CrossMark

It's Dust: Solving the Mysteries of the Intrinsic Scatter and Host-galaxy Dependence of Standardized Type Ia Supernova Brightnesses

Dillon Brout[1,3] and Daniel Scolnic[2]
[1] Department of Physics and Astronomy, University of Pennsylvania, Philadelphia, PA 19104, USA; dbrout@physics.upenn.edu
[2] Department of Physics, Duke University, Durham, NC 27708, USA; daniel.scolnic@duke.edu

Abstract

The use of Type Ia supernovae (SNe Ia) as cosmological tools has motivated significant effort to understand what drives the intrinsic scatter of SN Ia distance modulus residuals after standardization, characterize the distribution of SN Ia colors, and explain why properties of the host galaxies of the SNe correlate with SN Ia distance modulus residuals. We use a compiled sample of $\sim$1450 spectroscopically confirmed photometric light curves of SNe Ia and propose a solution to these three problems simultaneously that also explains an empirical 11σ detection of the dependence of Hubble residual scatter on SN Ia color. We introduce a physical model of color where intrinsic SN Ia colors with a relatively weak correlation with luminosity are combined with extrinsic dust-like colors ($E(B - V)$) with a wide range of extinction parameter values (R_V). This model captures the observed trends of Hubble residual scatter and indicates that the dominant component of SN Ia intrinsic scatter is variation in R_V. We also find that the recovered $E(B - V)$ and R_V distributions differ based on global host-galaxy stellar mass, and this explains the observed correlation (γ) between mass and Hubble residuals seen in past analyses, as well as an observed 4.5σ dependence of γ on SN Ia color. This finding removes any need to ascribe different intrinsic luminosities to different progenitor systems. Finally, we measure biases in the equation of state of dark energy (w) up to $|\Delta w| = 0.04$ by replacing previous models of SN color with our dust-based model; this bias is larger than any systematic uncertainty in previous SN Ia cosmological analyses.

Unified Astronomy Thesaurus concepts: Type Ia supernovae (1728); Cosmology (343); Interstellar dust (836); Circumstellar dust (236); Dark energy (351); Accelerating universe (12); Supernovae (1668)

1. Introduction

Studies in the last decade of research in cosmology with Type Ia supernovae (SNe Ia) have forewarned that the measurements of the equation of state of dark energy w will soon hit a systematic floor. Yet, such measurements (B14: Betoule et al. 2014, S18: Scolnic et al. 2018, B19b: Brout et al. 2019b; Jones et al. 2019) continually reach better levels of both statistical and systematic precision. This is due to the improvement of systematic uncertainties in survey and camera design, but also due to the possibility afforded by significantly larger samples to understand the systematics in the analysis. In the most recent analyses (S18, B19b), it has been found that systematic uncertainties in understanding the intrinsic scatter of standardized SN Ia brightnesses are at a similar level as or larger than uncertainties due to external photometric calibration. As calibration uncertainties have been dominant in past systematic error budgets, this moment marks a transition from a need to understand external issues independent of the supernovae to a need to also better understand SN Ia physics.

With current cosmological analyses of SNe Ia requiring millimag-level control of systematics, uncertainty over how to understand the intrinsic scatter of standardized SN Ia brightnesses, which is on the 0.1 mag level, is problematic. Practically, intrinsic scatter is measured as the excess scatter of SN Ia distance residuals to a best-fit cosmology after accounting for measurement noise. A holistic understanding of SN Ia intrinsic scatter and its underlying characterization has remained elusive, but its size has been found to depend on a wide variety of measurement components: redshift (e.g., B14),

wavelength range of the photometric observations (e.g., Mandel et al. 2011), host-galaxy properties (e.g., Uddin et al. 2017), and spectroscopic features (e.g., Fakhouri et al. 2015). Furthermore, Scolnic & Kessler (2016) showed that the relative amounts of chromatic versus achromatic components of the intrinsic scatter models were directly linked to the intrinsic SN Ia color population and reddening law; however, this study was unable to discriminate between different models.

After the discovery of the accelerating universe (Riess et al. 1998; Perlmutter et al. 1999), there were two commonly used light-curve fitters, MLCS2k2 (Jha et al. 2007) and SALT2 (Guy et al. 2010), that diverged in their approach to color and intrinsic scatter. MLCS2k2 attempted to model color based on dust with the possibility that each SN could have its own extinction law and assumed that a large amount of the intrinsic scatter was in color. The SALT2 model, on the other hand, was agnostic to any physical properties of the SN color and its relation to the intrinsic scatter. Cosmological analyses have since favored the SALT2 model because of its native spectral model to account for k-corrections and updated calibration, and it has been used in most recent cosmology analyses, including the Joint Light-Curve Analysis (JLA, B14), Pantheon (S18), the Dark Energy Survey 3 Year Sample (DES3YR, Brout et al. 2019a, B19b), and the Foundation + Pan-STARRS1 photometric analysis (Jones et al. 2019). However, despite the fact that MLCS2k2 has not been used in recent cosmological analyses, papers such as Scolnic et al. (2014b, 2018) and Mandel et al. (2017) have attempted to bridge the gap between SALT2 and MLCS2k2 methods by modeling a connection between the underlying population of color, dust, and reddening laws.

---

[3] NASA Einstein Fellow.

Still, SN Ia analyses that attempt to model dust using a cosmological sample have typically made the simplistic assumption that there is a single total-to-selective extinction parameter, R_V, that can be fixed at a single number. The parameter R_V is defined as $A_V/(A_B - A_V)$, where A_V is the extinction in the V ($\lambda V \sim 5500$ Å) band, and A_B is the extinction in the blue ($\lambda V \sim 4400$ Å) band. As R_V varies for different dust grain sizes and compositions, and galaxies have different dust properties, it is well known that different galaxies and different regions within galaxies exhibit a wide range of R_V values. In fact, while the Milky Way galaxy has an R_V on average of ~ 3.1, it has a distribution of at least $\sigma_{R_V} = 0.2$ (Schlafly et al. 2016). Additionally, different parts of the LMC and SMC have been found to have R_V values with a range of $R_V \sim 2$–5 (Gao et al. 2013; Yanchulova Merica-Jones et al. 2017). Furthermore, Salim et al. (2018) study the dust attenuation curves of 230,000 individual galaxies in the local universe, using GALEX, SDSS, and WISE photometry calibrated on the Herschel ATLAS, and they find that quiescent galaxies, which are typically high in mass, have a mean $R_V = 2.61$, and star-forming galaxies, which are lower in mass on average, have a mean $R_V = 3.15$.

Note that R_V has also been measured through large SN sample statistics and detailed studies of individual SNe, though often with varying sets of assumptions. Cikota et al. (2016) compiled 13 various studies of SN Ia samples from the literature, which determined a range of R_V values from ~ 1 to ~ 3.5. Cikota et al. (2016) determined R_V from nearby SNe, and for 21 SNe Ia observed in Sab–Sbp galaxies and 34 SNe in Sbc–Scp, they find $R_V = 2.71 \pm 1.58$ and $R_V = 1.70 \pm 0.38$, respectively. While so many past analyses have recovered $R_V < 2$ for studies of individual SNe (e.g., Wang et al. 2005; Krisciunas et al. 2006), these were often SNe Ia with high $E(B-V)$, and it was postulated that R_V may decrease with $E(B-V)$. However, Nobili & Goobar (2008) found from a sample of modestly reddened ($E(B-V) < 0.25$ mag) SNe Ia a small value of $R_V \sim 1$, and more recently, Amanullah et al. (2015) analyzed high-quality UV-NIR spectra of six SNe and found that SNe with high reddening indicated R_V values ranging from ~ 1.4 to ~ 2.8, and SNe with low amounts of reddening also indicated R_V values of ~ 1.4 and ~ 2.8. Importantly, Amanullah et al. (2015) stressed that the observed diversity in R_V is not accounted for in analyses that measure the cosmological expansion of the universe.

Since the low R_V values (<2) are not found in studies of the Milky Way, this has motivated various SN Ia studies to ascribe the dust to circumstellar dust around the progenitor at the time of the explosion (Wang 2005; Goobar 2008). However, an alternative interpretation could be that the low R_V values are caused by dust in the interstellar medium (Phillips et al. 2013). This understanding has been supported by Bulla et al. (2018a, 2018b), who constrained the location of the dust that caused the reddening in the SN Ia spectra to be, for the majority of the SNe that they observed, on scales of the interstellar medium, rather than the circumstellar surroundings. This could be due to cloud–cloud collisions induced by the SN radiation pressure (Hoang 2017), which produce small dust grains (Gao et al. 2015; Nozawa 2016).

While accounting for dust remains a challenge for current and future photometry cosmology analyses, this pursuit has often been done in parallel to the search for correlations between measured supernova luminosity after standardization and host-galaxy properties. Global and local properties of SN Ia

host galaxies such as stellar mass, star formation rate (SFR), stellar population age, and metallicity have all been shown to correlate with the distance modulus residuals after standardization (Hicken et al. 2009a; Lampeitl et al. 2010; Sullivan et al. 2010; Childress et al. 2013a; Rose et al. 2019). This correlation is often parameterized as a step function in host-galaxy stellar mass and is now commonplace in SN Ia cosmology analyses despite the lack of understanding of its physical underpinning or convincing evidence for exactly which host-galaxy property is most influential on SN Ia luminosity (e.g., Jones et al. 2018a; Scolnic et al. 2020). To explain this correlation, recent studies have suggested a potential relation between the luminosity of the SN and the progenitor, which can be related to the age of the galaxy or the local environment of the galaxy (Childress et al. 2013b; Rigault et al. 2013; Roman et al. 2018). However, as the aforementioned galaxy properties are all directly linked to dust properties, it is likely that the lack of dust modeling in SN Ia cosmology is related to the correlations between host-galaxy properties and standardized luminosities.

In this analysis, we show that there are clear limitations in SN Ia standardization techniques with a single color–luminosity correlation, but that these limitations can be addressed by inclusion of dust modeling with variation in R_V. This paper relies heavily on the work of Mandel et al. (2017), which follows closer to the framework of MLCS2k2 and developed a hierarchical Bayesian model to build a more rich understanding of SN color. Mandel et al. (2017) only used low-redshift data, did not account for selection effects, and assumed a fixed R_V extinction parameter; here we use a much larger data set across a wide redshift range and use survey simulations to forward-model what is done in Mandel et al. (2011), though with additional features to explain the discrepancies seen between simulations and data.

In Section 2, we present the data compilation, light-curve fitting and discrepancies between the data, and a simple understanding of SN color. In Section 3, we discuss how to differentiate between past models of SN color and our new dust-based color model. In Section 4, we show how the new model can explain the commonly seen correlation between distance modulus residuals and host-galaxy properties. In Section 5, we assess the impact on recovered cosmological parameters, and in Sections 6 and 7, we discuss further studies and conclusions.

2. Data Sample, Distance Moduli, and Description of SN Ia Colors

2.1. Data

We use a compilation of publicly available, spectroscopically classified photometric light curves of SNe Ia that have been used in past cosmological analyses and that have been calibrated to the SuperCal system (Scolnic et al. 2015). The low-redshift (low-z) SNe used here are made up of, in part, those used in B19b, which are from CSP (Stritzinger et al. 2010) and CfA3-4 (Hicken et al. 2009a, 2009b, 2012). At low z, we also include the recently released 180 low-z SNe from the Foundation sample (Foley et al. 2018). At high z, we include SNe from PS1 (Rest et al. 2014; Scolnic et al. 2018), SDSS (Sako et al. 2011), and SNLS (B14) as was done in the Pantheon analysis. Finally, we include data from the recently released DES 3 yr sample (Brout et al. 2019a), hereafter
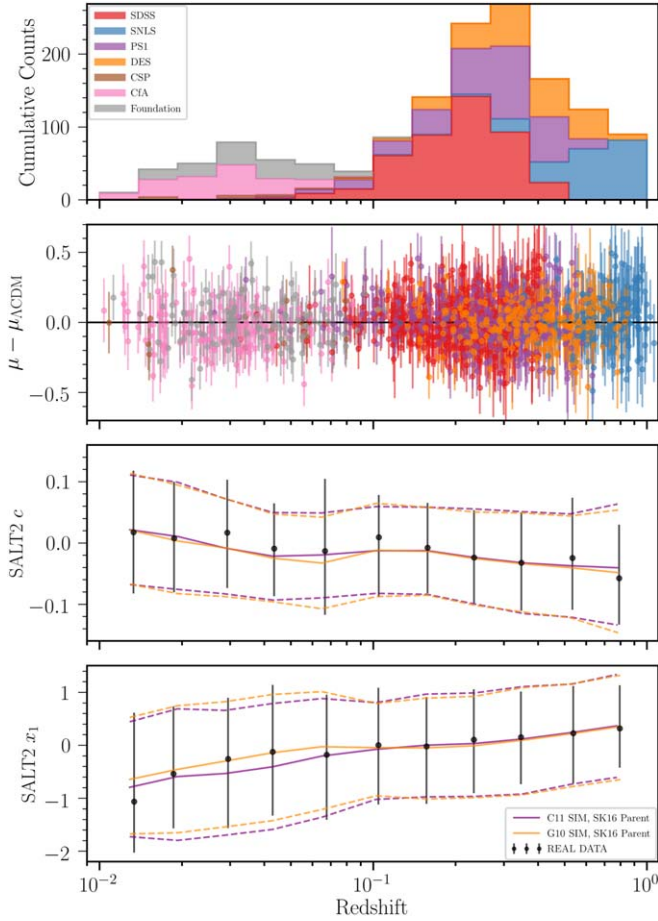
Figure 1. Top panel: stacked redshift histograms of each of the samples analyzed. Second: Hubble diagram residuals relative to a flat ΛCDM cosmology with $w = -1$. Third: mean and 68% intervals for the measured SALT2 color from the data, shown as blue points. Predictions from survey simulations are shown in purple for simulations with the C11+SK16 model and orange for simulations with the G10+SK16 scatter model. Bottom: same as third panel, but for the measured SALT2 stretch x_1.

