameters

The standardization nuisance parameters α, β, γ, and $\sigma_{\rm gray}$ defined in Equations (1) and (3) are shown in Table 2 for each of the scatter models used in this work. The best-fit α are similar across scatter models to within $\sim 1\sigma$. The best-fit β values differ across models owing to different treatments of SN Ia color; however, the values for the baseline dust model (BS21) and the P21 dust model are self-consistent.

As shown in Table 2, the additional $\sigma_{\rm gray}$ term for the BS21 and P21 models is found to be zero. As discussed in Section 2, this is consistent with the expectation that if the simulations correctly model the intrinsic scatter and noise of the data, the $\sigma_{\rm scat}(z, c, M_\star)$ term of Equation (3) is sufficient to describe the distance-modulus uncertainties with $\sigma_{\rm gray} = 0$. As discussed in Appendix A, for our G10 and C11 systematic treatment, $\sigma_{\rm scat}(z, c, M_\star)$ is set to 0, and therefore $\sigma_{\rm gray} \approx 0.10$ approximates the scatter, though it does not account for the observed color dependence.

Table 2 also shows that the best-fit host stellar-mass corrections (γ) are consistent with zero for BS21 and P21. This is in agreement with the findings of Popovic et al. (2021a), that modeling the intrinsic scatter in bias-correction simulations with correlations that match those in the observed data removes the need for ad hoc corrections in intrinsic brightness (i.e., $\gamma = 0$). This can also be seen in Figure 5. For the bias correction based on the G10 and C11 models that do not include any mass dependence, the resulting γ is $\sim$0.05 found at 7σ confidence.

4.2. The Hubble Diagram and Distance Covariance Matrix

4.2.1. The Hubble Diagram

The Pantheon+ Hubble diagram of 1701 SN Ia light curves compiled from 18 different surveys and ranging in redshift from 0.001–2.26 is shown in the top panel of Figure 4. The bottom panel of Figure 4 shows the residuals to the best-fit cosmology (Equation (10)). Best-fit cosmological parameters will be presented in the following subsections.

Shown in Table 2 is the total observed scatter (rms) in the Hubble diagram residuals to the best-fit model (bottom of Figure 4) for different scatter models. The BS21 model results in the lowest Hubble diagram rms and χ^2, a $> 5\sigma$ improvement determined from the difference in likelihoods relative to the G10 and C11 scatter models. Additionally, BS21 results in no significant residuals in the Hubble diagram as a function of SALT2 c, SALT2 x_1, and host galaxy mass (Figure 5). The observed scatter of $\sim$0.17 mag is larger than that seen in the original Pantheon, because Pantheon+ extends to lower redshifts and thus is more impacted by scatter induced by peculiar velocities. If we set the minimum redshift to 0.01, the total scatter is reduced to 0.15 mag, matching that of Pantheon. Finally, compared to the original BS21 analysis, P21 uses a more rigorous fitting process that is optimized to better characterize SN Ia colors and intrinsic scatter in addition to Hubble residuals. For this reason, the improvements of P21 are not solely described by the cosmological model likelihood $\mathcal{L}$ of Table 2. We therefore have included the use of P21 population parameters as a systematic uncertainty.

4.2.2. The Very Nearby Hubble Diagram

We note from Figure 4 that in the very nearby universe, $z < 0.008$ ($v < 2400$ km s^{-1}), the mean of the Hubble diagram residuals is positive by $\sim$5% at $\sim 2\sigma$ significance. This is seen *after* the use of peculiar-velocity maps from either 2M++ or 2MRS. A similar signal is also seen in the Hubble residuals of the Cepheid distances (Kenworthy et al. 2022). A bias of roughly this size and direction is expected in the presence of measurement errors and unmodeled peculiar velocities, which scatter more objects down from higher redshifts and a greater volume than from the reverse. This effect is significant only for the most nearby galaxies ($z < 0.008$). In Figure 4, we include the prediction (dashed line) for this bias assuming 250 km s^{-1} uncorrected velocity scatter (not a fit).

In the three-rung distance ladder utilized to measure H_0 by the SH0ES Team (R22) and in Equation (14) in this work, the nearby ($z < \sim$0.01) Hubble diagram is not used. Rather, only the distance moduli from such nearby SNe are used in the SN-Cepheid absolute distance calibration in the second rung. Furthermore, in the R22 measurement of the Hubble flow, only SNe with redshifts $z > 0.023$ are used in the third rung to limit sensitivity to peculiar velocities. This approach is insensitive to the volumetric redshift scatter effects, and there is no resulting impact on the R22 H_0. However, more local measurements of H_0 from, for example, a two-rung distance ladder using primary distance indicators like Cepheids and TRGB and their host redshifts (mostly at $z \leqslant 0.01$), are more sensitive to peculiar velocities and the volumetric bias they induce, and are likely to be biased low at the few percent level if not appropriately accounting for this expected bias (Kenworthy et al. 2022). For measurements of other cosmological parameters (e.g., w or Ω_M) with Pantheon+ described in the following subsections, the mean Hubble residual biases of the Low-z and Foundation sample are $\sim$2 mmag and $\sim$1 mmag, respectively, and are considered to be negligible.

4.2.3. The Distance Covariance Matrix

Built following Equation (7), the 1701×1701 systematic distance covariance matrix is shown in Figure 6. The sample is sorted by survey and redshift to help visualize the covariances. The Hubble diagram residuals (Equation (10)) that are used to build the covariance matrix are shown in Figure 7 for several example sources of systematic uncertainty. As discussed in Appendix C, the information used to create the Hubble diagram as well as the covariance matrix is publicly available,[38] and tools to read in this information are in CosmoSIS. The SDSS subsample contributions to the covariance matrix (Figure 6)

[38] pantheonplussh0es.github.io

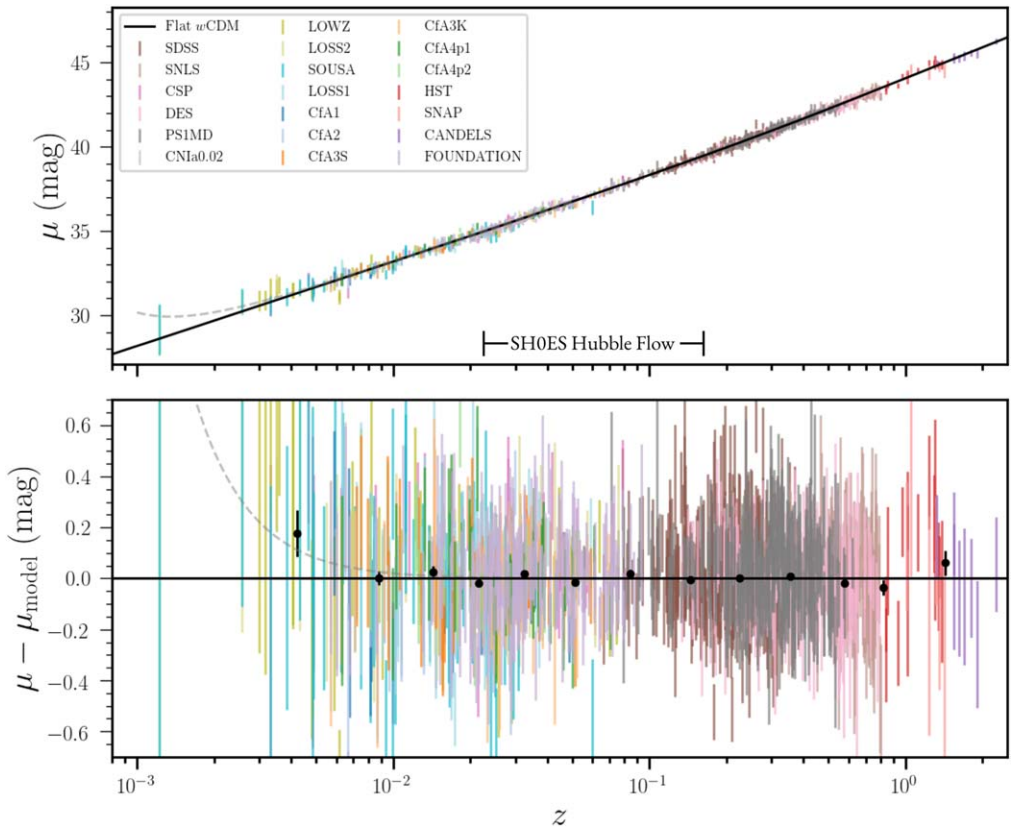

Figure 4. Top panel: the Pantheon+ "Hubble diagram" showing the distance modulus μ vs. redshift z. The 18 different surveys are each given different colors. Bottom panel: the distance-modulus residuals relative to a best-fit cosmological model with binned data for reference (black points). Both the data errors and the binned data errors include only statistical uncertainties. At $z < 0.01$, the sensitivity of peculiar velocities is very large, and the uncertainties shown reflect this uncertainty. The dashed line is the predicted Hubble residual bias stemming from biased redshifts due to volumetric effects in the very nearby universe (assuming 250 km s^{-1} uncorrected velocity scatter).

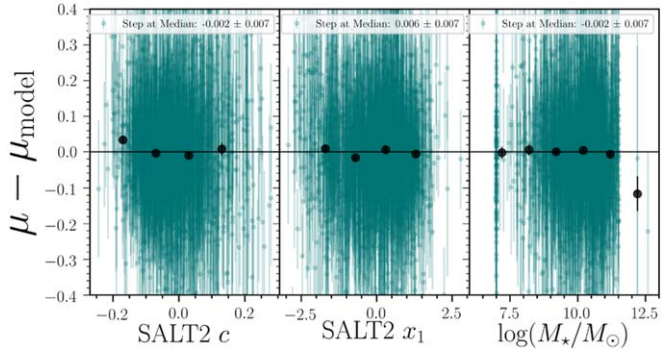

Figure 5. Pantheon+ sample Hubble diagram residuals (teal) to the best-fit cosmology ($\mu - \mu_{\rm model}$) for the baseline analysis as a function of SALT2 c, SALT2 x_1, and host-galaxy stellar mass $M_\star$. Distances (μ) follow Equation (1) and include α, β, $\delta_{\rm bias}$, and $\delta_{\rm host}$ corrections. Binned data are shown for reference (black). No significant residual correlations are seen.

stand out visually due to their strong spectroscopic selection function.

4.3. Constraints on Cosmological Parameters from Pantheon+ and SH0ES

Parameter constraints from the Pantheon+ SNe Ia and SH0ES Cepheid host absolute distances are shown in Table 3 for flat ΛCDM, ΛCDM, flat wCDM, and flat w_0w_aCDM. Unless otherwise stated, constraints on cosmological parameters include both statistical and systematic uncertainties.

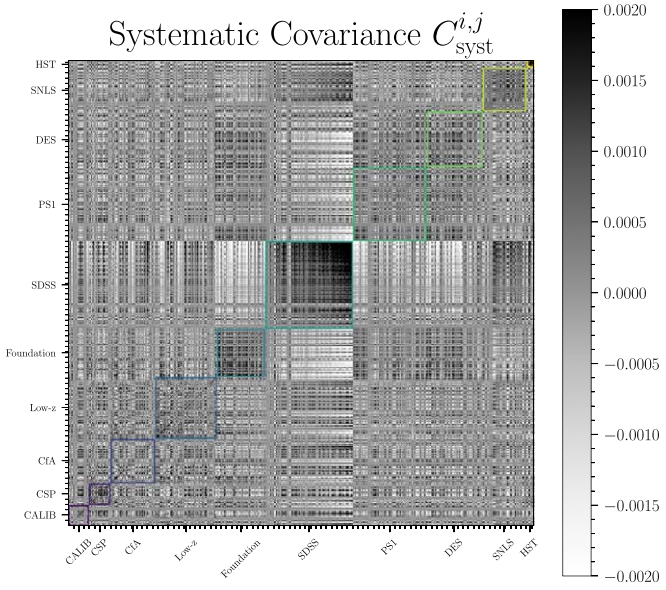

Figure 6. The systematic covariance matrix as defined in Equation (7). To show the inherent structure, the data set is sorted by survey and within each survey (colored boxes), by redshift. "CALIB" are the set of 81 SN light curves in the SH0ES Cepheid-calibrator galaxies. The shading corresponds to the size of the covariance in magnitudes.

