The Astrophysical Journal Letters, 952:L8 (6pp), 2023 July 20 https://doi.org/10.3847/2041-8213/ace34d

Binning is Sinning: Redemption for Hubble Diagram Using Photometrically Classified Type Ia Supernovae

R. Kessler[1,2], M. Vincenzi[3], and P. Armstrong[4]
[1] Kavli Institute for Cosmological Physics, University of Chicago, Chicago, IL 60637, USA
[2] Department of Astronomy and Astrophysics, University of Chicago, Chicago, IL 60637, USA
[3] Department of Physics, Duke University, Durham, NC 27708, USA
[4] Mt Stromlo Observatory, The Research School of Astronomy and Astrophysics, Australian National University, Stromlo, ACT 2601, Australia; kessler@kicp.uchicago.edu
Received 2023 June 9; revised 2023 June 30; accepted 2023 June 30; published 2023 July 18

Abstract

Bayesian Estimation Applied to Multiple Species (BEAMS) is implemented in the BEAMS with Bias Corrections (BBC) framework to produce a redshift-binned Hubble diagram (HD) for Type Ia supernovae (SNe Ia). BBC corrects for selection effects and non–SN Ia contamination, and systematic uncertainties are described by a covariance matrix with dimension matching the number of BBC redshift bins. For spectroscopically confirmed SN Ia samples, a recent "Binning is Sinning" article showed that an unbinned HD and covariance matrix reduces the systematic uncertainty by a factor of ∼1.5 compared to the binned approach. Here we extend their analysis to obtain an unbinned HD for a photometrically identified sample processed with BBC. To test this new method, we simulate and analyze 50 samples corresponding to the Dark Energy Survey (DES) with a low-redshift anchor; the simulation includes SNe Ia, and contaminants from core-collapse SNe and peculiar SNe Ia. The analysis includes systematic uncertainties for calibration and measures the dark energy equation of state parameter (w). Compared to a redshift-binned HD, the unbinned HD with nearly 2000 events results in a smaller systematic uncertainty, in qualitative agreement with BHS21, and averaging results among the 50 samples we find no evidence for a w-bias. To reduce computation time for fitting an unbinned HD with large samples, we propose an HD-rebinning method that defines the HD in bins of redshift, color, and stretch; the rebinned HD results in similar uncertainty as the unbinned case, and shows no evidence for a w-bias.

Unified Astronomy Thesaurus concepts: Supernovae (1668); Cosmology (343); Hubble diagram (759); Dark energy (351)

1. Introduction

Following the discovery of cosmic acceleration using a few dozen Type Ia supernovae (SNe Ia) (Riess et al. 1998; Perlmutter et al. 1999), increasingly large SN Ia samples have been used to improve measurements of the dark energy equation of state parameter, w. While the most precise w measurements are based on spectroscopically confirmed samples, photometric imaging surveys have been discovering far more supernovae than spectroscopic resources can observe. Existing SN surveys include the Sloan Digital Sky Survey-II (SDSS),[5] Supernova Legacy Survey (SNLS), Panoramic Survey Telescope and Rapid Response System-1 (PS1),[6] and Dark Energy Survey (DES);[7] future wide-area surveys that will overwhelm spectroscopic resources include the Legacy Survey of Space and Time (LSST)[8] and the Nancy Grace Roman Space Telescope.[9]

To make full use of these large SN Ia samples, photometric identification using broadband filters has been developed over the past decade. Photometric methods include template matching (Sako et al. 2011) and machine learning (Lochner et al. 2016; Möller & de Boissière 2020; Qu et al. 2021), and they determine the probability ($P_{\rm Ia}$) for each event to be an SN Ia. A framework to incorporate the resulting $P_{\rm Ia}$ was developed to measure cosmological parameters. This framework, called "Bayesian Estimation Applied to Multiple Species" (BEAMS; Kunz et al. 2007; Hlozek et al. 2012), was first used in an SN-cosmology analysis for the PS1 photometric sample (Jones et al. 2018).

As part of the DES SN-cosmology analysis, BEAMS was extended to "BEAMS with Bias Corrections" (BBC; Kessler & Scolnic 2017; hereafter KS17), a fitting procedure designed to produce a Hubble diagram (HD) that is corrected for selection biases and for contamination from core-collapse SNe (SNCC) and peculiar SNe Ia. BBC has been used in the SN Ia cosmology analysis for spectroscopic samples from Pantheon (Scolnic et al. 2018), DES (DES Collaboration 2019), and Pantheon+ (Brout et al. 2022a). BBC has also been used on a photometric sample from PS1 (Jones et al. 2019) and to examine contamination biases for the photometric DES sample (Vincenzi et al. 2023).