DES3YR. The redshift distribution of SNe Ia used in this work can be found in the top panel of Figure 1.

This analysis relies largely on the host-galaxy mass estimates provided by past analyses. We adopt the same masses released in the Pantheon sample, and references therein, for SDSS, PS1, SNLS, CSPDR2, and CfA. For DES3YR masses, we use the updated masses provided by Smith et al. (2020) and Wiseman et al. (2020). For the Foundation sample, we utilize masses derived in Jones et al. (2018b).

2.2. Light-curve Fits and Distance Modulus Determination

We fit the SNe with the SALT2 model as presented in Guy et al. (2010) and updated in B14. In SALT2, the SN Ia flux at a phase (p) and wavelength (λ) is given as

$$F(\mathrm{SN}, p, \lambda) = x_0 \times [M_0(p, \lambda) + x_1 M_1(p, \lambda) + \ldots] \times \exp[c\,\mathrm{CL}(\lambda)], \quad (1)$$

where the parameter x_0 describes the overall amplitude of the light curve, x_1 describes the observed light-curve stretch, and c describes the observed color of each SN. Also, M_0, M_1, and CL are global model parameters of all SNe Ia: M_0 represents the average spectral sequence (SED), M_1 is the SED variability,

and CL is the average color-correction law. The light-curve fits assume Fitzpatrick (1999) for Milky Way reddening. The mean observed c and x_1 for the data, binned over redshift, are shown in the bottom panels of Figure 1.

Distances are inferred following the Tripp estimator (Tripp 1998). The distance modulus (μ) to each candidate SN Ia is obtained using

$$\mu = m_B + \alpha_{\mathrm{SALT2}} x_1 - \beta_{\mathrm{SALT2}} c - M, \quad (2)$$

where m_B is the peak brightness based on the light-curve amplitude ($\log_{10}(x_0)$), and M is the absolute magnitude of an SN Ia with $x_1 = c = 0$; α_{SALT2} and β_{SALT2} are the correlation coefficients that standardize the SNe Ia and are determined following Marriner et al. (2011), in a process similar to what is done in B14. Marriner et al. (2011) minimize a χ^2 expression that depends on the Hubble residuals after applying the Tripp estimator (see Equation (2)) and normalize residuals by the quadrature sum of the measurement uncertainties and intrinsic scatter. The method separates the sample into redshift bins in order to remove the cosmological dependence of the fitting procedure. The procedure iterates to determine the intrinsic scatter σ_{int}, α, and β and the resultant distance modulus values.

In recent analyses with the Tripp estimator (S18, B19b), there is often additional additive terms δ_{bias}, the correction for distance biases calculated from survey simulations, and δ_γ, the correction due to the host-galaxy mass correlation. These additional corrections are not applied because new treatments for both of these terms are introduced in following sections.

Distance uncertainties are computed from the uncertainties in the light-curve fit parameters and their covariance (C):

$$\begin{aligned}
\sigma_\mu^2 = {} & C_{m_B, m_B} + \alpha_{\mathrm{SALT2}}^2 C_{x_1, x_1} + \beta_{\mathrm{SALT2}}^2 C_{c,c} \\
& + 2\alpha_{\mathrm{SALT2}} C_{m_B, x_1} - 2\beta_{\mathrm{SALT2}} C_{m_B, c} - 2\alpha_{\mathrm{SALT2}} \beta_{\mathrm{SALT2}} C_{x_1, c} \\
& + \sigma_{\mathrm{vpec}}^2 + \sigma_z^2 + \sigma_{\mathrm{lens}}^2 + \sigma_{\mathrm{int}}^2,
\end{aligned}$$

$$(3)$$

where σ_{vpec} is the distance modulus uncertainty due to peculiar velocities (250 km s^{-1}), σ_z is the distance modulus uncertainty due to the measured redshift uncertainty, σ_{lens} is the additional uncertainty from weak gravitational lensing ($0.055z$), and σ_{int} is determined such that the reduced χ^2 relative to a best-fit cosmology is 1.

Typical selection cuts are applied to the observed data sample as was done in B19b: we require a fitted color uncertainty < 0.05, fitted stretch uncertainty < 1, fitted light-curve peak date uncertainty < 2, and light-curve fit probability (from SNANA) > 0.01, and Chauvenaut's criterion is applied to distance modulus residuals, relative to the best-fit cosmological model, at 3.5σ. In total, after selection cuts, there are 1445 SNe in this sample.

2.3. Key Pillars of the Complexity of the Colors of SNe Ia

The complexity of the SN Ia color model is readily apparent after a simple Tripp standardization. Here, three critical features are presented in the observed data set that must be explained by models of SN Ia color and intrinsic scatter:

1. The distribution of observed SN Ia colors is shown in the top of Figure 2. There is a clear asymmetry, with an
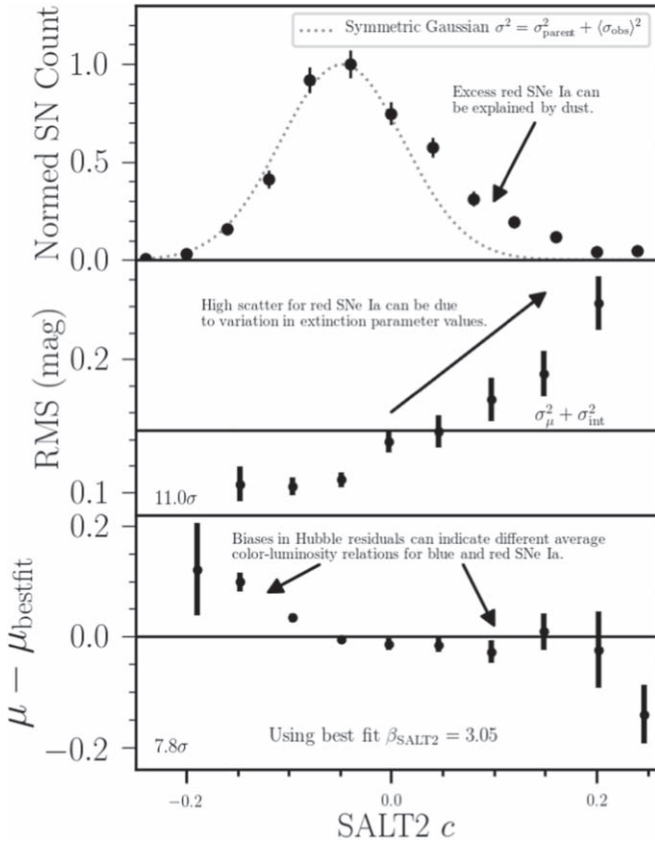
Figure 2. Top panel: observed color histogram from the full data sample, with symmetric Gaussian overlaid. Middle: rms of Hubble diagram residuals as a function of color. The rms is calculated after Tripp standardization and after subtracting the mean Hubble residual bias. Bottom: binned Hubble diagram residuals as a function of color, after Tripp standardization using the best-fit β_{SALT2}. In the bottom two panels, the significance of the deviation from a flat line is show in the bottom corner.

excess of red SNe in comparison with blue SNe, that is inconsistent with a symmetric Gaussian distribution.

2. The relation between the rms scatter of distance modulus residuals (with mean residual removed in each bin) as a function of SN Ia color is shown in the middle panel of Figure 2. There is an 11σ dependence relative to a flat line, where the redder SNe Ia ($c > 0.1$) exhibit nearly twice as much scatter (~ 0.18 mag) as the bluest SNe ($c < -0.1$), which exhibit ~ 0.1 mag scatter. This effect remains if any single survey is removed from the sample.

3. The relation between Hubble residual binned distance biases and SN Ia color is shown in the bottom panel of Figure 2. There is an $\sim 7.8\sigma$ dependence relative to a straight line. As shown in Figure 2, the recovered β_{SALT2} of the data is 3.05 ± 0.06.

The relation of increased scatter as a function of color has not been analyzed in a previous analysis. This paper is motivated by quantifying these observed features and building a model that can address all of them simultaneously.

2.4. Using Survey Simulations to Evaluate SN Ia Color and Intrinsic Scatter Models

For every model presented in this paper, 100 realizations of data-set-sized simulations are run. SNANA (Kessler et al. 2009) is used to simulate realistic samples of SNe Ia. These simulations account for observing cadence, observing conditions, noise properties, selection effects, cosmological effects, and astrophysical effects. A general description of the simulation methodology can be found in Kessler et al. (2019), and the survey-specific simulation details for SDSS and SNLS are described in Kessler et al. (2013); PS1, CSP, and CfA are described in S18; DES3YR is described in B19b; and Foundation is described in Jones et al. (2018b).

We define three metrics based on the three panels of Figure 2 that are pseudo χ^2 evaluations that assess agreement between simulations that assume an SN Ia model and the data. The first metric is defined as χ_c^2 for the agreement in color histograms of data (N_{data_c}) and survey simulations (N_{sim_c}) such that

$$\chi_c^2 = \sum_j (N_{c_j}^{\text{data}} - N_{c_j}^{\text{sim}})^2 / e_{nj}^2, \qquad (4)$$

and it is determined in bins of color (j) where e_{nj} is determined by Poisson statistics.

A second metric, the agreement in total Hubble diagram scatter (rms) between the data (rms$_{\text{data}}$) and survey simulations (rms$_{\text{sim}}$), is defined as χ_{rms}^2 over color bins such that

$$\chi_{\text{rms}}^2 = \sum_i (\text{rms}_{c_i}^{\text{data}} - \text{rms}_{c_i}^{\text{sim}})^2 / e_{ci}^2 \qquad (5)$$

and is determined in bins of color i and where e_{ci} are the errors determined from 100 realizations of the simulated data set. We use rms instead of intrinsic scatter as a metric because, for intrinsic scatter, the sensitivity of the different components of the error modeling is difficult to track.

A third metric is the agreement in distance modulus residuals between data ($\Delta \mu^{\text{data}}$) and survey simulations ($\Delta \mu^{\text{sim}}$), which can be expressed as $\chi_{\Delta \mu}^2$ over color bins such that

$$\chi_{\Delta \mu}^2 = \sum_i (\Delta \mu_{c_i}^{\text{data}} - \Delta \mu_{c_i}^{\text{sim}})^2 / e_{\mu i}^2 \qquad (6)$$

and is determined in bins of color i and where $e_{\mu i}$ are the errors derived from the data itself.

A fourth metric is the agreement between the recovered (Marriner et al. 2011) color–luminosity coefficients of simulations ($\beta_{\text{SALT2}}^{\text{sim}}$) and data ($\beta_{\text{SALT2}}^{\text{data}}$) such that

$$\chi_{\beta_{\text{SALT2}}}^2 = (\beta_{\text{SALT2}}^{\text{data}} - \beta_{\text{SALT2}}^{\text{sim}})^2 / (\sigma_{\beta_{\text{SALT2}}^{\text{data}}}^2 + \sigma_{\beta_{\text{SALT2}}^{\text{sim}}}^2) \qquad (7)$$

where $\sigma_{\beta_{\text{SALT2}}}$ is the uncertainty reported following Marriner et al. (2011).

Finally, we minimize the cumulative χ^2 in our fits:

$$\chi_{\text{Tot}}^2 = \chi_c^2 + \chi_{\text{rms}}^2 + \chi_{\Delta \mu}^2 + \chi_{\beta_{\text{SALT2}}}^2. \qquad (8)$$

The fitting process involves large simulations, which given current SN Ia sim/analysis infrastructure are prohibitive for out-of-the-box minimizers and Monte Carlo samplers. We therefore implement the following minimization algorithm:

1. Coarse grid minimization of model parameters for initial guess
2. Proposal of new model parameters
3. 100 simulations with one-dimensional perturbations around proposed parameters (no covariance)
4. Gradient descent
5. Repeat steps 2–5 (typically around 50 iterations).

Results using this algorithm and its limitations are discussed in Section 3.4. We note that this method does not account for

covariance between fitted parameters. This is the work of a future paper (Popovic et al. 2021), which incorporates significant infrastructure improvements.

3. Evaluating Models of SNe Ia Colors and Intrinsic Scatter

3.1. Previous Models of Intrinsic Scatter and Associated Intrinsic Color Populations

Recent studies have focused on two models of intrinsic scatter, which, to first order, can both be described by two parameters: the magnitude of chromatic and achromatic scatter. The two models are the "G10" scatter model (Guy et al. 2010), which ascribes 70% of the intrinsic scatter to coherent variation and 30% to chromatic (wavelength-dependent) variation, and the "C11" scatter model (Chotard et al. 2011), which ascribes only 25% of the intrinsic scatter to coherent variation but 75% to chromatic variation. Both of these models were trained on data: C11 was trained on spectra from the SNFactory (Aldering et al. 2002), and G10 was trained during the creation of the SALT2 model on a large subset of the light curves used in this analysis (Guy et al. 2010, B14).

