The Astrophysical Journal Letters, 896:L4 (7pp), 2020 June 10 https://doi.org/10.3847/2041-8213/ab94ad

Evidence for Cosmic Acceleration Is Robust to Observed Correlations between Type Ia Supernova Luminosity and Stellar Age

B. M. Rose[1], D. Rubin[2,3], A. Cikota[3], S. E. Deustua[1], S. Dixon[3,4], A. Fruchter[1], D. O. Jones[5], A. G. Riess[1,6], and D. M. Scolnic[7]

[1] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA; brose@stsci.edu
[2] Physics and Astronomy Department, University of Hawaii, Honolulu, HI 96822, USA
[3] Physics Division, E.O. Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA 94720, USA
[4] Department of Physics, University of California Berkeley, 366 LeConte Hall MC 7300, Berkeley, CA 94720, USA
[5] Department of Astronomy and Astrophysics, University of California, Santa Cruz, CA 92064, USA
[6] Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
[7] Department of Physics, Duke University, 120 Science Drive, Durham, NC 27708, USA
Received 2020 February 27; revised 2020 May 12; accepted 2020 May 15; published 2020 June 8

Abstract

Type Ia supernovae (SNe Ia) are powerful standardizable candles for constraining cosmological models and provided the first evidence of the accelerated expansion of the universe. Their precision derives from empirical correlations, now measured from >1000 SNe Ia, between their luminosities, light-curve shapes, colors, and most recently with the stellar mass of their host galaxy. As mass correlates with other galaxy properties, alternative parameters have been investigated to improve SN Ia standardization though none have been shown to significantly alter the determination of cosmological parameters. We re-examine a recent claim, based on 34 SN Ia in nearby passive host galaxies, of a 0.05 mag Gyr^{-1} dependence of standardized SN Ia luminosity on host age, which, if extrapolated to higher redshifts, would be a bias up to 0.25 mag, challenging the inference of dark energy. We reanalyze this sample of hosts using both the original method and a Bayesian hierarchical model and find after a fuller accounting of the uncertainties the significance of a dependence on age to be $\leqslant 2\sigma$ and $\sim 1\sigma$ after the removal of a single poorly sampled SN Ia. To test the claim that a trend seen in old stellar populations can be applied to younger ages, we extend our analysis to a larger sample that includes young hosts. We find the residual dependence of host age (after all standardization typically employed for cosmological measurements) to be consistent with zero for 254 SNe Ia from the Pantheon sample, ruling out the large but low significance trend seen in passive hosts.

Unified Astronomy Thesaurus concepts: Cosmological constant (334); Dark energy (351); Type Ia supernovae (1728); Observational cosmology (1146); Distance indicators (394)

Supporting material: machine-readable table

1. Introduction

Type Ia supernovae (SNe Ia), the thermonuclear explosion of carbon–oxygen white dwarfs, are precise cosmic distance indicators (Kowal 1968; Pskovskii 1969). Their observed variation in brightness can be empirically corrected (Rust 1974; Pskovskii 1977; Phillips 1993; Hamuy et al. 1996; Riess et al. 1996; Perlmutter et al. 1997). This allows their luminosity distances to be used to measure the expansion history of the universe and led to the discovery of cosmic acceleration caused by an unknown force, dark energy (Riess et al. 1998; Garnavich et al. 1998; Perlmutter et al. 1999). Since then, standardization methods have improved (Jha et al. 2007; Guy et al. 2010; Burns et al. 2011; Mosher et al. 2014) as have the resulting cosmological measurements (Suzuki et al. 2012; Betoule et al. 2014; Riess et al. 2018; Scolnic et al. 2018; DES Collaboration et al. 2019; Jones et al. 2019; Freedman et al. 2019).

As the number of SNe Ia at cosmological distances now exceed 1000, the selection criteria have become more stringent. Of all the SNe Ia observed, roughly 75% (Scolnic et al. 2018) are used for cosmology. Cosmologically useful SNe Ia are required to have sufficient data: adequate sampling in order to constrain the light curve and the decline rate. They must also pass "quality cuts;" i.e., their parameters should be nearest the centers of the population distributions and thus can be standardized through empirical correlations. Even after standardization, outliers exist, resulting in outlier rejection tools (Kunz et al. 2007; Rubin et al. 2015) and even the classification of a new class of transients (Foley et al. 2013).

The accuracy of SN Ia cosmological measurements require the absence of a redshift dependence of the standardized luminosity, which we refer to as luminosity evolution. The variation in peak luminosity of SNe Ia may be due to unknown properties of the progenitors. These could have three effects that concern cosmological measurements. First, these variations in progenitor properties can affect the population demographics. This results in a type of bias discussed in Scolnic & Kessler (2016). In addition, many progenitor properties that affect the peak luminosity are already corrected for by the empirical standardization process. Ultimately, luminosity evolution comes from a change in the progenitor system and peak luminosity that is not accounted for in our SN Ia models.

