# Precise cosmological analysis of Type Ia supernovae: the Hubble constant and dark energy

Bonnie R. Zhang

A thesis submitted for the degree of
Doctor of Philosophy in Astronomy
The Australian National University

June, 2018

# Declaration/Disclaimer

This thesis is an account of research undertaken between February 2014 and June 2018 at the Research School of Astronomy and Astrophysics in the College of Physical and Mathematical Sciences at the Australian National University in Canberra, Australia.

Except where acknowledged in the customary manner, the material presented in this thesis is, to the best of my knowledge, original and has not been submitted in whole or part for a degree in any university.

Bonnie R. Zhang
June, 2018

# Acknowledgements

This thesis could not have been completed without the vision, direction, and encouragement of my primary supervisor Brian Schmidt. Thanks Brian for your insight and support, for all of the scientific suggestions, for seeing me through to the end despite having a university to run, but most of all for your impressive ability to make me feel like everything will be alright.

I am also grateful to Tamara Davis, for showing me what it looked like to be brilliant, positive, and enthusiastic. Your genuine enthusiam for the science we are doing, and your dedication to the art of communicating clearly, set a wonderful example for me and undoubtedly many other students. I would also like to thank Chris Lidman and Tamara again for setting an example in your leadership and your efficiency, that has made OzDES the success it has been, and working in the Survey the joy that it was. Thank you to Anais for being not only a great collaborator, but also a great climbing/scuba/travel partner who has seen me through several injuries. I am immensely grateful to each and every person on my thesis panel: Brian, Tamara, Chris, Anais, Mike, Fang, and Richard, for your time and guidance, and good humour you've shown me. I have been beyond fortunate to have been surrounded by an as talented, supportive, uplifting, and inspiring group over the past four years.

Thank you to various physics and astronomy families: first my Physics Olympiad family, in particular Jake, Siobhan, Alix, Matt, and everyone else I've stayed up until morning testing exams with and tried unrecognisable international food with. I wouldn't be the same person today (as a scientist, teacher, student, person) without each of you, and I cannot begin to imagine what 2008–2016 would have looked like otherwise. Thanks to Mount Stromlo students over the years: thank you for being wonderful, funny, positive, cohesive, supportive – please never change. Thanks also to everyone else in the various supernova/cosmology/transient subgroups at Stromlo, for being fantastic people to work with. I am also grateful to astronomy/space department in Southampton, particularly the students and the supernova group, for your warm hospitality during November 2015 (and briefly in December 2016) – you have been an absolute joy to visit.

Thank you to my climbing partners, for holding my life in your hands and trusting me to do the same for you. For encouraging me to be braver and better, and reminding me of life beyond science. I am also grateful to my dear and valued friends for your company, support, and humour over the years.

I am most grateful to Dane: thank you for being brilliant, kind, and patient. Thanks for all the belays, adventures, and cocktails; for getting me out on weekends, and most recently, your self-sacrificing editing efforts. The past year and a half would have been nowhere near as much fun without you, and I could not have wished for a better person to have by my side.

Finally, to my parents Joanne and Kevin, and maternal grandparents Gong Gong and Por Por, I am indebted to you in a way only immigrant children can understand. Beyond this, I am thankful to you for giving me the freedom to stand up on my own, and for giving me the belief and confidence that I will make a right choice (if not 'the' right choice). If not for you, I would probably now be a doctor or banker or accountant or something. Thank for you doing everything for me except to view my achievements as a reflection of your own success.

# Abstract

The field of physical cosmology has advanced greatly in the last few decades. In this time, a holistic 'concordance' model of our Universe has been pieced together: an isotropic and homogeneous Universe which started very small, went through a brief inflationary period, and obeys General Relativity. Approximately a third of its energy content can be attributed to matter (mostly cold dark matter, with some baryonic), and the remainder to 'dark' energy, which appears to be a cosmological constant, tied to spacetime itself. Together, these components are referred to as the ΛCDM model.

Type Ia supernovae (SNe Ia) have been instrumental in reaching this understanding, playing a pivotal role in the late 1990s by signalling the Universe's accelerated expansion. Since then, SNe Ia and the other probes (baryon acoustic oscillations, weak lensing, and clustering) have been employed to improve constraints on the quantities that parametrise ΛCDM, and the equation-of-state parameter $w$ for dark energy. These efforts have converged in the Dark Energy Survey (2013–2018), which coordinates these probes. The paradigm for cosmology now is to diminish systematic errors to obtain precise measurements of cosmological parameters through observations far into the distant past.

More locally, the Hubble constant $H_0$ determines the expansion rate of the Universe at present; this normalises its distance scale. In the past several years, the precise value of $H_0$ has come into contention, with a significant discrepancy between results determined from a local distance ladder (with SNe Ia at the top), and those inferred from observations of the early Universe assuming the cosmological models that drive its expansion history. Taken at face value, this discrepancy could signal systematic errors in either measurement, or inaccurate assumptions about the model.

The efforts in this thesis use SNe Ia to address questions at both ends of the Universe's expansion history – the Hubble constant locally, and dark energy at higher redshifts – using the current best methods to account for the statistical and systematic terms which affect supernovae. These methods, were developed by the Supernova Legacy Survey (SNLS) and subsequent Joint Lightcurve Analysis (JLA), rely on covariance matrices to encapsulate correlated uncertainties in all lightcurve parameters, which are then propagated through probabilistic Bayesian parameter estimation methods to uncertainties in cosmological parameters.

In the works contained within this thesis, I have developed a framework for performing precise cosmological analysis of SNe Ia samples, including using covariance matrices to quantify systematic terms. I have applied to the aforementioned pertinent questions: the value of the Hubble constant and the nature of dark energy, using the data set in SH0ES (Riess et al., 2011) and intermediate Dark Energy Survey supernova data.

# Contents

Chapter 1

# Modern Cosmology

This first chapter will describe the current state of cosmology, encompassing the components of the standard $\Lambda$CDM 'concordance' cosmological model and how each of these was developed. Furthermore, the pressing questions still to be addressed at this point in time will be raised, and potential methods for studying them, including Type Ia supernovae (SNe Ia).

## 1.1 $\Lambda$CDM 'concordance' cosmology

This section narrates the development of the model of the Universe presently held most likely to be true, a so-called concordance of several components. In this model, the Universe is:

- expanding, having started very small and dense in a Big Bang; it possibly went through a brief inflationary period very early, then subsequently slowed down before eventually accelerating again at late times;
- spatially flat, isotropic, and homogeneous, with its gravity and spacetime described by General Relativity;
- having about a third of its energy content comprised of matter: a small portion of which is normal baryonic matter (i.e. interacting with light), and the remainder is cold dark matter (CDM);
- having the remainder of its energy content attributable to 'dark energy' (with negative equation of state, causing cosmic acceleration), which takes the form of a cosmological constant $\Lambda$, and is invariant in time;
- having formed its large scale structure gravitationally, with hierarchical growth of density perturbations.

Together, these components are given the shorthand $\Lambda$CDM. They were theorised, proposed, and later supported by both theory and observations. We detail these developments in the following subsections.

### 1.1.1 The expanding Universe

The theoretical conjecture and subsequent discovery of the expanding Universe came in the early 20th century, a time where astronomy and physics were being revolutionised. Previously, in the late 1800s, Maxwell's equations and the discovery of electromagnetism (that electromagnetic radiation is simply light, and encompasses the whole spectrum from radio waves to gamma rays), in combination with Newtonian mechanics, led to the illusion that our understanding of physics was complete. Einstein's *annus mirabilis* ('extraordinary year') works in 1905 shattered this illusion with not only the photoelectric effect demonstrating quantum behaviour, but the debut of relativity. The theory of special relativity (SR) built on thought experiments in different reference frames to highlight the incompatibility of electromagnetism with Newtonian relativism (i.e. transforms of reference frames). SR modified this mathematically, with a constant speed of light which could not be exceeded; this was valid for all inertial reference frames. This was extended to noninertial reference frames with general relativity (GR), which allowed for the effects of gravitation. GR contains the **equivalence principle**, the postulate that gravitational and inertial mass are equivalent: massive objects lead to the curvature of spacetime in the same way they experience it, through gravity. This was expressed mathematically in Einstein's field equations:

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} + \Lambda g_{\mu\nu} = \frac{8\pi G}{c^4}T_{\mu\nu} \tag{1.1}$$

which tie together the geometry and contents of spacetime. Specifically, the Ricci curvature tensor $R_{\mu\nu}$ and scalar curvature $R$ contain the spatial curvature, the stress-energy tensor $T_{\mu\nu}$ contains the source of this curvature, and the metric tensor $g_{\mu\nu}$ contains the spacetime metric. Here, the term $\Lambda$ is a 'cosmological constant' included to balance gravitational contraction, due to a preference for a steady state Universe, which was the paradigm at the time.

The Friedmann-Lemaître-Robertson-Walker (FLRW) metric (presented in Section 6.2, in Equation 6.3) solves Equation 1.1 exactly. From this, Friedmann's equations (Friedmann, 1922) (Equations 6.1 and 6.2) are derived, and describe the contents and scale evolution of an isotropic, homogeneous Universe within GR. Friedmann's equations allowed for an expanding universe; this was proposed explicitly by Lemaître (1927), proposing a proportionality between velocity and distance. This was soon observationally supported by Hubble's observations of 'extra-galactic nebulae' (Hubble, 1929), making use of Slipher's studies of spectra (Slipher, 1915). In the face of observed expansion, Einstein discarded $\Lambda$ and is often described has having lamented its inclusion as the 'greatest blunder' of his life.

Extrapolating expansion backward in time led naturally to an early universe that was extreme in density and temperature, beyond conditions have been observed or theorised, in which our understanding of physics may not hold. Within the Big Bang theory (BBT), the Universe started very small, dense, and hot – as a 'primeval atom' (Lemaitre, 1934) – that encompassed all of spacetime. Rather than conventional imaginings of expansion or explosion, spacetime itself expanded cohesively, without encroaching on an external space; similarly, there was no time before. This theory was not without controversy, derided by those who still clung to a steady state universe (Hoyle, 1948). Estimates of the state of the Universe at such extremes as the first millionth or so of the Universe is tenuous, requiring energies found in (and exceeding, prior to the first hundred-billionth of a second) particle accelerators. However, beyond this, the Universe could be viewed as a photon-baryon fluid, which as it cooled, resembled a blackbody. Applying the physics of blackbodies to the Universe as a whole led to prediction of a cosmic microwave background (CMB): a remnant afterglow everywhere from the 'surface of last scattering', when the Universe was last opaque. The prediction and subsequent discovery of the CMB is detailed further in Section 1.3.2.

### 1.1.2 Formation of structure; dark matter

In the almost 14 billion years since the Universe was in a hot plasma state, the contents of the Universe have coalesced into the now-observable large scale structure. Structure formation is an aspect of the Universe's development which is dependent on the effects of dark matter (DM). It is not a stretch to see how small fluctuations observed in the CMB could grow large, under gravity. However, as simulations of the early Universe – particularly $N$-body simulations of dynamical interactions (e.g. Blumenthal et al., 1984; Springel et al., 2005) – have improved, it has become increasingly clear that the level of structure visible today is much greater than could have occurred in the measured age of the Universe from those fluctuations alone, under the influence of only visible matter (e.g. Primack, 1997; Einasto, 2001; Bertone et al., 2005). Later with the suite of dynamical evidence from galaxies and galaxy clusters, considerations of structure formation have helped distinguish between models.

As a theory, dark matter has developed as a result of many disjoint pockets of observations. A thorough review of the literature over nearly the past century has been given in Bertone & Hooper (2016). We refer the reader to the above for a detailed study; here, we attempt to summarise these developments. In the beginning, early applications of the virial theorem to galaxy clusters found large velocity dispersions in the Coma (Zwicky, 1933) and Virgo (Smith, 1936) clusters. At the same time, Kapteyn (e.g. 1922); Oort (e.g. 1932); Zwicky (e.g. 1937) discussed possible unseen stars in our Galaxy with mass, thought to be faint or cold stars.

More insights were revealed by looking within galaxies themselves, specifically by examining circular rotation velocities of stars and gas in the spiral arms of galaxies. Rubin & Ford (1970) found that the rotation curve for M31 flattened out all the way to the outermost arms, rather than dropping off. A similar effect was noted in elliptical galaxies with their velocity dispersions (e.g. Faber & Jackson, 1976; Davies et al., 1983). Freeman (1970) fitted observed radial galaxy luminosities with theorised exponential disks to reconstruct mass distributions and rotation curves, and found that the latter peaked at larger radii than the luminosity could explain. These combined observations strongly indicated that the observed outermost regions of galaxies were not where the mass distribution stopped; a significant portion of massive material likely extends beyond the observed radius of the galaxy. This distribution of low-luminosity matter in outermost regions of a galaxy was called a **dark matter halo**.

Both galaxy clusters and galaxy rotation curves provided observational evidence for dark matter: in both, the motions of dynamical systems signified that they were influenced by more mass than was observed electromagnetically. By the 1990s it was generally believed that these effects were two sides of the same phenomenon, and pointed to some unseen mass (e.g. Einasto et al., 1974; Ostriker & Steinhardt, 1995). From a structure-formation perspective, DM is also necessary because the fluctuations in density observed in baryonic matter in the early Universe alone were insufficient to seed (in the age of the Universe so far) the hierarchical large scale structure (LSS) now observed. At the largest scales, filaments and sheets with voids in between, composed of superclusters, containing galaxy clusters. In the pre-existing view of regions of overdensity growing through gravitational attraction and collapse, something was missing with only the luminous matter observed so far.

Simulations reveal that the most satisfactory explanation is the presence of additional matter which has mass, ergo interacts gravitationally. Unlike ordinary or **baryonic matter**, DM does not interact electromagnetically. While it is crucial for structure formation by growing overdensities and attracting more matter, DM does not emit light. It is measured by a suite of cosmological methods (Section 1.3), while the relatively new field of astroparticle physics strives to directly detect hypothesised DM particles. Simulations have also been powerful for distinguishing between models; dim baryonic matter, termed massive compact halo objects (MACHOs), including brown dwarfs, neutron stars, and other less luminous but known objects, have been excluded (e.g. Alcock et al., 2000; Brandt, 2016).

Over time the term 'dark matter' has evolved from meaning matter which is not very bright to new exotic particles, such as axions, neutrinos, and/or weakly interacting massive particles (WIMPs). Today, dark matter involves many subfields of physics and astronomy, from particle physics to theory to galaxy dynamics through to cosmology.

The inflationary model for the very early Universe, initially theorised in Guth (1981) and recently reviewed in Baumann & Peiris (e.g. 2008); Linde (e.g. 2014), complements theories of dark matter in explaining observations. In this model, the Universe went through a period of rapid accelerated expansion, increasing in scale by some tens of orders of magnitude at some $10^{-36} - 10^{-32}$ seconds after Big Bang. The theory of inflation lies at the margins of the concordance model; by its construction, any configuration of the Universe before inflation would be wiped out and not observable. However, it is able to explain observations, notably including large scale homogeneity (including regions which otherwise could not have been in causal contact appearing to have reached thermal equilibrium), spatial flatness, and the lack of observed magnetic monopoles. The explosion in scale stretched fluctuations at a quantum scale to the perturbations in density necessary to seed structure formation, which later would grow gravitationally, amplified by cold DM.

### 1.1.3 Cosmic acceleration

Prior to the late 1990s, there had been hints of a yet unknown contribution to the energy density; these have been reviewed in Frieman et al. (e.g. 2008); Weinberg et al. (e.g. 2013). Most prominently, large scale structure interpreted as having formed with CDM pointed to a relatively low value of $\Omega_m$. Meanwhile, overall spatial flatness observed in the CMB (Section 1.3.2) and theoretically explained by inflation (Guth, 1981) implied that the total critical density should be unity, leaving a large portion of the Universe's energy content unaccounted for. With the other pieces of the concordance model having come together, some (e.g. Efstathiou et al., 1990; Ostriker & Steinhardt, 1995) suggested a cosmological constant was indicated prior to 1998. The acceleration of the Universe's expansion was the final piece in our understanding of the Universe today. The existence of dark energy was decisively cemented by two competing teams, the Supernova Cosmology Project (Perlmutter et al., 1999) and the High-Z Supernova Search (Schmidt et al., 1998; Riess et al., 1998). Both teams used SNe Ia on Hubble diagrams to measure the Universe's expansion at higher redshift than had been achieved up to that point, and each team independently found an accelerating universe.

The discovery of dark energy filled in a gap in our understanding of physics: an unknown component (later given the blanket term 'dark energy', much like the nomenclature 'dark matter') which causes the rate of the Universe's expansion to increase, contrary to the expectation that gravitation would lead to the Universe's slowing, and eventual collapse. As supernova cosmology matured as a field, subsequent efforts sought to measure this acceleration better in order to characterise dark energy, in surveys described in Section 2.4. Specifically, measurements are sought of cosmological parameters which include the relative energy densities $\Omega_m$ of matter and $\Omega_{\mathrm{DE}}$ of dark energy, and the equation-of-state parameter $w$. One hypothesis is that $w = -1$, and dark energy takes the form of a cosmological constant $\Lambda$, a constant term associated with empty space as initially predicted in Einstein's field equations. Initially $\Lambda$ was posed by Einstein as a curvature term, on the LHS of Equation 1.2 below (with the Einstein tensor $G_{\mu\nu}$ replacing the first two terms of the earlier formulation in Equation 1.1 and the $\Lambda$ term discarded)

$$G_{\mu\nu} = 8\pi T_{\mu\nu}, \tag{1.2}$$

but now it is usually considered as an energy contribution, i.e. a modification to the RHS. An alternative to the cosmological constant term is modified gravity: where the energy content remains the same, but gravity (the LHS), governed by GR, changes. Joyce et al. (2016) reviews theories of and tests for theories of modified gravity, which deviate from GR at larger scales, and in particular, violate the equivalence principle. These theories can be tested using large scale structure, with the growth rate of structure having different scale-dependence for theories of modified gravity compared to GR (Linder, 2005; Baker et al., 2014).

The quest to measure, categorise, and identify dark energy (or modified gravity) is one of the more pertinent open questions in modern cosmology. Current progress and future directions form a large part of this thesis; we continue to examine them but for now venture further into the unknown.

## 1.2 Questions near and far

Measuring the evolution of the physical size of the Universe over time via Hubble diagrams tells us how the Universe started, and hints at how it may end. Moreover, comparing observations to predictions inferred from models of the Universe allow us to probe the physics underlying the Universe's evolution. Two of the most significant discoveries governing our knowledge of the size scale and its evolution of the physical Universe were discovered in the past century: these were the aforementioned isotropic expansion (parametrised by the Hubble constant $H_0$), and the accelerating Universe, propelled by the thus far unidentified 'dark energy'.

The measurements of these two effects and the parameters describing them have been challenging, and have shaped our understanding of physical cosmology. The challenges facing their precise measurements have profound consequences for the concordance model, and potentially our understanding of physics. Now, almost twenty years after SNe Ia indicated the presence of dark energy, they remain vital as a cosmological probe, and further study may hold the answers to both questions. In 2018, the role of SNe Ia has moved well past discovery, and we are in a regime of performing precise measurements of cosmological quantities, in which longstanding approximations or ignored systematic effects must be reconsidered. We next turn to each of these questions, in both the local Universe and far into the past.

### 1.2.1 Dark energy as a cause for cosmic acceleration

Since the discovery of cosmic acceleration in 1998 (Riess et al., 1998; Schmidt et al., 1998; Perlmutter et al., 1999), the use of SNe Ia to measure distances has been paramount to probing cosmological models. Combining SN Ia data (redshifts from spectroscopy and distances from photometry) on a so-called Hubble diagram allows us to map the Universe's expansion history, thus extracting measurements of cosmological parameters. Subsequent studies (e.g. Wood-Vasey et al., 2007; Kessler et al., 2009b; Sullivan et al., 2011; Betoule et al., 2014; Scolnic et al., 2018) have vastly improved the statistical constraining power of the original measurement, pointing toward two conclusions: (i) dark energy is consistent with a cosmological constant (i.e. $w \sim -1$), and (ii) the error budget is now dominated by systematics, particularly in the photometric calibration. In the two decades since 1998, supernovae and other probes have cemented the ΛCDM concordance model (Chapter 1), and provide increasingly tight constraints on the numbers that parametrise this model.

The Supernova Legacy Survey (SNLS) performed an unprecedentedly large high-redshift (up to $z = 1$) SN Ia survey, discussed in detail in Section 2.4.3. SNLS found, in combination with complementary (orthogonal) cosmological probes $\Omega_m = 0.269 \pm 0.015$, only ∼4% baryonic matter, and $w = -1.061 \pm 0.069$ (Sullivan et al., 2011); or $w = -0.91^{+0.16}_{-0.20}$ from SNe Ia alone (Conley et al., 2011). These numbers cemented a dark-energy dominated Universe, albeit with significantly lower matter density than in Planck Collaboration et al. (2014) (which found $\Omega_m = 0.315 \pm 0.017$). The Joint Lightcurve Analysis (JLA) of SNLS and Sloan Digital Sky Survey (SDSS) supernovae improved on the inter-calibration between surveys and reduced errors relative to SNLS, finding $\Omega_m = 0.303 \pm 0.012$ and $w = -1.027 \pm 0.055$ combined with external probes (Betoule et al., 2014), soothing the discrepancy with Planck. SNLS and JLA have set the benchmark not only for cosmological measurements, but for the methods used in estimating systematics, particularly to reduce calibration uncertainties (a lesson from SNLS). The sample sizes in SNLS/JLA increased about tenfold in the decade since the accelerating Universe was discovered, and continue to dominate the known sample up until the Dark Energy Survey. Crucially, progress in the last decade highlighted that the error budget is no longer statistically dominated, having been overcome by numbers. Instead, systematic errors now form the bulk of the uncertainty, particularly photometric calibration.

In summary, we have known for two decades that dark energy exists, and evidence indicates that it is consistent with a cosmological constant. But we cannot conclusively say what physical form it may take. A goal at the forefront of modern cosmology is to observationally distinguish between different candidates for dark energy by measuring cosmological parameters, which include the equation-of-state parameter, $w$, for dark energy. Recent studies, SNLS and JLA, highlight that we are in a new era where statistical challenges have been overtaken, and instead the utmost priority is to precisely understand, account for, and minimise systematic uncertainties.

In this context, my work in this thesis involves the analysis of an Dark Energy Survey (DES) intermediate data set – the DES 3-year spectroscopic sample – using these methods to validate and compare (i) existing methods against newer analysis frameworks, and (ii) the constraining power of this data set against the existing JLA sample. In doing so we affirm that dark energy still has properties consistent with a cosmological constant Λ. This analysis is detailed in Chapter 6.

### 1.2.2 Local expansion: the Hubble constant

While studies of dark energy extend into the distant Universe, SNe Ia in the local Universe enable a measurement of its present expansion rate. The Hubble constant $H_0$ quantifies the proportionality between distance and recessional velocity, and has proven a challenge to measure ever since the discovery of the Universe's expansion almost a century ago, following the prediction of the latter in Friedmann's equations. As given in the Hubble law $v = H_0 D$ (first derived by Lemaître, 1927), $H_0$ sets the cosmic distance scale via the present expansion rate of the local Universe. First measurements of $H_0$ were approximate, made by comparing distances estimated for galaxies (see Section 1.3.1) to recessional velocities indicated by their redshifts. Today, the method for measuring $H_0$ via a distance ladder is generally similar, where the absolute peak magnitude of all SNe Ia is assumed constant (after corrections for secondary effects); alternate modes of measurement are detailed in subsequent sections. Making precise measurements of $H_0$ has been a continual challenge in observational cosmology, due to the difficulty of making accurate absolute distance measurements. The quest towards more accurate determinations of $H_0$ drove the Hubble Space Telescope (HST) Key Project (Freedman et al., 2001). The wider context for $H_0$ measurements up until 2011 had been reviewed in detail in Freedman & Madore (2010).

Recently, discrepant values obtained from local and global measurements have propelled the Hubble constant back into the spotlight. The tension between values of $H_0$ obtained from Planck observations of cosmic microwave background (CMB) anisotropies in the early Universe (propagated to the present time assuming size evolution of the Universe characterised by a standard ΛCDM model) and direct measurements from the local Universe lies at a $\sim 3\sigma$ level. Planck found $H_0 = 67.3 \pm 1.2$ $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Planck Collaboration et al., 2014) and $H_0 = 67.8 \pm 0.9$ $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Planck Collaboration et al., 2016), assuming a standard ΛCDM cosmology. $\sim 2.7\sigma$ lower than in Riess et al. (2011), In the local Universe, SN Ia and distance ladder measurements in the Supernovae nd $H_0$ for the Equation of State (SH0ES; most recently Riess et al., 2011, 2016) found a higher value of $H_0 = 73.8 \pm 2.4$ $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ in R11, and $H_0 = 73.0 \pm 1.8$ $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ in Riess et al. (2016). While the Planck measurement is dependent on an underlying cosmological model, the SN Ia-based measurement is model-independent. The precision of these values highlights the importance of the tension between the two modes of measurements.

In the context of this tension between local direct measurements and model-dependent values extrapolated from the early Universe, it is especially pertinent to independently re-examine current methods for obtaining a SN Ia-based measurement of $H_0$, incorporating the best practices (described in Section 3.4) for accounting for supernova systematics following the techniques embodied in SNLS and JLA. Thus, the reanalysis of the process in Riess et al. (2011, 2016) to determine $H_0$ using a sample of low-redshift SNe Ia, calibrated by a distance ladder of Cepheid variables – in turn calibrated by geometric methods including water megamasers in NGC 4258, detached eclipsing binaries in the Large Magellanic Cloud (LMC), and parallax measurements of Cepheids in the Milky Way – are necessary for understanding the tension in $H_0$, and form an important part of this thesis. We will affirm that the uncertainty in its value is still dominated by statistics, specifically limited by the small number of galaxies which anchor the supernova sample to an absolute magnitude/distance scale, by virtue of their containing both Cepheid variables and supernovae. Attention to detail in estimating this error is particularly important, as it reflects the seriousness of the discrepancy in early and late measurements of $H_0$.

## 1.3 Cosmological probes

This section contains an overview of the leading astrophysical tools used to measure cosmological parameters, including the Hubble constant. Apart from supernovae, we discuss other distance-based methods for measuring $H_0$: strong gravitational lensing and gravitational waves, as well as the CMB, baryon acoustic oscillations (BAO), weak lensing, and clustering.

### 1.3.1 Standard candles and other distance measures

Astronomical distances can be measured using **standard candles**, classes of objects with known or tractable absolute magnitude which, combined with their observed apparent magnitude, give a distance modulus. The techniques Hubble originally used in measuring the expanding Universe approximately followed this principle: if the brightest object in each 'nebula' (galaxy) was assumed to have the same absolute magnitude, then approximate distances to the galaxies could be determined. Objects now which act as standard candles typically have smaller dispersion. These include Cepheid variables and other pulsating variable stars such as RR Lyrae (described further in Section 4.2.1). Classical Cepheids are governed by the Leavitt Law (Leavitt & Pickering, 1912), tying their luminosities closely to their pulsation periods. The analysis in Chapter 4 rests on this relation, which will be presented in Section 4.3.2.

Type Ia supernovae are the canonical standard candle, with low dispersion and simple empirical form (Section 2.2.2), and are useful up to relatively high redshifts ($z > 1.5$ with current technology). On a larger scale, galaxy dynamics approximately indicate absolute luminosity, with velocity acting a proxy for size. The Tully-Fisher (Tully & Fisher, 1977)and Faber-Jackson (Faber & Jackson, 1976) relations relate galaxies' absolute luminosities, to the fourth power of either tangential rotation velocity (for spiral galaxies) or stellar velocity dispersion (for ellipticals); these galaxies are numerous and extend to higher redshifts, but have considerably higher scatter than supernovae.

These distances are often relative (in simple terms a difference between distances, or ratio between fluxes) rather than absolute, akin to the distinction between relative and absolute magnitudes (discussed in Section 3.1.1). Since each mode of measurement is useful only over a limited range of distances, multiple standard candles are tied together to form a so-called **distance ladder**. At the bottom of the ladder are absolute distances determined from geometric methods (i.e. trigonometric parallax, masers), only measurable at relatively small distances. Then nearby standard candles (i.e. Cepheid variables) give distances relative to this geometric scale; similarly, each rung of the ladder is calibrated on the previous. A recent approach modifies this by using baryon acoustic oscillations to set the distance scale in the early Universe, and using SNe Ia to extend this scale into lower redshift, rather than higher. This **inverse distance ladder** and the measurements it yields of $H_0$ are discussed in Section 5.2.2.

The use of supernovae as distance measures and cosmological probes, to address the questions raised in this chapter, will be the focus of the next chapter. It is using the distance ladder method above that SNe Ia, calibrated by Cepheids and geometric methods, are used to measure the Hubble constant, in the analysis in Chapter 4, in particular following equations in Section 4.3.1. For measuring dark energy, SN Ia magnitudes will be modelled as functions of redshifts and cosmological parameters, rather than distances derived from a distance ladder. As will be shown later (Sections 4.3.1 and 6.2), the absolute magnitudes of SNe Ia are degenerate with $H_0$. Thus while absolute distances are crucial for measuring $H_0$, relative magnitudes are sufficient for dark energy, as the cosmological parameters of interest do not depend on either $H_0$ or $M_B$. Nevertheless, good relative calibration remains vital for ensuring internal consistency within the supernova data (Section 3.1.1).

In addition to supernovae, two distance-based probes have provided measurements of $H_0$ from the local Universe: strong gravitational lensing, and the very recently detected gravitational waves (GW). The first observations of gravitational lensing (Eddington, 1919) and GW (Abbott et al., 2016), almost a century apart, have both been momentous discoveries which had been theoretically predicted by general relativity before they were searched for and discovered, in the latter case almost a century later. Now, these modes of measurement provide important independent validations of supernova-based measurements of the Hubble constant.

Gravitational lensing occurs due to the distortion of spacetime by massive bodies, curving the geodesic paths taken by light. This was first observed during a solar eclipse in 1919 when stars behind the Sun were deflected. Similarly, black holes, galaxies, and clusters can significantly bend spacetime, and if objects in their background are along a sufficiently close line of sight, their images will be distorted. In the strong regime, the distorted images are pronounced, appearing as multiple images, arcs, or Einstein rings. In these cases, light bends around a foreground object and travels along different paths. Not only does this produce multiple images, but the time taken differs, due to different gravitational potentials along each path. Thus, if the lensed image has a temporal component (e.g. quasars), the time delays can be used to calculate distances. The H0LiCOW program (Suyu et al., 2017) make a precise and blind measurement of $H_0$ this way by monitoring just five strongly lensed quasars, finding $H_0 = 71.9^{+2.4}_{-3.0}$ km s$^{-1}$ Mpc$^{-1}$ in flat ΛCDM (Bonvin et al., 2017).

The first discovery of gravitational waves in 2015 was a watershed moment, to be surpassed only by the first detection from a visible source, a binary neutron star merger, known as a **kilonova**. For the first time, spectra of the kilonova confirmed the presence of ejected heavy elements including platinum and gold synthesised during the NS merger. As the first concurrence of GW and electromagnetic radiation from the same object, GW170817 signalled a true beginning of multi-messenger astronomy, and was the first occurrence of a 'standard siren'. Schutz (1986) originally proposed the idea of measuring distances from (then theoretical) GW signals from inspiralling binary black holes and/or neutron stars, and developed further in Holz & Hughes (e.g. 2005). The phase of this gravitational wave reveals the 'chirp mass' of the binary system, which along with the redshift, allows the distance to be recovered from the amplitude. Four GWs had been discovered from the first, GW150914, up until GW170817, but the GW signal alone carried no redshift information. It was only through identification of the host galaxy that a precise measurement of redshift could be achieved, enabling a Hubble constant measurement of $70.0^{+12.0}_{-8.0}$ km s$^{-1}$ Mpc$^{-1}$.

### 1.3.2 Cosmic microwave background

The cosmic microwave background (CMB) has been the cornerstone piece of observational evidence for the Big Bang. Observations over the past two decades have also demonstrated the Universe's near spatial flatness, and constrained DM and parameters of ΛCDM. The CMB was theorised in 1946 (Gamow, 1946) as the afterglow of the Big Bang, a remnant from when the Universe turned transparent at recombination when it was approximately 380,000 years old. The last radiation, that was then coupled to baryons while the Universe was a blackbody, was freed (the 'surface of last scattering') and has since cooled to microwave wavelengths with a peak around $T \sim 3$ $K$. The first measurement of the CMB was by accident by Penzias & Wilson (1965), who detected a microwave signal in all directions, initially thought to be a noise source, that could not be eliminated. Subsequent satellite measurements of the CMB have increased significantly in resolution, with the COsmic Background Explorer (COBE; Smoot et al., 1992), Wilkinson Microwave Anisotropy Probe (WMAP; e.g. Hinshaw et al., 2003; Komatsu et al., 2009; Bennett et al., 2014), then Planck (Planck Collaboration et al., 2014, 2016) satellites.

The CMB angular power spectrum is the radio signal over the whole sphere, Fourier transformed and decomposed into spherical harmonics. The monopole term ($\ell = 0$) is the constant term with spherical symmetry (the aforementioned signal from recombination at $T = 2.7$ $K$); the higher order peaks, called **acoustic peaks**, represent the strength of the recombination signal at smaller angular scales. The initial theoretical prediction for the CMB was a nearly perfect blackbody spectrum, with tiny amplitude angular fluctuations that were the seeds for structure formation. In a triumph for cosmologists, these features were measured with remarkable agreement with COBE (Smoot et al., 1992; Mather et al., 1994). However, the low resolution of COBE of around 7° meant that only the largest scale terms were detected: the first acoustic peak, found in 2000 by MAXIMA (Hanany et al., 2000) and BOOMERanG (de Bernardis et al., 2000) balloon experiments, and later measured with extraordinary precision by WMAP (Hinshaw et al., 2003), occurs at $\ell \sim 200$. The locations and amplitudes of the higher order peaks − acoustic oscillations as photons, dark matter, and baryons all interact − can be theoretically predicted in terms of cosmological parameters, and are powerful in constraining these parameters. The angular sizes of the acoustic peaks constrain spatial curvature, while their amplitudes constrain densities $\Omega_b, \Omega_c$ of baryons and (cold dark) matter, respectively. CMB anisotropies directly constrain $\Omega_b h^2$ and $\Omega_c h^2$ (where $h = H_0/100$ km s$^{-1}$ Mpc$^{-1}$); they therefore provide a mode of measurement of the Hubble constant from the early Universe, modulo some assumptions. While inferences on acceleration are highly correlated with other parameters, the constraints on other parameters, particularly densities, make the CMB particularly powerful as a cosmological tool. In addition to geometry and constituents of the Universe, CMB observations have implications for inflationary models and primordial GWs (Komatsu et al., 2009).

### 1.3.3 Baryon acoustic oscillations

Baryon acoustic oscillations, the manifestation of the structure seen in the CMB within galaxy large scale structure, have emerged in the past decade as one of four leading probes for dark energy, as a complementary method to SNe Ia, at the opposite end of the Universe's timeline. While SNe Ia can be used to measure relative distances very well and constrain the present expansion rate, BAO provide an absolute distance scale at several discrete redshifts. This **acoustic length scale** $r_s$ is the comoving size of the sound horizon.

Fluctuations in density in the early Universe permeated as sound (compression) waves, which have been frozen at recombination, while photons escaped in the form of the CMB. These baryonic matter fluctuations appear as a compressed shell of overdense matter, at the characteristic scale $r_s$. Thus the early sound waves are now, at $z = 0$, imprinted in large scale structure as density fluctuations at this scale: these overdensities seeded the formation of galaxies at a preferred scale which can be measured today. The standard quantity in BAO measurements $r_s(z_d)$, is the sound horizon at the epoch of radiation drag (where baryons are free from the 'drag' of photons) in the early Universe, allowing their use as a **standard ruler**, analogous to the constant peak magnitude of SNe Ia allowing their use as standard candles. Its evaluation is a simple albeit model-dependent integral:

$$r_s(z_d) = \int_0^t \frac{c_s(t)}{a(t)}\, dt = \int_{z_d}^{\infty} \frac{c_s(z)}{H(z)}\, dz, \tag{1.3}$$

where $a(t)$ is the dimensionless scale parameter with first derivative $H(z)$ (the Hubble parameter). The integral in Equation 1.3 represents the distance a sound wave could have travelled from the Big Bang to shortly after the surface of last scattering. Thus, this quantity can be predicted accurately from theory using measurements from the CMB. Thus, like SNe Ia, BAO are part of the 'distance revolution' (Bassett & Hlozek, 2010), whereby objects of known size (magnitude) can be compared to their observed size (magnitude) to deduce accurate distances at high redshifts.

Standard rulers (objects of known size) have been proposed and theorised before BAO, including ultra-compact radio sources, double-lobed radio sources, galaxy clusters (Bassett & Hlozek, 2010). BAO are different from these sources in two ways: they are cosmological in origin, with a characteristic scale of structure originating from pre-recombination physics. To have a feature which has a size that can be modelled theoretically is unprecedented, within cosmological probes: $146.8 \pm 1.8 Mpc$ (Komatsu et al., 2009). This large scale, due to the relativistic sound speed $c_s$ before recombination, protects the BAO feature from non-linear structure formation at later times. The feature can be separated in transverse and radial directions, to delineate angular distance and Hubble parameter: the characteristic scales are $H(z)r_s$ and $D_A(z)/r_s$ respectively, where $D_A$ is the angular distance. Combining these gives the 'volume averaged' distance $D_V(z)$.