These scatter models cannot be used in survey simulations to predict color distributions or the trends of Figure 2 without an associated color population and a β_{SALT2} as defined in Equation (2). For both the G10 and C11 scatter models, Scolnic & Kessler (2016, hereafter SK16) determined the underlying color population such that when it was combined with measurement noise, the color scatter from the scatter model, and selection effects, the observed color distribution matched that seen for the data in the top panel of Figure 2. The underlying population was described by an asymmetric Gaussian, with three free parameters. The value of β_{SALT2} was determined by finding what input β_{SALT2} in the simulations would yield an output β_{SALT2} consistent with that found in the data from the methodology outlined in Section 2.2.

The number of parameters that describe the framework for one of these scatter models is six: two parameters for the spectral and coherent scatter, three parameters for the underlying population, and the value of β_{SALT2}. However, in order to explain inconsistencies between the low-z targeted sample and the high-z samples, SK16 determined the underlying population for each separately. Therefore, in total, a description of the full sample is described by nine parameters.

For the simulations with G10 and C11, a single input β_{SALT2} value is used for each one: $\beta_{SALT2} = 3.1$ and $\beta_{SALT2} = 3.8$ for G10 and C11, respectively. As explained in past analyses (Scolnic et al. 2014b; Scolnic & Kessler 2016; Kessler & Scolnic 2017), applying the 1D fitting procedure from Marriner et al. (2011) recovers an observed $\beta_{SALT2} \sim 3.1$ for both the G10 and C11 cases. Due to the larger amount of color scatter in the C11 model, the associated underlying color population of C11 appears much more like a sharp, dust-like exponential distribution (Scolnic et al. 2014b) than the one for G10. While it is unclear how to apply a physical interpretation to the G10+SK16 model, one possible interpretation for the C11+SK16 model is that there are two color–luminosity relations: one that relates the dust-like color to luminosity, and another with no luminosity correlation ($\beta = 0$) for the intrinsic color distribution. The populations used for the samples in Pantheon (Low-z, PS1, SNLS, SDSS) can be found in SK16, for Foundation in Jones et al. (2018b), and for DES3YR in B19b.

3.2. Evaluating Past SN Ia Scatter Models

As expected, because the SK16 populations were determined so that simulations would reproduce the observed color distribution of the data, simulations based on C11+SK16 and G10+SK16 show excellent agreement with the data (Figure 4), giving a χ_c^2 of 9.0 and 9.5, respectively (12 bins). The mean observed c and x_1 for the simulations, binned over redshift, are shown in the bottom panels of Figure 1 and are in similarly good agreement with the data. However, the agreement between data and simulations for both the rms (Figure 5(a)) and mean Hubble residuals (Figure 5(b)) is comparatively poor.

For the rms of Hubble residuals (Figure 5(a)), it is clear that neither G10+SK16 nor C11+SK16 can produce the trend observed in the data. We do see a nonlinear behavior predicted from the simulations for the C11+SK16 model, which ascribes more scatter to SN Ia chromatic variation and achieves a $\chi_{rms}^2 = 35$, whereas G10+SK16, which ascribes little color variation, achieves a $\chi_{rms}^2 = 68$. The relatively flat dependence of the rms on color as predicted from the G10+SK16 model shows that the trend in the data cannot be explained by a lower signal-to-noise ratio for SNe with redder colors.

The agreement between data and simulations for mean Hubble residuals (Figure 5(b)) is somewhat better for G10 +SK16 ($\chi_{\Delta\mu}^2 \sim 12$) but worse for C11+SK16 ($\chi_{\Delta\mu}^2 \sim 29$). As discussed in SK16 and used for the motivation of Kessler & Scolnic (2017), both models do predict the upturn in mean Hubble residuals for blue colors. Such distance modulus biases arise from the combination of asymmetric color distributions with color scatter and selection effects.

3.3. Parameterization of a New Dust-based Color Model

We present in Figures 2 and 3 a simple and more physical understanding of the trends seen: the redder colors can be explained by dust extinction, the high rms for red SNe Ia could be explained by variations in the extinction parameter, and Hubble residual biases for the blue and red SNe can be explained by different respective color–luminosity relations. Here, we follow Mandel et al. (2011, 2017), who build on the work of Jha et al. (2007) to create a model of SN color based on two components: (1) an intrinsic color component (c_{int}) related to luminosity by a correlation coefficient β_{SN} and (2) a dust component (E_{dust}) described by an exponential distribution of reddening values related to luminosity by the extinction ratio R_V. The observed color c_{obs} can be expressed as

$$c_{obs} = c_{int} + E_{dust} + \epsilon_{noise} \qquad (9)$$

where ϵ_{noise} is measurement noise. We expand on the model from Mandel et al. (2011) by allowing R_V to be described by a Gaussian distribution to reflect that a range of values are seen in the literature, rather than a single value. In total, the model has seven fundamental parameters:

1. $\bar{c}$: the mean of the intrinsic color distribution described by a symmetric Gaussian
2. σ_c: the 1σ width of the intrinsic color distribution described by a symmetric Gaussian
3. $\bar{\beta}_{SN}$: the correlation between intrinsic color and luminosity
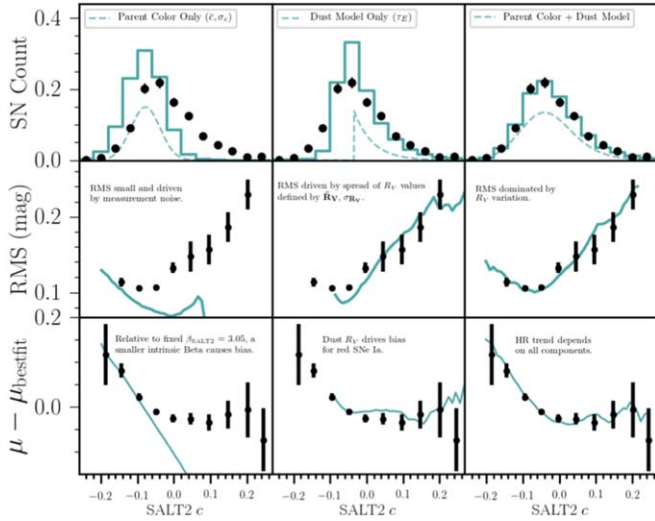
Figure 3. Explaining the BS20 model. Shown are parameterizations input to simulations (dashed teal), the simulated values after measurement noise and selection effects (solid teal), and the data set (black points) for the same three quantities (y-axes) as in Figure 2. Left: a simulation based solely on an intrinsic color distribution, described by a symmetric Gaussian, without dust. Middle: a simulation based solely on a delta function in intrinsic color and an exponential dust distribution. Right: a simulation with both intrinsic color Gaussian and dust distribution combined.

4. $\sigma_{\beta_{SN}}$: the 1σ width of the Gaussian distribution from which the correlation between intrinsic color and luminosity is drawn for each SN

5. $\bar{R}_V$: the center of the Gaussian distribution from which R_V values are drawn for each SN

6. σ_{R_V}: the 1σ width of the parent Gaussian R_V distribution

7. τ_E: the parameter describing the exponential distribution from which E_{dust} reddening values are drawn.

To set a "reddening-free" color, it is assumed that the intrinsic colors of SNe Ia can be determined with

$$P(c_{int}) = \frac{1}{\sqrt{2\pi}\,\sigma_c} e^{-(c_{int}-\bar{c})^2/2\sigma_c^2}. \quad (10)$$

The reddening for each SN is described by E_{dust} from Equation (9) and is related to the extinction of the SN by the standard equation

$$A_V = R_V * E_{dust} \quad (11)$$

where E_{dust} corresponds to $E(B-V)$.

The reddening values E_{dust} are drawn from an exponential distribution following Mandel et al. (2017) with probability density

$$P(E_{dust}) = \begin{cases} \tau_E^{-1} e^{-E_{dust}/\tau_E}, & E_{dust} > 0 \\ 0, & E_{dust} \leqslant 0 \end{cases} \quad (12)$$

where τ_E is a parameter in the model described above.

In addition, we draw from the distribution of possible values for R_V:

$$P(R_V) = \frac{1}{\sqrt{2\pi}\,\sigma_{R_V}} e^{-(R_V-\bar{R}_V)^2/2\sigma_{R_V}^2} \quad (13)$$

where $\bar{R}_V$ is the center of the Gaussian distribution of R_V, σ_{R_V} is the width, and individual R_V values below 0.5 are not allowed.

Finally, similar to Equation (13), values for β_{SN} are drawn for each SN using model parameters $\bar{\beta}_{SN}$ and $\sigma_{\beta_{SN}}$ such that

$$P(\beta_{SN}) = \frac{1}{\sqrt{2\pi}\,\sigma_{\beta_{SN}}} e^{-(\beta_{SN}-\bar{\beta}_{SN})^2/2\sigma_{\beta_{SN}}^2}. \quad (14)$$

In total, the change in observed peak brightness of an SN due to color can be expressed as Δm_B,

$$\Delta m_B = \beta_{SN} c_{int} + (R_V + 1)E_{dust} + \epsilon_{noise}, \quad (15)$$

where each observed parameter is unique to each SN. The coefficient $R_V + 1$ is used rather than R_V as in Equation (11), because to measure the change in m_B, the extinction parameter $R_B = R_V + 1$ is needed.

To describe one survey with this model, seven parameters are required. If one is to solve for parameters to describe all high-redshift and low-redshift surveys separately, then one additional parameter is needed: a separate τ_E for each. In total, this makes eight parameters. In contrast, as discussed previously, the G10+SK16 or C11+SK16 models require nine parameters when high-redshift and low-redshift samples are accounted for separately. Thus, the dust-based framework described here has fewer free parameters than those used in past cosmological analyses.

3.4. Results for the New Color Model

The parameters described in Section 3.3 can be fit from the photometric data itself using the four metrics (Equations (4), (5), (6), and (7)). Model parameters and their 1D uncertainties are shown in Table 1. We present the χ^2 surfaces from our iterative forward-modeling minimization process shown in Appendix A.1, and we show visually the degeneracies between model parameters in appendix Figure 10 (as well as for additional models). We note that the estimates of the uncertainties are limited by computational capability and thereby require the coarseness of the model grid. While some of the posteriors are not clearly Gaussian, we assume Gaussianity to determine the uncertainty. The only exception is when the posterior hits a cutoff (e.g., $\tau = 0$), in which case we report upper and lower uncertainties.

We find a mean reddening-free color of $\bar{c} = -0.084 \pm 0.004$ with an intrinsic color distribution of $\sigma_c = 0.042 \pm 0.002$ and a mean intrinsic color–luminosity correlation coefficient of $\bar{\beta}_{SN} = 1.98 \pm 0.18$. We find no evidence of β_{SN} variation (1.75σ significance) with $\sigma_{\beta_{SN}} = 0.35 \pm 0.20$. The recovered β_{SN} is smaller than the traditional $\beta_{SALT2} \sim 3$ found when assuming a single correction for the full SN Ia color and dust population simultaneously, and it shows a relatively weak correlation between intrinsic color and luminosity in comparison to the contribution due to dust. We find that the R_V distribution for the dust component is best described by $\bar{R}_V = 2.0 \pm 0.2$ and $\sigma_{R_V} = 1.4 \pm 0.2$. The value of $\sigma_{R_V} = 1.4$ indicates a wide range of R_V, though with a set floor of $R_V = 0.5$. Because a single color–luminosity relation is assumed in standardization, even though our simulations include a wide range of R_V values, we find that the measured R_V variation dominates the scatter of distance modulus residuals, contributing 0.095 to σ_{int}, the majority of the total σ_{int} (0.106). On the other hand, the measured variation in the intrinsic color–luminosity relation ($\sigma_{\beta_{SN}} = 0.35$) contributes 0.040 to the total σ_{int}.