From the Pantheon+ SNe Ia, for a flat ΛCDM model, we find $\Omega_M = 0.334 \pm 0.018$. We note that SH0ES (R22) utilizes Pantheon+ SNe at $z < 0.8$ to constrain the deceleration parameter

Table 3
Results for Cosmological Models

	Ω_M	Ω_Λ	H_0	w_0	w_a
Pantheon+ and SH0ES: *All Models*					
Flat ΛCDM	0.334 ± 0.018	0.666 ± 0.018	73.6 ± 1.1	⋯	⋯
ΛCDM	0.306 ± 0.057	0.625 ± 0.084	73.4 ± 1.1	⋯	⋯
Flat wCDM	$0.309^{+0.063}_{-0.069}$	$0.691^{+0.069}_{-0.063}$	73.5 ± 1.1	-0.90 ± 0.14	⋯
Flat w_0w_aCDM	$0.403^{+0.054}_{-0.098}$	$0.597^{+0.098}_{-0.054}$	73.3 ± 1.1	-0.93 ± 0.15	$-0.1^{+0.9}_{-2.0}$
External Probes (No SH0ES): *Flat wCDM*					
Planck and Pantheon+	$0.325^{+0.010}_{-0.008}$	$0.675^{+0.008}_{-0.010}$	$66.49^{+0.50}_{-0.83}$	$-0.982^{+0.022}_{-0.038}$	⋯
Planck and galaxy BAO and Pantheon+	$0.319^{+0.006}_{-0.007}$	$0.681^{+0.007}_{-0.006}$	$66.78^{+0.76}_{-0.50}$	$-0.974^{+0.025}_{-0.031}$	⋯
Planck and all BAO and Pantheon+	$0.316^{+0.005}_{-0.008}$	$0.684^{+0.008}_{-0.005}$	$66.87^{+1.00}_{-0.32}$	$-0.978^{+0.024}_{-0.031}$	⋯
External Probes (No SH0ES): *Flat w_0w_aCDM*					
Planck and Pantheon+	$0.318^{+0.012}_{-0.014}$	$0.682^{+0.014}_{-0.012}$	$67.4^{+1.1}_{-1.2}$	$-0.851^{+0.092}_{-0.099}$	$-0.70^{+0.49}_{-0.51}$
Planck and galaxy BAO and Pantheon+	$0.318^{+0.009}_{-0.006}$	$0.682^{+0.006}_{-0.008}$	$67.12^{+0.71}_{-0.69}$	$-0.878^{+0.063}_{-0.069}$	$-0.45^{+0.29}_{-0.32}$
Planck and all BAO and Pantheon+	$0.316^{+0.009}_{-0.005}$	$0.684^{+0.005}_{-0.009}$	$67.41^{+0.52}_{-0.82}$	$-0.841^{+0.066}_{-0.061}$	$-0.65^{+0.28}_{-0.32}$

Note. Summary of marginalized parameter constraints for Pantheon+ and other external probes. The mean and 68% confidence limit are provided for each cosmological parameter. A blank value indicates a parameter is not used in the cosmological fit.

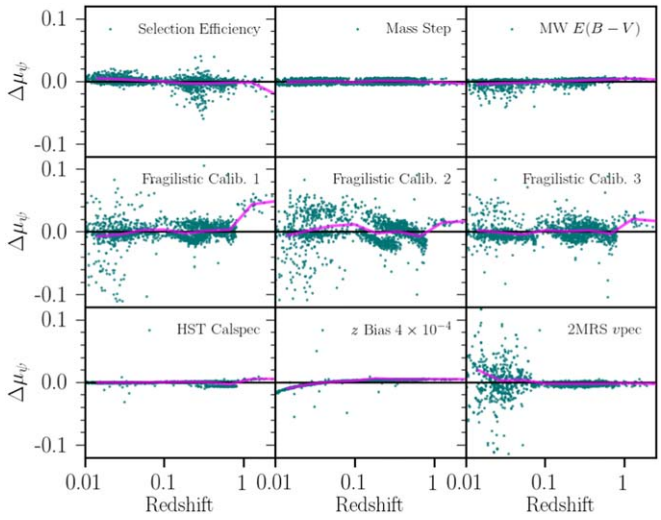

Figure 7. Visualizing the impact of a number of the top systematic uncertainties in this analysis. The μ residuals are described by Equation (6). Each of these systematics is explained in Section 3, and they are combined to form the covariance matrix shown in Figure 6. Fragilistic provides nine systematic sets of trained SALT2 models, zero-point solutions, and filter central wavelengths. Here we show the impact on distance of just the first three.

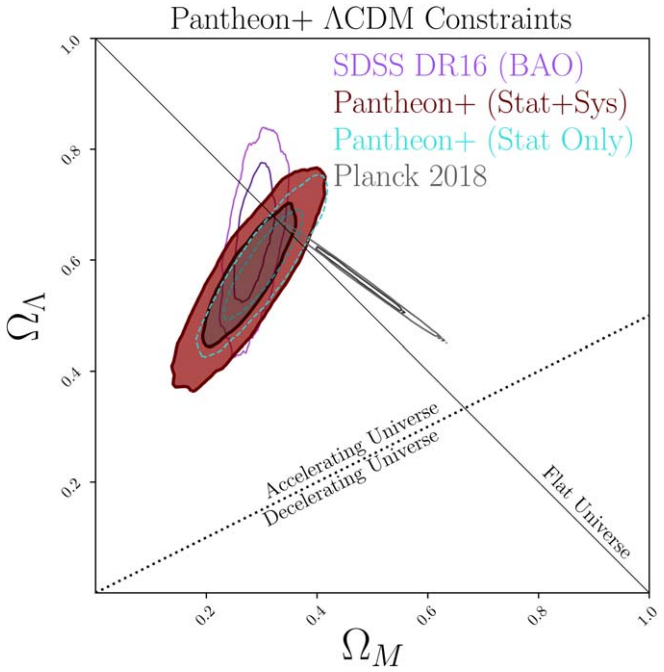

Figure 8. Confidence contours at the 68% and 95% levels for the Ω_M and Ω_Λ cosmological parameters for the ΛCDM from the Pantheon+ data set, as well as from the Planck and combined BAO data sets. The constraints from including both the statistical and systematic uncertainties (shaded red) are shown as well as when only statistical uncertainties are propagated (unfilled dashed). We include two lines for reference: one for a flat universe, where $\Omega_M + \Omega_\Lambda = 1$ and the other that indicates an accelerating universe.

and find $q_0 = -0.51 \pm 0.024$. In a flat universe $q_0 = \frac{3\Omega_M}{2} - 1$, which gives $\Omega_M = 0.326 \pm 0.016$, consistent with the result for Ω_M reported in this work. Results for H_0 from the inclusion of the SH0ES Cepheid host distances are discussed below.

The constraints on Ω_M and Ω_Λ for a ΛCDM model are shown in Figure 8. We find $\Omega_M = 0.306 \pm 0.057$ and $\Omega_\Lambda = 0.625 \pm 0.084$; a flat universe is within the 68% confidence region, and $\Omega_M = 0$ and $\Omega_\Lambda = 0$ are together rejected at 4.4σ using only the SNe.

For a flat wCDM model, from the SNe Ia alone (not including SH0ES Cepheid calibration), we find $\Omega_M = 0.309^{+0.063}_{-0.069}$ and $w = -0.90 \pm 0.14$, as shown in the third row of Table 3 and in the blue contour of Figure 9. This result is consistent within 1σ of the cosmological constant ($w = -1$).

For a flat w_0w_aCDM model, from the SNe Ia alone (not including SH0ES Cepheid calibration), we find $w_0 = -0.93 \pm 0.15$ and $w_a = -0.1^{+0.9}_{-2.0}$, as shown in the fourth row of Table 3 and in Figure 10. These results are again consistent with a cosmological constant.

Using distances and a stat+syst covariance matrix that extends to the Cepheid calibrators (Equation (15)) and combining the Pantheon+ SNe with the SH0ES Cepheid host distance calibration, we are able to robustly and simultaneously constrain

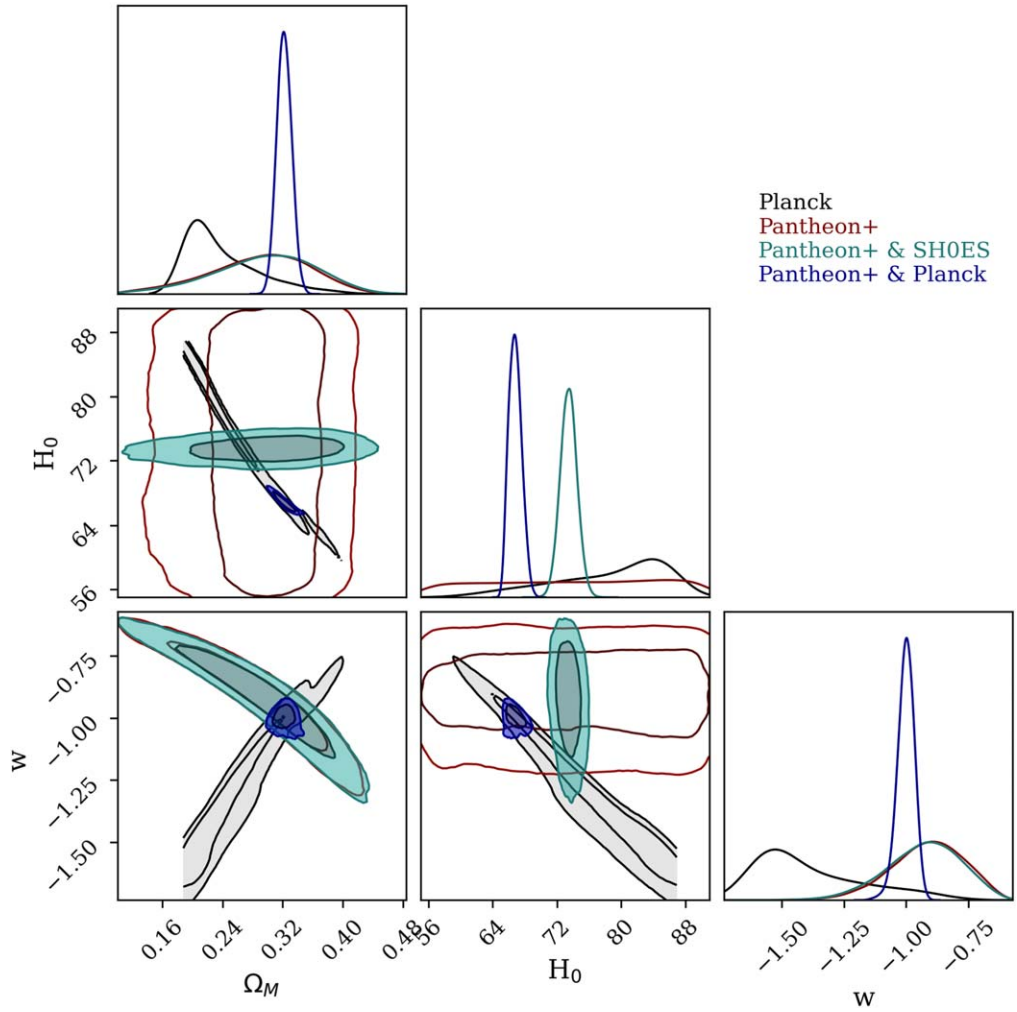

Figure 9. Sixty-eight percent and 95% confidence contours for flat wCDM for cosmological parameters Ω_M, H_0, and w. The contours from the Pantheon+ (red), Pantheon+ and SH0ES combined data set (teal), Planck Collaboration et al. (2020) TTTEEE-lowE constraints (gray) are shown. The combination of Planck and Pantheon+ (blue) is also shown, which is consistent with a cosmological constant. Planck constraints are bounded by $0.2 < \Omega_M < 0.4$ for computational speed. The histograms depict marginalized relative probabilities between probes.