As a proxy for a change in redshift or cosmic time, luminosity evolution can be constrained locally ($\lesssim$400 Mpc) by measuring differences in standardized SN Ia luminosity between galaxy types. Over the last decade large samples with strict quality control have revealed correlations between host galaxy properties and standardized peak luminosity at a modest level (e.g., Gallagher et al. 2008; Kelly et al. 2010; Sullivan et al. 2010; Lampeitl et al. 2010; Gupta et al. 2011; Rigault et al. 2013; Jones et al. 2015; Moreno-Raya et al. 2016; Uddin et al. 2017; Kim et al. 2018; Rigault et al. 2018; Jones et al. 2018; Rose et al. 2019). Each of these measurements agree in the direction of the host galaxy effect; it is clear that these do not agree by chance. Since the average galaxy changes with redshift and sample selection, it has become necessary to include such correlations in the standardization process to limit biases to the 1% level in distance (Rigault et al. 2013). The first recognized and most commonly used host property for such standardization is stellar mass (Kelly et al. 2010; Sullivan et al. 2010; Lampeitl et al. 2010). This standardization is referred to as the "mass step" because of the $\sim$0.06 mag change in average Hubble–Lemaître residual at $\sim 10^{10}\,M_{\odot}$. Hubble–Lemaître residuals are the difference between the measured luminosity distance and the expected distance from the best-fit cosmology.

Kang et al. (2020; hereafter K20) claim to have found a correlation between the ages of 34 early-type host galaxies—derived from spectral features—and SN Ia peak luminosity. If extrapolated to younger ages and higher redshifts, by convolving look-back time and SN Ia progenitor models, this correlation could cause a redshift-dependent luminosity evolution, $\Delta\mathrm{mag}/\Delta z > 0.2$ mag. The original discovery of accelerating cosmic expansion using SNe Ia (Riess et al. 1998; Perlmutter et al. 1999) ruled out such a large evolution in standardized luminosity by demonstrating consistency between SN Ia in early-type hosts and those in young, star-forming hosts.

K20's use of high signal-to-noise spectra to measure metallicities and ages of the host galaxies is impressive, however we have serious concerns about the cosmological interpretation. K20's finding of a correlation does not seem to be robust against different sample selections, or different assumptions about uncertainties. In addition, the application of the mass step correction drastically reduces the observed effect in external data. The motivation of the K20 work is well justified—correlations between SN Ia properties and their hosts exist, and these will need to be better characterized to significantly improve upon present cosmological measurements. However, in this work, we show that these correlations are not significantly limiting our current ability to use SNe Ia to measure the cosmological parameters of our universe.

2. Data and Techniques

Altogether Kang et al. (2016) and K20 observed 51 early-type, low-redshift SN Ia host galaxies, obtaining high signal-to-noise ratio (S/N) galactic spectra (S/N $\sim$ 175, Kang et al. 2016). The high-quality spectra allow for precise measurements of the SN Ia host galaxy properties. Most SN Ia host galaxy studies use photometry (e.g., Gupta et al. 2011; Jones et al. 2018; Rose et al. 2019), though some studies use lower signal-to-noise integral field unit spectra (Rigault et al. 2013, 2018). The SN Ia analyzed by K20 are archival, taking place from 1990 to 2010, and reanalyzed uniformly in the YONSEI SN catalog (Kim et al. 2019). The age measurement techniques used by K20 are well established (Faber et al. 1992; Worthey et al. 1994), and built on previous SN Ia research, such as Gallagher et al. (2008).

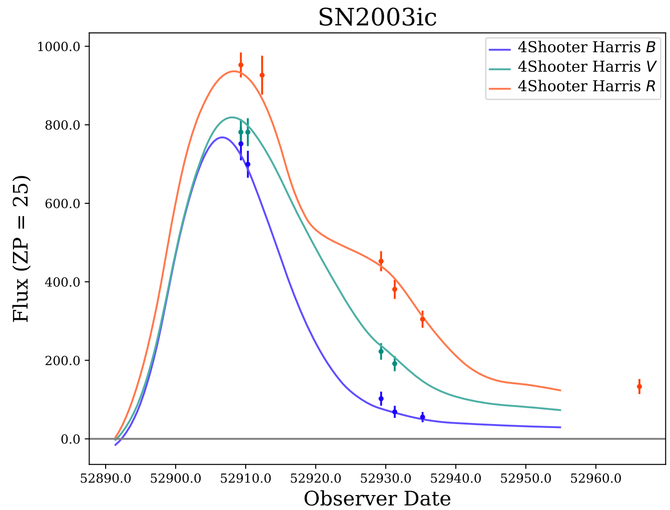

Figure 1. Light curve of SN2003ic. With no pre-maximum data points, it would commonly be removed from cosmological samples due to difficulties constraining the peak luminosity. However, this SN Ia alone takes the correlation seen in K20 from a 1.8σ to a 2.3σ significance. Original photometry from Hicken et al. (2009).

2.1. SN Ia Data Quality

K20 observed 51 SN Ia host galaxies. Via various cuts described in their paper, the fiducial analysis was performed with 34 SN Ia and their host galaxies. Using several definitions of SN Ia quality, we find 10 of the final 34 SNe Ia fail at least one quality cut. Using just the Joint Light-Curve Analysis (JLA) cosmology cuts (Betoule et al. 2014, Table 7), nearly 12% (four out of 34 objects) of the final sample are not of cosmological quality.