BAO are a 'statistical' standard ruler in that their signal is weak; like weak lensing, their constraining power lies in their statistics. Their detection is reliant on aggregate measurements of many galaxies over large cosmic volumes, potentially limited by cosmic variance. Large samples of objects, such as luminous red galaxies (LRGs) or quasars, with known redshifts are required for BAO measurements.

BAO were first theoretically proposed in Eisenstein & Hu (1998), stipulating the existence of a baryonic or matter power spectrum, separate from the CMB. Blake & Glazebrook (2003) demonstrated requirements for experimentally detecting this galaxy power spectrum. Soon after Eisenstein et al. (2005) measured BAO using SDSS luminous red galaxies (LRGs), followed by WiggleZ (Blake et al., 2011), 6dFGS (Beutler et al., 2011), and the Baryon Oscillation Spectroscopic Survey (BOSS; Alam et al., 2017; Beutler et al., 2017). Combined with CMB constraints, BAO allow the measurement of dark energy and $H_0$ by extrapolating the characteristic scale theoretically predicted in the early Universe to later times and comparing with observations at given redshifts.

### 1.3.4 Weak lensing and galaxy clusters

Gravitational lensing was introduced in Section 1.3.1, as the bending of spacetime, causing distortion of background light, due to an intervening mass in the foreground. In the strong regime, multiply lensed images appear, allowing direct use for measuring distances. Where this effect is weaker, individual distortions of galaxy images are small and not discernible, but instead, they can be studied statistically as an aggregate. This technique, **weak lensing**, is a powerful probe for large scale structure, lending its use to study dark matter and dark energy through their influences on structure formation. Weak lensing and its role as a probe for dark energy are reviewed in Huterer (e.g. 2010); Kilbinger (e.g. 2015); Mandelbaum (e.g. 2017).

The ellipticities of galaxy images change when lensed, even weakly; the term **cosmic shear** is given to the aggregate distortion of galaxy images due to large scale structure. This had been predicted theoretically by e.g. Miralda-Escude (1991); Kaiser (1992), but required large samples of foreground galaxies to overcome noise. As with BAO, large galaxy samples were needed to reach the statistical power required, in this case to detect cosmic shear. Detection occurred in 2000 in four independent studies (Bacon et al., 2000; Kaiser et al., 2000; Van Waerbeke et al., 2000; Wittman et al., 2000). Anything with great mass can act as a lens, including galaxies (specifically their dark matter haloes) and the greater cosmic web. By studying the effects these lenses have on the background objects, the mass distribution of clusters is revealed; meanwhile, this can be modelled to great precision. Tracing these images informs more generally the large scale structure, which can be compared to predictions from cosmological models and $N$-body simulations to infer the parameters which governed structure formation.

The most common figure-of-merit is the two-point correlation function in Fourier space, the lowest order statistic for the spatial distribution of galaxies; this can correlate weak lensing or galaxies independently and together. The foreground (lens) galaxy positions and lensing shear of background (source) galaxies, at higher redshift, are studied in in shear-shear ('cosmic shear'), galaxy-shear ('galaxy-galaxy lensing') and galaxy-galaxy correlations. These effects reflect the scale-dependence – the closer a pair of galaxies are, the greater the likelihood that they are affected by the same foreground image – of the distribution of mass at different redshifts, indirectly allowing us to measure dark energy, which impacts the clustering of matter (reducing it with accelerated expansion). The addition of redshift information, for example from OzDES (Section 2.4.5) for weak lensing in DES, allows for tomography.

## 1.4 Dark Energy Survey

The Dark Energy Survey (DES) is a multi-probe observational campaign seeking to provide the next generation of constraints on cosmic acceleration. DES observations are coordinated and shared to maximise science goals, between SNe Ia and three other probes: baryon acoustic oscillations, weak lensing, and galaxy clusters. The survey is broadly described in Dark Energy Survey Collaboration et al. (2016). The programme is running from 2013 – 2019, with supernovae observations having concluded in early 2018. Recently the first data release was published in Abbott et al. (2018). Key cosmological results from the first year of observations were published in Dark Energy Survey Collaboration et al. (2017) with combined results from galaxy clustering and weak lensing in Dark Energy Survey Collaboration et al. (2017), and measurement of the BAO scale to $z \sim 1$ in The Dark Energy Survey Collaboration et al. (2017). The survey has operated at Cerro Tololo Inter-American Observatory, on the 4-m Blanco telescope with the purpose-built Dark Energy Camera (DECam), a 570 Megapixel digital camera comprising 62 charge-coupled devices (CCDs). Over the six years, DES images have built up repeated observations of a composite 5000-square-degree area over the southern sky.

This ambitious project was the natural next step as each probe individually had matured to the point where they would benefit from the multitude of data from a targeted observing campaign, difficult to acquire from traditional observing programs. This approach is reflective of the evolution of astronomy toward wide-field surveys on purpose-built cutting edge instruments on large telescopes – particularly in cosmology. These surveys are often referred to as e.g. dark energy 'experiments', akin to large particle physics experiments and associated collaborations. They inherit many of the same characteristics: a singular instrument, collaborators in the hundreds or thousands spread across continents, with institutions working on separate parts of the experiment, complete with proprietary data, that has been guaranteed based on significant infrastructure (instrument and software engineering) investment. This 'experimental' approach is an evolution of the preceding Sloan Digital Sky Survey, which sought to bring about a wealth of data (imaging and spectroscopy of the entire sky above the equator), mined for science by a large team, augmented by many smaller teams with a range of scientific goals.

The key arms of science in DES (structured as working groups) differ greatly in methods and physics used, but benefit from the shared infrastructure. Challenges then are to determine common observing strategy, bearing in mind the requirements of the various science goals. The two prongs of the observational strategy (deep and wide) allow different goals. The deep survey searches a small area for time-domain science, consisting of a rolling search with a short cadence over ten smaller fields, with the aim of detecting SNe Ia and measuring their lightcurves. The wider survey covers large areas of the sky with deep exposures, for science where wider coverage is necessary – BAO, clustering, weak lensing – which may be limited by cosmic variance. As well as reaching these observational goals, much theoretical effort is required for simulations in various areas and to combine or cross-correlate different probes.

SNe Ia observations form a crucial part of DES, addressing the questions discussed in Section 1.2.1. While we have discussed DES generally here, its role in the chronology of SN Ia surveys is presented in Section 2.4.5; Chapter 6 will be dedicated to an analysis of DES supernovae, in the quest to measure dark energy.

## 1.5 Role of blind analysis

The analyses within this thesis are all performed blind, that is, with the final results of interest hidden from the people performing the analysis, team members, and all others, until all facets of the analysis are complete. More generally, blinding refers to the method of obscuring some aspect(s) of the analysis (usually a result) until it is complete. Within this paradigm, debugging, checking and critique of analysis methods all take place and are resolved before numerical results are revealed. Once the reveal takes place, no further modifications occur; if further changes to analysis are necessitated by some oversight, then the results following these changes cannot be referred to as 'blind'.

The overarching reasons for blind analysis are to prevent any human effects, particularly biases, on experimental results. Possible human influences, most often subtle and unconscious in its nature, can occur without experimenters even having any knowledge of it taking place. Thus, it is insufficient to simply state that we are perfectly rational, neutral, or unbiased. Human effects are complex, and can impact any study where there are judgements or decisions to be made – for example, choices of algorithms, or cuts, or a point at which to stop debugging and accept the analysis as complete – many of which are subtle and not obvious.

Knowledge of human factors and measures to preclude these effects have been long established in other scientific ventures and are by no means new; one common example is the placebo effect and double blind clinical trials. The placebo effect is the positive impact of a patient thinking they are receiving medical treatment, even when there is no medically active treatment. It is used as a control treatment in medical trials, to isolate the effect of a treatment under test, from the non-medical aspects of the treatment. The placebo effect is not as simple as being fooled by a sugar pill: the positive effects arise from factors including increase bodily awareness, psychological effects of being cared for, psychosomatic effects, all forming a complex web of interaction. Moreover, even an experimenter or doctor knowing the treatment group can impact the patient's progress, in subtle and unconscious ways, motivating the practice of double-blind studies, where neither subject nor experimenter is aware of the treatment group, for studies where humans are the subjects.

Such blinding techniques have been adopted in the physical sciences to exclude human effects, most commonly in large scale and difficult to replicate experiments, such as high-energy or particle physics experiments, or theoretical or new detections such as the Higgs boson and gravitational waves. Blinding is now standard practice in many particle physics experiments, in part motivated by studies of Particle Data Group (PDG) measurements of dozens of quantities over time, converging and/or demonstrating tight agreement and small scatter to extents not supported by their measured uncertainties (Roodman, 2003). Similar effects have been demonstrated in cosmology: the Croft & Dailey (2011) metaanalysis of measurements of cosmological parameters over several decades reveal clustering akin to PDG studies, particularly in densities $\Omega_m$ and $\Omega_\Lambda$. These low-scatter measurements either indicate overestimated errors, or could point to possible confirmation bias. Four years later, Maccoun & Perlmutter (2015) recommended that more fields, in particular cosmology, follow particle physics in adopting blind analysis. Blinding in cosmology is still relatively new: H0LiCoW (discussed in Section 1.3.1) and Zhang et al. (2017) (which forms the basis of Chapter 4) are examples of applying this to $H_0$; KiDS (Kuijken et al., 2015) is a notable example of end-to-end blinding in weak lensing, where three versions of the analysis (including two fake) were performed to the extent where paper drafts were written up. As experiments expand in scale and budget, the ability to replicate experiments drops to virtually zero. Astronomy is evolving as a field toward all-sky surveys with purpose-built telescopes, such as LSST. Science is done in larger collaborations with enormous surveys, increasingly resembling particle physics experiments. Then, with one chance to get crucial measurements right, blind analysis is necessary to make the most of each experiment.

In addition to the above practical reasons, there is the more fundamental argument that the numerical value of a result (and whether or not this agrees with expectations) alone does not reflect on the validity or correctness of the analysis method. Thus, it does not need to be known during the experiment or analysis, and should be obscured to the extent where it is possible to complete the full analysis without knowledge of the result.

Methods for blinding can occur at several points in an analysis chain or in combining multiple analyses, and range from simple to very complex. In Chapter 4 we take a relatively simple approach, where parameters (or more precisely, marginalised posterior distribution functions for parameters) are shifted or scaled by unknown but retrievable random offsets. These transformations are reversed at the time of unblinding to obtain true parameter values. These methods will be presented in more detail in Section 4.3.4. However, more complex approaches to blinding are necessary for measurements in Fourier space, in experiments such as clustering and weak lensing (Kuijken et al., 2015).

For the Hubble constant, the stakes are particularly high presently, in the face of the tension in its value based on different methods; its numerical value is currently the subject of intense scrutiny. As discussed in Section 1.2, new measurements of both $H_0$ and dark energy have the potential to reveal a fundamental gap in our standard $\Lambda$CDM concordance model and indicate new physics. The next step in probing the standard model with SNe Ia is to obtain more precise measurements, with a focus on detailed treatment of supernova systematics. Furthermore, the high stakes emphasise the necessity of precluding the effects of any human biases on analyses, where possible – recommending the application of blinding methods – established in medicine and commonplace in particle physics – to cosmological analyses.

## 1.6 Where this thesis sits

This thesis details our precise study of type Ia supernovae to address key fundamental questions in modern cosmology. We apply the current precise and well-understood supernova analysis framework (of SNLS/JLA) in a Bayesian context, consisting of a Monte Carlo maximal likelihood parameter estimation including, within the likelihood, covariance matrices for each systematic term. For each source of systematic error, covariance matrices encapsulate correlated uncertainties in supernova observables between all supernovae in the sample. We focus our efforts on exploring the current tension in the Hubble constant, and paving the way toward more precise measurements of dark energy and cosmological parameters, by analysing SNe Ia in the Dark Energy Survey. Overall, we set out to constrain the cosmic distance scale better than ever before, and question the standard cosmological model by probing the physics that underlies measurements of the Universe's expansion.

Looking into the future, these two scientific questions both have significant consequences. The Hubble constant tension may signify new physics, if it remains unresolved. Meanwhile there is much to be learned about dark energy with the imminent growth of data sets; if a cosmological constant is excluded by data being probabilistically inconsistent with $w \neq -1$, an alternate form of dark energy is required. Precision in both contexts are essential. The profound questions we now want to solve include: is the standard theory of cosmology valid? What is the physical form of dark energy? And what causes the tension in the Hubble constant, as measured in the nearby and distant universe, and where will this discrepancy lead us?

This thesis covers the work in the Zhang et al. (2017) redetermination of the Hubble constant and an analysis of the 3-year spectroscopic Dark Energy Survey data set. Having illustrated the context of this work in this introduction (Chapter 1), we will delve further into details of using SNe Ia in cosmology in Chapter 2, including methods and statistics. We will next face in Chapter 3 the general astronomical and statistical methods used throughout, including details of constructing covariance matrices to compute SN Ia systematic terms, which form an important part of the subsequent parts of the thesis. Chapters 4 and 5 focus on the Hubble constant, with details of the analysis in Zhang et al. (2017) and reflections on its place in the $H_0$ literature as well as future directions, respectively. Chapter 6 focuses on a cosmology fit using SNe Ia in the Dark Energy Survey, and Chapter 7 summarises and looks to the future.

Chapter 2

# Type Ia Supernovae

This chapter discusses the use of type Ia supernovae as a cosmological probe in more detail, after their introduction in Chapter 1. I will review the development of supernova cosmology, from the history of studying supernovae to the increasingly sophisticated surveys that have built larger samples. I review methods of retrieving distances from lightcurves, and the subtleties of studying SNe Ia as a collective sample.

## 2.1 Supernovae: an overview

In history and folklore, supernovae have held a significant place in the imaginations of human civilisations, due to their bright and transient nature. For fleeting periods of just days, these exploding stars have outshone entire galaxies. Supernovae in our Galaxy, bright enough to be observed with the naked eye, have been exceedingly rare; nevertheless a small handful have been recorded over millenia. Chinese astronomers had recorded supernova sightings as early as SN 185, calling them *ke xing*, or guest stars. There is substantial evidence (e.g. Hamacher, 2014) that older indigenous civilisations had observed novae and supernovae much earlier, interpreting these events in art and through stories. In the Western world the more notable Galactic supernovae have been SN 1572 ('Tycho's supernovae', known then for disrupting the view that stars existed in a permanent and ethereal sphere), and SN 1604 ('Kepler's supernovae'), the most recent in our Galaxy. Long after their explosions, supernovae have left in their place colourful nebulae known as **supernova remnants** such as the Crab Nebula (once SN 1054, recorded by the Chinese). Even in early times, supernovae were recognised as rare and extreme events. Unlike the planets, which also appeared in new places but had patterned movements and could be tracked, supernovae seemed to appear from nowhere, without warning.

In modern (20th century) western astronomy, larger telescopes, the advent of spectroscopy, and new understandings of stars (and their place within the Galaxy and wider Universe) have increased theoretical and observational interest in supernovae. Supernovae were postulated as the end products of stellar life cycles, through the convergence of theories of stellar nucleosynthesis and evolution. Observations of stellar populations on Hertzsprung-Russell diagrams or in colour-magnitude space, as well as chemical compositions of stars and supernovae ejecta deduced from spectral analysis, supported this theory.

Scientifically, supernovae grew to represent the confluence of disparate astrophysical interests: (i) the high-energy Universe: they are extremely violent, related to other relatively unknown stellar end-of-life products of neutron stars and black holes, and possible sources of cosmic rays and/or positrons; (ii) stellar nucleosynthesis: they are responsible for generating almost all of the metallic (i.e. heavier than helium) elements in the Universe; (iii) star formation and interstellar medium physics: they chemically enrich the interstellar medium by dispersing these metals, to be used in the formation of stars and solar systems.

The sheer superlative brightness of supernovae marks them as useful. They provide a unique way to probe stars (or their ends) in distant galaxies and at different epochs, a mode of study otherwise not available. Most evidently, supernovae are immensely useful for cosmology as standard candles, objects used to measure distance with remarkable precision far into the distant and past Universe. As will be discussed in Section 2.2, they are invaluable probes of the Universe's expansion history, from local recession through to cosmic acceleration.

### 2.1.1 Supernova types

Historically, the separation of supernovae into two broad categories has occurred phenomenologically based on the absence or presence of hydrogen lines in their spectra, type I and type II, respectively (Minkowski, 1941). Silicon and helium spectral lines are used to further separate type I supernovae with supernovae, those exhibiting Si II features classified as Ia, those with He I as Ib, and those with neither as Ic. Meanwhile, the further subclassification of type II supernovae is based on photometry. Most are classed as IIL, with linearly declining lightcurves, and IIP, with plateauing lightcurves.

However, spectral features of type Ib/c supernovae show them to be more similar to type II SNe than to Ias, in terms of astrophysical origins and explosion mechanisms. A more accurate separation of supernovae is into Ias – the result of a thermonuclear detonation of a carbon-oxygen (CO) white dwarf (WD) in a binary system – and core-collapse supernovae, comprising Ib/cs and type IIs alike. We will discuss Ias in more detail in Section 2.1.3, covering their observational signatures, theories for their progenitor systems and explosion mechanisms, and their standardisation and diversity. Here, we will briefly cover core-collapse supernovae (the main contaminants in samples of Ias gathered for cosmological studies), superluminous supernovae, and other peculiar supernovae which do not fit in these categories.

#### Core-collapse supernovae

Massive stars (typically with mass greater than $\sim 8M_{\odot}$) end their lives exploding as core-collapse supernovae (CC SNe). Along with violent black hole and/or neutron star mergers that produce gravitational waves (mentioned in Section 1.3.1), CC SNe are some of the most energetic astrophysical events known. Spectra of CC SNe show broad absorption lines indicating high velocities of ejecta (Filippenko, 1997), with hydrogen and helium lines indicating the makeup of the outer ejected shell: hydrogen lines in spectra of type II SNe suggests that the outer layers of the exploding star are dominated by hydrogen, while SNe Ib/c have been stripped of their hydrogen envelope, and SNe Ic have also lost their helium envelope. From a short time after explosion, type II SNe appear as largely featureless continua, except for type IIn supernovae, which exhibit narrow emission lines with little P Cygni absorption in their spectra. A smaller subclass, the type IIb, spectroscopically resemble Ibs except with hydrogen lines.

The core-collapse process begins towards the end of a massive star's life, as it runs out of fuel. Through nucleosynthesis, the material in the star's core increases in atomic number through successive nuclear fusion processes, until it runs out of fuel upon reaching iron: beyond this point, fusion is no longer energetically favourable. Unable to continue burning fuel, the star succumbs to inward gravitational pressure and begins collapse, until its core exceeds the Chandrasekhar mass (a commonality with SN Ia progenitors, and discussed in more detail in Section 2.1.3) where electron degeneracy pressure can no longer counteract gravitational collapse. This collapse accelerates as the core increases in temperature and more energy escapes with neutrinos. The inner core eventually rebounds when the strong nuclear force becomes dominant, resulting in a reflected shock wave which expands outward, accelerated by neutrinos and enabling further nucleosynthesis as a shell of extreme density passes through the outer layers of the star. The outer layers of the star are expelled, leaving at its centre a neutron star or a black hole (dependent on the initial stellar mass).

#### Superluminous and peculiar ('weirdo') supernovae

A newer class of transient, outshining known supernovae (Ia and core-collapse) by over two magnitudes, has only been studied in the past dozen years. The label 'superluminous supernova' (SLSN) has been applied to any supernova with absolute magnitude less (brighter) than $-21$ mag in any colour – compared with $\sim -19$ for SNe Ia and from $-14$ to almost $-19$ for CC SNe (distributions shown in Arcavi et al. (figure 1 of 2016)). SLSNe are of interest because their spectra resemble SNe IIn or Ic most closely, but

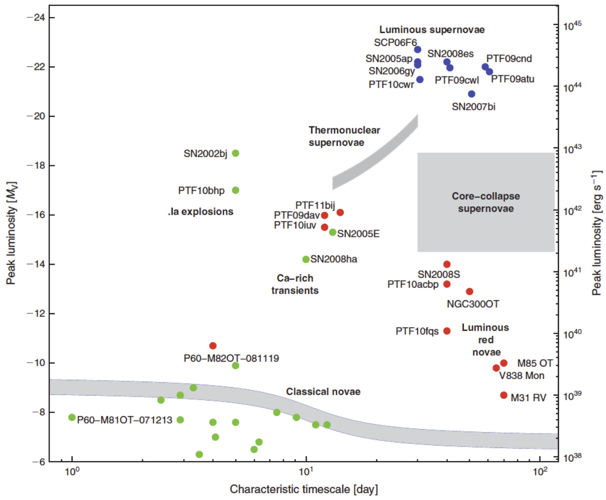


Figure 2.1 Optical transients as reviewed in (Kasliwal, 2012, figure 1), separated by peak luminosity and characteristic timescale. Recent (at the time) objects populating the luminosity gap between novae and SNe include .Ias and Ca-rich transients (Kasliwal et al., 2012), and numerous individual peculiar objects.

their lightcurves, in particular their extreme brightness, indicate that an entirely different power source (possible a magnetar) is required. Their extreme brightness compared to SNe Ib/c suggest that their progenitors are extremely massive, at the upper end of known stellar masses (e.g. Gal-Yam, 2012).

Beyond the common and less common classes of supernovae discussed thus far, many objects do not fit neatly into any of the aforementioned categories (Ia, core-collapse, and superluminous), yet have had lightcurves and spectra that clearly indicate they are supernovae. Notably, until around 2005, almost all known optical transients had fit into the categories of thermonuclear (type Ia) or core-collapse supernovae, and the much less luminous classical novae (burning on a WD driven by accretion). Many more 'weirdo' objects not fitting these categories have been discovered since, through searches in archival data or in SNLS (Section 2.4.3). These have been comprehensively reviewed in Kasliwal (2012), in particular the peculiar objects in the gap between novae and supernovae. These include, for example, so-called '.Ias' – explosions in ultra-compact binary WD systems (Bildsten et al., 2007; Shen et al., 2010) – and calcium-rich transients discussed in Kasliwal et al. (2012). These objects and others are separated by peak luminosity and characteristic timescale, in Figure 2.1 (Kasliwal, 2012, figure 1). In addition, specialised surveys over the past decade have focused on discovering peculiar objects, notably the Palomar Transient Factory (Rau et al., 2009; Law et al., 2009) and its successor, the Zwicky Transient Factory (Bellm, 2014). In the other direction in magnitude, a small handful of so-called 'superluminous-gap' objects in the gap between supernovae and SLSNe have been recently studied in Arcavi et al. (2016), to further understand the nature of these objects and SLSNe, as well as explosion mechanisms.

### 2.1.2 Supernova classification methods

There are several pieces of photometric evidence that indicate that an optical transient is likely a supernova (see Section 2.4.2). Nevertheless, the most reliable way to confirm an object's identity as a supernova, and to ascertain its subclass, is through spectroscopy. This is not to say that the photometric signals are useless: they are invaluable as evidence that a transient is a supernova candidate and that spectroscopic and other photometric resources should be dedicated to observing it in a timely manner.

In the past, telescope limiting magnitudes have meant that spectra were usually taken close to maximal brightness, near the same phase in the lightcurve. Hence, older spectral templates had not taken into account much spectral evolution of the supernova, and most observed spectra of any class of supernova appeared similar. In the past few decades improved technology has allowed both earlier detection of supernovae, and observations well after peak, where the supernova is several magnitudes dimmer. Thus, there are now three main variables which determine the appearance of supernova spectra: the identity (type) of the supernova, its phase, and its redshift (which roughly works out as a translation or transformation along the wavelength axis, but also includes observational effects – notably $K$-corrections, mentioned in Section 2.2.3). Of course, as indicated in Section 2.1.1, there are objects which are undeniably supernovae, yet fit in none of the primary types of supernovae (Ia, Ib/c, II); moreover in each category there is variation in the objects. The variation within Ias, and modern methods for ameliorating these corrections for use in cosmology, is described in Section 2.2.3. However as a first approximation, matches for redshift, type, and phase are quite successful.

Spectroscopic classification methods rely on libraries of templates, and matching observations to these libraries. Two primary routines for classification and determining redshifts have emerged in the 2000s: SNID (Blondin & Tonry, 2007) and `Superfit`.[1] SNID uses cross-correlation techniques (e.g. Tonry & Davis, 1979) to maximise a metric for goodness of correlation between input and template supernovae spectra. The redshift and phase of a test supernova can be determined by maximising this, using a library of various SN templates, and the cross-correlation value translates into an error on the redshift. `Superfit` is an alternative template-matching routine, with the additional ability to take into account a host galaxy contribution to the spectrum. The routine fits an input spectrum as the sum of supernova and galaxy components (the latter can be zero), testing over a range of redshifts, supernova types, phases, and with reddening corrections.

Recently, as machine learning (Section 3.3) techniques have benefited from a rapid increase in effectiveness, it has been applied to spectral template fitting. The Deep learning for the Automated Spectral classification of supernovae algorithm Muthukrishna et al., (in prep.) makes use of the `TensorFlow` (Abadi et al., 2015) machine learning framework and spectral templates in the Berkeley Supernova Ia Program (BSNIP; Silverman et al., 2012) to train a neural network to classify and determine redshifts and phases for supernovae.

Machine learning has also contributed to classification methods that remove reliance on spectra altogether, with obvious benefits as the capacity of surveys to detect greater numbers of transients continues to increase over time. Photometric classification methods which are reliant on cuts, machine learning, or both, have had remarkable success. Thee development of these methods has been closely related to the Supernova Legacy Survey Section 2.4.3 and machine learning methods, so we will postponed discussion until Section 3.3.3.

### 2.1.3 Type Ia supernovae physics

The study of supernovae works in two extremes: the very general (treating all SNe Ia as the same and correcting for, or smoothing out, their differences), and the very specific: detailed observations of individual objects and probing the physics which leads to these observations. Supernova cosmology as a necessity tends towards the former, but the latter are not to be discounted for the value they add to our understanding. From a cosmology perspective, the better we understand the systematic effects which lead to diversity, or physical processes behind observations, the better we can standardise objects. The homogeneity of Ia samples for cosmology has immense power as a statistical sample for placing on a Hubble diagram and constraining cosmological parameters (Sections 1.3.1 and 2.2), and yet the knowledge gained of individual supernovae as astronomical objects and former stars is vital to our knowledge of their explosion mechanisms, progenitors, and astrophysics in general. This knowledge will aid our analyses of supernovae into the future, including for cosmology, by allowing us to examine, estimate, and reduce systematic uncertainties.

In this section, we provide an overview of the physics of type Ia supernovae, and the questions currently relevant to their study, to give context to discussions of their standardisation in Section 2.2.

[1] https://github.com/dahowell/superfit; http://www.dahowell.com/superfit.html

### Observables: lightcurves and power sources

Without their uniformity, SNe Ia would not occupy their place as a powerful cosmological probe, and as precise distance indicators. The simplicity of this consistency is alluring, and founded on a sound theoretical basis: the hard limit of the Chandrasekhar mass. This seemingly places all SNe Ia at the same mass when they explode, resulting in remarkable homogeneity within SN Ia populations. The reality is slightly more subtle, and this view is an approximation. Like many approximations, it has served us well in cosmology: as a stepping stone to improve upon.

The **Chandrasekhar limit** (Chandrasekhar, 1931) at approximately $\sim$1.4 $M_\odot$ is a theoretically calculated maximum mass of a white dwarf. Being extremely dense (around a million times more dense than water), their extreme inward gravitational force is balanced by electron degeneracy pressure. Electrons are fermions, not able to multiply occupy the same state by the Pauli exclusion principle; consequently the higher energy electrons result in strong outward pressure exerted. However, at the Chandrasekhar mass, this pressure is no longer sufficient to counteract gravitational collapse. The star either undergoes collapse into a neutron star, or if fusion is triggered by an increase in mass (e.g. accretion of material in a binary system), it can explode as a type Ia supernova. We explore the commonly proposed scenarios for the circumstances in which this explosion takes place or is triggered, reviewed in Hillebrandt et al. (2013); Ruiz-Lapuente (2014); Maeda & Terada (2016).

Observational signatures have long established that SNe Ia explode as runaway thermonuclear explosions of a CO WD in a binary system with another star. The chemical compositions of SN Ia progenitor systems have been discerned from the evolution of spectral lines, as a function of velocity, which signal the nature of the ejecta (e.g. Branch et al., 1995). Spectra of SNe Ia are characterised by P-Cygni absorption lines superimposed on a blackbody continuum at early times, supporting a layered ejecta with iron-peak elements in the centre and intermediate mass elements such as silicon and sulfur on the outside (Stehle et al., 2005; Mazzali et al., 2007). The spectrum at late times (in the nebular phase) exhibits only forbidden lines from iron-peak elements with no continuum.

As with SN Ia spectra, the lightcurves of type Ia supernovae can be understood in terms of the physical mechanisms driving the explosion, in particular the radioactive decay of $^{56}$Ni. The decay of $^{56}$Ni $\rightarrow^{56}$Co $\rightarrow^{56}$Fe entirely powers the luminosity of SNe Ia, starting from the thermonuclear detonation at the centre of a white dwarf, initiating the synthesis of iron-peak elements in the dense centre of the WD (e.g. Kuchner et al., 1994; Howell, 2011; Hillebrandt et al., 2013). Over an average of $\sim$18 days, the pre-explosion supernova rises steeply in luminosity; while the amount of energy deposited in this stage is enormous, the centre of the WD is optically thick, and the radiation is mostly bound within the ejecta. As the density of the ejecta decreases, more energy is released outward, and the lightcurve reaches its maximum: at this point, the outward radiation matches the energy deposited by radioactive decay, and the slow decline of the lightcurve begins. The slope of the post-maximum decline in luminosity is tied to the half-life of $^{56}$Ni decay.

The standardisability of SN Ia lightcurves is the cornerstone of their use in cosmology. Applying corrections to luminosity for ‘stretch’ or slowness-to-fade (Phillips, 1993) is the next step in understanding the brightnesses of supernovae, after the initial approximation of constant brightness due to the Chandrasekhar mass. The stretch-luminosity relation, or Phillips relation, is grounded physically in the physical size of a WD at time of detonation; in simple terms, the more massive it is, the more luminous, and the more time it takes to undergo the same decline in magnitude. Recently, this has been examined in terms of ejected mass (i.e. total mass) by studying bolometric lightcurves (e.g. Scalzo et al., 2014; Piro & Nakar, 2014), in relation to observed decline rate and $^{56}$Ni mass lost. Developments in homogenising SN Ia lightcurves will be detailed in Section 2.2, from the first uses of supernovae as distance indicators, to current best methods.

### Progenitor channels and explosion mechanisms

The more uncertain factors in this picture are the identity of the other star and the mode of detonation. The WD may merge with another WD in the **double degenerate** scenario, or accrete from main sequence or red giant star in the **single degenerate** scenario, or accrete from an asymptotic giant branch (AGB) star in the **core degenerate** scenario. The paths through which the explosion can start also vary between deflagration and detonation models, or a combination; these are linked intrinsically to whether a WD has reached Chandrasekhar mass at explosion. Deflagration is thought to begin as a runaway flame near the centre of a WD as it reaches Chandrasekhar mass. Neither deflagration nor detonation by itself is sufficient to explain observations, particularly luminosities, and the correct amount of iron-peak elements produced (primarily $^{56}$Ni), i.e. amounts produced are too low and too high respectively. Proposed hybrid explanations include **delayed detonation**, where the explosion mechanism turns from deflagation to detonation at a critical density; if it fails to do so it results in a **failed deflagration**. Ignition may occur when two CO WDs combine in a **violent merger**, at or above Chandrasekhar mass (e.g. Pakmor et al., 2012). Alternatively, if the companion star has a helium shell, unstable triple-alpha reactions detonate at the surface and propagate inward to trigger a second detonation within the WD, in the **double detonation** model (Woosley & Weaver, 1994; Fink et al., 2010). In the **spin-up spin-down** scenario (Justham, 2011; Di Stefano & Kilic, 2012), where Roche lobe accretion from a donor star accelerates the spinning of the WD; the increased angular momentum allows the WD to exceed Chandrasekhar mass, and it detonates when it slows (e.g. due to magnetic braking). These competing models are explored and tested through a combination of theory and observations (binary population synthesis, modelling ejected (hence total initial) mass, nucleosynthesis codes, explosion simulation codes, spectral synthesis codes).

### Diversity and subclasses

The notion that type Ia SNe can be further subdivided was introduced in the early 1990s and supported by two anomalous Ias, SN 1991bg (Filippenko et al., 1992a; Leibundgut et al., 1993) and SN 1991T (Phillips et al., 1992; Filippenko et al., 1992b). The close occurrence of two outliers within a short period prompted questions of whether Ias were as homogeneous as they appeared, and what a typical Ia looks like; Branch et al. (1993) studied SNe Ia discovered up until then, using optical spectra to separate them into those resembling SN 1981B and a few similar SNe Ia, and the remainder, coined 'peculiar' Ias. The former category of spectroscopically normal (now often referred to as 'Branch-normal') Ias, later were parametrised using lightcurve fitters (Section 2.2.3) and used for cosmology fits. These subcollections of Ias are separated primarily by absolute peak magnitude, and decline rate: the quantity $\Delta m_{15}(B)$ (the dimming in $B$ band in the 15 days since peak). The separation of several subpopulations is shown in Figure 2.2 (Maeda & Terada, 2016, figure 1). Apart from normal Ias, the following subclasses of SNe Ia had been identified: '91bg'-like, '91T'-like, 'Iax', 'Ia-CSM', as well as overluminous SNe Ia exceeding even the brightness of 91T-like Ias. The label 'Iax' (Foley et al., 2013) was applied to fainter SNe Ia appearing like SN 2002cx (the first in the class), revealing a bluer continuum, Fe lines at higher ionization, lower line velocities than typical Ias, and not following the same Phillips relation (Phillips, 1993) between peak magnitude and decline $\Delta m_{15}$ (discussed more in Section 2.2.3). The Ia-CSMs (Silverman et al., 2013; Inserra et al., 2016) showed evidence of interaction with their circumstellar material (CSM), spectroscopically appearing like IIns (due to hydrogen from the CSM) with strong Balmer emission lines and a blue continuum, but show signatures of Ia ejecta expanding into dense CSM.

For a long time, efforts to connect the above progenitor models and explosion mechanisms to observations and simulations have been motivated by the hope of explaining a majority of observed SNe Ia with one answer. Particularly in the last decade, the number and quality of SN Ia data have shown this to simply not be possible: the range and variation over the collection of SN Ia observations have pointed the insufficiency of any single progenitor channel in explaining all observed Ias.

Significant developments have occurred in recent years: surveys have developed into more coordinated and automated operations and are thus more productive (Section 2.4), and supernovae have been discovered and followed up sooner, yielding better sampled lightcurves and spectroscopy at earlier epochs. These natural advances have been aided by the serendipitous occurrences of two very close SNe Ia, SN 2011fe (e.g. Nugent et al., 2011) and SN 2014J (e.g. Goobar et al., 2014), for which early and high-quality observations were available. In particular, the single degenerate (SD) progenitor channel has been probed and ruled out for SN 2011fe by HST deep images (Li et al., 2011b). Early spectra also showed unburnt carbon and oxygen, suggesting another CO WD (in the double degenerate scenario). A red giant was excluded as a companion of both SN 2011fe and SN 2014J, and a helium donor was also rejected for SN 2014J. Typically, SD scenarios are tested by searching historical (pre-explosion) images for a donor star as for SN 2011fe, or in our Galaxy, surviving companions long after (e.g. Kerzendorf et al., 2013); to date none of these searches have been successful. Although until recently SD progenitors have been favoured because of the comparative rarity of WD-WD binary systems, the above recent observations have made it clear that SD are insufficient as a sole progenitor channel for SNe Ia.

The diversity of SN Ia progenitors is also supported by the variation in pre-explosion WD masses: Scalzo et al. (2014) have inferred that between a quarter and a half of all normal Ias are inconsistent with Chandrasekhar mass, most of them well below the Chandrasekhar limit. This heterogeneity is also supported by Ruiz-Lapuente (2014), who summarise the variation in apparent progenitor channels observed.

These combined observational efforts have taken the challenge of identifying and explaining SN Ia progenitors into further subpopulations: no one channel or mechanism can explain all observed Ias, but perhaps the various physical mechanisms proposed can align with the separation of Ias into further subpopulations. To date, we have not detected a SD donor, i.e. a pre-explosion companion star observed where a SN Ia explosion later occured. After two neighbourhood (and spectroscopically normal) Ias this decade, the double degenerate channel in the form of violent mergers appear promising. Meanwhile, studies of SN Ia rates aim to relate different channels to observations. Ruiz-Lapuente (2014, figure 9) provide their following breakdown: around a third of observed Ias have unburnt carbon or exceed Chandrasekhar mass, from two merging WDs. Making up smaller proportions are SNe Iax, arising from probable failed deflagrations; 91bg-like Ias, most probably from violent mergers of smaller WDs; and Ia-CSMs, which are likely associated with 91T-like SNe Ia. SD explosions with main-sequence or subgiant donors seem to make up around 20% of observed Ias, while RG donors are much rarer. In summary, rather than treating all SNe Ia as one species from a singular astrophysical pathway, it is more appropriate to view them as the end products of several evolutionary paths, which all converge on the same phenomenon of an exploding CO WD with a WD or non-degenerate companion.

