A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s-1 Mpc-1 Uncertainty from the Hubble Space Telescope and the SH0ES Team

Adam G. Riess, Wenlong Yuan, Lucas M. Macri, Dan Scolnic, Dillon Brout, Stefano Casertano, David O. Jones, Yukei Murakami, Gagandeep S. Anand, Louise Breuval, et al.

The Astrophysical Journal Letters, 934:L7 (52pp), 2022 July 20 https://doi.org/10.3847/2041-8213/ac5c5b

A Comprehensive Measurement of the Local Value of the Hubble Constant with 1 km s^{-1} Mpc^{-1} Uncertainty from the Hubble Space Telescope and the SH0ES Team

Adam G. Riess[1,2], Wenlong Yuan[2], Lucas M. Macri[3], Dan Scolnic[4], Dillon Brout[5], Stefano Casertano[1], David O. Jones[6], Yukei Murakami[2], Gagandeep S. Anand[1], Louise Breuval[2,7], Thomas G. Brink[8], Alexei V. Filippenko[8,9], Samantha Hoffmann[1], Saurabh W. Jha[10], W. D'arcy Kenworthy[2], John Mackenty[1], Benjamin E. Stahl[8], and WeiKang Zheng[8]

[1] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218, USA
[2] Department of Physics and Astronomy, Johns Hopkins University, Baltimore, MD 21218, USA
[3] George P. and Cynthia W. Mitchell Institute for Fundamental Physics and Astronomy, Department of Physics & Astronomy, Texas A&M University, College Station, TX 77843, USA
[4] Department of Physics, Duke University, Durham, NC 27708, USA
[5] Center for Astrophysics, Harvard & Smithsonian, 60 Garden Street, Cambridge, MA 02138, USA
[6] Einstein Fellow, Department of Astronomy & Astrophysics, University of California, Santa Cruz, CA 95064, USA
[7] LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Université de Paris, 5 place Jules Janssen, F-92195 Meudon, France
[8] Department of Astronomy, University of California, Berkeley, CA 94720-3411, USA
[9] Miller Institute for Basic Research in Science, University of California, Berkeley, CA 94720, USA
[10] Department of Physics and Astronomy, Rutgers, the State University of New Jersey, Piscataway, NJ 08854, USA
Received 2021 December 10; revised 2022 February 18; accepted 2022 March 8; published 2022 July 22

Abstract

We report observations from the Hubble Space Telescope (HST) of Cepheid variables in the host galaxies of 42 Type Ia supernovae (SNe Ia) used to calibrate the Hubble constant (H_0). These include the complete sample of all suitable SNe Ia discovered in the last four decades at redshift $z \leqslant 0.01$, collected and calibrated from $\geqslant$1000 HST orbits, more than doubling the sample whose size limits the precision of the direct determination of H_0. The Cepheids are calibrated geometrically from Gaia EDR3 parallaxes, masers in NGC 4258 (here tripling that sample of Cepheids), and detached eclipsing binaries in the Large Magellanic Cloud. All Cepheids in these anchors and SN Ia hosts were measured with the same instrument (WFC3) and filters (*F555W*, *F814W*, *F160W*) to negate zero-point errors. We present multiple verifications of Cepheid photometry and six tests of background determinations that show Cepheid measurements are accurate in the presence of crowded backgrounds. The SNe Ia in these hosts calibrate the magnitude–redshift relation from the revised Pantheon+ compilation, accounting here for covariance between all SN data and with host properties and SN surveys matched throughout to negate systematics. We decrease the uncertainty in the local determination of H_0 to 1 km s^{-1} Mpc^{-1} including systematics. We present results for a comprehensive set of nearly 70 analysis variants to explore the sensitivity of H_0 to selections of anchors, SN surveys, redshift ranges, the treatment of Cepheid dust, metallicity, form of the period–luminosity relation, SN color, peculiar-velocity corrections, sample bifurcations, and simultaneous measurement of the expansion history. Our baseline result from the Cepheid–SN Ia sample is $H_0 = 73.04 \pm 1.04$ km s^{-1} Mpc^{-1}, which includes systematic uncertainties and lies near the median of all analysis variants. We demonstrate consistency with measures from HST of the TRGB between SN Ia hosts and NGC 4258, and include them simultaneously to yield 72.53 ± 0.99 km s^{-1} Mpc^{-1}. The inclusion of high-redshift SNe Ia yields $H_0 = 73.30 \pm 1.04$ km s^{-1} Mpc^{-1} and $q_0 = -0.51 \pm 0.024$. We find a 5σ difference with the prediction of H_0 from Planck cosmic microwave background observations under ΛCDM, with no indication that the discrepancy arises from measurement uncertainties or analysis variations considered to date. The source of this now long-standing discrepancy between direct and cosmological routes to determining H_0 remains unknown.

Unified Astronomy Thesaurus concepts: Hubble constant (758)

1. Introduction

The present expansion rate of the universe, the Hubble constant (H_0), sets its size and age scale, relating redshift (the direct consequence of expansion) to distance and time. The value of H_0 may be determined locally with measurements of distances and redshifts, and it can also be predicted from a cosmological model calibrated in the early universe (i.e., pre-recombination at redshift $z \geqslant 1100$) with measurements of the cosmic microwave background (CMB). The comparison of measured and predicted values of H_0 thus provides a crucial "end-to-end" test of the widest available range of the validity of cosmological models, from early times when the universe is dense and dominated by dark matter and radiation, to the present when it is dilute and dominated by dark energy.

1.1. The SH0ES Program

The Hubble constant is the most accessible parameter in the cosmological model. It can be estimated with a wide range of approaches and accuracies from limited knowledge of many types of astronomical sources, nearly all of which have been utilized in this endeavor over the past century. There have been >1000 estimates published since 1980, with one-third of those

in the last five years and 20% in the last two years—a recent quadrupling of the effort indicating the accelerating interest in H_0 (Steer 2020). However, past discrepancies internal to the body of local measurements reveal that systematic errors can dominate determinations of H_0, and that there is no reason to believe all efforts will regress to the mean or that a more accurate result can be derived from their median (Chen & Ratra 2011). Rather, to keep systematic uncertainties in check, it is necessary to pursue the most powerful, simplest, and most reliable tools, with strict attention paid to understanding, mitigating, and accounting for sources of measurement error.

Since the launch of the Hubble Space Telescope (HST) with a design goal of achieving a 10% determination of H_0, the leading approach to measuring it in the local universe (as indicated by the observing time competitively awarded by the community) has relied on imaging of Cepheid variable stars in the host galaxies of recent, nearby Type Ia supernovae (SNe Ia). Cepheids have been favored as primary distance indicators because they are very luminous ($M_V \approx -6$ mag at $P \approx 30$ days), are easy to identify thanks to their periodicity (Leavitt & Pickering 1912), obey a tight period–luminosity relation (P–L, the "Leavitt Law") that yields extremely precise distances (3% per source; Riess et al. 2019a, hereafter R19), and have well-understood physics (Eddington 1917; Bono et al. 1999). Other primary distance indicators that have been measured in the hosts of SNe Ia to determine H_0 include the tip of the red giant branch (TRGB; Freedman et al. 2019, hereafter F19) and Mira variable stars (Huang et al. 2020).

Cepheids also offer the most opportunities for obtaining strictly differential flux measurements—i.e., the use of the same facility to measure calibrator and source, a key requirement for eliminating zero-point errors. This is feasible through the use of HST to directly observe Cepheids in a large set of SN hosts and in geometric calibrators of Cepheid luminosities: the megamaser host NGC 4258, the Milky Way (hereafter MW) with plentiful parallaxes, and the Large Magellanic Cloud (hereafter LMC) via detached eclipsing binaries (DEBs). SNe Ia are favored to measure the Hubble expansion owing to their high precision (5% in distance per source), ubiquity, and deep reach, which reduces the impact of local flows.

The SH0ES program (Supernovae and H_0 for the Equation of State of dark energy) began in 2005 with a proposal in HST Cycle 15 to break the degeneracy among cosmological parameters used to model CMB data and the equation-of-state parameter $w = P/(\rho c^2)$, where P is the pressure and ρ is the mass density of dark energy. Its stated ambitious goal, based on the recommendation by Hu (2005), was to eventually reach a percent-level measurement of H_0, a goal approached if not fully reached in this work. This project was a "second-generation" effort to measure H_0 with HST from a distance ladder of Cepheids and SNe Ia using the then recently installed ACS (and later WFC3) instruments, following successful efforts during the 1990s by the "first-generation" HST Key Project on the Extragalactic Distance Scale (Freedman et al. 2001) and the SNe Ia Luminosity Calibration Program (Sandage et al. 2006), both of which primarily used WFPC2. The former searched for Cepheids in the hosts of numerous secondary distance indicators (excluding SNe Ia) while the latter focused on SN Ia hosts. The types of targets suitable for Cepheid searches and the observing sequences for use with HST were developed and first implemented by these ground-breaking programs.

The precision of the first-generation programs was ultimately limited by the lack of a precise geometric calibration of Cepheid variables, by the limited practical range of WFPC2 to measure Cepheids in SN Ia hosts (distance $D \lesssim 20$–25 Mpc), by the impact of reddening in the optical, and by the limited characterization of the Cepheid metallicity dependence at those wavelengths. The ability of SNe Ia to measure individual distances with ∼5%–10% precision and to sharply delineate the Hubble flow began with the use of light-curve versus luminosity relations (Phillips 1993), SN Ia colors (Riess et al. 1996; Tripp 1998; Phillips et al. 1999), and modern, digital samples (Hamuy et al. 1996; Riess et al. 1999).

Unfortunately, SNe Ia at $D \leqslant 20$ Mpc are rare, occurring about once per decade, with most of the few objects in this range observed up to a century ago using photographic technology. Such observations lacked the photometric precision, well-characterized bandpasses, and accurate determinations of host-galaxy backgrounds, SN light-curve shapes, and SN colors to take advantage of the new standardization methods. The tendency for intrinsically brighter SNe Ia with broader light curves to occur in (late-type) Cepheid hosts would also bias H_0 lower without light-curve standardization. A number of systematic differences in the first-generation calibration of SNe Ia by Sandage et al. (2006) were quantified by Riess et al. (2005, Table 16). These differences, totaling about 20%, arose from several effects which were amplified by small sample statistics: problematic SN Ia data such as photographic photometry, highly extinguished objects, and poorly sampled light curves; from photometric anomalies in WFPC2, such as the "long versus short" effect (Holtzman et al. 1995) and charge transfer efficiency (CTE; e.g., Whitmore et al. 1999); and from limited knowledge of the slope of the Cepheid P–L relation. The present geometric calibration of the distance to the LMC by Pietrzyński et al. (2019) using DEBs is also 7% smaller than the value assumed by Sandage et al. (2006) to calibrate Cepheids.

The SH0ES program has been designed to improve upon past determinations of H_0 by (1) extending the range of Cepheid observations with ACS and WFC3 to reach the hosts of a large sample of "ideal" SNe Ia, free from the preceding problems; (2) using near-infrared (NIR) observations of all Cepheids in SN Ia hosts with NICMOS and WFC3 to reduce the systematic uncertainty associated with the reddening laws for Cepheids and their hosts and the Cepheid metallicity dependence; and (3) calibrating Cepheids with new, geometric distances tied directly with HST to the Cepheids in SN Ia hosts to nullify zero-point uncertainties. "Ideal" or suitable SNe Ia for calibrating H_0 (given limited HST time) were defined by Riess et al. (2005) to be (1) observed before maximum light, (2) through low interstellar extinction ($A_V < 0.5$ mag), (3) with the same instruments and filters as the SNe Ia in the Hubble flow (at that time obtained by the Calán/Tololo and CfA surveys), and (4) to have typical light-curve shapes.[11] These characteristics are necessary to provide low dispersion in the Hubble flow, but they applied to only three Cepheid-calibrated SNe Ia from the first-generation projects (SNe 1981B, 1990N, and 1998aq).

Unlike the first-generation programs, which were granted long-term status and a large initial allocation of observing time with HST, the SH0ES program was proposed to the STScI

[11] These color and shape requirements translate in the Pantheon SN standardization (Scolnic et al. 2018) as $|c| < 0.2$ and $|x1| < 2$.

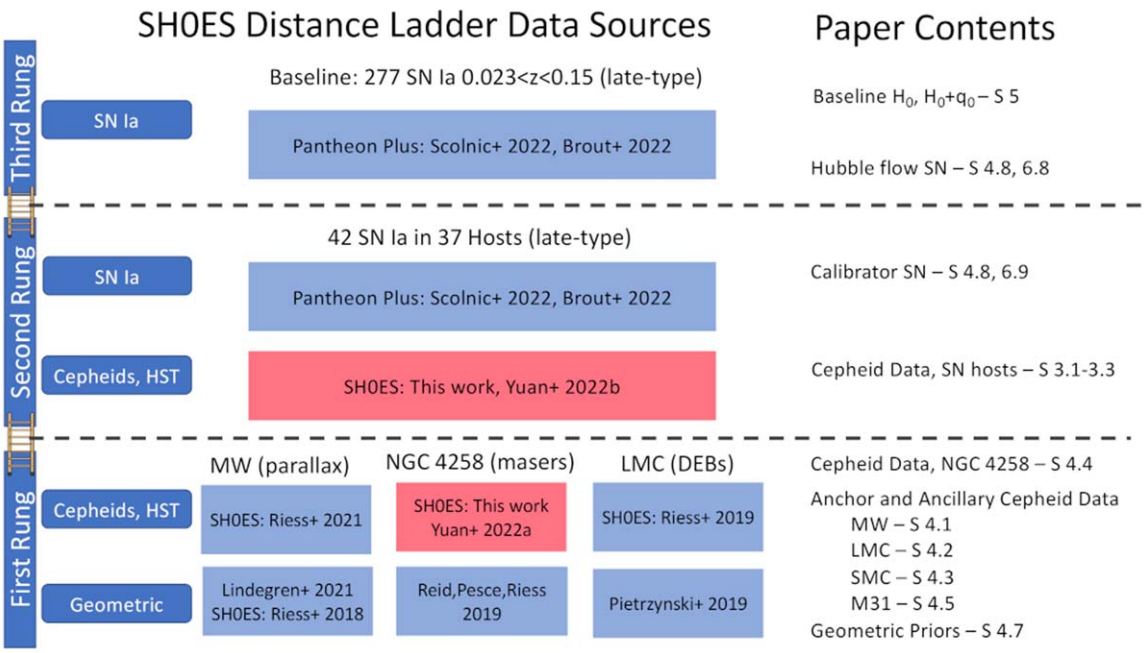

Figure 1. Sources of data for the distance ladder. The red block shows data from this work.

time-allocation committee year by year. The cumulative result, after 15 yr (cycles), has been to collect Cepheid observations in 37 hosts of 42 SNe Ia and calibrate them geometrically to Cepheids in the MW, the LMC, and NGC 4258, with a total of 18 individual HST proposals utilizing >1000 orbits. The source of the distance-ladder data is shown in Figure 1, and the source and sequence of all observations in SN Ia hosts are shown in Figure 2 and listed in Table 1. Observing Cepheids in hosts at $D \approx 40$–50 Mpc, double the range of first-generation observations and a factor of 8 increase in targets, was feasible owing to two features of WFC3: significantly better sampling of the point-spread function (PSF); with pixel sizes a factor of 2.5 smaller than the wide channel of WFPC2, which greatly reduced the impact of crowded backgrounds; and a white-light filter (*F350LP*) that, combined with the better sensitivity of WFC3, reduced the observing time required to identify Cepheids and measure their periods by a factor of ∼4. Calibrating those Cepheids differentially to bright Cepheids in the MW and LMC became feasible through the development of spatial scanning of HST and a rapid slew and guiding mode under gyroscopic control.

The first SH0ES results (Riess et al. 2009, hereafter R09) were based on Cepheids observed in the hosts of six ideal SN Ia calibrators using ACS (optical) and NICMOS (NIR), and one geometric anchor (NGC 4258) with a maser-based distance of 3% precision (Humphreys et al. 2013) and a large sample of Cepheids (Macri et al. 2006). The result was a 5% measurement of H_0, 74.2 ± 3.6 km s^{-1} Mpc^{-1}, which, combined with the 5 yr WMAP results (Komatsu et al. 2009), yielded $w = -1.12 \pm 0.12$ and was consistent with a value of $H_0 = 71.9 \pm 2.6$ determined from WMAP and ΛCDM alone.[12] The second iteration (Riess et al. 2011, hereafter R11) increased the calibrator sample to eight SNe Ia, observed and measured all Cepheids with WFC3 (both optical and NIR), and expanded the geometric calibration of Cepheids beyond NGC 4258 by including two additional independent anchors: the LMC, through various DEB-based distances (e.g., Pietrzyński et al. 2009), and the MW via parallaxes measured with the Fine Guidance Sensor on HST (Benedict et al. 2007). This resulted in $H_0 = 73.8 \pm 2.4$, which, coupled with the 7 yr WMAP results (Komatsu et al. 2011), yielded $w = -1.08 \pm 0.10$, or an estimate of the effective number of relativistic species of $N_{\rm eff} = 4.2 \pm 0.7$. This result was closely matched by a recalibration of the final HST Key Project results using the same MW parallaxes, different Cepheid measurements, and an updated Hubble diagram of SNe Ia, which yielded $H_0 = 74.4 \pm 2.2$ (Freedman et al. 2012, hereafter F12). Indications of any tension between the early and late universe at that time were $<2\sigma$ in significance.

1.2. The Hubble Tension

The first release of CMB data from the ESA Planck mission (Planck Collaboration et al. 2014) yielded $H_0 = 67.2 \pm 1.2$ in the context of ΛCDM, a then 2σ reduction relative to WMAP and a difference of 3σ from the results of R11 and F12. Reanalyses of the R11 data (Fiorentino et al. 2013; Efstathiou 2014; Zhang et al. 2017) produced essentially the same results as R11, with H_0 ranging from 72.5 to 76.0. The third iteration of SH0ES (Riess et al. 2016, hereafter R16) more than doubled the calibrator sample to 19 SNe Ia, used refined distance estimates to NGC 4258 and the LMC (Pietrzyński et al. 2013), and new Cepheid parallaxes measured by the SH0ES team using spatial scanning with WFC3 (Riess et al. 2014; Casertano et al. 2016) to reach a 2.4% determination of $H_0 = 73.2 \pm 1.7$, 3.4σ greater than the refined value from Planck+ΛCDM of 66.9 ± 0.6 (Planck Collaboration et al. 2016), a difference of ∼9% or 0.2 mag in units of $5 \log H_0$. An extensive number of reanalyses of the R16 data with many variations were undertaken (Cardona et al. 2017; Feeney et al. 2019; Follin & Knox 2018; Bovy 2018; Burns et al. 2018; Dhawan et al. 2018; Avelino et al. 2019), resulting in H_0 values

[12] To improve readability, the conventional units of H_0 will frequently be omitted in the rest of this paper. ΛCDM refers to the standard cosmological model with the cosmological constant (Λ) and cold dark matter (CDM).

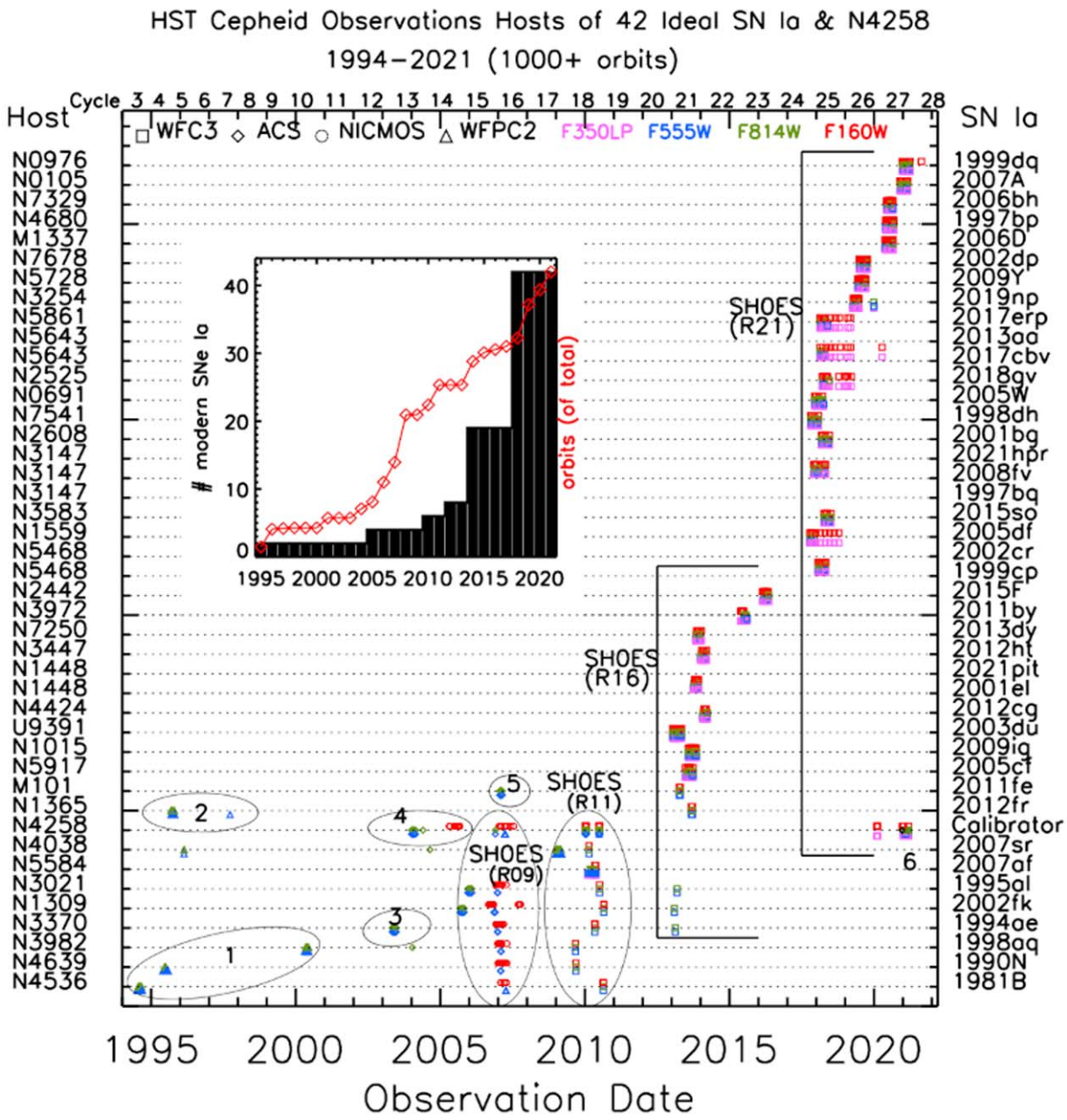

Figure 2. HST observations of Cepheids in 37 hosts of 42 "ideal" SNe Ia and NGC 4258, collected over 20 yr with four cameras and >1000 orbits of HST time. In most cases, 60 to 90 day campaigns in *F555W* and *F814W* or in *F350LP* were used to identify Cepheids from their light curves with occasional observations years later to identify longer-period Cepheids. NIR follow-up observations in *F160W* are used to reduce the effects of host-galaxy extinction, sensitivity to metallicity, and breaks in the *P*–*L* relation. Data sources: (1) HST Key Project, Freedman et al. (2001); (2) HST SN Ia Luminosity Calibration Program, Sandage et al. (2006); (3) Riess et al. (2005); (4) Macri et al. (2006); (5) Mager et al. (2013); (6) Yuan et al. (2022).

ranging from 73 to 74 and uncertainties from 2% to 2.5%. These analyses explored varying reddening laws, use of NIR SN data, use of alternative SN light-curve fitting, and hierarchical Bayesian statistics for data fitting, resulting in little change in H_0 or its uncertainty. Javanmardi et al. (2021) selected the Cepheid host from R16 with the largest sample of Cepheids (NGC 5584, also near the median sample host distance) and remeasured these starting from the archived HST pixels and using different methods to photometry, finding agreement with the R16 measurement to 1% precision and ruling out a significant methodological error in these measurements.

Since then, additional sources of MW parallax calibration of Cepheid luminosities have come from further use of HST WFC3 spatial scanning (Riess et al. 2018b), from Gaia DR2 Cepheid parallaxes with HST photometry (Riess et al. 2018a), from Gaia DR2 Cepheid binary companions and cluster hosts (Breuval et al. 2020), and from Gaia EDR3 parallaxes coupled with additional HST photometry (Riess et al. 2021, hereafter R21). Likewise, HST observations of 70 long-period Cepheids in the LMC (Riess et al. 2019b) and improved distance estimates to the LMC (Pietrzyński et al. 2019) and NGC 4258 (Reid et al. 2019) resulted in $H_0 = 73.2 \pm 1.3$, raising the difference with Planck+ΛCDM to 4.2σ. Other precise measures of H_0 in the local universe from the distance–

Table 1
Cepheid Observations with HST

Galaxy	SN(e) Ia	Exposure Time [s]				
		WFC3		All	NIR	UT
		NIR[a]	UVIS[b]	opt.[c]	Prop ID(s)	Date[d]
M101	2011fe	4846	3776	53072	12880	2013-03-03
Mrk 1337	2006D	13823	25994	25994	15640	2020-04-15
N0105	2007A	13270	34302	34302	16269	2020-10-23
N0691	2005W	9647	31413	31413	15145	2017-11-10
N0976	1999dq	15482	34312	34312	16269	2020-11-21
N1015	2009ig	14364	39336	39336	12880	2013-06-30
N1309	2002fk	6991	30020	89282	11570,12880	2010-07-24
N1365	2012fr	3617	31800	91200	12880	2013-08-06
N1448	2001el, 2021pit	6035	17562	17562	12880	2013-09-15
N1559	2005df	10058	22245	22245	15145	2017-09-09
N2442	2015F	6035	20976	20976	13646	2016-01-21
N2525	2018gv	9821	21177	21177	15145	2018-02-14
N2608	2001bg	9647	26942	26942	15145	2018-02-04
N3021	1995al	4426	29620	88722	11570,12880	2010-06-03
N3147	1997bq, 2008fv, 2021hpr	14470	37426	37426	15145	2017-10-28
N3254	2019np	8441	22106	22106	15640	2019-03-11
N3370	1994ae	4376	29820	88222	11570,12880	2010-04-04
N3447	2012ht	4529	19114	19114	12880	2013-12-15
N3583	2015so	9647	27001	27001	15145	2018-03-06
N3972	2011by	6635	19932	19932	13647	2015-04-19
N3982	1998aq	4017	14000	90840	11570	2009-08-04
N4038	2007sr	6794	20640	85684	11577	2010-01-22
N4258	Anchor	40234	10120	103690	11570	2020-01-02
N4424	2012cg	3623	17782	17782	12880	2014-01-08
N4536	1981B	2564	26000	95000	11570	2010-07-19
N4639	1990N	5379	16000	77480	11570	2009-08-07
N4680	1997bp	9647	25217	25217	15640	2020-04-24
N5468	1999cp,2002cr	14470	36566	36566	15145	2017-12-22
N5584	2007af	4929	74940	74940	11570	2010-04-04
N5643	2013aa,2017cbv	9052	24741	24741	15145	2018-01-16
N5728	2009Y	13823	26111	26111	15640	2019-05-05
N5861	2017erp	10058	21798	21798	15145	2018-01-13
N5917	2005cf	7235	23469	23469	12880	2013-05-20
N7250	2013dy	5435	18158	18158	12880	2013-11-08
N7329	2006bh	9044	24665	24665	15640	2020-05-06
N7541	1998dh	9647	26766	26766	15145	2017-09-21
N7678	2002dp	12058	32060	32060	15640	2019-05-25
U9391	2003du	13711	39336	39336	12880	2012-12-14

Notes.
[a] Obtained with WFC3/NIR and *F160W*;
[b] Obtained with WFC3/UVIS and *F555W*, *F814W*, or *F350LP* used to find and measure the flux of Cepheids;
[c] Includes time-series data from an earlier program and a different camera—see Figure 2;
[d] Date of first WFC3/NIR observation.

redshift relation generally range from 70 to 75, and those grounded in the pre-recombination version of ΛCDM range from 67 to 68, and have been extensively reviewed (Verde et al. 2019; Di Valentino et al. 2021; Shah et al. 2021). Because the tension is seen between different routes, which are comparable only via an accurate cosmological model, numerous possible theoretical explanations for the emergent "Hubble tension" have been proposed but no consensus has yet emerged (Di Valentino et al. 2021). Indeed, theoretical priors weigh heavily on these proposals or whether it may be considered "extraordinary" for the ΛCDM model to fail or pass this cosmic test. The test itself, however, is empirical and few would conclude it has yet been satisfactorily passed.

1.3. This Work

In this publication, we more than double the sample of SN Ia calibrators from 19 in R16 to reach 42 objects in 37 hosts. This increase is a milestone for a sample whose size has limited the precision to which H_0 can be locally measured. It provides the largest increase in size we can anticipate in the remaining lifetime of HST as it now includes all suitable SNe Ia (of which we are aware) observed between 1980 and 2021 at $z < 0.011$ and which slowly accrue at $\sim$1 yr^{-1}. We have also reprocessed and reanalyzed the NIR observations of Cepheids in the previous 19 hosts reported by R16 for consistency with the new sample. We make use of an automated pipeline (W. Yuan et al. 2022, in preparation, hereafter Y22b) to find the Cepheids in

18 new hosts (and reanalyze past hosts), which follows the steps developed manually for the first 19 hosts (Hoffmann et al. 2016, hereafter H16). We benefit from a factor of 3 increase in the sample of Cepheids within NGC 4258 discovered by observing four new fields with HST, fully reanalyzed using the same pipeline by (Yuan et al. 2022, hereafter Y22a). Extensive details concerning the analysis of SN Ia data are given by Scolnic et al. (2021) and Brout et al. (2021).

We present the formalism for measuring H_0 from the distance ladder in Section 2; new Cepheid data in Section 3 and in Y22a+Y22b; ancillary data used to measure H_0 in Section 4; our baseline, local determination of H_0 and a simultaneous measurement of H_0 and the expansion history with high-redshift SNe Ia in Section 5; extensions to the baseline and variants of the local measurement of H_0 in Section 6; discussion in Section 7; and conclusions in Section 8. Appendices provide further details on characterizing the spectral and photometric properties of the 42 calibrator SNe Ia (A), independent tests of the accuracy of Cepheid photometry (B), the Cepheid metallicity scale (C), and alternative applications of "Wesenheit" magnitudes (D).

2. Measuring the Hubble Constant

2.1. Distance-ladder Formalism

The Hubble constant is the present relation between redshift and distance, $cz = H_0D$, measured at cosmological distances where expansion is the dominant source of redshift. Here it is measured via a three-step (or three-rung) distance ladder employing a single, simultaneous fit between (1) geometric distance measurements to standardized Cepheid variables, (2) standardized Cepheids and colocated SNe Ia in nearby galaxies, and (3) SNe Ia in the Hubble flow. The fit is accomplished simultaneously by optimizing a χ^2 statistic to determine the most likely values of the parameters in the relevant relations. The data include measurements, their uncertainties, and their covariances as described in the next section. The parameters are the distances of all hosts and five additional parameters: the fiducial luminosity of SNe Ia and Cepheids, two parameters standardizing Cepheid luminosities (their dependence on period and metallicity), and H_0. This parameterization of the distance ladder can be expressed as a simple system of linear equations in a compact set of matrices as given below, useful for transmission, and for which the maximum-likelihood solution is easily found. The distance-ladder data and scheme are displayed in approximate form in Figure 1.

To briefly summarize the relevant relations, the distance modulus of a source is $\mu_0 = m_0 - M_0 = 5 \log D + 25$, with D the luminosity distance in Mpc, m the apparent magnitude (flux), M the absolute magnitude (luminosity), and the subscript 0 denoting a magnitude free of (or corrected for) intervening absorption by interstellar dust. The form of the dependence of Cepheid or SN Ia luminosity on observed characteristics (i.e., "standardization") has been well determined by prior work and is only briefly reviewed here. The dereddened Cepheid apparent magnitudes (also called "Wesenheit" magnitudes; Madore 1982) at mean phase will be described in Section 3.3 and are identified here as m_H^W (in a specific host like NGC 4258, those Cepheids would have magnitudes given as $m_{H,N4258}^W$). For the jth such Cepheid magnitude in the ith host given the period $P_{i,j}$ in days and metallicity $[O/H]_{i,j}$ relative to the Sun, we have

$$m_{H,i,j}^W = \mu_{0,i} + M_{H,1}^W + b_W (\log P_{i,j} - 1) + Z_W [O/H]_{i,j}, \quad (1)$$

where $M_{H,1}^W$ is the fiducial absolute magnitude (in the Wesenheit magnitude system of 3.4) of a Cepheid with $\log P = 1$ ($P = 10$ days) and solar metallicity, and the parameters b_W and Z_W (sometimes called γ in the literature) define the empirical relation between Cepheid period, metallicity, and luminosity. The $[O/H]_{i,j}$ is inferred at its galactocentric radial position as described in Section 3.5.

Any number of Cepheid hosts may have an independent geometric distance that contributes to the calibration of the Cepheid luminosity; e.g., for variables observed in the maser host NGC 4258 we adopt a distance modulus $\mu_{0,N4258}$, the best estimate of the distance, with formal uncertainty $\sigma(\mu_{0,N4258})$ from Reid et al. (2019). In this case the individual host distance parameter $\mu_{0,i}$ above is replaced with the external constraint, converting the apparent magnitudes to absolute,

$$M_{H,j}^W = m_{H,N4258,j}^W - \mu_{0,N4258} + \Delta\mu_{N4258}, \quad (2)$$

and introducing a new parameter $\Delta\mu_{N4258}$ as the difference from the measured and true distance with the additional, simultaneous constraint equation $0 = \Delta\mu_{N4258} \pm \sigma(\mu_{0,N4258})$. This definition allows the simultaneous use of multiple geometric "anchors" to calibrate the distance ladder; as we later show, it also allows the use of additional distance indicators for the same anchor, such as the tip of the red giant branch (TRGB) or Mira variable stars, while keeping track of their mutual dependence on the same geometric constraint.

A set of hosts of both SNe Ia and Cepheids connects the two distance indicators. Thus, for an SN Ia in the ith Cepheid host,

$$m_{B,i}^0 = \mu_{0,i} + M_B^0, \quad (3)$$

where $m_{B,i}^0$ is its maximum-light apparent magnitude that has been standardized (i.e., corrected for variations around the fiducial color, luminosity, and any host dependence; see Scolnic et al. 2021), M_B^0 is the fiducial SN Ia luminosity, and $\mu_{0,i}$ is the same parameter as in Equation (1). For SNe Ia, unlike Cepheids, the convention for keeping track of covariance in the standardization, as described by Scolnic et al. (2021), is to employ a set of standardized $m_{B,i}^0$, their uncertainties, and the covariance between any pair. This is an equally mathematically valid approach as keeping track of covariance through the standardizing relation for Cepheids in Equation (1).

The ladder is completed with a set of SNe Ia that measure the expansion rate quantified as the intercept, a_B, of the distance (or magnitude)–redshift relation. This is simply $a_B = \log cz - 0.2m_B^0$ in the low-redshift limit ($z \approx 0$) but given for an arbitrary expansion history and for $z > 0$ as

$$a_B = \log cz \left\{ 1 + \frac{1}{2}[1 - q_0]z - \frac{1}{6}[1 - q_0 - 3q_0^2 + j_0]z^2 + O(z^3) \right\} - 0.2m_B^0, \quad (4)$$

measured from a set of SNe Ia (z, m_B^0), where z is the redshift due to expansion, q_0 is the deceleration parameter, and j_0 is the jerk (see Visser 2004 for definitions). The determination of H_0 follows from

$$\log H_0 = 0.2M_B^0 + a_B + 5. \quad (5)$$

If the set of standardized SN Ia magnitudes in the hosts of Cepheids, which serve to calibrate M_B^0 (hereafter "calibrators" or CC SNe Ia), and those in the Hubble flow used to measure a_B (hereafter HF SNe Ia) have no common sources of uncertainty (i.e., no covariance), then Equations (3) and (4) and the ladder parameters they provide (M_B^0 and a_B) can be determined independently. This was the approach taken by R16. However, an increasingly thorough quantification of systematic uncertainties in SN Ia measurements, and the standardization undertaken as common practice for the determination of w (Scolnic et al. 2018), has demonstrated nontrivial covariance of SN Ia data, quantified following the approach of Conley et al. (2011) and Dhawan et al. (2020). We therefore undertake the optimization of Equations (3) and (5) simultaneously.