The precision and accuracy of SN Ia distances depends on the quality of the light-curves of the SN Ia. There are several SN Ia in the K20 sample with poorly sampled light curves, and light-curve fits for these SNe Ia will be problematic. The light curves for SN2007ap and SN2008af have no data prior to five days past maximum. SN2003iv and SN2007cp have fewer than four nights of observations, and SN2003ch and SN2003ic have fewer than seven. As an example, SN2003ic is shown in Figure 1. Finally both SN1993ac and SN2001ie have no data prior to five days post-maximum and fewer than seven nights of observations. Of these eight SN Ia with poorly sampled light curves, only SN2003ch and SN2007cp were removed from the final data set of 34 SN Ia used in K20.

To test assumptions used in cosmological analyses, it is not necessary to use a fully representative sub-sample, but they should at least all pass the typical quality cuts. Using JLA as an example, we see that four of the final 34 SN Ia do not pass quality cuts. SN2002do and SN2007au are best fit with the light curves shape parameter $x_1 < -3$. These fast decliners are outside the valid range of the SALT2 model (Guy et al. 2007, 2010) requiring alternative standardization methods (e.g., Garnavich et al. 2004). In addition, SN2002do, SN2006kf, and SN2008ia all have high Milky Way dust extinction, with $E(B-V)_{\mathrm{MW}} \geqslant 0.15$. The more dimming and reddening from Milky Way dust, the less accurate the SN Ia peak luminosity can be. For this reason, cosmological analyses typically use SNe Ia that are out of the plane of the Milky Way. Pantheon and Union perform very similar quality cuts (Scolnic et al. 2018; Suzuki et al. 2012, respectively). Other analyses include additional cuts on the phase coverage of the light curves,

Table 1
K20 SN Ia that Do Not Pass Typical "Quality Cuts"

	Light Curves Quality		JLA Cuts		Rest et al. (2014)	
	⩾1 Obs. $t < +5$ days	Total Num. Obs. >7	$\|x_1\| < -3$	$E(B-V)_{MW} \geqslant 0.15$	⩾1 Obs. $-10 < t < +5$ days	⩾1 Obs. $+5 < t < +25$ days
SN1993ac[a b]	X	X			X	
SN2001ie[a b]	X	X			X	
SN2002do[c]			X	X		
SN2003ic[a b]		X				X
SN2003iv[a]		X				
SN2006kf[c]				X		
SN2007ap[a b]	X				X	
SN2007au[b]			X			
SN2008af[a b]	X				X	
SN2008ia[c]				X		

Notes. From the final 34 SN Ia in the K20 sample, six SN Ia have poorly sampled light curves, four would not pass the JLA cuts, and five would not pass the Rest et al. (2014) phase coverage cuts.
[a] Has a poorly sampled light curve.
[b] Fails phase coverage cuts, defined in Rest et al. (2014).
[c] Fails JLA quality cuts, defined in Betoule et al. (2014).

expressed in terms of rest-frame days from maximum brightness. For example, Rest et al. (2014) required at least one observation between −10 and +5 days, at least one observation between +5 and +20 days, and at least five total observations between −10 and +35 days. There are 4 SNe Ia (SN1993ac, SN2001ie, SN2007ap, SN2008af) in the final sample of K20 that fail the first cut, and another SN Ia (SN2003ic) fails the second.

A summary of which SN Ia fails what cut can be seen in Table 1.

2.2. Standardization and Uncertainties

For SN Ia at low redshift (the K20 sample is at $z < 0.04$) there are several important uncertainties to consider: the uncertainty in the local peculiar motion(σ_v) and the unexplained scatter seen in SN Ia post standardization ($\sigma_{\rm unexplained}$).[8] If one accounts for expected flows using maps of large-scale structure on a SN Ia-by-SN Ia basis as undertaken by K20, a peculiar velocity uncertainty floor remains due to the unpredictable motions local to each host galaxy. Pantheon (Scolnic et al. 2018) calculated this to be $\sigma_v = 250$ km s^{-1}. The total distance uncertainty of a SN Ia comprises many individual uncertainties. A relevant example, based on the Pantheon analysis of Scolnic et al. (2018), is

$$\sigma_{\rm total}^2 = \sigma_N^2 + \sigma_z^2 + \boldsymbol{\sigma}_{\bf v}^{\bf 2} + \boldsymbol{\sigma}_{\bf unexplained}^{\bf 2} \quad (1)$$

where σ_N^2 is the photometric error of the SN Ia distance and σ_z^2 is the uncertainty from the redshift. In K20, $\sigma_{\rm unexplained}$ was misunderstood and σ_v was absent.