In the separation of SN Ia subclasses by peak absolute magnitude and decline rate $\Delta m_{15}$ in Figure 2.2, the Branch-normal and SNe Iax have each demonstrated a linear relationship between the two quantities. This Phillips relation (Phillips, 1993) naturally can be used to further reduce dispersion within SNe Ia. Later, Tripp (1998) extended this relation to correct the peak magnitude for for colour as well as for decline rate. These two quantities exist along a continuous scale, within the subsample of Ias deemed spectroscopically normal. Cosmological studies make use of these so-called Tripp relations to standardise SNe Ia further; these corrections are further explored and quantified in Section 2.2.3.

## 2.2 From lightcurves to distances

As introduced in Section 1.3.1, type Ia supernovae are immensely powerful cosmological tools as distance measures. Here we describe the nuances of using supernovae as standard candles, and their history, which was heavily intertwined with the development of supernova surveys in Section 2.4.

### 2.2.1 Early developments

The better understanding of stellar life cycles and white dwarfs (including the Chandrasekhar limit), as well as improved supernovae discovery and classification in the early 20th century, prompted the idea of placing supernovae on a Hubble diagram (introduced in Section 2.4): comparing redshifts to distance to measure expansion. The first Hubble diagram of supernova (then consisting of type I SNe, without

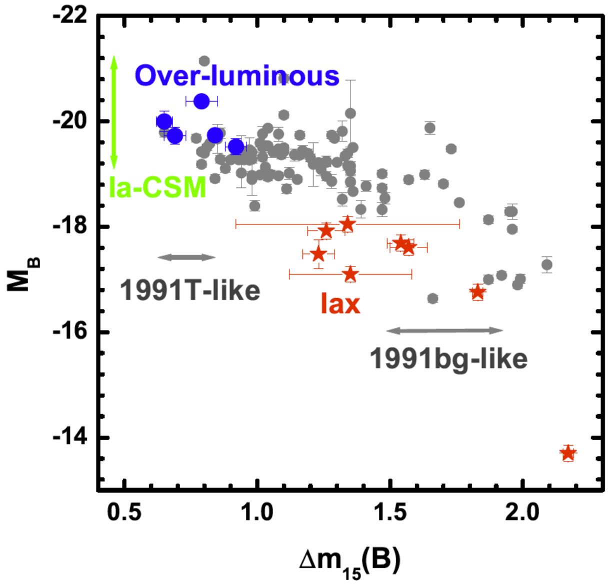


Figure 2.2 A collection of observed SNe Ia, separated by peak magnitude and decline rate (Maeda & Terada, 2016, figure 1). The grey circles are spectroscopically normal Ias obeying a Phillips relation. Blue circles are overluminous SNe Ia while red stars are SNe Iax. Grey arrows mark decline-rate ranges for 91T-like and 91bg-like SNe Ia, while the green are marks the magnitude range of Ia-CSMs.

separating out Ias from core-collapse) Kowal (1968) consisted of 22 supernovae, with a relatively large dispersion of 0.6 mag. However, this demonstrated the possibility of comparing redshifts and peak magnitudes of supernovae to measure expansion, and laid the groundwork to reduce this dispersion.

Further development occurred in the following decade, starting from the discovery of SN 1981B (Buta & Turner, 1983). This was helped by the advent of sensitive detectors (CCDs) and larger telescopes, resulting in better quality photometry and spectroscopy, as well as earlier discovery (Section 2.4). Later came the further separation of Ias from SNe Ib/c: it was clear that in addition to showing common silicon features in their spectra, Ias appeared much more similar to each other in the peak magnitudes and shapes of lightcurves (e.g. Leibundgut et al., 1991); separating out Ias from CC SNe reduced the dispersion to less than 0.25 mag (Branch & Tammann, 1992). The Phillips relation (Phillips, 1993) stipulated the dependence of the peak magnitude on the decline rate, reducing dispersion once corrected. In the 1990s observational efforts were greatly expanded, partly driven by the discovery of SN 1987A in the LMC; the start of systematic supernova surveys in search for larger homogeneous samples 2.4.1, and the definition of spectroscopically normal Ias in Branch et al. (1993).

SNe Ia would eventually be the canonical extragalactic standard candle, particularly for discovering dark energy. However, for some time type II supernovae also appeared promising as standard candles. The expanding photosphere method (Kirshner & Kwan, 1974; Schmidt et al., 1992) uses the constant temperature of SNe IIP during their plateau phase and ejecta velocity measured from spectra calculate the luminosity and hence absolute magnitudes, to measure distances.

### 2.2.2 SNe Ia as standard(isable) candles

A breakthrough for using SNe Ia to measure distance came in the Phillips relation. Phillips (1993) used Tully-Fisher distance estimates for host galaxies of nine SNe Ia to examine their peak absolute magnitudes in *BVI* and correlation with decline rate. The best correlation was found between the *B*-band peak magnitude and decline in 15 days, $\Delta m_{15}(B)$. Correcting for this 'brighter-slower' relation reduces dispersion in SNe Ia, greatly improving their utility as standard candles. This was confirmed in Hamuy et al. (1996), who also found a correlation with host galaxy morphology, where supernovae in more massive host galaxies were intrinsically brighter.

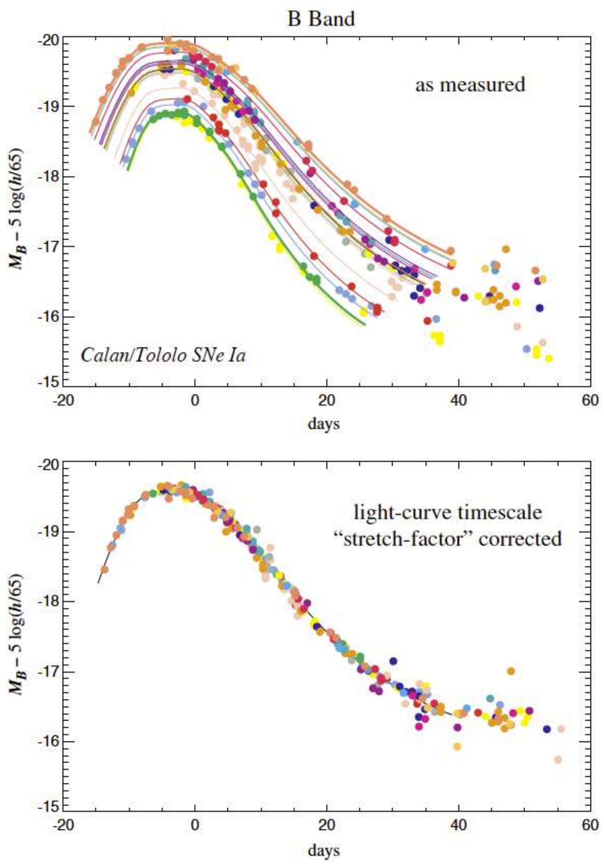


Figure 2.3 *B*-band lightcurves of CTSS supernovae as observed (top) and corrected for decline rate (bottom) (Kim, 2004), demonstrating the significant reduction in scatter than standardisation (correcting for the Phillips relation) enables.

Cosmologists using SNe Ia for distances made use of these corrections in different ways: the High-Z SN and SCP teams both used the brighter-slower relation in Phillips (1993); Hamuy et al. (1996) to reduce intrinsic scatter in their SN Ia populations, while High-Z also used MLCS (Riess et al., 1996) as an alternative method for deriving distances. Figure 2.3 (from Kim, 2004) shows this reduction in scatter with CTSS supernova lightcurves.

Prior to Riess et al. (1996), extinction was an effect that was understood to be potentially important, but not dealt with. Interstellar dust in our Galaxy and host galaxies of supernovae both dim and redden supernova light. Later, SNe Ia in the infrared were investigated as lower-dispersion standard candles than in classical optical bands, due to dust absorption and extinction having a substantially smaller effect at longer wavelengths (e.g. Mandel et al., 2017).

Another problem was the non-uniformity of spectra across all wavelengths (or beyond); in particular, the same portion of a supernova spectrum at different redshifts are necessarily observed in different passbands. The adjustment applied to correct for this is the $K$-correction: an expected ratio fluxes in different passbands, applied to observed supernova broadband magnitudes. $K$-corrections are defined in Nugent et al. (e.g. eq. 4 2002) in terms of integrals of transmissions curves with stellar SEDs over the filter passbands, and included in Hamuy et al. (1993); Kim et al. (1996) prior to the spectral template-specific calculations in Nugent et al. (2002). MLCS (Riess et al. (1996); Jha et al. (2007); more in Section 2.2.3) used in multiple colours to empirically parametrise lightcurves, with $K$-corrections added to either the template or observations, on top of reddening corrections.

Tripp (1998) further improved the standardisation of SN Ia lightcurves with the 'bluer-brighter' relation, finding good correlation between the $B$-$V$ colours and peak magnitudes of SNe Ia. Combined with the brighter-slower Phillips relation, these corrections were summarised in the Tripp metric: a peak absolute magnitude in $B$ linearly related to decline rate $\Delta M_{15}(B)$ and $B$-$V$ colour:

$$M_B = -19.48 + b(\Delta m_{15} - 1.05) + R(B - V) \tag{2.1}$$

While this two-parameter dependence was included in MLCS (Riess et al., 1996), this work made explicit how to calculate distances with respect to the decline rate and colour of SN Ia. These relations can be seen with SNLS data (Astier et al., 2006) in Figures 2.4 and 2.5, which show Hubble residuals (a proxy for magnitude, after correcting for distance and in this case the other correction term) against decline rate and colour, respectively.

### 2.2.3 Lightcurve fitting

The goal of a SN Ia lightcurve fitter is to turn raw observations of a supernova into a distance modulus by retrieving these parameters. The observations are a series of magnitudes in different filters at a number of phases; we also have the filter transmissions of the telescopes used for these observations. Methods for fitting lightcurves are based on parametrised template families; the two main classes are the MLCS and SALT2 iterations. The templates are developed empirically and compared to observations to determine their parameters (the 'training' process), then applied to observed data to retrieve observables, notably (for SALT2) the apparent $B$ magnitude at peak $m_B$, the 'stretch' or decline rate $X_1$ (approximately $\Delta m_B - 1$) and colour $C$, approximately $B$-$V$ at peak. These quantities allow the distance modulus of a SN Ia to be retrieved from its observable properties. In terms of these observables and the coefficients in the SALT2 framework, the Tripp relation then becomes:

$$\mu + M_B = m_B + \alpha X_1 - \beta C \tag{2.2}$$

The first family of fitter was the Multicolor Light Curve Shape (MLCS) method, to disentangle distance from intrinsic brightness and reddening of SNe Ia (Riess et al., 1996). MLCS uses a $\chi^2$-minimising model, including the $V$-band magnitude, unreddened colours ($B$-$V$, $V$-$R$, $V$-$I$), and reddening terms due to extinction for SNe Ia using observations in $BVRI$ bands. The colour information constrains

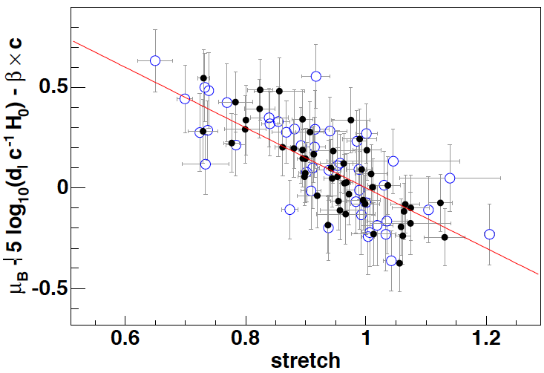


Figure 2.4 Hubble residuals as a function of stretch $s$, showing the brighter-slower relation for SNe Ia in SNLS (Astier et al., 2006, figure 9)

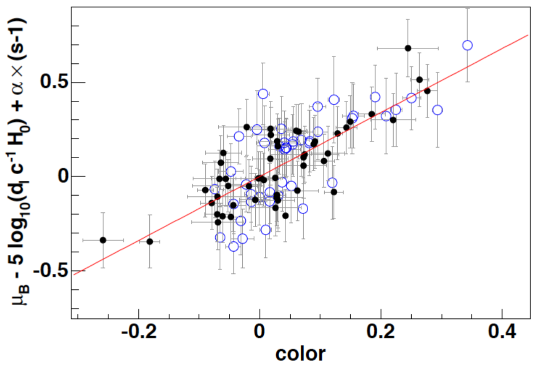


Figure 2.5 Hubble residuals as a function of colour $C$, showing the brighter-bluer relation for SNe Ia in SNLS (Astier et al., 2006, figure 10)

intrinsic colours of objects, and allows estimation of reddening and $K$-corrections. MLCS was originally trained on nine supernovae with independent distance and reddening measurements, and applied to 20 SNe Ia on a Hubble diagram. Jha et al. (2007) updated the method with added $U$-band data and updated extinction estimates to better separate reddening from intrinsic colour.

The SALT (Spectral Adaptive Lightcurve Template) lightcurve fitter (Guy et al., 2005) includes both spectroscopic and photometric training data. Their models describe luminosities of normal SNe Ia in several colours as a function of phase and redshift, in terms of some model parameters (again determined on the training set) and the three parameters for magnitude, colour, and decline rate. The SALT and MLCS families also differ in their treatment of extinction and $K$-corrections. Reddening due to galactic extinction (dependent on the ratio $R_V$ of total to selective absorption) is difficult to separate from intrinsic supernova colour. Thus, the single colour parameter $C$ encapsulates all of the colour information, including the supernova's intrinsic colour and host extinction. The SALT2 model, using empirically determined spline polynomials to model spectrophotometry of SNe Ia as a function of phase, takes into account all colour evolution, eliminating the need for $K$-corrections. A second release (SALT2; Guy et al., 2007) adds higher-redshift data from SNLS to the training set, which extends the useful redshift range of the fitter. Further retrainings came with the SNLS and JLA releases as (SALT2.2 Guy et al., 2010) and (SALT2.4; Betoule et al., 2014) respectively. Notably these retrainings have included supernovae from the SDSS-II survey (Sako et al., 2014) and high-$z$ SNe which have constrained the model better in the rest-frame ultraviolet region.

The parameters of the SALT2 model include the global coefficients $\alpha$ and $\beta$ for the stretch- and colour-dependences of the peak magnitude in Equation 6.18, which are included in cosmology fits (described later in Sections 4.3.1 and 6.2) as nuisance parameters to marginalise over. Within SALT2, the spectral sequence (evolution of the spectral energy density) of a SN Ia is parametrised by a vector $\theta$. To summarise, the flux at phase $p$ and wavelength $\lambda$ is expressed as the functional form

$$f(p,\lambda) = X_0\left(M_0(p,\lambda) + X_1 M_1(p,\lambda) + ...\right)\exp(C \times CL(\lambda)) \tag{2.3}$$

where $X_0$ is a normalisation constant, $M_0$ and $M_1$ are spectral sequences, and $CL$ is a colour law. The parameters $\theta$ of the model are coefficients of the third order $B$-spline polynomial basis for $M_0$ and $M_1$, and the third order polynomial describing $CL$ (Guy et al., 2007). The determination of $\theta$ through a training sample is described in (Guy et al., 2007, Section 4) and (Guy et al., 2010, appendix A.1). The uncertainty within the model (i.e. in the vector $\theta$) due to the finiteness of the training sample is discussed briefly in Section 3.4.1.

In our analyses in the subsequent chapters, we use SALT2 to fit for the set of parameters $\{m_B, X_1, C\}$

for each supernovae; these parameters then go into the cosmology fitting. We start with lightcurves in natural systems where available, and instruments used for observations: namely the system transmissions (including filter and instrument throughputs and atmosphere) and zero points (explained in Section 3.1.1) for each passband. Other inputs required for the lightcurve fitting are redshifts (heliocentric) and Milky Way extinction for each supernova. For some older supernovae where natural photometry is unavailable, we use the standard Landolt passbands and instruments as determined in SNLS and released with SALT2. The SALT2 model allows for custom colours as well as standard: for a set of observable parameters $\{m_B, X_1, C\}$, SALT2 can predict and simulate the flux or magnitude in any given passband, as a function of redshift and phase.

## 2.3 Supernova-specific systematic effects

As a statistical sample, type Ia supernovae are high-fidelity standard candles. However, as we have seen throughout this chapter, SNe Ia as astronomical objects are diverse and subject to systematics. Their measurable quantities (absolute brightness, observed colour and decline rate) are dependent on factors which correlate with their progenitors and environments. Numerous studies of these dependences and their astrophysical origins have been partly motivated with the view of reducing residual scatter in spectroscopically normal SNe Ia as a population. Observations of supernovae are also affected by more general observational influences such as calibration to different telescope magnitude systems, peculiar velocities, galactic extinction, and potential misclassification. In this section and Chapter 3, we discuss corrections to some of these systematic terms. Here we cover the correction terms broadly and discuss specifically the dependence of the peak magnitude $M_B$ of SNe Ia on host galaxy properties. Other supernova-specific systematic terms – potential contamination by non-Ias, and model-dependence of the lightcurve fitter – are left to Chapter 3 (Sections 3.4.7). More general observational considerations and correction terms (for peculiar velocities and selection bias) are left to to Chapter 3. These corrections, and systematic effects otherwise, have some measure of uncertainty, which must be treated with respect and accounted for carefully in our analyses; Section 3.4 is entirely devoted to our methods of doing so, using covariance matrices.

### 2.3.1 Corrections for SN Ia data

When trying to study cosmic expansion using SNe Ia, we plot their distance moduli against their redshifts on a Hubble diagram. These redshifts, along with a cosmological model, yields theoretical distance moduli (or similarly, apparent magnitudes) which can be compared with the observations. The data consist of the redshifts $z$ more or less directly measured from spectra, and distance moduli which are indirectly deduced from supernova lightcurves. The equations for fitting SNe Ia on a Hubble diagram for parameters of interest appear in Section 4.3.1 for the Hubble constant, and in Section 6.2 for dark energy; these extend on the Tripp correction in Section 2.2.2.

Before SNe Ia are analysed or placed on Hubble diagrams, necessary corrections (beyond the Tripp parameters) are modelled then applied to observed SN Ia quantities. These include corrections for the dependence on host galaxy properties, selection effects, and peculiar velocities of supernovae. These corrections are described in Sections 2.3.2, 3.1.2, and 3.1.3. The processes for estimating the uncertainties associated with these corrections are discussed in Section 3.4, and applied in Chapters 4 and 6.

### 2.3.2 Host galaxy dependence

The dependence of the intrinsic SN Ia brightness on properties of their host galaxies is well established, with numerous studies (e.g. Sullivan et al., 2010; Lampeitl et al., 2010; Kelly et al., 2010; Childress et al., 2013; Rigault et al., 2013) finding that more massive galaxies (correlated with higher metallicity and lower specific star formation rates) host SNe which are on average $\sim$0.08 mag brighter after the standard corrections for colour and lightcurve width have been applied. To mitigate the systematic error that this effect introduces to the cosmological analysis, we follow Sullivan et al. (2011) and subsequent analyses (Conley et al., 2011; Betoule et al., 2014) in using the host mass-step approach to account for this dependence. We adopt two discrete values for the SN Ia absolute magnitude, using the variable

$$M_B^* := \begin{cases} M_B, & \text{host galaxy mass} < 10^{10} M_\odot \\ M_B + \Delta M_B, & \text{host galaxy mass} > 10^{10} M_\odot. \end{cases} \tag{2.4}$$

The offset $\Delta M_B$ can either be included as a fit parameter (in Chapter 6) or fixed at an externally determined value (in Chapter 4). The our choice to fix $\Delta M_B$ in Chapter 4 will be expanded upon in Section 4.5.2, but is highly motivated by the tight agreement of values in the literature: Sullivan et al. (2010) report that SNe Ia in more massive and lower specific star formation rate (sSFR) galaxies are brighter by $0.06 - 0.09$ mag, using $\sim$400 supernovae in SNLS, combined with some low-$z$ and HST SNe (the later C11 compilation); C11 adopt the value of $\Delta M_B = -0.08 \pm 0.02$ derived from Sullivan et al. (2010, table 5) based on five choices for the host mass boundary. Similarly, Lampeitl et al. (2010) report $\Delta M_B = -0.10 \pm 0.04$ from $\sim$300 SDSS-II supernovae; values from Kelly et al. (2010, table 5) from a combined low-$z$ sample are in good agreement. More recently, Childress et al. (2013) find $\Delta M_B = -0.077 \pm 0.014$ based on a Nearby Supernova Factory sample.

Estimates of the host galaxy mass are typically derived using the `ZPEG` photometric redshift code (Le Borgne & Rocca-Volmerange, 2002, 2010) based on spectral energy densities from the PÉGASE.2 spectral synthesis code. Where possible, we obtain host galaxy masses from the literature – predominantly JLA, but also e.g. Neill et al. (2009); Kelly et al. (2010) for the supernovae in Chapter 4 – and check for consistency between multiple sources and our estimates. Some supernovae in Chapter 4 are missing host galaxy masses; we describe our procedure and motivation for handling these in Section 4.5.2. The propagation of uncertainties in this correction through to SN parameters is later described in Section 3.4.3.

## 2.4 Supernova surveys

This section explores the development of surveys of supernovae, predominantly type Ia. We trace progress in amassing supernova samples over the past few decades in roughly chronological order, starting from the beginnings of supernovae discovery, through the development of search and detection techniques, through to modern high-yield automated searches. We also discuss lessons learned in the past decade and their implications for current and future surveys.

### 2.4.1 First generation: from Calán/Tololo to CfA

The history of systematically searching for and measuring supernovae is closely intertwined with the notion that they can be used to measure distances; supernovae are also of intrinsic interest, given their relevance to several different areas of astrophysics (see beginning of this chapter). The history of using supernovae as standard candles is expounded further in Section 2.2.

Scientific study of supernovae developed in the early-to-mid 20th century, with theories of stellar evolution suggesting that stars could end their lives in explosions. Significant developments included constraints on theoretical maximum mass for a stable WD (Chandrasekhar, 1931), a spectral classification scheme (Minkowski, 1941), identification of the Crab Nebula as a supernova remnant (Baade, 1938), observational efforts to search for and collate supernovae in the next few decades (e.g. Zwicky, 1964; Kowal et al., 1971), the first attempts to place supernovae on a Hubble diagram (Kowal, 1968), the discovery of SN 1981B (Buta & Turner, 1983), and the first discovery of SN Ia at $z = 0.3$ (Norgaard-Nielsen et al., 1989).

Commencing in 1990, the Calán/Tololo Supernova Search (CTSS; Hamuy et al., 1993) paved the way in systematically searching for larger collections of supernovae, implementing the basic elements of modern supernova searches. CTSS aimed to produce an entirely CCD-observed distant (for the time) SN Ia sample for cosmological study, and to determine the dispersion in the peak magnitude. To search for supernovae, they imaged ∼25 fields covering ∼600 square degrees with photographic plates twice a month, physically blinking them to detect differences. CCD photometry, using larger telescopes at Cerro Tololo Inter-American Observatory (CTIO), were used for spectroscopic and photometric follow-up (for supernova classification, and obtaining lightcurves, respectively). Scheduling follow-up was understandably difficult, and heavily reliant on telescope time contributed by other (visiting) astronomers. Nevertheless, over 3.5 years the CTSS discovered 30 new SNe Ia up to $z = 0.1$. Using this sample, they were able to confirm the Phillips relation (Hamuy et al., 1996). Using this and with a cut on peak $B$-$V$ colour, CTSS further reduced dispersion in the Hubble diagram, and also established a relationship between host galaxy morphology and decline rate, later extended in Sullivan et al. (2006).

The legacy of CTSS as a first modern supernova survey remains, eventually forming a bulk of the SNe used to determine the Hubble constant in the HST Key project (Freedman et al., 2001) (discussed in Chapter 4). In the shorter term, Calán/Tololo had set up the higher-redshift searches that led to the discovery of the accelerating Universe – the High-Z Supernova Survey (Riess et al., 1998; Schmidt et al., 1998) and the Supernova Cosmology Project (SCP; Perlmutter et al., 1999) – serving as a low-redshift anchor for both. The search and follow-up methods pioneered in CTSS were developed further in both teams, and the lightcurve parametrisation lessons gleaned were vital to the success of both surveys. SCP had corrected for stretch using the Phillips relation, while the High-Z team used a new lightcurve fitting method, the Multicolor Light Curve Shapes (MLCS; Riess et al., 1996) algorithm, which takes into account colour excess due to dust.

Soon after, the Harvard-Smithsonian Centre for Astrophysics (CfA) supernova program (Riess et al., 1999; Jha et al., 2006) began, with the most recent data release (CfA4) in Hicken et al. (2012). Despite not actively searching, the unified follow-up program over two decades formed a core part of the collection of SN Ia photometry, particularly at low redshift. Early CfA supernova data were also used to develop MLCS further (MLCS2k2; Jha et al., 2007).

### 2.4.2 Searching for supernovae

The early search techniques for high-redshift supernovae in the High-Z SN Search and SCP were further extended, and made more automatic in subsequent years. Supernova searches are closely related to difference imaging, which is discussed in more detail below. The photographic search methods in CTSS, described in Section 2.4.1, evolved into digital searches, starting with the Berkleley Automated Supernova Search (Perlmutter et al., 1992), which was developed into the detection pipeline for SCP. An independent pipeline was developed for the High-Z SN Search (Schmidt et al., 1998), which formed the basis for the Beijing Astronomical Observatory Supernova Survey (BAOSS; Qiu et al., 1999; Li et al., 1999), and was later transferred to Lick Observatory Supernova Search (LOSS; Li et al., 2000; Filippenko et al., 2001).

The image subtraction methods (discusssed more below) were further developed in Alard & Lupton (1998) and later Becker (2015), and were crucial to all subsequent supernova searches. The computer-automated search of BAOSS in the late 1990s was the most productive supernova search to date, with 16 supernovae discovered in little over a year, most of which were discovered before maximum. LOSS further developed this software, originally adapted from the High-Z SN Search team, to image ∼5000 galaxies with short cadence (visiting each galaxy every 3–5 days in good weather), scheduled automatically using the Katzman Automated Imaging Telescope (KAIT). Supernova candidates are identified automatically with the aid of human scanners, and followed up in several filters; the most promising candidates have automatic additional photometry taken the same night. The short cadence means that supernovae were typically discovered pre-maximum; lightcurves were well-sampled including at early times, and in multiple colours.

Early efforts had relied on amateur astronomers for the discovery of supernovae, and used telescope time to obtain lightcurves. The techniques in the above wroks remain applicable now for supernova and transient discovery, at a much larger scale: high cadence searches, difference imaging, fast distribution of candidate information to encourage wider follow-up. Despite automated telescope operation and image reduction, the search process as a whole remained labour-intensive; in modern surveys many of these steps are automated. While CfA had observed a large number of SN Ia lightcurves, successfully monitoring newly discovered Ias (including by amateurs with no means for follow-up), LOSS was unique at the time in its effectiveness in discovering new supernovae, and was the most successful search to do so. The ability to search for SNe, and the improved/developed techniques for doing so, were pivotal for the next generation of supernova surveys, notably the first rolling supernova search in SNLS.

**Difference imaging and host subtraction**

The detection of supernovae and other optical transients such as asteroids relies on being able to observe an object not previously detected, typically through **difference imaging**. A new image is compared with a reference or template taken some time ago and subtracted. Only if the difference in the two images appears like a supernova then a range of spectroscopic and photometric follow-up resources are deployed. The primary photometric characteristics which signal that a candidate is likely a supernova are: appearing as a point source visually close to but not at the centre of a galaxy, and relatively blue colour. A point source at the centre is more likely to be an active galactic nucleus (AGN), and one with no visible identifiable host galaxy is more likely to be an asteroid, although SNe appearing to be hostless do seem to occur. Typically, however, asteroids tend to disappear after a detection, due to motion, and are generally much more red than supernovae. If multiple lightcurve points are recorded, a rise in brightness over a timescale of days (followed by a slower decline) is a fairly good identifier for a supernova.

Methods of comparing images (template subtraction) have improved substantially since the beginning, and become much more automated. In CTSS photographic plates had to be blinked manually with differences observed by eye; these made supernovae near centres of galaxies particularly difficult to identify. Similarly in BAOSS, images were digital but still compared by eye to rule out defects or asteroids. Not only was the operation of the telescope – from scheduling to carrying out observations, to protecting the telescope during inclement weather – fully automatic, but image subtraction was to a high degree automated also.

There are several basic components of difference imaging (subtracting a new image from a template), developed in the early surveys discussed earlier (CTSS, SCP, High-Z) and automated in BAOSS and LOSS, and later underpin the SNLS, ESSENCE, PTF, and most other supernova search programs in the world. A small survey telescope images the same areas of sky with high cadence (ideally once every few days, to maximise chances of early discovery). New images of the search area are compared to reference or template images to check for differences, by image subtraction (Alard & Lupton, 1998; Becker, 2015), and the difference is searched (through a combination of algorithmic and human classification) for candidates. Promising candidates are followed up by using larger telescopes to obtain spectroscopy, for classification, and additional photometry in multiple colours (possible using larger telescopes and/or longer exposure times) for quality multicolour lightcurves.

The difference imaging process (e.g. Perlmutter et al., 1992; Schmidt et al., 1998; Alard & Lupton, 1998; Filippenko et al., 2001; Becker, 2015; Scalzo et al., 2017) is is the most important algorithm, on which the success of the pipeline rests. There are several steps to match the reference and template images: matching the astrometry, the point spread function (PSF), and the intensity or zero points. The specifics of these practices are described fully in Filippenko et al. (e.g. 2001); Kessler et al. (e.g. 2015); Scalzo et al. (e.g. 2017). First, the coordinate systems are transformed by mapping positions of known stars. Then the point spread function of the reference image (typically narrower than the new image) is convolved to the new image (for SkyMapper, if the new image has smaller PSF and does not show a candidate, it replaces the ref, resulting in typically higher-quality or smaller-PSF reference images). Finally intensities of known stars are matched, so that an arithmetic subtraction of the new image from the reference should only return objects where there is a real change in brightness or a new object appears.

In reality, many objects appearing in the subtracted images are not potential supernovae. These can include cosmic rays (with very narrow PSFs); algorithmic artefacts such as bad subtractions, which appear obviously artificial to a human eye; and real astrophysical objects such as asteroids, AGN, and variable stars, in addition to supernova candidates. Only fields with potential supernova candidates are selected for follow-up; this vetting process requires many hours of human input and is labour-intensive. The application of machine learning as a classification aid is discussed in Section 3.3.

In SNLS and DES, a newer model-based approach to measuring supernova light separately from any host galaxy contribution has been explored and implemented. This method, called **scene modelling photometry** (discussed later in Section 6.4.2), is used for making precise photometric measurements, but not for searching images for supernovae.

### 2.4.3 Building higher redshift samples: SNLS, JLA

The High-Z SN Search and SCP had both demonstrated the immense value of high-$z$ SNe in the discovery of cosmic acceleration. The next step was obtaining more data (i.e. improving statistics) for better measurements of cosmological parameters.

This was achieved in the Supernova Legacy Survey (SNLS) in the early-to-mid 2000s, which produced an unprecedentedly large sample of high-redshift SNe Ia, all with spectroscopic classifications. SNLS remains the benchmark for a supernova survey at high redshift: in its first three years, SNLS found 242 SNe Ia at $z > 0.2$, totalling a joint sample of 472 spectroscopically confirmed SNe Ia when combined with low-redshift, SDSS-I, and HST supernovae (Guy et al., 2010; Sullivan et al., 2011; Conley et al., 2011, hereafter C11). The Joint Lightcurve Analysis (JLA) of SNLS with SDSS-II supernovae (Betoule et al., 2014, hereafter B14) increased the total number to 740, with the addition of the second SDSS Supernova sample.

As a data set, SNLS was unique and remarkably successful. It remained the most productive high-redshift supernova survey for over a decade until DES began. The volume of effort required to obtain spectra of every live supernova is touched upon later in Section 2.4.5. SNLS also pioneered **rolling search** techniques, where new images are turned into reference images for subsequent observations; this is now standard in transient searches.

Moreover, SNLS was the first SN Ia analysis/sample to estimate uncertainties by using covariance matrices to account for correlated systematics between different supernovae and quantify the impact on final uncertainties (Conley et al., 2011), an approach we have adopted throughout the analyses in this thesis. This method also assessed the relative importance of errors in terms of their contribution to the final errors in cosmological parameters, revealing that photometric calibration is far by the largest contribution to systematic uncertainty, much of this in the primary standard BD+17° 3248. The calibration chain, built upon this standard star, is discussed further in Section 3.1.1.

Thus, just as important as the addition of high-quality supernova data were the lessons learned from SNLS, which have since steered and shaped following surveys. In particular, SNLS highlighted the factors that must be considered for the next generation, which would be limited by systematics particularly in photometric calibration, and in the lower redshift sample, and for which spectroscopic classification of all transients would also be unfeasible. JLA was born from these lessons, and set out to revisit calibration chains of both surveys (Betoule et al., 2013), supplemented by new observations of standards, to reduce the calibration uncertainty. The methods demonstrated in SNLS and subsequently JLA have developed and provided methods of lightcurve fitting (SALT2), and computation and analysis of correlated SN Ia systematic errors with covariance matrices (not done before to the same level of rigour and detail), which in themselves have great legacy value. We continue using these methods in the work in this thesis (see also Scolnic et al. (2015)), and discuss some newer alternatives to dealing with SN Ia systematics in Section 3.2.3.

### 2.4.4 SkyMapper and other modern low-redshift surveys

In a supernova survey, a large range of redshifts (acting as a long lever arm on a Hubble diagram) increases the precision that can be achieved from a data set. Furthermore, within this range, increased numbers improve the statistical accuracy further. The high-redshift end of the sample is important because it tracks the Universe's size evolution further in the past, better constraining acceleration. Meanwhile, a robust and precisely calibrated low-redshift sample reduces statistical error in the whole sample, and is the foundation set of objects that the distant set is compared. Since SNLS and JLA, numerous surveys have concentrated on improved low-redshift samples including the Carnegie Supernova Program (CSP; Contreras et al., 2010; Stritzinger et al., 2011), which are necessary and important for reducing the zero-point error carried in the low-redshift anchor.

Techniques in amassing large quantities of high-redshift SNe Ia to place on Hubble diagrams have greatly advanced, with SNLS and subsequently Pan-STARRS and DES discovering and measuring supernovae frequently with high-quality photometry. However, these same surveys have revealed that a great contributor to errors in cosmological parameters is the inhomogeneity of supernova samples, particularly at low redshift. Low-redshift supernovae on a Hubble diagram are important because a large redshift span is needed to constrain deceleration (measured over a longer baseline), and as an anchor. The scale and shape of the Hubble diagram, showing the Universe's expansion history, are set at low-redshift; scatter within the low-redshift sample will translate to uncertainties in cosmological parameters. There is great value to improving the current low-redshift anchor which comprises supernovae from many sources and surveys, including CfA1–3, CSP, and Cálan/Tololo. Compared to the higher-redshift SNe, the uncertainties in the cross-calibration terms of these surveys are large. In addition, the selection functions are largely unknown, leading to potential selection biases which are difficult to estimate (discussed Section 3.4.2). The targeted nature of searches (Section 2.4.2) likely leads to supernovae from more massive hosts (discussed Section 3.4.3); moreover the resultant clustered distribution of supernovae can magnify uncertainties due to peculiar velocities (Section 3.4.4). Thus, errors in cosmological parameters will be reduced with a uniform, homogeneous, and untargeted low-redshift supernova sample.

The SkyMapper Transient Survey (SMT; Scalzo et al., 2017) had originally sought to address these shortcomings by performing an untargeted wide-field search of the southern sky, aiming to yield the largest yet SN Ia sample from a single telescope. The overlap of the SkyMapper search area with the DES footprint provided the possibility for direct cross-calibration with DECam. The 1.3-m wide-field SkyMapper telescope at Siding Spring Observatory (SSO) was built for the Southern Sky Survey (Keller et al., 2007), while searching for supernovae and other transients in bad-seeing time. The unique filter set – including Sloan-like $u$, and narrow custom $v$ band (Bessell et al., 2011) – enabled higher sensitivity to stellar parameters (surface gravity and metallicity). The rolling search largely took place in $gr$ bands with a 3–5 day cadence, with follow-up in $vgri$; a footprint of $\sim$2000 square degrees over the Southern sky was searched uniformly every week. Template 'reference' images were subtracted from new images (see Section 2.4.2) once coordinates, zero points, and PSFs were adjusted to match. The subtracted images were classified by eye, with the classifications used to further train the machine learning classifier on several occasions, as detailed in Scalzo et al. (2017). Supernova candidates were shared with PESSTO or followed-up on the ANU 2.3-m telescope at SSO spectroscopically; later automatic reporting to the SMT webpage and the IAU Transient Name Server website[2] were implemented.

I contributed to the day-to-day running of SMT over a period from 2014–2017, including modelling weather conditions at Siding Spring Observatory as an input to simulations of detection efficiency, submitting proposals for telescope time (the good-seeing portion) from 2016, performing simulations to optimise the cadence for better constraints, and occasional running/debugging of the pipeline. SkyMapper also hosted several successful citizen science projects, including the Zooniverse, and featuring in Stargazing Live and Supernova Sighting.