It is useful to expand the same set of Equations (1)–(5) in the form of matrices that organize the data into a vector of magnitude measurements y, the covariance matrix of standard errors of the magnitude measurements C, the equation matrix L, and the vector of free parameters q (hence the model, Lq) as follows:

$$y = \begin{pmatrix} m^W_{H,1} \\ .. \\ m^W_{H,\mathrm{nh}} \\ \hline m^W_{H,\mathrm{N4258}} - \mu_{0,\mathrm{N4258}} \\ m^W_{H,\mathrm{M31}} \\ m^W_{H,\mathrm{LMC}} - \mu_{0,\mathrm{LMC}} \\ \hline m^0_{B,1} \\ .. \\ m^0_{B,\mathrm{ncc}} \\ \hline M^W_{H,1,\mathrm{HST}} \\ M^W_{H,1,\mathrm{Gaia}} \\ 0 \\ 0 \\ 0 \\ \hline m^0_{B,1} - 5\log cz_1\{\} - 25 \\ .. \\ m_{B,\mathrm{nhf}} - 5\log cz_{\mathrm{nhf}}\{\} - 25 \end{pmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \end{matrix} \begin{matrix} \} \text{Cepheids in SN Ia hosts} \\ \\ \\ \} \text{Cepheids in anchors or non} - \text{SN Ia hosts} \\ \\ \\ \} \text{SNe Ia in Cepheid hosts} \\ \\ \\ \} \text{External constraints} \\ \\ \\ \\ \} \text{SNe Ia in the Hubble flow} \end{matrix}$$

$$C = \begin{pmatrix} \sigma^2_{\mathrm{tot},1} & .. & Z_{\mathrm{cov}} & Z_{\mathrm{cov}} & 0 & 0 & 0 & .. & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .. & 0 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ Z_{\mathrm{cov}} & .. & \sigma^2_{\mathrm{tot,nh}} & Z_{\mathrm{cov}} & 0 & 0 & 0 & .. & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .. & 0 \\ \hline Z_{\mathrm{cov}} & .. & Z_{\mathrm{cov}} & \sigma^2_{\mathrm{tot,N4258}} & 0 & 0 & 0 & .. & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .. & 0 \\ 0 & .. & 0 & 0 & \sigma^2_{\mathrm{tot,M31}} & 0 & 0 & .. & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .. & 0 \\ 0 & .. & 0 & 0 & 0 & \sigma^2_{\mathrm{tot,LMC}} & 0 & .. & 0 & 0 & 0 & 0 & 0 & 0 & 0 & .. & 0 \\ \hline 0 & .. & 0 & 0 & 0 & 0 & \sigma^2_{m_{\mathrm{B},1}} & .. & \mathrm{SN_{cov}} & 0 & 0 & 0 & 0 & 0 & \mathrm{SN_{cov}} & .. & \mathrm{SN_{cov}} \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ 0 & .. & 0 & 0 & 0 & 0 & \mathrm{SN_{cov}} & .. & \sigma^2_{m_{\mathrm{B,ncc}}} & 0 & 0 & 0 & 0 & 0 & \mathrm{SN_{cov}} & .. & \mathrm{SN_{cov}} \\ \hline 0 & .. & 0 & 0 & 0 & 0 & 0 & .. & 0 & \sigma^2_{\mathrm{HST}} & 0 & 0 & 0 & 0 & 0 & .. & 0 \\ 0 & .. & 0 & 0 & 0 & 0 & 0 & .. & 0 & 0 & \sigma^2_{\mathrm{Gaia}} & 0 & 0 & 0 & 0 & .. & 0 \\ 0 & .. & 0 & 0 & 0 & 0 & 0 & .. & 0 & 0 & 0 & \sigma^2_{\mathrm{grnd}} & 0 & 0 & 0 & .. & 0 \\ 0 & .. & 0 & 0 & 0 & 0 & 0 & .. & 0 & 0 & 0 & 0 & \sigma^2_{\mu,\mathrm{N4258}} & 0 & 0 & .. & 0 \\ 0 & .. & 0 & 0 & 0 & 0 & 0 & .. & 0 & 0 & 0 & 0 & 0 & \sigma^2_{\mu,\mathrm{LMC}} & 0 & .. & 0 \\ \hline 0 & .. & 0 & 0 & 0 & 0 & \mathrm{SN_{cov}} & .. & \mathrm{SN_{cov}} & 0 & 0 & 0 & 0 & 0 & \sigma^2_{m_{\mathrm{B},z,1}} & .. & \mathrm{SN_{cov}} \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ 0 & .. & 0 & 0 & 0 & 0 & \mathrm{SN_{cov}} & .. & \mathrm{SN_{cov}} & 0 & 0 & 0 & 0 & 0 & \mathrm{SN_{cov}} & .. & \sigma^2_{m_{\mathrm{B},z,\mathrm{nhf}},} \end{pmatrix}$$

where $\sigma^2_{\rm tot,1}$ representing the $n \times n$ covariance matrix for the Cepheids in the first host is expanded as

$$\sigma^2_{{\rm tot},j} = \begin{pmatrix} \sigma^2_{\rm tot,1,1} & .. & C_{1,1,n,\rm bkgd} \\ .. & .. & .. \\ C_{1,n,1,\rm bkgd} & .. & \sigma^2_{\rm tot,1,n} \end{pmatrix};$$

$$L = \begin{pmatrix} 1 & .. & 0 & 0 & 1 & 0 & 0 & \log P_{N,1} - 1 & 0 & [{\rm O/H}]_{N,1} & 0 & 0 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & 0 \\ 0 & .. & 1 & 0 & 1 & 0 & 0 & \log P_{N,\rm nh} - 1 & 0 & [{\rm O/H}]_{N,\rm nh} & 0 & 0 \\ \hline 0 & .. & 0 & 1 & 1 & 0 & 0 & \log P_{\rm N4258} - 1 & 0 & [{\rm O/H}]_{\rm N4258} & 0 & 0 \\ 0 & .. & 0 & 0 & 1 & 0 & 1 & \log P_{\rm M31} - 1 & 0 & [{\rm O/H}]_{\rm M31} & 0 & 0 \\ 0 & .. & 0 & 0 & 1 & 1 & 0 & \log P_{\rm LMC} - 1 & 0 & [{\rm O/H}]_{\rm LMC} & 1 & 0 \\ \hline 1 & .. & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & 0 \\ 0 & .. & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & .. & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .. & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .. & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & .. & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & .. & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & .. & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. & .. \\ 0 & .. & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \end{pmatrix}$$

Cepheids in SN Ia hosts

Cepheids in anchors or non − SN Ia hosts

SNe Ia in Cepheid hosts

External constraints

SNe Ia in the Hubble flow

$$q = \begin{pmatrix} \mu_{0,1} \\ .. \\ \mu_{0,\rm nh} \\ \Delta\mu_{\rm N4258} \\ M^W_{H,1} \\ \Delta\mu_{\rm LMC} \\ \mu_{M31} \\ b_W \\ M^0_B \\ Z_W \\ \Delta{\rm zp} \\ 5\log {\rm H}_0 \end{pmatrix}.$$

The number of Cepheid hosts is nh, the number of SNe Ia in these hosts is ncc, and the number of SNe Ia in the Hubble flow is nhf. Period uncertainties are comparatively negligible (Yuan et al. 2021). $\sigma_{\rm HST}$ and $\sigma_{\rm Gaia}$ denote the uncertainties in $M^W_{H,1,\rm Gaia}$ and $M^W_{H,1,\rm HST}$ as derived from parallaxes, respectively, while $\sigma_{\rm grnd}$ denotes the uncertainty in ground-based photometry. The term $\rm SN_{cov}$ is the covariance between SNe, $Z_{\rm cov}$ is the metallicity covariance given later in Equation (9), and $C_{i,j,k,\rm bkgd}$ is the background covariance given later in Equation (8). The χ^2 statistic is given as

$$\chi^2 = (y - Lq)^T C^{-1} (y - Lq) \quad (6)$$

and maximum-likelihood parameters are given as $q_{\rm best} = (L^T C^{-1} L)^{-1} L^T C^{-1} y$, while the standard errors and the covariance matrix of the parameters come from the matrix $cq = (L^T C^{-1} L)^{-1}$. The value of H_0 is derived from the final entry of q, 5 log H_0, and its error from the square root of the corner entry of cq. This provides a compact form for storing and transmitting the full data set used to determine H_0, to edit or augment it, and to enable others to determine its value. The above formalism is the same as used by R16 with the additions of SN and Cepheid covariance. We will also derive the parameters independently of the analytical solution in Section 5.1 by sampling the χ^2 statistic using a Markov Chain Monte Carlo (MCMC) approach to verify the analytical result with a different methodology.

3. Cepheid Observations in SN Ia Hosts and the Maser Host NGC 4258

3.1. Optical Cepheid Discovery

The SH0ES program has been selecting the SNe Ia that are most suitable for calibrating their fiducial luminosity (with selection criteria given in Section 1.1) through observations of Cepheids in their hosts. The results here include a complete sample of all such SNe Ia of which we are aware within $z < 0.011$ (40 objects), with the addition of two beyond this limit that are useful for testing the range of Cepheid distance measurements, for a total of 42 SN Ia calibrators.

Figure 2 and Table 1 show the sources of the HST data obtained for every SN Ia and host we measured, gathered from the indicated HST cameras, filters, and time periods. All of these publicly available data are readily obtained from the Mikulski Archive for Space Telescopes (MAST). The imaging data are used for both Cepheid discovery and their flux measurement. For the former, a campaign using a filter with a central wavelength in the visual band and ∼11–12 epochs with nonredundant spacings spanning ∼60–100 days is optimal to identify Cepheid variables by their unique light curves and large amplitudes (∼1 mag peak to trough) and accurately measure their periods (Madore & Freedman 1991; Saha et al. 1996; Stetson 1996). Image subtraction may find additional Cepheids (Bonanos et al. 2003), but these objects will be subject to greater photometric biases owing to blends which suppress their amplitudes and chances of discovery in

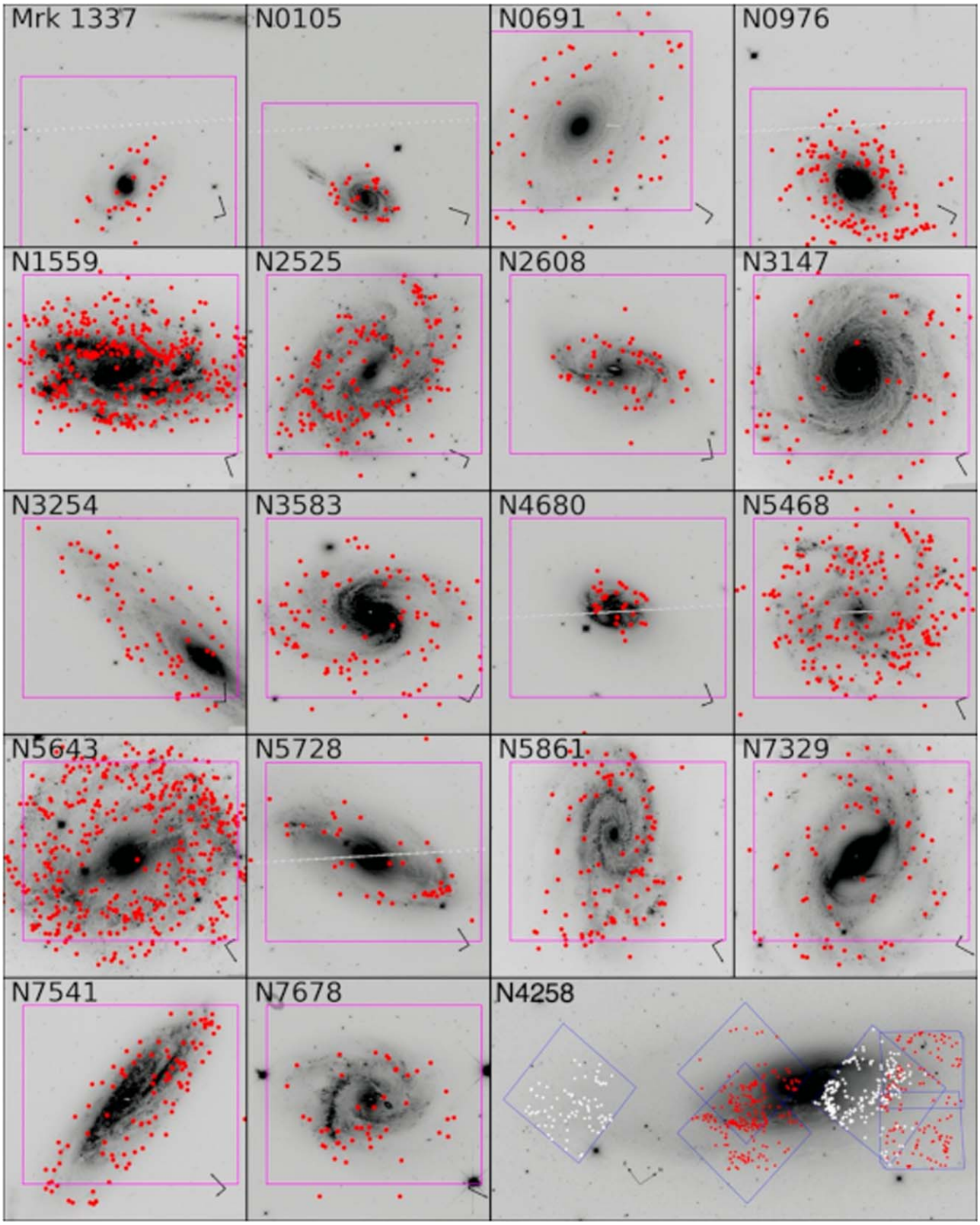

Figure 3. Images of 18 newly observed hosts of 23 SNe Ia and NGC 4258. Each image is of the Cepheid host indicated. The magenta outline shows the $2\farcm7 \times 2\farcm7$ field of the WFC3/NIR observations. Red dots indicate the positions of the Cepheids. Compass indicates north (long axis) and east (short axis). The NGC 4258 image shows Cepheids from Macri et al. (2006) in white and new ones from Yuan et al. (2022) in red.

time-series data (Ferrarese et al. 2000). The data were collected from ∼150 HST orbits with WFC3-NIR and ∼700 orbits obtained for the optical identification (350 from WFC3, 170 from ACS, and 180 from WFPC2), including ∼200 orbits from NICMOS superseded by WFC3-NIR, for a total of ∼1050 HST orbits. Additional observations of Cepheids in the MW and LMC anchors utilized ∼200 orbits or snapshots.

Table 2
WFC3/NIR Cepheids

Field	α (J2000)	δ	ID	P [days]	Color[a]	σ_{col}	$F160W$ [mag]	σ_{tot}	[O/H] [dex]	Note
M101	210.91148	54.357230	54672	6.853	0.91	0.11	24.04	0.44	0.05	HST
M101	210.88464	54.338980	110830	6.873	1.05	0.19	23.36	0.62	0.09	HST

Note.
[a] *F555W* – *F814W*.

Table 3
Properties of NIR *P*–*L* Relations

Galaxy	Number			$\langle P \rangle$ [days]	$\langle$[O/H]$\rangle$[a] [dex]
	FoV	meas.[b]	fit[c]		
M101	311	260	259	15.8	0.10
Mrk 1337	21	20	15	52.9	−0.18
N0105	32	8	8	41.5	−0.13
N0691	31	28	28	46.5	0.09
N0976	57	35	33	40.2	0.02
N1015	26	20	18	52.5	−0.03
N1309	57	53	53	54.1	−0.08
N1365	66	47	45	29.0	−0.14
N1448	90	77	73	35.2	−0.11
N1559	136	110	110	34.4	0.00
N2442	238	177	177	36.0	0.00
N2525	85	73	73	40.4	0.10
N2608	25	22	22	45.4	0.11
N3021	26	16	16	32.2	0.06
N3147	29	28	27	52.3	0.17
N3254	54	48	48	41.5	−0.18
N3370	82	73	73	42.5	−0.12
N3447	116	102	101	36.0	−0.16
N3583	62	54	54	41.6	−0.06
N3972	66	54	52	32.3	0.03
N3982	31	27	27	29.9	−0.14
N4038	38	29	29	53.6	0.03
N4424	17	10	9	31.1	0.06
N4536	45	41	40	36.0	−0.15
N4639	36	30	30	38.7	−0.01
N4680	18	11	11	55.1	−0.06
N5468	118	93	93	54.9	−0.10
N5584	196	167	165	36.8	−0.10
N5643	294	251	251	31.8	0.13
N5728	25	20	20	44.3	0.15
N5861	60	41	41	43.8	0.06
N5917	17	14	14	37.9	−0.30
N7250	30	21	21	37.3	−0.28
N7329	38	31	31	54.6	0.17
N7541	50	33	33	49.1	−0.12
N7678	21	16	16	42.8	0.02
U9391	36	33	33	39.6	−0.22
SN Total	2680	2173	2150	36.5	−0.01
N4258	555	451	443	14.4	−0.10
M31	⋯	55	55	19.1	−0.11
LMC[d]	⋯	342	339	13.3	−0.29
SMC	⋯	145	143	13.3	−0.72
Total All	⋯	3165	3129	⋯	⋯

Notes.
[a] Solar value: 12 + log [O/H] = 8.69, Asplund et al. (2009).
[b] Good-quality measurement, within the allowed color range, period above completeness limit.
[c] After 3.3σ outlier rejection (1.2% of sample).
[d] 69 of these are from HST and 270 from the ground.

Earlier efforts contribute ∼200 orbits of imaging of these hosts: 35 orbits from the HST Key Project, 105 from the SNe Ia Luminosity Program, 36 from Mager et al. (2013), and 50 from Macri et al. (2006), with the remaining ∼800 from SH0ES.

The procedure for identifying Cepheids from time-series optical data in visual or white-light bands has been described extensively (Saha et al. 1996; Stetson 1996; Riess et al. 2005; Macri et al. 2006; Hoffmann et al. 2016); details of the procedures followed for this sample are presented by W. Yuan et al. (2022, in preparation). These procedures utilize the DAO suite of software tools (Stetson 1987, 1994) for crowded-field PSF photometry and are similar to those used previously by the SH0ES team (Hoffmann et al. 2016) and to a large extent by the first generation of HST-based H_0 measurements. Past work has demonstrated that the use of different photometry algorithms yields a largely overlapping list of Cepheids with similar periods and photometry (Ferrarese et al. 1998).

The end result is a set of high-confidence Cepheids, which have passed the selection and quality cuts described by Y22b, with periods and mean colors (*F555W* – *F814W*) measured in the HST WFC3 photometric system. For each Cepheid, we estimate a precise position in the WFC3/NIR *F160W* images using a geometric transformation derived from the optical images using bright and isolated stars, with resulting mean position uncertainties for variables <0.03 pix. The positions of the Cepheids are indicated in Figure 3. Figure 4 shows color images of each Cepheid host.

We present composite light curves of all Cepheids with $10 < P < 80$ days in Figure 5 based on ∼60,000 individual *F555W* or *F350LP* photometric measurements, which identify these as bona fide classical Cepheids with the characteristic sawtooth light-curve shape of fundamental-mode pulsators. There are more subtle Cepheid light-curve features that are not apparent in individual examples at these distances. However, we can leverage the statistical power of the sample to look for these features as a strong validation test of the universality of Cepheids, as presented below.

3.2. Cepheid Validation: The Hertzsprung Progression

The shapes of Cepheid light curves change in subtle but characteristic ways (often referred to as the "Hertzsprung progression"; Hertzsprung 1926) as a function of period, visible only with high signal-to-noise-ratio (S/N) observations. Hertzsprung (1926) noted the characteristic "quick rise and slow decrease" for most periods but also a small range of periods (9–13 days) where the light curves are quite symmetric, and the presence of a "bump" or small local maximum whose phase occurs earlier with increasing period, appearing on the descending phase at $P \approx 6$ days, merging with the main peak at $P \approx 10$ days, and appearing on the rising phase at $P \approx 10$–15 days before disappearing (variables where the bump is apparent are sometimes referred to as "bump" Cepheids). While the general fast rise and slow decline are characteristics

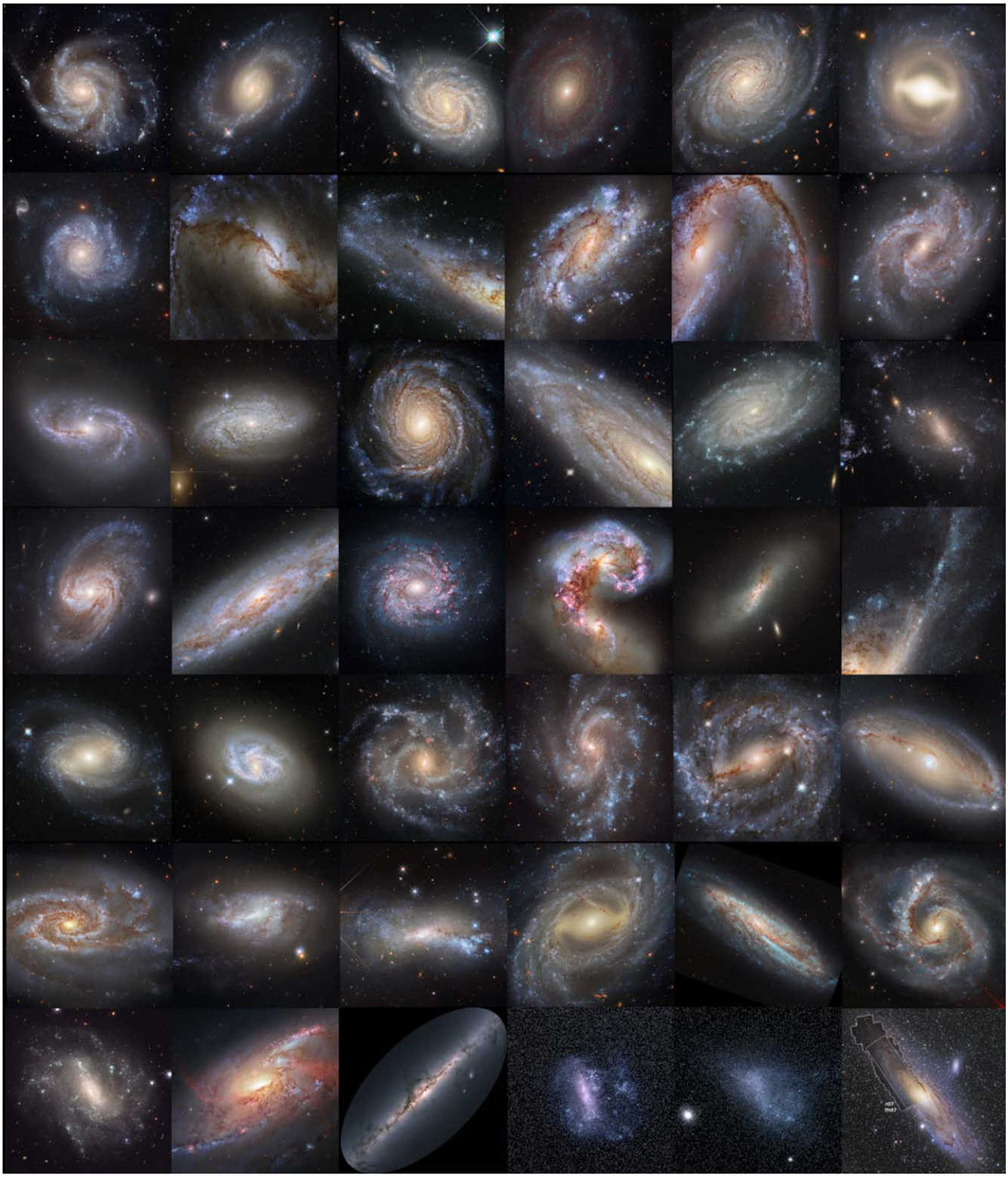

Figure 4. Pseudocolor images of all Cepheid-bearing galaxies analyzed in this work. From top left, 37 hosts of 42 SNe Ia presented in the same order as Table 1. The last row includes our three anchors (NGC 4258, MW, and LMC) and two supporting galaxies (SMC and M31). Galaxies are presented at arbitrary plate scales, though in most cases the panels encompass the entire ACS or WFC3/UVIS field of view. Credits: SN hosts and NGC 4258—ESA Hubble site; MW, LMC, and SMC—ESA Gaia site; M31—STScI.

of a star pulsating in the fundamental mode, the "bump" is understood through modeling to arise from a 2:1 resonance between the fundamental and second overtone (Cox 1993; Bono et al. 2002) and is a unique feature of classical Cepheids. Cepheids with $P > 40$ days occur at the high-mass end of the luminosity function and are thus rare, with the MW hosting only a few examples that show a broadening of the light curve and a decrease in amplitude.

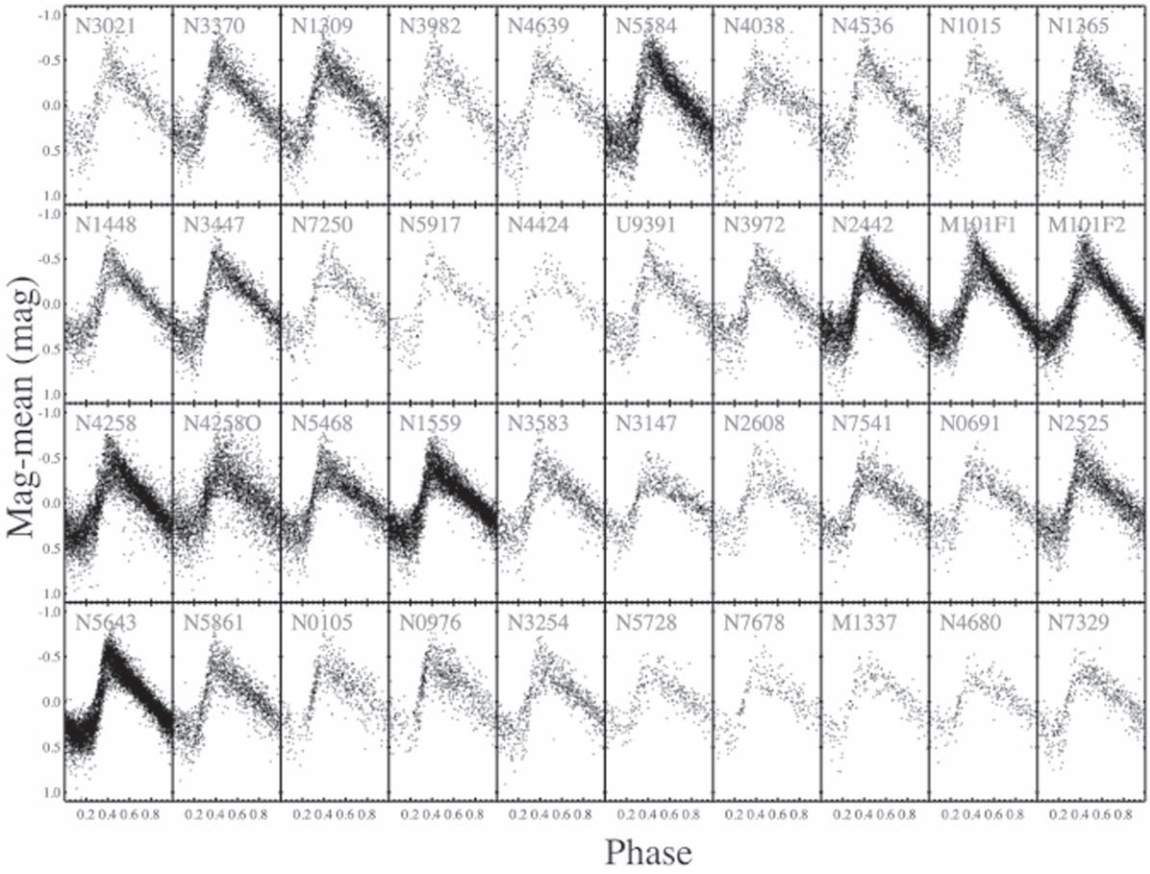

Figure 5. Composite visual (*F*555*W*) or white-light (*F*350*LP*) Cepheid light curves. Each HST Cepheid light curve with $10 < P < 80$ days is plotted after subtracting the mean magnitude and determining the phase of the observation.

By binning all of the extragalactic light curves for the 37 SN Ia hosts and NGC 4258 by period as shown in Figure 6, a striking display of this more subtle light-curve structure emerges. The resonance-induced bump is indicated on the descending phase at $P = 5$–9 days with the MW Cepheid DL Cas ($P = 7.6$ days) shown for comparison, the narrow-peaked symmetric curve at $P = 9$–14 days as in SS CMa ($P = 12.4$ days), and the resonance bump on the rising phase in the $P = 14$–19 days as in XZ Car ($P = 16.6$ days), before transitioning to sawtoothed curves in the next 3 bins ($P = 19$–35 days). Beyond this range, we see the gradual transition to flatter and more sinusoidal curves matching the MW sequence of SV Vul ($P = 45$ days), GY Sge ($P = 52$ days), and S Vul ($P = 69$ days).

While our pipeline selection requires candidate Cepheids to have amplitudes in the characteristic range of 0.2–1 mag, we impose no requirement regarding the presence of these more subtle light-curve features, which are not detectable at the data quality typical of single extragalactic Cepheids (Hoffmann et al. 2016). The presence of these features in aggregate light curves at the appropriate periods offers an additional level of scrutiny and validation of these sources as classical Cepheids, as well as the means for a more detailed comparison to those in the MW. The impressive similarity of these light-curve sequences demonstrates the universality of the physics producing these pulsating standard candles and provides an important test of their consistency along the distance ladder. The appearance of the Hertzsprung progression among the extragalactic sample demonstrates these objects are bona fide Cepheid variables.

3.3. NIR Photometry and Validation in NGC 4258

Following the same procedures described by R16, we NIR photometry of the Cepheids using their positions derived and fixed from the higher-resolution optical data. The NIR images of all Cepheids were reprocessed for this analysis using revised calibrations by the STScI pipeline (including new flat fields and geometric-distortion tables), and we were able to predict the expected NIR positions of the Cepheids found in the optical data with enhanced precision compared to our previous reduction (Hoffmann et al. 2016). Additional sources found in the NIR images are fit simultaneously with the Cepheids using an empirical model of the PSF derived from the solar-analog star P330E. The background calibration[13] and photometric uncertainties are determined from the retrieval of 100 locally placed artificial stars per Cepheid as described by R11 and R16.

In Figure 7 we show the distributions of the artificial-star measurements divided by their standard deviation, which demonstrates that they are well approximated by Gaussian distributions in magnitude space around their mean out to 3σ, a consequence of the log-normal distribution of underlying surface brightness fluctuations (primarily red giants, with larger and brighter collections increasingly rare) in the NIR. The mean difference between the mean and median of the artificial-star distributions, which vanishes for a true Gaussian, is 0.03σ. Averaged across all SN hosts, the difference between the mean

[13] Sometimes called a "crowding correction" because it adds the mean level of randomly superimposed sources to the initial background estimate derived from the unresolved or constant sky level.

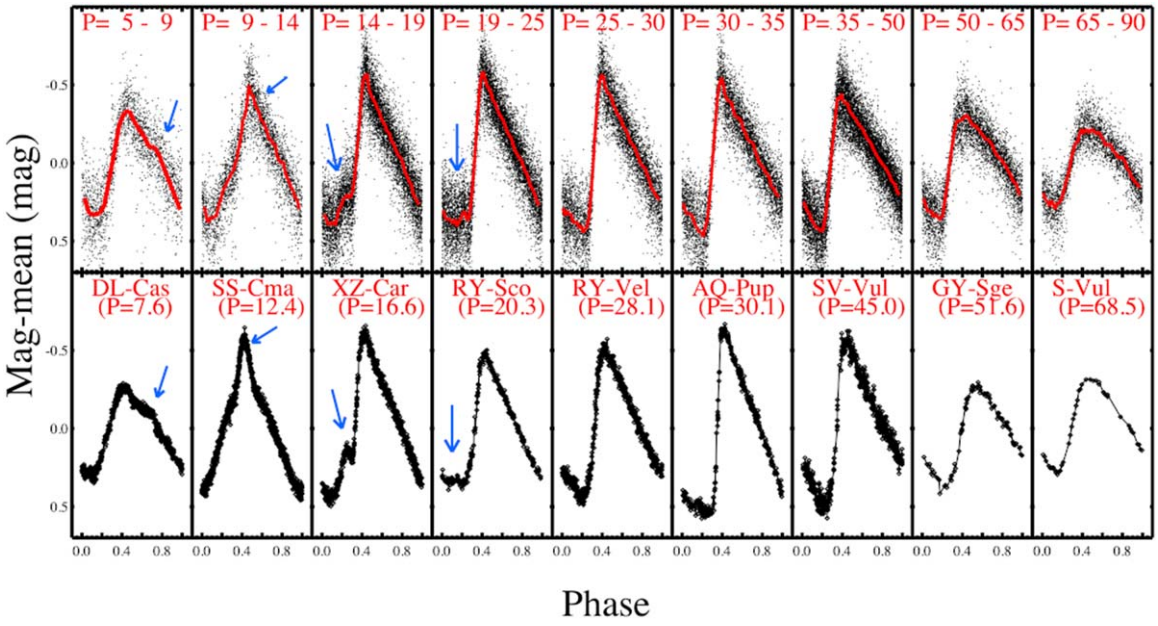

Figure 6. Composite visual (*F555W*) or white-light (*F350LP*) Cepheid light curves binned by period (upper) and compared with individual MW Cepheids near the middle period of the bin. The "Hertzsprung progression" (relation between light-curve shape and period) is apparent, including subtle features like the progression in phase of a resonance "bump" between the second overtone and fundamental pulsation for $P < 20$ days. The red line is a cubic spline constrained by the averages of bins in phase.

and median of the magnitude distribution drops to 0.01σ or 4 mmag, showing no apparent correlation between its third moment (skewness) and the host distance. The distribution for the geometric calibrator, NGC 4258, appears similar to that of the others, with a difference between the mean and median of 0.01 mag. These measurements justify the use of Gaussian statistics in magnitude space in the calculations to follow.

The accuracy of the background estimates owes to Cepheids being randomly superimposed on scenes, a consequence of our perspective whose local levels can be measured statistically. A caveat to this approach would be the presence of associated flux (colocated with the Cepheid), which becomes important only if it is then resolved for nearby Cepheids but not for distant ones. The level of such "associated flux" has been measured statistically from hundreds of Cepheids in M31 by Anderson & Riess (2018) to be ~7 mmag at distances beyond a few Mpc, and is due to associated open clusters that would not be resolved at those distances. This term is explicitly included here in the background estimates in order to compensate for this effect. An additional and direct consequence of a potential miscalibration of the background, independent of the Cepheid mean flux, would be a change in the apparent light-curve amplitude. Riess et al. (2020) determined that the NIR amplitudes of Cepheids in three SN hosts are fully consistent with those in the Milky Way, yielding an independent upper limit of 0.03 mag for the possible misestimation of the background if representative. The sensitivity of H_0 to the unresolved background can be further mitigated by the use of a distant anchor with similar background as the SN Ia hosts such as NGC 4258; this will be addressed in Section 4.

In Figure 8 we provide a strong test of the background estimation by comparing the Cepheid photometry in dense (inner) and sparse (outer) regions of NGC 4258. Because the Cepheids in both fields are at the same distance from us and the metallicity gradient in NGC 4258 is small (Bresolin et al. 2016), an apparent difference of the dereddened magnitudes would be a consequence of misestimating the background in the dense field. The difference in intercept (i.e., distance) is 0.01 mag and well within the indicated errors of the means, demonstrating that Cepheid PSF photometry is accurate in the presence of the same level of crowded backgrounds seen in the SN hosts.

In Appendix B we provide an independent validation of the PSF photometry using aperture photometry, a method that is accurate when the background is measured from the mean of pixels in concentric annuli and uncomplicated to apply, albeit less precise than the standard approach of using PSFs to model photometry. This test validates the mean PSF photometry in SN hosts to $\sigma \approx 0.02$ mag. In that appendix we further perform an additional "null test" of the background estimates by regressing them against the distance-ladder fit residuals, finding a dependence of 0.010 ± 0.014 mag per magnitude of source background (in the sense of overestimating the background but with no significance).

In Section 7.1 we review multiple tests of Cepheid PSF photometry in addition to six strong tests of background estimates in the presence of crowded backgrounds, all of which indicate that the Cepheid measurements are accurate.

3.4. Dereddened Magnitudes

The SH0ES program uses observations in HST filters at known Cepheid phases in optical (*F555W*, *F814W*) and NIR (*F160W*) bands to correct for the effects of interstellar dust and the finite width in temperature of the Cepheid instability strip. We employ NIR "Wesenheit" magnitudes (Madore 1982) to deredden Cepheids throughout, defined as

$$m_H^W = m_H - R\,(m_V - m_I), \tag{7}$$

where $m_H = F160W$, $m_V = F555W$, $m_I = F814W$ and $V - I = m_V - m_I$ in the HST system, and $R \equiv A_H/(A_V - A_I)$. Wesenheit magnitudes are not conventional magnitudes, which compare the brightness of one star to another; rather, they are used to compare the brightness of one standardized candle to another

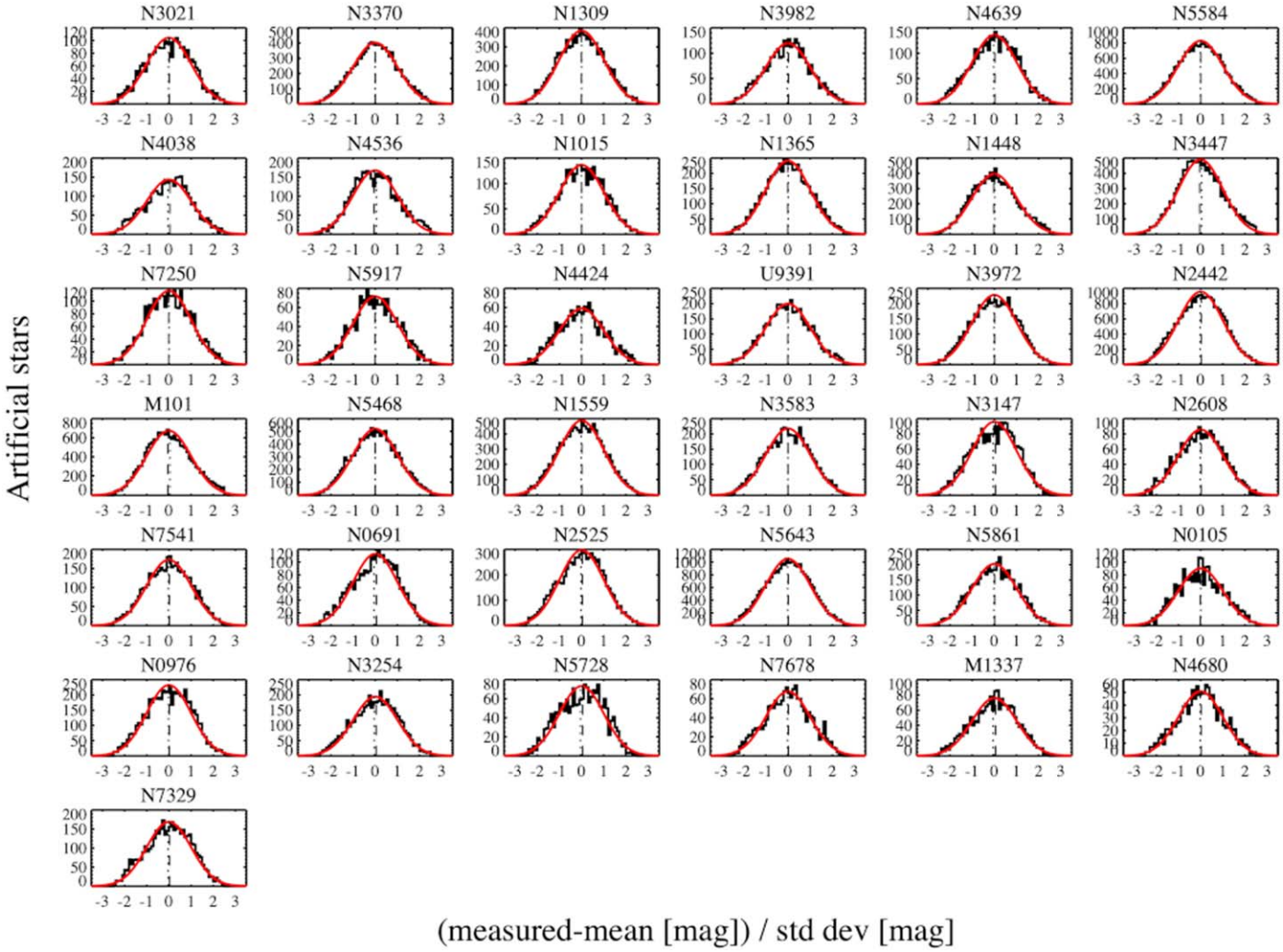

Figure 7. Distributions of recovered artificial stars in each host. For each real Cepheid, 100 artificial stars of the expected Cepheid magnitude (based on its period and the *P–L* relation) are added to the host images within an annulus with an outer radius of 2.″4 and measured following the same procedures as for the Cepheids to produce a distribution in magnitudes (log flux). The distributions for each host are combined after subtracting their mean and dividing by their standard deviation. They compare well to a Gaussian with a difference in median and mean (dashed and dotted lines) of $\sim 0.03\sigma$.

through the removal of their unequal extinction, as reviewed in Appendix D. While the value of R obtained from well-characterized extinction laws for these bands is $\sim$0.4, we note that the correlation between Cepheid intrinsic color and luminosity at a fixed period has the same sense as extinction (cooler is fainter), and is similar in size with an intrinsic value of $\sim$0.6 (as discussed in Section 6.2). Therefore, the value of R derived for extinction effectively also reduces the intrinsic scatter caused by the breadth of the instability strip. We analyze the sensitivity of H_0 to values of R derived from different extinction laws in Section 6.3. In Appendix D, we discuss pitfalls associated with varying R in Equation (7) between galaxies if the intrinsic color is not first subtracted from the observed color[14] (Follin & Knox 2018; Mortsell et al. 2021).

To avoid a magnitude bias, we include only Cepheids with periods above the completeness limit of detection in our primary fit for each host (Y22b). The measurements of Cepheids in SN Ia hosts are provided in Table 2, while Table 3 summarizes the properties of the resulting NIR *P–L* relations. We identify a number of improvements realized here since our previous Cepheid measurements in SN hosts $\sim$6 yr ago (R16, H16) and for NGC 4258, $\sim$16 yr ago (Macri et al. 2006).

1. Sky annuli: The size of the annulus used to estimate the level of the sky value of the region around the Cepheid (after subtracting all detected sources) in $F160W$ was reduced in size to inner and outer radii of 0.″24 and 0.″8 (from 0.″96 and 1.″44 in R16) based on extensive simulations. This determination of the sky level is more precise, although it requires taking into account the contribution to the sky from the wings of the PSF; the resulting offset of 0.008 mag is robustly determined from bright stars and corrected in the final photometry. The random sky error is propagated to the photometry error through the sampling of sky values in the artificial-star analyses).