The results of simulations with our model are presented in Figures 4 and 5. We find a $\beta_{SALT2} = 3.12 \pm 0.02$ when analyzed identically to the observed data set, which is

Table 1
Parameters Used for BS20 Model

Model	Sample	$\bar{c}$	σ_c	$\bar{\beta}_{\mathrm{SN}} - 1$	$\sigma_{\beta_{\mathrm{SN}}}$	$\bar{R}_V$	σ_{R_V}	τ_E
No Mass Split:								
Full	CfA, CSP, Foundation	-0.084 ± 0.004	0.042 ± 0.002	0.98 ± 0.18	0.35 ± 0.20	2.0 ± 0.2	1.4 ± 0.2	0.17 ± 0.04
Full	DES, PS1, SNLS, SDSS	-0.084 ± 0.004	0.042 ± 0.002	0.98 ± 0.18	0.35 ± 0.20	2.0 ± 0.2	1.4 ± 0.2	0.10 ± 0.02
Mass Split:								
High mass[a]	CfA, CSP, Foundation	-0.084 ± 0.004	0.042 ± 0.002	0.98 ± 0.18	0.35 ± 0.20	1.50 ± 0.25	1.3 ± 0.2	0.19 ± 0.08
High mass	DES, PS1, SNLS, SDSS	-0.084 ± 0.004	0.042 ± 0.002	0.98 ± 0.18	0.35 ± 0.20	1.50 ± 0.25	1.3 ± 0.2	0.15 ± 0.02
Low mass[b]	CfA, CSP, Foundation	-0.084 ± 0.004	0.042 ± 0.002	0.98 ± 0.18	0.35 ± 0.20	2.75 ± 0.35	1.3 ± 0.2	$0.01^{+0.05}_{-0.01}$
Low mass	DES, PS1, SNLS, SDSS	-0.084 ± 0.004	0.042 ± 0.002	0.98 ± 0.18	0.35 ± 0.20	2.75 ± 0.35	1.3 ± 0.2	0.12 ± 0.02

Notes.
[a] High mass: host $\log(M_*/M_{\mathrm{sun}}) > 10$.
[b] Low mass: host $\log(M_*/M_{\mathrm{sun}}) < 10$.
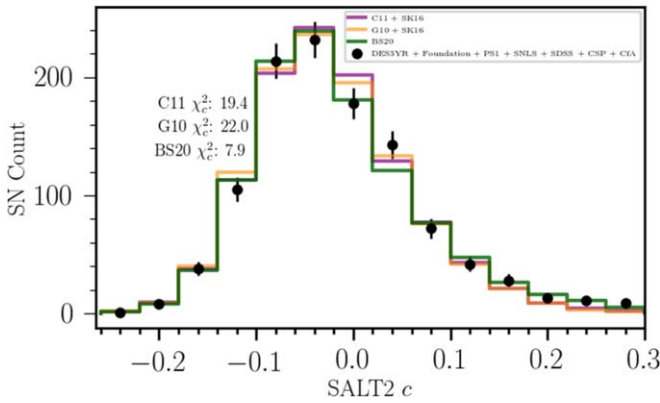
Figure 4. Histogram of the observed color values from data (points) and simulations (lines). As all models are fitted so that simulations match the data for this metric, good agreement between data and simulation is expected for all of the models.

consistent with that of the observed data set ($\beta_{\mathrm{SALT2}} = 3.04 \pm 0.06$). In Figure 4, we show that the BS20 model results in an observed SN Ia color distribution similar to that of the data ($\chi_c^2 \sim 8$). Furthermore, as shown in Figure 5(a), this model captures the increased rms scatter for the redder SNe ($\chi_{\mathrm{rms}}^2 \sim 7$), which is attributed to the variation of R_V. Finally, as shown in Figure 5(b), the BS20 model results in excellent agreement with observed Hubble residual biases ($\chi_{\Delta\mu}^2 \sim 6$).

In comparing χ^2 values between the different color scatter models for the three metrics in Table 2, the advancement of the BS20 model is clear, and with one less parameter, the improvement cannot be simply attributed to additional model complexity.

4. Dependence of the Host-mass Correlation with SNe Ia Luminosity on Color

4.1. Observed Trends of Color Metrics Based on Host-galaxy Stellar Mass

Many studies have found correlations between the Hubble residuals and various host-galaxy properties (Hicken et al. 2009a; Lampeitl et al. 2010; Sullivan et al. 2010; Childress et al. 2013a; Rigault et al. 2013; Roman et al. 2018; Rose et al. 2019). Here, we focus on the host-galaxy stellar mass as it is the most commonly used and most accessible and often yields some of the strongest correlations with Hubble residuals. In the top panel

of Figure 6(a), the rms versus SALT2 color plot as shown in Figure 5(a) is remade, but for the high and low host-mass subsamples separately. For the "dust-free" blue SNe ($c \sim -0.1$), there is little difference between the rms for SNe in low- and high-mass hosts. However, the rms increases with redder SN colors, and much more significantly for SNe in low-mass hosts.

As shown Figure 6(b), when splitting the data set into high and low host-mass subsamples, there is a distinct difference in the color dependence in the biases of Hubble residuals. For the "dust-free" blue SNe, the slope of the color–luminosity relation and the absolute Hubble residual biases for SNe in low- and high-mass host subsamples are identical. For the redder SNe, however, there are distinctly different color–luminosity relations, and there is as much as an ~0.15 mag difference in Hubble residuals. Overall, the subsamples are discrepant at greater than 5σ ($\chi^2/N_{\mathrm{bin}} = 57/10$) relative to each other.

Pursuing this further, we follow recent works like B19b and define γ as the mean difference in Hubble residuals given a split in host-galaxy properties:

$$\delta\gamma = \gamma \times [1 + e^{(\mathcal{M}_{\mathrm{host}} - \mathcal{M}_{\mathrm{step}})/0.01}]^{-1} - \frac{\gamma}{2}, \quad (16)$$

where $\mathcal{M} = \log(M_*/M_{\mathrm{sun}})$ and a log host-mass step location ($\mathcal{M}_{\mathrm{step}}$) of 10 is assumed. We determine γ for the sample in discrete color bins. This is shown in the bottom panel of Figure 6(a). As expected from the observations in Figure 6(b), for "dust-free" SNe Ia that are bluer than the intrinsic color $\bar{c}$, $\gamma = 0.003 \pm 0.029$, consistent with zero. However, for redder SNe, there is a significant $\gamma = 0.083 \pm 0.011$ as well as a 4.5σ increasing trend where $\Delta\gamma \sim 0.72 \pm 0.14 \times c$, showing that the typical γ values around 0.06 mag recovered in previous analyses are driven by the red SNe in the sample.

While many studies have shown that host mass and SN color are weakly correlated if at all (e.g., Sullivan et al. 2010), the dependence of γ itself on color has not been studied. As our model shows that redder colors can be described by dust, the differences between observed correlations between Hubble residuals and masses for different colors are all indicative of a dust-based explanation. We note that the trend seen in the bottom of Figure 6(a) is largely insensitive to whether distance bias corrections are applied. If we apply corrections based on Kessler & Scolnic (2017), the γ recovered is 0.0–0.02 mag lower per bin than that shown, which is discussed at length in Smith et al. (2020). The trend with rms is not affected by these
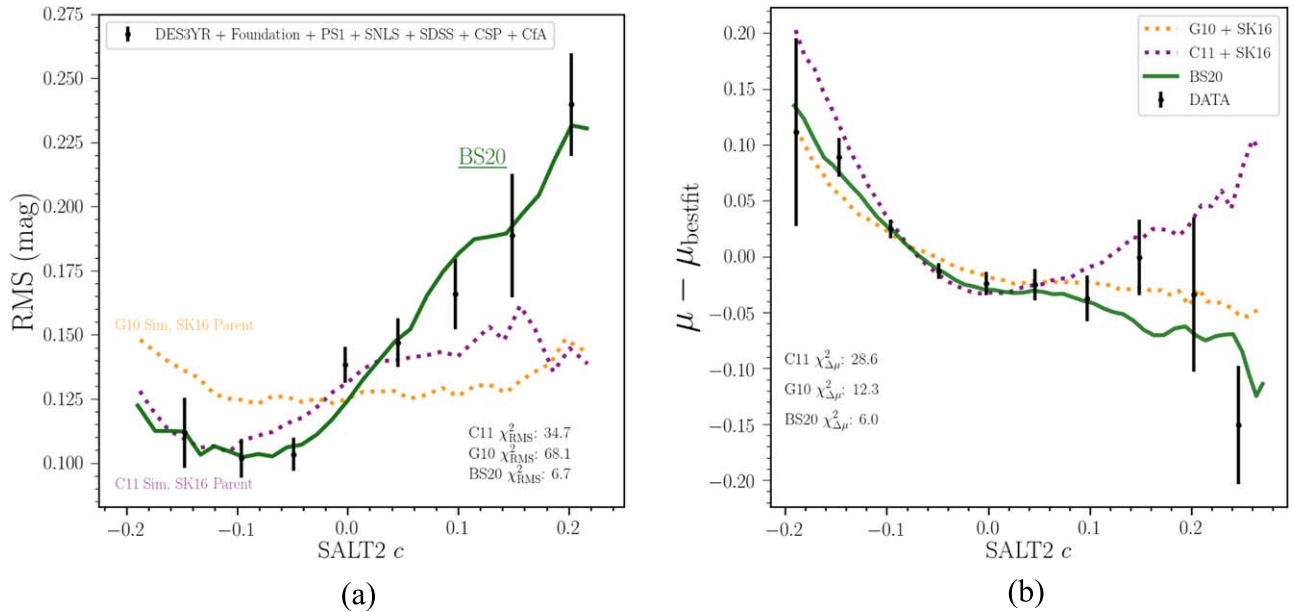
Figure 5. (a) Zero-mean rms of the Hubble residuals relative to ΛCDM vs. the observed color c of the SNe Ia. The data are shown in black points, and the predictions from simulations of the G10+SK16 and C11+SK16 models are shown in orange and purple dotted lines, respectively. The model created for this work, labeled BS20, is shown in green. Inset: same as main figure but for the intrinsic scatter term σ_{int} instead of rms. (b) Binned Hubble diagram residuals vs. color. Biases are seen in the observed data (black points) and predicted by the scatter models (solid/dotted lines). For the BS20 model used here, there is no split on host mass.

corrections because the rms measured per bin is calculated after a mean offset is removed, thereby effectively doing a correction similar to that in Kessler & Scolnic (2017).

4.2. Dust Modeling Explains the Mass Step

We repeat the process as described in Section 3 to determine the underlying dust-based color model, except now for the low- and high-mass host-galaxy subsamples separately. The fitted parameters are given in the "Mass Split" grouping of Table 1. Parameters that are intrinsic to the SNe Ia are fixed for both host-galaxy subsamples, while the dust distributions are allowed to vary for each subsample. We find that for SNe in low-mass hosts, $\bar{R}_V = 2.75 \pm 0.35$ with $\sigma_{R_V} = 1.3 \pm 0.2$, whereas for SNe in high-mass hosts, $\bar{R}_V = 1.50 \pm 0.25$ with $\sigma_{R_V} = 1.3 \pm 0.2$, suggesting that the peak $\bar{R}_V$ values differ by 2.9σ between low- and high-mass hosts. We note that σ_{R_V} is found to be the same between low- and high-mass hosts, though it is unclear what the physical motivation for this would be. After accounting for selection effects, the distribution shifts such that the average observed R_V for the detected SNe in the sample is 2.94 and 1.85 for low-mass hosts and high-mass hosts, respectively. In these simulations, 2% of all of the detected SNe have simulated R_V values greater than 5. The dust distribution for SNe in high-mass hosts that are discovered in high-z surveys is described with $\tau_E = 0.15 \pm 0.02$, whereas for low-mass hosts we find $\tau_E = 0.12 \pm 0.02$, and similarly for the low-z surveys the SNe can be described with $\tau_E = 0.19 \pm 0.08$, whereas for low-mass hosts it is $\tau_E = 0.01^{+0.05}_{-0.01}$.

We show in Figure 6(b) that simulations with these separate dust models do indeed each recover the trends in Hubble residuals, and consequently the trend seen in the bottom panel of Figure 6(a). Therefore, we conclude that modeling different dust properties for different galaxy populations can fully explain the net $\gamma \sim 0.06$ mag offset seen in past analyses, as well as the γ dependence on observed SN Ia color.

As shown from the data, applying a single offset (γ), as was done in past analyses, is incorrect. Furthermore, it has been unclear in past analyses why there should be any "step" behavior (Sullivan et al. 2010). Here, it is shown that the past step is an artifact of improper fitting and arises from significantly different R_V distributions for different types of galaxies.

5. Impact on Recovery of Cosmological Parameters

To understand the impact of these different models of SN Ia color on the recovery of cosmological parameters, both data and simulations are used. Before measuring cosmological parameters, we apply bias corrections following the methodology of Marriner et al. (2011) and B14 using large simulations with the three color models (G10+SK16, C11+SK16, BS20) to measure the dependence of distance biases on redshift, which are then applied as corrections to the data set or a simulated data set. Bias corrections following Kessler & Scolnic (2017) are not used because they have so far been only designed to work given a β_{SALT2} and a variation in c, but not R_V, nor variation thereof. Therefore, we apply bias corrections that assume a single β_{SALT2} and follow the same formalism that was used in the JLA analysis, and we do not split by host mass. This is done so a self-consistent comparison can be made against the impact of the G10+SK16 and C11+SK16 scatter models.

The impact of the bias corrections on the data is shown in Figure 7. The most noticeable differences between the corrections of BS20 versus the other scatter models are at $z < 0.1$ and $z > 0.8$, where selection has the greatest influence. Here, the differences in recovered distance modulus can change by up to $\sim$0.05 mag at low or high z depending on which color model is used. This difference is larger than any other systematic in past cosmology analyses (e.g., B19b).

As shown in Figure 7, we see the same effect with simulations as we do for data when simulating a sample of

Table 2
χ^2 for Each Evaluated Model

Scatter Model	Color Model	χ^2_c	χ^2_{rms}	$\chi^2_{\Delta\mu}$	$\chi^2_{\beta_{SALT2}}$	χ^2_{Tot}	Number of Parameters[a]
G10	SK16	22.0	68.1	12.3	1.6	104.0	9
C11	SK16	19.4	34.7	28.6	1.3	84.0	9
BS20 No mass split[b]	BS20	7.9	6.7	6.0	1.7	22.3	8

Notes.
[a] The number of parameters is counted in the text.
[b] The agreement as a function of mass, as shown in Figure 6(b), is not included in $\chi^2_{\Delta\mu}$, so the BS20 mass-split model is not listed in this table.

10,000 SNe with realistic proportions and distributions of SNe Ia from each survey. Here, the simulations of "data sets" are based on the BS20 model, but bias corrections are determined from the other models.