The Foundation Supernova Survey (Foley et al., 2018) on the Pan-STARRS1 telescope is moti-

[2] https://wis-tns.weizmann.ac.il/

vated by similar factors, and harnesses the existing infrastructure established by the Pan-STARRS Supernova Survey (PS1; Rest et al., 2014), in particular allowing uniform treatment with the exist PS1 supernovae. Given forecasts from Foundation, SMT has changed strategy to focus on supernova science rather than search for supernovae to build a low-redshift cosmology sample. Over the time the survey ran, from Science Verification in 2013 to late 2017, SkyMapper had discovered over 70 transients including over 40 classified SNe Ia and ∼20 other supernovae, including core-collapse SNe, a superluminous supernova (SN 2013hx), and several peculiar objects.

Other contemporaneous surveys of note were the ESSENCE Supernova Survey (Wood-Vasey et al., 2007), Carnegie Supernova Program (Contreras et al., 2010; Stritzinger et al., 2011) at low-redshift on the Swope Telescope; the SDSS SN survey (Kessler et al., 2009b; Sako et al., 2014) at intermediate redshifts (up to $z \approx 0.5$) and included in JLA; and the Pan-STARRS Supernova Survey (PS1; Rest et al., 2014) at similar redshifts to SDSS. Later, the PS1 and JLA samples were combined into Pantheon (Scolnic et al., 2018), using the revised low-redshift calibration (Supercal, discussed in Section 6.5.2) in Scolnic et al. (2015).

### 2.4.5 Future directions: DES and LSST

The Dark Energy Survey was described in Chapter 1 as a coordinated survey utilising four cosmological probes, including SNe Ia, to measure dark energy. The SN survey in DES is described in full detail in Chapter 5. In summary, over six years the Dark Energy Camera (DECam) images a 5000 square degree area of the sky, with around ∼2500 supernovae detected from ten smaller science fields totalling ∼30 square degrees. An analysis of an intermediate sample, the 3-year spectroscopically classified sample, forms Chapter 6. The 8-m wide-field Large Synoptic Survey Telescope (LSST) is the first very large telescope dedicated to time-domain astronomy, expected to see first light in 2019. With the ability to image the entire sky every several nights to new depths, LSST exceeds the boundaries of what is currently capable in terms of discovering and studying transient objects, even in DES. LSST represents the start of a new era in transient science, and DES paves the way for studying supernovae with LSST.

Many of the methods which will underpin LSST are developed within DES: automated pipelines, machine learning in scanning candidates and photometric classification, simulation-based methods to understand bias. The use of these techniques (described in Chapter 3) in DES characterise the shift in supernova analysis since JLA, beyond the increase in size. While the 3-year sample is analysed with existing ('JLA-like') and new methods alike, and relies on spectroscopic classification, the 5-year analysis is expected to use new techniques only (validated and compared using the 3-year sample) and include photometrically classified supernovae. These techniques, particularly photometric classification (Section 3.3.3) will continue to be relevant and useful in LSST.

My involvement in DES has been as a part of the OzDES collaboration, and the DES Supernova Working Group. My contributions have been to OzDES operations and infrastructure, as an OzDES builder, and leading the JLA-like analysis of the DES 3-year spectroscopic SN sample (Chapter 6).

#### OzDES

Obtaining spectra of supernovae, of sufficient signal-to-noise to classify them, is logistically complex and difficult for fairly self-evident reasons: they cannot be predicted (in time and space) so short of having ready access to telescope time follow-up cannot be prepared and must be arranged at short notice; they are very faint at high redshift, necessitating time on large (4-m and larger telescopes); they decline quickly, so there is only a short window in which follow-up is possible. This complexity and difficulty equally highlights the impressive scale and massive effort of SNLS (groundbreaking in its ∼500 SNe Ia, all spectroscopically confirmed), and means it is impossible to scale up for surveys the size of DES. Alternatives to spectroscopic redshifts and classification are clearly necessary. Photometric classification, verified by tests from SNLS, underpin the next generation of SN surveys, and is discussed in Section 3.3.3. Nevertheless, redshifts are still necessary for a Hubble diagram; current techniques in photometric redshifts are not comparable in accuracy to spectroscopic redshifts. For DES, this problem is circumvented by obtaining host galaxy redshifts instead, avoiding both limitations in time and brightness. By targeting galaxies in which SNe occur, their redshifts can be determined long after a supernova has faded.

The Australian Dark Energy Survey (OzDES; Yuan et al., 2015; Childress et al., 2017) is the spectroscopic counterpart to DES, taking place on the 4-metre Anglo-Australian Telescope (AAT) at Siding Spring Observatory over the same observing seasons as DES with an added short (12-night) sixth year (2013B–2018B). OzDES has benefited from the well-matched designs and specifications of the AAT and Blanco telescope as well as Australian expertise and instrumentation capabilities in spectroscopy.

OzDES plays a crucial role in obtaining spectra of DES objects, particularly obtaining redshifts of supernova host galaxies. This role is further expanded in Section 6.3.2. In addition to supernovae and their hosts, OzDES targets other DES objects including active galactic nuclei (AGN) for a reverberation mapping campaign, many luminous red galaxies (LRGs), galaxies in clusters, supporting a wider cross-section of DES science including the validation of photometric redshift methods. The first OzDES data release (Childress et al., 2017) contains redshifts from the first three years of OzDES, approximately half the lifetime of the survey. Over this period, $\sim$15000 redshifts had been obtained, including almost 100 supernovae and $\sim$2500 supernova hosts. Around 5700 supernova host galaxy redshifts are forecasted for the end of OzDES to support the later DES photometric supernova sample.

Chapter 3

# Methods

This chapter contains a summary of the methods used in this thesis, including general considerations in observational astronomy, Bayesian methods (particularly for estimating cosmological parameters), and the covariance matrices that form the bulk of our error estimation arsenal.

## 3.1 Observational considerations

### 3.1.1 Photometric calibration

To best scientifically study luminous objects, their brightness must be placed on a quantitative scale. This practice, and the resultant numbers, are referred to as **photometry**. The factors that affect a measured physical flux are numerous and complex, including many instrumental and atmospheric factors; the latter will necessarily change over the duration of observations. It is hence necessary to calibrate observations and the systems used to measure them, through intricate methods of photometric calibration. In particular, the importance of precise and accurate photometric calibration for supernova cosmology in particular will be stressed and explained later in this section.

Historically, astronomical magnitude was defined so that stars visible with the naked eye ranged from zeroth magnitude (brightest) to fifth magnitude; this has been formalised as negative logarithmic function of brightness, related to spectral flux density $f_\nu$,[1] defined at a single frequency $\nu$:

$$m = -2.5 \log_{10} f_\nu + ZP \tag{3.1}$$

where $ZP$ is a magnitude zero point whose definition depends on the magnitude system used. A magnitude system is a set of filter transmission curves (passbands) and zero points. Two broad classes of magnitude systems are the AB system (defined in Oke, 1965), on an absolute flux scale, and Vega-based magnitude systems, normalised to Vega ($\alpha$ Lyrae). In the AB system, the magnitude scale is normalised such that a theoretical object with flux of 3631 Janskys has magnitude zero. In practice, magnitudes are measured in broad-band systems, where a set of filters, each letting a range of frequencies through with some transmission efficiency [2] $T(\nu)$ is used to define the magnitude of an object in multiple colours. The magnitude is integrated as

$$m_{\mathrm{AB,b}} = -2.5 \log_{10} \left( \frac{\int_b T(\nu) f_\nu \; d\nu}{\int_b T(\nu) \; d\nu} \right) + 8.90, \tag{3.2}$$

over a passband $b$, where $f_\nu$ is in Janskys (with 1 Jy= $10^{-26}$erg$-$1s$^{-2}$Hz$^{-1}$). Common standard magnitude systems are the Johnson-Cousins '*UBVRI*' system, and the filters used in the Sloan Digital Sky Survey *ugriz* (Abazajian et al., 2003, 2004) – these systems and more are described in great detail in the Bessell (2005) review. The Johnson-Cousins system was first established in Johnson (1966) as a

[1] The **spectral irradiance** $f_\nu$ is defined as the energy radiated per unit surface area, per unit frequency. Spherical symmetry of a source is assumed, so this decreases with distance from the object squared, much like intensity.

[2] In Equation 3.2 it is assumed that $T(\nu)$ is defined per unit energy rather than per photon. Alternatively the transmission can be defined as photon-counting, requiring the transformation $T(\nu) \mapsto \frac{T(\nu)}{\nu}$.

series of broadband filters *UBVRIJHKLMN* from the ultraviolet to infrared ends of the visible spectrum, extending (alphabetically) into the infrared. Fluxes and colours in this system are normalised to Vega, and have been defined in terms of catalogues of stars distributed over the sky (e.g. Cousins, 1976; Landolt, 1983; Bessell, 1990; Landolt, 1992). Later, synthetic photometry has been used as an alternate and comparison method for calibration, by integrating model ('synthetic') spectra with transmission functions over instrumental passbands to determine theoretical magnitudes. In Vega-based magnitude systems, broadband magnitudes are close to their AB equivalents in Equation 3.2, differing by quantities known as **AB offsets**.

In recent years, the standard approach has trended toward publishing photometry in telescopes' natural systems (rather than the standard magnitude systems such as *UBVRI* as defined in Landolt (1992)), alongside properties of the magnitude system that would have been used for calibration to a standard system. These include transmission curves and zero points for each passband, and sometimes the standard star photometry used to determine the zero points. Doing so makes the calibration process more transparent, and reduces potential cross-calibration errors, particularly when studying composite data sets from multiple telescopes.

Photometry can be considered in absolute or relative terms. Absolute calibration is tying a set of observations, on a given night in a given magnitude system, to an absolute scale (AB or Vega). The magnitude and flux are related to absolute physical units, requiring some sources with models for physical energy or flux emitted by them. The HST CALSPEC system is based on three DA white dwarfs, which have atmospheres that could be modelled to determine their luminosities theoretically. Alternatives have been 'direct illumination' calibration experiments (DICE), using LED illuminators as known sources for calibration (Barrelet & Juramy, 2008; Regnault et al., 2016; Barrelet, 2016).

Relative calibration is the process of ensuring internal consistency within a set of observations taken by some instrument in some magnitude system, *before* tying this magnitude system to an absolute scale. Here, the absolute flux scale is not considered, because only ratios of fluxes (differences or offsets between magnitudes) are considered.

Calibration methods of wide field surveys (where observations were necessarily taken over many epochs, and over areas of the sky which do not overlap in a single exposure) have been based on the HST CALSPEC network through a calibration chain of non-variable stars. In this method, a network of 'secondary' standard stars are calibrated (i.e. have magnitudes in a given system determined) by the three DA white dwarfs, the 'primary standards'. Then for each science field, a set of 'tertiary' standard stars have their calibration tied to the secondary standards. The absolute magnitudes of these tertiary standards are fitted to determine zero points and colour transformations of the passbands in the natural magnitude system. These tertiary standards are observed on each night of scientific observations, to transfer the calibration of the CALSPEC network to the survey observations, and to determine the zero point offsets for that night. Examples of this framework are in SDSS, SNLS, Pan-STARRS (Padmanabhan et al., 2008; Betoule et al., 2013; Scolnic et al., 2015). SkyMapper currently utilises a similar calibration chain. The calibration of DES is via the Forward Global Calibration Method (FGCM), described in Section 6.4.1. The FGCM method defines the natural standard system of DECam; AB offsets are then used to transfer the CALSPEC zero points to the DECam system.

**Calibration in supernova surveys**

For supernovae, photometric calibration is especially relevant and important for several reasons: the magnitude is directly used to infer distance (i.e. they are standard candles), and it evolves rapidly over a timescale of days, on nights with different atmospheric conditions. Errors in relative calibration between nights will distort the measured lightcurve, and consequently the derived distances.

Photometry is made more complicated by the need to isolate light from a supernova, from light originating from its host galaxy. Methods involve host template subtraction, difference imaging, or in DES, Scene Modelling Photometry (Section 6.4.2); these are outlined in Section 2.4.2.

Moreover, magnitude systems are often calibrated using non-variable 'standard' stars, with substantially different spectral energy distributions (SEDs) to supernovae; extrapolating calibrations to SNe inherently involves a degree of uncertainty (this is addressed by model-based methods of calibration such as the Forward Global Calibration Method (FGCM) in DES, which is described in detail in Section 6.4.1 of Chapter 6). As found in SNLS and JLA, the challenge of photometric calibration is that it is both highly complex and subtle, and at present especially important because of its large contribution to the error breakdown in cosmological surveys; these findings have greatly influenced subsequent developments in cosmology: the cross-calibration in JLA between SNLS and SDSS, the careful calibration (Supercal) in Pan-STARRs and later, and FGCM as part of DES calibration efforts. FGCM and its associated uncertainties are discussed further in Section 6.5.2.

Ensuring accuracy and consistency in photometric calibration of a supernova survey is important for numerous reasons. When measuring dark energy, we are concerned with the shape of the Hubble diagram (the relative rate of change of expansion). While best efforts are made to perform accurate photometric calibration, offsets in the absolute scale will impact the peak SN Ia magnitude $M_B$ (a nuisance parameter for cosmology) but not the crucial cosmological parameters $w$ and $\Omega_m$ which we seek to measure. It is critical then the calibration is consistent between all supernovae within the sample, particularly between high-redshift and low-redshift supernovae. On the other hand, when measuring the Hubble constant, we require the present expansion rate in actual physical units. Thus, supernovae are anchored to an absolute scale via a distance ladder (Section 4.3), whereby SN Ia absolute magnitudes are calibrated by Cepheid variables, which are then calibrated by geometric methods. The work in Chapter 4 relies on this principle. In this context it is critical that the calibration between different rungs of the distance ladder is consistent and accurate. As Section 4.3.3 will demonstrate, the absolute SN Ia peak magnitude $M_B$ is degenerate with the Hubble constant $H_0$, so the absolute calibration is crucial. For finding $H_0$ the 'location' of the Hubble diagram on the vertical axis is most important, whereas for dark energy, the shape is most important.

### 3.1.2 Malmquist and selection biases

In any astronomical survey, the targets included in the sample are subject to factors which determine which objects are observed, detected, or selected for study. The studied sample may not be entirely representative of the whole population, especially in magnitude-limited surveys where the primary criterion for detection is exceeding a brightness threshold. These selection effects can lead to potential biases in results if using the sample to directly determine intrinsic properties of objects, or to indirectly determine cosmological parameters. The most notable selection effect to consider is Malmquist bias, or the preferential selection of intrinsically brighter objects.

At higher redshifts, any disparity in detection efficiency can become more significant, and can be complicated by or entangled with factors such as population evolution with redshift. In supernova surveys, where the objects studied are used to measure the scale of the Universe at different redshifts, any brightness- or redshift-dependent biases can have a significant impact on derived parameters. It is thus imperative to estimate and correct for such biases in supernova samples. Where limited by selection effects, studying only detected supernovae (without corrections) will lead to biases in observables, which imply SNe Ia are intrinsically brighter, bluer, and slower-declining. Over the past several years various approaches have been employed to account for selection effects.

#### Selection efficiency

Selection functions of high redshift surveys have two components: the detection itself due to being magnitude-limited (a strict cutoff can be well-simulated with Monte Carlo if the underlying distributions are well understood), and the spectroscopic follow-up, which is influenced by many more human and logistical factors. JLA used the selection functions modelled in Dilday et al. (2008) for SDSS, and

Perrett et al. (2010) for SNLS – these are functions of both peak magnitude, and colour in the case of SDSS (which favours bluer events). These are combined in data/MC simulations, described in Mosher et al. (2014, section 6.2), Betoule et al. (2013, section 5.3).

Methods to compute and correct for biases in $\mu$ as a function of redshift, e.g. those in Mosher et al. (2014), have limitations. For instance, the values of $\alpha$ and $\beta$ are determined without bias corrections to begin with, and underlying colour and stretch distributions assumed in these simulations are approximate. Scolnic & Kessler (2016) improve on these estimations by rigorously and iteratively determining these underlying colour and stretch distributions. They assume asymmetric underlying populations, and measure the migration of input parameters to observed distributions in surveys, given intrinsic scatter (various models), measurement noise, and detailed survey simulations. They also demonstrate that correcting for magnitude, stretch, and colour together as one parameter (distance modulus) is insufficient, and these individual biases should be treated independently. Improved understanding of underlying populations can be used for modelling more accurate corrections for biases from selection effects, through simulation.

#### Low-redshift selection functions

At low redshift, estimates of Malmquist bias are more uncertain. Unlike higher-redshift searches like SNLS and DES, these historically were not coordinated rolling searches detecting all transients to a given magnitude. Instead they were dependent on following up transients discovered by other sources (including amateurs). Moreover, selection is skewed, for example, by search strategies in LOSS, which targeted galaxies. It is more ambiguous than the higher-redshift surveys how appropriate data/MC simulations are. The approach in JLA was to consider both the magnitude-limited scenario, which can be modelled by data/MC simulations, and the volume-limited scenario, where selection biases have no impact, and use the difference between them as the uncertainty (Section 3.4.2) in the bias correction. The targeted discovery of supernovae in CfA3 and LOSS means they should not be magnitude limited; however, as observed in the JLA low-$z$ sample, the colour distribution grows more blue with redshift, suggesting that some selection effect is at play. Approximating the bias correction for our low-$z$ supernovae in Chapter 4 using the JLA approximation is justified as our supernova sample is similarly distributed to the low-$z$ sample in JLA. In Pantheon, the magnitude-limited case is used as a baseline for low-redshift samples, averaging **G10** and **C11** scatter models (explained in Section 3.4.7); we move towards using this bias correction for Chapter 6.

#### Modelling bias

Approaches to estimating Malmquist bias broadly involve Monte Carlo simulations of SN Ia lightcurves, using simulation packages such as SNCosmo (Barbary et al., 2016) and SNANA (Kessler et al., 2009a). Bias in observable parameters (magnitude, stretch, colour) due to selection effects can be estimated from these simulations, and corrected for prior to cosmology fitting. For these corrections to be accurate, the parent distributions of SNe Ia need to reflect what is observed in nature. These parent distributions, characterised by intrinsic scatter, colour distribution, and stretch distribution at any redshift bin, are used as inputs for simulating lightcurves; survey variables are then applied to predict what will be detected. Mosher et al. (2014) performs this process, and includes SALT2 fitting for 'detected' lightcurves, to infer biases in distance modulus due to selection effects, i.e. for each redshift bin, average difference between output distance moduli inferred from detected SNe Ia only and input values.

JLA (Betoule et al., 2014, section 5.3) relies on simulations in M14 to estimate bias corrections as part of end-to-end simulations of supernova surveys, which include the training of SALT2, lightcurve fitting, and bias corrections. These also incorporate the uncertainty associated with the lightcurve model (discussed in Section 3.4.7). The underlying colour and stretch distributions are approximated by averaging over realisations of the simulations. These distributions were combined with realistic survey simulations and adjusted to match output (post-selection) distributions of parameters (redshift, stretch, and colour) with observables. Input and output distance moduli for these simulations were compared, to retrieve a bias in distance modulus binned by redshift. Several inputs for intrinsic scatter models going into simulations (**G10**, **C11**, coherent); these are explained in Section 3.4.7.

Simulation-based approaches to supernova cosmology (e.g. Bayesian hierarchical methods, Approximate Bayesian Computation, and BEAMS (Kunz et al., 2007; Kessler & Scolnic, 2017)) are taking the approach of using statistical methods. Instead of trying to model the bias, these run a large number of Monte Carlo simulations to forward model the bias correction. These methods have a place in future dark energy studies including DES; and are further described in Section 3.2.3; some (e.g. BEAMS with bias corrections, or BBC, which will be mentioned in Chapter 6) have been validated on DES-like simulations.

### 3.1.3 Peculiar velocities

Peculiar velocities arise from motion other than from cosmological expansion, such as dipole or bulk flows, local galaxy infall, and higher order coherent flows. These perturb the observed redshifts via the Doppler effect,[3] and can impact cosmological analyses. Hui & Greene (2006) show that neglecting correlations between peculiar velocity uncertainties at low-redshift results in a greatly underestimated zero point uncertainty, and degrades the precision of the dark energy equation-of-state parameter $w$. Moreover, correlated SN peculiar velocities can bias cosmological results: Davis et al. (2011) show that neglecting coherent flows results in a shift of $\Delta w = 0.02$.

Thus, an effort to quantify the uncertainty induced by correlated motions is an essential part of any modern SN Ia cosmological analysis. Approaches to this include the addition of large ($300 - 400$ km s$^{-1}$) uncertainties in redshifts to account for peculiar velocities (Ganeshalingam et al., 2013; Hicken et al., 2009b), and attempts to correct for peculiar velocities. The latter first appeared in SNLS (C11), which corrects redshifts on a supernova by supernova basis for the (line-of-sight) peculiar velocity at the location of the SN, as determined from a velocity field.

To isolate the cosmological redshift $\bar{z}$, i.e. the redshift due to pure expansion, we correct the observed heliocentric redshift $z_h$ for the other terms (e.g. Davis et al., 2011): the peculiar velocities $\mathbf{v}_{\odot}^{\mathrm{pec}}$ due to the motion of the Solar system relative to the CMB, and $\mathbf{v}_{\mathrm{SN}}^{\mathrm{pec}}$ due to the motion of an individual SN, also relative to the CMB. Many SNe at low redshifts share some of the Local Group's motion; by correcting for the Solar system's motion we are also overcorrecting for the motion of nearby SNe, necessitating the second correction. For a SN at position $\boldsymbol{n}$ from the Sun, these redshifts and velocity are related by [4] (Davis et al., 2011)

$$\begin{aligned} 1 + z_h &= (1+\bar{z})(1+z_{\odot}^{\mathrm{pec}})(1+z_{\mathrm{SN}}^{\mathrm{pec}}) \\ &= (1+\bar{z})(1-\boldsymbol{v}_{\odot}^{\mathrm{pec}}\cdot\boldsymbol{n}/c)(1+\boldsymbol{v}_{\mathrm{SN}}^{\mathrm{pec}}\cdot\boldsymbol{n}/c). \end{aligned} \tag{3.3}$$

The difficulty in performing the correction above is in estimating $\mathbf{v}_{\mathrm{SN}}^{\mathrm{pec}}$. We follow the approach used in SNLS and JLA, which uses galaxy densities (from wide-field redshift surveys) to trace mass densities, and approximate gravitational fields and infall velocities using a linear biasing approximation. (Alternative approaches have been to consider covariances between all supernovae in each redshift bin, or to not correct for peculiar velocities at all. Investigations in the effects of systematic errors/shifts in redshift are presented in Andersen (2018).) We use the 2M++ velocity field (Carrick et al., 2015, http://cosmicflows.iap.fr/) from the combination of 2MASS and 6dFGRS redshifts, updated from the IRAS PSCz fields used in JLA.

Under the above assumptions of a linear regime, the mass density and galaxy number density are proportional via a linear bias factor $b$, i.e. $\delta_g = b\delta$. Then peculiar velocities are proportional to

[3] A supernova's peculiar motion changes not only its redshift but also its observed luminosity (Hui & Greene, 2006; Davis et al., 2011) as it experiences relativistic beaming. This in turn induces a deviation in the supernova's peak magnitude; however this is approximately an order of magnitude smaller than the change in redshift.

[4] The minus sign in front of $\boldsymbol{v}_{\odot}^{\mathrm{pec}}$ arises because we have defined it is the motion of the Sun relative to the CMB, rather than the other way around.

gravitational attraction:

$$\boldsymbol{v}_{\rm SN}^{\rm pec} = \frac{\beta^*}{4\pi}\int_0^{R_{\rm max}} \delta_g(\boldsymbol{r}')\frac{(\boldsymbol{r}'-\boldsymbol{r})}{|\boldsymbol{r}'-\boldsymbol{r}|^3}\, d^3\boldsymbol{r}' + \boldsymbol{U}, \tag{3.4}$$

where $\boldsymbol{U}$ represents a residual dipole (in 2M++ this is the dipole of the Local Group) with $\beta^* = \frac{f(\Omega_m)}{b} = 0.43 \pm 0.02$ (Carrick et al., 2015) where $f(\Omega_m) = \Omega_m^{0.55}$ for ΛCDM (Wang & Steinhardt, 1998). For a given density field $\delta_g$, the velocity field derived through Equation 3.4 can be parametrized by $\beta^*$, the ratio of the growth rate of density perturbations to the linear bias factor.

To apply corrections, we take the component of $\boldsymbol{v}_{\rm SN}^{\rm pec}$ in the radial direction (i.e. a dot product with its position) as in Equation 3.3. For our analyses in Chapters 4 and 6, we use $z_{\rm cor}$ converted using Equation 3.3 as an estimate of the CMB-frame redshift $\bar{z}$.

The peculiar velocity corrections we make here are reliant on predicted velocity fields, which are intrinsically approximate. In Section 3.4.4, we vary $\beta^*$, which parametrises the velocity field above, to estimate the uncertainty due to the peculiar velocity correction. We discuss and quantify these uncertainties, and provide a framework for propagating them to the SN fit parameters.

## 3.2 Bayesian methods for parameter estimation

Bayesian inference methods have become the standard for cosmological analyses for a number of reasons: the prevalence of incomplete astronomical data sets, the need to combine multiple heterogeneous data sets, and the benefits of quantifying beliefs as probabilistic distributions. The power of Bayesian analysis is the capability to deal with partial data, by updating existent or 'prior' beliefs with each new or additional data set. This method is mathematically described in Bayes' theorem:

$$P(\boldsymbol{\Theta}|\mathcal{D}, H) = \frac{P(\mathcal{D}|\boldsymbol{\Theta}, H)P(\boldsymbol{\Theta}|H)}{P(\mathcal{D}|H)} \tag{3.5}$$

$$\propto \mathcal{L}(\boldsymbol{\Theta})P(\boldsymbol{\Theta}|H) \tag{3.6}$$

where $H$ is a hypothesis or model, described by a vector $\boldsymbol{\Theta}$ of parameters, and $\mathcal{D}$ represents the data set as a whole (i.e. all observations which can be compared to the model). Equation 3.5 is the cornerstone of Bayesian analysis; simply put, it states that the **posterior** $P(\boldsymbol{\Theta}|\mathcal{D}, H)$ is proportional to the product of the **likelihood** $\mathcal{L}(\boldsymbol{\Theta}) = P(\mathcal{D}|\boldsymbol{\Theta}, H)$ and the **prior** $P(\boldsymbol{\Theta}|H)$.

These symbols can be broken down as *beliefs* about $\boldsymbol{\Theta}$ and $\mathcal{D}$: the prior probability distribution function (PDF) reflects the existing beliefs, before encountering data, about the parameters $\boldsymbol{\Theta}$. The PDF is defined over a subset of $n$-dimensional space (where $n$ is the number of parameters in $\boldsymbol{\Theta}$). The likelihood is the probability of the model parameters $\boldsymbol{\Theta}$ producing the observed data $\mathcal{D}$; that is, the probability of measuring $\mathcal{D}$, *given* $\boldsymbol{\Theta}$. A function of both the data and theoretical model, the likelihood is the mechanism for updating beliefs about the probability of model parameters $\boldsymbol{\Theta}$. The posterior PDF reflects the *updated* beliefs about the parameters, or the probability distribution of $\boldsymbol{\Theta}$, *given, or having measured*, the data $\mathcal{D}$. The **evidence** in the denominator reflects all observations, and is constant (independent of the hypothesised parameters), leading to the second, more commonly seen, form of Bayes' theorem as a proportionality relation (Equation 3.6). Often, the hypothesis $H$ is implied, and omitted from the above equations.

The common goal of most cosmological analyses is to glean more information about a model of the Universe. This can occur in several ways: testing the fit of a hypothesis model, selecting between multiple competing models, or constraining the parameters of a model. Bayesian inference is used for all of these; Hobson et al. (2014) reviews a myriad of applications of of Bayesian methods to cosmology. We concentrate on parameter estimation: unless otherwise specified, the problem we want to solve is to determine (i.e. obtain posterior PDFs of) parameters of the standard ΛCDM or $w$CDM cosmological model.

While Bayes' theorem appears simple, the computations involved are often complex, or even computationally impossible. Rather than applying Equation 3.5 directly, probabilistic methods are used to find the PDF, which rely on sampling efficiently and faithfully over the prior distribution, making use of Monte Carlo or other simulation methods, and approximating over large numbers (the central limit theorem). We discuss the most common approach, Markov Chain Monte Carlo (MCMC), in Section 3.2.2, first introducing the basics of parameter estimation.

### 3.2.1 Classical curve fitting

Perhaps the most well-known form of parameter estimation is fitting a line or curve to data. The simplest form this can take is a line $f(x) = mx + b$. For a data set $(\boldsymbol{x}, \boldsymbol{y})$, many lines parametrised by gradient-intercept pairs $(m, b)$ are tested, and one is found 'closest' to the data under some metric. This holds more generally for any curve and number of dimensions. We can formalise these by replacing the vector $\boldsymbol{x}$ of data with the matrix $\mathbf{X}$ whose columns are each an individual data point, and rows are different variables such that each observation is represented by vector $\boldsymbol{X_i}$. The curve or function to fit is generalised $f(\boldsymbol{X_i}, \boldsymbol{\Theta})$ for vector $\boldsymbol{X_i}$ of data and with model parameters $\boldsymbol{\Theta}$. Under a Euclidean metric (i.e. a least-squares fit), the statistic to minimise is

$$\sum_i (y_i - f(\boldsymbol{X_i}, \boldsymbol{\Theta}))^2, \tag{3.7}$$

where the sum is over all data points. The data points in Equation 3.7 have equal weight, whereas in practice, often some points have larger errors, and so the fit parameters should reflect a higher tolerance in these points. A modification of Equation 3.7 to account for this minimises the $\chi^2$ **statistic**:

$$\chi^2 = \sum_i \frac{(y_i - f(\boldsymbol{X_i}, \boldsymbol{\Theta}))^2}{\sigma_{y_i}} \tag{3.8}$$

However, Equation 3.8 still assumes that the errors $\sigma_{y_i}$ in the data are uncorrelated. Neglecting to account for correlated systematic errors can lead to biases, and/or gross underestimation of errors in results, an effect which motivates most of the work within this thesis. Correlations between data uncertainties can be written using a **covariance matrix**:

$$\mathbf{C}_{ij} = \langle \sigma_{y_i} \sigma_{y_j} \rangle. \tag{3.9}$$

Generalising Equation 3.8 further gives

$$\chi^2 = (\boldsymbol{y} - f(\mathbf{X}, \boldsymbol{\Theta}))\mathbf{C}^{-1}(\boldsymbol{y} - f(\mathbf{X}, \boldsymbol{\Theta}))^T \tag{3.10}$$

where the summation (as in Equation 3.8) is implicit. Assuming normal distributions, the likelihood can be written as a function of the data and covariance as

$$\mathcal{L}(\boldsymbol{\Theta}) = \exp\left(-\frac{\chi^2(\boldsymbol{\Theta})}{2}\right) \tag{3.11}$$

$$-\log \mathcal{L} = \frac{1}{2}(\boldsymbol{y} - f(\mathbf{X}, \boldsymbol{\Theta}))\mathbf{C}(\boldsymbol{y} - f(\mathbf{X}, \boldsymbol{\Theta}))^T. \tag{3.12}$$

$$\tag{3.13}$$

The covariance matrices described in Section 3.4 fit into this picture within the likelihood term, as a way of quantifying correlated uncertainties in the data.

### 3.2.2 Monte Carlo sampling methods

This section discusses some Bayesian methods for estimating cosmological parameters, specifically those which use Monte Carlo simulation to explore a parameter space. We use Markov Chain Monte Carlo (MCMC) extensively in the analyses in Chapters 4 and 6, and nested sampling in the higher-dimensional fits in Chapter 4. Section 6.5.1 contains a description of our use of MCMC in analysing DES data, while Section 4.4 refers to our fits using nested sampling with the MultiNest algorithm. In Bayesian inference, the goal is to estimate a posterior probability distribution function (PDF) from the data. This is often achieved by sampling or 'walking' the parameter space, given a likelihood term as a function of the data and model, and a prior distribution. Given a prior PDF reflecting the beliefs held before examining the data, and the likelihood, the posterior PDF can be retrieved through Monte Carlo methods.

For each parameter, **summary statistics** condense the information about parameters contained in a set of data into a smaller but sufficient subset, for example the parameter values at the mean, standard deviation, and maximal likelihood. When dealing with large parameters spaces, many parameters in $\Theta$ are necessary for computing the likelihood but are not parameters of interest; these are called **nuisance parameters** which are **marginalised** over, i.e. the posterior likelihood is integrated over all values of that parameter.

In higher dimensional parameter spaces, the computational expense of calculating and integrating the likelihood necessitates Monte Carlo techniques to statistically sample the parameter space.

#### Markov Chain Monte Carlo

MCMC is by far the most well known and widely used of Bayesian sampling techniques in cosmology and the wider sciences. A **Markov chain** is a discrete-time stochastic process where the motion at each point only depends on the stage before it or $n$ previous stages (that is, it has no memory), and 'Monte Carlo' refers to the broader technique of random sampling a distribution many times to solve a numerical problem. The combination of the two, MCMC, describes a class of methods now standard in cosmology (Lewis & Bridle, 2002), characterised by a series of points (**walkers**) which move through the parameter space as dictated by algorithms determined from the likelihood function, and converge to the desired posterior.

In all MCMC methods, a likelihood and prior PDF are required, the latter often taken from a family of tractable distributions e.g. uniform or gaussian. A number of samplers or **walkers** have their positions generated from the prior distribution. At each step of each walk, a transition to a new position is proposed by drawing from a proposal density distribution, and accepted or rejected according to some metric, determined from the likelihood. There are numerous algorithms for computing this process, the most common being the Metropolis-Hastings and Gibbs algorithms. The cumulative trajectories of the walkers are referred to as **chains**, and their positions in the parameter space as a distribution informs the posterior PDF. Sensitivity to the starting position of each walker can be mitigated by burn-in (discarding the first $N$ points of each chain), and various tests for convergence.

Limitations to MCMC include its reliance on an explicitly computed likelihood, which may be computationally complex; its potential dependence on a starting position (where it is possible to get stuck in a local extremum) and difficulties in testing for convergence. The common algorithms for determining transition, Metropolis-Hastings and Gibbs, can have suboptimal speed particularly in more dimensions; some of the following variants of MCMC (sometimes referred to as 'MCMC-like' techniques, or included within the MCMC umbrella) perform better in some circumstances.

#### Nested sampling and other variants

Nested sampling (Skilling, 2004) is a similar technique to MCMC, with the notable characteristic of using the likelihood function to map the many-dimensional parameter space into one dimension by determining a sequence of subspaces in the prior PDF (with size called 'prior mass') enclosed by contours of equal likelihood, each of which encloses the next. This effective reduction in dimensionality means nested sampling performs better than MCMC in higher-dimensional spaces; for this reason we choose it for fits with more than approximately six parameters. A common nested sampling algorithm, MultiNest (Feroz & Hobson, 2008; Feroz et al., 2009, 2013) (with Python implementation `PyMultiNest` described in Buchner et al. (2014)) can robustly retrieve posterior samples from distributions which may have multiple peaks or 'modes', or large degeneracies (Hobson et al., 2014). In MultiNest, the likelihood is evaluated at sample **live points** (similar to the 'walkers' in MCMC), drawn initially from the prior distribution. At each step the point with lowest likelihood is replaced with a point within the iso-likelihood contour. This way the live points are iteratively replaced, and the sequence of contours (specifically, their prior masses) monotonically shrinks until convergence, where the posterior PDF is recovered from the positions and histories of the set of live points, which are similar to MCMC chains.

A variant of MCMC called **simulated annealing** uses thermodynamic analogies to 'speed up' (increase the gradient of) walkers when in areas of low probability, with the effect of searching more broadly in those areas, while sampling the higher-probability areas with greater resolution. Similarly, Hamiltonian Monte Carlo models the 'energy' of the system using a Hamiltonian function and favours states with higher energy, improving performance compared to MCMC. Sequential (or Particle) Monte Carlo increases computational efficiency by pooling points in 'particles', where a transition kernel function determines moving to the next stage.

### 3.2.3 Beyond MCMC

The above Bayesian methods for parameter estimation, MCMC and nested sampling, have served astronomers well thus far; however, their limitations include needing theoretically prescribed functions in the likelihood, namely models of the data, including its systematic errors. Computing these explicitly and to the degree of precision required can be challenging in modern and future cosmological surveys, where it is especially important to be impervious to systematics. Throughout the analyses in this work, we make good estimates of both the data and systematics: a model for distances to spectroscopically normal SNe Ia assuming distances from a standard FLRW model of the Universe with $w$CDM or $\Lambda$CDM cosmology, with statistical and systematic errors for these data quantified by covariance matrices (described primarily in Section 3.4, but also in Sections 4.5.3 and 6.5.2). We work using these robust assumptions, but are aware that we rely on them, and also recognise some recent developments of Bayesian methods which are independent of these assumptions, or likelihood-free. We will describe the motivations for and details of two such methods: approximate Bayesian computation and Bayesian hierarchical methods.