When looking for a trend between Hubble–Lemaître residuals and a host galaxy property one can accidentally ignore cross correlations with the SN Ia standardization terms (Hamuy et al. 1995, 2000; Smith et al. 2020). Therefore, to further test the observed trend in K20, we sampled a simple standardization equation in the Bayesian hierarchical model UNITY[9] (Rubin et al. 2015; Rose et al. 2020). We used a typical Tripp-like linear standardization (Tripp 1998):

$$\mu = m_B - (M_B + \alpha x_1 + \beta c + \gamma a) \quad (2)$$

where μ, m_B, M_B are the distance modulus, apparent and absolute magnitude, respectively. The α, β, and γ parameters are the linear standardization coefficients corresponding to the SALT2 (Guy et al. 2007, 2010) light-curve shape (x_1) and color (c), along with the host galaxy age in gigayears (a). The parameters m_B, x_1, c, and a are unique for each SN Ia, whereas M_B, α, β, and γ are fit for simultaneously along with any cosmological parameters of interest. UNITY also simultaneously fits for the remaining unexplained scatter ($\sigma_{\rm unexplained}$) allowing for the additional term, γa, to explain more of the observed SN Ia variability but still tracking all uncertainties.

3. Re-examining the SN Ia–Age Correlation

3.1. The Impact of SN2003ic

The measured Hubble–Lemaître residual-age trend (K20, Figure 13) visually appears to be dominated by SN2003ic, the SN Ia with the oldest host. As seen in Figure 1 and addressed in Section 2.1, the light curve of SN2003ic is poorly sampled, including no pre-maximum measurements and only two epochs closely spaced in time to sample the the first 15 days of decline, the most valuable span of time for calibrating the light-curve decline rate. If SN2003ic were removed, the trend shifts from -0.051 ± 0.022 mag Gyr^{-1} (2.3σ) to a less significant -0.045 ± 0.024 mag Gyr^{-1} (1.8σ) using the original K20 data. Removing other poorly sampled SN Ia do not affect the trend as much as SN2003ic. A summary of each Hubble–Lemaître residual stellar age correlation discussed in this Letter can be found in Table 2.

3.2. Underestimating Uncertainties

K20 states that they fit their correlations using `LINMIX` (Kelly 2007). This methodology contains an "intrinsic random scatter," counter to the claim that K20 uses no intrinsic scatter. Our reproduction of their work was performed using the `linmix_err` package in IDL. We conclude that $\sigma_{\rm unexplained}$ was calculated by `LINMIX` and was ∼0.10 mag, as seen in

[8] K20 uses a common alternative name, intrinsic dispersion ($\sigma_{\rm int}$).
[9] https://github.com/rubind/host_unity

Table 2
Summary of Discussed Correlations between Hubble–Lemaître Residual and Stellar Age

Method	Correlation (mag Gyr^{-1})	Significance	Num. SN Ia
K20 fiducial analysis	−0.051 ± 0.022	2.3σ	34
K20 reproduction	−0.051 ± 0.023	2.3σ	34
K20 reproduction w/o SN2003ic	−0.045 ± 0.024	1.8σ	33
K20 plus 250 km s^{-1} velocity uncertainty	−0.047 ± 0.022	2.1σ	34
above plus 0.10 mag floor on $\sigma_{\rm unexplained}$	−0.046 ± 0.024	1.9σ	34
above w/o SN2003ic	−0.037 ± 0.025	1.5σ	33
UNITY	−0.035 ± 0.023	1.5σ	34
UNITY w/o SN2003ic	−0.013 ± 0.022	0.6σ	33
Spearman correlation coefficient	…	2.0σ	34
Pantheon Hubble–Lemaître residuals	−0.016 ± 0.031	0.5σ	27
Pantheon w/o SN2003ic	+0.008 ± 0.030	0.3σ	26

other works. Adding in the peculiar velocity uncertainty, σ_v, we calculate the significance of the age trend becomes 1.9σ–or 1.5σ when removing SN2003ic.

When re-analyzing the original data with UNITY, the significance of the K20 trend with age (γ) is reduced to 1.5σ. This suggests that the significance of any Hubble–Lemaître residual-host galaxy correlation is typically over estimated compared to when all parameters are simultaneously fit. As is necessary for an accurate error estimation, we included the non-diagonal covariance terms from the light-curve fitting; K20 only reported diagonal covariance terms. Due to this missing data and some inconsistencies between the values reported in K20 and the original YONSEI SN catalog (i.e. the x_1 value of SN2002G), we used the results from our own light-curve fits.

However, we are able to ignore the disputed uncertainties (σ_v and $\sigma_{\rm unexplained}$) and measure a correlation's significance directly from the scatter in the data. This is done via correlation coefficients. The Pearson correlation coefficient is the most common, but assumes both that the trend is linear and that each data set is normally distributed. Since the age values have been found to not be normally distributed (Childress et al. 2014; Rose et al. 2019), we use the Spearman rank-order correlation, which does not have this requirement. When using the final data set of K20, the Spearman correlation coefficient is $r_s = -0.35$, a 2.0σ non-zero result. This result is statistically consistent with the larger data set of Rose et al. (2019). Bypassing any question about the accuracy of the uncertainties, this trend appears only marginally significant.

Via several alternative analysis methods—both accounting for additional known uncertainties and bypassing them—we have seen the correlation is at most 2σ, but likely less. We conclude that there is no statistically significant trend with age in the K20 data.

3.3. Extrapolation to Constraints on Cosmological Parameters

Our next set of concerns are based around how K20 extrapolates a correlation with age to a bias in cosmology. As discussed previously, the correlations between Hubble–Lemaître residual and host galaxy age is dependent on a unique data set of K20 that is not typical of cosmological samples.