The BBC fit is performed in redshift bins to determine nuisance parameters and SN Ia distances that are independent of cosmological parameters, which enables more flexible use of cosmology-fitting programs. BBC therefore produces both a redshift binned and unbinned HD. Previous analyses using BBC took advantage of the binned HD to reduce computation time in cosmology-fitting programs. However, Brout et al.

[5] https://www.sdss.org/dr12/data_access/supernovae
[6] https://panstarrs.stsci.edu
[7] https://www.darkenergysurvey.org
[8] https://www.lsst.org
[9] https://roman.gsfc.nasa.gov

(2021, hereafter BHS21) showed that while the statistical uncertainty is the same using a binned or unbinned HD, the systematic uncertainty is ∼1.5 times smaller using an unbinned HD and unbinned covariance matrix. The uncertainty reduction is from an effect known as self-calibration (Faccioli et al. 2011).

BHS21 demonstrated the uncertainty reduction using BBC with a spectroscopically confirmed sample. Here we expand the use of unbinned HDs to photometric samples where BEAMS is used. In anticipation of very large samples in future analyses, we also explore the possibility of reducing computation time with a smaller HD and covariance matrix and still benefit from self-calibration: a rebinned HD in the space of redshift, color, and stretch. This choice of variables is motivated by the color-dependent systematic explored in BHS21. While the unbinned approach is optimal, the rebinned approach may be useful for the many intermediate simulation tests prior to unblinding.

We validate the unbinning and rebinning methods using simulations of DES that include SNe Ia, SNCC, and peculiar SNe Ia. The simulation and analysis presented here are similar to those in Vincenzi et al. (2023), and all analysis software used in this analysis is publicly available. The software for simulations, light-curve fitting, BBC, and cosmology fitting is from the SuperNova ANAlysis package (`SNANA`; Kessler et al. 2009).[10] The photometric classification software is from `SuperNNova` (SNN; Möller & de Boissière 2020).[11] For workflow orchestration we used `Pippin` (Hinton & Brout 2020).[12]

The outline of this Letter is as follows. The SALT2 and BBC formalism is reviewed in Section 2. The unbinning and rebinning procedures are presented in Section 3. The validation analysis is described in Section 4, and the validation results are given in Section 5.

2. Review of SALT2 and BBC

Using the SALT2 framework (Guy et al. 2010), BBC is a fitting procedure that delivers an HD corrected for selection effects and corrected for contamination. BBC incorporates three main features: (1) BEAMS (Kunz et al. 2007), (2) fitting in redshift bins to avoid dependence on cosmological parameters (Marriner et al. 2011), and (3) detailed simulation to correct distance biases (Kessler et al. 2019a).

For each event, a SALT2 light-curve fit determines the time of peak brightness (t_0), stretch (x_1), color parameter (c), and amplitude (x_0) with $m_B \equiv -2.5\log_{10}(x_0)$. Within BBC, the measured distance modulus is defined by the Tripp equation (Tripp 1998):

$$\mu = m_B + \alpha x_1 - \beta c + \mathcal{M} - \Delta\mu_{\rm bias}\,, \tag{1}$$

where α and β are the stretch- and color-luminosity parameters, $\mathcal{M}$ is a global offset, and $\Delta\mu_{\rm bias}$ is a distance-bias correction for each event, $\mu - \mu_{\rm true}$, determined from a large simulation. The $\Delta\mu_{\rm bias}$ value for each event is evaluated by interpolating in a five-dimensional space of $\{z, x_1, c, \alpha, \beta\}$. In Equation (1), there is an implicit SN index i for μ, m_B, x_1, c, and $\Delta\mu_{\rm bias}$; this index is suppressed for readability. To simplify this study, host–SN correlations have been ignored in Equation (1), and also in the simulations used for validation.

Following Section 5 of KS17 and making a few simplifications for this review, the BBC fit maximizes a likelihood of the form $\mathcal{L} = \prod_{i=1}^{N} \mathcal{L}_i$, where $\mathcal{L}_i$ for event i is

$$\mathcal{L}_i = P_{{\rm Ia},i} D_{{\rm Ia},i} + (1 - P_{{\rm Ia},i}) D_{{\rm CC},i}\,, \tag{2}$$

where $P_{{\rm Ia},i}$ is the photometric classification probability for event i to be an SN Ia. The SN Ia component of $\mathcal{L}_i$ is $D_{\rm Ia} \sim \exp[-\chi^2_{\rm HR}/2]$, where $\chi^2_{\rm HR} = {\rm HR}^2/\sigma^2_\mu$, HR is a Hubble residual described below, and σ_μ is the uncertainty on μ in Equation (1) as shown in Equation (3) of KS17. The non–SN Ia (contamination) component, $D_{\rm CC}$, is evaluated from a simulation.