#### Approximate Bayesian computation

There are benefits to forgoing an explicit likelihood term: it may be difficult to model (e.g. selection effects in SN Ia cosmology), and/or covariance matrices can be computationally expensive to evaluate and invert at every point. Approximate Bayesian Computation (ABC) circumvents this by using **forward model simulation**: for each set of model parameters in parameter space, reliably generating a realisation of the data. For supernovae, any simulation package that can generate realistic observations (lightcurves) from a cosmology and SN parameters can be used, with the most common examples being SNANA (Kessler et al., 2009a) and `sncosmo` (Barbary et al., 2016). Recently ABC has been formalised for cosmology in Jennings & Madigan (2017) and applied to supernovae in Jennings et al. (2016). In these works Sequential Monte Carlo is used for efficiency.

By going straight from the model to inferred 'model' data which can be compared with observed data, ABC allows **likelihood-free** parameter estimation. Instead, the aim is to model posterior truthfully without likelihood. A mathematical model for the data is still necessary, specifically for turning parameters, drawn from the prior, into simulated data; however the likelihood no longer needs to be calculated. The comparison is then between the forward-modelled data and the actual observed data. Points within the prior are accepted if they are sufficiently or 'approximately' close in distribution. Eventually, the distribution of accepted points is a model for the posterior (like MCMC, in finite time this is approximate). As with MCMC, a PDF for each model parameter is obtained by selectively rejecting sample points according to some metric. In continuous systems, distributions will match exactly with probability zero. Thus, the distributions – the data, and that forward-modelled from points drawn from the prior – are compared using a threshold. Some metric (e.g. Euclidean at some number of points in the distribution) can be used, along with a threshold $\epsilon$. Instead of comparing a full forward-modelled distribution it is often useful to use a number of summary statistics (e.g. the mean) which can be compared more easily.

**Bayesian hierarchical methods**

A limitation of maximal likelihood methods is that they only allow for errors in the 'Y' or dependent data (in this context the distance moduli $\mu$ of supernovae) to be taken into account (via covariance matrices), and not errors in the 'X' or independent data (the redshift $z$). This limitation was one motivation (Gull, 1989) for Bayesian hierarchical methods (BHM), which include layers of hierarchy in the model parameters. In BHM there is a distinction between observed quantities and hidden **latent variables**, to reflect that observation is inexact, and that there are degrees of knowledge obtained through observations. While performing simulations for BHM, the 'true' values of parameters are separated from their observed values. Similarly, intrinsic variation in observables is separated from observational uncertainties. A model in BHM is characterised by its **hyperparameters**, in this context the parameters of its prior distributions (rather than of the model), defined by informative priors (**hyperpriors**). Instances of BHM applied to supernovae have included March et al. (2011); Rubin et al. (2015); Shariff et al. (2016); Mandel et al. (2009), Wolf et al., (in prep.), Hinton et al., (in prep.); the mathematics of various applications of BHM is formalised in these works.

Of these, Hinton et al., (in prep.) and Wolf et al., (in prep.) describe methods applied to the DES data discussed in Chapter 6. Hinton et al., (in prep.) simulate layers of supernova observables including selection effects, separating observed quantities from simulated hyperparameters of the parent SN Ia populations and underlying cosmology. A likelihood function is computed explicitly, but in a large number of parameters; the many-dimensional parameter space is sampled using Hamiltonian Monte Carlo. In contrast, like ABC, BAMBIS also involves forward modelling SN Ia lightcurves at each point in parameter space. The model is allowed to include stochastic effects assessed through simulations from same point in parameter space, to realise sampling variations. Monte Carlo simulations rather than prescribed values are used for the model mean and covariance. The distributions (modelled and observed) are compared, and selection effects and sampling variance are taken into account in the MC simulations of underlying distributions.

## 3.3 Machine learning

The regression and deduction methods thus far discussed rely on a spectrum of statistical and computing methods implemented by humans. Even if a technique or algorithm is performed by a computer, each part of it had been taught by a person as a set of explicit instructions, known as **hard-coding**. However, in recent years, developments in artifical intelligence have advanced to the ability to program computers to solve problems and/or reveal insights without explicitly being programmed to do so. The resultant learning processes resemble the human practices of learning from examples and experiences, earning the label **machine learning**. This mimicry of human learning now forms a significant portion of artificial intelligence research, justifying the long-used term.

Advances in machine learning (ML) have enabled methods of problem solving which were previously computationally impractical or impossible. Especially as data sets in astronomy have grown exponentially, these techniques have been more widely utilised. In some cases, such as classifying galaxies, determining if a detection is real or bogus, or identifying a handwritten digit, the task would be clear to a human and only requires judgement (which remains supremely difficult to hard-code), and the utility of ML is simply a matter of scale and/or time. In other cases (the more subtle examples of, for example, examining patient data and predicting if a given patient will develop kidney disease, or classifying a supernova from some photometry points) the deduction processes involved largely remain a mystery and are not easily human-reproducible. Thus, ML techniques appear to provide a more 'top down' approach to fitting data: instead of prescribing a model for the data and selecting the model or its parameters to fit the data, they start from the data and empirically infer a body of knowledge which is able to make predictions. In other words, the algorithm works out the inside of the 'black box', for a process a human may not be able to perform. We next provide a general overview of ML, before focusing on astronomy-specific applications in Section 3.3.2.

### 3.3.1 Basic principles

There are two broad classes of ML problems, using supervised and unsupervised learning. In astronomy supervised learning is of the most use, as there is typically a prediction or classification to be made, aided by external input. The other kind of ML is unsupervised learning, where a machine seemingly gathers insights from a data set without guidance, or gleans the underlying structure of the data. The aim is typically to separate data into discrete categories (without needing to label the categories) or reduce the dimensionality of data. For example, **$K$-means clustering** places points for a data set with some metric and assigns data to a point, updating these points to the centroid of each cluster until convergence. Principal component analysis (PCA) is one way to represent a data set in fewer dimensions, by iteratively updating eigenvectors so as to reduce the data to fewer components while retaining almost all of its information.

Supervised learning is an extension of standard regression methods discussed in the previous section: with supervised learning there is a correct **label** (the answer $\boldsymbol{y}$ for the data $\mathbf{X}$ in the language of Section 3.2.1). Supervised learning can take the form of linear or logistic regression, with continuous or discrete predictions, respectively. Both operate on similar principles, similar to classical techniques e.g. least squares polynomial fitting, where a model is trained to make quantitative predictions. These predictions are continuous for linear regression, and otherwise composed with e.g. sigmoid functions to map to discrete results for logistic regression for classification problems. Training is the process of allowing the machine to learn this model from **training data** consisting of inputs $\mathbf{X}_{\text{train}}$ and the correct labels (values or classifications) $\boldsymbol{y}_{\text{train}}$. This involves fitting the data with some linear combination of features, which could be e.g. multivariate polynomials, and 'learning' the coefficients. The variable portion of this process, fine-tuned by humans, involves assessing the success of the learning using various metrics and criteria; the features and hyperparameters (the parameters of the learning process) must be human-selected using the training set. Following this, the model may be tested on disjoint **validation data** for further evaluation, before application to the further disjoint **test data** to make predictions.

Within the broader context of machine learning there are many methods and algorithms, and within each one, numerous tunable parameters that determine the success and efficiency of the algorithm, by many different metrics and diagnostics. These include the purity (classified positives which are not false positives) and efficiency (proportion of events detected), to account for possible contamination or misclassification. In terms of the rates of false positives (FPR) and false negatives (FNR), the purity is $1-$FNR, and efficiency is $1-$FPR. The trade-off between purity and efficiency is typically assessed using **receiver operating characteristic** (ROC) curve (FPR vs FNR), where a minimal **Area Under the Curve** (AUC) is optimal.[5] Other useful metrics are the accuracy (proportion of correct classifications) and rate of learning, measured using learning curves of accuracy (of training and validation sets) as a function of training set size. Potential traps include overfitting (training the model too well, so that it is very specific to the training data, and is not general enough to result in a good accuracy in the validation data, i.e. high variance) or underfitting (the model is not specific enough, so that it does not get better for either training or validation data with more training examples, often because not enough features are used, i.e. high bias); these can be diagnosed using learning curves. Examples of supervised learning algorithms are: **decision trees** (classification trees with a binary split at every attribute or feature; these are referred to as **boosted** if the speed is increased by using a gradient to decrease bias), **random forest** classifiers (similar to decision trees, with randomly selected features and training subsets, which are then averaged over many drawings), **Gaussian process regression** (explained in Section 3.3.4), and **neural networks** (multi-layered models which self-organise to process intermediate inputs and

[5] Some ROCs plot the efficiency, or true positive rate, against the FPR; in this case a maximal AUC (close to 1) is optimal.

outputs in an analogue to the brain, earning the name).

### 3.3.2 Applications in astronomy

In the past decade, the capacity of astronomical surveys to detect and measure vast numbers of objects (such as supernovae, galaxies, exoplanets, and so forth) has grown immensely. In the era of the Large Synoptic Survey Telescopic (LSST), astronomy has irrevocably become intertwined with data science. Many large supernova surveys, including the Dark Energy Survey and SkyMapper (discussed in Section 2.4.4, with machine learning applications described in Scalzo et al. (2017)), have already ventured in this direction, incorporating machine learning and data science techniques into survey operations. There are obvious benefits, of being able to robotically perform tasks which had previously been to complex and/or difficult to program explicitly, but are simple by eye for a human. Not only are such tasks tedious, but they will soon be impossible to perform in real time as data sets increase exponentially. In astronomy, obvious examples include object (e.g. galaxy) classification, real-bogus classification of images, and detecting objects in an image; some of these are common to wider sciences: for example, identifying infrared sources corresponding to radio jets from supermassive black holes in images[6] is – from a computing perspective – not dissimilar from identifying wildebeest in images of the savannah[7].

Thus, ML is the natural next step in a broader journey toward automation in astronomy, which began in the 20th century with CCDs replacing photographic plates, and the advent of robotic telescopes with automated observing procedures: pieces of technology that are now taken for granted. In recent times, automated SN imaging surveys including LOSS and SkyMapper robotically take all observations and reduce data with an entirely automated pipeline. Many tasks had previously required significantly more human attention and judgement. The benefit from having a computer perform them enables progression of astronomy, specifically the orders-of-magnitude increases in scale we now observe in galaxy and transient surveys. Human input is still needed in the training stage: to classify (i.e. give correct labels to) the training set and to validate and tune the training hyperparameters, to afterward to scan (i.e. eyeball and sanity-check) the machine classifications; however the input required is at a substantially smaller scale than before. Often, large volumes of human scanning (for training or verification) are outsourced to citizen scientists on platforms such as the Zooniverse[8]. This was first done for a SN survey in Galaxy Zoo Supernova (Smith et al., 2011); on several occasions, the SkyMapper Transient Survey had engaged citizen scientists for supernova classifications. With basic instructions, members of the public can classify images, supplementing ML classifications, and with the bonus benefits of science outreach and engagement of the general public. In supernova surveys, the two main areas where ML has most obvious utility are detection (Section 2.4.2) and photometric classification (Section 3.3.3). We will next discuss ML and other approaches to photometric classification; here we discuss detection with reference to SkyMapper, as part of the supernova discovery pipeline (discussed in Sections 2.4.2 and 2.4.4).

Difference imaging is used to compare new galaxy images to old, and the subtracted images are scanned for supernova candidates. Amongst the subtracted images which meet some crude criteria (e.g. some number of pixels with a number of counts), very few are supernova candidates, or even astrophysical in origin. The astrophysical (**'real'**) sources typically include asteroids, AGN, stars with real variability, and actual SN candidates. The remaining sources are **'bogus'**, including cosmic rays[9] which appear very differently, as a few saturated pixels rather than an image of an object. Apart from these, there are different bogus images, which are typically artefacts of the subtraction process, including bad subtractions, bad reference images (e.g. with some dead pixels, or at the edge of an image), or diffraction patterns. The application of ML to modern time-domain astronomy was developed in Bloom et al. (2012). Using PTF data, they trained ML classifiers for two of the most important and yet labour-intensive aspects of contemporary transient searches, the discovery and classification of candidates, and reached purity of 90% and efficiency of over 96%. Building on this proof of concept, Brink et al. (2013) further explored and tuned different supervised ML classifiers on PTF data, finding

[6] Radio Galaxy Zoo, https://radio.galaxyzoo.org/
[7] Wildebeest Watch, https://www.zooniverse.org/projects/aliburchard/wildebeest-watch
[8] https://www.zooniverse.org/
[9] Despite their name, these are for the most part produced within the instrument.

that a random forest (RF) algorithm outperforms other methods tested, including support vector machines (SVM) and logistic regression. They optimised the RF classifier, tuning hyperparameters and selecting features, using PTF training and validation data, finding an efficiency of over 92% (missed detection rate or FNR of 7.7%) for an accepted FPR of 1%. Building on these techniques, random forest real-bogus classifiers were implemented in SMT (Scalzo et al., 2017) and DES (Goldstein et al., 2015) detection pipelines. As mentioned in Chapter 2, machine learning has also recently been applied to spectral classification of supernovae Muthukrishna et al., (in prep.), but using artifical neural networks ('deep learning') with the TensorFlow framework (Abadi et al., 2015).

### 3.3.3 Photometric classification

The premise of using supernovae for cosmology without spectroscopic confirmation of type (such as in DES in its current form) relies on expectations that supernovae can be successfully classified with photometry alone. Jones et al. (2018) had recently presented such an analysis with Pan-STARRS supernovae. The efforts in DES to similarly present a photometric SN Ia sample for cosmology in the near future, will build on successful studies with SNLS data (which had photometry and spectroscopy of each supernova in its sample) that used cuts or machine learning to demonstrate the purity that could be achieved. Motivation for this is primarily limited photometric resources (telescope time, but also human power) as surveys rapidly expand by orders of magnitude, and become increasingly automated as capabilities (and investment in time in automating observations and reductions) increase.

The classification of supernovae from their spectra was introduced and discussed in Section 2.1.2 in the previous chapter. While spectroscopy remains the most reliable method for SN classification, obtaining spectra of 'live' transients is costly in terms both time and observational resources, as discussed in Section 2.4.5. As the scale of transient surveys continues to increase exponentially, they reach a point where obtaining a spectrum of each supernova becomes not just difficult, but impossible. The natural venture has thus been to validate alternate classification methods, using photometry alone.

SNLS (Section 2.4.3) has remained one of the best data sets for verifying photometric classification methods, consisting of a large set of SNe all with spectroscopic classifications, up to high redshift where signal-to-noise may be poor. Approaches to photometric classification of SNLS supernovae have involved both ML, and cuts. Bazin et al. (2011) details the deferred photometric pipeline (DPP), separate from the real-time SNLS analysis, and without spectroscopic information. The cuts on photometry described therein have been used to identify select analysis subsets of both simulated data and SN-like objects in SNLS; later comparison with spectroscopic classifications leads to estimates of $\sim$4% contamination using cuts alone. Other surveys have used photometric cuts, including SDSS (Campbell et al., 2013) and Pan-STARRS (Jones et al., 2017).

Using supervised ML, Möller et al. (2016) further evaluated the success of photometric classification using SNLS. Features extracted from the photometry include estimated (photometric) redshift and lightcurve parameters (e.g. rise time, fall time, amplitude) and their goodness-of-fit (i.e. of lightcurve parameters), and lightcurve fits in each band and their goodness-of-fit (i.e. of lightcurve fits). Classifiers tested include RF and other boosted decision trees (AdaBoost and XGBoost), which are evaluated on both simulated and real SNLS data in the deferred photometric pipeline. All three methods yield an AUC (as defined in Footnote 5) of at least 0.96, with XGBoost performing the best (at 0.98). An alternate ML method using hierarchical neural networks was developed by Karpenka et al. (2013) and applied to SNANA-simulated SN data (the Supernova Photometric Classification Challenge). These joint efforts to demonstrate that a purity of over 95% can reliably be achieved have underpinned subsequent efforts to measure cosmology using large photometric samples, notably Pan-STARRS, DES, and LSST.

### 3.3.4 Gaussian processes for machine learning

In statistics, a **Gaussian process** is any stochastic process described by a continuous function $f$ (indexed by a variable $x$), where behaviour at each point $f(x)$ is an independent normal random variable. Thus, they can be used to model real phenomena with unknown or estimated noise. In machine learning, Gaussian processes are a powerful supervised learning method for regression (Rasmussen & Williams, 2006).

The methods described in Section 3.2 are very effective for fitting a model to data (linear regression), in a paradigm where (i) great faith can be placed in a model $f(\mathbf{\Theta})$ and (ii) the errors or noise are perfectly well understood. In the absence of these optimal conditions, machine learning affords more flexibility and power: like other supervised learning methods, a model is learned from some training data and makes predictions from similar (but independent) data. The implicit model negates the need for assumptions on the form of the noise or an explicit likelihood function, an obvious advantage when these pieces are missing. An example is the problem of fitting a function to some set of points without restricting the solution to a particular family of curves, i.e. fitting over all possible curves, in a reasonable amount of time, while requiring that the resultant curve behaves 'nicely'.

In Gaussian process regression, the 'niceness' condition of a target function $f$ (which steers the posterior away from pathological behaviour) manifests in smoothness/continuity: if $x_i$ and $x_j$ are close together, then $f(x_i)$ and $f(x_j)$ should be too. This is implemented by using a 'kernel' function to model the covariance of the data (e.g. in Equation 3.10). The simplest mathematical description for this concerning two variables a multivariate gaussian function:

$$k_\alpha(x_i, x_j) = \exp\left(-\frac{(x_i - x_j)^2}{2\alpha^2}\right) \tag{3.14}$$

Diagonal errors are used to estimate the $\sigma_i$, while the kernel width $\alpha$ is a tunable hyperparameter. The prior functions are typically taken from a uniform distribution, i.e. the space is of all continuous functions with values in some finite range, at some finite number of points $x_i$. The kernel $K_\alpha = \delta_{ij}\sigma_i + k_\alpha(x_i, x_j)$ represents the joint multivariate gaussian distribution between any set of random variables $f(x_i)$; this takes the place of the covariance matrix $C$ in Equation 3.10, giving a **model-free likelihood function**. From this $n$-dimensional kernel and an e.g. uniform prior, a posterior set of functions (described by their values at these $x_i$, which have gaussian distributions, and specifically have means and widths) can be inferred.

Thus, in the spirit of Bayesian inference and other supervised ML techniques alike, GP regression allows functions to be modelled using training data, with minimal assumptions about the function and uncertainties: only a level of smoothness, empirical evidence (the input training data) and a prior (which can be uniform) are required. We demonstrate their use for modelling Malmquist bias estimates, as a function of redshift, for the DES 3-year SN sample in Section 6.5.2 in Chapter 6.

## 3.4 Covariance matrices

Our estimations of SN Ia systematics with the covariance matrix approach closely follow methods in SNLS and JLA, namely computing multiple covariance matrices, one for each systematic, to quantify correlations of systematic uncertainties between all SN Ia observables. These matrices sum to the $3N \times 3N$ matrix $\mathbf{C}_\eta$ which encompasses all covariances in the vector $\eta = \{m_{Bi}, X_{1i}, C_i\}_{1\leq i\leq N}$, which represents the $3N$ SN Ia observables (magnitude, stretch, colour, within the SALT2 framework presented in Section 2.2.3) for the $N$ supernovae in the sample.

For computing a likelihood term as in Equation 3.11, we require a covariance matrix for the quantity in the data that is compared to a model (in traditional terms a dependent or 'Y' variable). For a supernovae Hubble diagram this is the distance modulus $\mu$ (at a given redshift), although in Chapter 4 it is more natural to use the corrected apparent magnitude for a canonical SN Ia at that redshift, $m_B^\dagger$ as defined in Equation 4.22. To reach the $N \times N$ matrix $\mathbf{C}_\mu$, $\mathbf{C}_\eta$ is conjugated with the $N \times 3N$ matrix $\mathbf{A}$

(with $\mathbf{A}_{ij} = \delta_{3i,j} + \alpha\delta_{3i+1,j} + \beta\delta_{3i+2,j}$):

$$\mathbf{C}_\mu = \mathbf{A}\mathbf{C}_\eta\mathbf{A}^T + \mathrm{diag}\left(\frac{5\sigma_z}{z\log 10}\right)^2 + \mathrm{diag}(\sigma^2_{\mathrm{lens}}) + \mathrm{diag}(\sigma^2_{\mathrm{int,SN}}). \tag{3.15}$$

This matrix $\mathbf{C}_\mu$ goes directly into the likelihood in Equation 6.21; as unlike $\mathbf{C}_\eta$, it is dependent on the SN Ia stretch and colour coefficients $\alpha$ and $\beta$, and so it is recomputed at each step of the fit.

First we discuss the diagonal uncertainties which affect each SN individually, ascribed to uncorrelated uncertainties in redshift due to peculiar velocity uncertainties (distinct from the uncertainty in their corrections, described in Section 3.4.4), and perturbances in SN Ia magnitudes caused by gravitational lensing and intrinsic scatter in SN Ia absolute peak brightness (mentioned in Section 3.4.7). The values we adopt for these terms are based on those in C11 and B14, but vary between analyses (see Sections 4.5.3 and 6.5.2).

In simplest terms, each covariance matrix tracks how the observables in $\eta$ vary or are perturbed together when an underlying systematic (e.g. dust, calibration, Malmquist bias correction) is changed. This allows propagation of systematics through to parameters $\boldsymbol{\Theta}$ to be determined, by impacting the likelihood term, thereby affecting the widths of posterior distributions in the final parameters.

To understand $\mathbf{C}_\eta$, we separate it into statistical and systematic components, where the distinction between the two is explained further in Section 3.4.1. The contributions to $\mathbf{C}_{\mathrm{sys}}$ we consider are from uncertainties due to sources described earlier (Milky Way extinction, photometric calibration, and lightcurve model), and due to the corrections in Section 2.3.1 (peculiar velocity corrections, host galaxy mass dependence, and Malmquist bias corrections):

$$\begin{aligned}\mathbf{C}_\eta &= \mathbf{C}_{\mathrm{stat}} + \mathbf{C}_{\mathrm{sys}};\\ \mathbf{C}_{\mathrm{sys}} &= \mathbf{C}_{\mathrm{bias}} + \mathbf{C}_{\mathrm{cal}} + \mathbf{C}_{\mathrm{dust}} + \mathbf{C}_{\mathrm{host}} + \mathbf{C}_{\mathrm{model}} + \mathbf{C}_{\mathrm{nonIa}} + \mathbf{C}_{\mathrm{pecvel}}.\end{aligned} \tag{3.16}$$

We follow standard techniques to compute each covariance matrix, which involves enfolding partial derivatives of SN parameters with respect to each systematic, with the typical size of systematics:

$$\mathbf{C}_{\mathrm{sys}\,ij} = \sum_{k\in\kappa}\left(\frac{\partial\eta_i}{\partial k}\right)\left(\frac{\partial\eta_j}{\partial k}\right)(\Delta k)^2. \tag{3.17}$$

Here the sum is over all systematics $k$, each of size $\Delta k$. For some systematics (dust, calibration) we follow Equation 3.17 literally (e.g. Equation 3.21) by directly substituting in the systematic terms in $\kappa$ (the set of systematics). For others, we effectively work out the $\mathbf{C}_{\eta_{ij}}$ terms, sometimes by multiplying out the expected error vectors (denoted $\boldsymbol{E}$ below, with errors for $\eta$ from some systematic term). Then the matrix $\mathbf{C}_\eta$, with elements $\mathbf{C}_{\eta_{ij}} = \langle\sigma_{\eta_i}\sigma_{\eta_j}\rangle$, is equivalent to the matrix product of the error vector (a $3N\times 1$ matrix) with its own transpose (a $1\times 3N$ matrix), $\boldsymbol{E}\times\boldsymbol{E}^T$ (a $3N\times 3N$ matrix).

These calculations are intrinsically approximate (some more than others), yet even as estimates they are invaluable for gauging the contribution of each systematic term affecting SNe Ia, and affirming that we sufficiently account for each effect. Section 4.7.2 presents our assessment of these uncertainties.

In the remainder of the section we focus on individual systematic terms in a more general setting (i.e. cover common ground for the next few chapters), starting with systematic uncertainties in the correction terms: the Malmquist bias and host mass corrections to SN Ia magnitudes, and peculiar velocity corrections to redshift (which are then converted into magnitude), followed by the remaining systematic uncertainties.

### 3.4.1 Statistical uncertainties

The distinction between statistical and systematic errors blurs, as many uncertainties have sources for which both descriptors are appropriate. We adopt the separation used in C11, which defines statistical uncertainties as those that can theoretically be reduced by increasing the size of some data set. In this case the data sets are the measured supernovae on a Hubble diagram, and the training set used to define the SALT2 parameters (Guy et al. (2010), updated in B14). We separate these two terms into matrices $\mathbf{C}_{\mathrm{stat,diag}}$ and $\mathbf{C}_{\mathrm{stat,model}}$ respectively.[10]

The training of the SALT2 (initially SALT) model is described in Guy et al. (2005, 2007, 2010). The full model (Section 2.2.3) involves spectral time series of a family of spectroscopically normal SNe Ia, parametrised by their stretch and colour; there is inherently some statistical uncertainty in the parameters of this model due to the finiteness of the training sample. The calculation of these model parameters requires propagating statistical uncertainties in the lightcurve model through to supernova fit parameters $\eta$, as described in Guy et al. (appendix A3, 2010) and implemented in the `snpca` package.[11] We use the code `salt2_stat` from this package to directly compute $\mathbf{C}_{\mathrm{stat,model}}$.

The matrix $\mathbf{C}_{\mathrm{stat,diag}}$ arises from uncertainties in the observations of SN Ia lightcurves, which have finite sampling and measurement errors. These errors are directly determined from lightcurve fits, given as correlated uncertainties in $m_B, X_1, C$ (a $3 \times 3$ covariance matrix for each SN) reported in SALT2 outputs. The only correlations are between different observations in the same supernova, so only a diagonal strip of $\mathbf{C}_{\mathrm{stat,diag}}$ of width 3 is nonzero.

In summary, $\mathbf{C}_{\mathrm{stat}}$ is the sum of terms from statistical limitations of data sets. The two contributions to $\mathbf{C}_{\mathrm{stat}}$ carry statistical uncertainties from the finiteness of the training set used to define the SALT2 parameters (affecting the lightcurve model parameters, in $\mathbf{C}_{\mathrm{stat,model}}$), and of the measured supernovae (affecting the lightcurve fits, in $\mathbf{C}_{\mathrm{stat,diag}}$).

### 3.4.2 Malmquist bias correction

After fitting individual SN Ia lightcurves for quantities $m_B, X_1, C$, and before fitting the derived distance moduli for cosmology, each observed magnitude (or equivalently, distance modulus) is corrected for Malmquist or selection bias, according to methods in Section 3.1.2. The nature of the function $\delta\mu(z)$ varies between the analyses in Chapters 4 and 6, so we leave discussions of the specific functions for Sections 4.5.3 and 6.5.2. In each case, a function is usually evaluated at several discrete redshifts and its errors estimated; both the function and error are interpolated so they can be computed at each redshift value in the data sample. Thus, an error vector $\boldsymbol{E}_{\mathrm{bias}}$ of uncertainties in $\eta$ is populated (putting uncertainties into the $m_B$ entries, or every third entry). We treat the uncertainties in the bias corrections as correlated between all SN Ia pairwise, so the covariance matrix is

$$\mathbf{C}_{\mathrm{bias}} = \boldsymbol{E}_{\mathrm{bias}} \times \boldsymbol{E}_{\mathrm{bias}}^{T}, \tag{3.18}$$

that is, the matrix product of the uncertainty in the bias correction function $\delta\mu$ with its own transpose. Equivalently, each element $\mathbf{C}_{3i,3j;\mathrm{bias}}$ is the product of bias function uncertainties in $m_{B,i}$ and $m_{B,j}$.

### 3.4.3 Host mass dependence

As with the uncertainties in corrections for Malmquist bias and peculiar velocities in Sections 3.4.2 and 3.4.4 respectively, the contribution to the covariance matrix is a matrix product of the error vector. However not all supernovae have correlated errors, as specified below: an element $\mathbf{C}_{3i,3j;\mathrm{host}}$ of the covariance matrix is nonzero only when the $i$-th and $j$-th supernovae both have uncertainties in the host mass corrections.

[10] In B14 these two terms are combined as $\mathbf{C}_{\mathrm{stat}}$, while C11 sums $\mathbf{C}_{\mathrm{stat,diag}}$ and the three diagonal terms in our Equation 4.24 to their $\mathbf{D}_{\mathrm{stat}}$.

[11] Private SNLS communication.

The SN Ia magnitude zero point $M_B$ is corrected for the magnitude offset $\Delta M_B$ between high and low host galaxy stellar mass bins, as described in Section 2.3.2. The uncertainty in this correction is propagated to SN parameters in $\mathbf{C}_{\text{host}}$. As in SNLS and JLA, we treat the systematic associated with this correction as having two components: from potentially having attributed an individual supernova to the wrong host mass bin, and from the arbitrariness of the $10^{10}M_\odot$ cut. Both effects are discrete, so the computation of $\mathbf{C}_{\text{host}}$ differs from those of $\mathbf{C}_{\text{dust}}$ and $\mathbf{C}_{\text{cal}}$, which take partial derivatives with respect to continuous quantities.

In Chapter 4 we assume $\Delta M_B = -0.08$ mag, whereas it is a fit parameter in $\mathbf{\Theta}$ in Chapter 6. The size of this uncertainty is constant for all supernovae, so it suffices to use indicator functions: we write $\boldsymbol{B}$ for the indicator function of supernovae (or more precisely, their magnitude terms: the stretch and colour terms are unaffected) with their uncertainties in host mass spanning the $10^{10}M_\odot$ boundary. This uncertainty is uncorrelated between different supernovae, so its contribution to $\mathbf{C}_{\text{host}}$ is diagonal. We similarly write $\boldsymbol{H}_h$ and $\boldsymbol{H}_l$ for the indicator functions of SNe with host masses within an order of magnitude higher and lower respectively than the mass boundary (i.e. from $10^{10}$ to $10^{11}M_\odot$, and $10^9$ to $10^{10}M_\odot$ respectively) – each of these sets of supernovae is correlated with itself, but not the other. Then the covariance matrix is the sum of the diagonal part and the matrix product of these vectors:

$$\mathbf{C}_{\text{host}} = \Delta M_B^2(\boldsymbol{H}_h\boldsymbol{H}_h^T + \boldsymbol{H}_l\boldsymbol{H}_l^T + \text{diag}(\boldsymbol{B})). \tag{3.19}$$

### 3.4.4 Peculiar velocity correction

In SNLS and JLA, the uncertainty in redshift corrections for peculiar velocity (described in Section 3.1.3) is estimated by inserting perturbed values of $\beta^*$ (the ratio of the growth rate of density perturbations to the linear bias factor) and propagating the difference in the resultant $v_{\text{pec}}$ values to $z_{\text{cor}}$ (the peculiar velocity-corrected CMB-frame redshift, our best estimate for the cosmological redshift $\bar{z}$).

The velocity field predicts peculiar velocity corrections, which are applied to give $z_{\text{cor}}$ or $\bar{z}$ in Equation 3.3. The difference in redshift with different values of $\beta$ gives an uncertainty $\sigma_z$ in $z_{\text{cor}}$, which is then propagated to an uncertainty in distance modulus using Equation 4.21 or 6.22.

Then the $\sigma_{\mu,\text{pecvel}}$ terms in Equation 6.22 make up the nonzero terms (the terms corresponding to $m_B$ in the low-redshift sample) in the error vector $\boldsymbol{E}_{\text{pecvel}}$, which is then multiplied with its own transpose to give $3N \times 3N$ matrix $\mathbf{C}_{\text{pecvel}}$, as with $\mathbf{C}_{\text{bias}}$ above.

As described in Section 3.1.3 we have corrected individual SN redshifts for peculiar motion, using the 2M++ velocity field corrections. However, there is intrinsic uncertainty in these models, with variation between velocity fields generated from different galaxy density fields, and in some cases limited agreement between predicted and measured velocities (Springob et al., 2014; Scrimgeour et al., 2016).

Thus the significant contribution in the correction model itself must be taken into account. Below, we adopt the approach in C11 and B14, which is to use the uncertainty in the velocity field to inform $\mathbf{C}_{\text{pecvel}}$, the contribution to $\mathbf{C}_\eta$ from peculiar velocities. We emphasise that $\mathbf{C}_{\text{pecvel}}$ only accounts for uncertainties associated with peculiar velocity corrections (which are correlated), and not the uncorrelated (diagonal) scatter from random peculiar velocities, which is added as a diagonal term $\sigma_z$ to $\mathbf{C}_\mu$ (Equation 3.15, described in Section 3.4).

For a given density field $\delta_g$, the velocity field derived through Equation 3.4 can be parametrized by $\beta^*$, the ratio of the growth rate of density perturbations to the linear bias factor. In C11, $\beta^*$ is the systematic which encompasses the uncertainty in the peculiar velocity model; that is, $\mathbf{C}_{\text{pecvel}}$ is derived through Equation 3.17 with $k = \beta^*$. As this treatment of uncertainty lies within only one density field and model (that is, it does not account fully for velocities derived from different realisations/measurements of galaxy densities) we are conservative in using it; like C11 we perturb $\beta^*$ by five times its uncertainty[12]. Likewise we adopt $\beta^* = 0.43 \pm 0.02$ (Carrick et al., 2015) in the correction. To compute $\mathbf{C}_{\text{pecvel}}$, we measure the shift in $z_{\text{cmb}}$ when $\beta^*$ is set to 0.33 or 0.53 instead. The resultant difference in $z_{\text{cmb}}$ is propagated to $m_B$ using the derivative of Equation 4.9:

$$\sigma_{m_B} = \frac{5}{\log 10}\left(\frac{1}{z} + \frac{f'(z)}{f(z)}\right)\sigma_{z_{\text{cmb}}}. \tag{3.20}$$

This has no impact on the stretch and colour of SNe, so only the $m_B$ elements of $\mathbf{C}_\eta$ have non-zero entries from $\mathbf{C}_{\text{pecvel}}$.

### 3.4.5 Milky Way extinction

The covariance matrix $\mathbf{C}_{\text{dust}}$ is calculated by applying Equation 3.17 directly, where $\kappa$ contains a single systematic: the uncertainty in the Milky Way extinction. Perturbing the value of the extinction (encoded in the dust keyword @MWEBV in SALT2 inputs) and refitting lightcurves give the partial derivatives in Equation 3.21. Explicitly, the entries of the covariance matrix are:

$$\mathbf{C}_{\text{dust}ij} = \left(\frac{\partial\eta_i}{\partial E_{(B-V)}}\right)\left(\frac{\partial\eta_j}{\partial E_{(B-V)}}\right)(0.05 \times E_{(B-V)})^2. \tag{3.21}$$

In Chapter 4 we used the Schlegel et al. (1998, (SFD)) dust maps with a 20% relative uncertainty, with the intent of following JLA strictly (Section 4.5.3). For the later analysis in Chapter 6, we updated the dust map and uncertainty to to the Schlafly & Finkbeiner (2011) maps (equal to the SFD maps scaled by 0.86), with a 5% uncertainty; these choices are discussed in Section 6.5.2).

We verify that the partial derivatives of SN parameters $\eta$ with respect to Milky Way extinction are independent of the size of perturbation over a satisfactory range, and that our resultant $\mathbf{C}_{\text{dust}}$ is identical to the same matrix reported in B14 for the 60 SNe Ia in common.

### 3.4.6 Calibration

B14 and C11 emphasise the significant contribution of uncertainties in the calibration of individual surveys to the total error budget. We follow the methods therein and in Betoule et al. (2013) to reproduce the calibration covariance matrix relevant to our SN sample and the instruments used to observe them. Computing the calibration matrix $\mathbf{C}_{\text{cal}}$ relies on the same principle as in Section 3.4.5, but over multiple systematics. Calibration uncertainties are grouped into two types of systematics: uncertainties in the magnitude zero point (shifting the overall flux scale) and in the effective wavelength (shifting the transmission function in wavelength space), for each filter. These uncertainties, in the vector $\kappa$, depend on the instruments used to observe the data and training lightcurves, so are given in Chapters 4 and 6.

We directly consider the covariance matrix of calibration systematics $\mathbf{C}_{\kappa\mathbf{ij}} = \langle\sigma_{\kappa_i}\sigma_{\kappa_j}\rangle$, which captures the correlations between systematics in different instruments and passbands. Then Equation 3.17 is equivalent to $\mathbf{C}_{\text{cal}} = \mathbf{J}\cdot\mathbf{C}_\kappa\cdot\mathbf{J}^T$ where $\mathbf{J}_{ij} = \frac{\partial\eta_i}{\partial\kappa_j}$ is the Jacobian matrix, encoding partial derivatives of SN parameters. Then finding $\mathbf{C}_{\text{cal}}$ amounts to constructing $\mathbf{C}_\kappa$, and calculating $\mathbf{J}$ from first principles. Instruments in $\kappa$ are either those used for observations directly, and/or used for observing SNe Ia in the SALT2 training set; these two sets overlap as there are some low-$z$ supernovae in the SALT2 training samples included in our analyses in Chapters 4 and 6.

For computing the Jacobian $\mathbf{J}$, each systematic (element of $\kappa$, either a zero point or effective wavelength) is perturbed in the SALT2 instrument and the lightcurve is refit. The SALT2 instruments are either those used for observing the lightcurves, or involved in training SALT2. For the former, perturbing the model involves shifting the instrument zero points or effective wavelengths. The latter case concerns instruments used for observations of SN Ia lightcurves involved in training SALT2:

[12] Pike & Hudson (2005) find $\beta^* = 0.49 \pm 0.04$ so C11 vary $\beta^*$ between 0.3 and 0.7.

changing the systematic results in a different SALT2 model altogether (i.e. one that was trained with the systematic shift in question applied) which we then use for lightcurve fitting. Each SN lightcurve is fitted again to find the difference. The resultant Jacobian is smoothed following the process in footnote 9 of B14. The Jacobian part of computing $\mathbf{C}_{\text{cal}}$ is associated with the dependence of lightcurves on the calibration of magnitude systems.