Recent SN Ia cosmology analyses (Suzuki et al. 2012; Betoule et al. 2014; Rubin et al. 2015; Scolnic et al. 2018; DES Collaboration et al. 2019), all of which have demonstrated strong evidence for cosmic acceleration, account for the well-established change in average Hubble–Lemaître residual across a division in host stellar mass. This procedure reduces the effect of any new correlation with age, due to galaxy scaling relationships. However, the K20 sample uniquely isolates age from stellar mass and morphology. Many cosmological analyses include a parameter to marginalize over the uncertainty that this change in Hubble–Lemaître residual could be caused by another host galaxy property, such as age. This marginalization would further reduce the effect of a trend with age. In Rubin et al. (2015), this marginalization was done with a the redshift-dependent mass step, and the resulting best-fit cosmology slightly favored an age-like redshift dependence over a pure stellar mass effect. This drastically reduces the maximal bias on cosmological parameters possible from the correlation reported by K20.

Table 3
Hubble–Lemaître Residuals and Ages for 254 Low-redshift Pantheon SN Ia

SN Ia	HR (mag)	Uncertainty (mag)	Age (Gyr)
2001ah	−0.04	0.13	1.714
2001az	0.14	0.12	1.041
2001bf	−0.10	0.17	1.261
2001 da	−0.04	0.14	1.261
2001eh	−0.01	0.16	1.924

Note. Hubble–Lemaître residuals (HRs) are from Jones et al. (2018). Ages (light-weighted) are estimated using ZPEG. We used a fixed 15% uncertainty in this analysis.

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

To further investigate if a standard cosmological analysis that accounts for both host and selection effects may mitigate the effect of K20's trend on cosmological parameters, we replaced the Hubble–Lemaître residual with those calculated during the Pantheon analysis (Scolnic et al. 2018). There are 27 SNe Ia with both Pantheon Hubble–Lemaître residuals and Yonsei Evolutionary Population Synthesis (YEPS) host galaxy ages. We present the Hubble–Lemaître residuals and ages for the entire low-redshift Pantheon sample in Table 3. Using the standard cosmological correction for the host mass step and a new observational bias correction framework (BBC; Kessler & Scolnic 2017), the trend with Hubble–Lemaître residual becomes -0.016 ± 0.031 mag Gyr^{-1} or consistent with zero. Without SN2003ic, the trend with the Pantheon's Hubble–Lemaître residuals reverses direction ($+0.008 \pm 0.030$ mag Gyr^{-1}). This is also true for the two

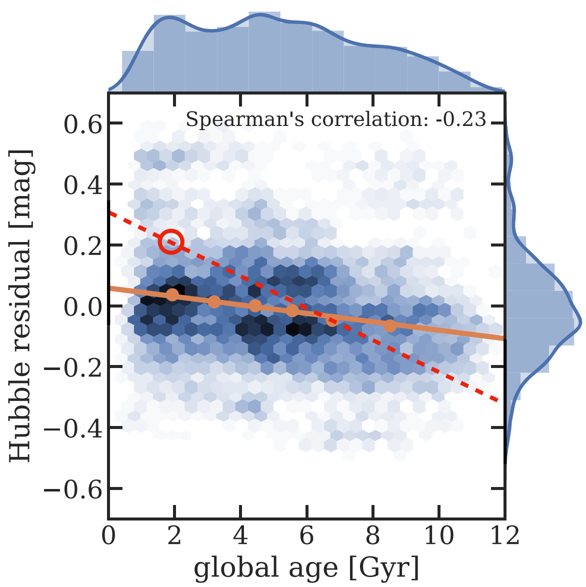

Figure 2. 2D density plot (darker colors indicate a higher density) depicting the probability of finding a SN Ia at a given Hubble–Lemaître residual and average stellar age. The global age is the mass-weighted average age from the galaxy spectral energy distribution (SED). Analysis details can be found in Rose et al. (2019). A linear fit to the data (orange line) is shown, along with six evenly filled bins (orange points). The extrapolated trend of K20 is shown as a red dashed line. The predicted average Hubble–Lemaître residual for young hosts is shown as the red circle. This prediction (red circle) is inconsistent with the measured average (orange point). The original data and figure are from Rose et al. (2019). We note that like K20, the Hubble–Lemaître residuals in Rose et al. (2019) do not include the mass step correction.

other age methods used in K20: going from K20 to Pantheon Hubble–Lemaître residuals the trend becomes consistent with zero. We conclude that using Hubble–Lemaître residuals that are standardized with the mass step results in an insignificant trend and therefore does not propagate to a bias in cosmological estimates.

3.4. Consistency with Other Data Sets

K20 ultimately applied their trend to cosmological distances by assuming it could be extrapolated to SNe Ia from younger stellar populations. This interpretation assumes that the physical mechanism is a smoothly varying process rather than discrete sub-populations as seen in Rigault et al. (2013) and Cikota et al. (2019). Indeed, it is quite possible that at all redshifts most SNe Ia are from young progenitors as SNe Ia in early-type hosts galaxies (typically dominated by old stars) make up only a small fraction of cosmological samples (Childress et al. 2014).