To remove the dependence on cosmological parameters in the BBC fit, we follow Marriner et al. (2011) and define the Hubble residual for the ith SN ($\rm HR_i$) as

$$\mathrm{HR}_i \equiv \mu_i - [\mu_{\rm ref}(z_i, \text{?}_{\rm ref}) + M_\zeta], \tag{3}$$

where $\mu_{\rm ref} = \mu_{\rm ref}(z_i, \text{?}_{\rm ref})$ is a reference distance computed from redshift z_i and an arbitrary choice of reference cosmology parameters denoted by $\text{?}_{\rm ref}$. Our choice for $\text{?}_{\rm ref}$ is flatness and

$$\text{?}_{\rm ref} \equiv \{\Omega_{\rm M} = 0.3,\, w = -1\}. \tag{4}$$

M_ζ are fitted distance offsets in redshift bins denoted by ζ. An important concept is that using a cosmological model for $\mu_{\rm ref}$ in Equation (3) is a convenience, not a necessity. For example, $\mu_{\rm ref}$ could be replaced with a polynomial function of redshift or any function that approximates the distance-redshift relation within each redshift bin.

The BBC fit determines α, β, γ, M_ζ, and an intrinsic scatter term ($\sigma_{\rm int}$) added to the distance uncertainties (Section 3) that results in a reduced χ^2 of one. The final binned HD is obtained as follows. First, each binned redshift (z_ζ) is computed from

$$z_\zeta = \mu^{-1}[\overline{\mu_{{\rm ref}\,\zeta}}], \tag{5}$$

where μ^{-1} is an inverse-distance function that numerically determines redshift from the weighted average of $\mu_{\rm ref}$ in redshift bin ζ. The weight for each event is σ_μ^{-2}. Next, the BBC-fitted distance in each redshift bin (μ_ζ) is

$$\mu_\zeta = \overline{\mu_{{\rm ref}\,\zeta}} + M_\zeta. \tag{6}$$

and the collection of $\{z_\zeta, \mu_\zeta\}$ is the binned HD corrected for selection effects and contamination. The uncertainty on μ_ζ is the BBC-fitted uncertainty for M_ζ. If a different choice of ? is used for $\mu_{\rm ref}$, the fitted M_ζ will change but the μ_ζ remain the same. For spectroscopically confirmed samples with all $P_{{\rm Ia},i} = 1$, the unbinned HD is the collection of $\{z_i, \mu_i\}$ where the μ_i are computed from Equation (1) using the BBC-fitted parameters and each distance uncertainty ($\sigma_{\mu,i}$) is computed from Equation (3) in KS17. This procedure is an approximation that we rigorously test (Section 5.1) with high-statistics simulations.

3. Unbinning and Rebinning after BBC Fit

For an unbinned HD, we use the BBC-fitted parameters and compute the distances defined in Equation (1). The unbinned distance uncertainties ($\sigma_{\mu,{\rm unbin},i}$), however, are not the naively computed distance uncertainties ($\sigma_{\mu,i}$) for a spectroscopically

[10] https://github.com/RickKessler/SNANA
[11] https://github.com/supernnova
[12] https://github.com/dessn/Pippin

confirmed sample. To determine $\sigma_{\mu,\mathrm{unbin},i}$, we require that the weighted average uncertainty in each redshift bin is equal to $\sigma_{M,\zeta}$, the BBC-fitted uncertainty on M_ζ;

$$1/\sigma^2_{M,\zeta} = \sum_{i\in\zeta} 1/\sigma^2_{\mu,\mathrm{unbin},i} \quad (7)$$

$$= \sum_{i\in\zeta} \mathcal{P}_{\mathrm{B(Ia)},i}/[S_\zeta \sigma_{\mu,i}]^2, \quad (8)$$

where i is the SN index within redshift bin ζ,

$$\mathcal{P}_{\mathrm{B(Ia)},i} = \frac{P_{\mathrm{Ia},i} D_{\mathrm{Ia},i}}{P_{\mathrm{Ia},i} D_{\mathrm{Ia},i} + (1 - P_{\mathrm{Ia},i}) D_{\mathrm{CC},i}} \quad (9)$$

is the BEAMS probability for event i to be an SN Ia, and S_ζ is a ζ-dependent uncertainty scale that is computed to satisfy Equation (8). We find that S_ζ is a few percent greater than 1 because of small correlations between the fitted parameters (α, β, M_ζ). Equation (8) is an ad hoc assumption and does not have a rigorous derivation. From Equations (7)–(8), the unbinned distance uncertainty is

$$\sigma_{\mu,\mathrm{unbin},i} = S_\zeta \sigma_{\mu,i} / \sqrt{\mathcal{P}_{\mathrm{B(Ia)},i}}. \quad (10)$$

As a crosscheck, the weighted average of the Hubble residuals ($\langle \mathrm{HR}_\zeta \rangle$) should be zero for each redshift bin:

$$\langle \mathrm{HR}_\zeta \rangle = \left[\sum_{i\in\zeta} \mathrm{HR}_i W_i\right] / \left[\sum_{i\in\zeta} W_i\right], \quad (11)$$

where $W_i = \sigma^{-2}_{\mu,\mathrm{unbin},i}$ and HR_i is defined in Equation (3).