H_0 and other cosmological parameters describing the expansion history. While we use SH0ES Cepheid data and covariance in this work, likewise, Pantheon+ distances and covariance are used in R22 in order to fit H_0 and q_0 in flat ΛCDM. As shown in the top Pantheon+ and SH0ES section of Table 3, for ΛCDM, flat wCDM, and flat w_0w_aCDM, we find $H_0 = 73.4 \pm 1.1$, 73.5 ± 1.1, and 73.3 ± 1.1 km s^{-1} Mpc^{-1}, respectively. We note that more complex models do not result in decreased H_0 constraining power from the SNe Ia + Cepheids, while this is not necessarily true for other cosmological probes (Section 4.4).

4.4. Constraints on Cosmological Parameters from Multiple Probes

In this work we combine the Pantheon+ SNe with external cosmological probes: CMB from Planck (Collaboration et al. 2020) TTTEEE-lowE and BAOs from SDSS Main Galaxy Sample (Ross et al. 2015), SDSS Baryon Oscillation Spectroscopic Survey (BOSS; Alam et al. 2017), SDSS Extended Baryon Oscillation Spectroscopic Survey (eBOSS) Luminous Red Galaxy (Bautista et al. 2020), SDSS eBOSS Emission Line Galaxies (Bautista et al. 2020), SDSS eBOSS Quasar (Hou et al. 2020), SDSS eBOSS Lyα (du Mas des Bourboux et al. 2020), all of which have been implemented in CosmoSIS. The aforementioned BAO constraints are denoted "all BAO"; we also provide constraints from the combination of a spectroscopic redshift galaxy-only subset of BAO probes denoted "galaxy BAO." We report constraints in Table 3 for combinations of data sets that are deemed compatible and discussed below.

For a flat wCDM model when combining Pantheon+ and Planck, we find $w = -0.982^{+0.022}_{-0.038}$ and $\Omega_M = 0.325^{+0.010}_{-0.008}$, and when further including all BAO, we find $w = -0.978^{+0.024}_{-0.031}$ and $\Omega_M = 0.316^{+0.005}_{-0.008}$, both of which are consistent with the cosmological constant at ∼3% (Figure 11). As can be seen in Figure 9, we do not include SH0ES in combinations with Planck because these measurements are incompatible (R22).

For a flat w_0w_aCDM model when combining Pantheon+ and Planck, we find $w_0 = -0.851^{+0.092}_{-0.099}$ and $w_a = -0.70^{+0.49}_{-0.51}$, and when combining Pantheon+, Planck, and BAO, we find $w_0 = -0.841^{+0.066}_{-0.061}$ and $w_a = -0.65^{+0.28}_{-0.32}$, which is moderately consistent (2σ) with a cosmological constant (Figure 12). We note that this result is not driven by any single probe. In Figure 10 we show constraints for Planck alone and for the combination of Planck & Pantheon+. While the broader model freedom of the flat w_0w_aCDM allows the Planck-alone H_0 to be consistent with 73 km s^{-1} Mpc^{-1} owing to degeneracy between H_0 and w_a (see

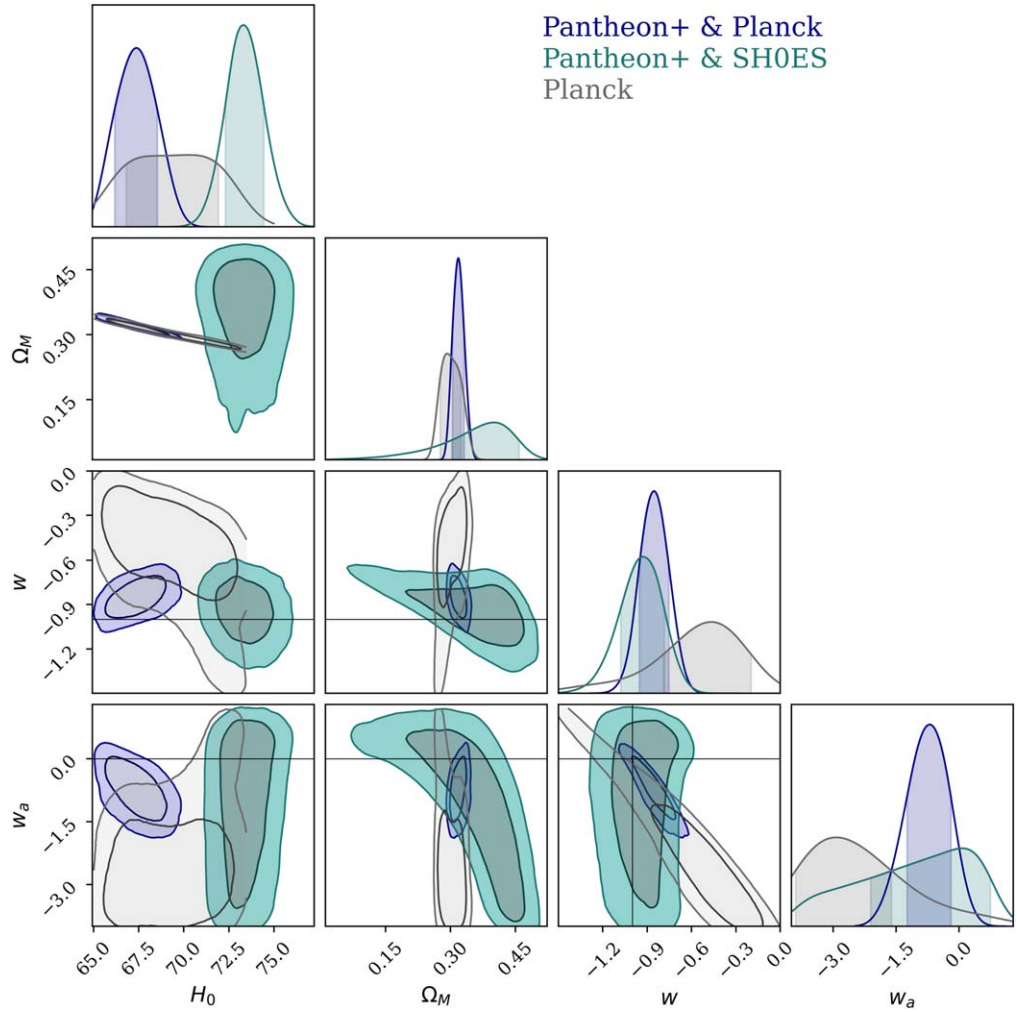

Figure 10. Constraints for flat w_0w_aCDM from the Pantheon+ data set in combination with SH0ES, Planck TTTEEE-lowE.

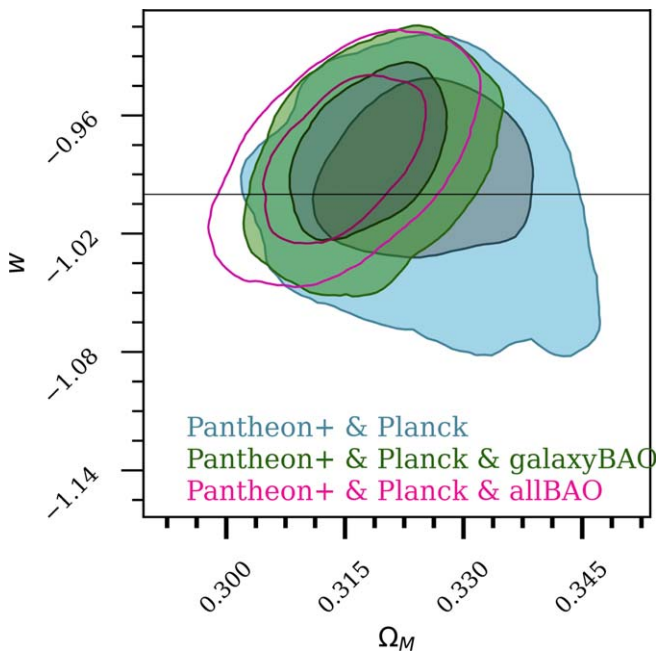

Figure 11. Constraints for flat w_0CDM from the Pantheon+ data set in combination with Planck and galaxy BAO or Planck and all BAO.

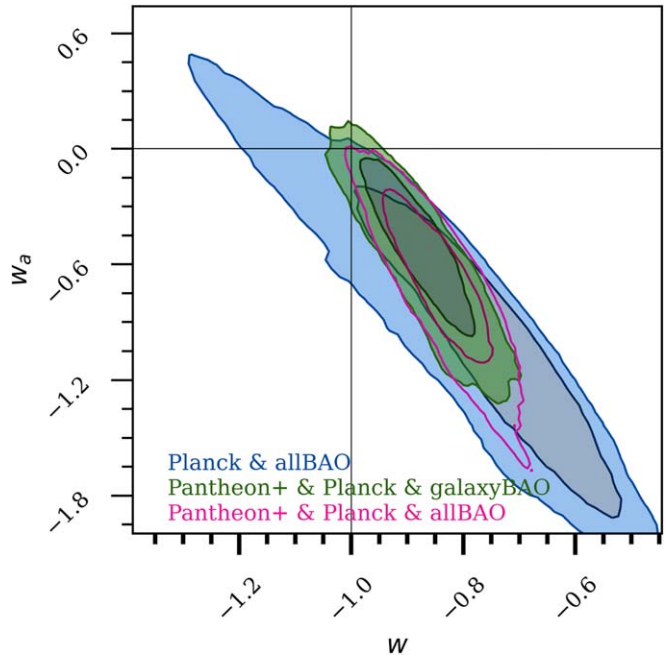

Figure 12. Constraints for flat w_0w_aCDM from the Pantheon+ data set in combination with Planck and galaxy BAO or Planck and all BAO.

Table 4
Sources of Uncertainty

Description	Baseline	Systematic (S_ψ)	σ_ψ	$\sigma_{w\mathrm{sys}}$	$\sigma_{w\mathrm{sys}}/\sigma_{w\mathrm{stat}}$	Δw_{sys}
All Systematics				0.019	0.79	−0.009
Calibration						
SALT2 Train[a]ZPT	Fragilistic best fit	10 covariance realizations	1/3 each	0.009	0.38	0.000
SALT2 Method	SALT2-B22	JLA SALT2 Surface	1/3	0.008	0.33	0.003
CSP Tertiary Stars	Krisciunas et al. (2017b)	Stritzinger et al. (2018)	1	0.003	0.13	−0.003
HST	Calspec 2020 update	5 mmag/7000 Å	3	0.003	0.13	−0.006
Redshifts						
v_{pec} Map	2M++	2M++ iLOS & 2MRS	0.7 each	0.002	0.08	0.005
Redshift Bias	No z-shift	10^{-4} z-shift	1	0.011	0.46	0.015
Astrophysics						
Intrinsic Variations	BS21 dust model	G10 and C11	0.7 each	0.002	0.08	−0.003
MW $E(B-V)$	Schlafly & Finkbeiner (2011)	4% Scaling	1.0	0.008	0.33	−0.010
MW Color Law	Fitzpatrick (1999)	Cardelli et al. (1989)	1/3	0.006	0.25	−0.006
Mass Step	Split at 10	Split at 10.2	1	0.001	0.04	0.000
Modeling						
Selection Efficiency	Nominal exposure time	20% increase	1	0.004	0.17	0.001
Populations	BS21 parameters	Three Sets of Params (P21)	0.6	0.000	0.00	0.003

Notes. A summary of the systematic uncertainties and the baseline component of the analysis as described in Section 3, the size of the systematic S_Ψ used to determine the impact of that systematic, the scaling of the systematic σ_ψ as constrained in this analysis, and the contribution to the total uncertainty in wCDM (can be compared to statistical uncertainty of 0.03), and the shift when allowing the uncertainty on the best-fit cosmological parameter. The last column shows the simplistic change in best-fit cosmology if a perturbation of size σ_ψ is applied with statistical-only uncertainties. The amount shown is different than seen for the combined shift for the best fit and increase of uncertainty given in the previous columns due to the self-calibration, as explained by Brout et al. (2021).
[a] ZPT denotes light-curve fitting zero-points.
[b] Constraints are combined with Planck Collaboration et al. (2020).