The interpretation in K20 implies that SN Ia in young hosts will have an average Hubble–Lemaître residual of $\sim$0.25 mag. This biased average Hubble–Lemaître residual is ruled out by the analyses of both Gupta et al. (2011) and Rose et al. (2019), who independently looked at data from the Sloan Digital Sky Survey (Sako et al. 2008; Campbell et al. 2013; Sako et al. 2018) using two distinct age estimators. An example of this discrepancy between external data and K20's predication can be seen in Figure 2. Measurements of the Hubble–Lemaître residuals for SNe Ia from young host galaxies place the prediction of K20 in the tail of the distribution.

The mass-weighted ages derived from the optical spectral energy distribution (SED) fitting of Rose et al. (2019) are not as precise for any one individual host galaxy as the K20's YEPS (Chung et al. 2013) ages derived from spectral features. However, when aggregated, SED based ages are statistically powerful until they reach the systematic limits of the stellar population models. Just like photometric redshifts, SED based ages can catastrophically fail for any one object, but in population studies they are a powerful tool.

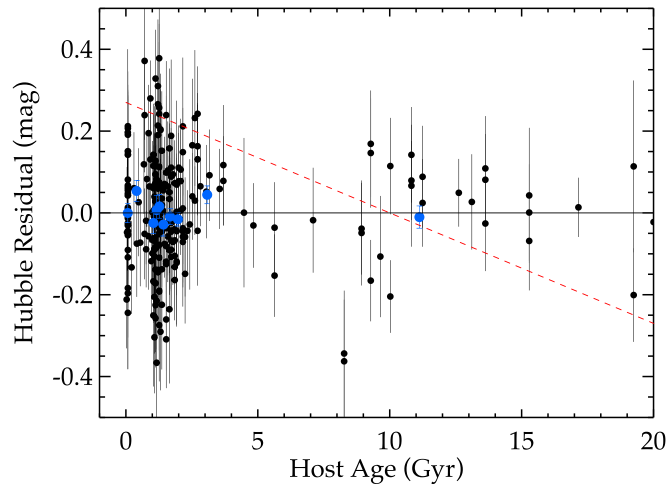

Figure 3. Relationship between the cosmological Pantheon sample's Hubble–Lemaître residuals and host galaxy age (light-weighted) for the low-redshift SNe Ia (black) of Jones et al. (2018). Blue points are bins of 25 SN Ia. Light-weighted ages are biased by bright young stars, reducing the range of observed ages and increasing the measured slope. The red dashed line is the trend seen in K20. The mass step that is applied to the Pantheon Hubble–Lemaître residuals would not drastically shift the trend from K20 because it was derived using only massive early-type galaxies. However, one can easily imagine concluding a low significance trend if one only used hosts with ages $>$2.5 Gyr (the last two bins), as expected in a passive only sample. There appears to be no systemic nor statistical bias in ages values where the K20 regression line would go through this external cosmological data set.

For a more direct and empirical test of the average Hubble–Lemaître residual of SNe Ia with young progenitors we analyzed the full sample of low-redshift SNe Ia in the Pantheon sample with host galaxy properties derived by Jones et al. (2018; $N = 254$). We measured the correlation between Hubble–Lemaître residual and host age as in K20. The Pantheon sample used light-weighted ages derived from SED fitting via ZPEG (Le Borgne & Rocca-Volmerange 2002), as described in Jones et al. (2018). ZPEG uses 15 star formation histories, the Salpeter initial mass function (Salpeter 1955), 200 stellar age bins, six metallicity bins, and marginalizes over $E(B-V)$ in order to fit the observed photometry. Figure 3 shows the expected result that the majority of low-redshift SNe Ia are seen in young hosts. The Hubble–Lemaître residuals seen in these hosts are strongly inconsistent with the $+0.25$ mag average residual predicted by extrapolating the trend proposed by K20. Indeed, only a small number of all SNe Ia (at any age) show residuals of $\gtrsim$0.25 mag, contrary to the prediction that this is the average Hubble–Lemaître residual for SNe Ia in young hosts. No bias in age or uncertainty (Gaussian, log-normal, or otherwise) would make the predicted trend match the data.

The age trend seen in the Pantheon sample is consistent with no trend. Light-weighted ages are biased young by bright stars, reducing the range of observed ages and increasing the measured slope. It is difficult to quantify this bias into an uncertainty on age. As such, these should only be treated as very crude estimates. Not surprisingly, by excluding the mass step correction the size of a trend with age more than doubles due to the aforementioned correlation between host mass and age, though this trend is still only significant at the 1.4σ level, similar to what was seen by Rose et al. (2019) for the SDSS data. When using the same light-curve standardization parameters ($\alpha = 0.15$, $\beta = 3.69$) as K20, but including the mass step and BBC corrections, the correlation only has a 0.6σ significance. If we restrict ourselves to early-type galaxies, as is the sample in K20, a very weak trend is found (1.2σ). No method of examining the Pantheon data set was able to find a significant uncorrected trend with age.