To determine cosmological parameters, we use CosmoMC (Lewis & Bridle 2002) and combine with cosmic microwave background (CMB; Planck Collaboration et al. 2020) constraints. In Table 3, the biases in cosmological parameters are given when simulated SN Ia data sets use different models of SN Ia color than the model used to determine the distance bias correction. We find that if the "true" model of SN Ia color is the dust-based model presented in Section 3.3 but the bias corrections are based on the G10+SK16 or C11+SK16 models, the propagated bias in w will be -0.025 and -0.040, respectively. Again, this bias is larger than any other systematic uncertainty reported in recent cosmological analyses.

In Table 3 we also show the differences in w for the real data when we apply bias corrections based on simulations using the three separate models of color: G10+SK16, C11+SK16, and BS20. Relative to BS20 bias corrections, there are changes in recovered w for G10+SK16 and C11+SK16 of -0.033 and -0.041, respectively, which are consistent with simulations. Interestingly, as shown in Figure 5(a), while C11 and BS20 better match the trend in the data, they produce the largest differences in w of $\sim$0.04.

6. Discussion

6.1. Dependence between R_V and Host-galaxy Properties

That the mass correlation can be explained by separate dust properties is now the only *direct* explanation for the correlation between host mass and distance modulus residuals. This possibility was briefly discussed in Mandel et al. (2017), which showed that if one changed the dust distribution (τ_E) for the SNe in low- and high-mass subsamples, one could remove one-third of the magnitude of γ, but not the whole effect. We follow this idea from Mandel et al. (2017), but add that the R_V distribution as well should be different for these subsamples. This can then explain the full γ and its color dependence. Our dust explanation aligns well with the observations in Burns et al. (2018) that at low z, the host-mass correlation with SN Ia luminosity is larger in the optical than in the near-infrared (NIR), where the correlation is consistent with zero. This should be the case if the correlation is tied to reddening, as the corresponding extinction ratio of R_V in the NIR is smaller. Furthermore, the range of R_V values is in good agreement with studies of R_V from individual SNe like in Amanullah et al. (2015). While the model shows that a fraction of SNe should have $R_V > 5$, we find that this is only 6% after accounting for selection effects.

This analysis makes a strong prediction that SNe in lower-mass galaxies have, on average, higher R_V values than SNe in higher-mass galaxies. As there are very few measurements of R_V in the interstellar medium of galaxies beyond the Milky Way, LMC, and SMC, it is difficult to find evidence that this trend would hold for galaxies themselves. Salim et al. (2018), who measured the dust attenuation curves of 230,000 individual galaxies in the local universe, found that quiescent galaxies, which are typically high in mass, have a mean $R_V = 2.61$, and star-forming galaxies, which are lower in mass on average, have a mean $R_V = 3.15$. This trend is in general agreement with our prediction.

The observation that global properties of the galaxy can impact the dust measured from the SNe is supported by Phillips et al. (2013) and Bulla et al. (2018b), who found that the dust responsible for the observed reddening of SNe Ia appears to be predominantly located in the interstellar medium of the host galaxies and not in the circumstellar medium associated with the progenitor system. This is also supported by Childress et al. (2013a), who showed that the color of SNe Ia is strongly tied to the metallicity of the host galaxy. For a future analysis, we encourage repeating this same exercise, but instead of using stellar mass, use metallicity, specific SFR, or local color; improved estimates of the dust distribution parameters would likely be obtained. For example, as shown in Sullivan et al. (2010), when measuring a single color–luminosity coefficient β_{SALT2} for different samples, there is an even bigger difference when splitting the sample for specific SFR than there is for mass. As our model constrains both the amount of dust and the properties of the dust itself, it is likely that different galaxy properties (e.g., distance to host and inclination; Galbany et al. 2012; Holwerda et al. 2015) will yield complementary insights about both of these components. We stress that our analysis does not limit the use of host-galaxy information in cosmological studies with SNe Ia, but rather proposes a new path forward.

Indirect explanations of γ have suggested that SNe from different progenitor systems have different luminosities, and the progenitor system can be potentially linked to the age of the host galaxy (Childress et al. 2013a). However, any model that assumes that the luminosity depends on progenitors does not predict the key observation in our analysis that the magnitude of γ depends on color. A progenitor-based explanation has motivated studies by Rigault et al. (2013), Childress et al. (2014), Jones et al. (2015, 2018a), and Roman et al. (2018), which focus on the local specific SFR, local mass, and local color. Some of these studies seem to indicate that measuring the local color produces the highest correlation with measured SN luminosity. In light of our dust-based SN Ia color model, a simple explanation is that the local host color yields insight about the amount of dust or dust properties at the position of
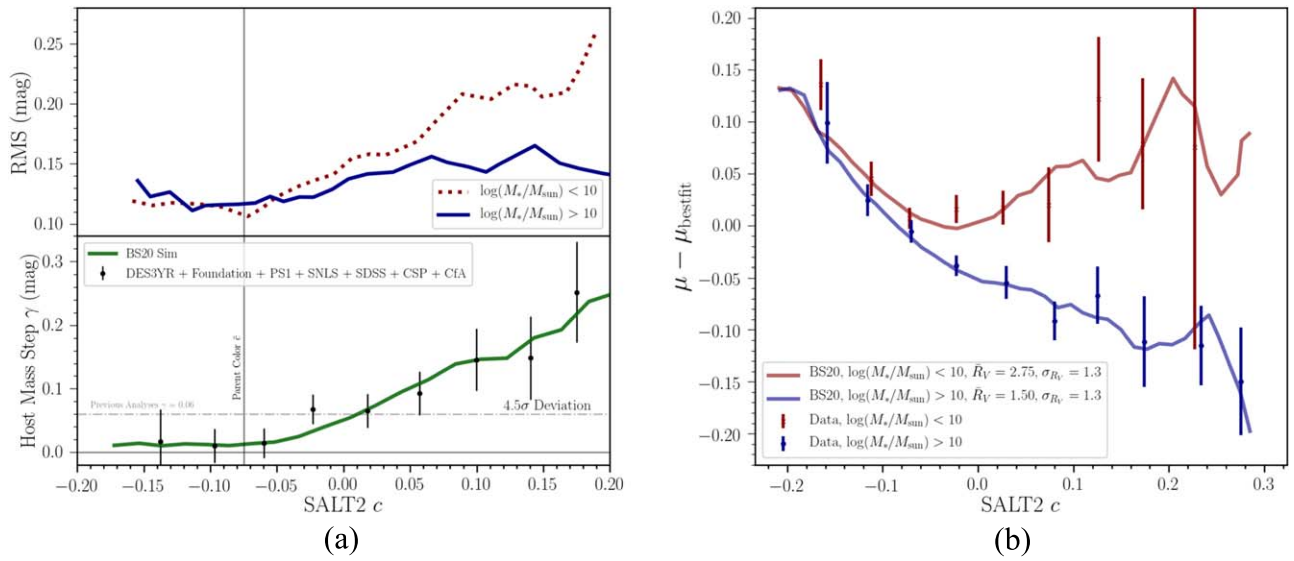
Figure 6. (a) Top panel: Hubble diagram scatter binned by SALT2 observed color and compared for SNe in host galaxies with low and high mass. Bottom: recovered values of γ for SNe Ia in high (log $(M_*/M_{\rm sun}) > 10$) vs. low (log $(M_*/M_{\rm sun}) < 10$) mass hosts, binned by SALT2 observed color. Predictions from the BS20 mass-split model are shown in green. Significance of the deviation from a constant γ of 0.06 is shown (4.5σ). (b) Binned Hubble diagram residuals vs. color split on host mass. Biases are shown for the observed data (points) and predicted using the scatter models (solid lines). The difference between the red and blue points has typically been found by marginalizing over color and finding a single step γ. The dust parameters causing the observed split are shown in the legend.
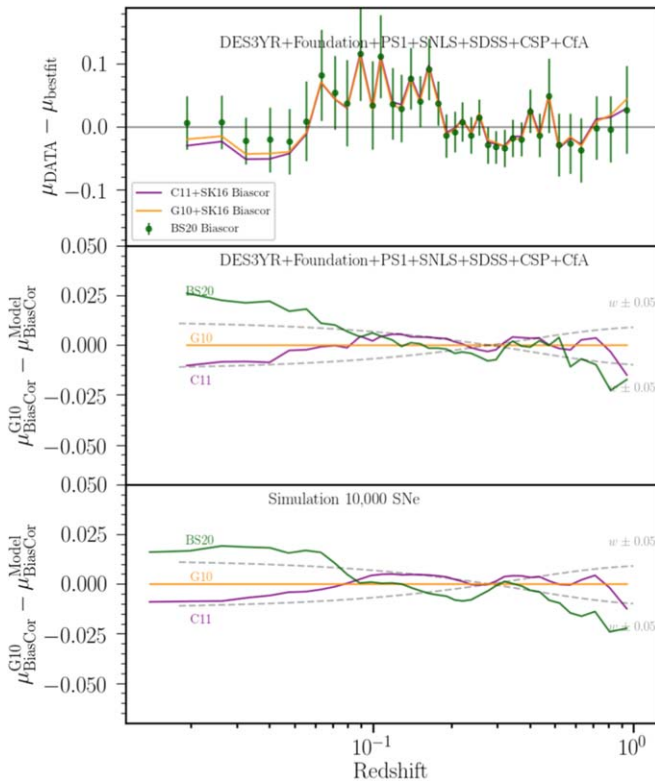
Figure 7. Top panel: Hubble diagram residuals of the compiled DES3YR +Foundation+PS1+SNLS+SDSS+CSP+CfA data set as a function of redshift. The data set is bias corrected with the three different models of SN Ia color. Error bars for G10+SK16 and C11+SK16 are not shown and are indistinguishable from those of BS20. Middle: impact of bias corrections on real data relative to distances computed using G10+SK16. Bottom: impact of bias corrections using simulated data relative to distances computed using G10+SK16.

the SN. Our model does not differentiate whether the dust is in the circumstellar surrounding that is still linked to the progenitor or in the interstellar medium that is not linked to

the progenitor, but we can rule out a luminosity dependence on the progenitor system.

Relatedly, many studies have found correlations between spectral features and Hubble residuals (Fakhouri et al. 2015; Siebert et al. 2020). Interestingly, Wang et al. (2009) split a sample of 158 SNe Ia based on whether their spectra indicate "normal velocity" or "high velocity" features and find $R_V = 2.36 \pm 0.07$ and 1.57 ± 0.07 for the two subsamples, respectively. Pan et al. (2015) show that the velocity of spectral features correlates with the mass of the host galaxies, such that high-mass host galaxies regularly have high-velocity SNe, so one would expect low R_V to be found for high-mass hosts. This is in great agreement with the results of our study, though we note that Foley & Kasen (2011) show that different R_V from Wang et al. (2009) depend on using SNe with very red colors $E(B-V) > 0.4$. As velocity features have typically been thought of as indicative of properties of the progenitor and circumstellar surrounding, it is unclear at what level this is causally connected versus correlated.

As discussed in the introduction, circumstellar dust surrounding SNe Ia has been used to explain low values of R_V (< 2; e.g., Goobar 2008). Circumstellar dust has also been used to explain similarly low R_V values found for core-collapse SNe (Nugent et al. 2006; Goobar 2008). This is supported by our findings that it is common for SNe Ia to have $R_V < 2$, though as part of a larger range of R_V from 0.5 to 8. The interplay between SN radiation and nearby dust, and what can be learned by detailed studies of individual SNe of different types, will be an important avenue for supporting or refuting this possible explanation.

6.2. Application of BS20 Model in Future Analyses

While we have shown that biases in w from our model relative to previous models would have been the largest systematic uncertainty of previous analyses, there is a clear path to utilizing this model for future analyses. In order to do so optimally, there are three necessary improvements. First, a full

Table 3
Results from Large ΛCDM Simulations and 1445 SNe Ia

SN Ia Color Model Data[a]	Host Dust Model Data	SN Ia Color Model 1D BiasCor[b]	Host Dust Model 1D BiasCor	SN Ia + Dust β_{SALT}	wCDM + Planck '16 Δw[c]
BS20	BS20	C11 + SK16 Parent	No host dust	3.07 ± 0.01	−0.04
BS20	BS20	G10 + SK16 Parent	No host dust	3.09 ± 0.01	−0.03
BS20	BS20	BS20	BS20	3.12 ± 0.01	0.00
Real data	Real data	C11 + SK16 Parent	No host dust	3.06 ± 0.06	−0.04
Real data	Real data	G10 + SK16 Parent	No host dust	3.05 ± 0.06	−0.03
Real data	Real data	BS20	BS20	3.06 ± 0.06	0.00

Notes.
[a] Data sets are based on large simulations of ∼10,000 SNe Ia. Each data set (row) is a unique statistical realization.
[b] Bias correction samples are large simulations of >1,000,000 SNe Ia.
[c] $\Delta w = w_{fit} - w_{BS20}$: this is relative to the last row (BS20) of each data set grouping.

Bayesian fit to solve for the intrinsic and extrinsic parameters, broken by survey, redshift range, or targeted versus untargeted, is needed. This could be facilitated by the recent advancements by Pippin (Hinton & Brout 2020). Future work can fully constrain and characterize systematic uncertainties on all nine dust and hyperparameters using a combination of χ^2 metrics. Second, this model should be integrated into the SALT2 training, which currently only accounts for one component of the observed SN Ia color. Evidence of the benefit of retraining is shown in the appendix Figure 12. It will be necessary in the future to attempt to train SALT2 based on intrinsic and extrinsic color components.