Figure 10), after combining Planck with Pantheon+, the H_0/w_a degeneracy is broken ($H_0 = 67.4^{+1.1}_{-1.2}$ km s^{-1} Mpc^{-1}). Therefore, the inclusion of SH0ES with Planck & Pantheon+ results in a Bayesian evidence ratio of −9, and we deem this set of probes incompatible and do not include them in Figure 10 nor in Table 3.

4.5. Impact of Systematics on Cosmological Parameter Fits

To understand the impact of systematic uncertainties, in Table 4 we group the systematics investigated in this work into four main categories: calibration/SALT2, redshifts, astrophysics, and modeling. The baseline, systematic treatments (S_ψ), and scaling priors (σ_ψ; as described in detail in Section 3) are summarized for each source. The final three columns of Table 4 relate to fits of the sample when combined with Planck Collaboration et al. (2020) in a flat wCDM model when isolating that systematic. We define both the change in the best fit (Δw_{sys}) and the systematic uncertainty contribution to w (σ_w^{sys}) as follows:

$$\Delta w_{\mathrm{sys}} = w_{\mathrm{sys}} - w_{\mathrm{stat}} \tag{16}$$

$$\sigma_w^{\mathrm{sys}} = \sqrt{\sigma_{w\mathrm{tot}}^2 - \sigma_{w\mathrm{stat}}^2}, \tag{17}$$

where w_{sys} and $\sigma_{w\mathrm{tot}}$ are the cosmological constraints when utilizing $C_{\mathrm{stat+sys}}$, and where w_{stat} and $\sigma_{w\mathrm{stat}}$ are the statistical-only constraints when utilizing C_{stat}.

We find that the final systematic uncertainty in w ($\sigma_{\mathrm{sys}} = 0.019$) is comparable to yet smaller (∼80%) than the statistical uncertainty, suggesting that the measurement is not systematics dominated. The largest contribution to the systematic error budget (0.011) is due to the potential for redshift-measurement bias. This is followed by the uncertainties in the Fragilistic calibration offsets and the resulting propagation to SALT2 model-training uncertainties and light-curve fitting uncertainties (0.009). Additionally important is the conservative uncertainty that was applied owing to the usage of the new SALT2 training methodology (0.008) as well as the uncertainty in the MW extinction maps (0.008).

Interestingly, numerous systematic uncertainties are found to be negligible (e.g., BS21 parameters, G10 versus C11) in the cosmological parameter budget. While certain systematics cause redshift-dependent trends as shown in Figure 7, they also change the relative scatter of the Hubble residuals. This can most easily be seen for the cosmological likelihood values ($\mathcal{L}$) for the distances with different intrinsic scatter models shown in Table 2. If the baseline analysis is significantly preferred (larger $\mathcal{L}$) by the data over one of the analysis variants, the impact of that systematic on cosmological constraints will be reduced, as is the case for intrinsic scatter.

As we have built a covariance matrix that includes the Cepheid calibrators, we can measure H_0 with and without systematic uncertainties. For flat ΛCDM, we find $H_0 = 73.6 \pm 1.1$ km s^{-1} Mpc^{-1}, and when considering only statistical uncertainties from the SNe alone (excluding Cepheid and physical distance calibration uncertainties), $\sigma_{H_0}^{\mathrm{stat+syst}} = 0.7$ km s^{-1} Mpc^{-1}, and $\sigma_{H_0}^{\mathrm{syst}} = 0.29$ km s^{-1} Mpc^{-1}. This suggests that SN systematic uncertainties are not dominating the constraint on H_0 and cannot explain the ∼7 km s^{-1} Mpc^{-1} difference between Planck and SH0ES.

In Figure 13 we show deviations to the best-fit H_0 for each individual source of systematic uncertainty relative to the baseline analysis and assuming ΛCDM. For reference, we also show the full SN contribution to the H_0 error bar (dashed). The deviations from the baseline (ΔH_0) are small and added in quadrature to 0.32 km s^{-1} Mpc^{-1}. We note that when assessing redshift-specific systematics, because model redshifts are not used for the SN-Cepheid calibration in Equation (14), they mainly impact the Hubble-flow SNe (third rung of the distance ladder).

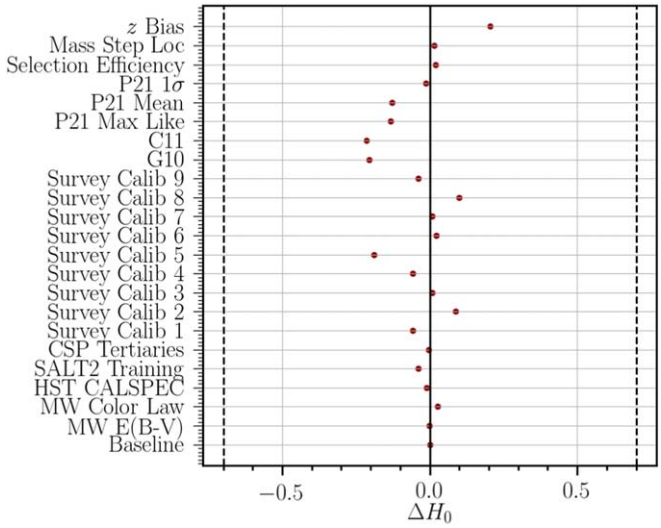

Figure 13. The impact on recovery of H_0, as explained in Section 2, of the systematic uncertainties described in Table 4. The units of these measurements are km s^{-1} Mpc^{-1}. The dashed lines are given at a ΔH_0 of 0.7, which is the entire contribution of the uncertainty in R22 from SN measurements.

Finally, to help visualize the impact of systematic uncertainties, we show in Figure 8 the constraints when including either statistical-only uncertainties or the combined statistical and systematic uncertainties. Error budgets for different cosmological parameterizations can be generated with the delineated files for systematics provided as part of this release.

4.6. Local Structure in the SN Ia Hubble Diagram

Large compilations of SN distances have provided impetus for searches of local structure, over/underdensities, and proper motion (e.g., Mathews et al. 2016; Soltis et al. 2019; Hu et al. 2020). As an initial study, we create sky maps of the SN Hubble diagram residuals (see Figure 14) and examine two specific areas on the sky that have been documented in the literature and have sufficient SN statistics in the Pantheon+ sample for study.

4.6.1. The CMB Kinematic Dipole

The motion of the Milky Way and solar system relative to the CMB rest frame ($v = 369.82$ km s^{-1}) is corrected for following Carr et al. (2021) and Peterson et al. (2022). The effect of the CMB dipole motion can be seen in the $z_{\rm HEL}$ sky map (middle-right panel of Figure 14), where $z_{\rm HEL}$ is the heliocentric redshifts. The $z_{\rm CMB}$ skymap (middle-left panel of Figure 14) has the CMB dipole-causing peculiar redshift removed, following Equation (7) of Peterson et al. (2022). The direction of the CMB dipole, $l = 264°$ and $b = 48°$ (red "o" in Figure 14), is shown for reference as well as its antipole (red "x").

As discussed in Section 3.1.3, we examine different velocity reconstructions due to local structure that include estimates of the bulk flow; these are the 2M++ (Carrick et al. 2015) and 2MRS (Lilow & Nusser 2021) corrections and are shown in the top row of Figure 14. These corrections also include the CMB dipole correction. Peterson et al. (2022) showed that the peculiar-velocity corrections overall reduce the Hubble residual scatter by $\sim 10\%$, and this is qualitatively confirmed in our maps. The heliocentric map shows a strong dipole as expected; the $z_{\rm CMB}$ map shows the dipole somewhat removed but with an overcorrection (as expected at low-z because local galaxies share some of our motion); and both $z_{\rm HD}$ maps show that the peculiar-velocity corrections have removed most of the overcorrection.

However, both reconstructions produce a small signal that can be seen in the maps in the direction opposite the motion causing the CMB dipole. This signal is found to be local, at $z < 0.02$, and grows with decreasing redshift until $z \approx 0.01$ (bottom-left panel of Figure 14). A possible reason that there is a residual signal in the negative dipole direction in both the $z_{\rm CMB}$ and peculiar-velocity corrected redshifts is that the MW motion is coupled with the motion of nearby galaxies in a way that is not yet sufficiently modeled. It is also likely that this is due to low-number statistics (this is only a 1σ deviation) and the uneven sky coverage (the SNe in this region are mostly clustered in Stripe-82). Lastly we note that the positive residuals are driven by SNe at $z < 0.02$, and thus are not included in the SH0ES (Riess et al. 2022) sample and inference of H_0.

4.6.2. The CMB Cold Spot

The "CMB cold spot," a 5° region of $-70\,\mu$K centered at ($l \sim 209°$, $b \sim -57°$), was first detected in data from the Wilkinson Microwave Anisotropy Probe (Vielva et al. 2004; Cruz et al. 2006), and subsequently in Planck data (Gurzadyan et al. 2014). Evidence for an underdensity aligned with the CMB cold spot was presented by Rudnick et al. (2007). Szapudi et al. (2015) and Kovács et al. (2021) subsequently found the Eridanus supervoid in the direction of the cold spot at $z \approx 0.15$. However, it is not clear if the alignment of Eridanus and the CMB cold spot is causal or coincidental.

We find a signal in the Pantheon+ Hubble diagram when examining SNe within a 20° radius of the location of the CMB cold spot (blue circle region in the top-left panel of Figure 14). The difference in Hubble diagram residuals as a function of redshift is shown in the bottom-right panel of Figure 14. There are nine SNe in this region of the sky with redshifts on the near side ($0.12 < z < 0.15$), and there are 14 SNe on the far side ($0.15 < z < 0.20$) of the proposed void at $z = 0.15$. There is a Hubble residual difference of -0.15 ± 0.06 mag between these two sets of SNe. For an estimate of the significance, we examine 1000 randomly selected 20° apertures across the sky with at least eight SNe in each of the near and far redshift ranges split on redshifts between 0.08 and 0.20, and find that deviations with a similar significance occur only 0.2% of the time. We note however, that there are not many independent regions that satisfy the selection criteria, and the vast majority of the SNe in the cold-spot selection come from the small deep-field patch within that region. Taking 100 random samples of 10° radius from the largest densely sampled region in Pantheon + (Stripe-82 region), we find no other patch has a significance that exceeds 1.6σ, making the Eridanus patch the most significant step at that redshift in our data.