[14] This might be pursued to allow the extinction law to vary in every host, but if the intrinsic color is not first subtracted, it has the unintended consequence of producing a large variation in the luminosity of the standard candle itself, which is unrelated to dust, is inconsistent with the premise of a distance ladder where stars (once standardized) have luminosities independent of the rung they live on, and most importantly is not supported by the data as shown in Appendix D.

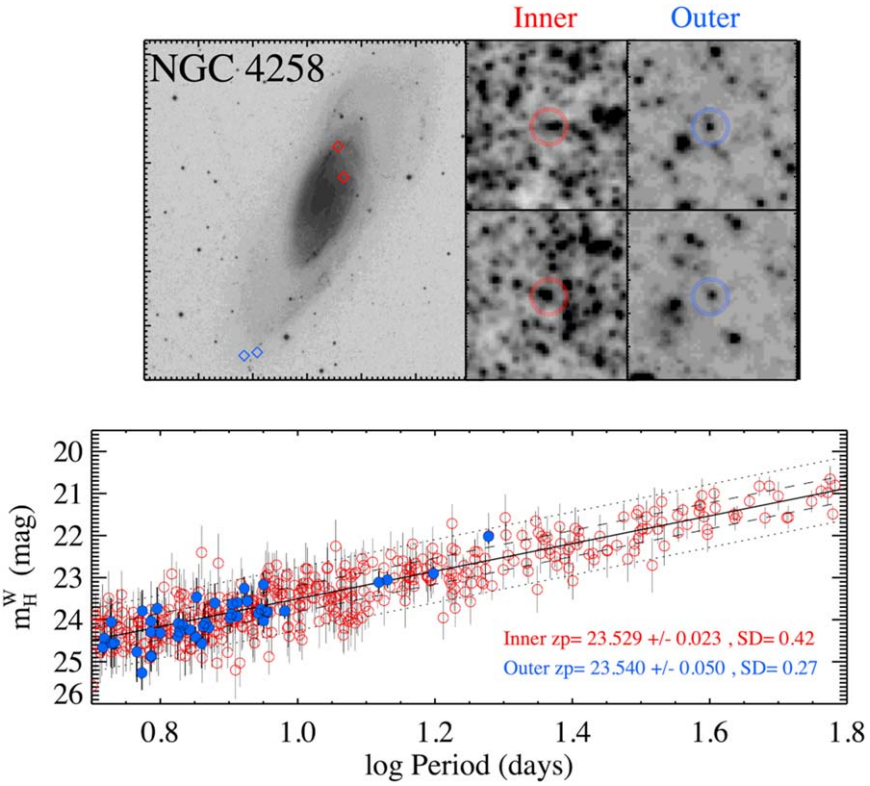

Figure 8. Comparison of Cepheids measured in a dense (inner) field (in red) and sparse (outer) field (in blue) of NGC 4258. Because these Cepheids are at the same distance, the comparison shows the accuracy of the background estimates, which differ in the mean over the same sampled range, $0.7 < \log(P/\mathrm{day}) < 1.2$, by 0.45 mag (relative to the Cepheids) yet yield a consistent intercept with $\sigma \approx 0.05$ mag. The difference in metallicity between the samples, $\Delta[\mathrm{O/H}] = 0.08$ dex, corresponds to a difference of 0.02 mag, smaller than the precision of this comparison.

2. Determination of the background and covariance: As in R16, artificial stars are added in the *F160W* images in the vicinity of each Cepheid at the apparent magnitude expected from its period and trial fits of the *P*–*L* relation. The difference between the input and recovered magnitudes is used to refine the initial estimate of the sky background and revise the Cepheid photometry. This process is necessarily iterative as the *P*–*L* relation used to predict the magnitude of a given Cepheid from its period is determined after correcting the photometry based on the retrieval of the artificial stars. Thanks to greater processing power, we increased the number of iterations since R16 and found that in some cases the iterations in our previous work were not adequate (as they had not fully converged). The new process fully converges with improved determination of the trial intercepts and slopes, and we use the final iteration to estimate a systematic uncertainty of 20% of the background (in units of the Cepheid magnitude correction) as the covariance "error floor" of any pair of Cepheids, *j*, *k* in the same *i*th host, given by

$$C_{i,j,k,\mathrm{bkgd}} = 0.2^2(\mathrm{bkgd}_{i,j})(\mathrm{bkgd}_{i,k}). \quad (8)$$

This provides a systematic uncertainty in the range of 0.03–0.06 mag for all Cepheids within each host, with the term "bkgd" representing the change in Cepheid magnitude due to the addition of the mean level of the crowded background from unresolved sources derived from the artificial stars. The artificial-star magnitudes are determined from the trial *P*–*L* relations of each host independently from every other host so there is no source of background covariance between different hosts.

3. Reference Files: The photometry benefits from the latest STScI data pipelines, including new flat fields (with better "blob" mapping), bias frames, long-history dark frames, pixel-based CTE corrections for optical data, and geometric-distortion corrections, yielding improved alignment between the optical and NIR frames.
4. Count Rate Nonlinearity (CRNL): We adopt a new calibration of the WFC3/NIR CRNL (Riess et al. 2019b). By convention, this is applied to Cepheids in anchors between their flux and the background level of Cepheids in SN hosts.
5. M101 long-period Cepheids: A reexamination of the Cepheids in M101, together with simulations, revealed that the baseline of the original monitoring campaign (carried out in 2006 and reported by Mager et al. 2013) was too short to provide reliable periods for apparent Cepheids with $P > 35$ days (equivalent to $1.2\times$ the time span of the observations). Two additional epochs obtained 7 yr later and separated by a week are insufficient to resolve the issue because at these periods the two epochs provide effectively only one phase measurement, and the prevailing period uncertainty of >0.5 days makes the phasing of the two sets unreliable. For that reason we exclude M101 Cepheids with $P > 35$ days (about 10% of the sample used by R16).
6. We correct Cepheid periods to rest-frame values owing to $(1+z)$ time dilation (Anderson 2019), a small (given our typical $z \approx 0.005$) but one-sided effect.

We note that the Cepheid color measurements, $V-I$, employed in Equation (7) to determine the baseline value of

H_0 here and in R16 (as well as most variants) are relatively insensitive to the previously noted improvements to the calibration of the Cepheid optical measurements realized in the last 6 or 16 years and included in Y22a and Y22b. These include the use of artificial stars to estimate the crowded backgrounds, revised flat fields, archival dark frames, updated geometric-distortion maps to rectify frames, and pixel-based CTE corrections, most of which were not implemented by H16 or by Macri et al. (2006), from which the N4258 data in H16 were derived. These improvements cancel to first order in the difference, $V-I$, as do changes in CCD sensitivity, which can decline on orbit or change abruptly when the electronics are refurbished as occurred for ACS in 2009. However, the use of optical Wesenheit magnitudes to determine H_0 without bias (as we explore in Section 6.13) requires fully calibrated optical magnitudes rather than only accurate optical colors (as in H16 and provided in R16). The fully calibrated optical magnitudes available in Y22a and Y22b are suitable for this purpose.

We do not attempt to quantify the impact of each of these improvements individually; however, in the aggregate, matching Cepheids within $1''$ of those in R16, the net change to Cepheid photometry in *F160W* is that 63% (37%) of Cepheids are fainter (brighter). We provide additional details of this comparison in Appendix B.

3.5. Cepheid Metallicities

As in R16 and H16, we measured radial gradients of the strong-line abundance ratios (R_{23}) of oxygen to hydrogen in H II regions in the Cepheid hosts. The optical spectra were obtained with the Low-Resolution Imaging Spectrometer (LRIS; Oke et al. 1995)) on the Keck I 10 m telescope on Maunakea, Hawaii. See R16 and H16 for details regarding the observations and data reduction.

We define the metallicity of each Cepheid to be the value of this linear function at its galactocentric radius. We have revised the calibration of these strong-line abundance measurements relative to those from Zaritsky et al. (1994, hereafter Z94) used by R16, taking advantage of more recent calibrations between R_{23} and the metallicity $12+\log[\mathrm{O/H}]$. We adopt the average of nine recent literature calibrations, with the transformations between the Z94 system and the newer ones given by Teimoorinia et al. (2021). Further details about the metallicity measurements and their uncertainties are given in Appendix C.

We also use direct abundance measurements derived from high-resolution spectra of Cepheids in the MW, LMC, and Small Magellanic Cloud (SMC). In Appendix C, we evaluate the consistency of the direct and radial (strong-line) abundance measures by comparing H II regions and Cepheid spectral abundances in the MW. While we use the average of nine strong-line abundance calibrations for our baseline results, we select the Pettini & Pagel (2004, hereafter PP04) calibration for comparison to the mean, as it is known to provide a good match to measured extragalactic stellar abundances (Bresolin et al. 2016, and F. Bresolin, 2021, private communication), and to estimate the uncertainty of the strong-line metallicity estimates. Specifically, we propagate a systematic uncertainty in the strong-line abundance scale as well as its covariance by including in our fit covariance matrix (given in Section 3) the product of the difference between the mean calibration and the PP04 calibration for the ith and jth Cepheids,

$$C_{i,j,\mathrm{syst}} = Z_W^2([\mathrm{O/H}]_{i,\mathrm{avg}} - [\mathrm{O/H}]_{i,\mathrm{PP04}})([\mathrm{O/H}]_{j,\mathrm{avg}} - [\mathrm{O/H}]_{j,\mathrm{PP04}}). \quad (9)$$

This approach requires an iteration to use the same value of Z_W in Equation (9) that is determined from the optimization of the global χ^2. The mean difference of $\sim$0.05 dex between the average and PP04 scale represents a systematic uncertainty in the abundance scale, propagated here in the covariance matrix, which is consistent with the empirical assessment in Appendix C

4. Anchor Constraints and Ancillary Data

The strongest constraints on the Cepheid *P–L* and *P–L–C* relations come, not surprisingly, from the nearest star-forming galaxies whose samples of Cepheids have better temporal sampling, higher resolution, wider wavelength coverage, and far greater S/N than we can expect to achieve from the distant Cepheids presented above that occupy the second rung of the distance ladder. An accurate and precise determination of H_0 requires leveraging such data to empirically constrain Cepheid properties. Ignoring such data naturally reduces the precision in H_0 and as a consequence the significance of any Hubble tension, but this is not a reasonable approach to determine the source of any tension. We adopt uncertainty estimates from the indicated, external sources as provided and without alteration, and will test their internal consistency within the distance ladder in the following sections.

Here we describe the data we use, in addition to those presented in Section 3. Some of these data, namely those described in Section 4.1 through 4.3, are anchor constraints—they provide direct information on the zero-point of the Cepheid *P–L* relation. For this purpose, we require (a) that Cepheids be observed directly in our three-filter HST photometric system, most importantly in the NIR with the WFC3 *F160W* filter, and (b) that the distance determination, to individual Cepheids or to their host, be purely geometric. Additionally, we make use of several data sets, to which we refer as ancillary data, which do not directly constrain the zero-point of the *P–L* relation but provide useful information on other characteristics of Cepheids; these include Cepheid measurements in nearby hosts with high-precision photometry (in similar filters but not in our standard HST system), Cepheids in hosts without a precision geometric distance, and information on SNe Ia in the Hubble flow.

4.1. Milky Way Cepheids

Trigonometric parallaxes to MW Cepheids offer a direct source of geometric calibration of their luminosities. We employ two samples with precise parallaxes and fluxes measured on the same HST system as the extragalactic Cepheids and with direct spectroscopic metallicity measurements. The first is a sample of eight MW Cepheids from Riess et al. (2018a) with parallaxes measured with HST WFC3 spatial scanning. The other contains 75 MW Cepheids with parallaxes from Gaia EDR3 as given by Riess et al. (2021). We do not use the sample from Benedict et al. (2007) that was part of R16 because their fluxes are too bright ($0 < m_H < 3$ mag) to be measured directly with HST nor have they been measured with good accuracy from the ground; the parallax sample from Gaia EDR3 provides superior information on MW Cepheids.

Because the distance uncertainties are measured (and Gaussian) in parallax space jointly and are not insignificant (>5%), the

Table 4
Ancillary Cepheid Data

Sample	Reference	N	$\langle P \rangle$ [days]	$\langle$[O/H]$\rangle$ [dex]	Photometry	Selection	Notes
MW Gaia EDR3 +HST	Riess et al. (2021)	66	12.5	0.13	HSTm_H, m_V, m_I	see ref.	$M^W_{H,1,\mathrm{Gaia}} = -5.903$, $\sigma_{\mathrm{Gaia}} = 0.024$[a], $Z_W = -0.20 \pm 0.12$
MW WFC3 SS	Riess et al. (2018a)	8	22.6	0.05	HST m_H, m_V, m_I	see ref.	$M^W_{H,1,\mathrm{HST}} = -5.810$, $\sigma_{\mathrm{HST}} = 0.054$[a]
LMC HST	Riess et al. (2019b)	70	16.0	−0.29[b]	HST m_H, m_V, m_I	see ref.	
LMC ground	Macri et al. (2015)	272	12.6	−0.29[b]	ground m_H, m_V, m_I	$P > 5$ days	m_H transformed to 2MASS[5]
SMC ground	Kato et al. (2007)	145	9.9	−0.70[b]	ground m_H, m_V, m_I	$P > 5$ days, $r < 0.6°$	m_H transformed to 2MASS[5]
M31 SH0ES	Li et al. (2021)	55	19.1	−0.11	HST m_H, m_V, m_I	$P > 4$ d	
M31 PHAT	Kodric et al. (2018)	463	10.5	0.12	HST m_H, m_J	$P > 4$ days	m_H, m_J transformed to m_H, m_V, m_I

Notes.
[a] Measured following Riess et al. (2021) with global fit P–L parameters.
[b] From Romaniello et al. (2021) and Romaniello et al. (2008) and $\langle$[Fe/H]$\rangle \sim \langle$[O/H]$\rangle - 0.06$.

direct transformation of parallax to distance modulus (i.e., distance to magnitudes) would yield a bias (often called the "Lutz–Kelker" bias, following Lutz & Kelker 1973). One can compensate for it by estimating approximate statistical corrections between the parallax and distance moduli for individual Cepheids, assuming the form of their spatial distribution in the MW using Bayesian inference (Bailer-Jones et al. 2021). However, it is simpler and more reliable to analyze the MW Cepheid data directly in parallax space, in order to retain the Gaussian parallax errors, and to derive their joint constraint on the Cepheid absolute-magnitude zero-point ($M^W_{H,1}$) following Riess et al. (2021), or similarly through the use of "astrometric-based luminosity" (Arenou & Luri 1999).

Given that the constraints from the parallaxes across a Cepheid sample are related by the P–L relation, the combined constraint is evaluated in parallax space and the resulting error in the mean is greatly reduced (for our two samples to 1% and 3%) so that the resulting mean constraint on the zero-point in magnitudes can then be approximated as Gaussian to better than 0.1%. We note that the 1% calibration from Gaia EDR3 includes marginalization over the Gaia parallax offset term as described in Riess et al. (2021) and that the measured $\sigma = 6$ μas uncertainty in the offset is the source of 0.9% uncertainty for the MW Cepheid sample due to its mean 650 μas parallax. These constraints are given in Table 4 for the P–L relation determined here. The constraints are not identical to those provided in Riess et al. (2021) because they employ the same P–L parameters, b and Z_W, that optimizes the global χ^2 for all Cepheids, not just MW Cepheids. Riess et al. (2021) showed that the global P–L parameters, including slope and metallicity dependence, are good fits for the MW Cepheids on their own.

4.2. Cepheids in NGC 4258

The R16 analysis used 143 Cepheids observed with HST in the maser host NGC 4258 with a geometric distance from Reid et al. (2019). With new HST imaging campaigns in four fields shown in Figure 3, we have more than tripled that sample to 443 Cepheids; details of their identification are given by Yuan et al. (2022). These Cepheids are included in Table 2.

4.3. Cepheids in the LMC

We include in the joint constraint the sample of 70 Cepheids in the LMC measured on the same HST three-band photometric system given by R19. These have also been corrected with the best-fit planar geometry model of the LMC to be at a single distance as described by R19. The distance to the center of the LMC is given by Pietrzyński et al. (2019) from a sample of 20 DEBs as $\mu_0 = 18.477 \pm 0.0263$ mag. Romaniello et al. (2021) has obtained spectra for 68 of these variables, demonstrating that they are consistent within the errors of having a single, common abundance of [Fe/H] $= -0.40$ dex. Sixty percent of the variables have sufficiently measured lines to further provide a mean spectroscopic abundance of [O/H] $= -0.29 \pm 0.02$ dex, which we adopt for the full LMC sample. LMC Cepheids are corrected for the WFC3 CRNL across the 5 dex between the Cepheid flux and the SN-host background (Riess et al. 2019a). The LMC Cepheids from HST also set the intrinsic scatter in m^W_H owing to the finite width of the instability strip to be 0.07 mag, which we include for all Cepheids.

The LMC also provides ancillary data, as defined above, that can be used independently of their inferred zero-point to refine the characterization of the P–L relation. For this purpose we use the ground-based sample of 785 LMC Cepheids from Macri et al. (2015) with m_H, m_V, and m_I photometry limited to 270 with $P > 5$ days. This photometry has been transformed to the HST system as described by Equations (10)–(12) in R16; however, we assign a common, systematic uncertainty of $\sigma_{\mathrm{grnd}} = 0.10$ mag to the transformed magnitudes, hence the simultaneous constraint takes the form $0 = \Delta zp \pm \sigma_{\mathrm{grnd}}$, where Δzp is a parameter describing the difference between the ground and HST zero-points to account for possible systematics associated with ground-based observations. Because the assigned systematic uncertainty is a factor of ∼10 larger than the mean of the smaller HST LMC sample (which has no such photometric system difference), the consequence is that the ground-based sample has negligible (100× less) weight in the tie to the LMC distance; therefore, the ground sample only helps constrain the slope of the P–L relation, still an invaluable contribution.

4.4. Cepheids in the SMC

An important development in support of the distance ladder is the recent measurement of a geometric distance to the SMC from 10 DEBs (Graczyk et al. 2020) with a precision of better than 2%. While this precision and method would make the SMC suitable as an additional anchor, two issues present limitations. The first is the considerable line-of-sight depth of the SMC, which can cause a 10% dispersion in star distances

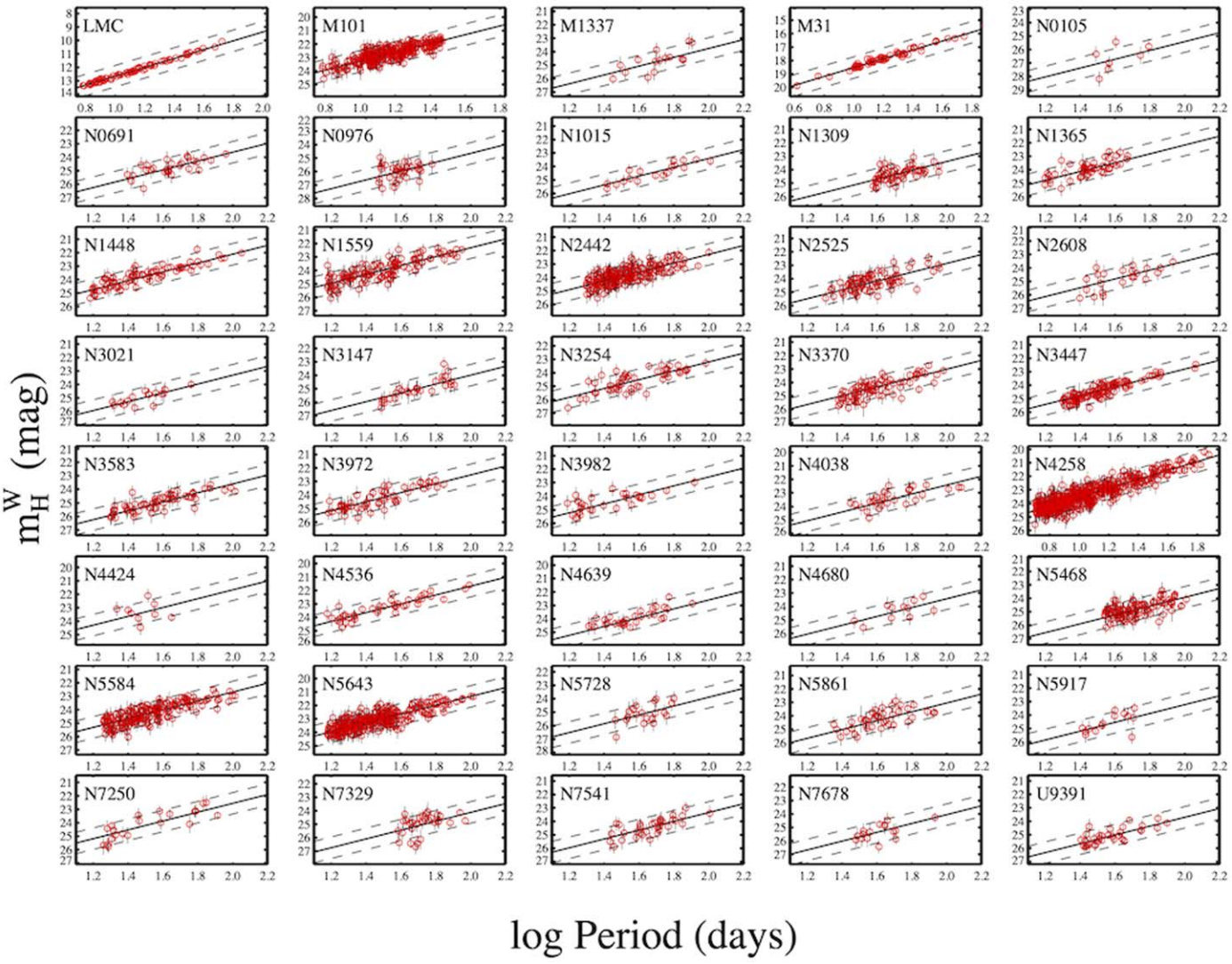

Figure 9. HST NIR Wesenheit Cepheid *P–L* relations. The Cepheid magnitudes are shown for 37 SN Ia hosts, M31, and 2 of the 3 possible distance-scale anchors (LMC and NGC 4258). The uniformity of the photometry and metallicity reduces systematic errors along the distance ladder. A single slope is shown and used for the baseline, but we also allow for a break (two slopes) as well as limited period ranges in some analysis variants.

across the full SMC structure and an offset between the DEBs and the Cepheids in the SMC. Here we follow the approach of Breuval et al. (2021) to make use of a sample of SMC Cepheids by (1) correcting for the depth by adopting the same geometry of the SMC as used by Graczyk et al. (2020) to characterize the DEBs and (2) limiting the Cepheids to the inner core of the SMC, a radius of $0\overset{\circ}{.}6$. This combination yields a Cepheid sample that, based on the SMC geometric model, is offset in depth by 2 mmag from the mean DEB distance (with an uncertainty of a small fraction of this) and has a modeled dispersion in depth of $\sigma = 0.024$ mag, while still providing a sample of 145 Cepheids (with $P > 5$ days) that can yield valuable constraints on the Cepheid metallicity term. Earlier studies (Gieren et al. 2018) were unable to employ the DEB-based distance from Graczyk et al. (2020) and did not focus on the core of the SMC (Wielgorski et al. 2017).

The other limitation is the lack of HST photometry for SMC Cepheids. Elsewhere we have been able to negate zero-point uncertainties in the distance scale by using Cepheids observed with the same instruments. However, the SMC Cepheids can still provide a powerful constraint on the Cepheid metallicity term, which is well-constrained using only differential, cross-calibrated ground-based photometry and the differential DEB distance measurement between the SMC and LMC, the latter of which is known better than the simple difference in the individual DEB cloud distances.

As discussed by Graczyk et al. (2020), most of the uncertainty in the DEB distance estimates to the LMC and SMC is systematic and propagates from the uncertainty in the surface brightness versus color calibration of red giants, the zero-point of the *V*-band and *K*-band photometry, and the uncertainty in the extinction law. Because the aforementioned SMC and LMC DEB measurements utilized the same relations, observational setup, and reduction methodologies, their differential distance from DEBs as given by Graczyk et al. (2020) is far better constrained to 0.500 ± 0.017 mag, independent of the absolute calibrations and their uncertainties in the DEB method. The use of consistently calibrated LMC and SMC ground photometry and this differential distance, even without reference to the absolute DEB distances, helps constrain the metallicity term as we will show in Section 6.2. Here we use the 2MASS Point Source Catalog (Cutri et al. 2003) to produce

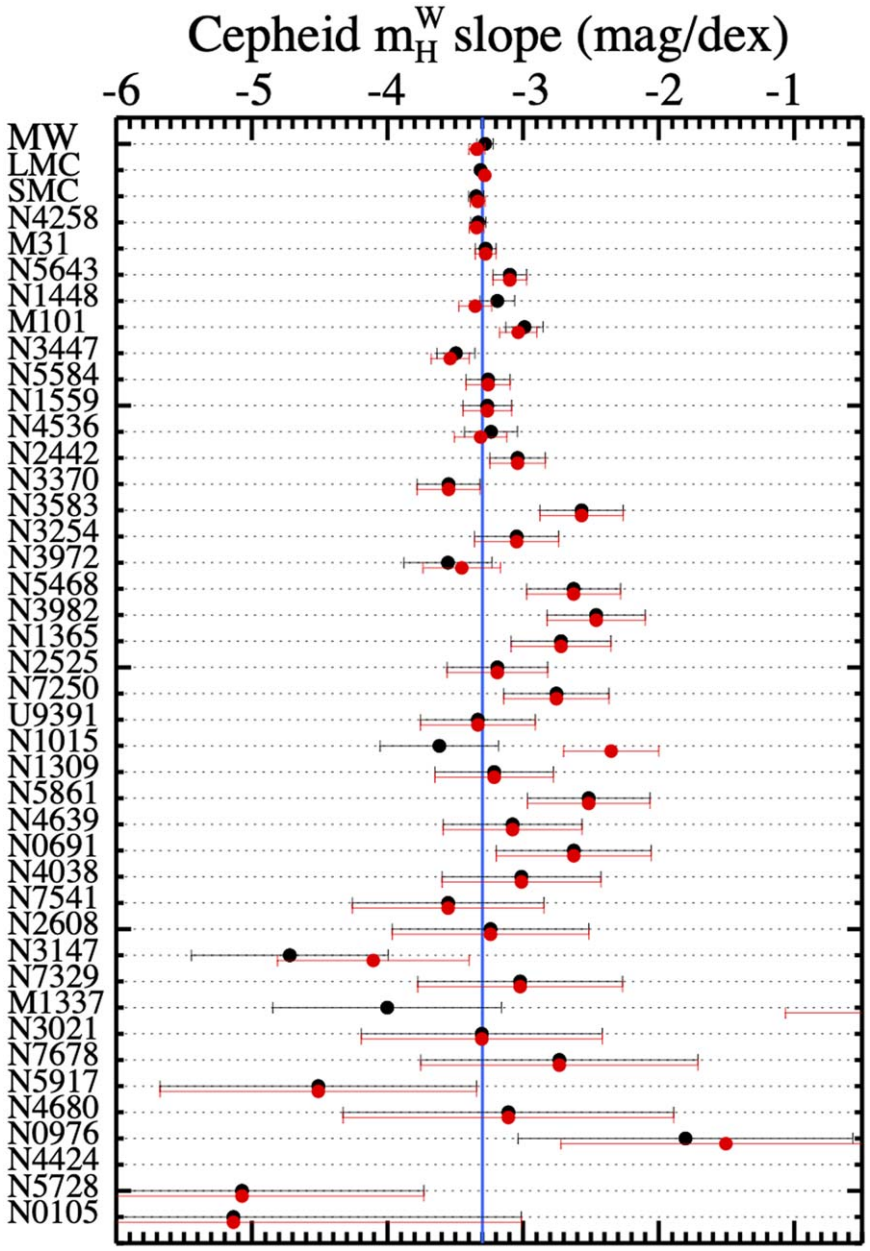

Figure 10. Independent determinations of the slope of the *P–L* relation in each host plotted in order of precision. Black (red) points are after (before) outlier rejection. We find that the slopes are consistent with a single value of −3.27 to −3.30 as indicated by the blue line. The most metal poor (SMC) and metal rich (MW) also have consistent slopes. Error bars are for the fit error only and do not include uneven period sampling or background covariance. The large change in slope for Mrk 1337 was due to the relatively small sample size and rejection of a few Cepheids far from its center.

a consistent calibration[15] of the *H*-band LMC Cepheid data from Macri et al. (2015) and the *H*-band SMC data from Kato et al. (2007), and m_V and m_I data from OGLE III (Soszynski et al. 2008), with data provided in Table 2. The precision of these transformations allows us to fully leverage the constraint on the distance difference from Graczyk et al. (2020).

We take the mean SMC Cepheid metallicity to be $\Delta[\mathrm{Fe/H}] = -0.41$ dex relative to the LMC or $[\mathrm{O/H}] = -0.70$ dex (Romaniello et al. 2021, 2008) ($[\mathrm{Fe/H}] = -0.76$). We add to this a systematic uncertainty (common covariance) in the mean LMC and SMC metallicity of 0.05 dex, as in Equation (9), each, separately, and relative to other hosts.

4.5. Cepheids in M31

In Li et al. (2021) we presented measurements of 55 Cepheids in M31 using the same three-filter HST system adopted elsewhere. Owing to their low dispersion and broad range in period, these Cepheids provide additional constraints on the slope of the *P–L* relation, independent of any prior knowledge of the distance to M31; this is how we employ these data for our baseline analysis. Given the M31 metallicity gradient of -0.023 dex kpc^{-1} (Zurita & Bresolin 2012), these Cepheids span only a narrow range of abundances with a mean $[\mathrm{O/H}] = -0.1$ dex and a standard deviation of 0.05 dex.

[15] Macri et al. (2015): LMC, $H_{2\mathrm{MASS}} = m_H + 0.0116 - 0.0054(J-K-0.4) - 0.0189(J-K-0.4)^2$, rms = 0.038 mag ($N \approx 34{,}000$). Kato et al. (2007): SMC, $H_{2m\mathrm{ASS}} = m_H - 0.0246 - 0.0228(J-K-0.4) + 0.0106(J-K-0.4)^2$, rms = 0.033 mag ($N \approx 14{,}000$).

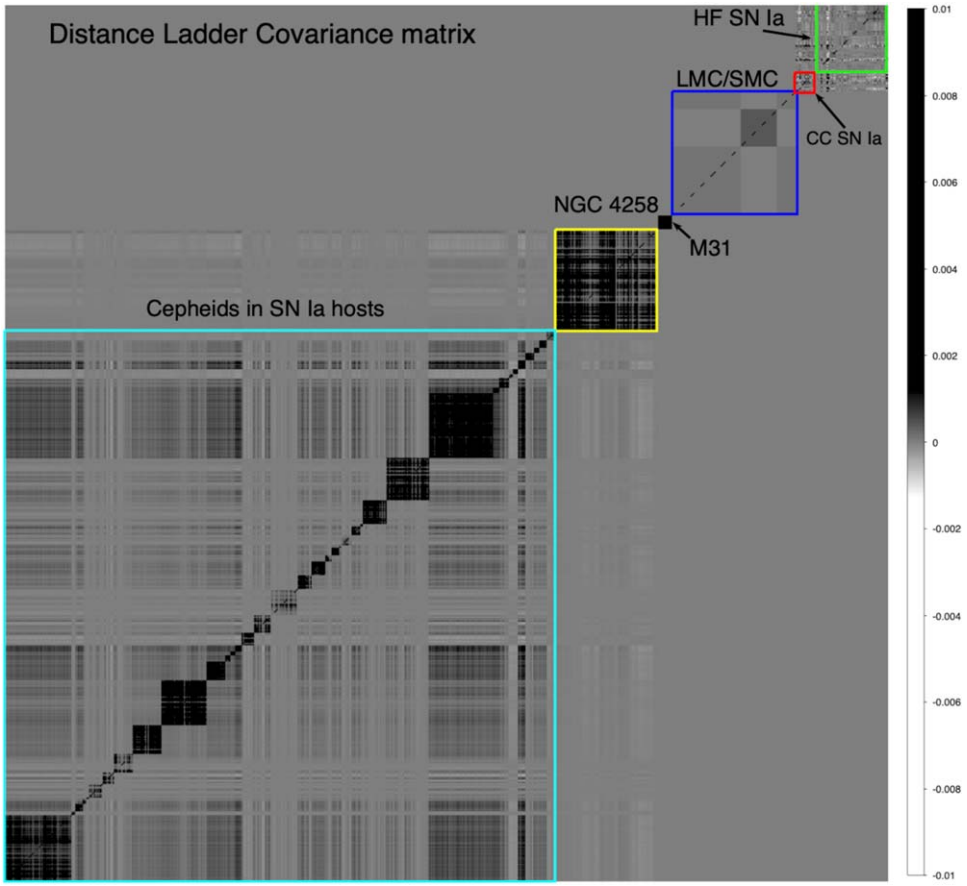

Figure 11. Distance-ladder covariance matrix. Graphical representation of the covariance matrix of measurements with values applicable to different sources indicated. The covariance matrix contains past sources of systematic error that propagate here as error in the fit. Among Cepheid hosts, square blocks along the diagonal indicate the background covariance within a host (Equation (8)) while the covariance pattern between hosts farther off the diagonals comes from the covariance of the metallicity scale (Equation (9)). The SN–SN covariance is characterized in the Pantheon+ analysis (Brout et al. 2021; Scolnic et al. 2021).

A much larger sample of M31 Cepheids is available from the HST PHAT Treasury program (Dalcanton et al. 2012), but the filters used to observe it do not correspond to the three used here, limiting its utility. The PHAT program observed these Cepheids with *WFC3 F160W* and a "wide-*J*" filter (*F110W*; Riess et al. 2012) defined a transformation to the m_H^W system. In R16 we included measurements for 375 PHAT Cepheids before the availability of those from Li et al. (2021). An expanded compilation from Kodric et al. (2018) includes 522 Cepheids from the PHAT program with $3 < P < 78$ days. We use the latter sample as an alternative to Li et al. (2021) in some variants of our baseline analysis in Section 6.5 because of its powerful leverage to examine evidence of a possible break in the *P–L* relation near $P \approx 10$ days.

4.6. Period–Luminosity Relations

Figure 9 shows the 40 individual Cepheid host m_H^W *P–L* relations (not including the MW as explained above) for a period range of $5 < P < 120$ days fitted with a common slope. Before combining the results of many hosts in a global analysis, we examine in Figure 10 the independently fitted slopes of the m_H^W *P–L* relations across all Cepheid hosts.

The slopes are all consistent (at the $\sim 2\sigma$ level) with a mean in the range of -3.27 to -3.30 mag dex^{-1}. The most tightly constrained slope comes from the LMC with -3.284 ± 0.017 mag dex^{-1}, mostly from the ground sample. Overall, we see no evidence to reject the null hypothesis of a single slope. There have been claims in the past of a break at $P \approx 10$ days, which we will consider as a variant of the primary analysis in the next section. However, the mean slopes from these data below and above $P = 10$ days are -3.33 ± 0.02 and -3.21 ± 0.06 mag dex^{-1} (respectively), a difference of 1.8σ; individual hosts with the strongest constraint (LMC and the M31 PHAT sample) show slope changes at $P = 10$ days in opposite directions. Furthermore, formal uncertainties in these slopes are somewhat underestimated because of the uneven sampling of periods between hosts. A Monte Carlo analysis (bootstrap resampling with replacement) from hosts with the largest samples of Cepheids shows variations owing to uneven period sampling increases the formal slope uncertainty typically by 10% and up to 35% for the LMC. We find that the metallicity dependence has a negligible effect on the mean slope. Therefore, in the following we will consider a single slope for $5 < P < 120$ days in our baseline analysis, but we analyze the impact of a break or limited period range on the determination of H_0 as variants of the baseline analysis.

4.7. Geometric Distance Priors

We make use of the geometric distances to NGC 4258 from masers (Reid et al. 2019) of $\mu_{\rm N4258} = 29.398 \pm 0.032$ mag, the LMC from DEBs (Pietrzyński et al. 2019) of $\mu_{\rm LMC} = 18.477 \pm 0.0263$ mag, and the SMC from DEBs (Graczyk et al. 2020) in the form of $\Delta\mu_{\rm LMC-SMC} = 0.500 \pm 0.017$ mag, as well as the MW Gaia EDR3 and HST spatial-

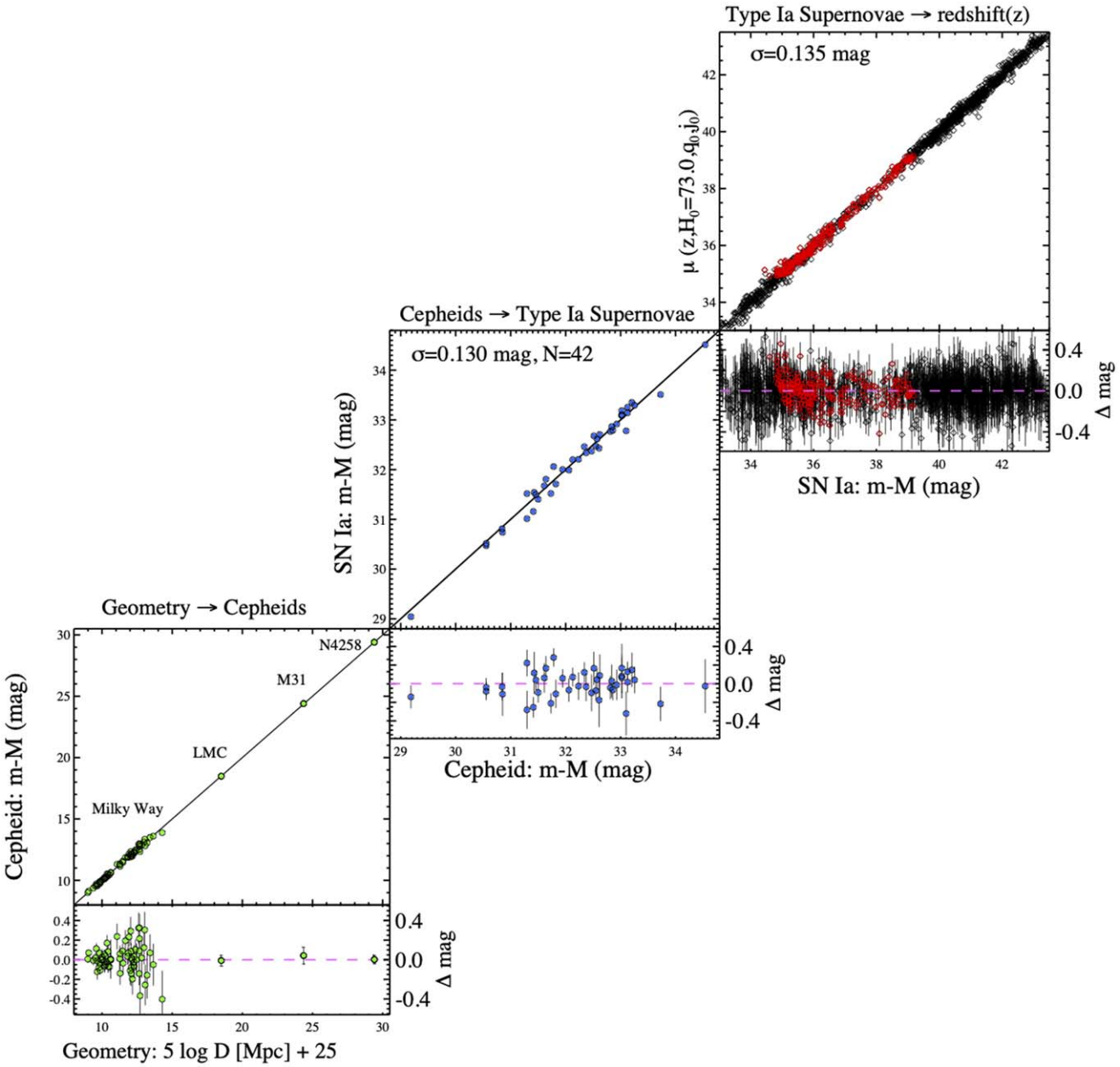

Figure 12. Complete distance ladder. The simultaneous agreement of distance pairs: geometric and Cepheid-based (lower left), Cepheid- and SN-based (middle), and SN- and redshift-based (top right) provides the measurement of H_0. For each step, geometric or calibrated distances on the abscissa serve to calibrate a relative distance indicator on the ordinate through the determination of M_B or H_0. Results shown are an approximation to the global fit as discussed in the text. Red SN points are at $0.0233 < z < 0.15$, with the lower-redshift bound producing the appearance of asymmetric residuals when plotted against distance.

scan parallaxes discussed in Section 4.1. These are all given in Table 4.