The other part of computing $\mathbf{C}_{\text{cal}}$ is in the estimation of $\mathbf{C}_{\kappa}$ which contains the sizes of uncertainties (correlated and uncorrelated) in the photometric calibration of all magnitude systems involved. Our approach to this starts from the method in JLA, described in Betoule et al. (2014, section 3.4) for the calibration procedure in Betoule et al. (2013). More details of computing $\mathbf{C}_{\kappa}$ are given in Chapters 4 and 6, particularly for the FGCM calibration method for DES.

### 3.4.7 Lightcurve fitter and residual scatter models

The process of going from SN Ia lightcurves to distance moduli using the SALT2 lightcurve fitter involves a number of assumptions, in particular about the parent population distributions and intrinsic scatter of SNe Ia. In JLA, these are taken into account by the work in Mosher et al. (2014) to perform numerous realisations of end-to-end simulations of SALT2 training and fitting, applied to realistic SNLS-like surveys. In particular, they investigate the difference or bias in $\mu$ when comparing different colour-dependent scatter models (e.g. **G10**, **C11** for those in Guy et al. (2010) and Chotard et al. (2011) respectively, or **COH** for coherent, with no colour-dependence), input models (**G10** for SALT2 surfaces in Guy et al. (2010), **G10'** – a second model trained using **G10** as a starting point, described in Mosher et al. (2014), or **H** for Hsiao et al. (2007) templates), and real versus ideal characteristics of the test and training sets.

As in JLA, we assume coherent intrinsic scatter, roughly binned by redshift for each survey, with $\sigma_{\text{int,SN}}$ values determined in JLA (fit using the restricted likelihood method in Betoule et al. (section 5.5, 2014)) of 0.08 for DES (similar to SNLS in redshift range) and 0.134 for nearby supernovae. We follow JLA in conservatively using the largest bias in distance modulus from Mosher et al. (2014) (from the **G10'**–**C11**–REAL–REAL model in figure 16 therein) to account for the systematic error from the SALT2 lightcurve fitter and model, and assuming a coherent intrinsic scatter. As in computing $\mathbf{C}_{\text{bias}}$ and $\mathbf{C}_{\text{nonIa}}$, the error vector $\boldsymbol{E}_{\text{model}}$ is computed for supernova magnitudes using only redshifts, and multiplied by its own transpose (as a matrix product, as in Equation 3.18) to find $\mathbf{C}_{\text{model}}$.

#### Non-Ia contamination

Despite each supernova in our sample having been classified a SN Ia from its spectrum, there is still potential for contamination from primarily SNe Ib/c, which have spectra that at certain epochs and wavelength ranges or redshifts can resemble SN Ia spectra, particularly when signal-to-noise is poor. Our estimation of the systematic error due to potential contamination from misclassified supernovae follows methods in SNLS and JLA, which are described in Conley et al. (2011).

Individual supernovae in the DES3YS sample are spectroscopically classified with two degrees of certainty: as 'Ia' or 'Ia*', for certain and highly probable SNe Ia, respectively. For calculating $\mathbf{C}_{\text{nonIa}}$, the effective bias from potential misclassification is computed as a function of binned redshift. This effective bias is the product of the SN Ia* fraction and the 'raw' bias, or expected magnitude bias from a single misclassified supernova. The latter, shown in the second column of Table 3.1, is taken directly from estimates in Conley et al. (section 5.5, 2011) derived from simulations which assumes that contaminants are SNe Ib/c, and using rates and luminosity distributions in Li et al. (2011a), also based on Richardson et al. (2002); Bazin et al. (2009). The third column of Table 3.1 shows the SN Ia* fraction for the DES3YS sample in that bin, taken to represent the probability of misclassification. The product of both is the effective bias, given in millimags in the fourth column.

For each redshift bin, the error going into the $m_B$ terms in $\boldsymbol{E}_{\text{nonIa}}$ is the product of the raw bias

Table 3.1 .

| Redshift[a] | Raw bias [b] (mag) | SN Ia* fraction | Effective bias (mmag) |
|---|---|---|---|
| 0.10 | 0.015 | 0.0294 | 0.44 |
| 0.26 | 0.024 | 0.173 | 4.2 |
| 0.41 | 0.024 | 0.102 | 2.4 |
| 0.57 | 0.024 | 0.120 | 2.9 |
| 0.72 | 0.023 | 0 | 0 |

[a] Lower boundary of redshift bin
[b] Taken directly from Conley et al. (table 14, 2011)

and the SN Ia* fraction; this is applied to all supernovae in the redshift bin. Then,

$$\mathbf{C}_{\mathrm{nonIa}} = \boldsymbol{E}_{\mathrm{nonIa}} \times \boldsymbol{E}^{T}_{\mathrm{nonIa}}. \tag{3.22}$$

For future DES supernova analyses where the sample consists of photometrically classified supernovae, the above method will need to be revised with more simulations involving intricacies of photometric classification, as well as possibly newer estimates of the raw magnitude bias due to SN Ib/c contamination.

Chapter 4

# A blinded redetermination of the Hubble constant

The following two chapters address the issue of the current tension in the Hubble constant (Section 1.2.2) using methods in Chapter 3. My work here has focused on determining the Hubble constant, using data in recent SN Ia based calculations in Riess et al. (2011) (hereafter R11). We approach these data sets using a renewed approach in modern SN Ia cosmology studies (Conley et al., 2011; Betoule et al., 2014), notably including using full covariance matrices to capture correlated supernova systematics, and a simultaneous fit of all data to a set of equations. This chapter is based around (and makes use of portions of text from) the refereed paper Zhang et al. (2017) (hereafter Z17) which sets up the analysis framework using the established R11 data, while Chapter 5 discusses the conclusions of this chapter, particularly the increased error, in the context of the unresolved tension in $H_0$.

The structure of this chapter of is similar to Z17: we start in Section 4.1 by motivating the work presented therein, and framing its place within the literature. This is followed by an outline of the data in Section 4.2 and establishing the equations for measuring $H_0$ in Section 4.3. In Section 4.4 we discuss the fitting the sample of Cepheid variables used for calibrating the supernova distance measurements and the systematics which affect them. Sections 4.5.4 presents preliminary SN Ia-only fits while Section 4.6 contains results from the combined simultaneous fits.

## 4.1 Introduction

In Section 1.2 we introduced the current discrepancy between values of the Hubble constant derived from local distance measurements, and from observations of the CMB in the early Universe extrapolated to the present time assuming ΛCDM. The effort to measure $H_0$ from a distance ladder have been led by SH0ES (Riess et al., 2009b, 2011, 2016). Numerous reanalyses of the SN Ia-based measurement have followed, many of which have focussed on the methods for the rejection of Cepheid outliers. Efstathiou (2014, hereafter E14) questions and revises the outlier rejection algorithm in R11, concluding $H_0 = 72.5 \pm 2.5$ km s$^{-1}$ Mpc$^{-1}$ assuming a null metallicity dependence of the Leavitt law. Recently, Cardona et al. (2017) uses Bayesian hyper-parameters to down-weight portions of the Cepheid data for both R11 and R16 data sets, finding $H_0 = 73.75 \pm 2.11$ km s$^{-1}$ Mpc$^{-1}$ for the R16 data. Moreover, the dependence of the intrinsic magnitude of SNe Ia on host galaxy properties has been explored in recent years (e.g. Sullivan et al., 2010). Rigault et al. (2013, 2015) find a relationship between peak brightness and star formation rate, and infer an overestimate of ∼3 km s$^{-1}$ Mpc$^{-1}$ in the R11 value of $H_0$ arising from the fact that the calibration set of SNe Ia exist in galaxies which necessarily contain Cepheids, hence are likely to be late-type galaxies. However, Jones et al. (2015) repeat the same analysis, with an increased sample size and the R11 selection criteria applied, and find no significant difference in the brightness of SNe Ia in star-forming and passive environments.

The CMB data in Planck has been reanalysed in Spergel et al. (2015), who find a similar value to Planck Collaboration et al. (2014), of $H_0 = 68.0 \pm 1.1$ km s$^{-1}$ Mpc$^{-1}$. Bennett et al. (2014) provides a CMB-based measurement which is independent of Planck, by combining data from WMAP9, the South

Pole Telescope (SPT) and Atacama Cosmology Telescope (ACT), and baryon acoustic oscillation (BAO) measurements from BOSS, finding a value of $H_0 = 69.3 \pm 0.7$ km s$^{-1}$ Mpc$^{-1}$ (with a slight increase to $H_0 = 69.7 \pm 0.7$ km s$^{-1}$ Mpc$^{-1}$ if SN Ia data from R11 are included), which is slightly less discrepant with SN Ia-based values. Strong lensing provides an independent but model-dependent local measurement of $H_0$: the Suyu et al. (2017, (H0LiCOW)) program studies time delays between multiple images of quasars in strong gravitational lens systems, and find $H_0 = 71.9^{+2.4}_{-3.0}$ km s$^{-1}$ Mpc$^{-1}$in flat $\Lambda$CDM. It is noteworthy that the H0LiCOW analysis was performed blind to derived cosmological parameters.

One of the greatest open questions in cosmology today is whether the tension in $H_0$ signifies new physics – where inconsistencies between results from supernovae and the CMB arise from the model-dependence of the measurement, and disappear when the correct model is used – or is the result of some systematic error in one or both measurements that has yet to be accounted for.

A genuine inconsistency in the value of the Hubble constant at low and high redshifts would have profound consequences. Therefore it is imperative to fully understand uncertainties in the measured values of $H_0$, and to preclude possible human biases on the result, as introduced earlier in Section 1.5. The most effective way of achieving the latter is to blind the value of $H_0$ throughout the analysis, as described in Section 4.3.4.

### 4.1.1 Revisiting SH0ES

We revisit the work in R11, with the aim of producing an end-to-end reanalysis of the data, starting from the published SN Ia photometry. The main differences in our analysis are: a simultaneous fit to all data sets, and the accepted methodology of recent supernova cosmology analyses (Conley et al., 2011; Betoule et al., 2014) for considering SN Ia systematics using covariance matrices. Both of these points carry significant differences from the R11 and R16 analysis chains, and have yet to be included in a reanalysis. Nor has the supernova data been revisited in its entirety. Thus, we are motivated by the desire to provide such a validation of the supernova data, and by the current relevance and importance of the Hubble constant, to produce in this work an independent, blinded, end-to-end reanalysis of the R11 data to determine $H_0$ and its uncertainty.

Numerous improvements over R11 have been made in R16, in the analysis as well as the size and quality of data. Changes to the outlier rejection and the Cepheid metallicity calculations have addressed some of the concerns raised in E14. The use of data from R11 is for a proof of concept, necessary for our blind analysis technique, and is followed in Z18 (in prep; Chapter 5) with the same analysis applied to R16 data.

In summary, we combine the framework for calibrating a SN Ia Hubble diagram with Cepheid variables, with the best estimates of supernova systematics via covariance matrices. We determine $H_0$ using the magnitude-redshift relation (i.e. a Hubble diagram) of low-redshift SNe Ia, with their zero point set by Cepheid variables in host galaxies of eight nearby SNe Ia, which are in turn calibrated by very long baseline interferometry (VLBI) observations of megamasers in NGC 4258, and other geometric distances to the Large Magellanic Cloud (LMC) and Cepheids in our Galaxy.

## 4.2 Data

In this section we break down the data in our analysis into three distinct sets, and outline each one. We discuss the use of Cepheid variables as standard candles, given the relation between their brightness and pulsation period (the Leavitt Law, equation 4.7) and the systematics that affect them. For supernovae, we refer to Chapter 2 for details of their use as distance indicators (from lightcurves to distances), and Chapter 3 for discussions of their systematics.

Our reanalysis of R11 concerns the following sets of data:

1. **Cepheid variables:** 570 in nine nearby galaxies (see Table 4.2), namely:

- 165 in the distance anchor NGC 4258, and
- 405 in eight galaxies that host recent nearby SNe Ia.

2. **Anchor ('nearby') supernovae:** 8 recent SNe Ia in the nearby galaxies (also in Table 4.2).

3. **Low-$z$ ('Hubble flow') SNe Ia:** 280 low-redshift ($z < 0.06$) SNe Ia from the CfA3 (Hicken et al., 2009a) and LOSS (Ganeshalingam et al., 2010) samples.

Together these three data sets allow us to calibrate our distance ladder. The galaxy NGC 4258 hosts the water masers that give us a precise absolute local distant measurement (Humphreys et al., 2013), and allows us to calibrate the Cepheids. As in R11, we also use the LMC and Milky Way (MW) as distance anchors in combination with NGC 4258, relying on independent distances measured from detached eclipsing binaries (Pietrzyński et al., 2013) to Cepheids in the LMC, and Hipparcos and HST parallax measurements of Cepheids in our Galaxy (van Leeuwen et al., 2007). The Cepheids in turn enable us to calibrate the absolute magnitudes of the eight supernovae that occurred in nearby galaxies with quality Cepheid measurements. These then allow us to calibrate the whole supernova sample, which ultimately gives most of the constraining power for our $H_0$ measurement. In practice we perform a global fit to all of these samples together. In the next section we outline the equations needed to relate all of these standard candles and extract a measurement of $H_0$ following the theory in Section 4.3.

Since the purpose of this paper is to provide an independent analysis of the data in R11, we adopt an identical sample in order to make a faithful comparison. We later (Chapter 5) use the same framework to analyse newer data sets including SNe Ia in the CfA4 survey (Hicken et al., 2012) and Cepheids in R16.

### 4.2.1 Cepheids

Cepheid variables are powerful distance indicators to nearby galaxies, which play an important role in this analysis by tying the calibration of the SNe Ia in this analysis to an absolute scale. They are pulsating supergiant stars, with luminosities well-characterised via the empirical Leavitt law (Leavitt, 1908; Leavitt & Pickering, 1912) – also commonly known as the Period-Luminosity relation; we will briefly describe the pulsation mechanism in Section 4.2.1. The equations for this, including the Leavitt law and apparent magnitudes of rungs of the distance ladder as a whole, appear in Section 4.3.2. The brightness and regular pulsation of Cepheid variables as well as their ease of discovery and classification make Cepheids reliable distance indicators in the nearby Universe, and the basis of the cosmic distance ladder (Freedman & Madore, 2010). However, they are subject to difficulties and systematics, which include crowding and confusion (necessitating outlier rejection), metallicity, and extinction; these are discussed further in Section 4.2.1.

The Cepheids in the nine galaxies in Table 4.2 were discovered or reobserved in the Supernovae and $H_0$ for the Equation of State (SH0ES) project (Riess et al., 2009b) on the HST, from Cycle 15. Infrared (F160W) observations of the SN Ia host galaxies were made using the Wide Field Camera 3 (WFC3). We refer the reader to R11, section 2 for descriptions of observations and data reduction. Our initial data set consists of 570 Cepheids from R11, table 2, excluding those marked 'low P'; this number is reduced to 488 if we adopt the $P < 60$ day cut on Cepheids, following E14.

We supplement the sample of 157 Cepheids in NGC 4258 with LMC and MW Cepheids, used as alternative anchors (discussed in Section 4.4.2). Persson et al. (2004) presents near-infrared photometry of 92 Cepheids, of which 53 have optical measurements in Sebo et al. (2002), which we use for determining Wesenheit magnitudes. Two of these 53 Cepheids have period greater than 60 days, which we exclude if we impose the period cut on the Cepheids in the supernova host galaxies. We also make use of 13 Cepheids in the Milky Way from van Leeuwen et al. (2007, table 2) (excluding Polaris, an overtone pulsator), which have combined parallaxes from Hipparcos and HST data.

We briefly mention those that affect our method, and refer the reader to Freedman & Madore

(2010, section 3.1) and references therein for further discussion of Cepheid systematics. In Sections 4.4.2 to 4.4.2 we test and report the dependence of the Leavitt law parameters on aspects of the Cepheid fit, namely outlier rejection, distance anchor, and cut on Cepheid period. Section 1.3.1 Cepheid variables are powerful distance indicators to nearby galaxies, which play an important role in this analysis by tying the calibration of the SNe Ia in this analysis to an absolute scale. They are well-characterised by their luminosity via the empirical Leavitt law (Leavitt, 1908; Leavitt & Pickering, 1912) – also commonly known as the Period-Luminosity relation. The equations for this, including the Leavitt law and apparent magnitudes of rungs of the distance ladder as a whole, appear in Section 4.3.2.

The Cepheids in the nine galaxies in Table 4.2 were discovered or reobserved in the Supernovae and $H_0$ for the Equation of State (SH0ES) project (Riess et al., 2009b) on the HST, from Cycle 15. Infrared (F160W) observations of the SN Ia host galaxies were made using the Wide Field Camera 3 (WFC3). We refer the reader to R11, section 2 for descriptions of observations and data reduction. Our initial data set consists of 570 Cepheids from R11, table 2, excluding those marked 'low P'; this number is reduced to 488 if we adopt the $P < 60$ day cut on Cepheids, following E14.

We supplement the sample of 157 Cepheids in NGC 4258 with LMC and MW Cepheids, used as alternative anchors (discussed in Section 4.4.2). Persson et al. (2004) presents near-infrared photometry of 92 Cepheids, of which 53 have optical measurements in Sebo et al. (2002), which we use for determining Wesenheit magnitudes. Two of these 53 Cepheids have period greater than 60 days, which we exclude if we impose the period cut on the Cepheids in the supernova host galaxies. We also make use of 13 Cepheids in the Milky Way from van Leeuwen et al. (2007, table 2) (excluding Polaris, an overtone pulsator), which have combined parallaxes from Hipparcos and HST data.

We briefly mention those that affect our method, and refer the reader to Freedman & Madore (2010, section 3.1) and references therein for further discussion of Cepheid systematics. In Sections 4.4.2 to 4.4.2 we test and report the dependence of the Leavitt law parameters on aspects of the Cepheid fit, namely outlier rejection, distance anchor, and cut on Cepheid period.

**Pulsating variables**

Cepheids fall within a class of stars called pulsating variables, with periodically varying luminosities driven by competing hydrostatic and gravitational forces (e.g. Eddington, 1917). Pulsating variables also include RR Lyrae stars and Mira variables, with varied periods and temperatures. Most lie within the **instability strip** on the Hertzsprung-Russell diagram, within a narrow range of temperatures or spectral classes. Within this region, stars undergo pulsations driven by the double ionization of helium; their outer He layers cycle between contraction/dimming (increasing temperature and density, ionizing to the more opaque He III) and expansion/brightening (as the increased opacity leads to further increasing temperatures). Within this strip, Cepheids and other pulsating variables oscillate between their extremes of brightness, size, and temperature.

Classical Cepheids are supergiant stars lying in the instability strip several magnitudes brighter than the main sequence, characterised by periods of days to hundreds of days, with rapid increases in brightness followed by slower declines. The brightness and regular pulsation of Cepheid variables as well as their ease of discovery and classification make Cepheids reliable distance indicators in the nearby Universe, and the basis of the cosmic distance ladder (Freedman & Madore, 2010). However, they are subject to difficulties and systematics, which include crowding and confusion (necessitating outlier rejection), metallicity, and extinction; these are discussed further in Section 4.2.1.

**Cepheid systematics**

Careful treatment of Cepheids starts with their discovery and identification, where crowding and confusion can lead to misidentification. Light from a Cepheid can be blended with nearby or background sources, and aliasing or sampling problems can cause the wrong period to be inferred. Thus, outliers from the Leavitt law fit must be identified and rejected. Moreover the intrinsic scatter in the Leavitt law must be taken into account in assessing the goodness-of-fit; outliers that are rejected should lie well outside the so-called instability strip.

The secondary dependence of Cepheid luminosities on atmospheric metallicity is an ongoing area of research, and remains contentious. This effect arises from changes in the atmospheres and structure of Cepheids with their chemical composition, which impacts colours and magnitudes. There is evidence of a mild metallicity dependence at optical wavelengths (Kennicutt et al., 1998; Sakai et al., 2004; Macri et al., 2006; Scowcroft et al., 2009), which is weaker in the infrared. In the LMC, using spectroscopic [Fe/H] measurements, Freedman & Madore (2011) find that $Z_H$ (the metallicity dependence in the $H$-band) is close to zero. Efstathiou (2014, section 3.2) argues that these LMC observations, along with theoretical considerations, give cause to applying an external prior on the metallicity dependence centred at $Z_W \sim 0$. We discuss this prior, which we find is inconsistent with the R11 data, in Section 4.4.2.

Historically the zero point of the Leavitt law has contentious due to uncertainties in parallax measurements. To circumvent this, more accurate absolute distances have been pursued, including VLBI measurements of water megamasers in NGC 4258 (Humphreys et al., 2013). Multiple distance anchors are also tested and combined to reduce the impact of any single distance anchor. The effects of varying and combining anchors is explored in this analysis in Section 4.4.2, following R11 and E14.

### 4.2.2 Supernovae

Our supernova data are identical to the data set in R11, consisting of eight 'nearby' SNe Ia in the galaxies hosting Cepheids (Table 4.2), and 280 unique Hubble flow 'low-$z$' SNe Ia from the 185 CfA3 (Hicken et al., 2009a) and 165 LOSS (Ganeshalingam et al., 2010) samples.[1] Details of photometry sources for the nearby supernovae are presented in Table 4.1. Natural photometry was not available for the oldest two, SN 1981B and SN 1990N. The most recent SNe are already in both CfA3 and LOSS, so we used combined photometry from both sources, and check these fits are consistent. The remaining three (SN 1994e, SN 1995al, SN 1998aq) were observed on the FLWO 1.2 m telescope with a variety of CCDs; we construct SALT2 instruments (including transmissions and zero points) using data from Jha et al. (2006).

CfA3 ran from 2001 to 2008 on the 1.2m telescope at FLWO almost entirely with the CfA3 4Shooter2 and Keplercam imagers (in $UBVri$ and $UBVRI$ filters respectively), while LOSS took place on the NICKEL and KAIT telescopes from 1998 to 2008 (in $BVRI$). Unlike more recent magnitude-limited surveys, CfA3 and LOSS targeted known galaxies and include SNe discovered by other sources, resulting in a more complex selection function and generally resulting in higher host galaxy masses (Section 2.3.2). We refer the reader to the above works for further details of observations. Newer low-$z$ SNe Ia samples have since been published, notably CfA4 (Hicken et al., 2012), Carnegie Supernova Project (CSP-II; Stritzinger et al., 2011), Pan-STARRS (Rest et al., 2014), Palomar Transient Factory (PTF; Rau et al., 2009), and La Silla-QUEST Supernova Survey (LSQ; Walker et al., 2015). However we retain the older CfA3-LOSS sample for this analysis to more faithfully compare our results to R11 and E14.

Photometry for the low-$z$ sample is sourced from Hicken et al. (2009b) and Ganeshalingam et al. (2013) in the natural systems of each filter set, with the exception of the CfA3 4Shooter2 and Keplercam $U$ filters for which reliable measurements do not exist – we use photometry in the standard Johnson-Cousins $UBVRI$ system as presented in Bessell (1990) for these passbands only, as well as the nearby SN 1981B and SN 1990N.

We use SALT2 (Guy et al., 2007) to fit these SN Ia lightcurves for the quantities $m_B, X_1, C$ (apparent peak magnitude, stretch, and colour respectively) according to details in Section 2.2.2. These

[1] There are 69 SNe in common between the samples; however SN 1998es was discarded because the lightcurve quality was so poor that the SALT2 lightcurve fit failed.

Table 4.1 Observations of nearby SNe Ia in Table 4.2, including sources of photometry, SALT2 instruments, magnitude systems (including filters) where available. Lightcurves of the two earliest supernovae were given as standard photometry only.

| SN Ia | Photometry source | Magnitude system and filters |
|---|---|---|
| SN 1981B | Buta & Turner (1983) | Standard *UBVR* |
| SN 1990N | Lira et al. (1998) | Standard *UBVRI* |
| SN 1994ae | Riess et al. (2005) | AndyCam[a] *BVRI* |
| SN 1998aq | Riess et al. (2005) | 4Shooter2/AndyCam *UBVRI* |
| SN 1995al | Riess et al. (2009a) | AndyCam *UBVRI* |
| SN 2002fk | CfA3[b] | 4Shooter2 *UBVRI* |
| | LOSS[c] | KAIT3/NICKEL *BVRI* |
| SN 2007af | CfA3 | Keplercam *BVri* |
| | LOSS | KAIT3/KAIT4 *BVRI* |
| SN 2007sr | CfA3 | Keplercam *BVri* |
| | LOSS | KAIT3/4 *BVRI* |

[a] A thin, back-illuminated CCD camera on the FLWO 1.2 m telescope (Jha et al., 2006).
[b] Hicken et al. (2009a).
[c] Ganeshalingam et al. (2010). Both CfA3 and LOSS photometry were available for the most recent three SNe Ia , so we used combined photometry from both sources.

quantities are then used to derive distances (Equation 4.8), following equations for apparent magnitudes in Section 4.3.2, hence populating a Hubble diagram.

## 4.3 Measuring the Hubble constant

In our determination of $H_0$ we rely on two standard candles: type Ia supernovae and Cepheid variables. These together prescribe a relative distance scale for the low-$z$ SNe Ia. The absolute calibration is given by the geometric maser distance of NGC 4258 from Humphreys et al. (2013). The Cepheid variables lie in this galaxy and eight other galaxies containing nearby SNe Ia, calibrating the supernovae.

In this section we discuss the specific details and equations for measuring the Hubble constant locally from a distance ladder of SNe Ia calibrated by Cepheid variables. We start with the cosmological equations for measuring $H_0$ from observations (Section 4.3.1), followed by equations for apparent magnitudes for each standard candle (Section 4.3.2). We then present the set of simultaneous equations we fit for (Section 4.3.3) as well as our blinding method.

### 4.3.1 Extracting $H_0$

In its traditional formulation Hubble's law states that the recession velocity of objects is proportional to their distance:

$$v(z) = H_0 D(z) \tag{4.1}$$

where the constant of proportionality $H_0$ represents the present expansion rate of the Universe, scaled by its size (i.e. $H_0 = \frac{\dot{a}}{a}$ where $a$ is the scale factor and overdot indicates differentiation with respect to time, $t$). Methods of determining $H_0$ typically involve taking the ratio of the two sides of Equation 4.1. We expand on the subtleties of this below.

The distance in Hubble's Law is related to the luminosity distance by

$$D(\bar{z}) = \frac{1}{1+z_h} D_L(\bar{z}, z_h), \tag{4.2}$$

where $\bar{z}$ and $z_h$ are the cosmological and heliocentric redshifts respectively. Equation 4.2 assumes a flat Universe; we generalise to other geometries in Equation 6.13 in Section 6.2.

The difference resulting from the same redshift $z = \bar{z}$ for both is negligible so we do not differentiate in our analysis. Calcino & Davis (2017, Section 4.2–4.3) quantify the effect of possible redshift systematic errors on the derivation of $H_0$ and find that a systematic redshift error as small as $\sim 2.6 \times 10^{-4}$ can result in a $\sim$0.3% bias in $H_0$. This is further discussed in Section 6.2, where both redshifts are included.

The luminosity distance $D_L(z)$ can be determined observationally (i.e. with no knowledge of cosmological parameters) using standard candles. These have known absolute magnitudes $M$, so taking the difference between $M$ and the apparent magnitude $m$ gives the distance modulus $\mu \equiv m - M$ and hence the luminosity distance $D_L \equiv 10^{\frac{\mu-25}{5}}$ Mpc. In practice the process of measuring distances is far from straightforward, and is outlined in Section 2.2.

On the left hand side, $v(z)$ is the predicted velocity due to expansion for a galaxy at redshift $z$. The exact expression for $v(z)$ is given by integrating the Universe's expansion up to redshift $z$:

$$v(z) = c \int_0^z \frac{dz'}{E(z')}, \tag{4.3}$$

where $E(z) \equiv H(z)/H_0$ is a function of cosmological parameters (expanded upon in Section 6.2 and defined in Peebles (1993)), and $v(z)$ is independent of $H_0$.[2]

At low redshifts the cosmological dependence of $v(z)$ is very weak and it is a good approximation to use a second order Taylor expansion in terms of the deceleration and jerk parameters $q_0$ and $j_0$[3]. Thus we follow R11 and use,

$$v(z) = \frac{cz}{1+z}\left[1 + \frac{1}{2}(1-q_0)z - \frac{1}{6}(1 - q_0 - 3q_0^2 + j_0)z^2\right]. \tag{4.4}$$

At low redshift Equations 4.3 and 4.4 both reduce to the familiar $v(z) \approx cz$. At moderate redshifts ($z < 0.1$), Equation 4.4 closely approximates most observationally reasonable cosmological models. We explored the uncertainty associated with assuming Equation 4.4 and the cosmology stated in Footnote 3, finding the impact to be small: varying either $\Omega_m$ or $w$ by 0.1 changes $H_0$ by 0.015 km s$^{-1}$ Mpc$^{-1}$ or 0.1 km s$^{-1}$ Mpc$^{-1}$, respectively, in the sense that an increase in $\Omega_m$ or $w$ causes an increase in $H_0$. The maximal difference in $\mathcal{M}$ induced by varying $q_0, j_0$ within values allowed by $1\sigma$ contours in Betoule et al. (2014) is an order of magnitude smaller than its statistical uncertainty.

Rearranging Equations 4.1, 4.2, and 4.4 gives the equation for $H_0$ as a function of observables, $z$ and $D_L$,[4]

$$\begin{aligned} H_0 &= \frac{v(z)(1+z)}{D_L(z)} \\ &= \frac{cz}{D_L(z)}\left[1 + \frac{1}{2}(1-q_0)z - \frac{1}{6}(1 - q_0 - 3q_0^2 + j_0)z^2\right]. \end{aligned} \tag{4.6}$$

[2]It is interesting to note that $v(z)$ is independent of $H_0$; it depends only on redshift and cosmological parameters such as $\Omega_m$ and $\Omega_\Lambda$. That may seem unintuitive, but it is velocity as a function of distance $v(D)$ that is function of $H_0$ (things that are moving faster have gone further). Velocity as a function of redshift $v(z)$ works differently since redshift is not proportional to distance. A galaxy's redshift is determined by how much the Universe has expanded since the light was emitted. That depends on the travel time, which does depend on the densities that cause the Universe to accelerate or decelerate (and thus for the light to take longer or shorter times to reach us), but not on $H_0$.

[3]We assume a standard $\Lambda$CDM cosmology with $\Omega_m \sim 0.3, \Omega_\Lambda \sim 0.7$, fixing $q_0 = -0.55$ and $j_0 = 1$.

[4]For non-zero curvature, Equation 4.6 becomes

$$H_0 = \frac{v(z)(1+z)}{D_L(z)} \frac{S_k(\chi)}{\chi}. \tag{4.5}$$

Table 4.2 Recent nearby SNe Ia and their host galaxies used in R11, along with observations of Cepheids in these galaxies.

| Galaxy | SN Ia | $N_{\text{Cepheids}}$ |
|---|---|---|
| NGC 4536 | SN 1981B | 69 |
| NGC 4639 | SN 1990N | 32 |
| NGC 3370 | SN 1994ae | 79 |
| NGC 3982 | SN 1998aq | 29 |
| NGC 3021 | SN 1995al | 26 |
| NGC 1309 | SN 2002fk | 36 |
| NGC 5584 | SN 2007af | 95 |
| NGC 4038 | SN 2007sr | 39 |
| NGC 4258 | - | 165 |
| Total | | 570 |

Thus determining $H_0$ amounts to comparing the velocity in Equation 4.4 – derived from the measured redshift – to the observed luminosity distance, measured with standard candles. The equations encapsulating this process are detailed in Sections 4.3.2 and 4.3.3.

### 4.3.2 Apparent magnitudes

**Cepheids magnitudes**

Our first data set, the Cepheid variables, allow us to infer distances to the nearby galaxies via the Leavitt law (also commonly known as the period-luminosity relation):

$$m_W = b_W(\log_{10} P - 1) + Z_W \Delta \log_{10}[O/H]_{ij} + M_W + \mu. \tag{4.7}$$

Equation 4.7 relates the apparent 'extinction-free' (Wesenheit) magnitude $m_W$,[5] period ($P$; in days), and metallicity of a Cepheid at distance modulus $\mu$. The slopes $b_W$ and $Z_W$ represent the dependence of the magnitude on period and metallicity; the zero point $M_W$ physically represents the Wesenheit magnitude of a Cepheid in our Galaxy (at a distance of 10 pc), with a period of 10 days. We use relative values of the metallicity ($\Delta \log_{10}[O/H]_{ij} := \log_{10}[O/H] - 8.9$) and period to pivot the fit near the data.

**SN Ia magnitudes**

Type Ia supernovae comprise our remaining data. A spectroscopically normal SN Ia has a lightcurve parametrized by its brightness (hence distance), observed colour and decline rate. These measures are represented by different quantities in various SN Ia frameworks; in SALT2 (Guy et al., 2007) these are the apparent magnitude $m_B$ at time of $B$-band maximum, 'stretch' $X_1$ and colour $C$ (roughly corresponding to $B - V$ at maximum), related by:

$$m_B = M_B - \alpha X_1 + \beta C + \mu \tag{4.8}$$

where $M_B$ is the canonical SN Ia absolute magnitude, and $\alpha, \beta$ are SALT2 nuisance parameters for the stretch and colour dependences.

SNe Ia in more massive galaxies are brighter after these standard corrections for colour and stretch, as discussed in Section 2.3.2. To account for this we replace $M_B$ in Equation 4.8 with the corrected absolute magnitude $M_B^*$ (defined in Equation 2.4), which can take two discrete values depending on the host galaxy mass: $M_B$ or $M_B + \Delta M_B$. We will fix $\Delta M_B$ and fit for the three global parameters $\{\alpha, \beta, M_B\}$; the reasoning for this choice will be detailed in Section 4.5.2.

[5] We use the quantity $M_W \equiv V - R_V(V - I)$ constructed in Madore (1982) from the Wesenheit function (van den Bergh, 1975), from $V$- and $I$-band absolute magnitudes. Assuming constant ratio $R_V$ of total to selective absorption, $M_W$ is independent of extinction. We fix $R_V = A_V/E(V - I) = 3.1$ as in R11.

Our second data set contains the eight 'nearby' SNe Ia in Table 4.2, with apparent magnitudes given by Equation 4.8 (with $M_B^*$ instead of $M_B$). A SN Ia and Cepheid in the same galaxy have common distance modulus $\mu$ in Equations 4.8 and Equation 4.7; thus, the Cepheids calibrate the nearby SNe, which in turn determine the SN Ia magnitude zero point $M_B$.

The much larger sample of 280 SNe Ia makes up our third data set. These 'low-$z$' supernovae originate from CfA3 and LOSS, with details in Section 4.2. Once we have calibrated their absolute magnitudes using the eight 'nearby' supernovae, we can use the theory derived in Section 4.3 to relate their measured magnitudes to the value of $H_0$. Assuming Equation 4.6 and writing $f(z) \equiv 1 + \frac{(1-q_0)z}{2} - \frac{(1-q_0-3q_0^2+j_0)z^2}{6}$, we have in place of Equation 4.8:

$$m_B = M_B^* - \alpha X_1 + \beta C + 5\log_{10}\left(\frac{czf(z)}{H_0}\right) + 25. \tag{4.9}$$

The way the magnitudes are combined to measure $H_0$ are then described in Section 4.3.3.

### 4.3.3 Simultaneous equations

Following Section 4.3.2, we present equations for fitting observed apparent magnitudes for all three data sets (Cepheids, SNe Ia in Cepheid hosts, and the Hubble flow SNe Ia) simultaneously. We first perform preliminary fits to only Equation 4.7 (Section 4.4) and to only Equation 4.12 (Section 4.2.2) to examine the dependence of nuisance parameters on choices such as distance anchor and rejection algorithm (for the Cepheid-only fit) and on SN Ia cuts (for the SN-only fit). Note that R11 performs these two steps to fix the Cepheid nuisance parameters $b_W$ and $Z_W$, and the SN Ia zero point $a_V$[6] from the Hubble flow SNe Ia; these quantities are then held fixed in fitting the nearby supernovae together with the Cepheids.

In our fit, all parameters are determined again, by fitting Equations 4.7, 4.8, 4.9 simultaneously for a combined fit to all Cepheid and SN Ia data. We rewrite these equations, making explicit the indexing: $i$ varies over the eight nearby galaxies (and the SNe Ia they contain), $j$ varies over Cepheids in these galaxies and NGC 4258, $k$ varies over the low-$z$ SNe.

$$m_{Wij} = b_W(\log_{10} P_{ij} - 1) + Z_W \Delta\log_{10}[O/H]_{ij} + M_W + \mu_{4258} + \Delta\mu_i \tag{4.10}$$
$$m_{Bi} = M_B^* - \alpha X_{1i} + \beta C_i + \mu_{4258} + \Delta\mu_i \tag{4.11}$$
$$m_{Bk} = M_B^* - \alpha X_{1k} + \beta C_k + 5\log_{10}(cz_k f(z_k)) - \mathcal{H}. \tag{4.12}$$

In Equation 4.12 we separate the intercept of Equation 4.9 into parameters $M_B^*$ (also appearing in Equation 4.11) and a constant term $\mathcal{H}$, which contains the same information as $H_0$:

$$\mathcal{H} := 5\log_{10} H_0 - 25. \tag{4.13}$$

We fit for all 16 parameters appearing in Equations 4.10–4.12; explicitly these are $\Theta = \{\alpha, \beta, \mathcal{H}, M_B, b_W, Z_W, M_W, \mu_{4258}, \Delta\mu_i\}$ where $i$ varies over the eight nearby galaxies. Note that we fit for $M_B$ instead of $M_B^*$ as the latter is not a constant. The distance moduli in Equations 4.10 and 4.11 are expressed as offsets $\Delta\mu_i \equiv \mu_i - \mu_{4258}$, relative to NGC 4258.