5. Discussion

This analysis is the latest in a series of papers that attempts to both grow the compilation of measured SN Ia light curves and improve on the systematic floor. The two most recent compilations and analyses are those of JLA and Pantheon, which, respectively, included $\sim 40\%$ and $\sim 60\%$ of the SN light curves analyzed here. As seen in Figure 1 of Scolnic et al. (2022),

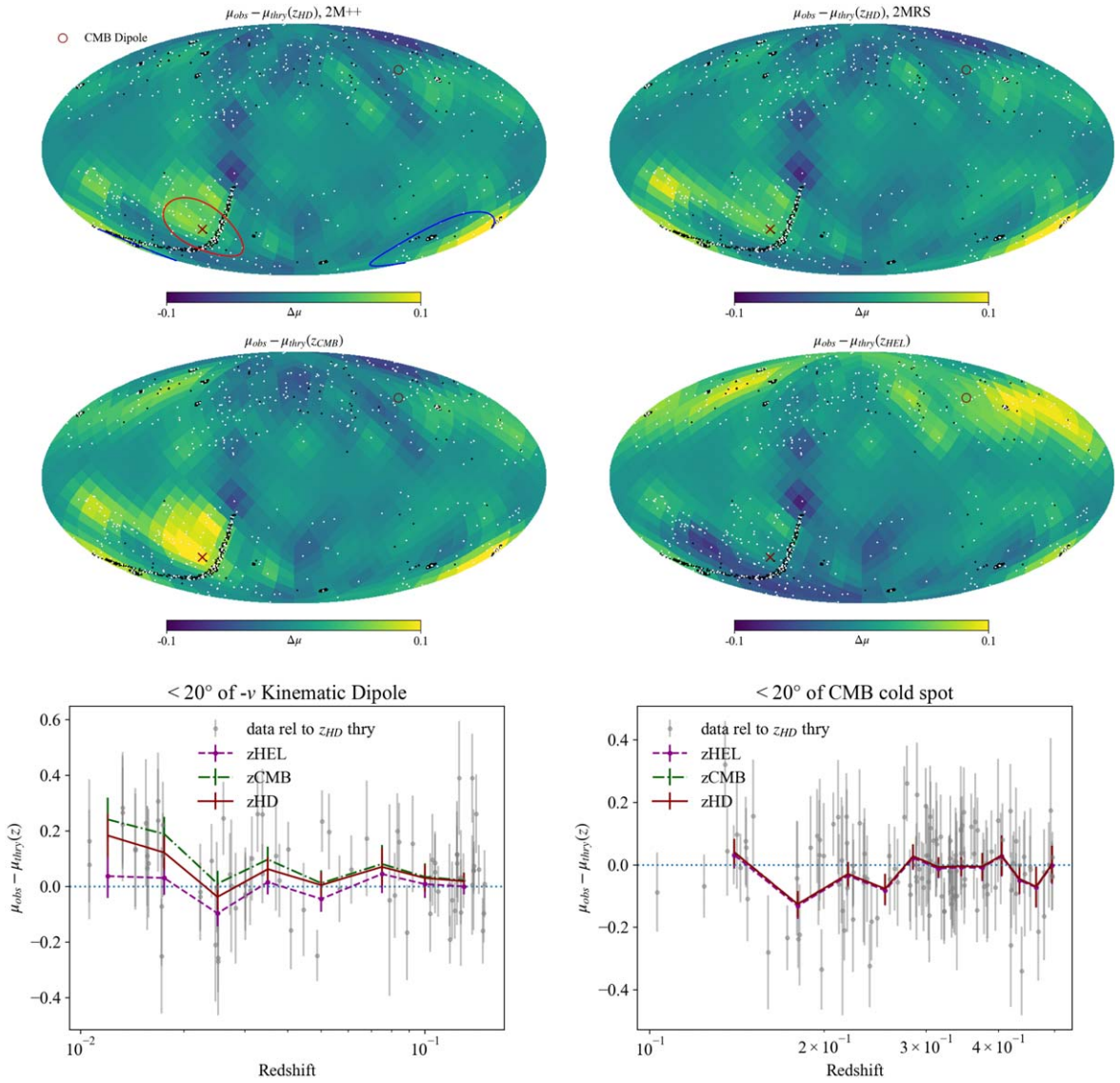

Figure 14. Healpix (NSIDE = 16) Hubble residual sky maps (the color bar is residual magnitudes) with 20° 2D-Gaussian kernel smoothing, and Hubble residuals for two selected apertures. $z > 0.01$ is applied. Dots show the locations of the SNe in the Pantheon+ sample, with white dots showing the nearby SNe ($z < 0.15$) and black dots showing the distant SNe ($z > 0.15$). Top left: Hubble diagram corresponding to the baseline analysis utilizing both $z_{\rm CMB}$ dipole corrections and 2M++ peculiar-velocity corrections. The circled regions designate the 20° regions centered on the negative CMB dipole (red) and CMB cold-spot directions (blue). The small circle in the top right (and "x" in bottom left) of each panel represents the direction (and opposite direction) of the motion causing the CMB dipole. Top right: same as top left, but instead using 2MRS peculiar-velocity corrections. Middle left) same as top left, but instead not applying any peculiar-velocity corrections. Middle right: same as top left, but instead applying neither peculiar-velocity corrections nor the CMB dipole correction. Bottom left: 20° region aligned with the (opposite) CMB dipole velocity depicting Hubble diagram residuals as a function of redshift. Bottom right: same as bottom left, but with aperture centered at the CMB cold spot ($l = 209°$, $b = 57°$), and over a higher redshift range.

the majority of the statistical increase for Pantheon+ is in the addition of numerous low-redshift samples extending down to $z = 0.001$. However, the largest differences in the Hubble diagram are not solely the result of statistical increase, but rather due to improvements in our methodology.

We show in Figure 15 the difference in inferred distance-modulus values (marginalized over M) for the Pantheon+ sample relative to the assumptions used in the JLA analysis, for the three most significant improvements presented in this work. First is the update in the flux cross-calibration to the Fragilistic solution, which impacts both the training of the SALT2 model and the zero-points used in light-curve fitting. Second is the impact from updating the MW extinction curve used in JLA (Cardelli et al. 1989) to the Fitzpatrick (1999) relation that is used here. Third is the change resulting from improved modeling of the SN Ia intrinsic scatter; while in this work we adopt the BS21 model, we include the models developed for JLA (G10 and C11) as systematics. Each of these changes has been motivated externally by previous works (e.g., Schlafly & Finkbeiner 2011; Brout & Scolnic 2021; Brout et al. 2022); however, they nonetheless cause shifts in $d\mu/dz$ of ∼0.05, or ∼0.04 in w. Finally, because all of three of these changes have the same sign of $d\mu/dz$ slope, rather than canceling each other, when combined in this work, they result in a ∼0.1 difference in the constraint on w relative to JLA (after combining with CMB).

As discussed by Scolnic et al. (2019), the constraining power of large samples of SNe Ia extends beyond inferences of H_0 and w/Ω_M. Large compilations of low-z SNe Ia enable precision measurements of the local growth-of-structure, typically parameterized by $f\sigma_8$ (e.g., Huterer et al. 2017; Stahl et al. 2021).

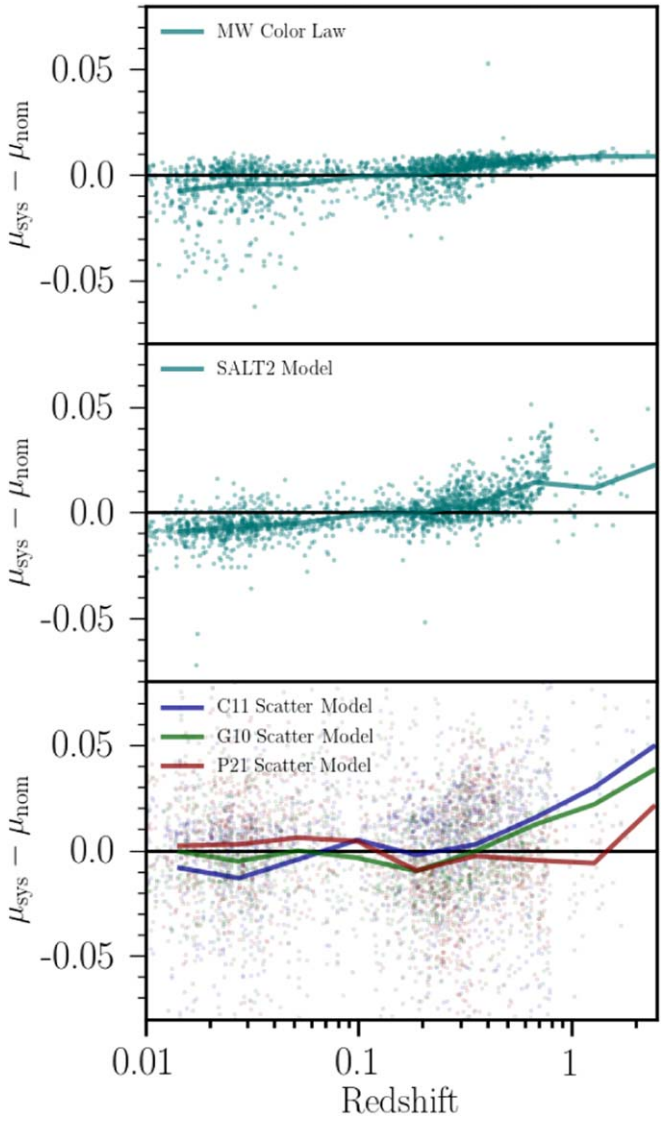

Figure 15. Largest differences in analysis compared to Betoule et al. (2014) and Scolnic et al. (2018b). Top: updating the extinction curve used in the light-curve fitting from CCM to F99. Middle: updating the SALT2 model, as discussed in Brout et al. (2022). Bottom: changing the baseline assumption for the intrinsic scatter to the P21, G10, and C11 models.

Work is ongoing for this measurement using the Pantheon+ sample (S. Boruah et al. 2022, in preparation), which will include validation with simulations as well as propagation of the covariance matrix, which previously would have had a limited effect on σ_8 calculations owing to smoothing/binning over redshift.

While in Section 4 we show a Healpix map of Hubble residuals across the sky, there are additional and related tests of anisotropy that can be performed with these data. Previous analyses of the first Pantheon sample (e.g., Andrade et al. 2018; Brownsberger et al. 2019; Colin et al. 2019; Soltis et al. 2019) typically search for radial or hemispherical residuals across the sky. The addition of statistics in the low-redshift sample and improved accounting in Pantheon+ would particularly strengthen these types of studies. A search for matter over/underdensities was performed by Colgáin (2019), which varied the minimum and maximum redshift in the original Pantheon sample and redetermined cosmological constraints. Colgáin (2019) found for Pantheon that Ω_M could be <0 for a low maximum z of $\sim$0.15, though with only $\sim 2\sigma$ difference

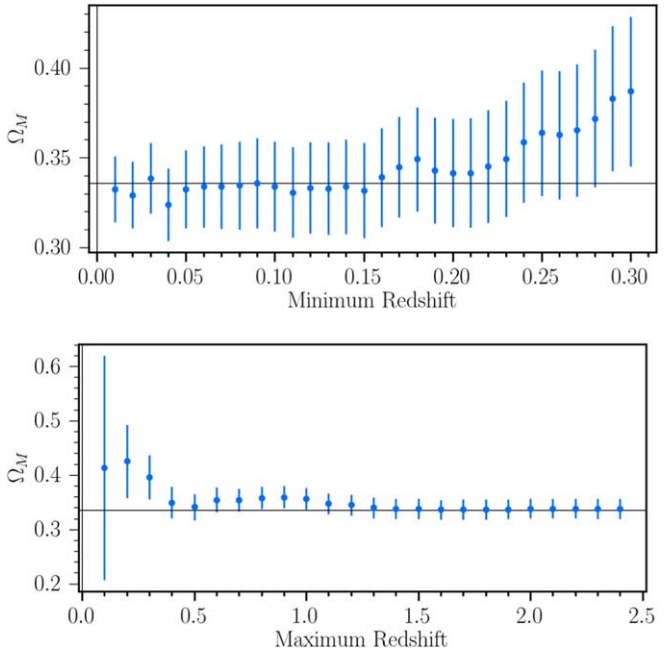

Figure 16. Constraints on Ω_M in flat ΛCDM when the bounds of the redshift range of the sample are changed. In the top panel, the minimum redshift is varied. The nominal minimum redshift is 0.01 for Pantheon+ cosmology fits without SH0ES. In the bottom panel, the maximum redshift is varied. The nominal maximum redshift is 2.4 for all fits.

compared to the value of Ω_M from the full sample. We show a similar test in Figure 16 and find relatively stable values of Ω_M with no signs of the underdensity seen by Colgáin (2019).