To limit the size of the HD and still benefit from reduced systematics, we propose rebinning in the space of redshift, stretch, and color, denoted by $\boldsymbol{\zeta} = \{z,\ x_1,\ c\}$. The distance modulus in each 3D $\boldsymbol{\zeta}$ cell is a weighted average of distances in the cell,

$$\mu_{\boldsymbol{\zeta}} = \sum_{i\in\boldsymbol{\zeta}} \mu_i W_i / \sum_{i\in\boldsymbol{\zeta}} W_i, \quad (12)$$

and following Equation (7) the uncertainty on $\mu_{\boldsymbol{\zeta}}$ is

$$1/\sigma^2_{\mu,\boldsymbol{\zeta}} = \sum_{i\in\boldsymbol{\zeta}} 1/\sigma^2_{\mu,\mathrm{unbin},i}. \quad (13)$$

4. Validation I: Simulation and Analysis

We test the unbinning and rebinning procedure (Section 3) by analyzing 50 simulated data-sized samples that closely follow Vincenzi et al. (2023). Each simulation corresponds to the 5 yr DES photometric sample for events with an accurate spectroscopic redshift of the host galaxy, combined with a spectroscopically confirmed low-redshift (LOWZ) sample ($z < 0.1$). The LOWZ sample uses the cadence and signal-to-noise ratio (S/N) properties for the Carnegie Supernova Project (CSP),[13] Center for Astrophysics (CFA3, CFA4; Hicken et al. 2009, 2012), and Foundation Supernova Survey (Foley et al. 2018).

The simulated models include:

1. SNe Ia generated from the SALT2 model in Guy et al. (2010) using trained model parameters from Betoule et al. (2014);
2. SNCC generated from spectral energy distribution (SED) templates in Vincenzi et al. (2019);
3. Peculiar SN Iax using the SED model from Kessler et al. (2019b)[14] and extinction correction from Vincenzi et al. (2021);
4. Peculiar 91bg-like SNe Ia using the SED model from Kessler et al. (2019b); and
5. ΛCDM with $\Omega_\mathrm{M} = 0.311$, $\Omega_\Lambda = 0.689$, and $w = -1$.

All simulated events are analyzed as follows:

1. Use the SALT2 light-curve fit for each event to determine $\{t_0,\ m_B,\ x_1,\ c\}$.
2. Apply selection requirements (cuts):
   (a) At least two passbands with maximum S/N >5;
   (b) At least one observation before t_0;
   (c) At least one observation >10 days after t_0 (rest frame);
   (d) $|x_1| < 3$ and $|c| < 0.3$;
   (e) Fitted uncertainties $\sigma_{x1} < 1.0$ and $\sigma_{t0} < 2$ days;
   (f) SALT2 light-curve fit probability >0.001; and
   (g) Valid bias correction in the BBC fit (see below).

For the 50 samples, the average number of events passing cuts is 1897 (1622, 172, and 103 for DES, Foundation, and LOWZ, respectively).

3. Determine P_Ia using the "SuperNNova" photometric classification (Möller & de Boissière 2020)[15] based on recurrent neural networks.
4. Use the BBC fit to determine redshift-binned HD corrected for selection effects and non–SN Ia contamination. We use 20 z-bins, with bin size proportional to $(1+z)^3$ so that there is finer z-binning at lower redshift.
5. Create statistical+systematic covariance matrix as in Conley et al. (2011).
6. Use methods from Section 3 to produce an unbinned HD and two rebinned HDs. The first rebinned HD has 2 stretch and 4 color bins (Rebin2x4), and the total number of HD bins is $20 \times 2 \times 4 = 160$. The second rebinned HD has 4 stretch and 8 color bins (Rebin4x8) and a total of 640 HD bins.
7. Perform cosmology fit using a fast minimization program that combines an SN Ia HD with a cosmic microwave background prior that uses an R-shift parameter computed from the same cosmology parameters as in the simulated samples. To match the constraining power from Planck Collaboration et al. (2020), the R-uncertainty is $\sigma_R = 0.006$. We fit for Ω_M and w using the wCDM model, and we also fit for Ω_M, w_0, and w_a using the $w_0 w_a$CDM model.
8. For the $w_0 w_a$CDM model, the figure of merit (FoM) is computed based on the dark energy task force (DETF) definition in Albrecht et al. (2006),

$$\mathrm{FoM} = [\sigma(w_0) \times \sigma(w_a) \times \sqrt{1 - \rho^2}]^{-1}, \quad (14)$$

where ρ is the reduced covariance between w_0 and w_a.