Third, as discussed in Section 5, BBC5D from Kessler & Scolnic (2017) is not capable of bias-correcting two effective color-correlation coefficients (β_{SN} and R_V). Additionally, future analysis should not be correcting for an observed γ. Rather, dust distributions should be fit to different subsamples of host galaxies, and using this information, distance bias corrections can be computed as a function of observables (c, x_1, z, and host galaxy properties). A future approach to such bias corrections (Popovic et al. 2021) would be similar to that of Kessler & Scolnic (2017). If done properly, we predict that there will be no residual γ in the distance modulus residuals. Doing so will also improve the comparison of cosmological constraints in Section 5, where we had to assume naive mass-independent bias corrections. Ultimately one should compare the impact of the bias corrections from the two-mass model to the bias corrections from the G10 and C11 models when a luminosity correction due to host mass is applied.

The difference in rms for "dust-free" blue colors (rms ∼ 0.1) versus redder colors (rms ∼ 0.18) is striking. The statistical weight of these different SNe Ia when constraining dark energy with our improved color model shows that a blue SN Ia is ∼3× more constraining than a red SN. The blue SNe Ia exhibit an rms at the same level as NIR SN Ia standardized luminosities (Mandel et al. 2011). With tighter color measurement cuts and rstate-of-the-art samples (SNLS, DES), we have seen that the "dust-free" rms can even be as low as 0.08. In addition, as shown in Kessler & Scolnic (2017), bias corrections are much smaller for blue SNe Ia than for red SNe Ia. Additionally, because γ is found to be consistent with zero for the blue SNe, it appears that there are numerous advantages to using a sample of solely blue SNe. As LSST (Ivezić et al. 2019) and WFIRST (Spergel et al. 2015; Hounsell et al. 2018) will discover thousands of SNe Ia in this unextincted regime ($c \sim -0.1$), the
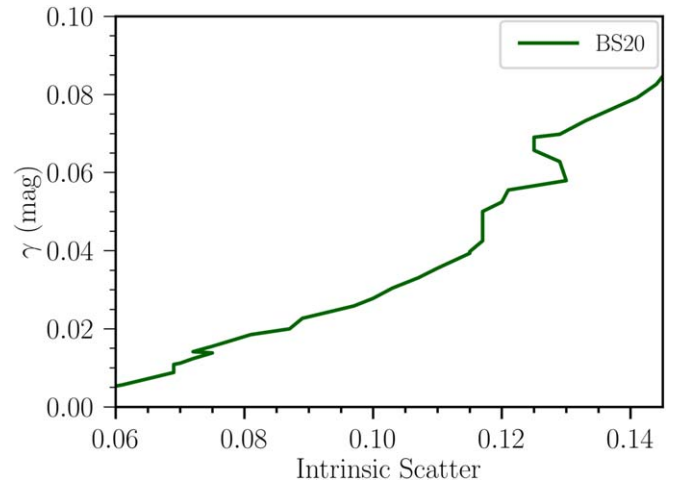
Figure 8. Simulations with the BS20 dust-based model predict a correlation between γ and observed intrinsic scatter. This correlation was originally seen for real data in B19b.

vast difference in constraining power and intrinsic scatter for the blue SNe Ia compared to the red SNe Ia should be considered in planning the survey strategy.

B19b showed an interesting trend that the magnitude of the recovered intrinsic scatter from various SN samples is correlated with the γ recovered from that sample. They also remarked that the σ_{int} value of the low-z sample was more than 3σ discrepant from that of the DES3YR sample. However, we show in Figure 8 that this behavior arises naturally from the BS20 model: both σ_{int} values are indeed consistent with a dust-based interpretation, and the relation between recovered σ_{int} and γ is a direct prediction of the BS20 model. This is because the different distributions of observed colors for each sample imply different amounts of dust, different amounts of intrinsic scatter, and different magnitudes of γ.

With our new model, we showed that the bias in recovered w due to assuming the incorrect scatter model is ∼0.04, larger than any other systematic uncertainty quantified in recent SN Ia cosmology analyses. In Brout et al. (2019b), the systematic uncertainty ascribed to this issue was determined by averaging distance modulus values after applying bias corrections based on both the G10+SK16 and C11+SK16 models, thus halving the difference between them, but they still found it to be one of the largest at $\sigma_w = 0.017$. As the sensitivity of cosmological

parameters to different scatter models is so large, we emphasize that this issue cannot be ignored in any future cosmological analysis. This statement is true for analyses of w and for analyses of H_0 as well. Dhawan et al. (2020) estimate biases due to scatter models to be on the level of 0.5%–1.0% in H_0. As the H_0 measurement has different systematic sensitivity than w due to the comparison of SNe in calibrator galaxies versus Hubble flow galaxies, we recommend that these two samples have similar demographics of blue and red SNe. A full systematics treatment, as done in Dhawan et al. (2020), should be done using the new dust-based SN Ia color model described in this paper. Furthermore, we note that past discussions (e.g., Rigault et al. 2013; Jones et al. 2018a) about potential biases in H_0 should be reconsidered in light of this paper's findings.

7. Conclusion

In this paper, we introduced a new, physical two-component color model of SNe Ia with an intrinsic component modeled as a simple symmetric Gaussian that correlates with SN Ia luminosity and an extrinsic component that can be modeled by a dust distribution that is tied to extinction by a wide R_V distribution. This model has fewer free parameters than previous models of SN Ia color and a more physical motivation that better matches the data. Our findings suggest that the dominant component of observed SN Ia intrinsic scatter is due to R_V variation of the dust around the SN. We also show that there is a 4.5σ dependence on color of the correlation of host mass with distance modulus residuals. Strikingly, this shows that previously observed host-galaxy property correlations with SN Ia luminosity are driven by the redder SNe of the sample. This also suggests a dust-based explanation for the host-galaxy property correlations. By allowing our model to have different parameters for the dust distributions of SNe in high-mass versus low-mass host galaxies, we show that the correlation between distance modulus residuals and host-galaxy stellar mass can be attributed to a 2.9σ difference in $\bar{R}_V$ between low and high mass.

By finding that the previously seen host-galaxy correlation with SN Ia luminosity after standardization is actually due to differences in dust, and not due to possible variation in the luminosity based on progenitor systems, we find that that there is a tremendous amount of leverage to continue to improve cosmological analyses with studies of larger samples, measurements covering larger wavelength ranges, more host-galaxy properties examined, and improved dust models. Our study shows that so many disparate analyses of SNe Ia are actually intricately connected, and unifying these studies will provide

tremendous improvements to measurements of the expansion of the universe.

Appendix

A.1. Model–Data Agreement and Parameter Sensitivity

Here we present the χ^2 surfaces from our iterative forward-modeling minimization process in Figure 9, and we show the impact of the various color-dust model parameters has on the three metrics. We list those variants here:

1. "BS20"—the main model proposed in this work
2. "No Dust"—a model with a narrow intrinsic color distribution and a weak ($\bar{\beta}_{SN} = 2$) correlation between color and luminosity
3. "Only Dust"—a model with only a dust distribution and a delta function for the intrinsic color distribution
4. "G10+SK16"—described in Section 3.1
5. "C11+SK16"—described in Section 3.1
6. "C11+SK16 + $\bar{\beta}_{SN}$ variation"—a model similar to the "C11+SK16" one, except we allow the β_{SN} to vary to reproduce the rms for redder colors
7. "BS20, $\sigma_{\beta_{SN}} = 0$"—the nominal BS20 model, except β_{SN} values are drawn from a delta function with value $\bar{\beta}_{SN}$
8. "BS20, $R_V + 0.5$"—the nominal BS20 model, except we shift our R_V distribution by the full sample by 0.5
9. "BS20, $\tau_E - 0.5$"—the nominal BS20 model, except we reduce τ_E to describe the dust distribution by 0.05
10. "BS20, $\bar{\beta}_{SN} + 0.5$"—the nominal BS20 model, except we increase $\bar{\beta}_{SN}$ by 0.5
11. "BS20, $\bar{\beta}_{SN} = 0$"—the nominal BS20 model, except we set β_{SN} to be 0. This effectively describes the intrinsic
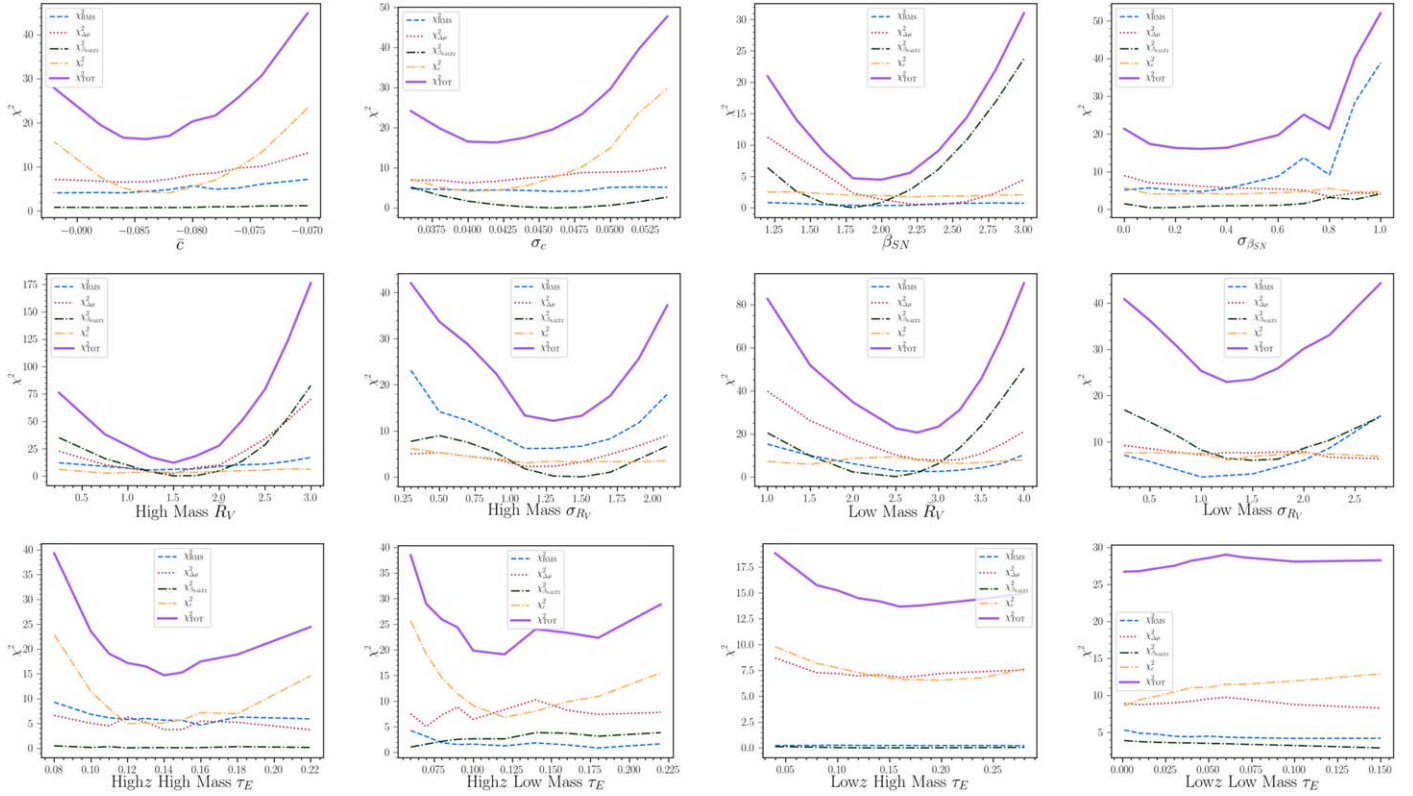
Figure 9. The χ^2 surfaces for each of the metrics in Equations (4), (5), (6), (7), and (8) are shown for each of the fitted parameters in the BS20 model. Each parameter is varied separately, with all other parameters held at their best-fit values.

color distribution as color scatter, similar to what is in the C11+SK16 model.

12. "BS20, No R_V variation"—the nominal BS20 model, except the variation in R_V is removed.

We show the results from using these different variants in Figure 10. We include on the bottom panel the recovered β_{SALT2} for each case because as some variants may have a good χ^2 in the three metrics, the recovered β_{SALT2} is far from that of the data (~ 3.05). It is important to note that besides the BS20, G10+SK16, and C11+SK16 models, none of the other models are fit to match the data.

A.2. Observed Correlations with SALT2 x_1

SN Ia cosmology analyses that measure correlations between SN light-curve parameters and host-galaxy mass regularly find a correlation between host-galaxy stellar mass and x_1 (e.g., Sullivan et al. 2010; Scolnic et al. 2014a). This correlation is shown in Figure 11(a) for our compiled data set, and from this, we expect trends with x_1 similar to what we observed with host stellar mass in Section 4. While there is no dependence of the rms of distance modulus residuals on x_1 (Figure 11(b)) seen in the data or predicted from simulations, we do see similar trends with color when splitting on x_1 (Figure 11(c)) as we do when splitting on $\mathcal{M}_{host}$. We also compute a Hubble residual step when splitting on x_1:

$$\delta\kappa = \kappa \times [1 + e^{(x_1 - x_{1_{step}})/0.01}]^{-1} - \frac{\kappa}{2}, \qquad (A1)$$

where an SN Ia stretch step location ($x_{1_{step}}$) of -0.5 is assumed, although $x_{1_{step}}$ values between -0.5 and $+0.5$ provide good discrimination between subsamples according to our three metrics. We determine κ for the sample in discrete color bins (Figure 11(d)). When deriving $\delta\kappa$ for the full sample with a single $x_{1_{step}}$ split, we find $\delta\kappa = 0.032 \pm .011$ mag, roughly half the size of the step when splitting by host stellar mass. As shown in Figure 11(d), similarly to host mass, the magnitude of κ depends on color; there is a 3.9σ deviation relative to a single step.