Investigation of the main goal of this work (constraints from SNe Ia alone for a flat wCDM model) results in stat+syst uncertainties of $^{+0.058}_{-0.063}$ and 0.13 for Ω_M and w, respectively. This represents a factor of two improvement in figure of merit over the original Pantheon (stat+syst uncertainties 0.072 and 0.22 for Ω_M and w). This cannot be explained solely by statistical improvements, but rather is also due to a leap in systematics methodology over the original Pantheon and JLA. As shown by Brout et al. (2021), cosmology uncertainty budgets are improved by a factor of $\sim$1.5 when not binning or smoothing data and covariance. In Appendix B we discuss and show a binned error budget for comparison and find a similar factor of 1.5 improvement from this choice alone. In examining the unbinned error budget in Table 4, it can be seen that several systematics are no longer impacting SN Ia cosmology analyses as strongly as had previously been thought. One such example is the negligible size of the parent population systematic despite including three additional sources of scatter model uncertainty, as was also seen by Popovic et al. (2021a). This, as well as the reduction of a number of other systematics in comparison to their size in binned analyses (also shown in Appendix Table 6), is due to the power of the large data sets themselves to self-constrain the size of systematic uncertainties when the systematic itself is not solely degenerate with the cosmological model parameterization. This is especially important because it brings this work from potentially being dominated by systematics to rather being dominated by statistical uncertainties. Furthermore, as shown by Brout et al. (2021), as data sets grow in size, many systematics will continue to shrink without any additional effort. Lastly, it is important to note that approaches such as the approximate Bayesian computation method given by Jennings et al. (2016) will not be able to make use of this self-constraining benefit unless additional parameters are included to

allow the data themselves to scale the input sizes of the systematic uncertainties (S_{sys} in Brout et al. 2021).

While the SN Ia mass step has received much attention in the last decade, we find here that its contribution to the error budget is exceedingly small. Unlike previous analyses, the mass-step treatment in this work is based on an SN color- and dust-dependent model (BS21). We find that this more physical model results in smaller scatter in the Hubble diagram (Table 2) and better χ^2 relative to cosmological models, which then results in smaller systematic uncertainties. We note that properties of SN Ia host galaxies other than stellar mass have been seen to correlate with SN Ia Hubble diagram residuals. Star formation rate, specific star formation rate (sSFR), stellar-population age, and metallicity have all been shown to correlate to varying degrees with the distance-modulus residuals after standardization (Lampeitl et al. 2010; Sullivan et al. 2010; Childress et al. 2013b; Rigault et al. 2013; Rose et al. 2019). For this reason, using sSFR values presented by S22, we also examined the size of an sSFR step in the subset of the Pantheon+ sample for which we have obtained sSFR measurements ($z < 0.2$). Without applying any bias corrections, we find a significant step in sSFR (across the median sSFR) of 0.031 ± 0.011. However, after applying the nominal set of dust and mass-based bias corrections (BS21) used in this analysis, we find a step in sSFR of 0.008 ± 0.011, consistent with zero. This is likely due to galaxy properties (i.e., stellar mass) being linked to dust properties, and that applying a dust-mass correction is accounting for most, if not all, of the correlations with sSFR and is also tracing the dust distribution.

Going forward, statistical constraints on w and Ω_M from SNe will improve significantly owing to upcoming data sets from SN programs of the DES (D'Andrea et al. 2018), Zwicky Transient Facility (Dhawan et al. 2022), Young Supernova Experiment (Jones et al. 2021), Legacy Survey of Space and Time (The LSST Dark Energy Science Collaboration et al. 2018; Sánchez et al. 2022), Nancy Grace Roman Telescope (Hounsell et al. 2018), etc. It is likely that these future data sets will improve the statistical precision by a factor of 100 (Scolnic et al. 2018a).

The size of systematic errors on cosmological parameter estimates matched the statistical errors for JLA and the original Pantheon. Systematic uncertainties in this work have been reduced in comparison to Pantheon, and while their impact is still significant, it is no longer the dominant component of the total uncertainty. With the coming surveys, systematics will also likely improve alongside the increase in statistics, as has been the case for previous analyses over the last two decades, and as expected from the impact of the systematic "self-calibration" described in Brout et al. (2021).

As shown in the systematics error budget Table 4, the dominant sources of systematic uncertainty are now from (1) the combination of SALT2 training and calibration of surveys, (2) potential redshift-measurement biases, and (3) Milky Way dust systematics. Fortunately there are paths forward for each of these. For survey flux calibration, dedicated programs are needed, and there are currently multiple paths underway to improve the fundamental calibration of SN Ia samples and how they are tied to various other samples (e.g., Regnault et al. 2015; Stubbs & Brown 2015; Narayan et al. 2019). There is also ongoing work (G. Taylor et al. 2022, in preparation) to train the SALT2 model with more photometric systems, which has already shown promising improvements to systematic uncertainties and the ability to constrain the rest-frame U band. The systematic from the redshift-measurement floor has the potential to be reduced using improved cosmology fitting methodology, although the extent to which the data itself can constrain the size of this floor remains unproven. Alternatively, future large surveys can use multiple spectroscopic instruments and redshifting codes to mitigate potential sources of redshift-measurement bias. The Pantheon+ sample is especially sensitive to Milky Way dust systematics because of the differences in the samples used for low and high redshift. At low redshift, to obtain sufficient statistics in a volume limited sample, we have used SNe across the sky and with up to 0.2 in MWEBV, whereas the high-redshift surveys have been carried out in low extinction regions of the sky (MWEBV < 0.05). Future surveys of larger volumes will be able to mitigate this with a plethora of both low and high redshift in low MW extinction regions on the sky.

Throughout this work, there are a number of upstream components of this analysis that impact downstream analysis steps; i.e., new calibration (or MWEBV maps/color law) motivates new SALT2 training, which motivates new fitting of the SN parent populations, which motivates new bias corrections. The Pippin framework (Hinton & Brout 2020), used extensively in this work, was intentionally developed to automate and asynchronize this multistep type of analysis; however, it has yet to incorporate aspects such as the SALT2 retraining (Taylor et al. 2021) or population fitting (Popovic et al. 2021a). Likely, this framework will need to expand for future analyses.

There is an alternate approach to obtaining cosmology constraints from SNe that has been gaining traction over the last decade. Bayesian hierarchical models have been developed that utilize bias-corrected observables (Shariff et al. 2016) and that incorporate selection effects directly into the model (Rubin et al. 2015) or likelihood (Hinton et al. 2019). However, unlike BBC in combination with CosmoSIS, these methods have not been validated with large realistic simulations. As noted in Appendix C, we release, as part of this analysis, 10 realistic simulations of the Pantheon+ data set for such validations.

While constraints on w should easily improve with upcoming large SN samples, the road to improving constraints on H_0 is more challenging. There are a limited number of SNe Ia that will explode in the near future within a $\sim$40 Mpc radius, a constraint due to HST discovery limits of Cepheids. At roughly one SN Ia per year, it will take several decades to double the current sample of 42 SNe calibrated by SH0ES Cepheid hosts. Fortunately, we find that the systematics in the measurement of H_0 from the SNe are at a scale of $0.3\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, as shown in Figure 13. This is consistent with the general finding of Brownsberger et al. (2021), who showed how robust H_0 is to systematic uncertainties in comparison to the relatively calibration-sensitive constraints of w_0 or Ω_M. Lastly, there is ongoing work that combines the progress used here by Peterson et al. (2022) and applies it to a "two-rung" distance-ladder analysis, in which SNe are excluded from the distance ladder (Kenworthy et al. 2022).

6. Conclusion

This work is the culmination of a number of supporting analyses as part of the Pantheon+ effort. In this work, we summarize the various inputs and analyses required to combine the supporting works and ultimately measure distances and cosmological parameters. For the first time, we are able to measure the cosmic expansion history and the local distance

ladder H_0 simultaneously. We combine our results with additional external probes. Importantly, we release a number of data and analysis products to facilitate reproducing our work by the community. This includes a joint covariance of SNe used for measurements of H_0 and w.

For our main results, we find $\Omega_M = 0.334 \pm 0.018$ in flat ΛCDM from SNe Ia alone. For a flat w_0CDM model, we measure $w_0 = -0.90 \pm 0.14$ from SNe Ia alone and $w_0 = -0.978^{+0.024}_{-0.031}$ when combining SNe with constraints on the CMB and all BAO; both are consistent with a cosmological-constant model of dark energy. We also present the most precise measurements to date on the evolution of dark energy in a flat w_0w_aCDM universe, and measure $w_a = -0.1^{+0.9}_{-2.0}$ from Pantheon+ alone and $w_a = -0.65^{+0.28}_{-0.32}$ when combining with CMB and BAO data. Finally, while nominal constraints on H_0 are presented in a companion paper by the SH0ES team (R22), we perform joint constraints of H_0 with expansion history and find $H_0 = 73.5 \pm 1.1$ in flat wCDM, and we show how systematic uncertainties in measurements of the SN component of the distance ladder cannot account for the current level of the "Hubble tension."

D.S., D.B., and A.R. thank the John Templeton Foundation for their support of grant No. 62314. D.B. acknowledges support for this work provided by NASA through NASA Hubble Fellowship grant HST-HF2-51430.001 awarded by the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS5-26555. D.S. is supported by DOE grant DE-SC0010007, the David and Lucile Packard Foundation, and NASA under contract No. NNG17PX03C issued through the WFIRST Science Investigation Teams Programme. We acknowledge the generous support of Marc J. Staley, whose fellowship partly funded B.E.S. while contributing to the work presented herein as a graduate student. A.V.F. is grateful for support from the TABASGO Foundation, the Christopher R. Redlich Fund, the U.C. Berkeley Miller Institute for Basic Research in Science (in which he was a Miller Senior Fellow), and many individual donors. S.N. thanks the STFC Ernest Rutherford Fellowship for support via grant ST/T005009/1. L.K. thanks the UKRI Future Leaders Fellowship for support through the grant MR/T01881X/1. This work was completed in part with resources provided by the University of Chicagos Research Computing Center. The Katzman Automatic Imaging Telescope (with which the LOSS samples were obtained) and its ongoing operation were made possible by donations from Sun Microsystems, Inc., the Hewlett-Packard Company, AutoScope Corporation, Lick Observatory, the NSF, the University of California, the Sylvia & Jim Katzman Foundation, and the TABASGO Foundation. Research at Lick Observatory is partially supported by a generous gift from Google.

Simulations, light-curve fitting, BBC, and cosmology pipeline are managed by `PIPPIN` (Hinton & Brout 2020). Contours and parameter constraints are generated using the CHAINCONSUMER package (Hinton 2016). Plots are generated with Matplotlib (Hunter 2007). We use astropy (Price-Whelan et al. 2018), SciPy (Virtanen et al. 2020), and NumPy (Oliphant 2006). Analysis and visualizations provided in part by https://github.com/bap37/Midwayplotter.

D.B. thanks his spouse Isabella and their future daughter for their support, as the due date is rapidly approaching!