Here we consider 70 systematic uncertainties that include the following:

1. For each of the 34 passbands, shift the zero-point using the uncertainty from Brout et al. (2022b).

[13] https://csp.obs.carnegiescience.edu

[14] https://github.com/RutgersSN/SNIax-PLAsTiCC

[15] https://github.com/supernnova

Table 1
STD$_\mu$ for Different ? Choices

? Variant	STD$_\mu \times 10^4$
$\Omega_M = 0.20$	1.0
$\Omega_M = 0.25$	1.7
$\Omega_M = 0.35$	0.4
$\Omega_M = 0.40$	0.7
$w = -1.2$	0.7
$w = -1.1$	0.5
$w = -0.9$	0.3
$w = -0.8$	0.6
p0 ($\mu_{i,?} = 40$)	59
p3[a]	6.8
p6[b]	3.1

Notes.
[a] Third-order polynomial fit to the ΛCDM model.
[b] Sixth-order polynomial fit to the ΛCDM model.

2. For each of the 34 passbands, shift the filter transmission wavelength using the uncertainty from Brout et al. (2022b).
3. Correlated zero-point shift, 0.00714λ/micron, corresponding to the Hubble Space Telescope calibration uncertainty for primary reference C26202.
4. Galactic extinction uncertainty is 4%.

For each of the 68 zero-point and wavelength systematics, the SALT2 model is retrained and the shift is propagated in the simulated data. M. Vincenzi et al. (2023, in preparation) present the complete set of systematic uncertainties that includes calibration covariances.

5. Validation II: Results

5.1. BBC Sensitivity to Reference Cosmology

We begin by evaluating the sensitivity of fitted BBC distances to the reference cosmology ?$_{\rm ref}$ defined in Equation (3). Using the wCDM model to vary ?$_{\rm ref}$, we vary Ω_M up to ± 0.1 with fixed $w = -1$, and we vary w by up to ± 0.2 with fixed $\Omega_M = 0.3$. We define a sensitivity metric to be

$$\mathrm{STD}_\mu = \mathrm{STD}(\mu_{i,?} - \mu_{i,?_{\rm ref}}), \quad (15)$$

where STD is the standard deviation, $\mu_{i,?_{\rm ref}}$ are unbinned distances (Equation (1)) using ?$_{\rm ref}$ (Equation (4)) in the BBC fit, and $\mu_{i,?}$ are unbinned distances from using a different ? in the BBC fit. Results for eight wCDM model variants are shown in Table 1, and we find STD$_\mu \sim 10^{-4}$ mag, which is about 1000 times smaller than the intrinsic scatter.

The last three rows of Table 1 are based on a polynomial function of redshift for $\mu_{i,?}$ to illustrate the BBC performance with poorer $\mu_{i,?}$ estimates. A constant $\mu_{i,?}$ (p0) results in STD$_\mu$ that is more than 1 order of magnitude larger than for the wCDM models, but is still well below 0.01 mag and thus works surprisingly well for such a poor $\mu_{i,?}$ estimate. Using third- and sixth-order polynomial fits to the baseline ΛCDM model (p3 and p6) works much better than constant $\mu_{i,?}$, but still not quite as well as for the wCDM models.

5.2. Uncertainty Scale and HR Check

For the Ia+CC samples, the uncertainty scale (S_ζ in Equations (8) and (10)) versus redshift bin is shown in the upper panel of Figure 1, averaged over the 50 data samples. The average S_ζ value is ∼1.01 at all redshifts, and is thus a small correction. The Hubble residual crosscheck ($\langle \mathrm{HR}_\zeta \rangle$ defined in Equation (11)) versus redshift is shown in the lower panel of Figure 1. The values are within 0.001 mag of the expected value of zero.