4.8. SN Magnitudes

We adopt standardized SN Ia magnitudes m_B^0 from the Pantheon+ analysis (Scolnic et al. 2021; Brout et al. 2021), where the value $m_{B,i}^0$ is a measure of the maximum-light apparent B-band brightness of an SN Ia in the ith host at the time of B-band peak, corrected to the fiducial color and luminosity determined for each SN Ia from its multiband light curves and a light-curve-fitting algorithm. We use the uncertainties and covariance of m_B^0 as given by the Pantheon+ analysis. The SN Ia covariance matrix has substantial off-diagonal terms and is displayed in Figure 11. An improvement in the current analysis of calibrator SNe Ia over R16 is our use of multiple SN light-curve data sets for most calibrators, 77 sets in all for 42 SNe Ia, a mean of ∼2 independent sets per SN, reducing measurement errors (not intrinsic scatter, which is covariant among multiple samples of the same SN) by a mean factor of 1.4. These data sets are given in Scolnic et al. (2021).

5. The Local Value of H_0, the Baseline

Our calibrator sample contains 42 SNe Ia in the 37 Cepheid hosts presented in the previous section and 277 SNe Ia in the Hubble flow, all the objects at $0.0233 < z < 0.15$ from the Pantheon+ sample that pass the same quality cuts and are in late-type hosts like the Cepheid calibrators. Criteria for

Table 5
Fits for H_0

Fit	Variant	$\chi^2_{\rm dof}$	N	H_0	b	γ	M^0_W	M^0_B	a_b
1	Baseline	1.03	3445	73.04 1.01	−3.299 0.015	−0.217 0.046	−5.894	−19.253	0.714158
	Cepheid-clipping Variants: Section 6.1								
2	global $\sigma_{\rm clip} = 3.3$	1.03	3446	73.19 1.01	−3.298 0.015	−0.216 0.046	−5.891	−19.249	0.714174
3	individual P–L $\sigma_{\rm clip} = 3.3$	0.99	3370	73.25 1.02	−3.296 0.015	−0.201 0.045	−5.893	−19.248	0.714315
4	tight:one-by-one MAD $\sigma_{\rm clip} = 3.0$	0.99	3429	73.12 1.01	−3.299 0.015	−0.220 0.045	−5.893	−19.251	0.714175
5	tight:global $\sigma_{\rm clip} = 3.0$	0.99	3432	73.22 1.01	−3.298 0.015	−0.215 0.045	−5.891	−19.248	0.714194
6	loose:global $\sigma_{\rm clip} = 5.0$	1.16	3475	73.34 1.01	−3.294 0.016	−0.221 0.049	−5.888	−19.244	0.714183
7	loose:one-by-one MAD $\sigma_{\rm clip} = 5.0$	1.15	3474	73.42 1.01	−3.295 0.016	−0.222 0.048	−5.888	−19.242	0.714178
8	loose:individual P–L $\sigma_{\rm clip} = 5.0$	1.05	3397	73.35 1.01	−3.296 0.015	−0.202 0.046	−5.892	−19.244	0.714257
9	none	1.23	3481	73.41 1.01	−3.290 0.016	−0.206 0.050	−5.885	−19.242	0.714248
	Geometric Anchors Variants: Section 6.2								
10	N4258	1.06	3454	72.51 1.54	−3.294 0.015	−0.204 0.051	−5.905	−19.269	0.714174
11	Milky Way	1.03	3446	73.02 1.19	−3.298 0.015	−0.208 0.051	−5.895	−19.254	0.714179
12	LMC	1.03	3446	73.59 1.36	−3.298 0.015	−0.208 0.051	−5.878	−19.237	0.714178
13	N4258+MW	1.03	3446	73.00 1.09	−3.298 0.015	−0.207 0.050	−5.895	−19.254	0.714171
14	N4258+LMC	1.03	3446	73.35 1.17	−3.299 0.015	−0.211 0.050	−5.885	−19.244	0.714179
15	MW+LMC	1.03	3446	73.25 1.05	−3.298 0.015	−0.216 0.046	−5.889	−19.247	0.714169
	Cepheid Dust-color Treatment Variants: Section 6.3								
16	Fitzpatrick 99 law $R_V = 2.5$	1.03	3446	73.24 1.00	−3.291 0.015	−0.209 0.046	−5.850	−19.247	0.714171
17	CCM law $R_V = 3.1$	1.04	3445	73.09 1.00	−3.310 0.015	−0.226 0.046	−5.957	−19.252	0.714183
18	Nataf law $R_V = 3.3$	1.03	3445	73.32 0.99	−3.276 0.015	−0.204 0.046	−5.804	−19.245	0.714141
19	R_W free global	1.03	3446	73.40 1.02	−3.286 0.016	−0.204 0.046	−5.835	−19.242	0.714167
20	intrin. col. subtr. F99 $R_V = 3.3$	1.03	3446	73.13 1.01	−3.201 0.015	−0.218 0.046	−5.604	−19.250	0.714170
21	intrin. col. subtr. F99 $R_V =$ free	1.03	3446	73.34 1.01	−3.201 0.015	−0.206 0.046	−5.582	−19.244	0.714170
22	intrin. col. subtr. R_V(host mass-SFR)	1.05	3446	73.85 1.02	−3.202 0.015	−0.228 0.046	−5.605	−19.229	0.714230
23	None (A_H values assumed to cancel)	1.11	3437	74.78 1.03	−3.188 0.016	−0.115 0.048	−5.434	−19.201	0.714086
	P–L Break/Span Variants: Section 6.4								
24	Break at $P = 10$ days	1.03	3444	72.68 1.04	−0.10 0.05	−0.222 0.046	−5.887	−19.264	0.714151
25	Only use $P > 10$ days	1.06	3004	73.15 1.11	−3.337 0.023	−0.169 0.051	−5.861	−19.250	0.714158
26	Only use $P < 60$ days	0.93	3699	73.99 1.04	−3.261 0.011	−0.236 0.044	−5.898	−19.226	0.714276
	M31 Cepheid Sample Variants: Section 6.5								
27	M31 PHAT sample	1.02	3854	73.21 1.00	−3.297 0.013	−0.234 0.044	−5.893	−19.248	0.714184
28	M31 PHAT sample +Break at $P = 10$ days	1.02	3852	72.74 1.03	−0.07 0.04	−0.240 0.044	−5.891	−19.262	0.714156
	Metallicity Variants: Section 6.6								
29	no metallicity dependence	1.04	3446	73.52 1.01	−3.296 0.015	...	−5.860	−19.239	0.714194
30	PP04 metallicity scale	1.04	3446	72.84 1.00	−3.297 0.015	−0.166 0.042	−5.883	−19.259	0.714162
	TRGB Inclusion Variants: Section 6.7								
31	adds EDD TRGB+N4258tip	1.03	3457	72.76 0.95	−3.301 0.015	−0.197 0.045	−5.891	−19.262	0.714253
32	adds CCHP TRGB+N4258tip	1.03	3457	72.29 0.94	−3.304 0.015	−0.208 0.046	−5.894	−19.276	0.714254
	Hubble-flow Sample Variants: Section 6.8								
33	all host types $0.0233 < z < 0.15$	1.03	3652	73.32 0.99	−3.298 0.015	−0.216 0.046	−5.891	−19.246	0.714479
34	highz:all host types $0.0233 < z < 0.80$	1.00	4483	73.68 0.98	−3.298 0.015	−0.216 0.045	−5.891	−19.244	0.716225
35	skip local alltypes $0.06 < z < 0.15$	1.04	3318	73.35 1.06	−3.298 0.015	−0.217 0.046	−5.891	−19.245	0.714311
36	highz:skip local alltypes $0.06 < z < 0.8$	1.00	4149	73.90 1.01	−3.298 0.015	−0.217 0.045	−5.891	−19.242	0.716991
37	highmass:hubble flow host logmass >10	1.04	3304	72.97 1.04	−3.298 0.015	−0.217 0.046	−5.891	−19.251	0.713297
	Calibrator Sample Variants: Section 6.9								
38	complete calibrator sample $z < 0.011$	1.03	3446	73.30 1.02	−3.298 0.015	−0.217 0.046	−5.891	−19.245	0.714191
39	complete sample $z < 0.011$+TRGB	1.04	3458	72.88 0.95	−3.301 0.015	−0.196 0.045	−5.888	−19.258	0.714297
40	highmass:calibrator host logmass >10	1.03	3445	73.54 1.10	−3.296 0.015	−0.214 0.046	−5.891	−19.238	0.714201
41	use least-crowded half	1.02	3446	73.34 1.16	−3.297 0.015	−0.223 0.046	−5.892	−19.246	0.714597
42	use most-crowded half hosts	1.02	3445	73.35 1.37	−3.293 0.015	−0.215 0.046	−5.891	−19.243	0.714063
43	only 19 hosts from R16	1.02	3446	73.47 1.17	−3.297 0.015	−0.229 0.046	−5.893	−19.242	0.714505

Table 5
(Continued)

Fit	Variant	$\chi^2_{\rm dof}$	N	H_0	b	γ	M^0_W	M^0_B	a_b
44	only hosts since R16	1.02	3445	73.07 1.31	−3.293 0.015	−0.210 0.046	−5.890	−19.252	0.714178
45	closer half hosts $m_b < 13$	1.03	3446	73.07 1.16	−3.295 0.015	−0.219 0.046	−5.891	−19.252	0.714191
	Excluded SN Surveys Variants: Section 6.10								
46	No SDSS SNe	1.03	3446	72.90 1.02	−3.298 0.015	−0.217 0.046	−5.891	−19.249	0.712428
47	No CSP SNe	1.02	3445	73.43 1.06	−3.298 0.015	−0.217 0.046	−5.891	−19.246	0.715130
48	No literature SNe	1.03	3446	73.47 1.05	−3.298 0.015	−0.213 0.046	−5.890	−19.240	0.714072
49	No LOSS SNe	1.02	3447	73.26 1.04	−3.297 0.015	−0.213 0.046	−5.891	−19.251	0.715022
50	No Swift SNe	1.03	3445	73.09 1.02	−3.297 0.015	−0.217 0.046	−5.891	−19.248	0.713551
51	No CfA1/2 SNe	1.03	3446	73.03 1.03	−3.298 0.015	−0.218 0.046	−5.891	−19.250	0.713464
52	No CfA3/4 SNe	1.03	3447	73.31 1.02	−3.298 0.015	−0.214 0.046	−5.891	−19.245	0.714257
53	No foundation SNe	1.01	3446	73.46 1.03	−3.298 0.015	−0.217 0.045	−5.891	−19.241	0.714196
54	No pre-2000 SNe	1.02	3446	73.20 1.09	−3.297 0.015	−0.218 0.046	−5.891	−19.245	0.713530
	SN-fitting Variants: Section 6.11								
55	SN scatter monochromatic	1.00	3444	73.54 1.08	−3.296 0.015	−0.223 0.045	−5.892	−19.238	0.714067
	Peculiar-velocity Variants: Section 6.12								
56	2MRS	1.03	3446	73.12 1.01	−3.298 0.015	−0.216 0.046	−5.891	−19.249	0.713775
57	CMB frame z	1.04	3446	72.70 1.00	−3.298 0.015	−0.216 0.046	−5.891	−19.249	0.711260
58	$q_0 = -0.52$	1.03	3446	73.19 1.01	−3.298 0.015	−0.216 0.046	−5.891	−19.249	0.714149
59	$q_0 = -0.52$ all highz	1.00	4483	73.65 0.98	−3.298 0.015	−0.216 0.045	−5.891	−19.244	0.715950
	Optical Wesenheit Variants: Section 6.13								
60	optical Wesenheit clipping=one-by-one	0.94	3626	72.70 1.03	−3.299 0.010	−0.248 0.041	−5.858	−19.264	0.714413
61	optical Wesenheit clipping=global	0.94	3626	72.90 1.03	−3.291 0.010	−0.247 0.041	−5.857	−19.259	0.714417
62	optical Wesenheit F99 $R_V = 2.5$	0.92	3618	73.20 1.03	−3.230 0.010	−0.202 0.041	−5.623	−19.249	0.714359
63	optical Wesenheit CCM $R_V = 3.1$	0.97	3626	72.46 1.03	−3.335 0.010	−0.270 0.042	−6.020	−19.272	0.714469
64	optical Wesenheit N4258 only	0.92	3623	74.85 2.31	−3.291 0.010	−0.211 0.045	−5.797	−19.201	0.714429
65	optical Wesenheit MW only	0.93	3623	71.93 1.15	−3.291 0.010	−0.211 0.045	−5.883	−19.288	0.714421
66	optical Wesenheit LMC only	0.92	3623	74.26 1.39	−3.291 0.010	−0.211 0.045	−5.814	−19.218	0.714416
67	optical Wesenheit+TRGB	0.94	3638	72.15 0.94	−3.301 0.010	−0.239 0.041	−5.858	−19.281	0.714385

inclusion in the sample used to measure the Hubble flow and variations are further considered in Section 6.7.

Our baseline (best) analysis includes the full array of constraints on our model and what we assess to be the optimal choices (considering a wide range of alternatives in Section 6) in the treatment of the available data. Its results, derived from the optimization of χ^2 in Equation (6), are given in the first line of Table 5, which also provides the best-fit parameters. This fit gives a $\chi^2_{\rm dof} = 1.03$ with $N = 3445$ degrees of freedom, with Cepheid slope, metallicity, and luminosity parameters $b_W = -3.299 \pm 0.015$ mag/dex, $Z_W = -0.217 \pm 0.046$ mag/dex, and $M^W_{H,1} = -5.894 \pm 0.017$ mag. These parameters are similar to those found by R16 and updated by R19, with small increases in the absolute values of the slope of the P–L relation (from -3.26 ± 0.03 in R16) and the metallicity dependence (from -0.17 ± 0.06 in R19). Additional "nuisance" parameters in the baseline fit are $\Delta\mu_{\rm N4258} = -0.013 \pm 0.022$ mag and $\Delta\mu_{\rm LMC} = 0.010 \pm 0.019$ mag. Table 6 provides individual host distances and SN parameters. Figure 12 displays the baseline data, fit, and residuals, while Figure 13 provides more details about the central panel of the preceding figure.

The fiducial SN absolute-magnitude parameter applicable to the Pantheon+ standardization (Scolnic et al. 2021; Brout et al. 2021) is $M^0_B = -19.253 \pm 0.027$ mag. The value of the Hubble constant derived from the baseline fit is $H_0 = 73.04 \pm 1.01$ km s^{-1} Mpc^{-1}. Including an additional systematic uncertainty from the analysis variants as discussed in Section 6.14 yields

$$\rm H_0 = 73.04 \pm 1.04 km\ s^{-1}\ Mpc^{-1} (baseline\ with\ systematics). \quad (10)$$

The determination of H_0 in the measured space is $5 \log H_0 = 9.318 \pm 0.031$ (with systematics), where the errors are symmetric in log space and slightly asymmetric for H_0 with errors of (-0.98 and $+1.03$ for the fit). The full difference from the Planck+ΛCDM prediction of $H_0 = 67.4 \pm 0.5$ (Planck Collaboration et al. 2020) in units of $\Delta 5 \log H_0$ is 0.176 ± 0.035 mag (errors in quadrature), a difference of 5.0σ (one in 3.5 million).

The dispersion (weighted by the measurement errors) between the 42 calibrator SNe Ia and Cepheids is $\sigma = 0.130$ mag, which is equivalent to (albeit slightly lower than) the 0.135 mag dispersion between the SN magnitudes and redshifts of the Hubble-flow sample. We would expect these dispersions to be comparable (as they are) because the additional sources of scatter independent of SNe and applicable to the two comparisons (Cepheid distance errors and peculiar velocities) are both subdominant to SN scatter. From these dispersions and the good global χ^2, we conclude that there is no unexplained variance in the baseline fit beyond the intrinsic scatter of SNe Ia and the intrinsic width of the Cepheid instability strip.

Table 6
Approximations for Distance Parameters

N	Host	SN	$m^0_{B,i}$	σ	$\mu_{\rm Ceph}$[a]	σ (mag)	$M^0_{B,i}$	σ	$\mu_{\rm host}$[b]	σ	R[c]
1	M101	2011fe	9.7800	0.115	29.194	0.039	−19.414	0.121	29.178	0.041	0.44
2	M1337	2006D	13.655	0.106	32.922	0.124	−19.267	0.163	32.920	0.123	0.43
3	N0105	2007A	15.250	0.133	34.531	0.253	−19.281	0.286	34.527	0.250	0.37
4	N0691	2005W	13.602	0.139	32.847	0.109	−19.245	0.177	32.830	0.109	0.42
5	N0976	1999dq	14.250	0.103	33.719	0.151	−19.468	0.183	33.709	0.149	0.30
6	N1015	2009ig	13.350	0.094	32.570	0.075	−19.220	0.120	32.563	0.074	0.46
7	N1309	2002fk	13.209	0.082	32.546	0.060	−19.337	0.102	32.541	0.059	0.38
8	N1365	2012fr	11.900	0.092	31.379	0.057	−19.479	0.108	31.378	0.056	0.46
9	N1448	2001el	12.254	0.136	31.290	0.037	−19.036	0.141	31.287	0.037	0.41
10	N1448	2021pit	11.752	0.200	31.290	0.037	−19.538	0.203	31.287	0.037	0.41
11	N1559	2005df	12.141	0.086	31.501	0.062	−19.361	0.106	31.491	0.061	0.35
12	N2442	2015F	12.234	0.082	31.457	0.065	−19.223	0.105	31.450	0.064	0.36
13	N2525	2018gv	12.728	0.074	32.067	0.100	−19.339	0.124	32.051	0.099	0.39
14	N2608	2001bg	13.443	0.166	32.629	0.155	−19.186	0.227	32.612	0.154	0.39
15	N3021	1995al	13.114	0.116	32.475	0.160	−19.361	0.198	32.464	0.158	0.34
16	N3147	2021hpr	13.843	0.159	33.044	0.165	−19.201	0.229	33.014	0.165	0.45
17	N3147	1997bq	13.821	0.141	33.044	0.165	−19.223	0.217	33.014	0.165	0.45
18	N3147	2008fv	13.936	0.200	33.044	0.165	−19.108	0.259	33.014	0.165	0.45
19	N3254	2019np	13.201	0.074	32.332	0.077	−19.130	0.107	32.331	0.076	0.52
20	N3370	1994ae	12.937	0.082	32.123	0.052	−19.186	0.097	32.120	0.051	0.33
21	N3447	2012ht	12.736	0.089	31.939	0.035	−19.203	0.096	31.936	0.034	0.40
22	N3583	ASASSN-15so	13.509	0.093	32.808	0.081	−19.299	0.123	32.804	0.080	0.34
23	N3972	2011by	12.548	0.094	31.644	0.090	−19.096	0.130	31.635	0.089	0.46
24	N3982	1998aq	12.252	0.078	31.724	0.072	−19.472	0.106	31.722	0.071	0.36
25	N4038	2007sr	12.409	0.106	31.615	0.117	−19.207	0.158	31.603	0.116	0.57
26	N4424	2012cg	11.487	0.192	30.856	0.130	−19.369	0.232	30.844	0.128	0.87
27	N4536	1981B	11.551	0.133	30.838	0.051	−19.287	0.142	30.835	0.050	0.49
28	N4639	1990N	12.454	0.124	31.818	0.085	−19.364	0.150	31.812	0.084	0.48
29	N4680	1997bp	13.173	0.205	32.606	0.208	−19.433	0.292	32.599	0.205	0.48
30	N5468	1999cp	13.880	0.080	33.120	0.075	−19.240	0.110	33.116	0.074	0.39
31	N5468	2002cr	13.993	0.072	33.120	0.075	−19.127	0.104	33.116	0.074	0.39
32	N5584	2007af	12.804	0.079	31.775	0.053	−18.971	0.095	31.772	0.052	0.36
33	N5643	2013aa	11.252	0.079	30.570	0.050	−19.318	0.093	30.546	0.052	0.39
34	N5643	2017cbv	11.208	0.074	30.570	0.050	−19.362	0.089	30.546	0.052	0.39
35	N5728	2009Y	13.514	0.115	33.117	0.206	−19.602	0.236	33.094	0.205	0.58
36	N5861	2017erp	12.945	0.107	32.232	0.101	−19.287	0.147	32.223	0.099	0.33
37	N5917	2005cf	13.079	0.095	32.362	0.122	−19.284	0.154	32.363	0.120	0.47
38	N7250	2013dy	12.283	0.178	31.628	0.126	−19.344	0.218	31.628	0.125	0.49
39	N7329	2006bh	14.030	0.079	33.274	0.116	−19.244	0.140	33.246	0.117	0.39
40	N7541	1998dh	13.418	0.128	32.504	0.120	−19.086	0.176	32.500	0.119	0.30
41	N7678	2002dp	14.090	0.093	33.196	0.155	−19.106	0.181	33.187	0.153	0.39
42	U9391	2003du	13.525	0.084	32.849	0.068	−19.324	0.108	32.848	0.067	0.44

Notes. $m^0_{B,i}$ and error given is an approximation, the simple average from 1 or more SN surveys.
[a] Approximate, Cepheid-based distances derived without inclusion of the SNe in each host.
[b] Cepheid-based distances derived without inclusion of any SNe in any host.
[c] Empirical host reddening ratio derived from mass, SFR, using EDA, see Section 6.3.

5.1. Markov Chain Monte Carlo Sampling

To check the preceding calculation using a different methodology, we performed Markov Chain Monte Carlo (MCMC) sampling. Our likelihood function is

$$p(q|L,\,C,\,y) \propto p(q)\,p(L,\,C,\,y|q), \tag{11}$$

where the $p(L,\,C,\,y|q)$ is a Gaussian likelihood based on the covariance matrix C; if C is diagonal, $C_{ij} = \sigma_i^2\,\delta_{ij}$, it has the form

$$p(L,\,C,\,y|q) = \prod_i \frac{1}{\sqrt{2\pi}\,\sigma_y} \exp\left[\frac{-(y_i - y_{{\rm fit},i})^2}{2\sigma_y^2}\right]. \tag{12}$$

Because the covariance matrix C is independent of q in our setup, Equation (11) can be simplified to

$$p(q|L,\,C,\,y) \propto p(q)\exp\left(-\frac{1}{2}\sum_i \chi_i^2\right). \tag{13}$$

For the calculation of χ^2, we followed the method described in Equation (6). For numerical accuracy, we substituted C^{-1} with an inverse of Cholesky-decomposed lower triangular matrix $\tilde{C}_L^{-1}$ (which satisfies $\tilde{C}_L\tilde{C}_L^{-1} = I$ and $\tilde{C}_L\tilde{C}_L^T = C$):

$$\chi^2 = (y - L\cdot q)\tilde{C}_L^{-1}(y - L\cdot q). \tag{14}$$

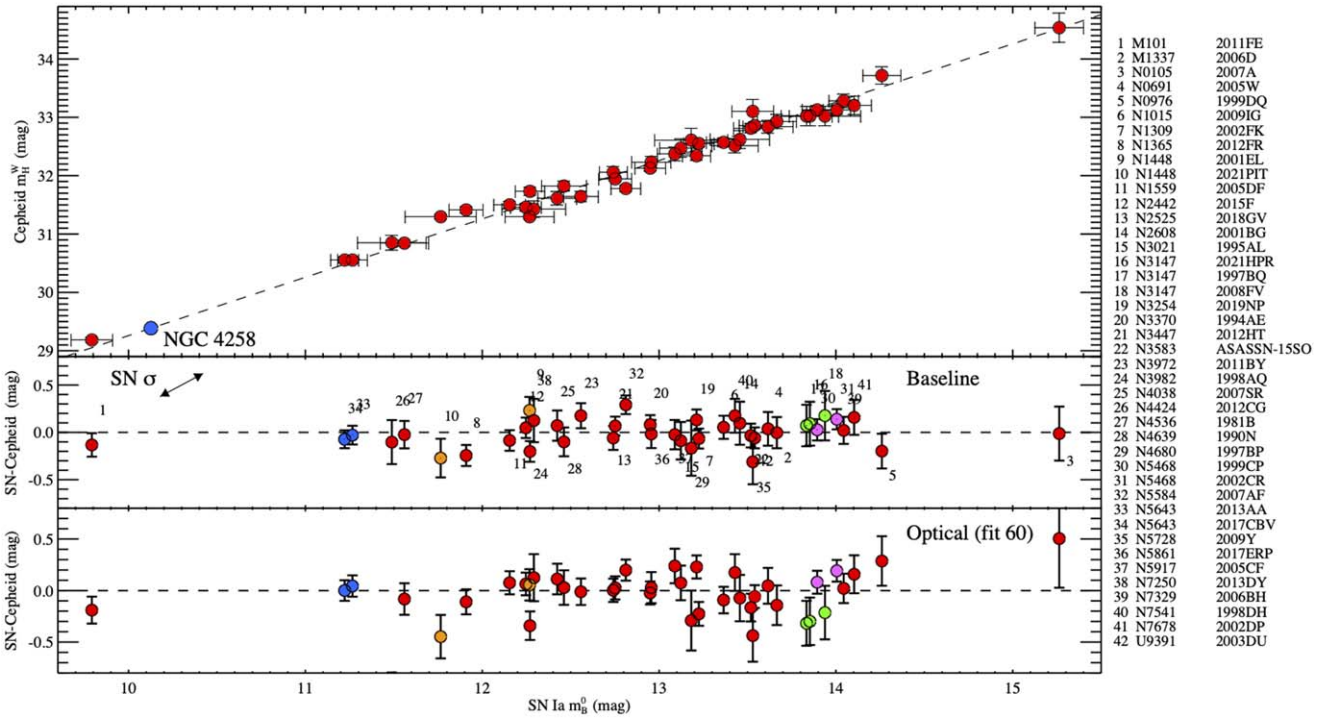

Figure 13. Comparison of SN and Cepheid distance measures on the second rung of the distance ladder, and residuals (middle). Because the residual panel includes SN measures on both axes, residuals are convolved by the SN error arrow as indicated. An alternate plot with Cepheids on the abscissa is shown in Figure 12. SNe that share a common host are identified with various colors. We find no significant trend ($<1.5\sigma$) in this space. Results shown are an approximation to the global fit as discussed in the text. The lower panel replaces the NIR-based Cepheid magnitudes with the optical only from Fit 60.

As a "neutral" prior, we chose the uniform distribution

$$q_i \sim \text{Uniform}(\mu_i - 10\sigma_i, \mu_i + 10\sigma_i), \tag{15}$$

which is centered at the results from the analytical result. Its width is set to be sufficiently (20×) larger than the estimated standard deviation from linear regression σ. We find that this range is more than enough to allow the determination of median and standard deviation of the resulting population from μ and σ.

We used the `Python` package `emcee` (Foreman-Mackey et al. 2013) to perform sampling under conditions mentioned above. The initial states are uniformly distributed in the prior range (Equation (14)) for a total of 100 walkers. The convergence is monitored using `emcee`'s recommended method to estimate the autocorrelation time τ, which employs a method originally proposed by Goodman & Weare (2010). The burn-in time is set to be 5τ to allow chains to fully converge.[16]

The convergence of the sampled distribution is checked visually and by the estimated autocorrelation time value. We required the total number of chains after burn-in N to satisfy $N > 100\tau$ before any visual inspection for convergence is performed. The samples used in the final analysis are contained well within the inner region of the initial states for all parameters, indicating that our choice of prior size did not affect the final result.

The results for selected parameters are shown in Figure 14. All parameters exhibit Gaussian-like posterior probability density functions (PDFs), which indicates that the results from the linear-regression method are a good representation of the results from this project. The locations of Gaussian-equivalent percentiles (i.e., the locations at which the same relative volume is met) in our sample suggest that the samples in H_0 are more broadly distributed toward larger values (right-hand side; RHS) than smaller ones (left-hand side; LHS), a consequence of Gaussian errors in magnitudes of $5\log H_0$, with the mean distances between each line $\bar{\sigma}_R = 1.06$ on the RHS and $\bar{\sigma}_L = 0.98$ on the LHS (toward lower H_0), similar to the analytical results.

More than 120 million samples were used for the baseline fit to delineate the position of the 5σ confidence interval for H_0, as shown in Figure 15. The calculated values of σ, whose mean is $\bar{\sigma}_{\rm all} = 1.02$, contains the estimated uncertainty from the linear-regression method $\sigma_{\rm reg} = 1.01$, hence showing a full consistency between two methods. Using these three values ($\bar{\sigma}_L$, $\bar{\sigma}_{\rm all}$, and $\sigma_{\rm reg}$), and regarding the 50th percentile value as the most probable value for the first two cases, our results are $H_{0,\rm MCMC-L} = 73.04 \pm 0.98$, $H_{0,\rm MCMC-all} = 73.42 \pm 1.02$, and $H_{0,\rm reg} = 73.04 \pm 1.01$ km s^{-1} Mpc^{-1}.

The significance of the discrepancy between our results and those of Planck Collaboration et al. (2020), $H_{0,\rm Planck} = 67.4 \pm 0.5$, are 5.1σ, 4.9σ, and 5.0σ, respectively.

5.2. Simultaneous Constraints on H_0 and the Expansion History

This section illustrates an approach to simultaneously measure H_0 and the expansion history for an arbitrary form of H(z), using the full covariance of the data sets.

The "near-field" determination of H_0 in the preceding section is quite insensitive to specific knowledge of the recent expansion history because the mean redshift at which the measurement is made, $\langle z \rangle = 0.055$, is low. Yet to make the measurement more

[16] We used the final estimation of $\tau(N = N_{\rm max})$, where the relative change of estimated value becomes small enough: $d\tau/dN/\tau \ll 0.001\%$.

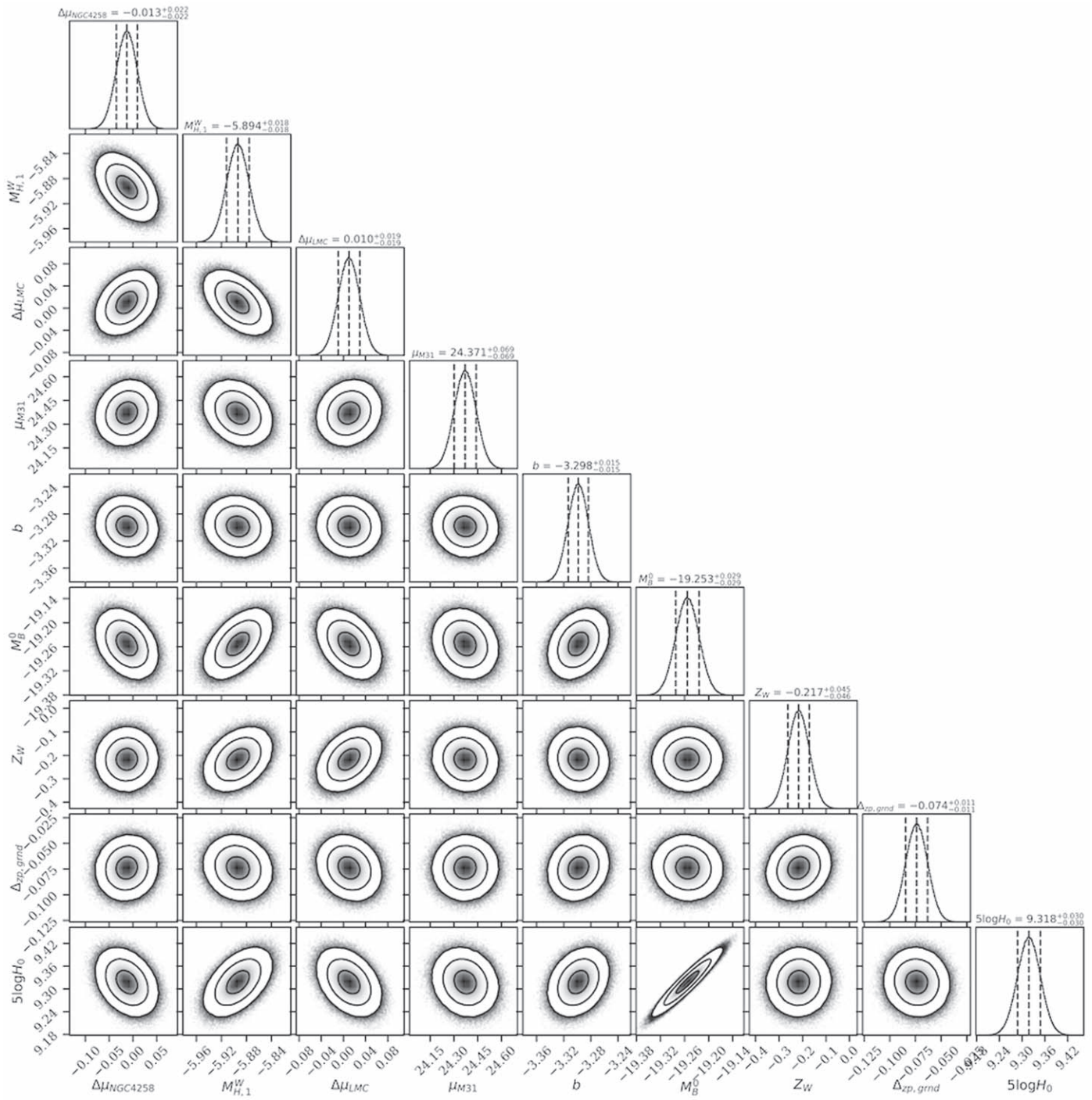

Figure 14. MCMC sampling of the χ^2 statistic of the global fit showing all free parameters except the individual host distances. Contours are 1σ, 2σ, and 3σ confidence regions. We find that the means and uncertainties agree very well with the analytical minimization of χ^2.

precise at even these small redshifts, in the prior section we accounted for the derivative of the expansion history in Equation (4) through the empirical derivative of H(z) measured from higher-redshift SNe Ia, q_0. We set $q_0 = -0.55$ as this is historically a good fit to high-redshift SNe Ia.[17] This empirical approach is fully independent of the CMB, so an independent comparison of H_0 to the CMB with ΛCDM is appropriate.

However, to consider less-conventional expansion histories such as a rapid change in $H(z)$ together with the measured value of H_0, as may be undertaken in the effort to find a resolution of the Hubble tension, it might be necessary and it is certainly more reliable to explicitly account for the dependence of H_0 on the form of the expansion history, $H(z)$, at low redshifts (Camarena & Marra 2021; Efstathiou 2021). Furthermore, there is covariance in the measurements of SNe Ia, whether in Cepheid hosts or at moderate redshifts (Scolnic et al. 2018; Dhawan et al. 2020), and it is necessary to account for this when SNe Ia are used simultaneously to measure H_0 and the recent expansion history.

[17] It is also the expectation for a consensus ΛCDM with $\Omega_M = 0.3$ and $\Omega_\Lambda = 0.7$.

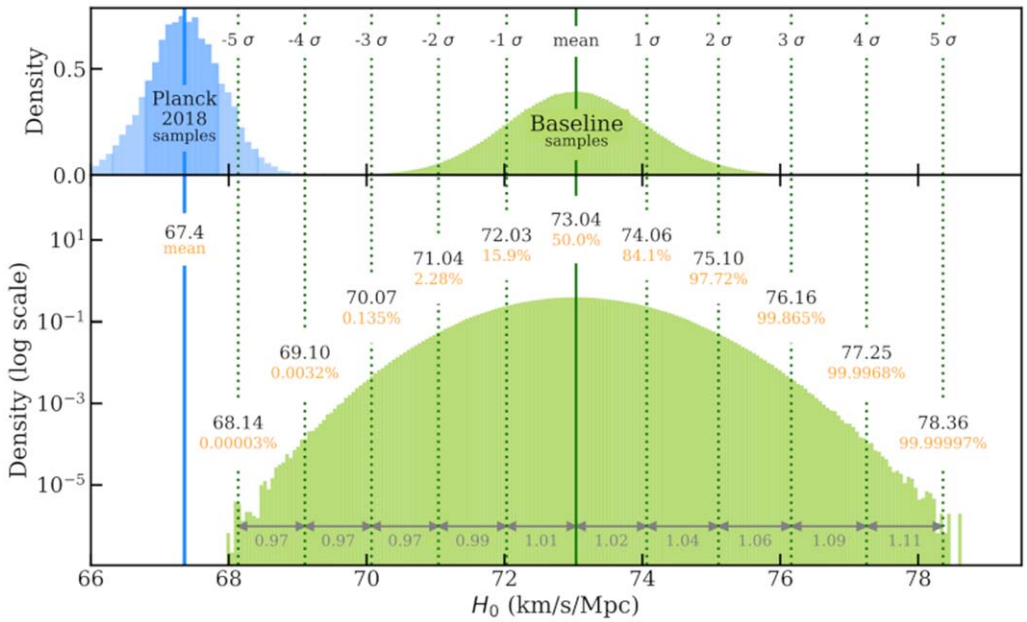

Figure 15. Extended MCMC sampling of the posterior for H_0 to measure out to the $\pm 5\sigma$ confidence level. The upper panel shows the probability density for the baseline from SH0ES and from the Planck Collaboration et al. (2020) chains. The bottom panel shows the log of the probability density to improve the ability to see the tails. We note some asymmetry to the distribution, with intervals on the low-H_0 side a little smaller than that on the high side, as the measurements are Gaussian in magnitude and in 5 log H_0, hence slightly skewed in H_0.

It is important to recognize that standardized SN Ia data can provide only relative distance measurements between all SNe measured within the same standardization scheme, with SN parameters, uncertainties, and covariance with values relevant within the context of the algorithm used to standardize the SNe. Therefore, to avoid inconsistencies between SN standardization schemes or loss of knowledge of measurement covariance, one would ideally make full use of all relevant SN data simultaneously to determine absolute quantities. A straightforward and reliable path is to use the set of absolute distances to SN hosts (their uncertainties and covariance) derived from only the first two rungs without the use of any SN data, together with a consistently standardized set of SNe (in these hosts and in the Hubble flow), to determine H_0 and $H(z)$ simultaneously, the so-called "forward" direction. Or, one could alternatively follow an "inverse" approach, starting with CMB data and using the best-fit ΛCDM model to calibrate the distance–redshift relation of SNe including the predicted distances of nearby hosts of SNe Ia and Cepheids (or TRGB).[18] Here we follow the forward approach.

For the set of SN hosts with calibrated distances (using only the first two rungs of the distance ladder), the dereddened absolute distance modulus μ_0 is given by

$$\mu_{0,\rm host} - m_B = M_B, \tag{16}$$

with terms given in Equation (3). We then have a second set of SNe with cosmological redshifts $z > 0$,

$$m_B = 5 \log cz \left\{ 1 + \frac{1}{2}[1 - q_0]z - \frac{1}{6}[1 - q_0 - 3q_0^2 + j_0]z^2 + O(z^3) \right\} - 5 \log \mathrm{H}_0 + M_B + 25. \tag{17}$$

The two leading terms on the RHS of Equation (17) can be replaced with any empirical or cosmological model for H(z), such as the example of ΛCDM with the dark energy equation-of-state parameter w and mass density Ω_M,

$$m_B = 5 \log \left[c\mathrm{H}_0^{-1}(1+z) \int_0^z \frac{dz'}{E(z')} \right] + 25 + M_B, \tag{18}$$

where

$$E(z) \equiv \left\{ \Omega_m (1+z')^3 + (1-\Omega_m) \times \exp\left[+3 \int_0^{\ln(1+z)} d\ln(1+z')(1+w(z')) \right] \right\}^{1/2}. \tag{19}$$

The LHS of Equations (16) and (17) (or 16 and 18) are measured quantities, either $\mu_{0,\rm host}$ or m_B, and the free parameters M_B, H_0, and q_0 (or w and Ω_M for Equation (18) instead of Equation (17)) are determined by simultaneously optimizing these.