We conclude that the linear extrapolation to young ages is inconsistent with external data. Seeing no significant trend in a cosmological data set, we find no evidence for a significant unaccounted for bias in the cosmic acceleration signal from SN Ia.

4. Conclusions

Kang et al. (2020) claim that an empirically-determined dependence of SN Ia host age and luminosity derived from a small sample of early-type host galaxies can be extrapolated to large samples and young ages to account for the majority of the cosmic acceleration signal. However, we find that this trend is not robust to reanalysis. The first issue is that 12% of the final sample, would not pass the JLA cosmological quality cuts, meaning that a large fraction of the data does not have reliable Hubble–Lemaître residuals.

The inclusion of standard error sources, clearly present in SN Ia residuals, reduces the significance of the dependence to $<2\sigma$. Bypassing any need for formal uncertainty accounting, the Spearman rank-order coefficient only sees a correlation at 2σ. Further, the removal of the single SN Ia with the oldest host and a poorly sampled light curve, SN2003ic, reduces the significance to 1.5σ. Finally, by doing a full re-fit and Bayesian hierarchical analysis that also marginalizes over the correlations in the standardization coefficients, we find the trend falls to a 1.5σ or 0.6σ significance with and without SN2003ic, respectively.

If this correlation exists, the propagation to a bias in cosmological parameters is not direct or simple. When replacing the Hubble–Lemaître residuals from K20 with those used in the Pantheon analysis, we see that the standard practice of applying a host galaxy mass correction leaves only a very weak and insignificant relation between Hubble–Lemaître residuals and inferred age.

Finally, comparing the claimed trend against large, recent cosmological samples, which include young hosts, the trend is strongly ruled out.

The recent results of K20, upon re-examination, do not justify calling into question the presence of dark energy. However, we do concur with their closing remarks: the redshift dependence of SN Ia remains an important challenge for future precision dark energy measurements and requires ongoing studies.

The authors thank Peter Garnavich, Young-Wook Lee, and Young-Lo Kim for their time reading and commenting on an early draft. In addition, we thank the two referees whose comments improved the clarity, and hopefully impact, of this Letter. B.M.R., D.R., A.C., S.E.D., S.D., and A.F. acknowledge support from NASA through grant NNG16PJ311I. D.O.J. is supported by a Gordon and Betty Moore Foundation postdoctoral fellowship at the University of California, Santa Cruz. D.M.S. is supported by DOE grant de-sc0010007, the David and Lucile Packard Foundation, and in part by NASA through grant NNG17PX03C.

Software: corner.py (Foreman-Mackey 2016), kde_corner (https://github.com/rubind/kde_corner), Matplotlib (Hunter 2007), Numpy (van der Walt et al. 2011), Pandas (McKinney 2010), pystan (Riddell et al. 2018), SciPy (Virtanen et al. 2020), stan (Carpenter et al. 2017), UNITY (Rubin et al. 2015), ZPEG (Le Borgne & Rocca-Volmerange 2002).

ORCID iDs

B. M. Rose https://orcid.org/0000-0002-1873-8973
D. Rubin https://orcid.org/0000-0001-5402-4647
A. Cikota https://orcid.org/0000-0001-7101-9831
S. E. Deustua https://orcid.org/0000-0003-2823-360X
S. Dixon https://orcid.org/0000-0003-1861-0870
A. Fruchter https://orcid.org/0000-0002-6652-9279
D. O. Jones https://orcid.org/0000-0002-6230-0151