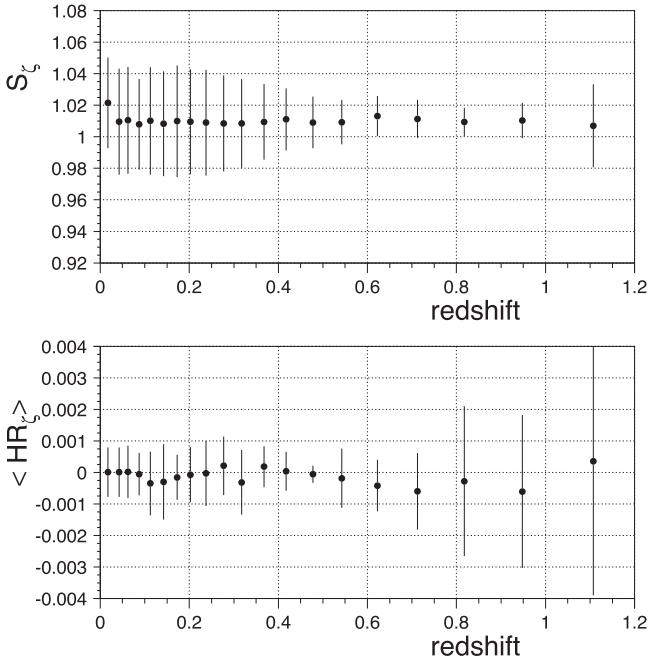

Figure 1. S_ζ versus redshift (upper panel) and $\langle \mathrm{HR}_\zeta \rangle$ versus redshift (lower panel). The error bars show the rms among the 50 simulated data samples.

5.3. Bias Results for wCDM

We define the w-bias to be $\Delta w \equiv w - w_{\rm true}$, and we define $\langle w\text{-bias} \rangle$ to be the average over the 50 simulated data samples. The $\langle w\text{-bias} \rangle$ uncertainty is the standard deviation of the Δw values divided by $\sqrt{50}$. We begin with an SN Ia–only sample that has no contamination and $P_{\rm Ia} = 1$ for all events. The $\langle w\text{-bias} \rangle$ results are shown in the upper half of Table 2 with no systematics (i.e., only statistical uncertainties), with systematics included, and with four binning options: (i) binned, (ii) unbinned, (iii) rebinned with two stretch bins and four color bins (Rebin2x4), and (iv) rebinned with four stretch bins and eight color bins (Rebin4x8). A bias consistent with zero at the 2σ level (N_σ in Table 2) is considered to be unbiased. All $\langle w\text{-bias} \rangle$ results are unbiased except for a 2.4σ bias for Ia-Only using an unbinned HD and systematics. By averaging results over 50 samples, the bias uncertainty and constraint is nearly 1 order of magnitude smaller than the average w-uncertainty ($\langle \sigma_w \rangle$ in Table 2) for a single data sample.

The primary motivation for an unbinned HD is to reduce the total uncertainty. With systematics, the average total uncertainty ($\langle \sigma_w \rangle$ column in Table 2) is reduced by ∼7% compared to the binned result. Naively subtracting the no-syst contribution in quadrature, the systematic uncertainties are 0.0182 and 0.0143 for the binned and unbinned, respectively, resulting in an ∼20% reduction in the systematic uncertainty. The rebinned uncertainties are comparable to that of the unbinned result.

Our 20% reduction in systematic uncertainty is smaller than the 50% reduction in BHS21 because we did not include the intrinsic scatter systematic that is reduced by more than a factor

Table 2
Average w-bias, Significance, and Uncertainty vs. Redshift Binning Option (wCDM)

SN Types	Syst	Bin Option	$\langle w\text{-bias}\rangle$	N_σ[a]	$\langle\sigma_w\rangle$
Ia-Only	None	Binned	$+0.0002 \pm 0.0030$	0.1	0.0249
		Unbin	$+0.0004 \pm 0.0030$	0.1	0.0250
		Rebin2x4[b]	$+0.0004 \pm 0.0030$	0.1	0.0249
		Rebin4x8[c]	$+0.0005 \pm 0.0030$	0.2	0.0250
Ia-Only	All	Binned	$+0.0008 \pm 0.0031$	0.3	0.0308
		Unbin	-0.0069 ± 0.0029	2.4	0.0288
		Rebin2x4	$+0.0009 \pm 0.0031$	0.3	0.0293
		Rebin4x8	$+0.0010 \pm 0.0030$	0.3	0.0286
Ia+CC	None	Binned	$+0.0020 \pm 0.0030$	0.7	0.0250
		Unbin	$+0.0026 \pm 0.0030$	0.9	0.0250
		Rebin2x4	$+0.0022 \pm 0.0030$	0.7	0.0250
		Rebin4x8	$+0.0022 \pm 0.0030$	0.7	0.0250
Ia+CC	All	Binned	$+0.0024 \pm 0.0032$	0.7	0.0309
		Unbin	-0.0044 ± 0.0030	1.5	0.0288
		Rebin2x4	$+0.0025 \pm 0.0031$	0.8	0.0294
		Rebin4x8	$+0.0019 \pm 0.0032$	0.6	0.0286

Notes.
[a] Absolute value of $\langle w - \text{bias}\rangle$ divided by its uncertainty.
[b] Rebinning with two stretch bins and four color bins.
[c] Rebinning with four stretch bins and eight color bins.

of 3. For the calibration systematics used in both analyses, the reduction in BHS21 (see their Table 2) is similar to ours.