Following this approach, the set of 37 values of $\mu_{0,\rm host}$ for 42 SNe Ia given in Table 6 (or 40 values for 46 SNe Ia including the TRGB data as in Section 6.7) and their covariance matrix,

[18] A variation of the inverse approach is to predict M_B and compare that with local hosts as suggested by Efstathiou (2021), an approach that is equivalent in principle but the dependence of M_B on the method of SN standardization makes it less widely applicable.

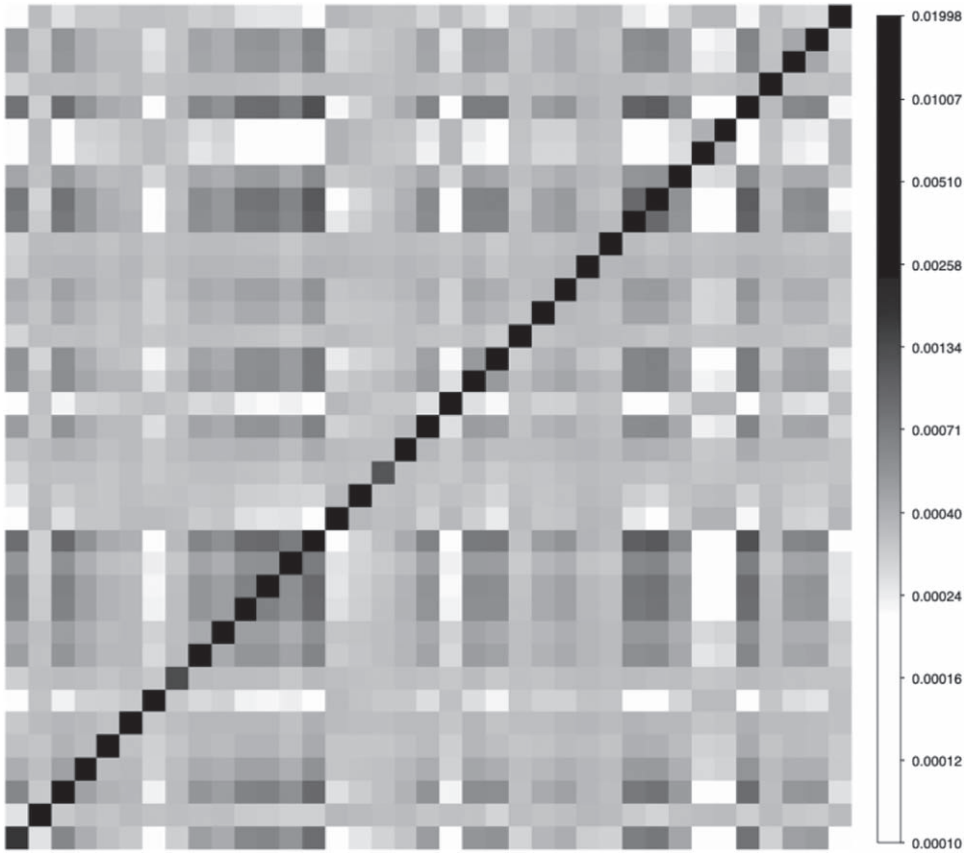

Figure 16. Host distance ($\mu_{\rm host}$) covariance matrix. This is the covariance for the baseline set of 37 hosts derived only from the first two rungs and without the use of any SN Ia data. The nonzero off-diagonal terms result from common anchors, common parameters of the Cepheid *P*–*L* relation, and covariance of the metallicity scale. The systematic uncertainty in all host distances may be measured from the square root of the level of the off-diagonal elements and is 0.019 mag, or 0.9% in distance.

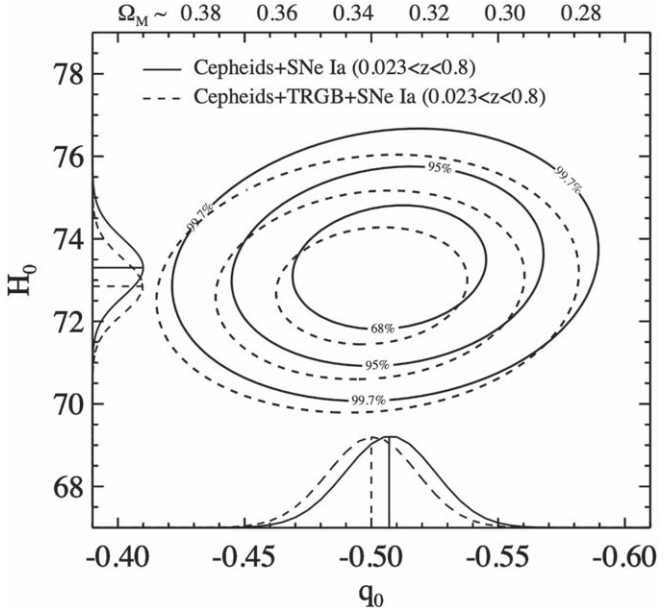

Figure 17. Simultaneous measurement of H_0 and expansion history. Here the expansion history is fit with a single free parameter, q_0, though as discussed in the text, it can be fit with any form of $H(z)$ using the set of $\mu_{\rm host}$ and a set of consistently standardized SNe Ia in these hosts and in the Hubble flow.

derived from the first two rungs (i.e., without the use of any SN data), are taken as input. The covariance matrix of $\mu_{0,\rm host}$ is shown in Figure 16. In addition, the values of m_B for the SNe in these hosts and in the Hubble flow (including their redshifts) and the covariance matrix between all SN measurements are taken as additional data and the two constraining relations are solved simultaneously.

As an example, in Figure 17 we follow this approach, marginalizing over the SN standardization parameter, M_B, and simultaneously measuring H_0 and q_0. As expected, H_0 will be quite uncorrelated with q_0 (or other cosmological parameters) unless a cosmological model produces a much more rapid change in $H(z)$ at $z \ll 1$ than the polynomial in Equation (17). For this case we add SNe from Pantheon+ at $0.15 < z < 0.8$ to the prior analysis, finding $q_0 = -0.51 \pm 0.024$ and $H_0 = 73.30 \pm 1.04\ {\rm km\ s^{-1}\ Mpc^{-1}}$. The result is very similar to our baseline result, and the added freedom in the expansion history has had little impact on the uncertainty. Brout et al. (2021) give results for a Friedman–Robertson–Walker (FRW) expansion history governed by w and Ω_M. We provide the SN-independent host distances and their covariance matrix at pantheonplussh0es.github.io to allow for other forms of $H(z)$.

6. Extensions or Variants of the Baseline Analysis

The baseline analysis was identified as providing the most accurate model of the data that is also the most economical in terms of the number of free parameters. Here we review 12 categories of alternatives or extensions to the baseline analysis, selected to explore the sensitivity of the results to additional considerations and systematic uncertainties. These are given as

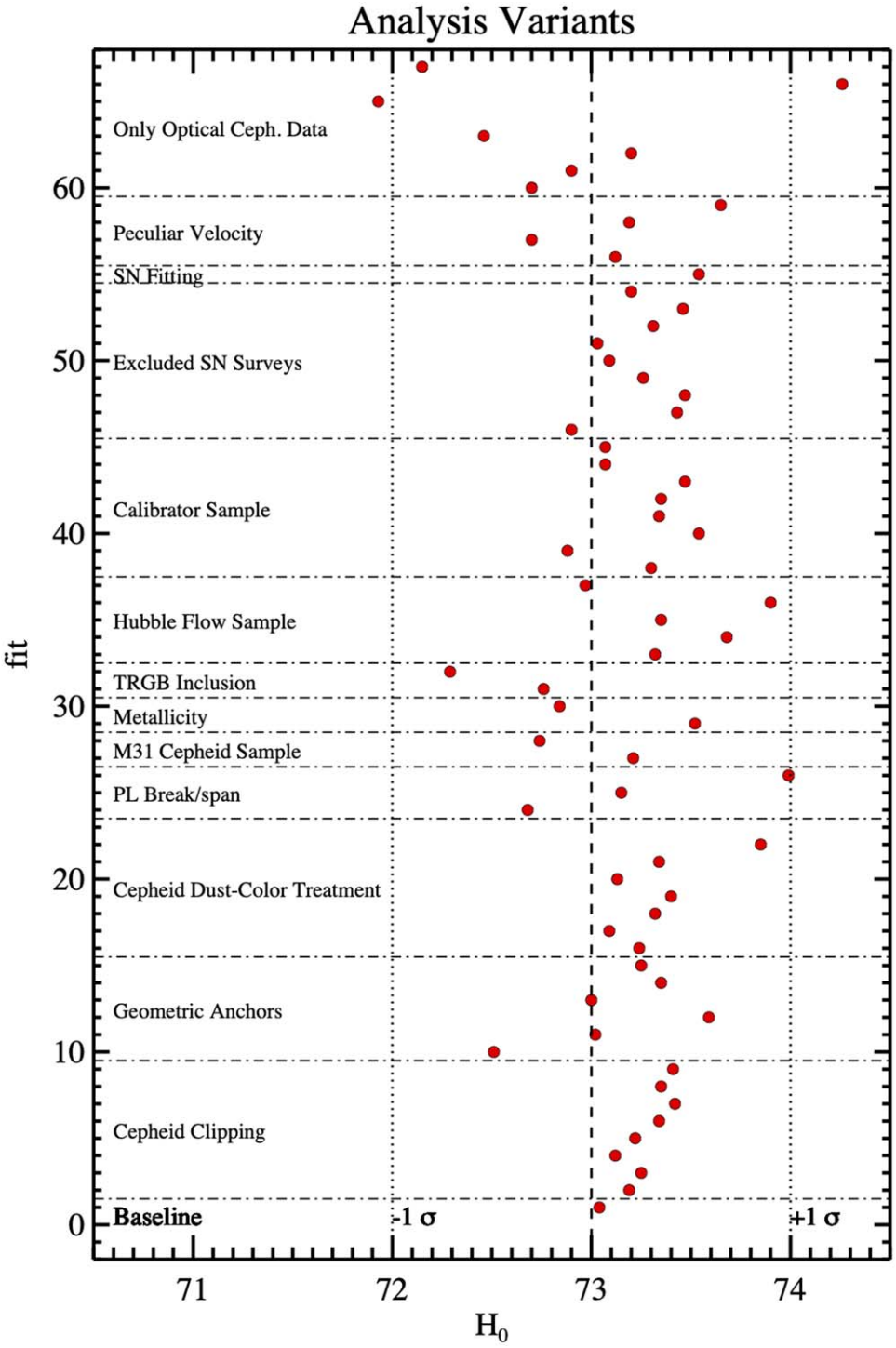

Figure 18. Display of 67 fits in 12 categories of alternatives or extensions to the baseline as shown in Table 5.

67 fits listed in Table 5, shown graphically in Figure 18, and summarized in Section 6.14.

6.1. Cepheid Clipping

The optical selection of our Cepheid sample is discussed by H16 and Y22b. As in R16, we include only Cepheids with colors *F814W–F160W* within 0.8 mag of the median color in each host to remove blends with unresolved sources of comparable luminosity and different color (e.g., red giants, blue supergiants, unresolved star clusters). This is a useful criterion as it is independent of distance and period, insensitive to reddening, and anchored to the physical properties of Cepheids (i.e., stars with spectral types F–K). We still may expect a small number of outliers owing to fully blended yellow supergiants, or the sample may include a small number of objects erroneously identified as Cepheids.

Our baseline analysis removes Cepheids that deviate from the global fit at $>3.3\sigma$ (Chauvenet's criterion), iteratively discarding the single largest outlier (i.e., MAD algorithm) until

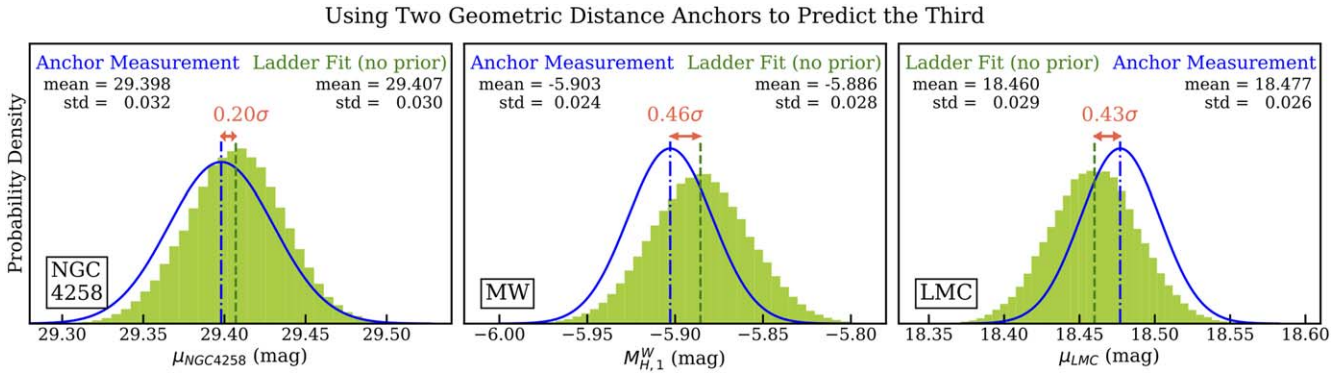

Figure 19. Comparison of each geometric anchor distance with its expected value based only on the Cepheids it hosts and the distance ladder calibrated by the geometric distance of the other two anchors. The green histograms are MCMC samples from Fits 13, 14, and 15 showing the expected distance of an anchor whose independent distance was excluded from the analysis. The independent measurement (blue curve) and prediction (green) are in good agreement in all cases to $<1\sigma$.

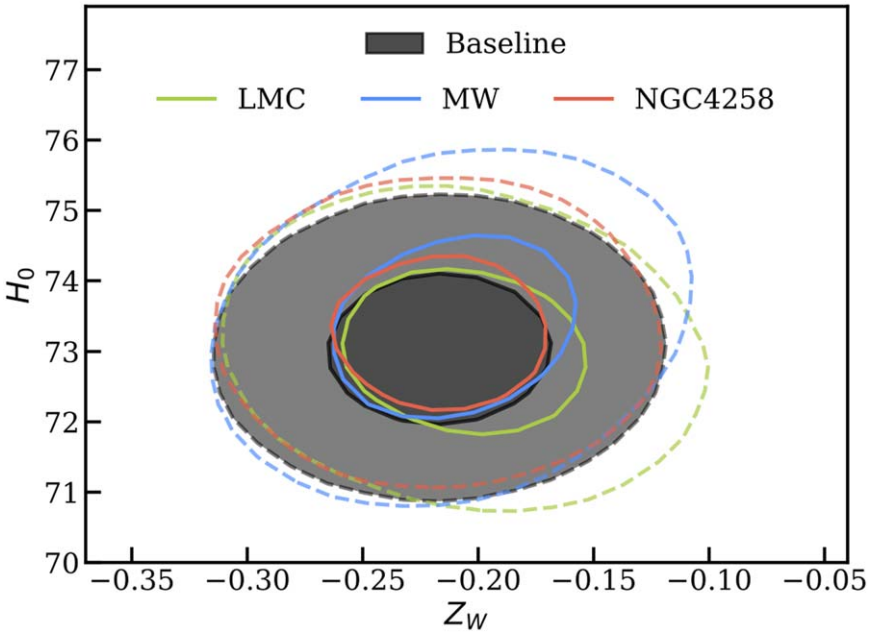

Figure 20. Marginalized posterior covariance between H_0 and the metallicity term, Z_W, for the three two-anchor cases and the baseline fit. The metallicity term is well constrained, with a substantial tightening owing to the differential DEB distance between the two clouds (LMC and SMC) from Graczyk et al. (2020). The term has little correlation with H_0 because the anchor abundances span the range in the SN hosts, but this term provides for the consistency between the anchors.

none remain above the threshold. The fraction of such objects is 1.2% (these outliers are available upon reasonable request). Fits 2–9 explore other rejection approaches: global (i.e., all objects above the threshold are removed as opposed to the most deviant followed by recalculation of the fit), from the individual *P–L* relations, and with tighter or looser thresholds (3σ or 5σ), as well as no rejection. The median H_0 of these alternatives to the baseline clipping is larger by 0.2 km s^{-1} Mpc^{-1}. Because the global outlier removal (Fit 2) is faster to calculate and gives results within 0.15 km s^{-1} Mpc^{-1} of the MAD baseline, we use this approach for most other fits unless we explicitly state the use of MAD. For the number of degrees of freedom here, fits with $\chi^2_{dof} > 1.07$ are considered not good (probability to exceed $>3\sigma$), which applies to Fits 6, 7, and 9 that remove few or no Cepheid outliers.

6.2. Geometric Anchors: Consistency with Metallicity

Fits 10–15 provide the results of including the geometric distance measurement(s) of only one or at most two anchors, rather than all three. The goal here is to explore the possibility of an unexpected error in one of the external geometric distance constraints. In Figure 19 we compare the geometric distance estimate to each anchor with the value modeled using only its Cepheids and the geometric distance of the other two anchors. As shown, the external measured distances are consistent with their Cepheids at -0.2σ for NGC 4258, $+0.4\sigma$ for the LMC, and -0.5σ for the MW. Their internal consistency (despite having different mean abundances) is a consequence of the Cepheid metallicity dependence.

The consistency of the anchors is most readily seen by comparing their intercepts and metallicity as shown in Figures 20 and 21. Here we determine the intercepts from the Cepheid data table for each host, NGC 4258, LMC, SMC, and the MW using their geometric distances for an absolute measure; see also Breuval et al. (2021) for a similar analysis. It is clear that the consistency of the anchors is a direct consequence of the Cepheid metallicity dependence, which has been greatly refined since R16. In our primary analysis we find a metallicity dependence on [O/H] of -0.217 ± 0.046 mag/dex. In R16 our

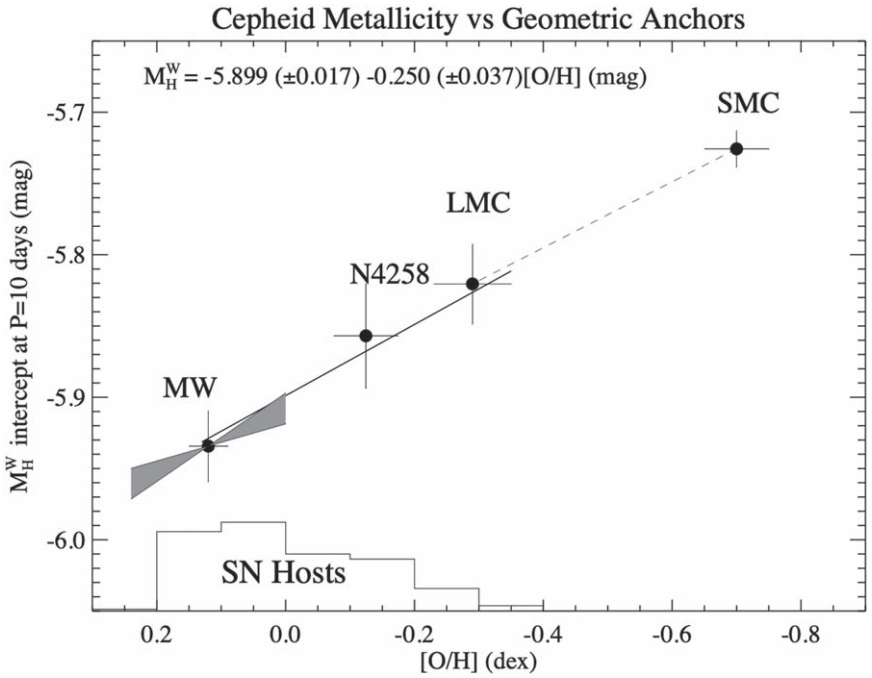

Figure 21. The metallicity term and consistency of the geometric anchors. The mean slope of four hosts with geometric distances is plotted against the intercept of the Wesenheit magnitude *P–L* relation. The DEB distance difference between the SMC and LMC from Graczyk et al. (2020) as discussed is independent of the calibration uncertainties of the DEB method, making this link, the dashed line, robust and independent of the other anchors (a linear fit to the four points is very good, though not unexpectedly so with $P = 10\%$ to be better). The gray constraint on the MW comes from the breadth of metallicities and individual Gaia EDR3 parallaxes. The metallicities of the Cepheids in SN hosts span the range of the anchors, making the value of H_0 insensitive to the value of the metallicity dependence, with a change in H_0 of 0.2 units for a change in Z_W of 0.1 mag dex^{-1}.

primary result was $\gamma = -0.13 \pm 0.07$ (with variants values ranging from -0.24 to -0.08) and in R19 it was -0.17 ± 0.06. The total χ^2 for the three anchors in 21 is 0.18, which for a line fit leaves one remaining degree of freedom, and the likelihood to find agreement this good or better is 33% and thus nominal but not surprising.

A number of recent developments have tightened this constraint considerably while broadening its range. The DEB distance for the SMC from Graczyk et al. (2020) as discussed in Section 4.4 provides a differential measurement between the Cepheids in the LMC and SMC, which constrains the metallicity dependence. As shown in Figure 21, this constraint alone gives $\gamma = -0.22 \pm 0.05$, similar to the values and uncertainties found by Breuval et al. (2021) for [Fe/H], and falling along the line that joins the other two anchors. It is one of the strongest constraints available because it comes from the difference in the DEB distances to each Cloud, a measure that has a small uncertainty owing to calibration cancellation and that does not depend on the comparison between the LMC and the other two anchors (Graczyk et al. 2020). In addition, the constraint internal to the MW Gaia EDR3 parallaxes alone (-0.22 ± 0.09) indicates a similar value.[19] The global fit also makes use of the internal metallicity gradients in the SN hosts to constrain Z_W. Individually, these are not constraining, with a median uncertainty greater than 1 and a minimum of 0.3. However, combined, these are supportive of the results from the local galaxies albeit less constraining, yielding -0.13 ± 0.11 mag dex^{-1} by combining their independent fits with uncertainties in both axes.

Abundance measurements for 68 of the 70 LMC Cepheids used here (Romaniello et al. 2021) show that they are consistent with a single value, and the lack of any measurable breadth in metallicities negates the ability to measure an abundance dependence internal to the LMC as claimed in prior analyses (Freedman & Madore 2011) and further discussed by Romaniello et al. (2021). Together, these developments provide a consistent result of a ~ -0.2 mag dex^{-1} metallicity dependence in the NIR that also provides accord among the anchors and is little changed by excluding any anchor as shown in Figure 20.

It is important to note that excluding knowledge of the geometric distance to an anchor (e.g., the DEB distance to the LMC) as we do in Fits 10–15 does not exclude the Cepheids in that anchor, which remain extremely valuable (at any distance) for constraining the global properties of Cepheids. Rather, when excluding knowledge of an anchor distance, we allow that distance to become a free parameter which may subsequently be compared to the external geometric estimate. In this case, a parameter such as the slope of the *P–L* remains constrained by the excluded anchor Cepheids because these Cepheids are only consistent with a single distance to the host with an accurate value of their slope.

6.3. Variants with Color and the Reddening Ratio, R

Here we explore variants in how the Cepheid color, $V - I$, is used in their distance determination. Our baseline analysis

[19] R21 derived -0.20 ± 0.13 from 66 Cepheids with HST photometry. Ripepi et al. (2021) used a larger ground-based sample of $N = 317$ fundamental and first-overtone pulsators to derive -0.37 ± 0.09 on the ground system with similar filters, steeper by $\sim 1.5\sigma$ than R21 but less applicable here owing to the presence of overtones and objects with low accuracy. Groenewegen (2018) used 205 MW Cepheids with Gaia DR2 parallaxes, somewhat less precise than those in EDR3, to derive a NIR Wesenheit abundance term of -0.204 ± 0.14. Here we identify 211 Cepheids with basic quality cuts, fundamental-mode only, $V - I < 2$ mag, $m_G > 6$ mag, $P > 3$ days, and Gaia GOF (goodness of fit) <10, transforming the ground magnitudes to the HST filter system to obtain -0.22 ± 0.09 with a slope of -3.29 as indicated in Figure 21.

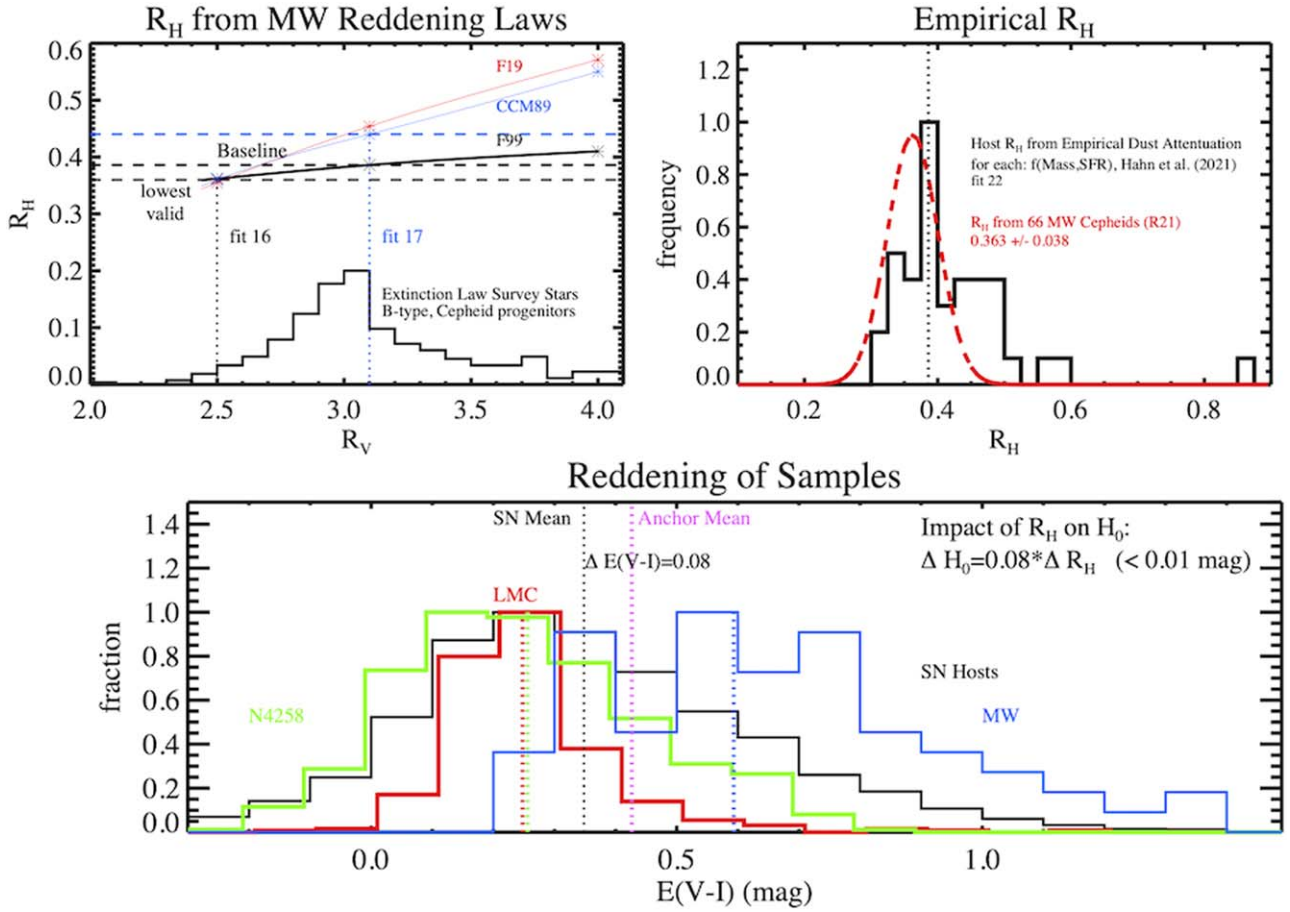

Figure 22. Reddening ratios and impact on H_0. The upper-left panel shows various reddening laws and their free parameter R_V used to estimate the appropriate ratio R_H for dereddening the Cepheids and the range over which this is applicable. The upper-right panel shows the empirical determination of R_H using the MW Cepheids from R21 and individual estimates parameterized by host, type, mass, and SFR from SDSS galaxies in Hahn et al. (2022). The lower panel shows the color excess, $E(V-I)$, for Cepheid samples. The net change to H_0 due to differential Cepheid reddening on the ladder is $0.08\,R_H$ (0.03) mag, with realistic differences in R_H producing changes <0.01 mag or <0.5%.

derives the Wesenheit parameter R (sometimes refereed to as R_H to avoid confusion with the optical reddening law parameter, R_V) from the Fitzpatrick (1999) law with reddening parameter $R_V = 3.3$, assuming MW-like reddening for the sample of late-type hosts and thus $R = 0.386$. Fits 16, 17, and 18 (respectively) change the reddening parameter to $R_V = 2.5$ or adopt different laws—Cardelli et al. (1989, with $R_V = 3.1$) or Nataf et al. (2016, appropriate for the inner halo). The relationship between R, R_V, and the reddening laws is shown in the upper-left panel of Figure 22. Fit 19 allows the value of R for all hosts to be a common but free parameter. The value derived from MW Cepheids for R is 0.36 ± 0.04 (Riess et al. 2021) and for the full sample of Cepheids here is 0.34 ± 0.02. These fits yield a very similar value for H_0; including R as a free parameter has little impact on the uncertainty.

If one wishes to allow for differences in the reddening law or the value of R across different hosts, it is necessary to first subtract the intrinsic color of Cepheids using their empirical period–color (P–C) relation in order to separate the component of the color that results from dust reddening; see Appendix D for details. Sandage et al. (2004) dereddened MW Cepheids to determine $\langle (V-I)_0 \rangle = 0.256 \log P + 0.497$. We used the LMC reddening maps from Skowron et al. (2021) to deredden the Cepheids in the catalog of Macri et al. (2015) to derive a P–C relation of $\langle (V-I)_0 \rangle = 0.238 \log P + 0.513$, very similar to the MW one. From the same data we derived the intrinsic P–C relation from the LMC Cepheids as above to be $\Delta m_H = 0.635(\pm 0.021)\ \Delta(V-I)$.

We use a mean relation of $\langle (V-I)_0 \rangle = 0.25 \log P + 0.50$ to first subtract the intrinsic color from all Cepheids using their periods, and then in Fit 20 we substitute the Wesenheit magnitudes in Equation (7) for $m_H - R\ E(V-I)$ with $R = 0.386$, which reduces H_0 by 0.1 km s^{-1} Mpc^{-1}. In Fit 21 we allow this reddening ratio to be a free parameter with little change from the result given in Fit 19.

We can also consider different values for the reddening ratio, R, for different hosts. One might derive these as the value that optimizes the relation between colors and magnitudes within each host. However, after subtracting the intrinsic P–C relation, the small residual color span, coupled with relatively large color measurement uncertainties, does not provide any meaningful constraint on the individual values of R for a given host beyond the few nearest galaxies as demonstrated in Appendix D. In addition, Appendix D shows that determining an unbiased estimate of R requires accounting for uncertainties in both axes (in this case color and brightness with the statistical issue discussed by Tremaine et al. 2002), and failing to do so leads to large underestimates of R and its uncertainty, as seen (for example) in Perivolaropoulos & Skara (2021) and Mortsell et al. (2021). We find typical uncertainties in individual values of R to be $\sim$1 and thus uninformative. Likewise, Follin & Knox (2018) also concluded that such color

data are "insufficient to make a completely data-driven inference on [individual] R," and they used a "wide prior" of 0.39 ± 0.1, finding distant hosts consistent with that prior and little impact on H_0. It is also crucial to recognize that the empirical reddening law is only valid for use at $R_V > 2.5$, as stated by Fitzpatrick et al. (2019), because sight lines with $R_V < 2$ are not seen among the hundreds of massive MW stars used to determine the law as shown in Figure 22 (indeed, the Rayleigh-scattering limit for absorption corresponds to $R_V \geqslant 1.5$). Thus, there is no empirical support for $R < 0.3$ from these laws or from the stars (significantly, of the same B type that later become Cepheids) that inform them.

An alternative approach with better grounding is to derive individual values of R for each host based on its specific properties and the empirically determined correlation of these properties with dust attenuation. We use the empirical dust attenuation (EDA) framework from Hahn et al. (2022), which derives individual extinction laws for hosts as a function of their mass, type, and star formation rate (SFR) as determined to best match colors to SDSS galaxy observations. These individual estimates of R are given in Table 6 and shown in Figure 22, and have a mean of 0.42 and dispersion of 0.10, matching well the prior used by Follin & Knox (2018). The smallest value for any host is 0.30 for NGC 7541, demonstrating that values lower than this become unrealistic as discussed further in Appendix D. Fit 22 uses these modeled values of R from Hahn et al. (2022) and results in an increase in H_0 of $0.8\,\mathrm{km\ s^{-1}\,Mpc^{-1}}$.

It is clear why varying the reddening ratio in hosts is not effective at changing H_0. The mean value of $E(V-I)$ in the SN hosts is 0.35 mag, as shown in Figure 22. For the Cepheids in the anchors MW (Gaia EDR3), LMC, and NGC 4258, it is 0.58, 0.23, and 0.24 mag (respectively), for a mean anchor value (weighted by the precision of the anchor distance) of 0.43 mag. Thus, the net difference in $E(V-I)$ between anchors and SN hosts is 0.08 mag, and hence the full impact on H_0 of correcting for Cepheid reddening is 0.03 mag or $\sim$1.5%, with perturbations to this procedure changing the result by a smaller amount. As shown in Figure 22, the strongest empirical evidence has the characteristic value of R less than 0.1 from the baseline; hence, e.g., $\Delta H_0 = 0.08 \Delta R$ (mag) <0.01 mag. In Fit 23 we discard the use of color altogether, representing a reasonable assumption that the Cepheid extinction at 1.6 μm is modest and to first approximation cancels along the distance ladder, and find that H_0 goes up by 2%.

6.4. Form of the P–L Relation

In Section 4.6 and Figure 10 we measure the slopes of the P–L relations of each host as well as the mean slope above and below $P = 10$ days, finding no clear evidence of a break. It is important to note that we have not included Cepheids with $P < 5$ days (either in the anchors or SN hosts) in our analyses. Such Cepheids could provide additional support for a break; however, Cepheids with overtone pulsations become much more common below this period, obey a different P–L relation, and may be confused with fundamental Cepheids.

However, in R16 the baseline was to allow for a break at $P = 10$ days, in accordance with prior claims of a break in the optical in the LMC (Sandage et al. 2004; although no evidence for one was found by R16). To allow for additional comparison to R16, in Fit 24 we allow for a break at $P = 10$ days with independent slopes above and below this pivot. The two slopes have a slope difference of 0.10 ± 0.05 as given in Table 5, which lowers H_0 by $0.45\,\mathrm{km\ s^{-1}\,Mpc^{-1}}$, with the difference largely driven by the ground-based samples from the LMC and SMC, whose Cepheids at lower periods have a shallower slope by 0.12 ± 0.08 and 0.34 ± 0.20 (respectively). To further explore the evidence for a break, we expand the sample of Cepheids in M31 by including those from Kodric et al. (2018, sample III) in Fit 28. Because in M31 the opposite occurs (shorter-period Cepheids have a steeper slope by 0.20 ± 0.12), the inclusion of both samples reduces the difference in slope to 0.07 ± 0.04 and H_0 is lower than the baseline by only $0.3\,\mathrm{km\ s^{-1}\,Mpc^{-1}}$. A free-form analysis of a slope change (allowing each host to have two individual slopes) reduces the sample evidence of a break to $<1\sigma$, with only three hosts (the LMC, SMC, and M31) providing any significant weight at $P < 10$ days (MW Cepheids with Gaia parallaxes provide no indication of a change in slope). The lack of significant evidence of a break (or change in slope) is why the baseline analysis uses the single-slope parameterization. Another option here is to exclude Cepheids with $P < 10$ days, which we do in Fit 25, resulting in an 0.1 unit increase in H_0.

6.5. M31 Variants

Fits 27 and 28 exchange the Li et al. (2021) sample of M31 Cepheids for the PHAT sample from Kodric et al. (2018) with a 10 fold increase in the number of variables, but with filters that are transformed to the set of three used elsewhere rather than directly observed with them. The main value of this change is to gain further traction of a possible break at $P = 10$ days as discussed in the prior section.

6.6. Variants of Metallicity

In Fit 29 we ignore the metallicity term and find that it has little impact, with H_0 rising by $0.3\,\mathrm{km\ s^{-1}\,Mpc^{-1}}$. The metallicity term has little impact on H_0, as shown in the marginal confidence region of Figure 20. The reason can be seen in Figure 21, where the distribution of Cepheid metallicities in SN hosts matches well the range of the three anchors; thus, the metallicity term does not move the SN hosts with respect to the mean of the anchors. It does, however, provide for consistency of the Cepheid and anchor distances.

Appendix C gives further details about the Cepheid metallicity scale we adopt, the average of nine recent and well-characterized relations (Teimoorinia et al. 2021). In Fit 30, we replace the use of the mean metallicity scale with one of these systems (Pettini & Pagel 2004)[20] based on [O III] and [N II] lines, which we show in Appendix C provides consistent metallicities for MW H II regions with those at the same radius derived from spectra of MW Cepheids.

6.7. Inclusion of TRGB

The TRGB offers an additional, independent rung between anchors and SN hosts (Freedman et al. 2019; Anand et al. 2021, and references therein). The inclusion of TRGB distances can in principle improve the calibration of SN Ia hosts and add to the number of such objects on the second rung. However, before including TRGB distances, it is important to determine if they

[20] The PP04 calibration was suggested by F. Bresolin (private communication, 2021) and indicated in Bresolin et al. (2016) to provide a good match to extragalactic-star spectral abundances, making it a good reference system.