References

Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22
Burns, C. R., Stritzinger, M., Phillips, M. M., et al. 2011, ApJ, 141, 19
Campbell, H., D'Andrea, C. B., Nichol, R. C., et al. 2013, ApJ, 763, 88
Carpenter, B., Gelman, A., Hoffman, M. D., et al. 2017, J. Stat. Software, 76, 1
Childress, M. J., Wolf, C., & Zahid, H. J. 2014, MNRAS, 445, 1898
Chung, C., Yoon, S.-J., Lee, S.-Y., & Lee, Y.-W. 2013, ApJS, 204, 3
Cikota, A., Patat, F., Wang, L., et al. 2019, MNRAS, 490, 578
DES Collaboration, Abbott, T. M. C., Allam, S., et al. 2019, ApJL, 872, L30
Faber, S. M., Worthey, G., & Gonzales, J. J. 1992, in IAU Symp. 149, The Stellar Populations of Galaxies, ed. B. Barbuy & A. Renzini (Cambridge: Cambridge Univ. Press), 255
Foley, R. J., Challis, P. J., Chornock, R., et al. 2013, ApJ, 767, 57
Foreman-Mackey, D. 2016, JOSS, 1, 24
Freedman, W. L., Madore, B. F., Hatt, D., et al. 2019, ApJ, 882, 34
Gallagher, J. S., Garnavich, P. M., Caldwell, N., et al. 2008, ApJ, 685, 752
Garnavich, P. M., Bonanos, A. Z., Krisciunas, K., et al. 2004, ApJ, 613, 1120
Garnavich, P. M., Jha, S., Challis, P., et al. 1998, ApJ, 509, 74
Gupta, R. R., D'Andrea, C. B., Sako, M., et al. 2011, ApJ, 740, 92
Guy, J., Astier, P., Baumont, S., et al. 2007, A&A, 466, 11
Guy, J., Sullivan, M., Conley, A., et al. 2010, A&A, 523, A7
Hamuy, M., Phillips, M. M., Maza, J., et al. 1995, AJ, 109, 1
Hamuy, M., Phillips, M. M., Suntzeff, N. B., et al. 1996, AJ, 112, 2408
Hamuy, M., Trager, S. C., Pinto, P. A., et al. 2000, AJ, 120, 1479
Hicken, M., Wood-Vasey, W. M., Blondin, S., et al. 2009, ApJ, 700, 1097
Hunter, J. D. 2007, CSE, 9, 90
Jha, S., Riess, A. G., & Kirshner, R. P. 2007, ApJ, 659, 122
Jones, D. O., Riess, A. G., & Scolnic, D. M. 2015, ApJ, 812, 31
Jones, D. O., Riess, A. G., Scolnic, D. M., et al. 2018, ApJ, 867, 108
Jones, D. O., Scolnic, D. M., Foley, R. J., et al. 2019, ApJ, 881, 19
Kang, Y., Kim, Y.-L., Lim, D., Chung, C., & Lee, Y.-W. 2016, ApJS, 223, 7
Kang, Y., Lee, Y.-W., Kim, Y.-L., Chung, C., & Ree, C. H. 2020, ApJ, 889, 8
Kelly, B. C. 2007, ApJ, 665, 1489
Kelly, P. L., Hicken, M., Burke, D. L., Mandel, K. S., & Kirshner, R. P. 2010, ApJ, 715, 743
Kessler, R., & Scolnic, D. 2017, ApJ, 836, 56
Kim, Y.-L., Kang, Y., & Lee, Y.-W. 2019, JKAS, 52, 181
Kim, Y.-L., Smith, M., Sullivan, M., & Lee, Y.-W. 2018, ApJ, 854, 24
Kowal, C. T. 1968, AJ, 73, 1021
Kunz, M., Bassett, B. A., & Hlozek, R. A. 2007, PhRvD, 75, 103508
Lampeitl, H., Smith, M., Nichol, R. C., et al. 2010, ApJ, 722, 566
Le Borgne, D., & Rocca-Volmerange, B. 2002, A&A, 386, 446
McKinney, W. 2010, Data Structures for Statistical Computing in Python, https://conference.scipy.org/proceedings/scipy2010/mckinney.html
Moreno-Raya, M. E., López-Sánchez, Á R., Mollá, M., et al. 2016, MNRAS, 462, 1281
Mosher, J., Guy, J., Kessler, R., et al. 2014, ApJ, 793, 16
Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565
Perlmutter, S., Gabi, S., Goldhaber, G., et al. 1997, ApJ, 483, 565
Phillips, M. M. 1993, ApJL, 413, L105

Pskovskii, Y. P. 1969, SvA, 12, 750
Pskovskii, Y. P. 1977, SvA, 21, 675
Rest, A., Scolnic, D., Foley, R. J., et al. 2014, ApJ, 795, 44
Riddell, A., Hartikainen, A., Lee, D., et al. 2018, stan-dev/pystan, Zenodo, doi:10.5281/zenodo.1456206
Riess, A. G., Casertano, S., Yuan, W., et al. 2018, ApJ, 861, 126
Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, ApJ, 116, 1009
Riess, A. G., Press, W. H., & Kirshner, R. P. 1996, ApJ, 473, 88
Rigault, M., Brinnel, V., Aldering, G., et al. 2018, arXiv:1806.03849
Rigault, M., Copin, Y., Aldering, G., et al. 2013, A&A, 560, A66
Rose, B. M., Dixon, S., Rubin, D., et al. 2020, ApJ, 890, 60
Rose, B. M., Garnavich, P. M., & Berg, M. A. 2019, ApJ, 874, 32
Rubin, D., Aldering, G., Barbary, K., et al. 2015, ApJ, 813, 137
Rust, B. W. 1974, PhD thesis, Oak Ridge National Lab.
Sako, M., Bassett, B., Becker, A., et al. 2008, AJ, 135, 348
Sako, M., Bassett, B., Becker, A. C., et al. 2018, PASP, 130, 064002
Salpeter, E. E. 1955, ApJ, 121, 161
Scolnic, D., & Kessler, R. 2016, ApJL, 822, L35
Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, ApJ, 859, 101
Smith, M., Sullivan, M., Wiseman, P., et al. 2020, MNRAS, 494, 4426
Sullivan, M., Conley, A., Howell, D. A., et al. 2010, MNRAS, 406, 782
Suzuki, N., Rubin, D., Lidman, C., et al. 2012, ApJ, 746, 85
Tripp, R. 1998, A&A, 331, 815
Uddin, S. A., Mould, J., Lidman, C., Ruhlmann-Kleider, V., & Zhang, B. R. 2017, ApJ, 848, 56
van der Walt, S., Colbert, S. C., & Varoquaux, G. 2011, CSE, 13, 22
Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nat. Methods, 17, 261
Worthey, G., Faber, S. M., Gonzalez, J. J., & Burstein, D. 1994, ApJS, 94, 687