When examining Hubble diagram residual biases in bins of color (Figure 11(e)), simulations using the dust and R_V distributions that were fit in Section 4.2 roughly predict the residuals when splitting on x_1. This indicates that x_1 and $\mathcal{M}_{host}$ yield similar information about the dust properties. However, upon studying the mean Hubble residual bias with color, as shown in Figure 11(f), we find that the information from x_1 and $\mathcal{M}_{host}$ are complementary in potentially constraining R_V because the difference in Hubble residuals from the subsample of low x_1 values and large host mass values (purple) in comparison to those from high x_1 values and small mass values (orange) is larger than simply splitting on host mass (data points) as was done in Section 4. This finding is consistent with studies like Rose et al. (2019), who argue that combinations of various host-galaxy properties and light-curve parameters could further improve the standardizability of SNe Ia brightnesses, as well as with Galbany et al. (2012), who find that x_1 is a good discriminator of galaxy morphology.
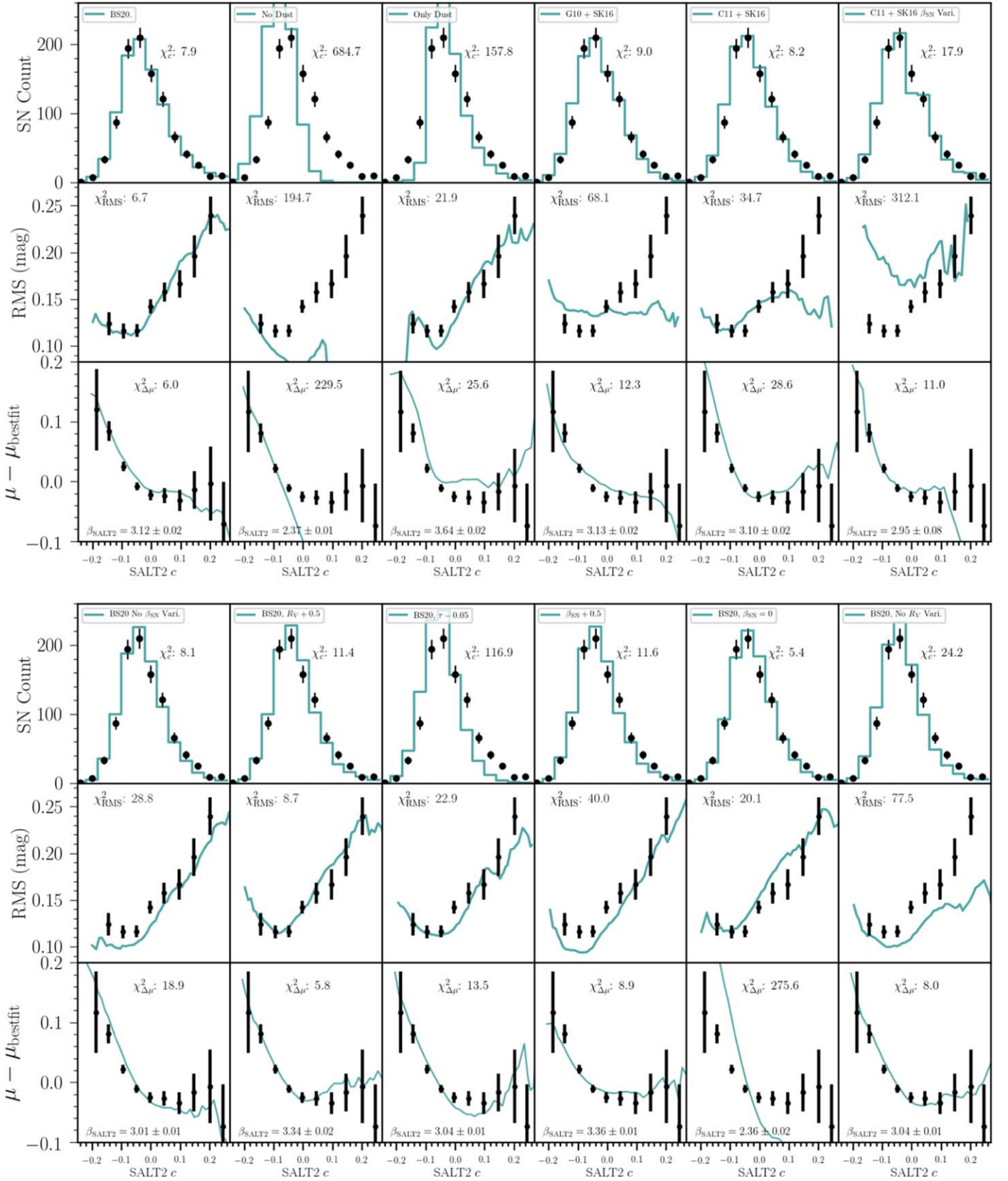
Figure 10. Similar to what is shown in Figures 4 and 5, but here we show the agreement between data and simulations for variants on our main BS20 mass-split model as well as for the G10+SK16 and C11+SK16 based models. The χ^2 for each metric is given, and in the bottom panel, we show the recovered β_{SALT2} to be compared with that from the data ($\sim$3.05). The sensitivity to these variations shows the high constraining power of the metrics.
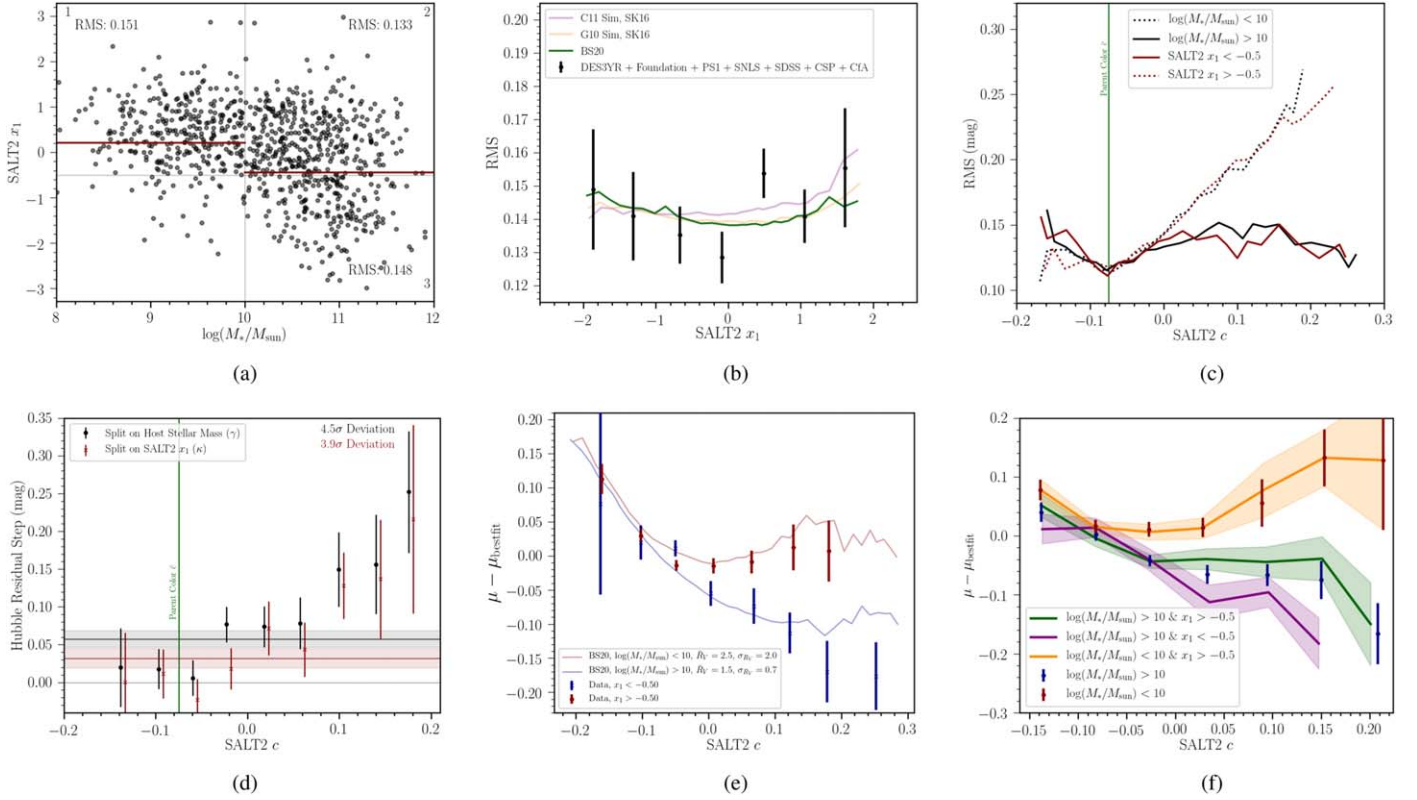
Figure 11. (a) Correlation between observed SALT2 x_1 and host-galaxy stellar mass. The trend shows a difference in weighted-average x_1 values (red) when split on $\mathcal{M}_{\rm step} = 10$. Hubble residual zero-mean rms values are reported for subsets of the data. (b) rms of Hubble diagram zero-mean residuals vs. SALT2 x_1. No dependence is seen. (c) rms of Hubble diagram residuals vs. SALT2 c when splitting on both $\log (M_*/M_{\rm sun})$ (black) and SALT2 x_1 (red). (d) Host-mass step as a function of observed color now with γ (black) and κ (red) overlaid. Simple averages and 1σ uncertainties are shown with horizontal lines. The significance of the deviation from the respective horizontal lines is reported in the text. (e) Binned Hubble diagram residuals for the data set when splitting on $x_1 = -0.5$ (points). The predictions using dust and R_V distributions from Section 4.2 are overlaid (lines). (f) Binned Hubble diagram residuals from three sectors of the data set (lines) corresponding to the three quadrants shown in panel (a). Overlaid are the binned Hubble diagram residuals used when only splitting on mass.
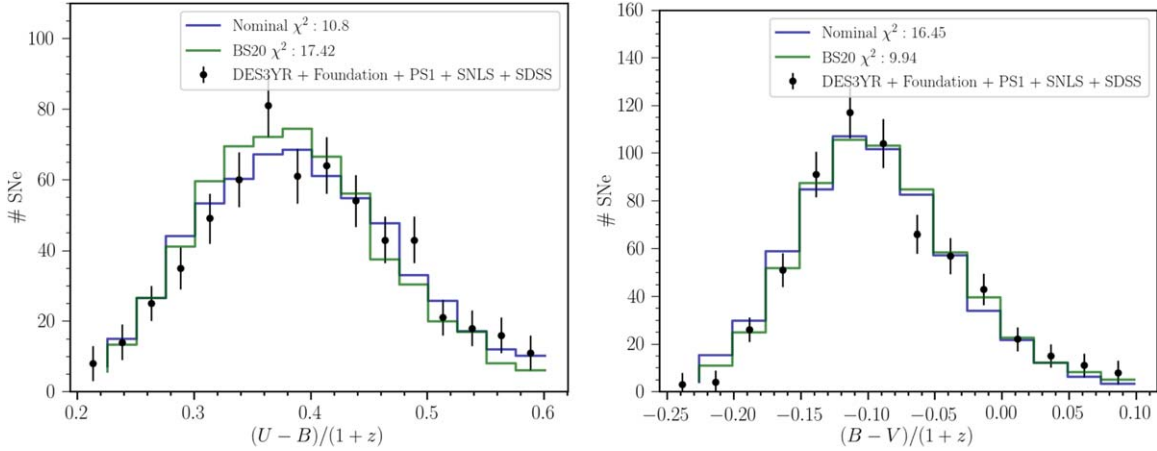
Figure 12. Predicted distribution of rest-frame colors in $(U - B)/(1 + z)$ and $(B - V)/(1 + z)$ from BS20 and the G10+SK16 model (nominal), where z is the observed redshift of the SNe. The χ^2 defined in Equation (3) is reported. Lesser agreement in $(U - B)$ for the BS20 model motivates retraining of the light-curve model in a future study.

A.3. SALT2 Color Law

In the discussion, we explain that a future analysis should retrain the color law(s) to match the data, rather than rely on our a posteriori model selection. In Figure 12, we derive the predicted distribution of peak rest-frame colors from our model and from a nominal SALT2-based color model (G10+SK16) and compare to the data. This is done by k-correcting

observations to the rest frame using the SALT2 spectral model. We show that the BS20 model better predicts the observed $(B - V)$ distribution, while G10+SK16 better predicts the observed $(U - B)$ distribution, both by similar amounts in χ^2. As the BS20 model selection had little sensitivity to the rest-frame UV colors, this evaluation is not surprising. It is moderately surprising, however, that given the lack of UV sensitivity in the metrics, the BS20 does as well as

it does in the UV. Still, we argue that a proper retraining of the light-curve model that incorporates flexibility for discrimination between the intrinsic SN Ia color law and dust color laws is needed in the future. Interestingly, Amanullah et al. (2015) shows that in the UV, a Fitzpatrick $R_V = 2.2$ matches observations of nearby SNe significantly better than the SALT2 color law does. As such, we expect that retraining based on our model can improve the plot shown here.