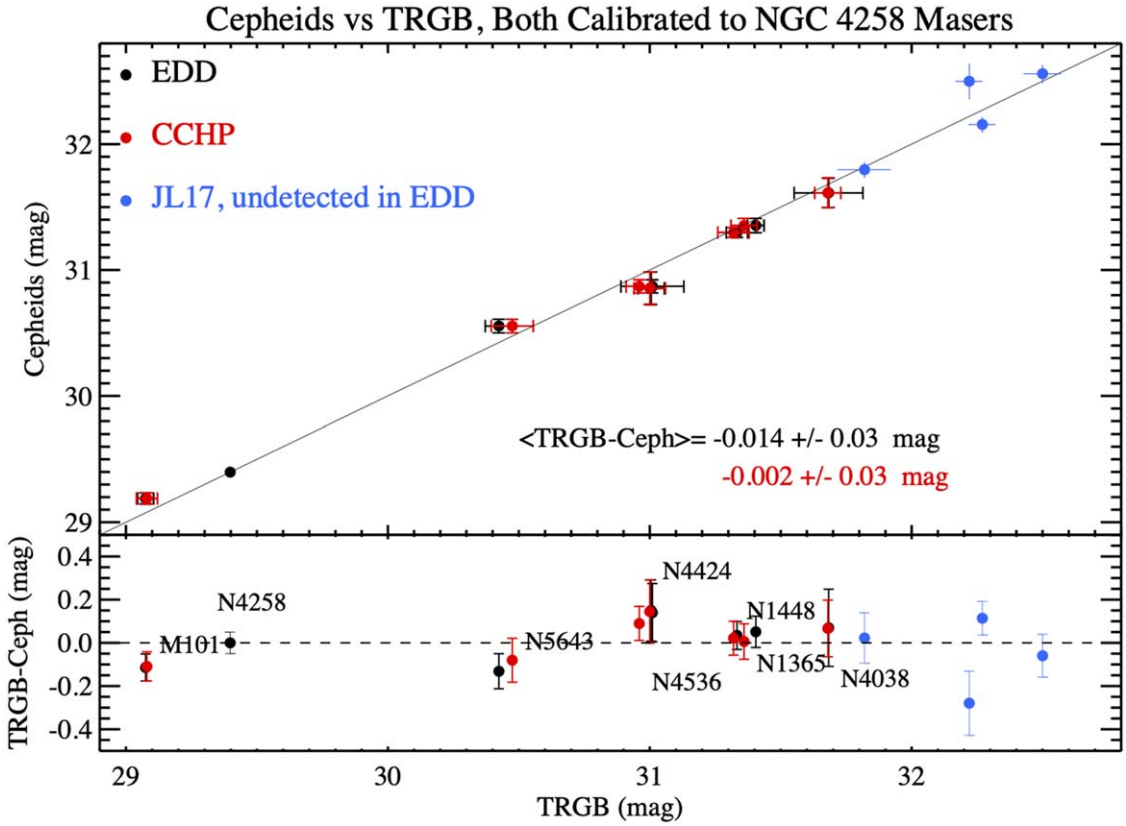

Figure 23. Cepheid distances from this work and TRGB distances from the CCHP (F21) and the EDD (Anand et al. 2021) groups to the same SN hosts calibrated by the same geometric anchor distance to NGC 4258, all measured with HST. The seven hosts indicated are all of those measured by these three groups. In blue we show four more-distant hosts measured by Jang & Lee (2017) for which evidence of a TRGB was not detected by Anand et al. (2021). The Cepheid and TRGB distances are consistent, allowing us to combine them (including for three additional TRGB hosts: NGC 1316, 1404, and 4526) simultaneously in Fits 31 and 32.

are consistent with those employed here for Cepheids.[21] We use the TRGB samples measured by the CCHP (Freedman 2021, hereafter F21) and EDD (Anand et al. 2021) groups.

To determine the consistency of the two distance indicators, we calibrate both using the same geometric anchor, NGC 4258, whose Cepheids and TRGB were both observed with HST, making this a purely differential comparison to a set of the same SN hosts. In Figure 23 we compare the distances from each method for all seven SN Ia hosts with Cepheid and TRGB measurements, which are available from both CCHP and EDD. The mean difference between Cepheids and TRGB is -0.014 ± 0.030 and 0.002 ± 0.030 mag for EDD and CCHP, respectively (we have not included the four more-distant SN Ia hosts measured by Jang & Lee 2017 and included in F19 because Anand et al. 2021 could not identify any reliable TRGB for these, but we plot them in Figure 23), and conclude these are consistent (see Section 7.2 for further discussion). For consistency with the preceding Cepheid-only analyses, we do not include SNe Ia with TRGB measures that do not pass the SN quality cuts employed above ($|c| < 0.15$ and $|x1| < 2$), which affects four objects (SNe 1989B and 1998bu with $c = 0.3$ corresponding to $A_V \approx 1$ mag, and SN 1981D with $c = 0.2$, also with high reddening; SN 2007on with $x1 = -2.2$). Ten TRGB distances for 11 SNe Ia are thus added: the 7 hosts with Cepheid distances (and 8 SNe) in Figure 23 (M101 and NGC 1365, 1448, 4038, 4424, 4536, and 5643) and 3 additional hosts without Cepheids with 4 additional SN calibrators (SNe 1980N and 2006dd in NGC 1316, SN 2011iv in NGC 1404, and SN 1994D in NGC 4526).

As we did for the Cepheids, we use TRGB distances where the anchor and SN hosts made use of the same telescope and instrument to negate telescope zero-points. In this case we first use Table 2 from Anand et al. (2021), which includes the calibration based on new observations of NGC 4258 from our program (GO 16198) that also employ the same filters (*F606W* and *F814W*), ACS electronics, and similar level of CTE to the SN hosts. We do not include TRGB zero-points in the Magellanic Clouds or the MW because these have not been measured directly on the HST system (owing to the impracticable area). For the joint analysis we include an additional parameter for the TRGB luminosity, M_I, which is optimized in the fit. We note that the TRGB constraint is included as available for an SN host simultaneously with the Cepheid constraint through the addition of the relation for the ith SN host,

$$m_{I,\mathrm{TRGB},i} = \mu_{0,i} + M_{I,\mathrm{TRGB}}, \tag{20}$$

and the calibrating relation

$$m_{I,\mathrm{N4258}} - \mu_{0,\mathrm{N4258}} = \Delta\mu_{\mathrm{N4258}} + M_{I,\mathrm{TRGB}}, \tag{21}$$

which adds a single free parameter, $M_{I,\mathrm{TRGB}}$.

Including TRGB (EDD), Fit 31 lowers H_0 by 0.3 units to 72.76 ± 0.95 km s^{-1} Mpc^{-1} and reduces the overall error by 5%, yielding a value of $M_{I,TRGB} = -4.003 \pm 0.025$ mag, similar to that found by Anand et al. (2021). The TRGB SN

[21] It is not possible to make an absolute statement of whether TRGB and Cepheid distances are consistent because this will depend on how each is calibrated and measured. Here we focus on a specific implementation of each.

host and Pantheon SN data yield 71.5 km s^{-1} Mpc^{-1} before the sample is tripled by combining with the SN-Cepheid hosts with the increase in H_0 of ∼1 km s^{-1} Mpc^{-1}, consistent with the shot noise of each subsample as discussed in detail in Section 7.2. We note that because the EDD TRGB parameterization includes a color dependence, the value of $M_{I,\mathrm{TRGB}}$ quoted here corresponds to their fiducial, blue TRGB with $V-I=1.23$ mag. For Fit 32 we replace the EDD TRGB measurements with the CCHP set as given by F19 and F21 for the same SN hosts and by Jang et al. (2021) for NGC 4258 TRGB[22] and without a color dependence to match the CCHP implementation. This reduces H_0 by 0.45 units from the EDD-based result (0.75 units below the baseline). The difference between the EDD- and CCHP-based result is a direct consequence of the 0.04 mag brighter measurement of the tip in NGC 4258 by Jang et al. (2021) compared to EDD, the significance of which is 1.9σ as further discussed in Section 7.2. The Jang et al. (2021) measurement in NGC 4258 in concert with the other constraints yields $M_{I,\mathrm{TRGB}}=-4.025\pm0.023$ mag for the above parameter. (We note that a direct comparison of the $M_{I,\mathrm{TRGB}}$ parameter to that observed for NGC 4258 requires adding to this $\Delta\mu_{\mathrm{N4258}}$ in Equation (2), which for the baseline is −0.013 mag.) The mean result from the two TRGB implementations is 72.53 ± 0.95 km s^{-1} Mpc^{-1}, which we adopt as a reference value for TRGB inclusion.

6.8. Hubble-flow SN Sample Variants

Here we explore fits utilizing different selection criteria for inclusion in the Hubble-flow sample. The goal of the selection of SNe and their hosts for the Hubble-flow sample is to match as well as possible the same criteria used to collect the calibrator sample to guard against the presence and imbalance between samples of additional host or SN characteristics currently known or unknown that may correlate with Hubble residuals. The enhanced size of the calibrator sample, now with 42 SNe Ia, reduces the likelihood of a chance imbalance of such properties.

The host selection for the baseline Hubble-flow samples requires visual identification from the best available optical imaging to be a spiral type in the full range of Sa–Sd, making them likely hosts of massive star formation in the last 0.1 Gyr —the primary criterion for our targeting them for Cepheid observations. This excludes highly inclined hosts (>75°) as these are complex targets and were not considered for finding Cepheids nearby. The host requirements are in addition to the SN Ia quality cuts, which are relatively tight in order to match the calibrators ($|c|\leqslant0.15$, $|x1|<2$), with the first observation earlier than 5 days after maximum light, distance measurement error of ⩽0.2 mag, with outliers $>3.5\sigma$ from the Hubble flow removed. The redshift range, as in R16, is $0.0233<z<0.15$. We refer to this sample as "LZSPI" (low-z, spiral). In Section 7.2 we will further analyze the host properties of these samples to confirm their balance in mass, SFR, and specific star formation rate. In Appendix A we compare the colors and light-curve shapes of the SN samples, confirming the similarity of the calibrator and selected Hubble-flow sample.

In Fit 33 we expand the Hubble-flow sample by removing any limitation on the host and including all SNe that pass the Pantheon+ quality cuts (which allow $|c|\leqslant0.3$ and $|x1|<3$), nearly doubling the sample to 482 SNe. It is noteworthy that there is no change in H_0, suggesting that after Pantheon+ standardization, SNe in different host types provide consistent distances. In Fit 34, we expand the sample to high redshift ($z<0.8$) and more than 1300 SNe, which raises H_0 by 0.5 km s^{-1} Mpc^{-1}. The use of such high-redshift SNe requires knowledge of q_0 as discussed in Sections 5.2 and 6.12. Fits 35 and 36 use these two expanded samples but exclude $z<0.06$ to circumvent concerns about the presence of a hypothesized local void structure. These raise H_0 by ∼0.3 km s^{-1} Mpc^{-1}.

6.9. Calibrator SN Sample Variants

Here we explore variations of the calibrator sample. A complete, volume-limited sample is desirable, as it has the simplest and most unbiased selection function. Such a sample is limited by the volume $z\leqslant0.011$ of suitable calibrators between 1980 and 2021 and can be selected by excluding two SNe Ia from the baseline sample at higher redshifts (SNe 1999dq and 2007A), with results given in Fit 38. Fit 39 contains the same with the added TRGB distance measures. Fit 40 limits the hosts to only high-mass galaxies, $\log(M/M_\odot)>10$.

In Section 7 we discuss tests of the Cepheid background. Here we undertake another test of the background with Fits 41 and 42, which divide the calibrator sample into halves with lower-than and higher-than-median background. The difference in H_0 is 0.0 km s^{-1} Mpc^{-1}, which is less than the independent shot noise of each half (1.3% or 0.9 units for each), and thus there is no indication of a difference in H_0 for high and low backgrounds. Fits 43 and 44 respectively use only the same 19 SNe from R16 or only the 23 SNe Ia added here, yielding a difference of 0.4 units. Fit 45 includes only the nearer half, defined from SNe as having $m_B<13$ mag (the median of the sample, corresponding to $D<28$ Mpc), as nearer hosts offer greater spatial resolution, lowering H_0 by 0.1 units.

6.10. Excluded SN Survey Variants

The SN standardized magnitudes have been drawn from the Pantheon+ sample, which is based on more than a dozen past SN surveys. The Pantheon+ sample recalibrates each survey photometrically to a common reference using the standard-star measurements in the fields of each SN to negate the impact of survey calibration errors. In addition, most SNe observed in the Cepheid host sample are matched by SNe in the Hubble-flow sample observed by the same SN survey with no one survey having a dominant share. As demonstrated by Brownsberger et al. (2021), "gray" photometric survey errors strongly and beneficially cancel by populating both samples with SNe from the same surveys. Brownsberger et al. (2021) find that survey miscalibration and incomplete cancellation would affect the H_0 measured from the survey mix used here at $\sigma=0.15$ km s^{-1} Mpc^{-1} even for extremely large survey zero-point errors of ∼0.1 mag, and

[22] The Jang et al. (2021) measurement in NGC 4258 used observations with a different filter, *F555W*, and different electronics than the *F606W* and the refurbished electronics used for all SN hosts and by Anand et al. (2021). More significantly, the state of CTE degradation on ACS was also markedly different between these observations of NGC 4258. In 2003–2005, a typical TRGB star in the halo would have suffered CTE losses of ∼0.04 mag. At the time of the SN-host observations in 2015–2019, such stars would have lost ∼0.14 mag, more similar to the observations analyzed by Anand et al. (2021). While pixel-based CTE rectification in the STScI pipeline attempts to account for such losses, the differential loss over the 15 yr would be ∼0.1 mag and the uncertainty in the correction would be a sizable fraction of that, likely a few hundredths of a magnitude. For this reason it is preferred to use data more closely spaced in time.

more likely 0.06 units for realistic survey calibration errors of 0.025 mag.

To further explore the sensitivity of H_0 to survey errors, Fits 46–54 present the results excluding each of the major surveys contributing to the sample. While none of these exclusions change the baseline H_0 by more than ∼0.3 units, we note that excluding the CSP sample raises H_0 by 0.3 units, with most of this (0.2) seen in the change in the Hubble intercept (a_B). Using a Hubble-flow sample exclusively from the CSP (while using a mix of surveys for the calibrator sample) yields a lower value of H_0 by 0.5 units. While this is consistent with the sample shot noise, it will reduce H_0 for the CCHP results, which use only the CSP sample for the Hubble flow. Brownsberger et al. (2021) find that this sample asymmetry is expected to produce errors of 0.8 units for realistic calibration errors of 0.025 mag; it is discussed further in Section 7.2.

6.11. SN-fitting Variants

Fit 55 changes the way the intrinsic scatter of SN colors is modeled from an empirical approach that includes both a component intrinsic to SNe Ia and another due to host dust as given by Brout & Scolnic (2021) and is further described in the Pantheon+ sample (Brout et al. 2021; Scolnic et al. 2021). It is a simpler description, where the intrinsic scatter of SNe Ia is monochromatic (i.e., dispersion only in the luminosity, not the color) from Guy et al. (2010) and used in the JLA analysis (Betoule et al. 2014) and in the first Pantheon compilation (Scolnic et al. 2018). This fit raises H_0 by 0.3 units. The calibrator sample has no preference for either method, yielding similar dispersion between SN and Cepheid distances.

6.12. Velocity Variants

Here we provide variations related to values of the redshifts as implemented in Equation (4), which benefit from the combined and improved values from Pantheon+ (Carr et al. 2021). Peterson et al. (2021) provide a comprehensive overview and comparative analyses of various predictions of empirical cosmic flows (or peculiar velocities). They found important improvements to the Hubble-flow residuals by (1) replacing SN-host redshifts with their host-galaxy group redshift (when available), and (2) using local density maps to account for motions induced by local gravity. The latter are provided by constrained realizations of the peculiar-velocity field by Carrick et al. (2015), Said et al. (2020), and Lilow & Nusser (2021) based on 2M++ (Lavaux & Hudson 2011) and 2MRS (Huchra et al. 2012; Macri et al. 2019), respectively. The baseline included both the group-redshift replacement and the 2M++ corrections from Peterson et al. (2021). Fit 56 exchanges the 2M++ values for those from 2MRS, which provide a comparable improvement in residuals and lower H_0 by 0.05 units. A noteworthy change is seen in Fit 57, which forgoes the flow corrections and leaves the redshifts in the CMB frame, reducing H_0 by 0.5 units. However, as shown by Peterson et al. (2021) for 585 SNe Ia with $z < 0.08$, the tightening of the Hubble diagram (from $\sigma = 0.17$ mag to <0.15 mag, or a decrease in χ^2 of 100) gives evidence in favor of these corrections which is too strong to ignore. Further, the increase in H_0 that comes with the decrease in residuals runs counter to the hypothesis that the SN sample lives inside a large-scale void that artificially raises the local value of H_0 (Kenworthy et al. 2019).

While peculiar flows cause a small perturbation in H_0 measured from SNe Ia at $0.0233 < z < 0.15$, they would produce a greater uncertainty if we forgo the use of SNe and measure H_0 directly from only the first two rungs—that is, from the Cepheid-host redshifts (which are not used in the three-rung distance ladder). The baseline sample host redshifts have a mean of $cz = 2000$ km s^{-1}, with many <1000 km s^{-1}. Elsewhere (W. D. Kenworthy et al. 2022, in preparation), we present an analysis of H_0 from this two-rung ladder, which importantly accounts for the spatial covariance of the local peculiar flows, largely limiting the available precision from this route to 3%–4% and demonstrating the value of SNe Ia for the third rung for measuring H_0.

Fits 58 and 59 raise q_0 from −0.55 to −0.52 (equivalent to raising Ω_M in flat ΛCDM from 0.30 to 0.32) for either the local sample of spiral hosts or the sample with all hosts and for $z < 0.8$, with little impact on H_0 relative to these samples at $q_0 = -0.55$. In Section 5 we considered a free-form fit for H(z) using q_0 as a free parameter simultaneous to the determination of H_0.

6.13. Optical Wesenheit Variants

Fits 60–67 use an optical-only Wesenheit magnitude, substituting for Equation (7), $m_I^W = m_I - R(V - I)$, and thus discarding the NIR observations. The Fitzpatrick (1999) reddening law with $R_V = 3.3$ yields $R = 1.19$ in the HST passband system ($m_V = F555W$, $m_I = F814W$). The optical Wesenheit has the advantage of lower "sky" backgrounds (and their fluctuations) but the disadvantage of higher reddening (and sensitivity to the form of the reddening law). The baseline fit with the optical Wesenheit yields 72.70 ± 1.07 km s^{-1} Mpc^{-1}, similar to the baseline fit. However, the optical Wesenheit is somewhat noisier when compared to the SN distances with a relative dispersion of 0.16 mag (versus ∼0.13 mag with the NIR data). We also see larger variations in the anchors and the color variants as seen in Figure 18, with the scatter among optical-based variants that is three times greater than the NIR-based results and comparable to the statistical uncertainties. This illustrates the rationale by the SH0ES program for pursuing NIR observations for Cepheids.

Both of these differences are expected consequences of variations in the reddening law in the optical. For example, for hosts whose Cepheids have a mean $E(V - I) = 0.4$ mag, a difference between a Fitzpatrick (1999) reddening law with $R_V = 2.5$ and $R_V = 3.3$ causes a difference in distance of only 0.01 mag for the NIR Wesenheit but 0.09 mag for the optical one, which can explain the aforementioned noise. In R16 we concluded that future improvements must rely on NIR data until additional studies of variations in reddening laws in the optical were available. The situation has not improved in that regard. While optical-only Wesenheit data have yielded similar values of H_0 (Freedman et al. 2012; Riess et al. 2016), their larger systematic uncertainties make them unsuitable to pursue the percent-level determination of H_0 we approach, and their further analysis is not pursued here. Nevertheless, none of the optical Wesenheit fits result in a noteworthy change to H_0.

6.14. Summary of Variants and Error Budget

Our baseline determination of H_0 lies 0.2 km s^{-1} Mpc^{-1} (20% of the uncertainty) below the median of all analysis

Table 7
H_0 Error Budgets (%), Terms Approximated from the Global Fit

Term	Description	Riess+ (2016)			Riess+ (2019)			This work		
		LMC	MW	4258	LMC	MW	4258	LMC	MW	4258
$\sigma_{\mu,\mathrm{anchor}}$	Anchor distance	2.1	2.1	2.6	1.2	1.5	2.6	1.2	1.0[a]	1.5[b]
$\sigma_{\mathrm{PL,anchor}}$	Mean of P–L in anchor	0.1	...	1.5	0.4	...	1.5	0.4	...	1.0
$R\sigma_{\lambda,1,2}$	zero-points, anchor-to-hosts	1.4	1.4	0.0	0.1	0.7	0.0	0.1	0.1[a]	0.0
σ_Z	Cepheid metallicity, anchor-hosts	0.8	0.2	0.2	0.9	0.2	0.2	0.5	0.15	0.15
	subtotal per anchor	2.6	2.5	3.0	1.5	1.7	3.0	1.4	1.0	1.8
All Anchor subtotal			1.6			1.0			0.7	
$\sigma_{\mathrm{PL}}/\sqrt{n}$	Mean of P–L in SN Ia hosts		0.4			0.4			0.4	
$\sigma_{\mathrm{SN}}/\sqrt{n}$	Mean of SN Ia calibrators (# SN)		1.3 (19)			1.3 (19)			0.9 (42–46)	
σ_{m-z}	SN Ia m–z relation		0.4			0.4			0.4	
σ_{PL}	P–L slope, Δ log P, anchor-hosts		0.6			0.3			0.3	
statistical error, $\sigma_{\mathrm{H_0}}$			2.2			1.8			1.3	
Analysis systematics[c]			0.8			0.6			0.3	
Total uncertainty on $\sigma_{\mathrm{H_0}}$ [%]			2.4			1.9			1.35	

Notes.
[a] Riess et al. (2021).
[b] Reid et al. (2019).
[c] Uncertainties labeled in past analyses as "systematics" related to the metallicity scale, Cepheid background/crowding corrections, and SN systematics are formally included here in the covariance matrix in Figure 11 and thus propagate there as part of the complete uncertainty. Following past work, we measure the remaining systematic errors as the standard deviation of analysis variants presented in each work as in the dispersion in Figure 18 and as discussed in Section 6.14. All terms here are approximations derived from the global fit.

variants, indicating that it is a good proxy for the set. In R16 we measured the dispersion of 23 variants and identified that as a systematic error. In this analysis we have moved previous sources of systematic uncertainty into the covariance matrix to include them formally, and thus most of the variants presented here were intended to gauge sensitivities in the analysis (e.g., excluding a data source) rather than true uncertainties. Nevertheless, we measure the dispersion of the NIR variants (see Figure 18) around a 3σ-clipped mean to be 0.3 km s^{-1} Mpc^{-1} and conservatively add this in quadrature as characterizing additional systematic uncertainties to yield a full uncertainty in H_0 of 1.04 km s^{-1} Mpc^{-1}, or 1.4%.

None of the variants appear to offer a particularly promising route to solving the Hubble tension, with none shifting the value of H_0 much below the full error interval. The lowest value of H_0 comes from Fit 65, 71.93 ± 1.19 km s^{-1} Mpc^{-1}, from discarding the NIR Cepheid data and two anchors (leaving only the Milky Way); the highest is from Fit 22, 74.85 ± 2.33 km s^{-1} Mpc^{-1} and comes from discarding the use of Cepheid colors to account for extinction. Both of these fits represent suboptimal accounting of dust.

Table 7 and Figure 24 present the full error budget derived (in approximate form) from the global fit and in comparison to prior analyses from the SH0ES team. The fractional reduction presented here is the largest seen since that between Riess et al. (2011) and Riess et al. (2016).

7. Discussion

Here we address a number of considerations using the preceding fits.

7.1. Accuracy of Cepheid Photometry

The accuracy of Cepheid photometry is important to the measurement of H_0. The dependence of photometry on calibration has been negated by the use of the same photometric system throughout the above measurements. Here we review a number of tests of the relative accuracy for extragalactic Cepheids.

1. Replication of PSF photometry by others: Measurements of NIR PSF Cepheid photometry using the same raw pixels but different software and methods have been directly compared to those used here by Javanmardi et al. (2021) in NGC 5584 and by Yuan et al. (2021) in NGC 4051, both finding good agreement with reported differences in distance moduli of 0.024 ± 0.046 and 0.00 ± 0.04 mag (respectively).[23] While this is short of a full replication of all hosts, it is sufficient to exclude a large methodological error in Cepheid photometry as a primary source of the $\sim$0.2 mag H_0 tension.
2. Replication of PSF photometry with apertures: In Appendix B we use aperture photometry, an independent method that is simple, highly reproducible, and accurate, albeit less precise than the standard approach of using PSFs to model photometry in crowded fields. For this validation, we employ aperture photometry for which the

[23] Comparing the Cepheids in Table 3 in NGC 5584 in common with Javanmardi et al. (2021) but revised here since R16, we find they are consistent in the mean with a difference in the Wesenheit of 0.07 ± 0.05 mag; their results are brighter, which if applied to all SN hosts would raise H_0 by 2 km s^{-1} Mpc^{-1}.

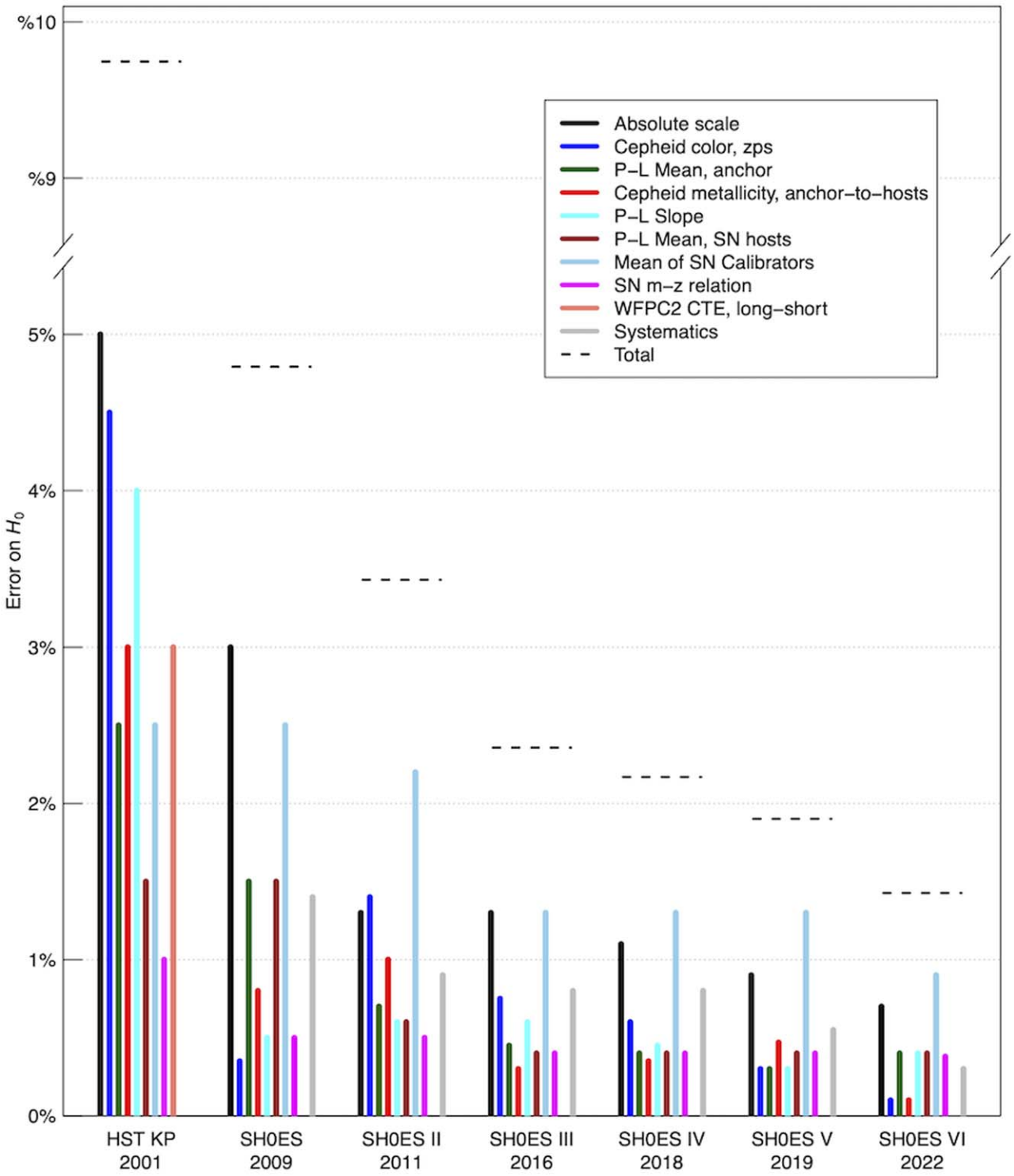

Figure 24. Comprehensive error budgets for six iterations of the SH0ES measurement of H_0 including this one (2022) and the Key Project from 2001 (Freedman et al. 2001). The greatest improvement here is realized by the increase in the number of SN calibrators, to decrease all terms to <1%. The combined error is indicated by horizontal dashed lines.

background is measured from the mean pixel value (not the mode of the background pixels) in an annulus centered on the Cepheids. This approach does not depend on artificial-star measurements to determine the variable background because the pixels in the annulus include the mean source contribution to the background. We compare the photometry and find good agreement in their means with a difference (PSF minus aperture) of 0.008 ± 0.010 mag for the Cepheids in SN Ia hosts and 0.002 ± 0.030 mag for the Cepheids in NGC 4258.

There are a number of strong tests of the accuracy of the background estimates presented here.

1. The Cepheids in NGC 4258 have a similar mean level of crowded backgrounds as in the SN Ia hosts (see Figure B2), almost fully negating a systematic underestimate or

overestimate of the background on the determination of H_0 when NGC 4258 is the sole anchor. The crowded background is similar because although NGC 4258 is three to four times closer than the mean SN host, its Cepheids have been mined from fields that are closer to the dense center by a similar factor. This results in $H_0 = 72.51 \pm 1.54$ km s^{-1} Mpc^{-1}, similar to the baseline result.

2. In Figure 8, we compare the P–L relations of Cepheids in the dense, inner (high background) and sparse, outer region (low background) of NGC 4258, finding a negligible difference of 0.01 mag.
3. In Appendix B we regress the background with the distance-ladder fit residuals and find a dependence of 0.010 ± 0.014 mag per magnitude of source background (in the sense of overestimating the background and consistent with no misestimate). The background misestimate trends required to explain the tension are strongly excluded as shown in Figure B2.
4. The optical background is nearly an order of magnitude smaller than that in the NIR owing to the higher resolution, smaller pixels, and lower flux from red giants. The baseline results are consistent with those from the optical Wesenheit, which do not include the NIR data, $H_0 = 72.70 \pm 1.30$ km s^{-1} Mpc^{-1}.
5. There is no significant trend with distance and distance difference between Cepheids and SNe Ia (zero at $<1.5\sigma$); see Figures 12 and 13.
6. Splitting the calibrator sample by background (Fits 41 and 42) yields no difference in H_0. Splitting in distance (determined by the SN Ia, Fit 45) yields a difference of 0.1 units. Both are consistent with no trend based on the shot noise of half the sample (1.0 units).

Finally, external to this paper, an additional and unavoidable consequence of the miscalibration of the background, independent of Cepheid mean flux, would be a change in apparent light-curve amplitude. Riess et al. (2020) compared the NIR amplitudes of Cepheids in three SN hosts and in the MW, found them to be consistent, and provided a quantitative limit of any misestimate of background to be 0.03 mag.

7.2. Consistency of TRGB and Cepheid Distance Scales

In Figure 23 we presented a comparison of distances measured with Cepheids and TRGB to seven SN Ia hosts—the set that allows for a purely differential and direct comparison by employing the same geometric calibration source (NGC 4258), with data in both host and calibrator obtained with the same telescope (HST) and setup to negate zero-point and geometric calibration errors. This comparison further employed the two most widely used methods for measuring the TRGB, edge-detection (F19) and luminosity-function fit (EDD). These Cepheids and TRGB measures are consistent with each other, with a mean difference of -0.002 ± 0.03 mag (CCHP; Freedman 2021) and -0.014 ± 0.03 mag (EDD; Anand et al. 2021). With no mean difference between methods as a starting point, we explore the broader question of the sources of the difference in the value of H_0 from distance ladders measured through the use of either method.

In Table 8 we expand on the results from Anand et al. (2021), who compared the value of H_0 from EDD, 71.5 km s^{-1} Mpc^{-1}, and CCHP, 69.8 km s^{-1} Mpc^{-1}, derived using the same images to measure the TRGB and the same geometric calibration source, NGC 4258. For consistency with the comparison presented by Anand et al. (2021) of EDD and CCHP, which was provided in units of magnitudes (for differences in 5 log ΔH_0), we retain this unit below and in Table 8 in which a difference of 0.03 mag corresponds to a difference of $\sim$1.0 km s^{-1} Mpc^{-1}.

Table 8
Sources of Differences in H_0 between EDD, CCHP, and SH0ES (in Magnitudes)

Term	ΔF19	ΔF21	ΔEDD
EDD versus CCHP TRGB			
1. Zero-point (NGC 4258)	0.06	0.06	⋯
2. No TRGB detected by EDD (four hosts)	0.00	0.00	⋯
3. NGC 1404	0.01	0.01	⋯
4. NGC 5643*	0.01	⋯	⋯
5. $\Delta m_{\rm TRGB}$*	−0.03	−0.02	⋯
TRGB subtotal	0.05	0.05	⋯
SH0ES Cepheids versus EDD TRGB			
6. +SN 2021pit	−0.02	−0.02	−0.02
7. −SN 2007on	−0.01	−0.01	−0.01
8. −3 red SNe Ia	0.00	0.00	0.00
9. +33 SNe Ia in 30 hosts	0.06	0.06	0.06
Calibrator set change subtotal	0.03	0.03	0.03
10. LMC+MW anchors	0.02	0.02	0.02
11. Flows	0–0.02	0–0.02	0–0.02
12. Hubble-flow surveys	0–0.02	0–0.02	0–0.02
Total	∼0.10	∼0.10	∼0.05

Notes. "*" weighted by SN, not by host. ΔF19, ΔF21 = differences between Freedman et al. (2019) or Freedman (2021) and EDD, respectively. ΔEDD = differences between EDD and SH0ES. Descriptions of individual entries: (1) measured zero-point calibration of the TRGB in NGC 4258; (2) Anand et al. (2021) did not detect the TRGB in four SN-host galaxies; (3) inclusion of NGC 1404 (not directly measured by F19; EDD include SN 2007on which F19 included and F21 excluded); (4) inclusion of NGC 5643 (not available in F19); (5) mean difference in measured values for the TRGB of the remaining hosts; (6) SN 2021pit not available in EDD; (7) exclude SN 2007on in NGC 1404 owing to $x1 < -2$; (8) exclude 3 SNe with $c > 0.15$; (9) addition of 33 SNe Ia; (10) addition of LMC and MW anchors; (11) corrections for cosmic flows in Pantheon+ raises H_0 by 0.02 mag relative to CSP SNe (see Peterson et al. 2021); (12) EDD used both Pantheon or CSP, the latter lowers H_0 by 0.02 mag relative to Pantheon+ average of surveys. The uncertainties in all terms above are 0.01 mag.

The largest contribution (0.04 mag of the full 0.05 mag difference) arises from the difference in the calibration of the TRGB measured in NGC 4258 applicable to (i.e., in relation to) the mean SN host, where CCHP derived $M_{I,\rm TRGB} = -4.05$ mag for a blue TRGB and EDD derived $M_{I,\rm TRGB} = -4.01$ mag, color-corrected to the fiducial, blue TRGB with $F606W - F814W = 1.2$ mag. Although additional sources in the LMC and MW have been used to support the TRGB calibration by F21, we use NGC 4258 as the reference here because it is the only source where the TRGB calibration is available directly on the HST system measured in a manner consistent with the TRGB in SN hosts.[24] The difference in the EDD and CCHP measurement of

[24] Although we do not recommend mixing photometric systems, to demonstrate the consequence of including the ground-based zero-points of TRGB as well as the Jang et al. (2021) measure of the tip of NGC 4258 we can add a strong prior of $M_{I,\rm TRGB} = -4.05 \pm 0.02$ mag to Fit 31, in which case we find $H_0 = 72.00 \pm 0.86$ km s^{-1} Mpc^{-1}, very close to the mean of the baseline here and the F21 value of H_0, simply weighted by the size of each SN sample that passes quality cuts in each set, and a posterior result of $M_{I,\rm TRGB} = -4.040 \pm 0.015$ mag.

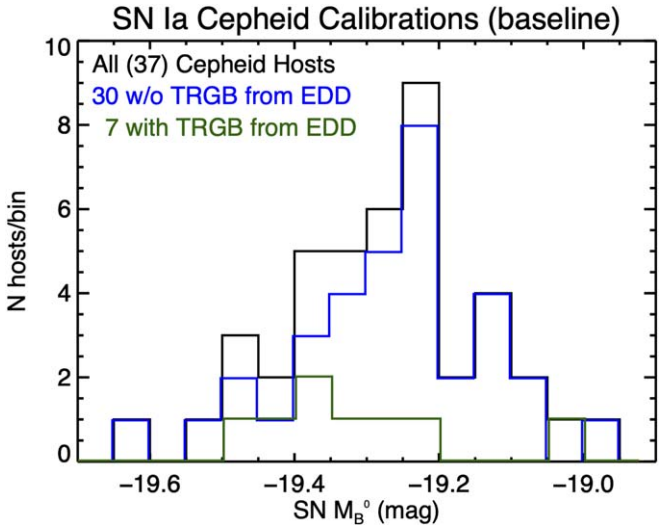

Figure 25. Histogram of the 42 SN Ia absolute magnitudes of SNe Ia calibrated from Cepheids in 37 hosts. The 9 SNe Ia in 7 hosts for which a TRGB distance is available from both EDD and CCHP are shown in green (calibrated here with Cepheids), and the 33 SNe Ia in 30 hosts without a TRGB measure from both in blue (calibrated here with Cepheids). The difference in their means of 0.08 ± 0.05 mag is consistent with the shot noise of the SN samples as discussed in Section 7.2, and combined with other differences between the TRGB calibrator set (lines 6–8) produce a net difference in H_0 of 1 km s^{-1} Mpc^{-1} (0.03 ± 0.05 mag) as shown in Table 8, which is less than 1σ from the shot noise of the two samples.

the tip in NGC 4258 is persistent, having been found for two different fields of NGC 4258, being significant at the $\sim 2\sigma$ level and is readily evident in Figure 3 of Anand et al. (2021), where the apparent location of the TRGB edge of $m_{F814W} = 25.372 \pm 0.014$ mag given by Jang et al. (2021) appears much brighter than the edge highlighted at 25.43 ± 0.025 mag. A comparison of each group's color–magnitude diagrams (i.e., photometry) might identify the cause of the difference, but these data are only available from Anand et al. (2021). Nevertheless, we offer the option for either TRGB result in Fits 31 and 32.

The second half of Table 8 provides the sources of difference between the EDD TRGB analysis (based on 16 SNe Ia) for which $H_0 = 71.5$ km s^{-1} Mpc^{-1} and our baseline analysis with TRGB (based on 46 SNe Ia), starting with the same geometric calibrator (NGC 4258). The primary change is in the makeup of the Cepheid and TRGB study calibrator samples with 34 SNe Ia added here and 4 subtracted, a net change of 30 objects producing a net increase in H_0 of 0.03 mag or 1 km s^{-1} Mpc^{-1}. The first three rows break out changes related to calibrators available from both TRGB analyses, which together lower H_0 by 0.03 mag. The first change is due to the addition of SN 2021pit, which recently appeared in NGC 1448 and is added here. SN 2007on has a light-curve shape ($x1 = -2.2$), which falls outside the quality range of $|x1| < 2$ imposed in our baseline analysis and is excluded. There are three SNe Ia (SNe 1981D, 1989B, and 1998bu) used by CCHP and EDD that are redder ($c > 0.15$ and $A_V > 0.5$ mag) than our baseline quality range cut and have not been included in this or any past SH0ES analyses. The next row adds 33 SNe Ia in 30 Cepheid hosts, tripling the sample from the preceding line's 13 SNe Ia in 10 hosts (with EDD and CCHP TRGB measurements and not excluded by SN quality cuts) to 46 SNe Ia in 40 hosts (for Fit 30, which includes TRGB-only hosts). This step raises H_0 by 0.06 mag or 2 km s^{-1} Mpc^{-1}. This difference is fully seen internal to the set of Cepheid-only SN Ia calibrations (i.e., independent of absolute anchors or TRGB distances) in the baseline fit as shown in Figure 25, which compares the two sets of SNe as measured only by Cepheid*s*. The 9 SNe in 7 EDD TRGB hosts are brighter than the 33 SNe Ia in 30 hosts without EDD and CCHP TRGB by 0.08 ± 0.05 mag as measured by Cepheids, which is consistent at 1.6σ with the sampling noise of the two sets (0.04 mag and 0.02 mag random error, respectively). As seen in Fits 38–45, this difference does not correlate with potentially relevant changes to the calibrator sample including splitting in distance, background level, or newness of the measurements, so we conclude that this difference of 1.6σ is a not-unexpected result of the increase in sample statistics.