5.4. Bias Results for w_0w_aCDM

For the w_0w_aCDM model, the bias summary is shown in Table 3. The biases are consistent with zero at the 2σ level, and the bias uncertainty and constraint are nearly 1 order of magnitude smaller than the average single-sample uncertainty ($\langle\sigma_{w0}\rangle$ and $\langle\sigma_{wa}\rangle$ in Table 3). For the Ia+CC sample, the average FoM is $\langle\text{FoM}\rangle = 45$ for the binned option, and the unbinned option increases $\langle\text{FoM}\rangle$ to 55. The Rebin2x4 option results in $\langle\text{FoM}\rangle = 51$ that is between the binned and unbinned $\langle\text{FoM}\rangle$. The Rebin4x8 option results in $\langle\text{FoM}\rangle = 54$ that is very close to the unbinned $\langle\text{FoM}\rangle$.

5.5. Binning Option Consistency

Without systematics, wCDM w-results for all binning options, with and without contamination, agree to within <0.001. For the w_0w_aCDM model and Ia-Only, all binning options agree to within 0.001 and <0.01 for w_0 and w_a, respectively. With contamination, the rebin results differ by 0.003 and 0.01 for w_0 and w_a, respectively, suggesting a subtle bias with the rebin procedure.

With systematics, the binned and rebinned wCDM results agree to within 0.001 in w, while the unbinned w-result differs significantly by 0.006 ± 0.001, which corresponds to 20% of the total uncertainty. This comparison is the same with and without contamination. For the w_0w_aCDM model, the binned and rebinned results agree to within 0.002 in w_0; the w_a results agree to within <0.01 for Ia-Only and differ by up to nearly 0.03 with contamination. The unbinned results differ by $\sim$0.01 and 0.06 for w_0 and w_a, respectively. While all binning options show unbiased cosmology results, there is evidence for a small difference between the binned the unbinned results. This difference is present with our without contamination.

5.6. Impact of $\mathcal{P}_{B(Ia)}$ Term

To check the impact of the $\mathcal{P}_{B(Ia)}$ term in Equation (10), we forced $\mathcal{P}_{B(Ia)} = 1$ and reevaluated the w_0w_aCDM bias for an unbinned HD. We find more than 10σ bias, which illustrates the necessity of accurately evaluating this term.

6. Conclusion

As a follow-up to the original binned Hubble diagram (HD) from BBC, we have developed methods to derive cosmological results from an unbinned HD and from a rebinned HD in the space of redshift, stretch, and color. Averaging analysis results from 50 simulated data samples, we find biases consistent with

Table 3
Average $w0$, w_a-bias, Significance, Uncertainty, and FoM vs. Redshift Binning Option (w_0w_aCDM)

SN Types	Syst	Bin Option	$\langle w_0\text{-bias}\rangle$	N_σ[a]	$\langle\sigma_{w0}\rangle$	$\langle w_a\text{-bias}\rangle$	N_σ[b]	$\langle\sigma_{wa}\rangle$	$\langle\text{FoM}\rangle$
Ia-Only	None	Binned	-0.005 ± 0.014	0.3	0.101	-0.022 ± 0.063	0.3	0.482	78
		Unbin	-0.004 ± 0.014	0.3	0.100	-0.023 ± 0.063	0.4	0.480	78
		Rebin2x4[c]	-0.005 ± 0.014	0.4	0.101	-0.017 ± 0.064	0.3	0.482	78
		Rebin4x8[d]	-0.006 ± 0.014	0.4	0.101	-0.015 ± 0.064	0.2	0.481	78
Ia-Only	All	Binned	-0.003 ± 0.015	0.2	0.138	-0.049 ± 0.069	0.7	0.634	45
		Unbin	$+0.008 \pm 0.014$	0.5	0.122	-0.122 ± 0.066	1.9	0.575	56
		Rebin2x4	-0.002 ± 0.015	0.1	0.130	-0.047 ± 0.069	0.7	0.595	51
		Rebin4x8	-0.001 ± 0.014	0.1	0.125	-0.044 ± 0.065	0.7	0.576	55
Ia+CC	None	Binned	-0.010 ± 0.014	0.7	0.102	$+0.013 \pm 0.065$	0.2	0.481	77
		Unbin	-0.010 ± 0.014	0.7	0.101	$+0.013 \pm 0.065$	0.2	0.479	78
		Rebin2x4	-0.013 ± 0.014	0.9	0.102	$+0.025 \pm 0.066$	0.4	0.481	77
		Rebin4x8	-0.013 ± 0.014	0.9	0.102	$+0.025 \pm 0.065$	0.4	0.480	77
Ia+CC	All	Binned	-0.004 ± 0.015	0.3	0.139	-0.039 ± 0.069	0.6	0.631	45
		Unbin	$+0.005 \pm 0.015$	0.3	0.123	-0.100 ± 0.067	1.5	0.574	55
		Rebin2x4	-0.003 ± 0.016	0.2	0.131	-0.037 ± 0.072	0.5	0.595	51
		Rebin4x8	-0.008 ± 0.015	0.5	0.126	-0.012 ± 0.069	0.2	0.575	54