Equations 4.10–4.12 assume a distance anchor of NGC 4258. The use of the LMC and MW as alternate or additional anchors is explored, and discussed in Section 4.4.2. We impose a strong gaussian prior $\mu_{4258} = 29.404 \pm 0.066$ on the distance, measured from VLBI observations of megamasers in Humphreys et al. (2013)[7] whenever NGC 4258 is used as an anchor, and similarly $\mu_{\rm LMC} = 18.494 \pm 0.049$

[6] This is equivalent to $0.2\mathcal{M}$ in our analysis.

[7] This distance is slightly higher than the older value $\mu_{4258} = 29.31$ assumed in Riess et al. (2012); this increase acts to decrease $H_0$ relatively.

if the LMC is included.

### 4.3.4 Blinding parameters

Our priority is to hide the value of $H_0$ so as to not influence its result, so we blind the parameter $\mathcal{H}$ which contains equivalent information.[8] We also blind the SN Ia magnitude zero point $M_B$ which has the most interaction with $\mathcal{H}$, and is the best-constrained in the literature, relative to other parameters in $\Theta$. We implement these blinds in the analysis and data respectively. For any likelihood function containing $\mathcal{H}$ (i.e. involving the low-$z$ SNe) we make the shift $\mathcal{H} \mapsto \mathcal{H} + \mathfrak{o}_H$ for an offset $\mathfrak{o}_H$. Meanwhile we effectively shift $M_B$ by adding another offset $\mathfrak{o}_M$ to all SN magnitudes $m_B$. Both offsets $\mathfrak{o}_H$ and $\mathfrak{o}_M$ are unknown real numbers, randomly drawn from normal distributions and never printed. These are seeded by distinct known numbers to ensure that the offsets are constant and can be retrieved. Our method allows the recovery of the true unblinded values by simply subtracting the offsets once the blind is lifted.

We choose to not blind the other parameters which appear in the preliminary Cepheid- or SN-only fits, primarily because these parameters do not have strong enough priors from the literature to introduce human bias. Moreover, the variation we observe in the preliminary values of the nuisance parameters $\{b_W, Z_W, \alpha, \beta\}$ is useful for informing which preliminary fits to carry forward to the global fits. Knowing the preliminary nuisance parameters will not bias our results because 'best' versions of the preliminary fits are not chosen; instead we select a representative sample of these fits and use the scatter to quantify the systematic uncertainties.

## 4.4 Cepheid-only fit

Our fit to all Cepheid data is based on E14 with the difference that we do not assume the SN Ia zero point (the quantity $a_V$ in R11 and E14) or indeed any SNe data. This is because we intend to fit the Cepheids separately from the SN data, whereas E14 calculates values of $H_0$ from the Cepheid fits, assuming SN Ia data from R11. All Cepheid data are fitted to the Leavitt law (Equation 4.10) with MultiNest (Section 3.2.2). The 12 parameters of fit include the three nuisance parameters $\{b_W, Z_W, M_W\}$, the strongly constrained distance $\mu_{4258}$, and the eight distance modulus offsets $\{\Delta\mu_i\}$. We set an external gaussian prior on $\mu_{4258}$, and by default place uniform priors for all other parameters over generous intervals. The $\chi^2$ function for the Cepheid fit is a function of $\{b_W, Z_W, M_W, \mu_{4258}\}$, and $\{\Delta\mu_i\}$, and takes the form

$$\chi_c^2 = \sum_{ij} \frac{(\hat{m}_{Wij} - m_{Wij,\mathrm{mod}})^2}{\hat{m}^2_{Wij,\mathrm{err}} + \sigma^2_{\mathrm{int,C}}}. \tag{4.14}$$

Here $m_{Wij,\mathrm{mod}}(b_W, Z_W, M_W, \mu_{4258}, \{\Delta\mu_i\})$ is the model magnitude of the $j$-th Cepheid in galaxy $i$ (given by Equation 4.10) and $\sigma_{\mathrm{int,C}}$ is the intrinsic scatter in Cepheid magnitude, from the width of the instability strip. For clarity, measured quantities are denoted with hats to distinguish them from model quantities. The logarithm of the likelihood $\mathcal{L} = e^{-\chi_c^2/2}$ and the priors on the fit parameters are inputs for MultiNest. We use 1000 live points in MultiNest and confirm that the precision is sufficient.[9]

### 4.4.1 Results of Cepheid-only fit

The results of all Leavitt law fits, for all combinations of distance anchor, outlier rejection, and upper period limit, are presented in Table 4.3 (table D4). The details of these choices are given in Section 4.4.2, along with the effect they have on fit results. The variation in the fits is visualised in Figure 4.1 in $b_W, Z_W$-space. The choice of these two parameters is obvious as they characterise the Leavitt law and are solely influenced by the Cepheid sample – all other parameters in $\Theta$ are influenced by the SN data,

[8] We fit for the parameter $\mathcal{H} := 5\log_{10} H_0 - 25$, which is linear in magnitude (Equation 4.13), instead of $H_0$.

[9] For selected fits we repeat the outlier rejection and fitting steps, and find that the scatter in final parameters within ten runs is $< 1\%$ of the statistical uncertainty.

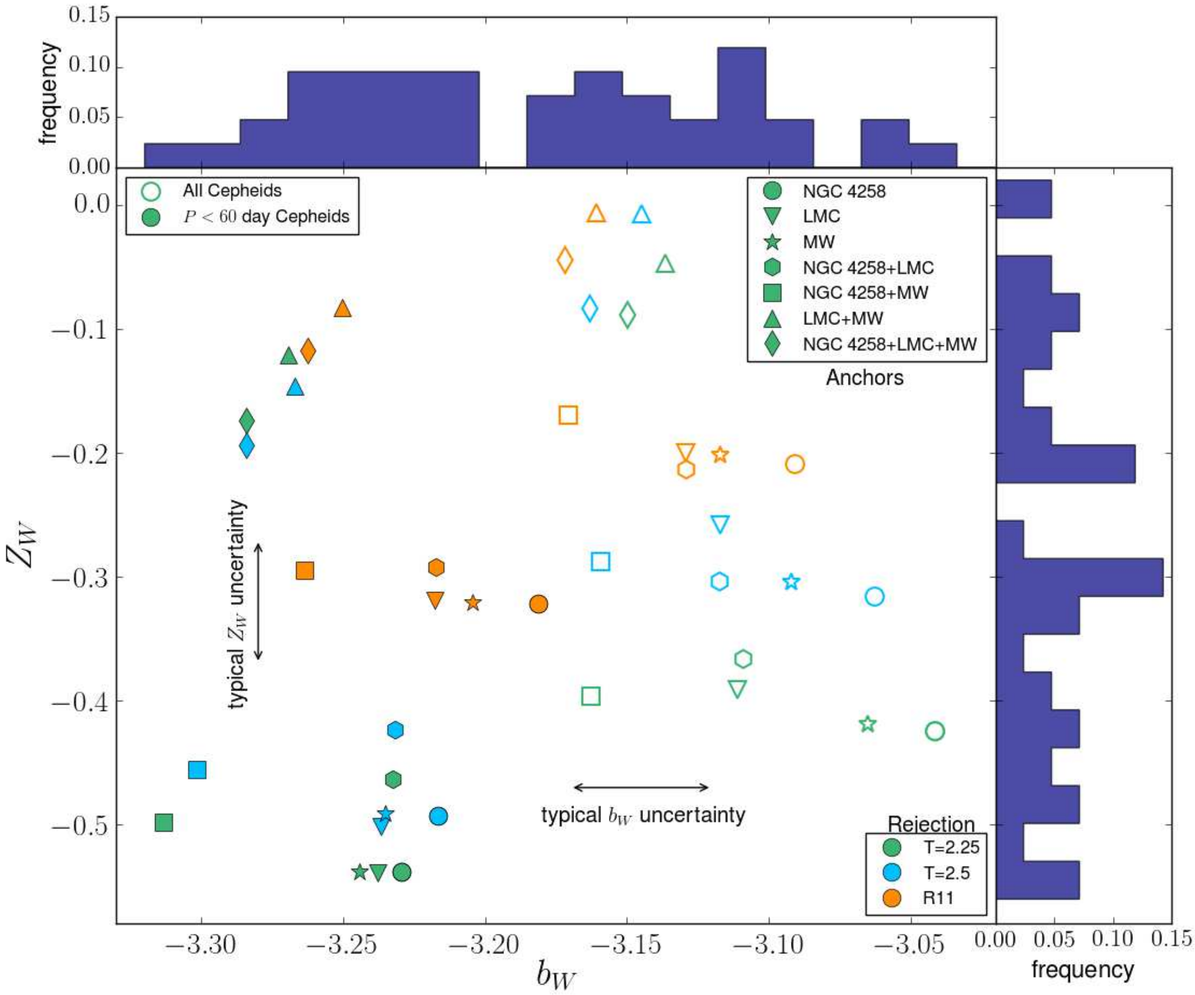


Figure 4.1 The best-fitting values for $b_W$, $Z_W$ from all Cepheid-only fits to the Leavitt law (Equation 4.10), assuming various distance anchors and rejection algorithms, with and without a cut on the period. The different markers represent these properties as indicated in the legends, with the colour representing the outlier rejection algorithm, shape representing the distance anchor, and solidness reflecting the period cut. We consider all seven combinations of distance anchor galaxies NGC 4258, LMC, and MW (Section 4.4.2), and all three rejection algorithms (Section 4.4.2). This figure shows: (i) including the longer-period Cepheids increases both $b_W$ and $Z_W$ (empty markers lie up and to the right of solid markers) (ii) Systematic variation in parameters with distance anchor (e.g. for each choice of period cut, the NGC 4258 + MW anchor gives the lowest $b_W$ and the NGC 4258-only anchor gives the highest); meanwhile fits with both the LMC and MW as anchors (diamonds and upward triangles, with and without NGC 4258 respectively) are clustered tightly, indicating that these two galaxies together provide a strong constraint on both parameters. (iii) The R11 rejection results in less negative $Z_W$ and to a lesser extent $b_W$ (reflected in orange markers concentrated in the upper-right portion of the figure), while the E14 algorithm with rejection threshold $T = 2.5$ (turquoise) results in higher $Z_W$ compared to $T = 2.25$ (green) for fits other than those with both the LMC and MW anchors. The typical uncertainties, indicated by the arrows, are $\sim$0.05 for $b_W$ and $\sim$0.1 for $Z_W$ for most fits, but can be larger for some anchors or rejection algorithms. Evidently the scatter arising from varying the distance anchor, cut on period, and rejection far exceeds the statistical uncertainty. The histograms in the margins display distributions of $b_W$ and $Z_W$ values over all fits. The histogram for $b_W$ shows that values are clustered around $b_W \sim -3.25$ for fits with a $P < 60$ day cut (reflective of the influence of the LMC Cepheids) and $b_W \sim -3.10$ for fits without. The histogram for $Z_W$ shows a spread centred at $Z_W \sim -0.3$, dependent on distance anchor; fits with both the LMC and MW anchors lie with $-0.2 < Z_W < 0$.

Table 4.3. Results of the Cepheid-only fits described in Section 4.4 from each combination of distance anchor, rejection and period cut for each Cepheid fit. The best-fitting Cepheid parameters $\{b_W, Z_W, M_W\}$ are given, as well as the number of Cepheids remaining after rejection and intrinsic scatter.

| Rejection | Distance anchor | $P < 60$d | $N_{\rm cepheids}$ | $\sigma_{\rm int,C}$ | $b_W$ | $Z_W$ | $M_W$ |
|---|---|---|---|---|---|---|---|
| $T = 2.25$ | n4258 | Y | 439 | 0.17 | -3.23 (0.07) | -0.54 (0.13) | -6.03 (0.07) |
| $T = 2.5$ | n4258 | Y | 463 | 0.27 | -3.22 (0.08) | -0.49 (0.14) | -6.06 (0.07) |
| R11 | n4258 | Y | 379 | 0.21 | -3.18 (0.07) | -0.32 (0.14) | -6.05 (0.08) |
| $T = 2.25$ | LMC | Y | 439 | 0.17 | -3.24 (0.05) | -0.54 (0.13) | -6.16 (0.07) |
| $T = 2.5$ | LMC | Y | 464 | 0.27 | -3.24 (0.05) | -0.50 (0.14) | -6.14 (0.08) |
| R11 | LMC | Y | 379 | 0.21 | -3.22 (0.05) | -0.32 (0.14) | -6.07 (0.07) |
| $T = 2.25$ | MW | Y | 439 | 0.17 | -3.24 (0.07) | -0.54 (0.13) | -5.83 (0.05) |
| $T = 2.5$ | MW | Y | 463 | 0.27 | -3.24 (0.07) | -0.49 (0.15) | -5.83 (0.05) |
| R11 | MW | Y | 379 | 0.21 | -3.20 (0.07) | -0.32 (0.15) | -5.82 (0.05) |
| $T = 2.25$ | n4258+LMC | Y | 439 | 0.18 | -3.23 (0.05) | -0.46 (0.11) | -6.10 (0.05) |
| $T = 2.5$ | n4258+LMC | Y | 466 | 0.28 | -3.23 (0.05) | -0.42 (0.12) | -6.10 (0.05) |
| R11 | n4258+LMC | Y | 379 | 0.21 | -3.22 (0.05) | -0.29 (0.12) | -6.05 (0.05) |
| $T = 2.25$ | n4258+MW | Y | 437 | 0.17 | -3.31 (0.06) | -0.50 (0.12) | -5.89 (0.04) |
| $T = 2.5$ | n4258+MW | Y | 464 | 0.27 | -3.30 (0.07) | -0.46 (0.14) | -5.90 (0.04) |
| R11 | n4258+MW | Y | 379 | 0.21 | -3.26 (0.06) | -0.30 (0.14) | -5.89 (0.04) |
| $T = 2.25$ | LMC+MW | Y | 435 | 0.16 | -3.27 (0.05) | -0.12 (0.10) | -5.91 (0.05) |
| $T = 2.5$ | LMC+MW | Y | 464 | 0.28 | -3.27 (0.05) | -0.15 (0.11) | -5.92 (0.06) |
| R11 | LMC+MW | Y | 379 | 0.21 | -3.25 (0.05) | -0.08 (0.11) | -5.90 (0.06) |
| $T = 2.25$ | n4258+LMC+MW | Y | 434 | 0.16 | -3.28 (0.05) | -0.17 (0.10) | -5.95 (0.04) |
| $T = 2.5$ | n4258+LMC+MW | Y | 463 | 0.27 | -3.28 (0.05) | -0.19 (0.10) | -5.96 (0.04) |
| R11 | n4258+LMC+MW | Y | 379 | 0.21 | -3.26 (0.05) | -0.12 (0.11) | -5.94 (0.04) |
| $T = 2.25$ | n4258 | N | 521 | 0.2 | -3.04 (0.05) | -0.42 (0.12) | -6.10 (0.07) |
| $T = 2.5$ | n4258 | N | 540 | 0.26 | -3.06 (0.06) | -0.32 (0.13) | -6.11 (0.07) |
| R11 | n4258 | N | 444 | 0.21 | -3.09 (0.06) | -0.21 (0.13) | -6.08 (0.07) |
| $T = 2.25$ | LMC | N | 523 | 0.21 | -3.11 (0.04) | -0.39 (0.12) | -6.12 (0.07) |
| $T = 2.5$ | LMC | N | 544 | 0.28 | -3.12 (0.04) | -0.26 (0.13) | -6.06 (0.07) |
| R11 | LMC | N | 444 | 0.21 | -3.13 (0.04) | -0.20 (0.13) | -6.04 (0.07) |
| $T = 2.25$ | MW | N | 521 | 0.2 | -3.07 (0.05) | -0.42 (0.12) | -5.80 (0.05) |
| $T = 2.5$ | MW | N | 539 | 0.26 | -3.09 (0.05) | -0.30 (0.13) | -5.81 (0.05) |
| R11 | MW | N | 444 | 0.21 | -3.12 (0.06) | -0.20 (0.13) | -5.81 (0.05) |
| $T = 2.25$ | n4258+LMC | N | 523 | 0.21 | -3.11 (0.04) | -0.37 (0.11) | -6.09 (0.05) |
| $T = 2.5$ | n4258+LMC | N | 539 | 0.26 | -3.12 (0.04) | -0.30 (0.11) | -6.08 (0.05) |
| R11 | n4258+LMC | N | 444 | 0.21 | -3.13 (0.04) | -0.21 (0.12) | -6.05 (0.05) |
| $T = 2.25$ | n4258+MW | N | 520 | 0.2 | -3.16 (0.05) | -0.40 (0.12) | -5.89 (0.04) |
| $T = 2.5$ | n4258+MW | N | 538 | 0.26 | -3.16 (0.05) | -0.29 (0.13) | -5.90 (0.04) |
| R11 | n4258+MW | N | 444 | 0.21 | -3.17 (0.06) | -0.17 (0.13) | -5.89 (0.04) |
| $T = 2.25$ | LMC+MW | N | 519 | 0.2 | -3.14 (0.04) | -0.05 (0.10) | -5.89 (0.05) |
| $T = 2.5$ | LMC+MW | N | 546 | 0.29 | -3.14 (0.04) | -0.01 (0.10) | -5.89 (0.06) |
| R11 | LMC+MW | N | 444 | 0.21 | -3.16 (0.04) | -0.01 (0.10) | -5.88 (0.05) |
| $T = 2.25$ | n4258+LMC+MW | N | 517 | 0.19 | -3.15 (0.04) | -0.09 (0.09) | -5.95 (0.04) |
| $T = 2.5$ | n4258+LMC+MW | N | 544 | 0.28 | -3.16 (0.04) | -0.08 (0.10) | -5.96 (0.04) |
| R11 | n4258+LMC+MW | N | 444 | 0.21 | -3.17 (0.04) | -0.04 (0.10) | -5.94 (0.04) |

even the zero point $M_W$. Figure 4.1 allows us to identify which of the Cepheid fits lie at the edges of the parameter space. The resultant scatter observed in Figure 4.1 far exceeds the statistical uncertainties reported in Table 4.3. Therefore it is paramount that the systematic associated with varying the choices made in Sections 4.4.2 through 4.4.2 is propagated carefully through the entire analysis process.

The choice of whether or not to apply the upper period limit of $P < 60$ days has the most effect on the parameters, especially $b_W$. Figure 4.1 reveals clearly the impact of including the longer-period Cepheids on parameters $b_W$ and $Z_W$, most notably splitting Figure 4.1 down the middle vertically, i.e. by Leavitt law slope. Both parameters are smaller in magnitude by $\sim$0.1 when the longer-period Cepheids are included, indicating a weaker dependence of Cepheid magnitude on both period and metallicity. For the slope $b_W$, this difference dominates the statistical uncertainty and any other variation in $b_W$, whereas for $Z_W$ the resultant change from changing the period cut is comparable in size to the dependence on rejection algorithm, and the statistical uncertainty.

When the longer-period Cepheids are included, each of $b_W$ and $Z_W$ is better-constrained by the distance anchor, and the rejection algorithm, respectively: this is reflected in the vertical lines of empty markers with the same shape, and near-horizontal lines of markers with the same colour. That is, when the $P < 60$ day cut is applied, the fit results are more sensitive to the choice of rejection algorithm and distance anchor. However, even without the cut, there remain strong dependences of $Z_W$ on rejection, and of $b_W$ on distance anchor.

Within each choice of period cut, the slope $b_W$ varies systematically with distance anchor: the NGC 4258 + MW and NGC 4258 anchors result in the lowest and highest $b_W$ respectively, with results form the other anchor combinations lying in between. The fits with both the LMC and MW in the anchors (upward triangles and diamonds) have the least spread in both parameters. With the exception of these fits, the data suggest a reasonably strong metallicity dependence with $-0.5 < Z_W < -0.2$. As noted above, the results are sensitive to rejection algorithm, with the R11 rejection resulting in less negative values for $Z_W$ (and for $b_W$ with the $P < 60$ day cut), followed by the E14 rejection with $T = 2.5$ to a smaller extent.

We observe (Table 4.3) that there is little difference in values for $M_W$ between fits with and without the $P < 60$ day cut, with the difference decreasing to zero for fits anchored on both n4258 and the Milky Way. However we defer further comment on $M_W$ (as well as $\{\Delta\mu_i\}$) to the discussion of global fit results. As $M_W$ is a magnitude zero point, and the $\Delta\mu_i$ are affected by the nearby SNe, the values of these parameters have potential to be influenced by the SN Ia data, and are expected to change with their inclusion.

#### Results compared to R11 and E14

We compare our fits in Table 4.3 to equivalent results in R11 and E14: our fits with R11 rejection and no period cut are compared to bolded fits in R11, table 2, and we compare our fits with the $P < 60$ day cut to the results in E14, tables 2–4 without priors on $b_W$ and $Z_W$. Relative to R11, our $b_W$ values with LMC-only or MW-only anchors are slightly lower in magnitude ($\sim -3.12$ instead of $-3.19$ in R11, a $\sim 1\sigma$ difference). Moreover our fits with LMC + MW anchors result in a lesser metallicity dependence ($-0.2 < Z_W < -0.1$ instead of $Z_W \sim -0.3$ in E14) – however uncertainties in $Z_W$ in these E14 fits exceed 0.1, and our $Z_W$ values (without the period cut) are supported by R11. Aside from these differences our results are in good agreement with R11 and E14, lying well within ranges allowed by statistical uncertainties. We retain 444 Cepheids when adopting the rejection flagged in R11, table 2 (close to the minimum of 448 reported in table 4 of R11) and only 379 with the $P < 60$ day restriction. Applying the E14 rejection algorithm, our fits consistently result in lower numbers of remaining Cepheids by $10 - 20$, and consequently slightly lower $\sigma_{\mathrm{int,C}}$. It is worth noting that our methodology differs from E14 (and R11) in that we do not involve any SNe in the fit (omitting the third term in equation 14 of E14), whilst E14 includes the SN fit results by assuming a value of $a_V$ taken from R11. Presumably the complex ways of probing the multi-dimensional parameter space are leading to differences, albeit slight, between this work, R11, and E14, that cannot be easily reconciled. We believe a solution for the future is for authors

Table 4.4 Summary of selected Cepheid fits to carry forward to global fit (i.e. rejection, anchors, and period cut used). The positions of the best-fitting values for $b_W$ and $Z_W$ in the $b_W, Z_W$-plane (represented in Figure 4.1) are also given, as well as the symbols for these fits in Figure 4.1. The top half of the table (solid symbols) lists fits with the $P < 60$ day cut, whilst the bottom half (empty symbols) contains fits without.

| | Rej ($T$) | Anchor[a] | Symbol |
|---|---|---|---|
| top left | 2.25 | n4258+LMC+MW | solid green diamond |
| top left | 2.5 | n4258+LMC+MW | solid turquoise diamond |
| top left | R11 | n4258+LMC+MW | solid orange diamond |
| top left | R11 | LMC+MW | solid orange $\Delta$ |
| middle | 2.5 | n4258+LMC | solid turquoise hexagon |
| middle | R11 | MW | solid orange star |
| lower left | 2.25 | n4258+MW | solid green square |
| lower | 2.25 | LMC | solid green $\nabla$ |
| lower | 2.25 | n4258 | solid green circle |
| top | 2.25 | n4258+LMC+MW | empty green diamond |
| top | 2.5 | n4258+LMC+MW | empty turquoise diamond |
| top | R11 | n4258+LMC+MW | empty orange diamond |
| top | R11 | LMC+MW | empty orange $\Delta$ |
| top | 2.25 | LMC+MW | empty green $\Delta$ |
| middle | 2.25 | n4258+MW | empty green square |
| right | R11 | n4258 | empty orange circle |
| right | 2.5 | n4258 | empty turquoise circle |
| lower right | 2.25 | n4258 | empty green circle |

[a] For typographic ease we abbreviate ‘NGC’ to ‘n’.

to provide code and data sets used for calculations as part of publication, that can be used to better understand differences.

**Selection for global fits**

The choice of Cepheid fits to carry forward to the global fit is informed by their results, i.e. Leavitt law slope $b_W$ and metallicity dependence $Z_W$, as these parameters are only influenced by the Cepheid sample and are very minimally affected by the SN data. We are interested in the effect the choice of Cepheid sample (through varying aspects of the fit such as distance anchor, rejection, and upper period limit) has on these parameters in the global fit. In particular it is essential to quantify the systematic uncertainty in $\mathcal{H}$ with varying these choices.

We select 18 fits in total, summarised in Table 4.4. To span the full range of uncertainty induced by various Cepheid fits, we select fits at extremes of the parameter space (Figure 4.1), with a selection of anchors and rejection algorithms. The combination of all three distance anchors has the most constraining power, so we include all of these fits to quantify the uncertainty within them.

Each fit has an associated set of best-fitting parameters with uncertainties, as well as (unless using the R11 rejection) the values of the intrinsic scatter and rejection threshold, which together uniquely define a set of Cepheids remaining after outlier rejection. These then make up the Cepheid data and priors for some parameters in $\Theta$, going into the global fit (Section 4.6).

### 4.4.2 Dependence of Cepheid-only fit

**Outlier rejection**

We perform fits in two ways: either assuming the outlier rejection in Riess et al. (2011), or following the rejection method in Efstathiou (2014). The R11 algorithm rejects Cepheids from each galaxy (rather than from the global fit), based on their deviation from the best Leavitt law fit. This rejection does not take into account the size of the Cepheid uncertainties, so that points with small residuals but large uncertainties are selectively accepted (E14, section 3.1). Consequently a large fraction of the total number of Cepheids is rejected, including a set of subluminous low-metallicity Cepheids (later corrected in R16, as discussed in Section 4.4.2). Moreover, the intrinsic scatter is overestimated, resulting in a low reduced $\chi^2$.

These limitations in the R11 rejection are addressed in the upgraded algorithm in E14, which rejects a Cepheid from the global fit if its magnitude deviates from the global fit by more than the quantity $T\sqrt{m_{W,\mathrm{err}}^2 + \sigma_{\mathrm{int,C}}^2}$ for a threshold $T$ (set to 2.25 or 2.5), where $m_{W,\mathrm{err}}$ and $\sigma_{\mathrm{int,C}}$ are the uncertainty in the individual Cepheid's measurement and the intrinsic scatter $\sigma_{\mathrm{int,C}}$ respectively. This process is iterative, with $\sigma_{\mathrm{int,C}}$ recalculated at each step (such that $\chi_c^2$ per degree of freedom $\sim$1) with increments of 0.1, where the sum in $\chi_c^2$ is always over only the Cepheids in NGC 4258 and SN hosts (i.e. not the LMC or MW). The rejection at each iteration is based on the best fit determined in the previous iteration, i.e. the mean and $1\sigma$ uncertainty of the posterior distribution.

Initially $\sigma_{\mathrm{int,C}}$ is set to 0.3. Then in each iteration we perform the following steps:

1. perform a MultiNest fit to all remaining Cepheids, to find marginalised posterior distributions;
2. find and remove outliers based on these parameters;
3. compute the new $\sigma_{\mathrm{int,C}}$ for these parameters and the updated Cepheid sample.

These steps are repeated until convergence, i.e. until no Cepheids are rejected in the second step.

The variation in fit results from different outlier rejection is presented in Figure 4.1 and Section 4.4.1. In general the R11 rejection results in less negative values of both $b_W$ and $Z_W$, attributable to the aforementioned subluminous and low-metallicity subsample that it rejects.

The fit is forced to be good for all three rejection algorithms: $\sigma_{\mathrm{int,C}}$ is engineered to result in $\chi^2/\mathrm{DoF}\sim 1$. Thus the algorithms cannot be compared statistically; the outlier rejection method has the drawback of not allowing the uncertainty on $\sigma_{\mathrm{int,C}}$ to be estimated, and the related consequence that we (by construction) cannot assess goodness-of-fit. Alternative statistical methods used in recent SN Ia analysis can surpass these limitations, including Bayesian hierarchical models (March et al., 2011; Shariff et al., 2016), the alternate Bayesian framework in Rubin et al. (2015), and Approximate Bayesian Computation Jennings et al. (2016). Notably, these have been applied to determining $H_0$ from the R11 and R16 data sets in Cardona et al. (2017).

### Distance anchors

Our equations in Section 4.3 assume NGC 4258 is the only distance calibrator. We can generalise these equations to allow for combinations of the three anchor galaxies in R11, adding Cepheids in the LMC and MW (data described in Section 4.2). As these additional Cepheids do not have metallicity measurements, we adopt the mean values from E14 of $12 + \log_{10}[O/H]$ of 8.5 and 8.9 for LMC and MW Cepheids respectively. Here, we test the dependence of the Cepheid parameters on the distance anchor. For Cepheids in the LMC and MW the Leavitt law (Equation 4.7) takes the forms:

$$m_{W,\mathrm{LMC}j} = b_W(\log_{10} P_j - 1) - 0.4 Z_W + M_W + \mu_{\mathrm{LMC}} \tag{4.15}$$

$$m_{W,\mathrm{MW}j} = b_W(\log_{10} P_j - 1) + M_W. \tag{4.16}$$

We consider combinations of NGC 4258, LMC, and MW (seven in total) as distance calibrators. If NGC 4258 is not included, then no prior for $\mu_{4258}$ is imposed in MultiNest. However, the likelihood $\mathcal{L}$ still depends on $\mu_{4258}$, which is indirectly constrained through the other anchors and $M_W$, and hence remains a fit parameter in $\Theta$. If the LMC is used as an anchor then it is necessary to include $\mu_{\mathrm{LMC}}$ as a parameter in $\Theta$; this always has a (gaussian) prior set to reflect the Pietrzyński et al. (2013) measurement from eclipsing binaries of $\mu_{\mathrm{LMC}} = 18.494 \pm 0.049$. The likelihood $\mathcal{L}$ is affected as a term $\chi^2_{\mathrm{LMC}}$ (Equation 4.17) is added to Equation 4.14; similarly if the MW is used as an anchor then $\chi^2_{\mathrm{MW}}$ (Equation 4.18) is added.

We assume $\sigma_{\rm int,C} = 0.113$ and 0.1 for the LMC and MW respectively following E14.

$$\chi^2_{\rm LMC} = \sum_j \frac{(\hat{m}_{W,{\rm LMC}j} - m_{W,{\rm LMC}j,{\rm mod}})^2}{\hat{m}^2_{{\rm LMC,err}j} + \sigma^2_{\rm int,C}} \tag{4.17}$$

$$\chi^2_{\rm MW} = \sum_j \frac{(\hat{m}_{W,{\rm MW}j} - m_{W,{\rm MW}j,{\rm mod}})^2}{\hat{m}^2_{{\rm MW,err}j} + \sigma^2_{\rm int,C}}. \tag{4.18}$$

A modification to the above is necessary if both the LMC and MW are used as distance anchors, to account for the calibration uncertainty between ground-based and HST photometry. We do this using the covariance matrix $\mathbf{C}_{{\rm LMC+MW}ij} = (\hat{m}^2_{Wi} + \sigma^2_{\rm int,C})\delta_{ij} + \sigma^2_{\rm cal}$ with $\sigma_{\rm cal} = 0.04$ (R11). Instead of $\chi^2_{\rm LMC} + \chi^2_{\rm MW}$, we add the term

$$\begin{aligned}\chi^2_{\rm LMC+MW} =& (\hat{\boldsymbol{m}}_{\boldsymbol{W},\mathbf{MW}} - \boldsymbol{m}_{\boldsymbol{W},\mathbf{MW},\mathbf{mod}}) \cdot \mathbf{C_{LMC+MW}}^{-1} \cdot \\ & (\hat{\boldsymbol{m}}_{\boldsymbol{W},\mathbf{MW}} - \boldsymbol{m}_{\boldsymbol{W},\mathbf{MW},\mathbf{mod}})^T\end{aligned} \tag{4.19}$$

to the $\chi^2_c$ term going into $\mathcal{L}$. Here, bolded quantities represent vectors over all LMC and MW Cepheids.

The results of varying the distance anchor are discussed in Section 4.4.1. Briefly, the inclusion of both the LMC and MW anchors constrains both $b_W$ and $Z_W$ more tightly.

**Longer-period cepheids**

The data include Cepheids with period ranging from ∼10 days, to over 200 days, and Cepheids of all periods are included in Leavitt law fits in R11 (except for those Cepheids marked 'low P' in R11, table 2). Bird et al. (2009) examine longer-period ($P > 80$ day) Cepheids and find that these Cepheids obey a flatter Leavitt law, with a less negative period dependence $b_W$. Accordingly, recent studies of the Leavitt law (e.g. Freedman et al., 2011; Scowcroft et al., 2011) have excluded Cepheids with period greater than 60 days. Similarly, E14 in their reanalysis of the R11 data have imposed the same upper limit on Cepheid period because of the observed change in slope. It is pragmatic to follow these examples in only using Cepheids over a period range where the slope remains constant. However, it is also useful to include the full range of periods to accommodate the change in slope and for the sake of comparison with R11. Thus, rather than making an argument to include the $P < 60$ day cut or not, we perform fits with and without an upper limit on the period and examine the results of both.

E14 reasons that while including longer-period Cepheids decreases the magnitude of the Leavitt law slope $b_W$, there is little impact on $H_0$ (Efstathiou, 2014, appendix A), so they only include $P < 60$ day Cepheids in their fits. Our priors on $b_W$ differ slightly from E14 (discussed in Section 4.4.2), and we are interested in the variation of Cepheid parameters with the choice of period cut (as with distance anchor and rejection algorithm in previous sections), so we test the dependence of fit results on the inclusion of an upper limit on period. Results of including longer-period Cepheids are lesser dependence on Cepheid slope and metallicity dependence (less negative $b_W$ and $Z_W$), as described in Section 4.4.1.

**Slope and metallicity priors**

We test and discuss the gaussian priors on $b_W$ and $Z_W$ described in E14 (but not mentioned in R11), and explain our choices for our fits. E14 performs Cepheid fits with and without gaussian priors centred at $b_W = -3.23$ and $Z_W = 0$, motivated by expectations of what the slope and metallicity dependence should be. We test the same priors in our fits but ultimately decide to not include these different priors as one of the variables in our fit, for reasons which follow.

Out of all the Cepheid data, the LMC Cepheids constrain the slope $b_W$ most tightly. Given the relative paucity of data on the Leavitt law, we always include this information on the Leavitt law in all our fits, independent of whether the LMC is used as a distance anchor. For the fits where the LMC is not included as an anchor, we impose the same gaussian prior on the slope as in E14: $\langle b_W \rangle = -3.23$, $\sigma_{b_W} = 0.10$. If the LMC is used as an anchor, there is already a contribution to the likelihood from these Cepheids, so it is inappropriate to reuse this information as a prior. Then, the inclusion of the prior on $b_W$ is prescribed by the distance anchor.

The metallicity priors in E14 are motivated by the observed strong dependence of the Cepheids' period on metallicity, in contrast with expectations that $Z_W{\sim}0$, based on theoretical considerations and measurements in the LMC (Freedman & Madore, 2011; Efstathiou, 2014, section 3.2, and references therein). However, the R11 sample of Cepheids demonstrates a strong metallicity dependence, with values of $Z_W$ around $-0.3$ or $-0.5$ for the R11 and E14 rejection algorithms respectively. The difference between values for $Z_W$ from the two approaches to outlier rejection (detailed in Section 4.4.2) can be traced to a sample of low metallicity Cepheids that are rejected by cuts in R11 but not E14. This systematic difference (discussed in E14, section 3.2) arose from the erroneous extrapolation of metallicity gradients to large radii, and was later corrected in R16. Including both the LMC and MW as distance anchors reduces the magnitude of the metallicity dependence $Z_W$. As we have observed that the R11 Cepheid data do not support the $Z_W{\sim}0$ priors (weak or strong) in E14, it is most appropriate to exclude these gaussian priors in our analysis.

## 4.5 Supernovae

In this section, we put the low-$z$ supernova on a Hubble diagram. First, we discuss the corrections to SN Ia observables, and the supernova systematics to consider. Then we perform a preliminary fit of only these supernovae to Equation 4.12, in analogy with the Cepheid-only fit in Section 4.4, to identify the dependence of the SN parameters on the different cuts in Section 4.5.1.

We refer the reader to Chapter 2 for a broader view of SNe Ia in cosmology, particularly Section 2.2 for a description of using supernovae (and other standard candles) for distance measurements. The treatment of SN Ia systematics follows the covariance matrix methods in SNLS and JLA, discussed in Section 3.4, outlined and with specifics discussed in Section 4.5.3. The data used here were described in Section 4.2.2, and the equations for SN Ia apparent magnitudes in Section 4.3.2 and now expanded.