Combining all of the additions and subtractions to the calibrator sample, the net change between the calibrator samples with both TRGB measures versus the baseline+TRGB here is 0.03 mag and the combined result of all membership changes is less than 1σ from the shot noise of the two samples. We note that if we use only the (average) TRGB results of the 13 good-quality SNe, discarding 33 good-quality SNe with only Cepheid measurements, we get 71.0 ± 2.5 km s^{-1} Mpc^{-1} as expected (near the mean of CCHP and EDD), but this uses less than a third of the SN sample and has a higher shot noise, while still consistent with the baseline.

As the sample size from both methods increases, we would expect them to regress to an increasingly similar mean. Including the two other anchors (MW and LMC) raises H_0 by 0.02 mag (and results in $M_{I,\mathrm{TRGB}} = -4.00 \pm 0.024$ mag), though each of the three anchors is consistent with the others at the 0.5σ level, as shown in Figure 19.

Finally, there are two developments based on recent analyses of SNe in the Hubble flow, which yield an increase in the value of H_0 relative to the CCHP TRGB analysis in F19 and F21, independent of the calibrator sample. The CCHP measurement relies on a calibrator SN sample from several SN surveys, but selects a Hubble-flow SN sample from only the CSP SN survey. Regardless of whether the CSP survey is better calibrated than others in an absolute sense, Brownsberger et al. (2021) show that the use of similar surveys for both SN samples reduces errors in H_0 arising from survey miscalibration to 0.1 km s^{-1} Mpc^{-1} owing to error cancellation, but to only 0.8 km s^{-1} Mpc^{-1} for the CCHP reliance on one survey compared to many. Scolnic et al. (2021) find that the CSP sample measures fainter compared to the mean of all other surveys by 0.025 mag, which matches what we find when we remove CSP as shown by Fit 47 in Table 5, where H_0 then increases by 0.4 km s^{-1} Mpc^{-1}. With only 5 of 19 calibrators in F19 observed with the CSP survey, the incomplete cancellation of this survey difference would be 0.6 km s^{-1} Mpc^{-1}. In addition, the CCHP analysis of SNe Ia does not account for cosmic flows expected from local density maps, the best of which (2M++ or 2MRS) as shown by Peterson et al. (2021) reduces the Hubble diagram χ^2 for $\sim$500 SNe Ia by $>$50 units, decreases the scatter, and raises H_0 by $\sim$0.5 km s^{-1} Mpc^{-1}. These two terms at the bottom of Table 8 should be included in a direct comparison between the results here and those from CCHP. The combination of both balancing SN survey errors in both samples and accounting for peculiar velocities would raise the CCHP value by 1.1 units and is hard to ignore. As expected, if we use only CSP for the Hubble flow and also do not account

Table 9
Host Properties of SN Ia Samples

Sample	N	$\mathrm{Log}(M/M_\odot)$	Log SFR	Log sSFR
Baseline Calibrator	37	10.3 SD = 0.6	0.3 SD = 0.5	−9.8 SD = 0.7
LZSPI HF	276	9.8 SD = 0.8	0.3 SD = 0.7	−9.3 SD = 0.4
LZBRD HF	482	10.1 SD = 0.8	0.1 SD = 0.9	−9.6 SD = 0.7
HZBRD HF	1354	N/A	N/A	N/A
LZSPI HF $\log(M/M_\odot) > 10$	132	10.4 SD = 0.5	0.4 SD = 0.7	−9.7 SD = 0.6

for cosmic flows, the baseline H_0 goes down by 1.5 km s^{-1} Mpc^{-1}.

To summarize, we find that TRGB and Cepheids give consistent distances when using the same anchors and consistent procedures, so that comparisons are meaningful. There are two main sources of the difference in H_0 between the CCHP implementation in F21 and the baseline here. (1) As given by Anand et al. (2021), a (net) difference of 0.04 mag between the EDD and CCHP implementation of TRGB in F21 arises from the difference in the apparent location of the tip in NGC 4258 and is the source of the difference between $H_0 = 69.8$ and 71.5 km s^{-1} Mpc^{-1} (i.e., TRGB-only results) in these two studies. (2) The other, corresponding to a net increase of 1 km s^{-1} Mpc^{-1} (0.03 mag) between the EDD TRGB and the baseline, is due to differences in the SN Ia calibrator samples including a reduction of 1 km s^{-1} Mpc^{-1} owing to our exclusion of 4 SNe Ia that fail quality cuts and an increase of 2 km s^{-1} Mpc^{-1} owing to the fainter mean seen for 30 SNe in 33 Cepheid-only hosts compared to 9 SNe in 7 TRGB+Cepheid hosts, a difference seen internal to Cepheid measurements and consistent with a statistical fluctuation due to the combined changes at the 1σ level (see Figure 25). We expect that additional TRGB data will result in these distributions agreeing, as we see no reason for a difference besides shot noise. Relative to F21, the sample difference change is 1.3 km s^{-1} Mpc^{-1} (0.04 mag) with other differences from the baseline owing to the absence of an accounting for peculiar flows as described by Peterson et al. (2021) and not including SNe from multiple surveys in the Hubble flow to cancel zero-point differences among the calibrator surveys as described by Brownsberger et al. (2021). Combined, these raise H_0 by ∼1.1 km s^{-1} Mpc^{-1}, and we see no reason not to include these in an SN Ia–derived measurement of H_0 given their strong empirical support. The combined Fit 30 yields 72.76 ± 0.95 km s^{-1} Mpc^{-1} for EDD and 72.29 ± 0.94 km s^{-1} Mpc^{-1} for CCHP, and we cite the mean of the two (72.53 ± 0.95 km s^{-1} Mpc^{-1}) as representative of the combination of Cepheids and TRGB.

7.3. Consistency of SNe Ia on Second and Third Rungs

In Appendix A we show that the calibrator SNe Ia are spectroscopically all prototypical and photometrically have a distribution of light-curve shapes and colors that are well matched to the selected Hubble-flow sample. The use of tighter-than-typical quality cuts ($|c| < 0.15$, $|x1| < 2$) ensures that the sample comparison is insensitive to the standardization method, as it makes little difference in the sample means.

In R11 the Hubble-flow sample consisted of SNe Ia without limitations placed on the properties of their hosts (with about two-thirds coming from spirals). Rigault et al. (2015) suggested that the calibrator sample, all with spiral hosts and thus greater mean SFR, could introduce a bias in H_0 if Hubble-flow residuals (even after accounting for an empirical host-mass dependence as done by R11 and R16) presented a residual correlation with host SFR (either the global rate of the host or local to the SN). R16 addressed this sample difference by including as an analysis variant a Hubble-flow sample composed of only spiral or globally star-forming hosts, a precaution (regardless of whether such a correlation exists) conservatively adopted here as the baseline. Because the size of the Hubble-flow sample is so much larger than the calibrator sample, it is sensible to cull the former to match the selection of the latter to control even yet-undiscovered systematics with little cost to the precision of H_0. It is also important to recognize that the size of a correlation of Hubble residuals with a host property depends on the method of SN standardization and the SN sample host selection and that it is quite possible for a specific combination of method and sample to show a significant correlation that does not exist for different methods and samples.

Jones et al. (2018) measured the correlation of the local and global SFR and specific SFR (sSFR) with the Hubble residuals from Pantheon SNe (Scolnic et al. 2018) used by R16 and found little or no significant correlation with implied corrections (if significant) at the 0.3 km s^{-1} Mpc^{-1} level. Here we use two SN standardization methodologies from Pantheon+, one of which has no correlation with host mass (Brout & Scolnic 2021) in the baseline, and Fit 55, which has a 0.045 mag step at $\log(M/M_\odot) = 10$, increasing H_0 by 0.3 km s^{-1} Mpc^{-1}.

In Table 9 we compare the global properties of galaxies in the calibrator and Hubble-flow samples including mean mass, SFR, and SSFR. The baseline Hubble-flow and calibrator samples have the same mean SFR (0.3), consistent with their matched selection (the mean SFR of the early-type hosts in Pantheon+ is −0.6). The mean mass is also similar, with the calibrators higher by 0.5 dex—a difference smaller than the dispersion of either sample. In Appendix A we compare the distributions of host masses for various samples. To produce a late-type Hubble-flow sample with mass exceeding the calibrator sample, Fit 37 limits the LZSPI to $\log(M/M_\odot) > 10$ and lowers H_0 by 0.3 km s^{-1} Mpc^{-1}. We conclude that the calibrator and baseline Hubble-flow samples are well matched in mass and SFR.

We would not expect any other host properties, especially any local to the SN that was not a selection criterion, to significantly differ between samples. For example, Anderson et al. (2015) measured the relative strength of Hα at the sites of 98 SNe Ia in exclusively late-type, star-forming hosts, including by chance 20 of the 38 selected for Cepheid measurements, and the fraction with detected local Hα is similar for calibrators and Hubble-flow hosts (30% versus 45%). Indeed, it would be very hard to understand how a difference in local SN-host properties could occur between the two sets of hosts with matched selection. While additional host or SN properties beyond the ones we have used to measure H_0 may be used to

improve SN distance estimates now or in the future, the matching of SN and host samples employed here and the large sample size for each would mitigate any significant effect on the determination of H_0.

7.4. State of the Hubble Tension

Our baseline determination of H_0 is 73.04 ± 1.04 km s^{-1} Mpc^{-1} (with systematics), which exceeds the Planck+ΛCDM result by 5σ. It may be of interest and it is straightforward to calculate a combined value of H_0 that is free of measurement interdependencies and has lower uncertainty using additional redshift–magnitude relation data. We start with the combined Cepheid and TRGB result, 72.53 ± 0.99 km s^{-1} Mpc^{-1} (or Fits 31 and 32), and to this we add two recent measures that provide enough information that allow us to have a consistent calibration and redshift frame while avoiding double use of data. Pesce et al. (2020) measured six masers in the Hubble flow; excluding the nearest from that set, NGC 4258 (because its maser distance is used here), and quoting their result in the 2M++ frame (same as our baseline) gives $H_0 = 72.1 \pm 2.7$ km s^{-1} Mpc^{-1}. Blakeslee et al. (2021) measured the IR surface brightness fluctuation distances with HST in 63 galaxies calibrated by the TRGB in the 2M++ frame. Accounting for the small difference in TRGB zero-point ($M_I = -4.014 \pm 0.025$ mag for the mean of CCHP and EDD) and used there ($M_I = -4.03$ mag) yields $H_0 = 74.0 \pm 3.0$ km s^{-1} Mpc^{-1}. The combination of these independent and independently consistent measures gives $H_0 = 72.61 \pm 0.89$ km s^{-1} Mpc^{-1} (or $H_0 = 72.42 \pm 0.89$ km s^{-1} Mpc^{-1} with the CCHP TRGB in Fit 32 and $H_0 = 72.80 \pm 0.89$ km s^{-1} Mpc^{-1} from the EDD TRGB with Fit 31), a local determination with 1.2% precision, which is also 5.2σ greater than the Planck+ΛCDM result. Other combinations may be determined but require care to avoid measurement inconsistencies in calibration, redshift frame, or double use of any data.

There has been a wide variety of ideas proposed to resolve the Hubble tension, including (but not limited to) an episode of scalar-field dark energy before recombination, the presence of additional species of neutrinos (perhaps with interactions), decaying dark matter, the presence of primordial magnetic fields, a changing electron mass, decaying or interacting dark matter, a breakdown of the Friedmann–Lemaître–Robertson–Walker metric or general relativity, and so on; we direct the reader to recent reviews, such as Di Valentino et al. (2021) and Schöneberg et al. (2021). These proposals range from moderately successful to unsuccessful with no clear resolution. Some of the more successful ideas mitigate the tension through a similar mechanism, such as increasing $H(z)$ in the early universe so that recombination occurs earlier, thereby shrinking the sound horizon which is the fundamental scale of the CMB (and also of baryon acoustic oscillations). In many scenarios this will produce additional features in the CMB, which are either incompatible with the data or make them appear more plausible (Hill et al. 2021; Poulin et al. 2021). The presence of unaccounted systematics in early- or late-universe measurements have also been suggested, but in Section 6 we comprehensively reviewed those pertaining to the route presented here, with none showing indications of validity.

Both the late- and early-universe data present formidable obstacles to hypotheses involving "new physics" or new systematics owing to the rigor and redundancy of the measurements; any proposals require specificity to see which may be viable. However, opportunities for progress on this problem exist on many fronts. We anticipate gains from improved characterization of $H(z)$, the use of new facilities to refine the local measurements (e.g., LIGO and JWST) and the early-universe measurements (e.g., CMB Stage 4 and the Simons Observatory), neutrino experiments, as well as from new theoretical insights.

8. Conclusions

1. Our baseline determination of H_0 is 73.04 ± 1.04 km s^{-1} Mpc^{-1} (with systematics) from a Cepheid-only calibration of 42 SNe Ia with good SN data quality, or 72.53 ± 0.99 km s^{-1} Mpc^{-1} combining Cepheid and TRGB for a total of 46 SN Ia calibrators with good SN data quality.
2. The measurement exceeds the Planck+ΛCDM result by 5σ (one in 3.5 million), making it implausible to reconcile the two by chance.
3. An exhaustive study of variations in the analysis and systematic uncertainties including 67 variants of analyses reveals no indication of significant inconsistencies within the measurement or promising sources of unrecognized error. The dispersion of the 59 NIR variants is 0.3 km s^{-1} Mpc^{-1} and is conservatively adopted as an additional systematic uncertainty. The mean of the variants is 73.25 km s^{-1} Mpc^{-1}, which is higher than the baseline by 0.2 km s^{-1} Mpc^{-1}.
4. We find the dispersion between 42 SN Ia and Cepheid relative distance measures is $\sigma = 0.130$ mag, similar (albeit lower) than the $\sigma = 0.135$ mag dispersion of SNe Ia in the Hubble-flow sample and yielding no evidence of excess noise in Cepheid distance measurements.
5. We find that Cepheid and TRGB distance measures are consistent when starting and ending from the same hosts (i.e., between rung one and two). We highlight a net difference of 1.3 km s^{-1} Mpc^{-1} (or 0.04 ± 0.02 mag) between measurements by two groups of the location of the tip in NGC 4258 and the resulting calibration of the TRGB (we use the mean of both), and a net 1 km s^{-1} Mpc^{-1} (0.03 mag) higher value of H_0 from the change (tripling) of the SN calibrator sample, which is consistent with the ± 0.05 mag shot noise of the subsamples.
6. We find that each of the three independent geometric anchors is consistent with the distance predicted by its Cepheids and the other two anchors. Improvements in the calibration of the Cepheid metallicity dependence have tightened this conclusion.
7. The SNe between the second and third rungs of the ladder are hosted by galaxies of the same late type (i.e., spiral) with the same or similar mean SFR and mass. Their color and shape distributions are also highly consistent. The calibrator set contains a complete sample of all suitable SNe Ia (i.e., with good data quality) at $z < 0.011$ in the last four decades. We see no indication of differences between the mean properties of the samples' hosts, nor a reason with matched selection that such would exist and impact H_0.
8. Extragalactic Cepheids appear to have a uniform relation between period and luminosity consistent with a single slope and fine structure in their light curves that

resembles those in the MW at the same period (i.e., the Hertzsprung progression).
9. The constraint provided by the distance ladder presented here is well approximated by the derived value of H_0, except in the case of models that introduce rapid, unexpected, late-time changes in $H(z)$ (relative to either ΛCDM or low-order fits to $H(z)$), or perhaps some forms of new physics. For such models, we advise replacing H_0 with the absolute SN Ia host distances derived from the first two rungs, the set of SN Ia standardized magnitudes in these hosts, SN Ia magnitudes in the Hubble flow, and their covariance, and provide an example that yields a joint constraint of $H_0 = 73.30 \pm 1.04$ km s^{-1} Mpc^{-1} and $q_0 = -0.51 \pm 0.024$.
10. The source of this long-standing, significant discrepancy between the local and cosmological routes to determining the Hubble constant remains unknown.

We thank Graeme Addison, Fabio Bresolin, George Efstathiou, and Doron Kushnir for helpful conversations related to this work. We are grateful to Peter Challis for sharing an unpublished spectrum of SN 2008fv. The ESA Hubble office provided most of the color composites shown in Figure 4. An anonymous referee provided a thoughtful report that led to improvements in this paper.

This research was supported by NASA/HST grants GO-12879, GO-12880, GO-13334, GO-13335, GO-13344, GO-15145, GO-15146, and GO-15640 from the Space Telescope Science Institute (STScI), which is operated by the Association of Universities for Research in Astronomy, Inc., under NASA contract NAS5-26555. Some of the data presented in this paper were obtained from the Mikulski Archive for Space Telescopes (MAST) at the STScI. The specific observations analyzed can be accessed at doi:10.17909/fkay-8z97.

L.M.M. acknowledges additional support from the Mitchell Institute for Fundamental Physics & Astronomy at Texas A&M University. D.O.J. acknowledges support from NASA Hubble Fellowship grant HF2-51462.001 awarded by the Space Telescope Science Institute. A.V.F.'s group at UC Berkeley is also grateful for financial assistance from NSF grant AST-1211916, the TABASGO Foundation, the Christopher R. Redlich Fund, the Miller Institute for Basic Research in Science (in which A.V.F. is a Miller Senior Fellow), and numerous individual donors.

Some of the data presented herein were obtained at the W. M. Keck Observatory, which is operated as a scientific partnership among the California Institute of Technology, the University of California, and NASA; the observatory was made possible by the generous financial support of the W. M. Keck Foundation. We thank Brad Tucker for designing the multislit masks used for the Keck LRIS spectroscopy of H II regions in the host galaxies.

Appendix A
Properties of Calibrator SNe Ia

The members of the SN Ia calibrator sample were first identified when selecting hosts for observing Cepheids with HST. We have attempted to follow the criteria for selection given by Riess et al. (2005) to provide reliable SN Ia distances: objects observed before maximum light, through low interstellar extinction ($A_V < 0.5$ mag), with modern (i.e., nonphotographic) photometry, and with typical light-curve shapes. Here we analyze the spectra of the calibrators to confirm their initial selection.

A.1. Spectral Properties Using DeepSIP

We consider two samples of calibrators. The first contains 40 SNe Ia and is a complete, volume-limited ($z < 0.011$) sample of all suitable SNe Ia seen over the years 1980–2021 in spiral hosts (a requirement for finding Cepheids, with results for H_0 provided for these alone in Fit 38). A second sample contains just two SNe Ia (SN 1999dq and SN 2007A) from a program targeting hosts that are located at greater redshifts and are more luminous, in an effort to reach the Hubble flow directly with Cepheids. Using the following quantitative spectroscopic analysis, we find that all 42 SNe Ia are in the normal range.

We obtain spectra directly from the Open Supernova Catalog[25] (OSC; Guillochon et al. 2017), which aggregates data from numerous sources (including notable low-redshift SN Ia spectroscopy releases; e.g., Silverman et al. 2012; Blondin et al. 2012; Folatelli et al. 2013; Stahl et al. 2020b). From the OSC-retrieved spectra, we select—per SN Ia—the nearest-to-maximum-light spectrum having (i) full coverage[26] of the characteristic Si II λ 6355 absorption feature, and (ii) an S/N of at least 10 per pixel.[27] If these criteria cannot be met by the available spectra, we reduce the S/N threshold to 5 pixel^{-1}; if still no spectra satisfy the criteria, we remove the Si II λ 6355 coverage requirement. If no spectra are available even after this relaxation, our automated algorithm flags the SN in question for manual intervention. Following this approach, 37 spectra are obtained with our full criteria satisfied, one is obtained with our most relaxed criteria, and two fail.

We resolve these two failures as follows.

1. SN 2008fv was spectroscopically classified as a normal SN Ia in CBET 1522 (Challis 2008). Though never published, we have obtained this spectrum from the author.
2. SN 2021hpr has three spectra available on the Transient Name Server.[28] We take the spectrum that best satisfies our full criteria as stated above.

We also override successful OSC acquisitions in several cases where more suitable spectra (i.e., those that better match our criteria) are available elsewhere. The spectra in our final set have a median phase of -0.1 days, with the earliest at -11.7 days and the latest at 14.5 days. A full accounting of relevant metadata is available upon reasonable request.

A key concern in studies that utilize SN Ia distances (such as this one) is ensuring that the objects used are indeed "normal" SNe Ia in the sense that they can be standardized using width–luminosity relations (Phillips 1993). Here we take the conservative approach of excluding subluminous SN 1991bg–like (Filippenko et al. 1992b; Leibundgut et al. 1993) and overluminous SN 1991T–like objects (Filippenko et al. 1992a; Phillips et al. 1992), thereby defining "normal" SNe Ia as those that fall within the central, most well-studied and precise region of the Phillips relation. Moreover, we confirm the normalcy of our selected objects spectroscopically by employing the

[25] sne.space
[26] As in Stahl et al. (2020b), we define "full coverage" as having a wavelength minimum below 5750 Å and a maximum above 6600 Å.
[27] Ties when two spectra have exactly the same phase are broken by taking the one with broader wavelength coverage.
[28] https://www.wis-tns.org/

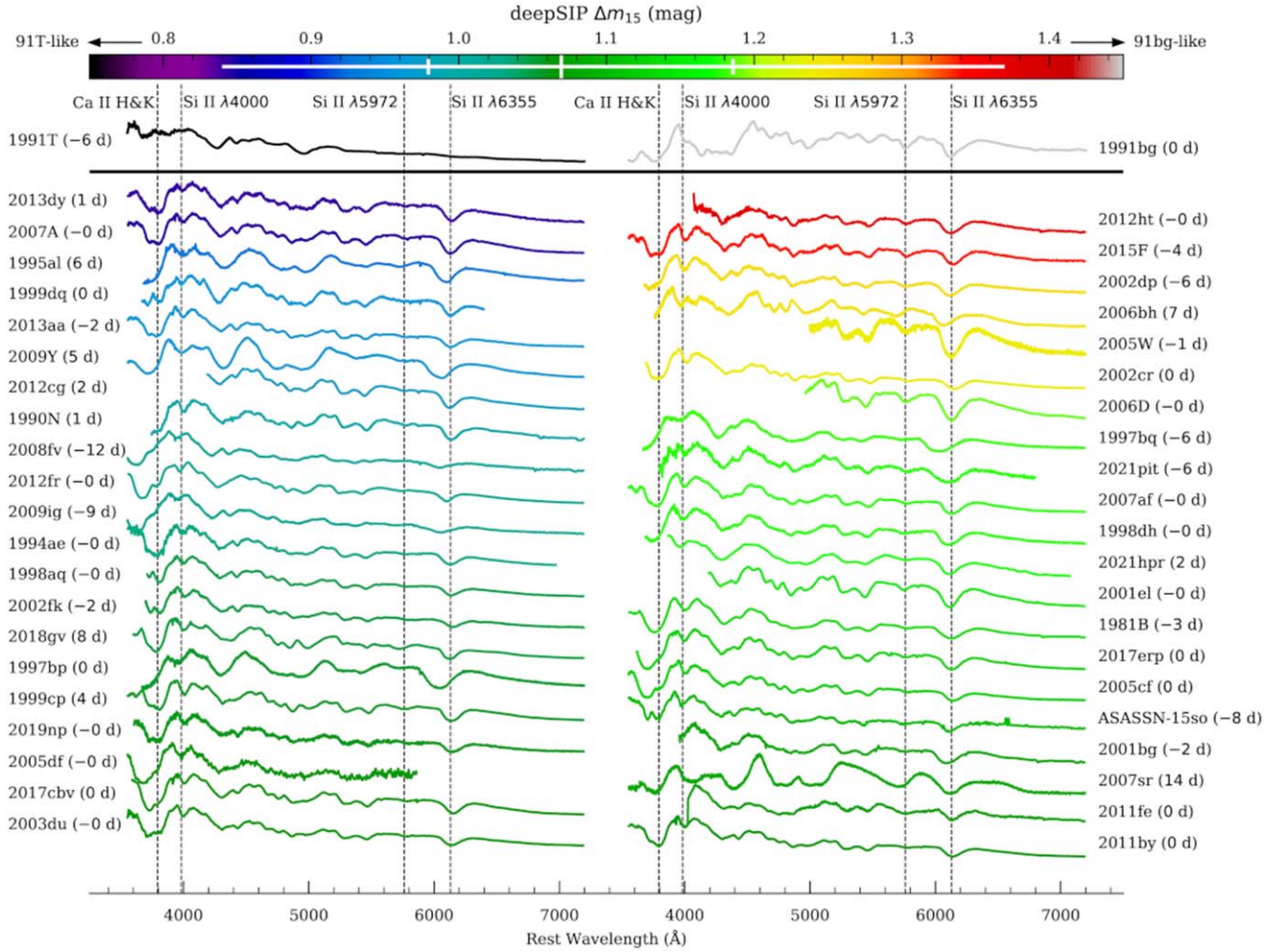

Figure A1. Spectra (scaled f_λ) of all 42 SNe Ia in the calibrator sample, color-coded by their deepSIP score calibrated to the photometric parameter, Δm_{15}. As discussed in the text, we find all calibrator SNe to be in the normal range. Peculiar objects SN 1991T and SN 1991bg bracket the extremes and are shown for comparison.

`deepSIP` package (Stahl et al. 2020a), which provides highly effective, trained convolution neural networks that, among other things, can (i) classify if a spectrum belongs to an SN Ia with a rest-frame phase between −10 days and 18 days and light-curve shape (parameterized by `SNooPy`'s Δm_{15} parameter; see Burns et al. 2011 for more details) between 0.85 and 1.55 mag, a conservatively narrow window corresponding to "normal" objects, and (ii) predict (with uncertainties) quantitative Δm_{15} values. Because we know the rest-frame phase of each spectrum in our sample from the light-curve-derived times of maximum brightness, distinctions made by the aforementioned classifier provide a direct probe if the SNe Ia in our sample are spectroscopically normal. The calibrator SNe Ia have DeepSIP Δm_{15} values between 0.84 and 1.37 mag.

We find that only two spectra do not satisfy this test, one of which (SN 2008fv) is expected owing to it being the sole case in our sample where the spectrum is at a phase earlier than −10 days. The other case (SN 1997bq) is most certainly spectroscopically normal (see Blondin et al. 2012). The fact that this single false negative is obtained is not unexpected because in developing `deepSIP`, Stahl et al. (2020a) tuned the decision threshold of the model with an eye to their subsequent scientific use case in which false positives (i.e., classifying an SN Ia as normal when it actually is not) represent a far worse error than false negatives (i.e., failing to classify an SN Ia as normal when it actually is). We visualize the entire spectral sample along with representative near-maximum-light spectra of SN 1991bg and SN 1991T for reference, color-coded by `deepSIP`-predicted Δm_{15} value, in Figure A1.

A.2. Photometric Properties

In Figure A2 we show the distributions of the SALT II color (c) and shape ($x1$) parameters, as well as the host masses, for the calibrator sample and for three Hubble-flow samples: (1) the baseline $0.0233 < z < 0.15$ spiral sample and the tighter quality cuts $|c| < 0.15$, $|x1| < 2$; (2) the same redshift range for all host types and the Pantheon+ standard quality cuts $|c| < 0.3$, $|x1| < 3$; and (3) a sample of all types to $z < 0.8$ and standard quality cuts. It can be seen that the samples are well matched in the mean, with the baseline sample better matched in breadth to the calibrator sample owing to the tighter quality cuts. As a result of this investigation, we conclude that each SN in the calibrator sample is unambiguously normal and thus can be reliably standardized to match well the baseline Hubble-flow sample.

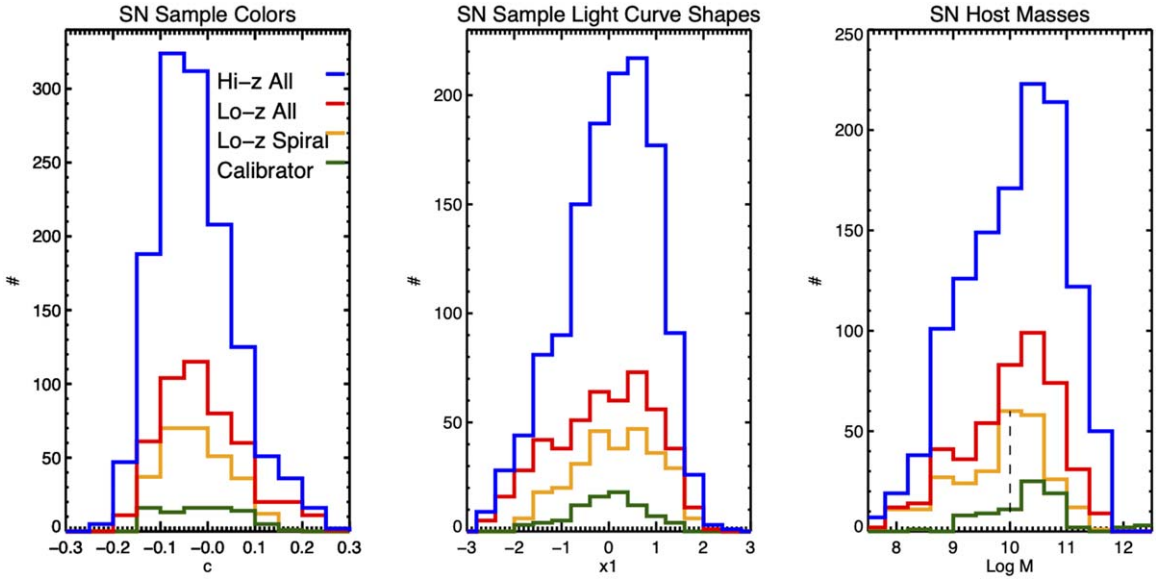

Figure A2. SN Ia and host properties for the calibrator sample (green) and three Hubble-flow samples. The color (left panel) and light-curve shape (middle) distributions are in good agreement as discussed in the text, particularly for the baseline sample (Lo-z Spiral, orange) whose quality cuts limits are $|c| < 0.15$ and $|x1| < 2$. The right panel shows the host mass including a truncation of the Lo-z sample (dashed line) with values for this and other host properties in Table 9.

Appendix B
Independent Tests of Cepheid Photometry

B.1. Aperture Photometry

The photometry of Cepheids presented here is derived from a set of standard procedures referred to in the astronomical literature as "scene modeling" or "crowded-field photometry." These methods use an empirical description of the PSF to model unresolved sources (i.e., stellar profiles) by comparing image pixels to a model constructed from the superposition of PSFs, each with its own X and Y coordinate and amplitude as well as the constant level of the background. The initial set of sources and their positions may be derived from the image or a catalog.

This approach offers greater precision than fixed-aperture photometry (i.e., summing flux in an aperture), as it can separate the blended flux of distinct sources and additionally improves the S/N by the optimal weighting of source pixels. Bias resulting from the inability to resolve nearly coincident background sources from the Cepheid can be determined statistically by adding artificial stars of known flux to the scene. This bias is really just the mean level of the fluctuating background. Uncertainties are measured from the distributions of recovered artificial stars. Frequently used software packages that enable this approach include DAOphot, DoPHOT, DolPhot, HSTphot, and Romaphot. In general, the use of these different packages has been shown to yield similar results subject to the settings for which these packages are employed.

However, it is valuable to have a robust cross-check of the accuracy of the photometry measured with these techniques using a simpler approach that is easily replicated by others without reference to any specific piece of software. Such a method is aperture photometry, which, though less precise, provides a strong test of the accuracy the Cepheid photometry reported above.

Since aperture photometry cannot readily separate superimposed sources in dense regions, a few considerations described here are necessary to produce accurate aperture photometry. (1) Small apertures are required, here set to 1.4 drizzled pixels (0.″11) in radius, along with an aperture correction derived from our PSF model to estimate the flux outside the aperture. (2) We limit the comparison to one-third of the Cepheid sample with the lowest surface brightness, a sample large enough to measure H_0 but which still has less background and thus greater precision. (3) We use a mean of the pixels in an annulus around each Cepheid to determine the sky, rather than the mode or median, which is more commonly used in sparse regions. This last step is very important for deriving accurate aperture photometry in the presence of a fluctuating background. In dense regions, the mean of the sky is an unbiased estimator of the level beneath a source, as it includes both the spatially constant background as well as the mean level of superimposed sources. In contrast, the mode or median of the crowded sky or any statistic calculated after fitting and subtracting visible stars necessarily underestimates the sky under the target source where we cannot resolve or subtract blended sources. We assume the position of the Cepheid was previously identified from sparse, high-contrast optical images, as is the case for the SH0ES program.

We measured aperture photometry of the Cepheids using the same images (before removing any background level) used for the PSF-based photometry and the same, fixed Cepheid positions. The sky value was set to the simple mean of the sky pixels in an annulus between radii of 15 and 25 pixels from each Cepheid.[29] Figure B1 displays a comparison of the aperture and PSF photometry for the Cepheids as a function of

[29] If anomalously bright Cepheids are removed as outliers from the final sample it is necessary to apply the same threshold to the sky pixels before calculating their mean to avoid a measurement bias. In practice, it is common to exclude Cepheids brighter than the P–L by $\leqslant 3\sigma$ or >1 mag to provide a more robust P–L relation. Thus, we remove sky pixels from the determination of the mean, which would cause an outlier of this size, i.e., those $\geqslant 2.5\times$ the level of the central Cepheid pixel. For the images of Cepheids studied here this fraction is typically a few percent of the brightest Cepheids and sky pixels. To avoid a similar bias in PSF photometry, it is necessary to apply this same threshold when measuring the simple mean of artificial stars to determine the mean background correction.

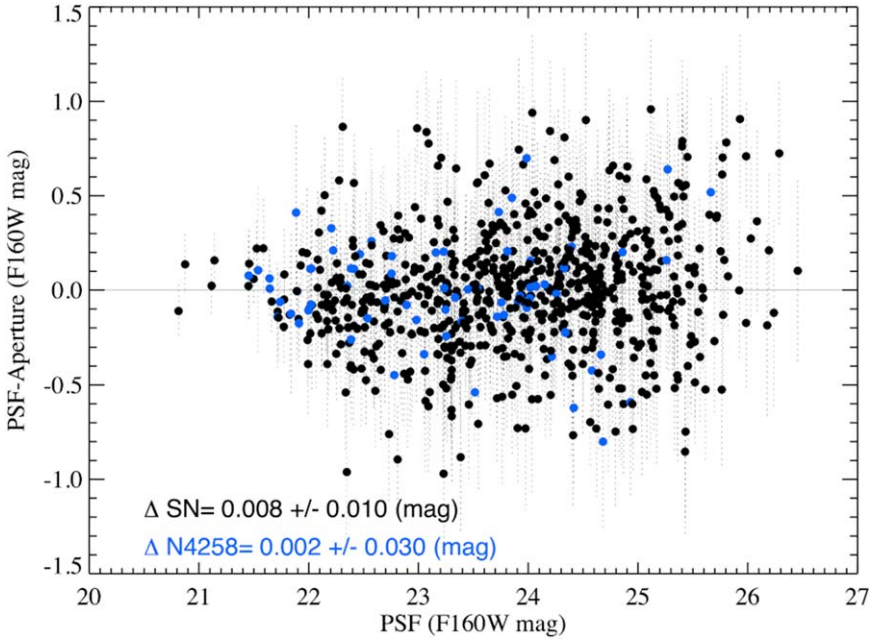

Figure B1. Comparison of PSF photometry (with backgrounds from artificial stars) for Cepheids in SN hosts (black) and NGC 4258 (blue) used here and basic small-aperture photometry with backgrounds from the mean pixel in a local annulus as discussed in the text.

PSF magnitude, showing agreement within the sample mean errors.

We can make a few additional observations about the comparison that clarifies the relationship between the statistics used to measure the sky in aperture photometry and the crowded background corrections used to account for bias in PSF photometry. Because PSF photometry sets the sky as a "floor" level added to modeled sources, the PSF sky level will be similar to the mode (i.e., the most common value or peak of the distribution of sky annulus pixels) of the sky pixels used in (uncrowded) aperture photometry, the uniform level of pixels without apparent sources. Therefore, the difference in aperture photometry calculated from a sky level using the mode or mean of the sky pixels will be similar to the crowded background correction in PSF photometry. In sparse fields, the mode and mean are equivalent, and the crowded background bias is negligible.

The asymmetric distribution of pixel levels in the sky annulus also explains the useful feature that in log-normal or magnitudes, the Cepheid uncertainties that are dominated by the asymmetric distribution of sky pixels are relatively Gaussian to a few standard deviations, as shown in Section 3.3.

B.2. Impacts of Improvements in Photometry

As discussed in Section 3.4, we itemize half a dozen improvements in Cepheid photometry from the last time these Cepheids were measured between ∼6 and ∼16 yr ago. Matching Cepheids by position, the error-weighted mean of the matched Cepheids are fainter in SN hosts by 0.06 mag and in NGC 4258 by 0.04 mag (a net difference of 0.02 mag when NGC 4258 calibrates the Cepheids and H_0). For the LMC, Cepheids became fainter by 0.03 mag between the ground sample used by R16 and their replacement by direct HST observations in R19. The mean change in *F555W* – *F814W* was below 0.01 mag for Cepheids in SN hosts and in the outer field of NGC 4258 but 0.06 mag in the inner field relative to the previous calibration undertaken by Macri et al. (2006; and provided in H16), which lacked accounting for the crowded background level, a larger effect in *F814W* than in *F555W*, and pixel-based CTE rectification. We cannot report the change in photometry for MW Cepheids because we use different MW parallax samples here than those in R16.

For context, R16 propagated a zero-point uncertainty between ground and HST system photometry of $\sigma = 0.03$ mag, which is the same size as the changed observed for the LMC sample upon direct observation with HST. The net change between the Cepheids in the SN hosts and NGC 4258 or the LMC, the quantity that determines H_0, is 0.02, comparable to the systematic uncertainty in R16 of 0.026 mag for the NGC 4258 anchor and smaller than the overall uncertainty in H_0 of 0.052 mag. A more noteworthy change appears between the MW-only anchor results where the replacement with the Gaia EDR3 results plus HST photometry with the sample from Benedict et al. (2007) and their ground-based photometry reduced H_0 by 0.086 mag, with some of this change related to the aforementioned update of the Cepheid photometry in SN hosts, though the quadrature sum error for only the parallax samples is ∼0.06 mag, making this change not surprising. For additional context for the LMC, we note that the DEB distance also decreased by ∼0.015 mag and the metallicity term by 0.033 mag (same sense) since R16. These factors explain the net increase in H_0 for the LMC.

B.3. Background Dependence of Fit Residuals

An accurate assessment of the background flux is critical to the use of standard candles. The Cepheid backgrounds are determined statistically by adding and measuring artificial stars in random positions local to the Cepheid scenes as described above. Here we provide an additional null test of the background estimates.