Notes.
[a] Absolute value of $\langle w_0 - \text{bias}\rangle$ divided by its uncertainty.
[b] Absolute value of $\langle w_a - \text{bias}\rangle$ divided by its uncertainty.
[c] Rebinning with two stretch bins and four color bins.
[d] Rebinning with four stretch bins and eight color bins.

zero and bias constraints almost 1 order of magnitude smaller than the single-sample uncertainty. We also find that using an unbinned HD results in a reduced total uncertainty consistent with BHS21. This conclusion holds for both the wCDM and w_0w_aCDM models, and we find the same results with or without photometric contamination.

Using a rebinned HD with two stretch and four color bins (Rebin2x4), we recover unbiased cosmology results and also benefit from the reduced uncertainty in the unbinned HD. Using more bins (four stretch and eight color bins), there is still no bias and the total uncertainty is similar to the unbinned case. With ∼2000 events in the DES unbinned HD, the rebinned cosmology-fitting speed is only a factor of few faster compared to the unbinned case. With anticipated future samples in the 10^4–10^5 range, the rebinned HD size does not increase and therefore the cosmology-fitting speed improvement will be much more significant.

While our unbiased results are encouraging, we note that Mitra et al. (2023) reported a significant cosmology bias using an unbinned HD from a simulated LSST data sample of pure SNe Ia. We therefore recommend repeating our bias tests on simulated data for future analyses.

Acknowledgments

R.K. is supported by DOE grant DE-SC0009924. P.A acknowledges parts of this research were carried out on the traditional lands of the Ngunnawal people. We pay our respects to their elders past, present, and emerging. P.A. was supported by an Australian Government Research Training Program (RTP) Scholarship. We acknowledge the University of Chicago's Research Computing Center for their support of this work.

ORCID iDs

R. Kessler https://orcid.org/0000-0003-3221-0419
P. Armstrong https://orcid.org/0000-0003-1997-3649

References

Albrecht, A., Bernstein, G., Cahn, R., et al. 2006, arXiv:astro-ph/0609591
Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22
Brout, D., Hinton, S. R., & Scolnic, D. 2021, ApJL, 912, L26
Brout, D., Scolnic, D., Popovic, B., et al. 2022a, ApJ, 938, 110
Brout, D., Taylor, G., Scolnic, D., et al. 2022b, ApJ, 938, 111
Conley, A., Guy, J., Sullivan, M., et al. 2011, ApJS, 192, 1
DES Collaboration 2019, ApJL, 872, L30
Faccioli, L., Kim, A. G., Miquel, R., et al. 2011, APh, 34, 847
Foley, R. J., Scolnic, D., Rest, A., et al. 2018, MNRAS, 475, 193
Guy, J., Sullivan, M., Conley, A., et al. 2010, A&A, 523, A7
Hicken, M., Challis, P., Jha, S., et al. 2009, ApJ, 700, 331
Hicken, M., Challis, P., Kirshner, R. P., et al. 2012, ApJS, 200, 12
Hinton, S., & Brout, D. 2020, JOSS, 5, 2122
Hlozek, R., Kunz, M., Bassett, B., et al. 2012, ApJ, 752, 79
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. 2018, ApJ, 857, 51
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., Narayan, G., Avelino, A., et al. 2019b, PASP, 131, 094501
Kessler, R., & Scolnic, D. 2017, ApJ, 836, 56
Kunz, M., Bassett, B. A., & Hlozek, R. A. 2007, PhRvD, 75, 103508
Lochner, M., McEwen, J. D., Peiris, H. V., Lahav, O., & Winter, M. K. 2016, ApJS, 225, 31
Marriner, J., Bernstein, J. P., Kessler, R., et al. 2011, ApJ, 740, 72
Mitra, A., Kessler, R., More, S., Hlozek, R. & The LSST Dark Energy Science Collaboration 2023, ApJ, 944, 212
Möller, A., & de Boissière, T. 2020, MNRAS, 491, 4277
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
Qu, H., Sako, M., Möller, A., & Doux, C. 2021, AJ, 162, 67
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009
Sako, M., Bassett, B., Connolly, B., et al. 2011, ApJ, 738, 162
Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, ApJ, 859, 101
Tripp, R. 1998, A&A, 331, 815
Vincenzi, M., Sullivan, M., Firth, R. E., et al. 2019, MNRAS, 489, 5802
Vincenzi, M., Sullivan, M., Graur, O., et al. 2021, MNRAS, 505, 2819
Vincenzi, M., Sullivan, M., Möller, A., et al. 2023, MNRAS, 518, 1106