Appendix A
Additional Formalism for Distance and Uncertainty Estimation

As shown in BS21, SN Ia scatter has both a color and host-mass dependence (increasing scatter) and a redshift dependence that arises from selection effects (decreasing scatter). In this work we introduce a new method of accounting for the uncertainties using the scatter model predictions. We include $\sigma_{\rm scat}(z,\ c,\ M_\star)$ from simulations as an additive uncertainty inside Equation (3) rather than the multiplicative uncertainty $f(z,\ c,\ M_\star)$ on the computed $\sigma_{\rm meas}$ that has been used in past analyses. The $\sigma_{\rm scat}(z,\ c,\ M_\star)$ term is computed from simulations that use the choice of scatter model. The BBC process, after correcting distances for selection effects, determines the magnitude of $\sigma_{\rm scat}(z,\ c,\ M_\star)$ in each $z,\ c,\ M_\star$ bin by requiring that the observed-simulated distance reduced χ^2 in each bin is unity. If the simulations using a model of intrinsic scatter fully describe the observed scatter in the data, the uncertainty modeling term in Equation (3), $\sigma_{\rm scat}(z,\ c,\ M_\star)$, will cause $\sigma_{\rm gray}$ to be 0.

In the case of the decrease in observed scatter at high redshift arising from only intrinsically bright/blue events being selected at the limits of the telescope (Kessler et al. 2015), we instead apply it as a downscaling of $f(z,\ c,\ M_\star)$ of the reported measurement uncertainty and set $\sigma_{\rm scat}(z,\ c,\ M_\star) = 0$. Conversely, for bins of z, c, and $M_\star$ with χ^2 greater than unity, the necessary $\sigma_{\rm scat}(z,\ c,\ M_\star)$ is applied, and f is set to 1. The resulting $f(z,\ c,\ M_\star)$ and $\sigma_{\rm scat}(z,\ c,\ M_\star)$ found from the simulations are applied to the Pantheon+ data.

The method and dimensionality for the application of bias corrections are dependent on the adopted scatter model. Table 5 summarizes the differences between the two main methods used in this work, the first of which is applied when assuming the BS21/P21 scatter model, and the other when assuming the G10 or C11 scatter model. The main difference between these groups of scatter models, as discussed in Section 3, is whether the intrinsic scatter is driven by diversity in the reddening ratios R_V of the light curves, which affects the application of bias corrections. For both analysis paths, we follow the methodology introduced by Popovic et al. (2021b).

Table 5
Distance Bias (and Uncertainty) Estimation for Scatter Models

	G10/C11	BS21/P21
Dimensionality	7D (z, x_1, c, $M_\star$, γ, α, β)	4D (z, x_1, c, $M_\star$)
Mass-step correction	γ a fitted parameter	γ corrected for within $\delta_{\rm bias}$ (γ and $\delta_{\rm host}$ consistent with zero)
Intrinsic Scatter Floor	$\sigma^2_{\rm floor} = \sigma^2_{\rm gray}$	$\sigma^2_{\rm floor} = \sigma^2_{\rm scat}(z_i, c_i, M_\star) + \sigma^2_{\rm gray}$, applied when $f(z, c, M_\star) > 1$
Selection Effects	$f(z, c)$	$f(z, c, M_\star) \leqslant 1$, applied when $\sigma^2_{\rm scat}(z_i, c_i, M_\star) = 0$

Note. The formalism for four-dimensional and seven-dimensional bias corrections is described by Popovic et al. (2021b) that depend on the intrinsic scatter model assumed—either G10/C11 or BS21/P21. The statistical and intrinsic scatter uncertainties from Equation (3) are shown here; the other uncertainty components from Equation (3) are independent of the scatter model.

Appendix B
Binned Systematic Error Budget

In Table 6 we show a systematic error budget that is nearly identical to what was performed in Table 4, except that the data set (ΔD) and covariance matrix ($C_{\rm stat+syst}$) are binned in 20 redshift bins. This error budget is similar to the methodology performed in the most recent SN cosmology analyses where binned covariance matrices were used (e.g., Pantheon and DES3YR; Brout et al. 2019b) and where smoothed data vectors and matrices (which were shown to be equivalent to binned) were used (JLA). The total systematic error when binning is a factor of 1.5 larger (0.029) than when not binning the data set (0.019).

Systematics that improve the most with unbinned matrices are those with smaller $\sigma w_{\rm sys}^{\rm unbinned}/\sigma w_{\rm sys}^{\rm binned}$. Binned analyses collapse valuable information in the Hubble diagram down to a single dimension, redshift. We find that as expected, the redshift bias systematic does not improve much at all. This is because systematics that only exhibit redshift dependence are degenerate with cosmological model parameters and cannot be self-constrained by the data as easily. Systematics that exhibit dependence in other parameters (such as SN color) can be drastically reduced in SN Ia cosmological parameter error budgets when not performing binned analyses.

Table 6
Comparison of Binned and Unbinned Systematic Error Budgets

Description	[a]$\sigma w_{\rm sys}^{\rm binned}$	[a]$\sigma w_{\rm sys}^{\rm unbinned}$	$\sigma w_{\rm sys}^{\rm unbinned}/\sigma w_{\rm sys}^{\rm binned}$
All Systematics	0.029	0.019	0.66
Calibration			
SALT2 Train and ZPT[b]	0.019	0.009	0.47
SALT2 Method	0.009	0.008	0.88
CSP Tertiary Stars	0.005	0.003	0.60
[d]HST	0.002	0.003	1.50
Redshifts			
[c]$v_{\rm pec}$ Map	N/A	0.002	N/A
Redshift Bias	0.012	0.011	0.92
Astrophysics			
Intrinsic Variations	0.009	0.002	0.18
MW $E(B-V)$	0.012	0.008	0.67
MW Color Law	0.007	0.006	0.86
Mass Step	0.001	0.001	1.00
Modeling			
Selection Efficiency	0.008	0.004	0.50
Populations	0.011	0.000	0.00

Notes.
[a] Constraints are combined with Planck prior.
[b] ZPT denotes light-curve fitting zero-points.
[c] Due to implementation methodology of this systematic, it has not been performed in the binned case.
[d] The increase in the "HST" systematic is likely due to noise, as the values are very small for both binned and unbinned.

Appendix C
Products

The following data products that are provided in part by the full suite of Pantheon+ supporting papers are released publicly in machine-readable format[39] at pantheonplussh0es.github.io and as part of SNANA and CosmoSIS (where noted).

1. Light-curve Photometry, Redshifts, and Host-galaxy Properties; from S22 and Carr et al. (2021)
2. Trained SALT2-B22 Model; from Brout et al. (2022)
3. SALT2 Fit Parameters; from S22
4. Ten Catalog Level Simulations of Pantheon+ Light-curve Fit Parameters; this work
5. SN/Host Redshifts and Peculiar Velocities; from Carr et al. (2021)
6. SN Distance Modulii and Redshifts; this work; Carr et al. (2021); see Table 7 [40]
7. SN Distance Covariance; this work
8. Cepheid Host Distances; from R22
9. Cepheid Host Distance Covariance; from R22
10. SN Ia + Cepheid Host Cosmology Likelihood; this work
11. SN Cosmology Chains; this work

Table 7
Pantheon+ Hubble Diagram

CID	Survey	$z_{\rm HD}$	$\sigma_{z\rm HD}$	$z_{\rm CMB}$	$z_{\rm HEL}$	$m_B^{\rm corr}$	$\sigma_{m_B{\rm corr}}^{\rm diag}$	c	σ_c	x_1	σ_{x_1}	m_B	σ_{m_B}
2011fe	LOSS2	0.00122	0.00084	0.00122	0.00082	9.746	1.516	−0.108	0.040	−0.548	0.134	9.584	0.033
2011fe	SOUSA	0.00122	0.00084	0.00122	0.00082	9.803	1.517	−0.033	0.038	−0.380	0.086	9.784	0.035
2012cg	LOSS2	0.00256	0.00084	0.00256	0.00144	11.470	0.782	0.101	0.018	0.492	0.024	11.816	0.024
2012cg	SOUSA	0.00256	0.00084	0.00256	0.00144	11.492	0.799	0.122	0.039	0.713	0.084	11.880	0.036
1994DRichmond	LOWZ	0.00299	0.00084	0.00299	0.00187	11.523	0.881	−0.112	0.026	−1.618	0.050	11.533	0.032
1981B	LOWZ	0.00317	0.00084	0.0035	0.00236	11.542	0.614	−0.005	0.031	−0.445	0.165	11.664	0.034
2013aa	SOUSA	0.00331	0.00085	0.00478	0.00411	11.207	0.594	−0.104	0.054	0.513	0.152	10.891	0.106
2013aa	CSP	0.00331	0.00085	0.00478	0.00411	11.300	0.580	−0.158	0.036	0.633	0.139	10.844	0.100
2017cbv	CSP	0.00331	0.00085	0.00478	0.00411	11.148	0.578	−0.126	0.032	0.617	0.053	10.773	0.094
2017cbv	CNIa0.02	0.00331	0.00085	0.00478	0.00411	11.258	0.578	−0.096	0.035	0.819	0.066	10.914	0.099

(This table is available in its entirety in machine-readable form.)

[39] Will be made available after publication.

[40] $\sigma_{m_B{\rm corr}}^{\rm diag}$ in Table 7 is the error on standardized magnitude from the diagonal of the covariance matrix. It is for plotting purposes only and not to be used for cosmological fits.

ORCID iDs

Dillon Brout https://orcid.org/0000-0001-5201-8374
Brodie Popovic https://orcid.org/0000-0002-8012-6978
Adam G. Riess https://orcid.org/0000-0002-6124-1196
Anthony Carr https://orcid.org/0000-0003-4074-5659
Rick Kessler https://orcid.org/0000-0003-3221-0419
Tamara M. Davis https://orcid.org/0000-0002-4213-8783
Samuel Hinton https://orcid.org/0000-0003-2071-9349
David Jones https://orcid.org/0000-0002-6230-0151
W. D'Arcy Kenworthy https://orcid.org/0000-0002-5153-5983
Erik R. Peterson https://orcid.org/0000-0001-8596-4746
Khaled Said https://orcid.org/0000-0002-1809-6325
Patrick Armstrong https://orcid.org/0000-0003-1997-3649
Antonella Palmese https://orcid.org/0000-0002-6011-0530
Helen Qu https://orcid.org/0000-0003-1899-9791
Benjamin M. Rose https://orcid.org/0000-0002-1873-8973
Bruno Sanchez https://orcid.org/0000-0002-8687-0669
Christopher W. Stubbs https://orcid.org/0000-0003-0347-1724
Charlotte M. Wood https://orcid.org/0000-0003-4773-4602
Peter J. Brown https://orcid.org/0000-0001-6272-5507
Rebecca Chen https://orcid.org/0000-0003-3917-0966
Ken Chambers https://orcid.org/0000-0001-6965-7789
David A. Coulter https://orcid.org/0000-0003-4263-2228
Mi Dai https://orcid.org/0000-0002-5995-9692
Georgios Dimitriadis https://orcid.org/0000-0001-9494-179X
Alexei V. Filippenko https://orcid.org/0000-0003-3460-0103
Ryan J. Foley https://orcid.org/0000-0002-2445-5275
Saurabh W. Jha https://orcid.org/0000-0001-8738-6011
Robert P. Kirshner https://orcid.org/0000-0002-1966-3942
Anais Möller https://orcid.org/0000-0001-8211-8608
Seshadri Nadathur https://orcid.org/0000-0001-9070-3102
Yen-Chen Pan https://orcid.org/0000-0001-8415-6720
Armin Rest https://orcid.org/0000-0002-4410-5387
Cesar Rojas-Bravo https://orcid.org/0000-0002-7559-315X
Masao Sako https://orcid.org/0000-0003-2764-7093
Matthew R. Siebert https://orcid.org/0000-0003-2445-3891
Mat Smith https://orcid.org/0000-0002-3321-1432
Benjamin E. Stahl https://orcid.org/0000-0002-3169-3167