Because the individual backgrounds are determined locally as part of measuring the photometry, they are not part of the model. A null test of the level of background is to analyze the

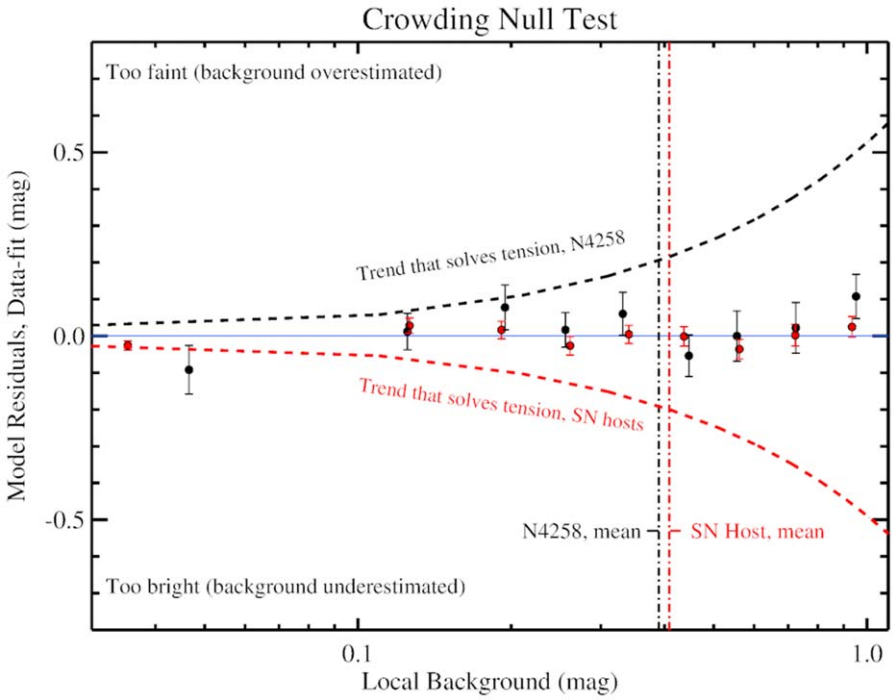

Figure B2. Mean differences between Cepheid photometry and distance-ladder fit versus the local background measured by artificial stars (in units of the Cepheid magnitudes), binned in equal numbers of Cepheids. Red is for Cepheids in SN Ia hosts and green for Cepheids in the anchor NGC 4258, with vertical lines indicating the mean background of each set. An underestimate of Cepheid backgrounds in SN Ia hosts at a level of 0.2 mag for the mean Cepheid background or an overestimate of Cepheid backgrounds in NGC 4258 by 0.2 mag for the mean Cepheid (indicated), or half that level for both simultaneously, would be required to solve the tension, but these possibilities are inconsistent with the data, which are consistent with no overestimates or underestimates for either set. Furthermore, it is hard to imagine how an overestimate or underestimate would not similarly affect (and therefore cancel) Cepheids in both NGC 4258 and SN Ia hosts.

correlation of background with the model residual of each Cepheid. If backgrounds are significantly systematically overestimated or underestimated, we would expect a correlation in this space as the residual would be a function of the (misestimated) background. We show the results of this test in Figure B2 for the baseline fit, binning the residuals in small ranges of background values. These bins show no significant trend with background. A linear fit gives a relation with slope 0.015 ± 0.014 mag per mag of background in the sense of overestimating the background by 0.006 ± 0.005 mag at the mean background level in the SN hosts and in NGC 4258 (as displayed in that figure). The figure also shows the expected trend if the background were underestimated enough to produce a 0.2 mag change in H_0 for the mean host (i.e., to solve the tension), which is very far from the data (and would also apply to the Cepheids in NGC 4258 and thus not address the discrepancy for that anchor in Fit 10).

For those who would like to check the accuracy of independently-measured extragalactic Cepheid PSF photometry, we can provide (upon reasonable request) images of hosts with artificial Cepheids of known brightness and position added to the frames. Reproducing the known photometry would be an important test of any photometry algorithm that can be completed blindly and in advance of an independent determination of H_0.

Appendix C
Cepheid Metallicity

Cepheid abundances beyond the Magellanic Clouds are generally derived from radial metallicity gradients measured from the ratios of strong emission lines in H II regions, such as the R_{23} strong-line diagnostic based on oxygen and hydrogen (Zaritsky et al. 1994, hereafter Z94). The conversion of R_{23} to $12 + \log[\mathrm{O/H}]$ in the Z94 formula was based on photoionization models that themselves were calibrated at a time when the standard solar metallicity was believed to be $12 + \log[\mathrm{O/H}] = 8.93$ (Anders & Grevesse 1989). Since Asplund et al. (2005, 2009), the solar abundance was revised downward by 0.24 dex to 8.69 ± 0.05. Over the same period, empirical transformations between strong-line measures like R_{23} were also revised downward by ~0.2–0.3 dex, making them more consistent with stellar measures referenced to the Sun (Pettini & Pagel 2004; Kewley & Ellison 2008; Tremonti et al. 2004; Dopita et al. 2016; Curti et al. 2017). Bresolin et al. (2016) have shown good agreement between extragalactic oxygen abundance measurements from young stars and from H II regions using these more modern systems (particularly for Pettini & Pagel 2004, hereafter, PP04 O3N2; F. Bresolin, private communication). Collisionally excited lines generally underestimate stellar abundances by 0.2 dex for reasons discussed by Carigi et al. (2019). Using 30,000–50,000 galaxy spectra from SDSS DR7, Teimoorinia et al. (2021) derived third-order polynomial conversions between the Z94 scale and nine different strong-line metallicity calibrations developed since 2004.

For our measurement of H_0, we employ two strong-line abundance systems. The first is the simple mean of all nine recent (since 2004) diagnostics of Teimoorinia et al. (2021). This revises the Z94 scale down by a mean of 0.28 dex. The change is not quite constant, as the mean produces flatter galaxy gradients than Z94, revising the Z94 values down an additional 0.05 dex at >9.0 and less at <8.8 by 0.04 dex.

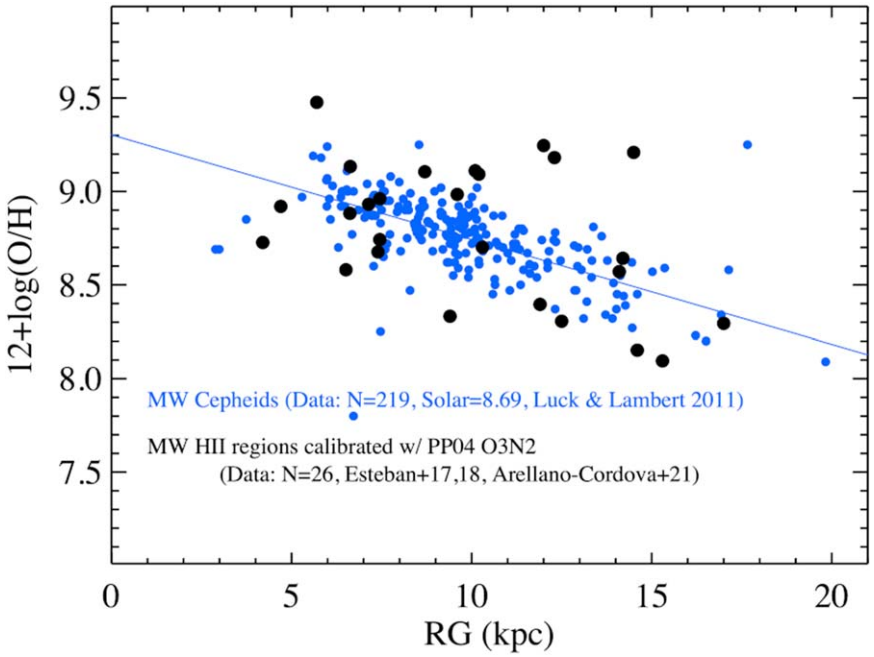

Figure C1. Comparison of oxygen abundance gradients in the MW inferred by spectra of Cepheids (blue) and from the PP04 calibration of H II regions (black) with data sources indicated in the plot. The mean levels are found to be consistent to $\sigma < 0.1$ dex, and we propagate systematic differences in scales in the covariance matrix (as discussed in the text).

The second is the PP04 O3N2 diagnostic, which uses the dereddened abundances of [O III] λ5007 relative to Hβ λ4861 and [N II] λ6583 relative to Hα λ6563 to measure the log of the ratio of [O III] to [N II], O3N2, with the calibration $12 + \log[\mathrm{O/H}] = 8.73 - 0.32 \times \mathrm{O3N2}$. The difference between the Z94 diagnostic and the PP04 O3N2 diagnostic is very nearly constant, reducing the oxygen abundance scale by 0.23 dex (0.235 dex at Z94 >9.0 and 0.222 at <8.8). The PP04 conversions of Teimoorinia et al. (2021) are valid for oxygen abundance on the Z94 scale >8.35, encompassing all Cepheid hosts, with a typical stated precision of 0.03 dex. This change not surprisingly mirrors the drop in solar abundance of 0.24 dex from the revisions of Asplund et al. (2005, 2009) and the related recalibration of strong-line diagnostics. For this reason, it has made little difference for Cepheid measurements whether to use the prior Z94 calibration referenced to the Anders & Grevesse (1989) solar abundance or the revised (mean of) the recent systems referenced to the Asplund et al. (2005, 2009) solar abundance.

When calibrating Cepheids in SN Ia hosts to those in NGC 4258, both sets of metallicities are derived using the same strong-line abundance scale so that differences in abundance calibration do not affect the difference in abundance. However, LMC and MW Cepheid abundances are generally measured through direct spectroscopy of the stars, so when these calibrators are employed it is important to test the consistency of direct and indirect methods for measuring Cepheid abundances. This is most readily accomplished in the MW where both measures are available. Luck & Lambert (2011) provide oxygen abundances for 219 MW Cepheids on the Asplund et al. (2005) solar calibration of 8.69 from which they measured a gradient of $12 + \log[\mathrm{O/H}] = 9.303\ (\pm 0.028) - 0.056(\pm 0.003)$ dex kpc^{-1} as shown in Figure C1.

To compare to these, we derived the strong-line abundances from the three largest samples of MW H II regions from Esteban et al. (2017), Esteban & García-Rojas (2018), and Arellano-Córdova et al. (2021). We used the tabulated values of [O III] λ5007, Hβ, Hα, and [N II] λ6583 in those references to calculate the ratio O3N2 and from that the PP04 calibration to determine $12 + \log[\mathrm{O/H}]$ as shown in Figure C1. The abundance gradient from the Cepheids as seen in that figure is a good match to the H II regions. A formal fit gives the PP04 stellar scale $= 0.035 \pm 0.06$ dex. From the Teimoorinia et al. (2021) transformations we find the mean difference between the PP04 and the nine recent systems, PP04$-$ave $= 0.047$ dex, demonstrating that the average of the nine systems is right on the stellar scale to a precision of $\sim$0.06 dex. In Section 3.5, Equation (9), we employ a metallicity covariance in the fit covariance matrix measured as the difference between the mean of the metallicity diagnostics and the PP04 O3N2 diagnostic, which propagates both a 0.05 dex scale uncertainty, and also the differences that depend on the range of metallicity and rise to 0.1 dex at $12 + \log[\mathrm{O/H}] = 8.8$ (i.e., $[\mathrm{O/H}] = +0.11$) and up to 0.15 dex at 9.0. We also present the determination of H_0 on both reference systems as well for the exclusion of any dependence of Cepheid luminosity on metallicity to bracket the full range.

Appendix D
Use and Misuse of Wesenheit Magnitudes

The Wesenheit magnitude used to measure distances in Equation (7) is not a conventional magnitude (i.e., not a ratio of flux to a standard star), but rather a "dereddened magnitude" (see Appendix C of Madore & Freedman 1991 for a review), the ratio of two standardized candles with the same fiducial luminosity would be expected to have in the absence of reddening. The Wesenheit formalism takes advantage of the candles having the same mean luminosity and color (before being reddened by dust) to preserve their difference in distance (in magnitudes) without requiring knowledge of either the (absolute) intrinsic luminosity or color.

This method, and potential pitfalls with alternative uses of these magnitudes, can be better understood by decomposing the Cepheid color into its constituent parts. We can define the apparent Cepheid color,

$$V - I = (V - I)_0 + \Delta(V - I) + E(V - I), \quad (22)$$

where $(V-I)_0$ is the color (at a given period) in the middle of the Cepheid instability strip, $\Delta(V-I)$ is the displacement in color from the midline of the instability strip (with an accompanying displacement in mean luminosity owing to Stefan's law projected into the observed bandpasses, redder is fainter) and is quantified empirically in the Cepheid *P–L–C* relation. For a statistical sample of Cepheids beyond period completeness (which provides Cepheids uniformly distributed over the instability strip, mixed by the comparatively small evolution time to cross the strip compared to the larger differences in formation times), the mean $\langle \Delta(V-I) \rangle = 0$ mag. Hence, the relative difference in distance (in magnitudes) between samples of Cepheids is statistically preserved by subtracting from the apparent magnitudes of each any fixed number, R, which multiplies the (same) intrinsic color of the midlines of their instability strips (i.e., $R(V-I)_0$). If the value of R is chosen to be the reddening ratio derived from an accurate reddening law of dust, then the subtracted term will also remove the absorption by dust (i.e., $R\,E(V-I)$) from each Cepheid for dissimilar values of $E(V-I)$. This is the function of a Wesenheit magnitude. A "happy coincidence" is that Stefan's law is quite parallel to the reddening law (redder is fainter or dustier), so this operation will also reduce the small dispersion due to the width of the instability strip (Madore & Freedman 1991). As given in Section 6.3, we derived the intrinsic *P–L–C* (Period-Luminosity-Color) relation from the LMC Cepheids as above to be $\Delta H = 0.635\ (\pm 0.021)\ \Delta(V-I)$, the same sense and similar in scale to reddening by dust for which $R \approx 0.4$.

It is not valid to use a different reddening ratio (law) in different hosts by simply varying the value of R for each host using the same Wesenheit system. Using a different value of R for different Cepheids (or for Cepheids in different hosts) in Equation (7) without first removing the intrinsic color from the apparent color produces a change in the measured differences in the distance of Cepheid samples of size $R_1(V-I)_0 - R_2(V-I)_0$, where $R_1 \neq R_2$, which is unrelated to dust but is rather a large, ad hoc change in the standard candle luminosity itself as a fraction of its color. Because the typical $(V-I)_0$ color is a factor of a few times larger than the typical $E(V-I)$, this change in Cepheid luminosity will dominate the dust correction and produce a bias that is a rather arbitrary fraction of a Cepheid's color. This is a shortcoming in the analysis presented by Mortsell et al. (2021).

It is also important to recognize that Cepheids are well-understood yellow supergiant stars that will only pulsate over a very narrow range in surface temperature of $\log T = 3.7$ ($T = 5000$ K), which sets their intrinsic color of $(V-I)_0 \approx 0.8$ mag (at $P = 10$ days) at the instability strip of the temperature–magnitude diagram of stars. This is the surface temperature where its internal structure supports pulsations by the kappa mechanism (change in internal opacity in a deep partial ionization zone from ionization and recombination of He+) so that the value of $(V-I)_0$ is well constrained by the physics that causes a Cepheid to be formed.

Thus, varying R in the Wesenheit definition does not appear to follow any specific hypothesis. To consider the hypothesis of different reddening laws in different hosts, it is necessary to first subtract the intrinsic color as described in Section 6.3 and by Follin & Knox (2018).

D.1. Errors in Color and Fitting

Because the value of R is tightly constrained by the sample of MW Cepheids to 0.363 ± 0.038 (and better together with the Cepheids in other nearby hosts such as M31 and M101), the error in color $(V-I)$ can be propagated to an error in H through $H \propto R(V-I)$ ignoring the small uncertainty in R and added in quadrature in the covariance matrix of errors. However, this is no longer true if R is poorly constrained by the data and the errors in color are large, as would be the case if we tried to determine R independently for each individual SN host after subtracting the intrinsic color. Because the measurement errors in color are significant, to determine individual host values of R without bias, it is necessary to consider errors in both axes (H-band magnitude and $V-I$ color) to optimize R as discussed for a similar problem by Tremaine et al. (2002) or follow the constrained approach of Follin & Knox (2018).

Unfortunately, beyond the nearest few galaxies, the data are quite inadequate for constraining R owing to the combination of a small range of color and comparatively large color uncertainties. Figure D1 shows the available Cepheid data for a nearby host with considerable measurable extinction (M31), one of the two nearest SN Ia hosts (NGC 5643) for which the data are just sufficient, and for a typical SN Ia host (UGC 9391) where the data are insufficient. Accounting only for uncertainty in the dependent variable produces a biased result and underestimates the uncertainty when the data quality is low, as is readily seen in Monte Carlo simulations compared to the input value of R.

In these cases, the breadth of colors is largely attributed to the measurement errors rather than simply extinction. Accounting for uncertainty in both axes, we find that the value of R is quite unconstrained beyond the nearest few hosts. A bootstrap resampling of the data shows that the constraints on R become uninformative beyond the nearest two SN Ia hosts with a mean uncertainty for the rest of $\sim$1.5. The lower-right panel of Figure D1 shows the likelihoods for different host R values. We find that the nearest few large spirals have R consistent with the MW, but there is no meaningful constraint or well-defined peak for the rest. We can combine all such constraints as shown in the figure with a result that is unsurprisingly not well defined. Mortsell et al. (2021) and Perivolaropoulos & Skara (2021) did not include the color errors, which our analysis shows leads to an underestimate of the values of R and their uncertainties for nearly all SN hosts. From a Bayesian perspective, the addition of a free parameter, R, for each host ($\sim$40 free parameters) is not justified by the improvement in the fit over the global R, as also suggested in the analysis by Mortsell et al. (2021) and Follin & Knox (2018). An informed approach to modeling individual host R values without so many poorly constrained free parameters can be found in Section 6.3, Fit 22, and Hahn et al. (2022).

The EDA framework also provides estimates of the mean extinction seen midway through the SN hosts (i.e., the average location of a Cepheid) in our NIR bandpass from the same host

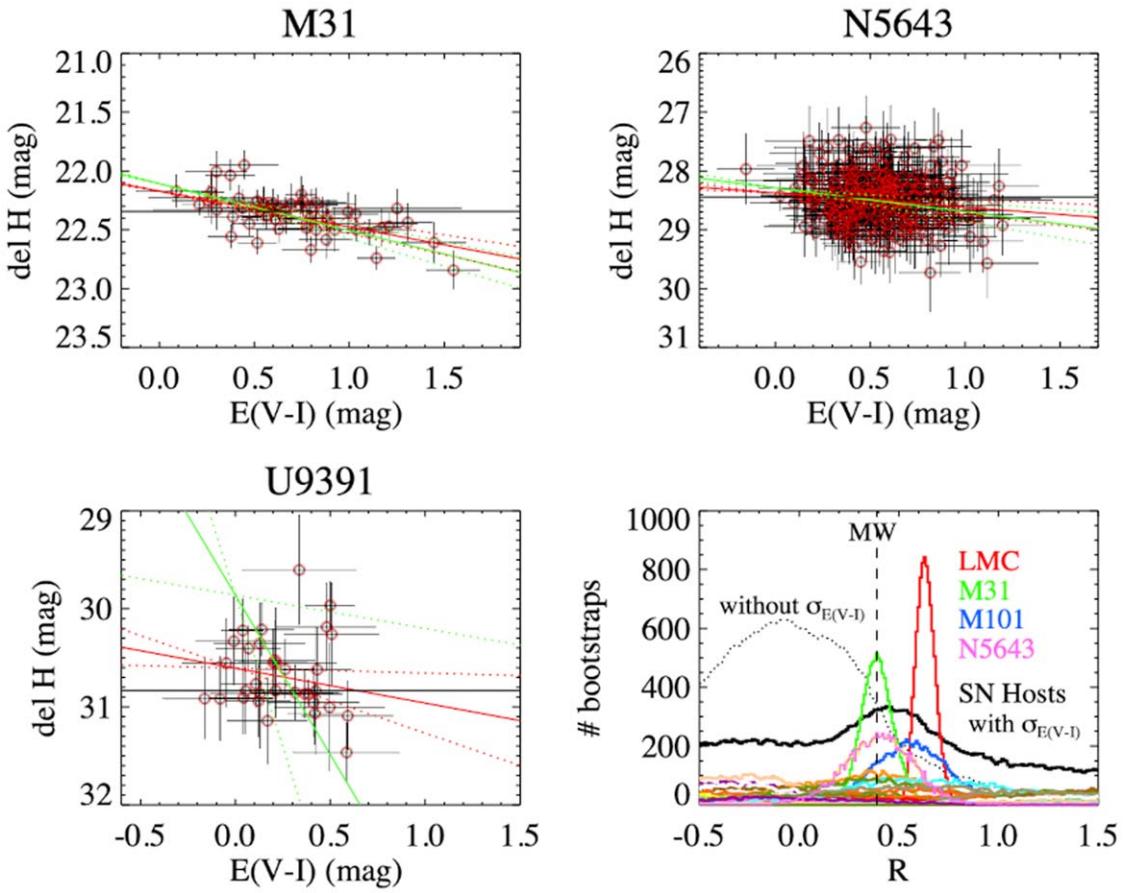

Figure D1. Measuring the reddening ratio R in each host. For each host we subtract the intrinsic period–color relation as given in the text to provide $E(V-I)$ and correct the H-band magnitudes for the P–L relation (period and metallicity). The ideal case, M31, has a large span of extinction (because it is highly inclined) and small measurement errors (because it is nearby) to produce a good constraint on the slope, R, whether we consider errors in only the dependent variable (red line and dotted for uncertainty) or both axes (green line). $R = 0$ is shown as a black line. NGC 5643 is one of the two nearest SN Ia hosts and has the most Cepheids; it yields some constraint on R along with M101. In contrast, the rest of the SN hosts have data similar to UGC 9391, which produces an uninformative constraint on R and shows a large difference in its value and uncertainty by ignoring uncertainties in both axes. The lower right shows the bootstrap resampling for all hosts, with a different color for each, showing that a meaningful constraint is only seen for the few nearest cases, with the others relatively flat. The measurable cases are consistent with the MW R shown as a black dashed line, as are the sum of all SN hosts. Ignoring errors in $E(V-I)$ underestimates R and its uncertainty, as discussed in the text.

properties used to determine R above but now including the host inclination. These values range from $A_H = 0.07$ mag (NGC 4424, which at type Sa is the earliest spiral in our sample) to 0.30 mag with a mean of 0.18 mag and a dispersion of 0.05 mag. This mean is similar to the empirical result of the product of the SN-host mean $E(V-I)$ and R, which yields $\langle A_H \rangle = 0.14$ mag, indicating that the level of extinction inferred from the Cepheid reddening is the expected amount.[30]

ORCID iDs

Adam G. Riess https://orcid.org/0000-0002-6124-1196
Wenlong Yuan https://orcid.org/0000-0001-9420-6525
Lucas M. Macri https://orcid.org/0000-0002-1775-4859
Dan Scolnic https://orcid.org/0000-0002-4934-5849
Dillon Brout https://orcid.org/0000-0001-5201-8374
David O. Jones https://orcid.org/0000-0002-6230-0151
Yukei Murakami https://orcid.org/0000-0002-8342-3804
Gagandeep S. Anand https://orcid.org/0000-0002-5259-2314
Louise Breuval https://orcid.org/0000-0003-3889-7709
Thomas G. Brink https://orcid.org/0000-0001-5955-2502
Alexei V. Filippenko https://orcid.org/0000-0003-3460-0103
Samantha Hoffmann https://orcid.org/0000-0002-4312-7015
Saurabh W. Jha https://orcid.org/0000-0001-8738-6011
W. D'arcy Kenworthy https://orcid.org/0000-0002-5153-5983
John Mackenty https://orcid.org/0000-0002-4312-7015
Benjamin E. Stahl https://orcid.org/0000-0002-3169-3167
WeiKang Zheng https://orcid.org/0000-0002-2636-6508

[30] A more detailed comparison would require modeling of the Cepheid selection in the presence of variable depth of dust and is beyond the scope here.

References

Anand, G. S., Tully, R. B., Rizzi, L., Riess, A. G., & Yuan, W. 2021, arXiv:2108.00007
Anders, E., & Grevesse, N. 1989, GeCoA, 53, 197
Anderson, R. I. 2019, A&A, 631, A165
Anderson, R. I., & Riess, A. G. 2018, ApJ, 861, 36
Anderson, R. I., Sahlmann, J., Holl, B., et al. 2015, ApJ, 804, 144
Arellano-Córdova, K. Z., Esteban, C., García-Rojas, J., & Méndez-Delgado, J. E. 2021, MNRAS, 502, 225
Arenou, F., & Luri, X. 1999, in ASP Conf. Ser., 167, Harmonizing Cosmic Distance Scales in a Post-HIPPARCOS Era, ed. D. Egret & A. Heck (San Francisco, CA: ASP), 13

Asplund, M., Grevesse, N., & Sauval, A. J. 2005, in ASP Conf. Ser., 336, Cosmic Abundances as Records of Stellar Evolution and Nucleosynthesis, ed. I. Barnes, G. Thomas, & F. N. Bash (San Francisco, CA: ASP), 25
Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, ARA&A, 47, 481
Avelino, A., Friedman, A. S., Mandel, K. S., et al. 2019, ApJ, 887, 106
Bailer-Jones, C. A. L., Rybizki, J., Fouesneau, M., Demleitner, M., & Andrae, R. 2021, AJ, 161, 147
Benedict, G. F., McArthur, B. E., Feast, M. W., et al. 2007, AJ, 133, 1810
Betoule, M., Kessler, R., Guy, J., et al. 2014, A&A, 568, A22
Blakeslee, J. P., Jensen, J. B., Ma, C.-P., Milne, P. A., & Greene, J. E. 2021, ApJ, 911, 65
Blondin, S., Matheson, T., Kirshner, R. P., et al. 2012, AJ, 143, 126
Bonanos, A. Z., Stanek, K. Z., Sasselov, D. D., et al. 2003, AJ, 126, 175
Bono, G., Castellani, V., & Marconi, M. 2002, ApJ, 565, L83
Bono, G., Marconi, M., & Stellingwerf, R. F. 1999, ApJS, 122, 167
Bovy, J. 2018, Statistics and Inference, Course 2018, Univ. of Toronto
Bresolin, F., Kudritzki, R.-P., Urbaneja, M. A., et al. 2016, ApJ, 830, 64
Breuval, L., Kervella, P., Anderson, R. I., et al. 2020, A&A, 643, A115
Breuval, L., Kervella, P., Wielgórski, P., et al. 2021, ApJ, 913, 38
Brout, D., & Scolnic, D. 2021, ApJ, 909, 26
Brout, D., Taylor, G., Scolnic, D., et al. 2021, arXiv:2112.03864
Brownsberger, S., Brout, D., Scolnic, D., Stubbs, C. W., & Riess, A. G. 2021, arXiv:2110.03486
Burns, C. R., Stritzinger, M., Phillips, M. M., et al. 2011, AJ, 141, 19
Burns, C. R., Parent, E., Phillips, M. M., et al. 2018, ApJ, 869, 56
Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245
Camarena, D., & Marra, V. 2021, MNRAS, 504, 5164
Cardona, W., Kunz, M., & Pettorino, V. 2017, JCAP, 3, 056
Carigi, L., Peimbert, M., & Peimbert, A. 2019, arXiv:1906.07804
Carr, A., Davis, T. M., Scolnic, D., et al. 2021, arXiv:2112.01471
Carrick, J., Turnbull, S. J., Lavaux, G., & Hudson, M. J. 2015, MNRAS, 450, 317
Casertano, S., Riess, A. G., Anderson, J., et al. 2016, ApJ, 825, 11
Challis, P. 2008, CBET, 1522, 1
Chen, G., & Ratra, B. 2011, PASP, 123, 1127
Conley, A., Guy, J., Sullivan, M., et al. 2011, ApJS, 192, 1
Cox, A. N. 1993, in IAU Coll. 139, Convection and the Bump Cepheid Resonance (Cambridge: Cambridge Univ. Press), 261
Curti, M., Cresci, G., Mannucci, F., et al. 2017, MNRAS, 465, 1384
Cutri, R. M., Skrutskie, M. F., van Dyk, S., et al. 2003, yCat, II/246
Dalcanton, J. J., Williams, B. F., Lang, D., et al. 2012, ApJS, 200, 18
Dhawan, S., Brout, D., Scolnic, D., et al. 2020, ApJ, 894, 54
Dhawan, S., Jha, S. W., & Leibundgut, B. 2018, A&A, 609, A72
Di Valentino, E., Mena, O., Pan, S., et al. 2021, CQGra, 38, 153001
Dopita, M. A., Kewley, L. J., Sutherland, R. S., & Nicholls, D. C. 2016, Ap&SS, 361, 61
Eddington, A. S. 1917, Obs, 40, 290
Efstathiou, G. 2014, MNRAS, 440, 1138
Efstathiou, G. 2021, MNRAS, 505, 3866
Esteban, C., Fang, X., García-Rojas, J., & Toribio San Cipriano, L. 2017, MNRAS, 471, 987
Esteban, C., & García-Rojas, J. 2018, MNRAS, 478, 2315
Feeney, S. M., Mortlock, D. J., & Dalmasso, N. 2019, MNRAS, 476, 3861
Ferrarese, L., Silbermann, N. A., Mould, J. R., et al. 2000, PASP, 112, 177
Ferrarese, L., Bresolin, F., Kennicutt, R. C. J., et al. 1998, ApJ, 507, 655
Filippenko, A. V., Richmond, M. W., Matheson, T., et al. 1992a, ApJL, 384, L15
Filippenko, A. V., Richmond, M. W., Branch, D., et al. 1992b, AJ, 104, 1543
Fiorentino, G., Musella, I., & Marconi, M. 2013, MNRAS, 434, 2866
Fitzpatrick, E. L. 1999, PASP, 111, 63
Fitzpatrick, E. L., Massa, D., Gordon, K. D., Bohlin, R., & Clayton, G. C. 2019, ApJ, 886, 108
Folatelli, G., Morrell, N., Phillips, M. M., et al. 2013, ApJ, 773, 53
Follin, B., & Knox, L. 2018, MNRAS, 477, 4534
Foreman-Mackey, D., Hogg, D. W., Lang, D., & Goodman, J. 2013, PASP, 125, 306
Freedman, W. L. 2021, ApJ, 919, 16
Freedman, W. L., & Madore, B. F. 2011, ApJ, 734, 46
Freedman, W. L., Madore, B. F., Scowcroft, V., et al. 2012, ApJ, 758, 24
Freedman, W. L., Madore, B. F., Gibson, B. K., et al. 2001, ApJ, 553, 47
Freedman, W. L., Madore, B. F., Hatt, D., et al. 2019, ApJ, 882, 34
Gieren, W., Storm, J., Konorski, P., et al. 2018, A&A, 620, A99
Goodman, J., & Weare, J. 2010, Comm App Math Comp Sci, 5, 65
Graczyk, D., Pietrzyński, G., Thompson, I. B., et al. 2020, ApJ, 904, 13
Groenewegen, M. A. T. 2018, A&A, 619, A8
Guillochon, J., Parrent, J., Kelley, L. Z., & Margutti, R. 2017, ApJ, 835, 64
Guy, J., Sullivan, M., Conley, A., et al. 2010, A&A, 523, A7
Hahn, C., Starkenburg, T. K., Anglés-Alcázar, D., et al. 2022, ApJ, 926, 122
Hamuy, M., Phillips, M. M., Suntzeff, N. B., et al. 1996, AJ, 112, 2408
Hertzsprung, E. 1926, BAN, 3, 115
Hill, J. C., Calabrese, E., Aiola, S., et al. 2021, arXiv:2109.04451
Hoffmann, S. L., Macri, L. M., Riess, A. G., et al. 2016, ApJ, 830, 10
Holtzman, J. A., Burrows, C. J., Casertano, S., et al. 1995, PASP, 107, 1065
Hu, W. 2005, PhRvD, 71, 047301
Huang, C. D., Riess, A. G., Yuan, W., et al. 2020, ApJ, 889, 5
Huchra, J. P., Macri, L. M., Masters, K. L., et al. 2012, ApJS, 199, 26
Humphreys, E. M. L., Reid, M. J., Moran, J. M., Greenhill, L. J., & Argon, A. L. 2013, ApJ, 775, 13
Jang, I. S., & Lee, M. G. 2017, ApJ, 836, 74
Jang, I. S., Hoyt, T. J., Beaton, R. L., et al. 2021, ApJ, 906, 125
Javanmardi, B., Mérand, A., Kervella, P., et al. 2021, ApJ, 911, 12
Jones, D. O., Riess, A. G., Scolnic, D. M., et al. 2018, ApJ, 867, 108
Kato, D., Nagashima, C., Nagayama, T., et al. 2007, PASJ, 59, 615
Kenworthy, W. D., Scolnic, D., & Riess, A. 2019, ApJ, 875, 145
Kewley, L. J., & Ellison, S. L. 2008, ApJ, 681, 1183
Kodric, M., Riffeser, A., Seitz, S., et al. 2018, ApJ, 864, 59
Komatsu, E., Dunkley, J., Nolta, M. R., et al. 2009, ApJS, 180, 330
Komatsu, E., Smith, K. M., Dunkley, J., et al. 2011, ApJS, 192, 18
Lavaux, G., & Hudson, M. J. 2011, MNRAS, 416, 2840
Leavitt, H. S., & Pickering, E. C. 1912, HarCi, 173, 1
Leibundgut, B., Kirshner, R. P., Phillips, M. M., et al. 1993, AJ, 105, 301
Li, S., Riess, A. G., Busch, M. P., et al. 2021, ApJ, 920, 84
Lilow, R., & Nusser, A. 2021, MNRAS, 507, 1557
Luck, R. E., & Lambert, D. L. 2011, AJ, 142, 136
Lutz, T. E., & Kelker, D. H. 1973, PASP, 85, 573
Macri, L. M., Ngeow, C.-C., Kanbur, S. M., Mahzooni, S., & Smitka, M. T. 2015, AJ, 149, 117
Macri, L. M., Stanek, K. Z., Bersier, D., Greenhill, L. J., & Reid, M. J. 2006, ApJ, 652, 1133
Macri, L. M., Kraan-Korteweg, R. C., Lambert, T., et al. 2019, ApJS, 245, 6
Madore, B. F. 1982, ApJ, 253, 575
Madore, B. F., & Freedman, W. L. 1991, PASP, 103, 933
Mager, V. A., Madore, B. F., & Freedman, W. L. 2013, ApJ, 777, 79
Mortsell, E., Goobar, A., Johansson, J., & Dhawan, S. 2021, arXiv:2105.11461
Nataf, D. M., Gonzalez, O. A., Casagrande, L., et al. 2016, MNRAS, 456, 2692
Oke, J. B., Cohen, J. G., Carr, M., et al. 1995, PASP, 107, 375
Perivolaropoulos, L., & Skara, F. 2021, PhRvD, 104, 123511
Pesce, D. W., Braatz, J. A., Reid, M. J., et al. 2020, ApJL, 891, L1
Peterson, E. R., Kenworthy, W. D., Scolnic, D., et al. 2021, arXiv:2110.03487
Pettini, M., & Pagel, B. E. J. 2004, MNRAS, 348, L59
Phillips, M. M. 1993, ApJL, 413, L105
Phillips, M. M., Lira, P., Suntzeff, N. B., et al. 1999, AJ, 118, 1766
Phillips, M. M., Wells, L. A., Suntzeff, N. B., et al. 1992, AJ, 103, 1632
Pietrzyński, G., Thompson, I. B., Graczyk, D., et al. 2009, ApJ, 697, 862
Pietrzyński, G., Graczyk, D., Gieren, W., et al. 2013, Natur, 495, 76
Pietrzyński, G., Graczyk, D., Gallenne, A., et al. 2019, Natur, 567, 200
Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A16
Planck Collaboration, Aghanim, N., Ashdown, M., et al. 2016, A&A, 596, A107
Planck Collaboration, Aghanim, N., Akrami, Y., et al. 2020, A&A, 641, A6
Poulin, V., Smith, T. L., & Bartlett, A. 2021, PhRvD, 104, 123550
Reid, M. J., Pesce, D. W., & Riess, A. G. 2019, ApJL, 886, L27
Riess, A. G., Casertano, S., Anderson, J., MacKenty, J., & Filippenko, A. V. 2014, ApJ, 785, 161
Riess, A. G., Casertano, S., Yuan, W., et al. 2021, ApJL, 908, L6
Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. 2019a, ApJ, 876, 85
Riess, A. G., Narayan, G., & Calamida, A. 2019b, Calibration of the WFC3-IR Count-rate Nonlinearity, Sub-percent Accuracy for a Factor of a Million in Flux, Instrument Sci. Rep. WFC3 2019-1, STSci
Riess, A. G., Fliri, J., & Valls-Gabaud, D. 2012, ApJ, 745, 156
Riess, A. G., Press, W. H., & Kirshner, R. P. 1996, ApJ, 473, 88
Riess, A. G., Yuan, W., Casertano, S., Macri, L. M., & Scolnic, D. 2020, ApJL, 896, L43
Riess, A. G., Kirshner, R. P., Schmidt, B. P., et al. 1999, AJ, 117, 707
Riess, A. G., Li, W., Stetson, P. B., et al. 2005, ApJ, 627, 579
Riess, A. G., Macri, L., Casertano, S., et al. 2009, ApJ, 699, 539
Riess, A. G., Macri, L., Casertano, S., et al. 2011, ApJ, 730, 119
Riess, A. G., Macri, L. M., Hoffmann, S. L., et al. 2016, ApJ, 826, 56
Riess, A. G., Casertano, S., Yuan, W., et al. 2018a, ApJ, 861, 126
Riess, A. G., Casertano, S., Yuan, W., et al. 2018b, ApJ, 855, 136
Rigault, M., Aldering, G., Kowalski, M., et al. 2015, ApJ, 802, 20

Ripepi, V., Catanzaro, G., Molinaro, R., et al. 2021, MNRAS, 508, 4047
Romaniello, M., Primas, F., Mottini, M., et al. 2008, A&A, 488, 731
Romaniello, M., Riess, A., Mancino, S., et al. 2021, A&A, 658, A29
Saha, A., Sandage, A., Labhardt, L., et al. 1996, ApJ, 466, 55
Said, K., Colless, M., Magoulas, C., et al. 2020, MNRAS, 497, 1275
Sandage, A., Tammann, G. A., & Reindl, B. 2004, A&A, 424, 43
Sandage, A., Tammann, G. A., Saha, A., et al. 2006, ApJ, 653, 843
Schöneberg, N., Abellán, G. F., Pérez Sánchez, A., et al. 2021, arXiv:2107.10291
Scolnic, D., Brout, D., Carr, A., et al. 2021, arXiv:2112.03863
Scolnic, D. M., Jones, D. O., Rest, A., et al. 2018, ApJ, 859, 101
Shah, P., Lemos, P., & Lahav, O. 2021, A&ARv, 29, 9
Silverman, J. M., Foley, R. J., Filippenko, A. V., et al. 2012, MNRAS, 425, 1789
Skowron, D. M., Skowron, J., Udalski, A., et al. 2021, ApJS, 252, 23
Soszynski, I., Poleski, R., Udalski, A., et al. 2008, AcA, 58, 163
Stahl, B. E., Martínez-Palomera, J., Zheng, W., et al. 2020a, MNRAS, 496, 3553
Stahl, B. E., Zheng, W., de Jaeger, T., et al. 2020b, MNRAS, 492, 4325
Steer, I. 2020, AJ, 160, 199
Stetson, P. B. 1987, PASP, 99, 191
Stetson, P. B. 1994, PASP, 106, 250
Stetson, P. B. 1996, PASP, 108, 851
Teimoorinia, H., Jalilkhany, M., Scudder, J. M., Jensen, J., & Ellison, S. L. 2021, MNRAS, 503, 1082
Tremaine, S., Gebhardt, K., Bender, R., et al. 2002, ApJ, 574, 740
Tremonti, C. A., Heckman, T. M., Kauffmann, G., et al. 2004, ApJ, 613, 898
Tripp, R. 1998, A&A, 331, 815
Verde, L., Treu, T., & Riess, A. G. 2019, NatAs, 3, 891
Visser, M. 2004, CQGra, 21, 2603
Whitmore, B. C., Zhang, Q., Leitherer, C., et al. 1999, AJ, 118, 1551
Wielgorski, P., Pietrzynski, G., & Gieren, W. 2017, European Physical Journal Web of Conf. 152, Wide-Field Variability Surveys: A 21st Century Perspective — 22nd Los Alamos Stellar Pulsation (Les Ulis: EDP Sciences), 07011
Yuan, W., Macri, L. M., Riess, A., et al. 2022, ApJ, arXiv:2203.06681
Yuan, W., Macri, L. M., Peterson, B. M., et al. 2021, ApJ, 913, 3
Zaritsky, D., Kennicutt, R. C., Jr., & Huchra, J. P. 1994, ApJ, 420, 87
Zhang, B. R., Childress, M. J., Davis, T. M., et al. 2017, MNRAS, 471, 2254
Zurita, A., & Bresolin, F. 2012, MNRAS, 427, 1463