A.4. Similar but Nonviable Model

Recent studies proposed that there is more than one parent population of SN Ia color (Milne et al. 2013; Pan et al. 2014; Stritzinger et al. 2018; Jiang et al. 2018; Bulla et al. 2020; Kelsey et al. 2021; Gonzalez-Gaitan et al. 2020) with differing β values. To illustrate, we follow the optimization process described in Section 2.4 and fit for parameters of a two-component color population model. We assume that SN Ia colors are drawn from a bluer Gaussian ($c_{\rm int}^{\rm blue}$, $\sigma_c^{\rm blue}$) and a redder Gaussian ($c_{\rm int}^{\rm red}$, $\sigma_c^{\rm red}$). A larger $\sigma_c^{\rm red}$ in comparison to $\sigma_c^{\rm blue}$ results in the asymmetric histogram seen in the data (Figure 13 top). Each color population is given its own color–luminosity relation ($\beta^{\rm blue}$ and $\beta^{\rm red}$) to explain the color-dependent residuals in the Hubble diagram (Figure 13 bottom). In order to explain the scatter floor in the data (Figure 13 middle), we must include a $\sigma_{\rm coh}$ term, and to explain the increasing rms as a function of color, we allow for $\sigma_\beta^{\rm blue}$ and $\sigma_\beta^{\rm red}$.

A simulation with the best-fit population parameters is shown in the pink curves of Figure 13. We find that while the color distribution (top) and the trend of Hubble residuals with color (bottom) are reproduced, we are not able to reproduce the trend of rms scatter with color (middle). We note that for the multiple-population model, the color at which the Hubble diagram scatter is smallest occurs at the mean observed color. This is the expectation arising from the assumption of two symmetric parent populations. However, this is not what is observed in the data. In the data, the color corresponding to the minimum scatter in the Hubble diagram is -0.085, far bluer than the mean observed color (-0.016); this is indicative of a one-sided color variance (i.e., dust). Our best-fit dust model is shown for comparison (teal) and agrees with this expectation.

Furthermore, we note that this multiple-population model is more complicated in its physical underpinnings than a model where there is an intrinsic color and dust. For this new model, there remains no unanimously accepted reason for why there would be two populations of SNe with differing colors and differing color–luminosity relations. Additionally, this model has one more parameter (nine) than the model we propose (eight), and despite its added complexity, the agreement with the observed data is poorer.

Lastly, as is explained in Section 4.2, an additional correlation between dust-model parameters (R_V) and host properties can fully explain the mass step. Likewise, if one is motivated to use the multiple-population model to explain the mass step, additional complexity must be added. From Figure 6(b), there are clearly three separate slopes in the Hubble residuals versus color when split on host-galaxy mass. Therefore, in the multiple-population methodology, this implies that three separate populations of SNe with three β values would be needed to explain the mass step. Bimodal populations motivated by two different progenitor scenarios are inconsistent with this finding of three separate slopes.
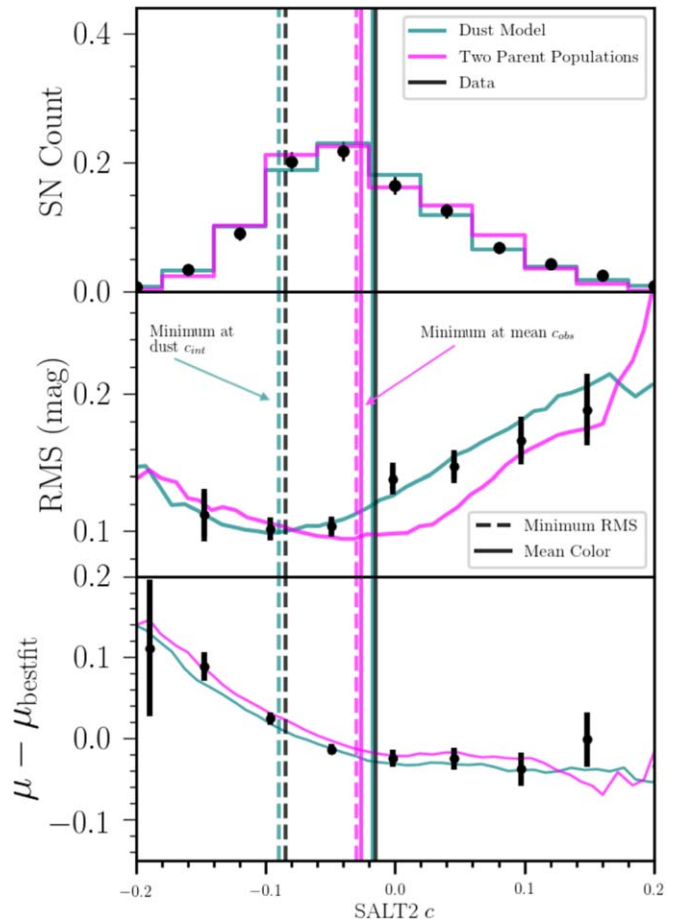
Figure 13. Same as Figure 3 but with the inclusion of a multiple-color population model (pink). The multiple-color population model reproduces the trends observed in the data for the color distribution (top panel) and Hubble diagram residuals (bottom), but does not reproduce the scatter in the Hubble diagram as a function of color (middle). The minimum scatter in the data (vertical black dashed line) and in the dust model (vertical teal dashed line) occurs bluer than the mean observed colors (black/teal solid vertical lines). However, this is not the case for the multiple-color population model, which has a minimum scatter (pink dashed vertical line) occurring at the mean observed color (pink solid vertical line).

ORCID iDs

Dillon Brout ● https://orcid.org/0000-0001-5201-8374

References

Aldering, G., Adam, G., Antilogus, P., et al. 2002, Proc. SPIE, 4836, 61
Amanullah, R., Johansson, J., Goobar, A., et al. 2015, MNRAS, 453, 3300
Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22
Brout, D., Sako, M., Scolnic, D., et al. 2019a, ApJ, 874, 106
Brout, D., Scolnic, D., Kessler, R., et al. 2019b, ApJ, 874, 150
Bulla, M., Goobar, A., Amanullah, R., Feindt, U., & Ferretti, R. 2018a, MNRAS, 473, 1918
Bulla, M., Goobar, A., & Dhawan, S. 2018b, MNRAS, 479, 3663
Bulla, M., Miller, A. A., Yao, Y., et al. 2020, ApJ, 902, 48
Burns, C. R., Parent, E., Phillips, M. M., et al. 2018, ApJ, 869, 56
Childress, M., Aldering, G., Antilogus, P., et al. 2013a, ApJ, 770, 107
Childress, M., Aldering, G., Antilogus, P., et al. 2013b, ApJ, 770, 108
Childress, M. J., Wolf, C., & Zahid, H. J. 2014, MNRAS, 445, 1898
Chotard, N., Gangler, E., Aldering, G., et al. 2011, A&A, 529, L4
Cikota, A., Deustua, S., & Marleau, F. 2016, ApJ, 819, 152
Dhawan, S., Brout, D., Scolnic, D., et al. 2020, ApJ, 894, 54
Fakhouri, H. K., Boone, K., Aldering, G., et al. 2015, ApJ, 815, 58
Fitzpatrick, E. L. 1999, PASP, 111, 63
Foley, R. J., & Kasen, D. 2011, ApJ, 729, 55

Foley, R. J., Scolnic, D., Rest, A., et al. 2018, MNRAS, 475, 193
Galbany, L., Miquel, R., Östman, L., et al. 2012, ApJ, 755, 125
Gao, J., Jiang, B. W., Li, A., Li, J., & Wang, X. 2015, ApJL, 807, L26
Gao, J., Jiang, B. W., Li, A., & Xue, M. Y. 2013, ApJ, 776, 7
Gonzalez-Gaitan, S., de Jaeger, T., Galbany, L., et al. 2020, arXiv:2009.13230
Goobar, A. 2008, ApJL, 686, L103
Guy, J., Sullivan, M., Conley, A., et al. 2010, A&A, 523, A7
Hicken, M., Challis, P., Jha, S., et al. 2009a, ApJ, 700, 331
Hicken, M., Challis, P., Kirshner, R. P., et al. 2012, ApJS, 200, 12
Hicken, M., Wood-Vasey, W. M., Blondin, S., et al. 2009b, ApJ, 700, 1097
Hinton, S., & Brout, D. 2020, JOSS, 5, 2122
Hoang, T. 2017, ApJ, 836, 13
Holwerda, B. W., Reynolds, A., Smith, M., & Kraan-Korteweg, R. C. 2015, MNRAS, 446, 3768
Hounsell, R., Scolnic, D., Foley, R. J., et al. 2018, ApJ, 867, 23
Jha, S., Riess, A. G., & Kirshner, R. P. 2007, ApJ, 659, 122
Jiang, J.-a., Doi, M., Maeda, K., & Shigeyama, T. 2018, ApJ, 865, 149
Jones, D. O., Riess, A. G., & Scolnic, D. M. 2015, ApJ, 812, 31
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
Ivezić, Ž., Kahn, S. M., Tyson, J. A., et al. 2019, ApJ, 873, 111
Kelsey, L., Sullivan, M., Smith, M., et al. 2021, MNRAS, 501, 4861
Kessler, R., Bernstein, J. P., Cinabro, D., et al. 2009, PASP, 121, 1028
Kessler, R., Brout, D., Crawford, S., et al. 2019, MNRAS, 485, 1171
Kessler, R., Guy, J., Marriner, J., et al. 2013, ApJ, 764, 48
Kessler, R., & Scolnic, D. 2017, ApJ, 836, 56
Krisciunas, K., Prieto, J. L., Garnavich, P. M., et al. 2006, AJ, 131, 1639
Lampeitl, H., Smith, M., Nichol, R. C., et al. 2010, ApJ, 722, 566
Lewis, A., & Bridle, S. 2002, PhRvD, 66, 103511
Mandel, K. S., Narayan, G., & Kirshner, R. P. 2011, ApJ, 731, 120
Mandel, K. S., Scolnic, D. M., Shariff, H., Foley, R. J., & Kirshner, R. P. 2017, ApJ, 842, 93
Marriner, J., Bernstein, J. P., Kessler, R., et al. 2011, ApJ, 740, 72
Milne, P. A., Brown, P. J., Roming, P. W. A., Bufano, F., & Gehrels, N. 2013, ApJ, 779, 23
Nobili, S., & Goobar, A. 2008, A&A, 487, 19
Nozawa, T. 2016, P&SS, 133, 36
Nugent, P., Sullivan, M., Ellis, R., et al. 2006, ApJ, 645, 841
Pan, Y.-C., Sullivan, M., Maguire, K., et al. 2014, MNRAS, 446, 354
Pan, Y. C., Sullivan, M., Maguire, K., et al. 2015, MNRAS, 446, 354
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Phillips, M. M., Simon, J. D., Morrell, N., et al. 2013, ApJ, 779, 38
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
Popovic, B., Brout, D., Kessler, R., Scolnic, D., & Lu, L. 2021, arXiv:2102.01776
Rest, A., Scolnic, D., Foley, R. J., et al. 2014, ApJ, 795, 44
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009
Rigault, M., Copin, Y., Aldering, G., et al. 2013, A&A, 560, A66
Roman, M., Hardin, D., Betoule, M., et al. 2018, A&A, 615, A68
Rose, B. M., Garnavich, P. M., & Berg, M. A. 2019, ApJ, 874, 32
Sako, M., Bassett, B., Connolly, B., et al. 2011, ApJ, 738, 162
Salim, S., Boquien, M., & Lee, J. C. 2018, ApJ, 859, 11
Schlafly, E. F., Meisner, A. M., Stutz, A. M., et al. 2016, ApJ, 821, 78
Scolnic, D., Casertano, S., Riess, A., et al. 2015, ApJ, 815, 117
Scolnic, D., & Kessler, R. 2016, ApJL, 822, L35
Scolnic, D., Rest, A., Riess, A., et al. 2014a, ApJ, 795, 45
Scolnic, D., Smith, M., Massiah, A., et al. 2020, ApJL, 896, L13
Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, ApJ, 859, 101
Scolnic, D. M., Riess, A. G., Foley, R. J., et al. 2014b, ApJ, 780, 37
Siebert, M. R., Foley, R. J., Jones, D. O., & Davis, K. W. 2020, MNRAS, 493, 5713
Smith, M., Sullivan, M., Wiseman, P., et al. 2020, MNRAS, 494, 4426
Spergel, D., Gehrels, N., Baltay, C., et al. 2015, arXiv:1503.03757
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, ApJL, 864, L35
Sullivan, M., Conley, A., Howell, D. A., et al. 2010, MNRAS, 406, 782
Tripp, R. 1998, A&A, 331, 815
Uddin, S. A., Mould, J., Lidman, C., Ruhlmann-Kleider, V., & Zhang, B. R. 2017, ApJ, 848, 56
Wang, L. 2005, ApJL, 635, L33
Wang, X., Filippenko, A. V., Ganeshalingam, M., et al. 2009, ApJL, 699, L139
Wang, X., Wang, L., Zhou, X., Lou, Y.-Q., & Li, Z. 2005, ApJL, 620, L87
Wiseman, P., Smith, M., Childress, M., et al. 2020, MNRAS, 495, 4040
Yanchulova Merica-Jones, P., Sandstrom, K. M., Johnson, L. C., et al. 2017, ApJ, 847, 102