References

Abbott, T. M. C., Allam, S., Andersen, P., et al. 2019, ApJL, 872, L30
Alam, S., Ata, M., Bailey, S., et al. 2017, MNRAS, 470, 2617
Andrade, U., Bengaly, C. A. P., Santos, B., & Alcaniz, J. S. 2018, ApJ, 865, 119
Astier, P., Guy, J., Regnault, N., et al. 2006, A&A, 447, 31
Bautista, J. E., Paviot, R., Vargas Magaña, M., et al. 2020, MNRAS, 500, 736
Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22
Bohlin, R. C., Hubeny, I., & Rauch, T. 2020, AJ, 160, 21
Brout, D., Hinton, S. R., & 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., Taylor, G., Scolnic, D., et al. 2022, ApJ, 938, 111
Brown, P. J., Breeveld, A. A., Holland, S., Kuin, P., & Pritchard, T. 2014, Ap&SS, 354, 89
Brownsberger, S., Brout, D., Scolnic, D., Stubbs, C. W., & Riess, A. G. 2021, arXiv:2110.03486
Brownsberger, S. R., Stubbs, C. W., & Scolnic, D. M. 2019, ApJ, 875, 34
Burns, C. R., Ashall, C., Contreras, C., et al. 2020, ApJ, 895, 118
Burns, C. R., Parent, E., Phillips, M. M., et al. 2018, ApJ, 869, 56
Calcino, J., & Davis, T. 2017, JCAP, 2017, 038
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Carr, A., Davis, T. M., Scolnic, D., et al. 2021, arXiv:2112.01471
Carrick, J., Turnbull, S. J., Lavaux, G., & Hudson, M. J. 2015, MNRAS, 450, 317
Chen, P., Dong, S., Kochanek, C. S., et al. 2020, arXiv:2011.02461
Chen, R., Scolnic, D., Rozo, E., et al. 2022, arXiv:2202.10480
Childress, M., Aldering, G., Antilogus, P., et al. 2013a, ApJ, 770, 108
Childress, M., Aldering, G., Antilogus, P., et al. 2013b, ApJ, 770, 107
Chotard, N. E. A. 2011, A&A, 529, L4
Colgáin, E. Ó. 2019, JCAP, 2019, 006
Colin, J., Mohayaee, R., Rameez, M., & Sarkar, S. 2019, A&A, 631, L13
Conley, A., Guy, J., Sullivan, M., et al. 2011, ApJS, 192, 1
Cruz, M., Tucci, M., Martínez-González, E., & Vielva, P. 2006, MNRAS, 369, 57
D'Andrea, C., Smith, M., & Sullivan, M. 2018, AAS Meeting, 231, 245.10
Davis, T. M., Hinton, S. R., Howlett, C., & Calcino, J. 2019, MNRAS, 490, 2948
Dhawan, S., Brout, D., Scolnic, D., et al. 2020, ApJ, 894, 54
Dhawan, S., Goobar, A., Smith, M., et al. 2022, MNRAS, 510, 2228
du Mas des Bourboux, H., Rich, J., Font-Ribera, A., et al. 2020, ApJ, 901, 153
Fitzpatrick, E. L. 1999, PASP, 111, 63
Foley, R. J., Scolnic, D., Rest, A., et al. 2018, MNRAS, 475, 193
Gall, C., Stritzinger, M. D., Ashall, C., et al. 2018, A&A, 611, A58
Ganeshalingam, M., Li, W., Filippenko, A. V., et al. 2010, ApJS, 190, 418
Gilliland, R. L., Nugent, P. E., & Phillips, M. M. 1999, ApJ, 521, 30
Gurzadyan, V. G., Kashin, A. L., Khachatryan, H., et al. 2014, A&A, 566, A135
Guy, J., Sullivan, M., Conley, A., et al. 2010, A&A, 523, A7
Handley, W. J., Hobson, M. P., & Lasenby, A. N. 2015, MNRAS, 453, 4385
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
Hinton, S., Kim, A., Davis, T., et al. 2019, ApJ, 876, 15
Hinton, S. R. 2016, JOSS, 1, 00045
Hogg, D. W. 1999, arXiv:astro-ph/9905116
Hou, J., Sánchez, A. G., Ross, A. J., et al. 2020, MNRAS, 500, 1201
Hounsell, R., Scolnic, D., Foley, R. J., et al. 2018, ApJ, 867, 23
Hu, J. P., Wang, Y. Y., & Wang, F. Y. 2020, A&A, 643, A93
Hunter, J. D. 2007, CSE, 9, 90
Huterer, D., Shafer, D. L., Scolnic, D. M., & Schmidt, F. 2017, JCAP, 2017, 015
Jennings, E., Wolf, R., & Sako, M. 2016, arXiv:1611.03087
Jha, S., Kirshner, R. P., Challis, P., et al. 2006, AJ, 131, 527
Jones, D. O., Foley, R. J., Narayan, G., et al. 2021, ApJ, 908, 143
Jones, D. O., Riess, A. G., Scolnic, D. M., et al. 2018a, ApJ, 867, 108
Jones, D. O., Scolnic, D. M., Foley, R. J., et al. 2019, ApJ, 881, 19
Jones, D. O., Scolnic, D. M., Riess, A. G., et al. 2018b, ApJ, 857, 51
Jönsson, J., Sullivan, M., Hook, I., et al. 2010, MNRAS, 405, 535
Kawabata, M., Maeda, K., Yamanaka, M., et al. 2020, ApJ, 893, 143
Kelsey, L., Sullivan, M., Smith, M., et al. 2021, MNRAS, 501, 4861
Kenworthy, W. D., Riess, A. G., Scolnic, D., et al. 2022, ApJ, 935, 83
Kessler, R., Bernstein, J. P., Cinabro, D., et al. 2009, PASP, 121, 1028
Kessler, R., Brout, D., Crawford, S., 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
Kovács, A., Jeffrey, N., Gatti, M., et al. 2021, MNRAS, 510, 216
Krisciunas, K., Contreras, C., Burns, C. R., et al. 2017b, AJ, 154, 211
Krisciunas, K., Suntzeff, N. B., Espinoza, J., et al. 2017a, RNAAS, 1, 36
Lampeitl, H., Smith, M., Nichol, R. C., et al. 2010, ApJ, 722, 566
Lewis, A., & Bridle, S. 2002, PhRvD, 66, 103511
Lilow, R., & Nusser, A. 2021, MNRAS, 507, 1557
Mathews, G. J., Rose, B. M., Garnavich, P. M., Yamazaki, D. G., & Kajino, T. 2016, ApJ, 827, 60
Milne, P. A., Brown, P. J., Roming, P. W. A., et al. 2010, ApJ, 721, 1627
Narayan, G., Matheson, T., Saha, A., et al. 2019, ApJS, 241, 20
Oliphant, T. E. 2006, A Guide to NumPy, Vol. 1 (USA: Trelgol Publishing)
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Peterson, E. R., Kenworthy, W. D., Scolnic, D., et al. 2022, ApJ, 938, 112
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
Popovic, B., Brout, D., Kessler, R., & Scolnic, D. 2021a, arXiv:2112.04456
Popovic, B., Brout, D., Kessler, R., Scolnic, D., & Lu, L. 2021b, ApJ, 913, 49
Price-Whelan, A. M., Sipőcz, B., Günther, H., et al. 2018, AJ, 156, 123
Regnault, N., Guyonnet, A., Schahmanèche, K., et al. 2015, A&A, 581, A45
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009

Riess, A. G., Kirshner, R. P., Schmidt, B. P., et al. 1999, AJ, 117, 707
Riess, A. G., Macri, L. M., Hoffmann, S. L., et al. 2016, ApJ, 826, 56
Riess, A. G., Nugent, P. E., Gilliland, R. L., et al. 2001, ApJ, 560, 49
Riess, A. G., Rodney, S. A., Scolnic, D. M., et al. 2018, ApJ, 853, 126
Riess, A. G., Strolger, L.-G., Casertano, S., et al. 2007, ApJ, 659, 98
Riess, A. G., Strolger, L.-G., Tonry, J., et al. 2004, ApJ, 607, 665
Riess, A. G., Yuan, W., Macri, L. M., et al. 2022, ApJL, 934, L7
Rigault, M., Copin, Y., Aldering, G., et al. 2013, A&A, 560, A66
Rose, B. M., Garnavich, P. M., & Berg, M. A. 2019, ApJ, 874, 32
Ross, A. J., Samushia, L., Howlett, C., et al. 2015, MNRAS, 449, 835
Rubin, D., Aldering, G., Barbary, K., et al. 2015, ApJ, 813, 137
Rudnick, L., Brown, S., & Williams, L. R. 2007, ApJ, 671, 40
Said, K., Colless, M., Magoulas, C., Lucey, J. R., & Hudson, M. J. 2020, MNRAS, 497, 1275
Sako, M., Bassett, B., Connolly, B., et al. 2011, ApJ, 738, 162
Sánchez, B., Kessler, R., Scolnic, D., et al. 2022, ApJ, 934, 96
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
Schlegel, D. J., Finkbeiner, D. P., & Davis, M. 1998, ApJ, 500, 525
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., Kessler, R., Brout, D., et al. 2018a, ApJL, 852, L3
Scolnic, D., Perlmutter, S., Aldering, G., et al. 2019, Astro2020: Decadal Survey on Astronomy and Astrophysics, 270
Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018b, ApJ, 859, 101
Shariff, H., Jiao, X., Trotta, R., & van Dyk, D. A. 2016, ApJ, 827, 1
Smith, M., D'Andrea, C. B., Sullivan, M., et al. 2020a, AJ, 160, 267
Smith, M., D'Andrea, C. B., Sullivan, M., et al. 2020b, AJ, 160, 267
Soltis, J., Farahi, A., Huterer, D., & Liberato, C. M. 2019, PhRvL, 122, 091301
Stahl, B. E., de Jaeger, T., Boruah, S. S., et al. 2021, MNRAS, 505, 2349
Stahl, B. E., Zheng, W., de Jaeger, T., et al. 2019, MNRAS, 490, 3882
Stritzinger, M., Burns, C. R., Phillips, M. M., et al. 2010, AJ, 140, 2036
Stritzinger, M. D., Shappee, B. J., Piro, A. L., et al. 2018, ApJ, 864, L35
Stubbs, C. W., & Brown, Y. J. 2015, MPLA, 30, 1530030
Sullivan, M., Conley, A., Howell, D. A., et al. 2010, MNRAS, 406, 782
Suzuki, N., Rubin, D., Lidman, C., et al. 2012, ApJ, 746, 85
Szapudi, I., Kovács, A., Granett, B., et al. 2015, MNRAS, 450, 288
Taylor, G., Lidman, C., Tucker, B. E., et al. 2021, MNRAS, 504, 4111
The LSST Dark Energy Science Collaboration, Mandelbaum, R., Eifler, T., et al. 2018, arXiv:1809.01669
Tripp, R. 1998, A&A, 331, 815
Tsvetkov, D. Y., & Elenin, L. 2010, PZ, 30, 2
Tully, R. B. 2015, AJ, 149, 171
Vielva, P., Martínez-González, E., Barreiro, R. B., Sanz, J. L., & Cayón, L. 2004, ApJ, 609, 22
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, NatMe, 17, 261
Zhang, T., Wang, X., Li, W., et al. 2010, PASP, 122, 1
Zuntz, J., Paterno, M., Jennings, E., et al. 2015, A&C, 12, 45