The lightcurve fitting process is as described in Section 2.2.2. In contrast to R11, we choose to only use SALT2; one reason for this is that our framework for assessing SN Ia systematic uncertainties with covariance matrices (Sections 3.4 and Section 4.5.3) follows that in the SNLS-SDSS Joint Lightcurve Analysis (hereafter JLA; Betoule et al., 2014), which relies on the SALT2 model. In addition, SALT2 is the most modern fitter and used ubiquitously in cosmology analyses; thus our use allows for easier comparison and greater consistency. While R11 test the effects of fitting lightcurves with both SALT2 and MLCS2k2 (Jha et al., 2007) lightcurve fitters (both discussed in Section 2.2.3), we use SALT2 only. This is justified, as the latest version SALT2.4 (described in Betoule et al., 2014) was released in parallel with simulations in Mosher et al. (2014) which assess and quantify the uncertainty associated with the choice of lightcurve fitter (and the lightcurve model itself) in covariance matrices (Section 3.4). Hence it is unnecessary to use of multiple fitters to assess the the aforementioned systematic uncertainty.

### 4.5.1 Cuts

We make quality cuts on our SN Ia sample to eliminate potential biases from poorly constrained lightcurves and peculiar events, and to remain within the bounds of the SALT2 model. With the intent of replicating the sample in R11 as closely as possible, we broadly follow the cuts described in CfA3 (Hicken et al., 2009b) and LOSS (Ganeshalingam et al., 2013), also using cuts in SNLS and JLA – described in Guy et al. (2010, section 4.5), Conley et al. (2011, section 2.1), Betoule et al. (2014, section 4.5) – as guidance or as alternate cuts. In summary our criteria are as follows:

- low Milky Way extinction $E(B-V) < 0.2$

- exclude local SNe Ia not in the Hubble flow $z > 0.01$
- goodness-of-fit from SALT2 $\chi^2/\mathrm{DoF} < 8$
- first detection by +5 days, relative to $B$-band maximum
- exclude stretch outliers $|X_1| < 3$
- exclude colour outliers $|C| < 0.5$
- well-constrained stretch $\sigma_{X_1} < 0.8$
- well-constrained colour $\sigma_C < 0.1$.

The above encompass cuts in CfA3 and LOSS, with stricter cuts on the date of first detection and lightcurve goodness-of-fit (originally at +10 days and $\chi^2/\mathrm{DoF} = 15$ in CfA3 respectively), and with additional cuts on the uncertainties in $X_1$ and $C$ to further exclude supernovae which have large uncertainties in their stretch or colour. Our cuts are also informed by visual inspection of individual lightcurves and their SALT2 fits, particularly in placing boundaries for the lightcurve goodness-of-fit, uncertainties in stretch and colour, and date of first detection. In summary, we exclude supernovae at very low redshift (i.e. not yet in the Hubble flow), significantly extinguished by Milky Way dust, detected too late, with poorly constrained stretch and colour. We also exclude SNe Ia with poor SALT2 fits, and SNe that are too blue or red or have very fast or slow decline to exclude peculiar objects and ensure our sample fit within the SALT2 model.

Furthermore we test some alternate cuts, including some suggested in JLA and original CfA3/LOSS cuts which we have changed above. We repeat the SN-only fit with these cuts to test the effect on the SN fit parameters, carrying some through to the global fit. In particular, we follow R11 in raising the low-redshift cut to $z = 0.0233$,[10] and test strengthening or relaxing the lightcurve goodness-of-fit threshold to $\chi^2/\mathrm{DoF} < 5$ or $\chi^2/\mathrm{DoF} < 15$, and relaxing the date of first detection to +10 days. Following JLA we examine the effects of imposing a stricter bound on the colour ($|C| < 0.3$), the uncertainty on the stretch ($\sigma_{X_1} < 0.5$), and Milky Way extinction. These tests are important as the influence of these alternate cuts on the fit results is not straightforward or obvious; moreover no particular cut is necessarily more valid than the others. We discuss these results and their significance in Section 4.5.5. Histograms showing $X_1$ and $C$ distributions for several cuts are included in appendix B5 of Z17; figure B4 therein, for the $z > 0.0233$ cut, is included here as Figure 4.4.

### 4.5.2 Correction terms

We correct the SN Ia absolute magnitudes (or equivalently distance moduli) for host galaxy mass and Malmquist bias, according to Sections 2.3.2 and 3.1.2 respectively, using the bias correction function for the low-$z$ supernovae in JLA for the latter. Meanwhile the CMB-frame redshifts are corrected for peculiar velocities, as described in Section 3.1.3.

For making the host mass correction in Section 2.3.2, we adopt a fixed value of $\Delta M_B = -0.08$ for the magnitude offset between low- and high-mass host galaxies, based on values outlined in Section 2.3.2, from Sullivan et al. (2010); Lampeitl et al. (2010); Childress et al. (2013). We had also considered fitting for the magnitude offset using our data in the SN-only fit: using only our Hubble flow SNe Ia, we found a larger absolute offset of $\Delta M_B = -0.15 \pm 0.07$, with a large uncertainty and some degeneracy with $\mathcal{M}$. Our choice to use the fixed external value is informed by the tightness of constraints on $\Delta M_B$ from external data sets, and consistency between several independent samples, as discussed in Section 2.3.2. Our estimate of $\Delta M_B$ was intended as an observation of the constraining power of our dataset for this parameter, and was made without taking into account many of the subtleties in the above studies. Finally, while our uncertainty is much larger, our value is ultimately consistent with the literature. Thus, the decision to use the reference value is both justified and prudent.

[10] This is to reduce possible bias from local coherent flows, or a possible local underdensity (a so-called 'Hubble bubble'). The latter is discussed in R11 and Conley et al. (2011); however there is no conclusive evidence for its existence.

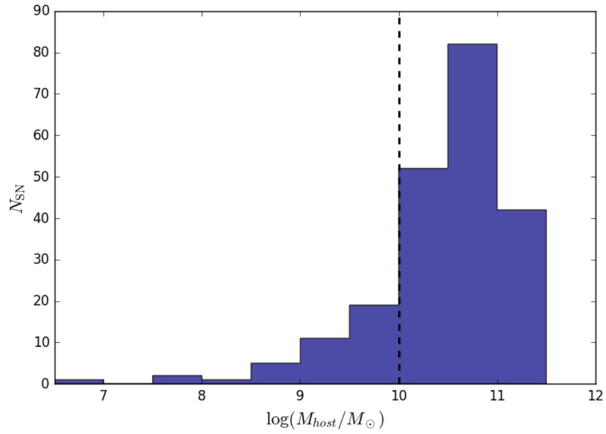


Figure 4.2 Histogram of host galaxy masses of low-$z$ SNe Ia where available (total of 220 SNe). The dashed line indicates the boundary $10^{10}M_\odot$ which splits the absolute SN Ia magnitude (Equation 2.4).

The host galaxy stellar masses are obtained from the literature where available. The eight nearby galaxies all have stellar masses given in Neill et al. (2009). Out of the 280 low-$z$ SNe in the Hubble flow, we have mass estimates for 220 of them, with 168 directly from literature and 72 estimated from photometry. Of the 168 estimates, 77 are from JLA and 71 from a combination of Sako et al. (2014); Childress et al. (2013); Neill et al. (2009); Kelly et al. (2010). We were able to derive mass estimates for 72 of the remaining galaxies using SDSS photometry, following standard methods in Sullivan et al. (2006); Smith et al. (2012).

Despite our efforts, there remain 60 low-$z$ supernovae where neither masses nor photometry are available. We assign these Hubble flow supernovae to the high-mass bin, with a large error which spans the $10^{10}M_\odot$ boundary, following C11 and B14. Our choice of the high-mass bin differs from the choice of the low-mass bin for host galaxies with unknown masses in JLA (B14, section 5.2), explained by differences in the samples. The JLA SNe without identified host masses were discovered in magnitude-limited surveys SNLS and SDSS, and were considered hostless in high-redshift deep imaging stacks, justifying their assignment to the low-mass bin. In contrast, the CfA3 and LOSS SNe Ia in our sample predominantly exist in more massive galaxies, as a consequence of having been discovered by targeted searches in known galaxy clusters – clearly supported by Figure 4.2, which shows that 181 out of the 220 SNe with known masses lie in the high-mass bin. Thus, it is reasonable to assign the remaining 60 SNe to this bin with a large uncertainty.

### 4.5.3 Supernova systematics

The supernovae in our analysis are subject to numerous systematic uncertainties discussed in Chapter 2. Section 3.4 describes our approach to accounting for SN Ia uncertainties follows methods in JLA, which are largely based on those in the Supernova Legacy Survey (hereafter SNLS; Conley et al., 2011). These use individual covariance matrices for each systematic, tracking correlated uncertainties between different SN quantities (i.e. $m_B, X_1, C$), between different supernovae. Advantages of the covariance matrix method over the more traditional method of adding systematics in quadrature are discussed in Conley et al. (2011, section 4); these include the ability to fully capture correlations in uncertainties, and the ease of including or reproducing uncertainties in further analyses.

In Section 3.4, we introduced the total covariance matrix $\mathbf{C}_\eta$ of correlated uncertainties in SN Ia observables for a sample as the sum of matrices, each representing a source of error:

$$\mathbf{C}_\eta = \mathbf{C}_{\rm stat} + \mathbf{C}_{\rm bias} + \mathbf{C}_{\rm cal} + \mathbf{C}_{\rm dust} + \mathbf{C}_{\rm host} + \mathbf{C}_{\rm model} + \mathbf{C}_{\rm pecvel} \tag{4.20}$$

where $\mathbf{C}_{\rm nonIa}$ is omitted for a low-redshift sample (in JLA, the portion of this matrix for the low-$z$ SNe Ia was identically zero). The matrix $\mathbf{C}_\eta$ in converted into $\mathbf{C}_{m_B^\dagger}$ according to Equation 4.24, which fits into the likelihood. In the following, we discuss specifics of calculating the individual covariance matrix terms, particularly deviations or additions to Section 3.4, then provide specific details of calculating the calibration uncertainties, then explain our approach to calculating the total covariance matrix for the nearby supernovae (i.e. those in galaxies with Cepheid measurements).

The diagonal terms in Equation 4.24, due to uncertainties in individual supernovae due to peculiar velocity scatter, weak lensing magnification, and intrinsic variation, were introduced at the start of Section 3.4. For these terms, we adopt the same values as in C11 and B14, of $c\sigma_z = 150\ \mathrm{km\,s^{-1}}$, $\sigma_{\rm lens} = 0.055z$, and $\sigma_{\rm int,SN} = 0.12$. In computing covariance matrices for the statistical errors $\mathbf{C}_{\rm stat}$, and the systematic errors from uncertainties in lightcurve model ($\mathbf{C}_{\rm model}$) and in the host galaxy dependence correction ($\mathbf{C}_{\rm host}$) are computed exactly according to descriptions in Sections 3.4.1, 3.4.7, and 3.4.3 respectively. For some other terms, we follow Section 3.4 closely with minor changes or comments.

**Dust, peculiar velocities, Malmquist bias**

The computation of the covariance matrix $\mathbf{C}_{\rm dust}$ associated with uncertainties in the dust extinction values follows Section 3.4.5 closely, using dust maps from Schlegel et al. (1998), and assuming a relative uncertainty of 20% (following B14, and increased from the 10% uncertainty in C11).

The uncertainty in peculiar velocity corrections, in $\mathbf{C}_{\rm dust}$, is computed according to the method described in Section 3.4.4. For translating errors in redshift into magnitude, we use the older equation:

$$\sigma_{\mu,\rm pecvel} = \frac{5\sigma_z}{z\log(10)}. \tag{4.21}$$

The uncertainty in the bias correction, in $\mathbf{C}_{\rm bias}$, follows Section 3.4.2, using the bias polynomial and its error were taken directly from Betoule et al. (2014). We adopt the estimate of the bias (for low-$z$ supernovae) in Betoule et al. (2014, section 5.3), which adopts a magnitude limited selection function, and uses the difference between the resultant bias and an unbiased regime as the uncertainty in the correction (the covariance matrix $\mathbf{C}_{\rm bias}$ in Section 3.4).

Some details of computing each of $\mathbf{C}_{\rm dust}$, $\mathbf{C}_{\rm pecvel}$, and $\mathbf{C}_{\rm bias}$ are later revised in Chapter 6 to reflect our beliefs about the best ways to estimate these quantities; updated calculations are described later when they apply. The methods in this section were applied in the spirit of following JLA exactly; some of these have evolved since.

**Calibration**

The calculation of calibration uncertainties which make up $\mathbf{C}_{\rm cal}$ were broadly described in Section 3.4.6. The vector $\kappa$ of systematic terms associated with calibration are enumerated in Table 4.5. The instruments and passbands to consider in $\kappa$ are those used for observing the low-$z$ SNe: 4Shooter2 and Keplercam for CfA3, and KAIT1–4 and Nickel for LOSS, and those involved in the training of SALT2 (i.e. used to observe the SNe in the training sample). The latter, and sizes of systematics in these passbands, are given in B14, table 5. It is necessary to include these training instruments and passbands as they influence measured magnitudes of training SNe hence the SALT2 model.

To find $\mathbf{C}_\kappa$ we start with the same matrix from JLA and reindex it according to Table 4.5, appending the LOSS instruments and removing HST instruments NICMOS and ACSWF (which do not contribute to the SALT2 training sample). We approximate the elements of $\mathbf{C}_\kappa$ for LOSS instruments as diagonal: this is exact for the $\lambda_{\rm eff}$ elements, and a good approximation for the zero point. Using Ganeshalingam et al. (2010) as a guide, we take the zero point and $\lambda_{\rm eff}$ uncertainties to be 0.03 mag and 10 Å respectively. As LOSS observations (with KAIT1–4 and Nickel) of SNe Ia were not used for SALT2 training, only SNe in the sample with LOSS measurements have nonzero partial derivatives with respect to these instruments.

Table 4.5 Systematics in $\kappa$

| Instrument | Filters | ZP index | $\lambda_{\rm eff}$ index |
|---|---|---|---|
| MegaCam | *griz* | 0–3 | 50–53 |
| Standard | *UBVRI* | 4–8 | 54–58 |
| KeplerCam | *UsBVri*[a] | 9–13 | 59–63 |
| 4Shooter2 | *UsBVRI* | 14–18 | 64–68 |
| Swope | *ugriBV* | 19–24 | 69–74 |
| SDSS | *ugriz* | 25–29 | 75–79 |
| KAIT1–4 | *BVRI* | 30–45 | 80–95 |
| NICKEL | *BVRI* | 46–49 | 96–99 |

[a] *Us* indicates the standard Landolt *U* passband, derived from Bessell (1990).

**Nearby SNe**

For the nearby SNe (i.e. those in galaxies containing Cepheids), we only include $\mathbf{C}_{\rm diag}$ and $\mathbf{C}_{\rm host}$. The host mass correction has the potential to shift the magnitude scale by up to $\sim$0.08 mag, and is important in the context of the dependence of SN Ia magnitude on host galaxy properties (Section 2.3.2). The other correction terms, for Malmquist bias and peculiar velocities, are redshift-dependent effects hence irrelevant for this sample. The remaining covariance matrices are not tied to the corrections in Section 4.5.2, and are more precise than warranted, given the inhomogeneity and larger uncertainties in these data, so we neglect them.

### 4.5.4 SN-only fit

To clearly separate the data and model in Equation 4.12 we define the quantity $m_B^\dagger$ for the apparent SN magnitude corrected for stretch and colour:

$$\begin{aligned} m_B^\dagger &:= m_B + \alpha X_1 - \beta C, \\ \text{with} \quad \hat{m}_{B\rm mod}^\dagger &= 5\log_{10}(czf(z)) + M_B^* - \mathcal{H}. \end{aligned} \tag{4.22}$$

Explicitly the $\chi^2$ function for the low-$z$ SN fit is

$$\chi^2_{\rm SN} = (\boldsymbol{\hat{m}_B^\dagger} - \boldsymbol{m_{B\rm mod}^\dagger}) \cdot \mathbf{C}_{\mathbf{m_B^\dagger}}{}^{-1} \cdot (\boldsymbol{\hat{m}_B^\dagger} - \boldsymbol{m_{B\rm mod}^\dagger})^T \tag{4.23}$$

where the matrix $\mathbf{C}_{\mathbf{m_B^\dagger}}$ (dependent on $\alpha$ and $\beta$) is derived from covariances in all SN parameters $\{m_B, X_1, C\}$ in $\mathbf{C}_\eta$ (independent of $\alpha$ and $\beta$), via conjugation with the matrix $\mathbf{A}$:

$$\mathbf{C}_{m_B^\dagger} = \mathbf{A}\cdot\mathbf{C}_\eta\cdot\mathbf{A}^T + \mathrm{diag}\left(\frac{5\sigma_z}{z\log 10}\right)^2 + \mathrm{diag}(\sigma^2_{\rm lens}) + \mathrm{diag}(\sigma^2_{\rm int,SN}). \tag{4.24}$$

This is identical to $\mathbf{C}_\mu$, defined in Equation 3.15, used in Chapter 6 – except framed in terms of a corrected apparent magnitude, rather than a distance modulus.

It is evident from Equations 4.12 and 4.22 that the SN-only fit is degenerate: we cannot constrain both $M_B$ and $\mathcal{H}$ simultaneously; the nearby SNe are necessary to constrain $M_B$. Instead we fit for the difference $\mathcal{M} := M_B - \mathcal{H}$, adopting the blinds for each $M_B$ and $\mathcal{H}$ noted in Section 4.3.4 i.e. with the transformations $m_B \mapsto m_B + \mathfrak{o}_{M_B}$ in Equations 4.11 and 4.12 and $\mathcal{H} \mapsto \mathcal{H} + \mathfrak{o}_\mathcal{H}$ in the likelihood incorporating Equation 4.23 (a function of both $M_B$ and $\mathcal{H}$ through Equation 4.22). The marginalised posterior distributions (mean and $1\sigma$ width) for $\alpha$ and $\beta$ are presented in Table 4.6 and plotted in Figure 4.3; these results are dependent on the choice of quality cuts on the SN sample described in Section 4.5.1

Table 4.6 Results of preliminary SN-only fits for various cuts.

| SN cut | $N_{\rm SN}$ | $\alpha$ | $\beta$ | $\mathcal{M}$ |
|---|---|---|---|---|
| default | 171 | 0.164 (0.013) | 3.07 (0.14) | −3.240 (0.036) |
| higher $\chi^2$ | 175 | 0.167 (0.013) | 3.12 (0.13) | −3.244 (0.036) |
| lower $\chi^2$ | 163 | 0.158 (0.013) | 3.04 (0.16) | −3.256 (0.038) |
| $z > 0.0233$ | 96 | 0.163 (0.016) | 2.73 (0.17) | −3.252 (0.038) |
| stricter $C$ | 160 | 0.158 (0.015) | 2.93 (0.18) | −3.238 (0.037) |
| str. $\sigma_{X_1}, \sigma_C$ | 164 | 0.171 (0.014) | 3.10 (0.14) | −3.232 (0.038) |
| str. $\sigma_{X_1}$ | 165 | 0.171 (0.013) | 3.10 (0.15) | −3.245 (0.037) |
| str. $E(B-V)$ | 166 | 0.167 (0.013) | 3.06 (0.15) | −3.241 (0.036) |
| $t_{1st} < +10$d | 187 | 0.165 (0.013) | 3.11 (0.14) | −3.234 (0.035) |

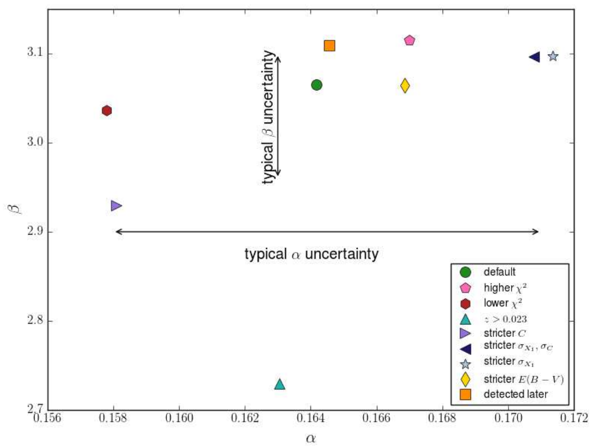


Figure 4.3 The best-fitting values for $\alpha, \beta$ from all SN-only fits to Equation 4.12, assuming various cuts on the low-$z$ SNe. The different markers represent the cuts described in Section 4.5.1. The typical statistical uncertainties are indicated by the arrows. The variation in $\alpha$ is comparable to the statistical uncertainty, and the same is true for $\beta$ if we disregard the higher low-redshift cut.

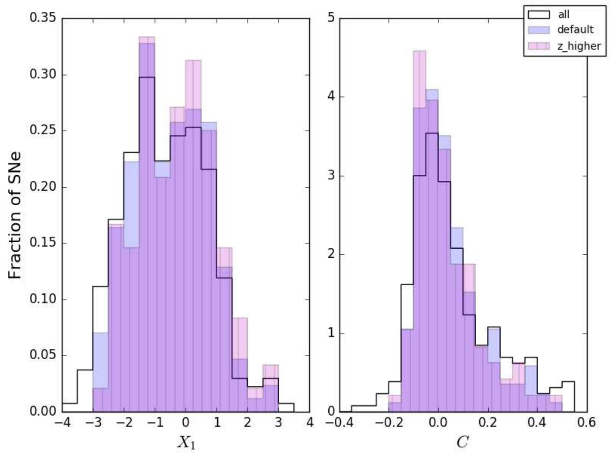


Figure 4.4 Normalised histograms of the $X_1, C$ distributions with a higher low-redshift cut of $z > 0.0233$. The proportion of higher $X_1$ (slower-declining) supernovae is marginally higher with the $z > 0.0233$ cut, but otherwise the relative distributions appear very similar.

### 4.5.5 Results of SN-only fit

The results of the SN-only $\chi^2$-minimising fit are presented in Table 4.6, while Figure 4.3 shows the differences in fits with various SN cuts lie in the $\alpha, \beta$-plane (this is analogous to Figure 4.1, which displays the numerous Cepheid fits in $b_W, Z_W$-space). We discuss the dependence of the fit results on the various cuts, and select cuts with results spanning the parameter space to carry forward to the global fit to assess the associated systematic uncertainty.

The notable outlier is the higher low-redshift cut ($z > 0.0233$), effecting a much lower value of $\beta$ than the other cuts. This cut, along with the stricter colour and stricter goodness-of-fit cuts, results in lower $\alpha$ also. The lowest and highest values of $\alpha$ correspond to lower $\chi^2$, and stricter $\sigma_{X_1}$ respectively. Figure 4.4 shows normalised $X1$ and $C$ distributions for the $z > 0.0233$ cut: there are marginally slower-declining SNe compared to the default, but overall the distributions appear similar. From Figure 4.4, it does not appear that the discrepant fit results from this cut are the result of a change in the colour (or stretch) distribution of the sample, due to selecting bluer SNe at higher redshifts. Indeed, our tests with jackknifed samples (described below) indicate this is likely the result of removing a large portion of the sample (over 40% relative to the default). We disregard the $z > 0.0233$ cut in our reported error for the nuisance parameter $\beta$ in Equation 4.27 because of the large number of supernovae removed from the sample, and the lack of evidence that this result reflects a systematic or selection effect.

Excluding the $z > 0.0233$ cut, the variation in $\alpha$ and $\beta$ with the different cuts we test appears only slightly larger than the typical statistical uncertainties in these parameters (Figure 4.3).

We use jackknife resampling to assess the statistical significance of the dependence of results on the cuts in Table 4.6. For several cuts (the lower lightcurve $\chi^2$/DoF, higher redshift cut and stricter cuts on colour or uncertainties in stretch and colour) we draw subsamples of size $N_{\mathrm{SN}}$ (Table 4.6) of the 171 SNe selected by the default cut. For each cut we compare the systematic change in fit results (parameters $\alpha, \beta, \mathcal{M}$) from the new cut to results from repeated jackknifed subsamples of size $N_{\mathrm{SN}}$ and their scatter. These reveal a systematic variation of $1 - 3\sigma$ from the default for almost all combinations of parameters and cuts (where $\sigma$ is the scatter within the numerous jackknifed subsamples). Thus the differences between rows of Table 4.6 cannot be solely attributed to shot noise, and the variation due to different cuts must be propagated to the global fit (Section 4.6) and treated as a contribution to the total systematic uncertainty. However, we will find that the variation from the choice of SN cut is dwarfed by the analogous source of uncertainty from the choice of Cepheid fits.

## 4.6 Global fit

This section contains our final simultaneous fits to all Cepheid and SNe Ia data. We set out parameters and equations for this fit, and present fit results for all parameters, including the dependence of results on choices within the individual Cepheid and supernova data sets. We summarise our uncertainties, and discuss their increase compared to other analyses of the same data. Finally, we break down the statistical and systematic contributions to the uncertainty budget.

### 4.6.1 Equations

We fit all Cepheid and supernova data simultaneously to Equations 4.10, 4.11, 4.12 as described in Section 4.3.3. We minimise a global $\chi^2$ function (a function of $\Theta = \{\alpha, \beta, M_B, \mathcal{H}, b_W, Z_W, M_W, \mu_{4258}, \{\Delta\mu_i\}\}$), which has contributions from the Cepheids and low-$z$ SNe remaining after cuts (given in Equations 4.14 and 4.23 respectively), and also an equivalent term to $\chi^2_{\rm lowz}$ for the eight nearby SNe:

$$\chi^2_{\rm global} = \chi^2_c + \chi^2_{\rm SN} + \chi^2_{\rm nearby} \tag{4.25}$$

$$\chi^2_{\rm nearby} = (\hat{\boldsymbol{m}}^\dagger_{\boldsymbol{B}} - \hat{\boldsymbol{m}}^\dagger_{\boldsymbol{B}\rm mod}) \cdot \mathbf{C}_{\mathbf{m}^\dagger_{\mathbf{B}},\mathbf{n}}{}^{-1} \cdot (\hat{\boldsymbol{m}}^\dagger_{\boldsymbol{B}} - \hat{\boldsymbol{m}}^\dagger_{\boldsymbol{B}\rm mod})^T. \tag{4.26}$$

The bolded quantities in Equation 4.26 are vectors, over the eight nearby SNe Ia. The terms contributing to the nearby covariance matrix $\mathbf{C}_{\mathbf{m}^\dagger_{\mathbf{B}},\mathbf{n}}$ are covariances between SALT2 quantities $m_B, X_1$, and $C$, and the diagonal intrinsic scatter $\sigma_{\rm int,SN}$.

The global simultaneous fit is 16- or 17-dimensional (without and with the LMC included as a distance anchor respectively), and performed using MultiNest as described in Section 3.2.2. We are ultimately interested in $\mathcal{H}$, which contains the value of $H_0$. However to demonstrate degeneracies and correlations between parameters, we display in Figures 4.5 and 4.6 marginalised contour plots of the posterior distribution of an example fit (with $T = 2.25$, NGC 4258+LMC+MW anchor, $P < 60$ day cut Cepheid fit and default SN cuts).[11] The former posterior distribution is marginalised over the eight $\Delta\mu_i$, while the latter is also marginalised over $\mu_{4258}$ and the SN and Cepheid parameters which are strongly constrained by initial fits: $\{\alpha, \beta, b_W, Z_W\}$. Figure 4.5 shows a strong positive correlation between $M_B$ and $\mathcal{H}$ as expected from their degeneracy in the low-$z$ SN sample (Equation 4.12), and less apparent correlations between the 'zero point-like' parameters $\{\mathcal{H}, M_B, M_W, \mu_{4258}\}$. In contrast, the five other parameters $\{\alpha, \beta, b_W, Z_W, \mu_{4258}\}$ each are largely independent of the other parameters (Figure 4.5).

We repeat the global fit for each of 18 Cepheid fits in Table 4.4 and six supernova cuts determined in Section 4.5.5 from Figure 4.3. Each Cepheid fit and SN cut corresponds to a subset of the total sample to use in the global fit, and associated values of best-fitting parameters, as well as $\sigma_{\rm int,C}$ for the Cepheids. In total there are 108 fits; the analysis of these results and the variation therein follows.

### 4.6.2 Results of global fit

The best-fitting values and uncertainties of parameters in $\Theta$ are determined for each of 108 fits. Some tables of results in Appendix D of Z17 are omitted here for brevity; we refer the interested reader to table D3 of Z17 for the full results for parameters in $\Theta$. Figure 4.7 displays these fits in various subspaces of the 16- or 17-dimensional space spanned by $\Theta$, focussing on parameters $\{\alpha, \beta, b_W, Z_W, M_W, M_B, \mathcal{H}\}$. We discuss the dependences that this figure shows (which motivate the averaged tables and figures later), then present results for the nuisance parameters and the parameters of interest: the SN Ia peak absolute magnitude $M_B$ and (proxy for the) Hubble constant $\mathcal{H}$, which are degenerate with each other.

In the remainder of the section we depart from the distinction we make between statistical and systematic uncertainties in Section 3.4: the uncertainties returned by MultiNest, reported in table D3 of Z17, simply the $1\sigma$ widths of the PDFs, do not distinguish between the statistical and systematic

[11]Figures 4.5, 4.6, and 9 were created with the `ChainConsumer` package (Hinton, 2016).

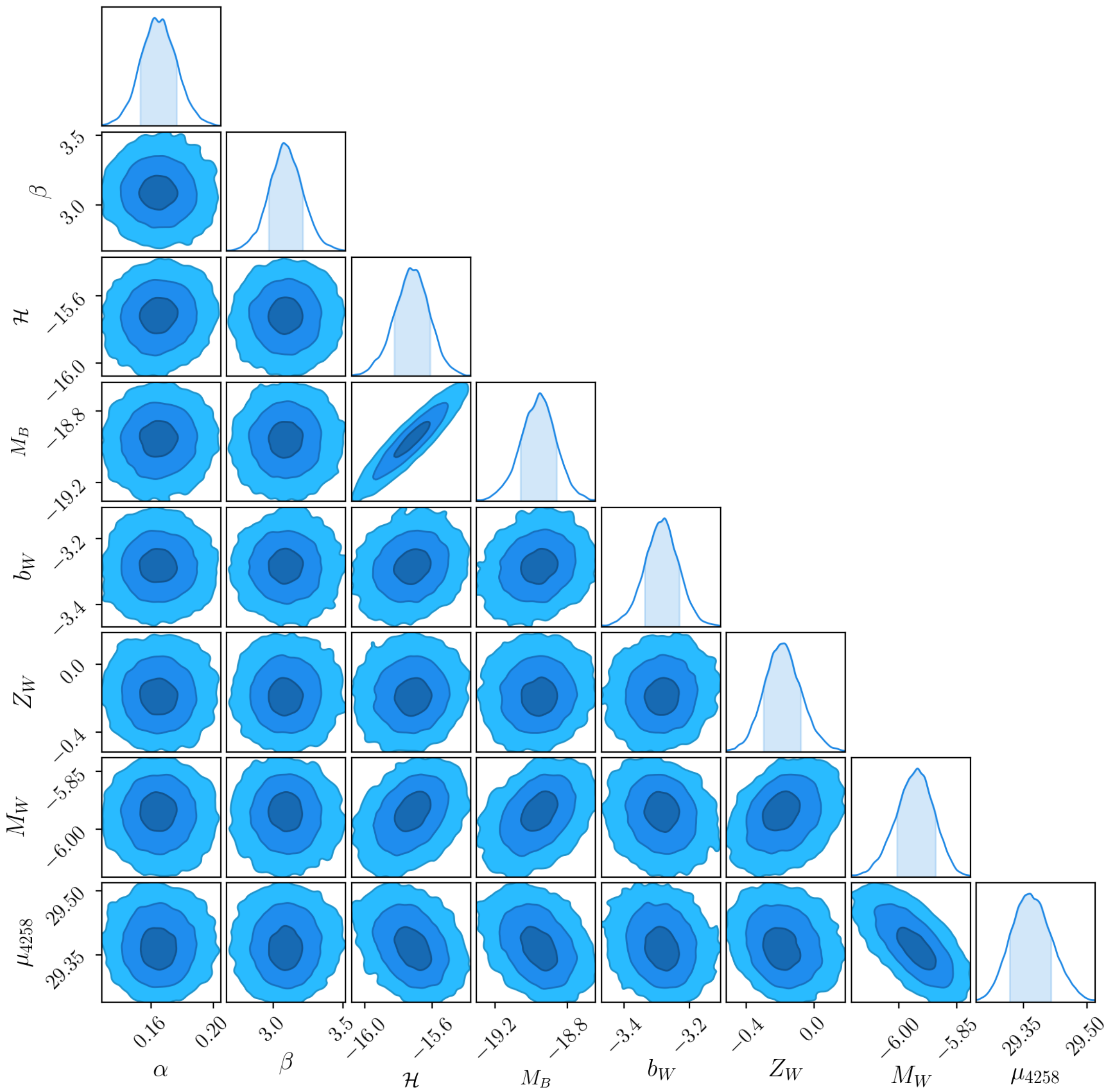


Figure 4.5 Constraints on parameters in $\Theta$ from an example global MultiNest fit (with $T = 2.25$, NGC 4258+LMC+MW anchor, $P < 60$ day cut Cepheid fit, default SN cuts) marginalised over $\{\Delta\mu_i\}$. The shaded regions in the PDFs represent $1\sigma$ levels, and the $1\sigma$, $2\sigma$, and $3\sigma$ regions are shown in the contours. Note the strong degeneracy between $\mathcal{H}$ and $M_B$, and slightly weaker degeneracies between $\mathcal{H}, M_B, \mu_{4258}$, and $M_W$. The other parameters appear uncorrelated.

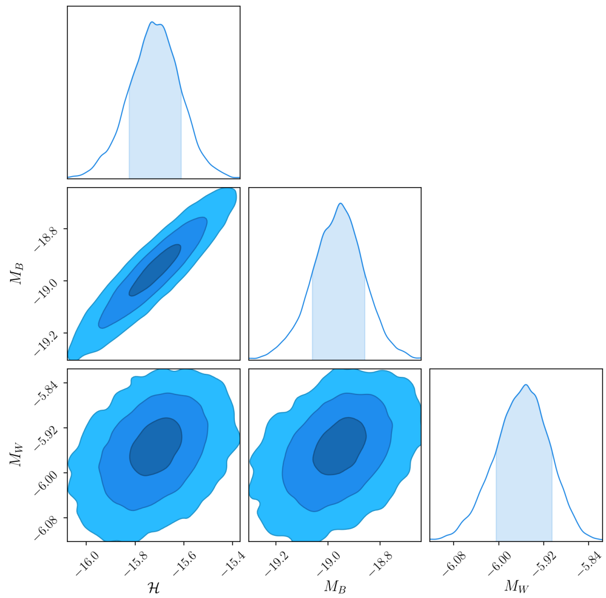


Figure 4.6 The same fit as Figure 4.5, also marginalised over $\{\alpha, \beta, b_W, Z_W, \mu_{4258}\}$. This shows the three parameters that are the most highly correlated.

components of covariance matrices input into the fit in the likelihood. Henceforth, we refer to this uncertainty from the MultiNest fit as statistical, and the variation observed in e.g. Figure 4.7 between global fits with differing supernova cuts or Cepheid fits as systematic.

### 4.6.3 Dependence of parameters

Figure 4.7 highlights the following dependence of parameters on properties of the global fit:

1. The Cepheid parameters $\{b_W, Z_W, M_W\}$ depend only on the choice of Cepheid fit (carried forward from Section 4.4.1), reflecting the variation observed in Figure 4.1. Thus there is negligible scatter in values for these parameters between fits with the same Cepheid data, regardless of the SN cut (Figure 4.7(a), (b)).

2. Similarly, the SN parameters $\{\alpha, \beta\}$ depend most strongly on cuts, and minimally on Cepheid fit, although there is more scatter than in $\{b_W, Z_W\}$. On average fits without an upper period cut on the Cepheids result in slightly lower $\alpha$ by $\sim$0.01, for each SN cut (Figure 4.7(c)).

3. The Cepheid and SN zero points $M_W$ and $M_B$ both depend predominantly on the Cepheid fit (Figure 4.7(b)), reflecting the fact that the SNe Ia are calibrated on the Cepheids. While $M_W$ depends directly on the Cepheid data (Equation 4.10), the influence on $M_B$ is through its interaction with $M_W$ via the distance modulus offsets $\{\Delta\mu_i\}$ (Equations 4.10 and 4.11). We note that $M_W$ has negligible dependence on SN cut, whereas $M_B$ varies slightly with the choice of SN cut (with a spread of $\sim$0.01 within each choice of Cepheid fit).

4. As mentioned at the start of Section 4.5.4, $\mathcal{H}$ is degenerate with $M_B$. Figure 4.7(d) shows this degeneracy between the parameters, and that the difference $\mathcal{M} = M_B - \mathcal{H}$ lies on a straight line. Within each choice of Cepheid fit there is slight systematic dependence only on the choice of SN cut. There is no systematic difference between these parameters from fits with and without a cut on Cepheid period.

5. $\{\Delta\mu_i\}$: the values of the distance modulus offsets from the global fit depend significantly on the Cepheid fit, as shown in Figures 4.8 and 4.9.

In summary, it is expected that the SN cuts determine parameters $\{\alpha, \beta\}$, and the Cepheid fits determine parameters $\{b_W, Z_W, M_W\}$. However the interaction of the 'zero point-like' parameters is more

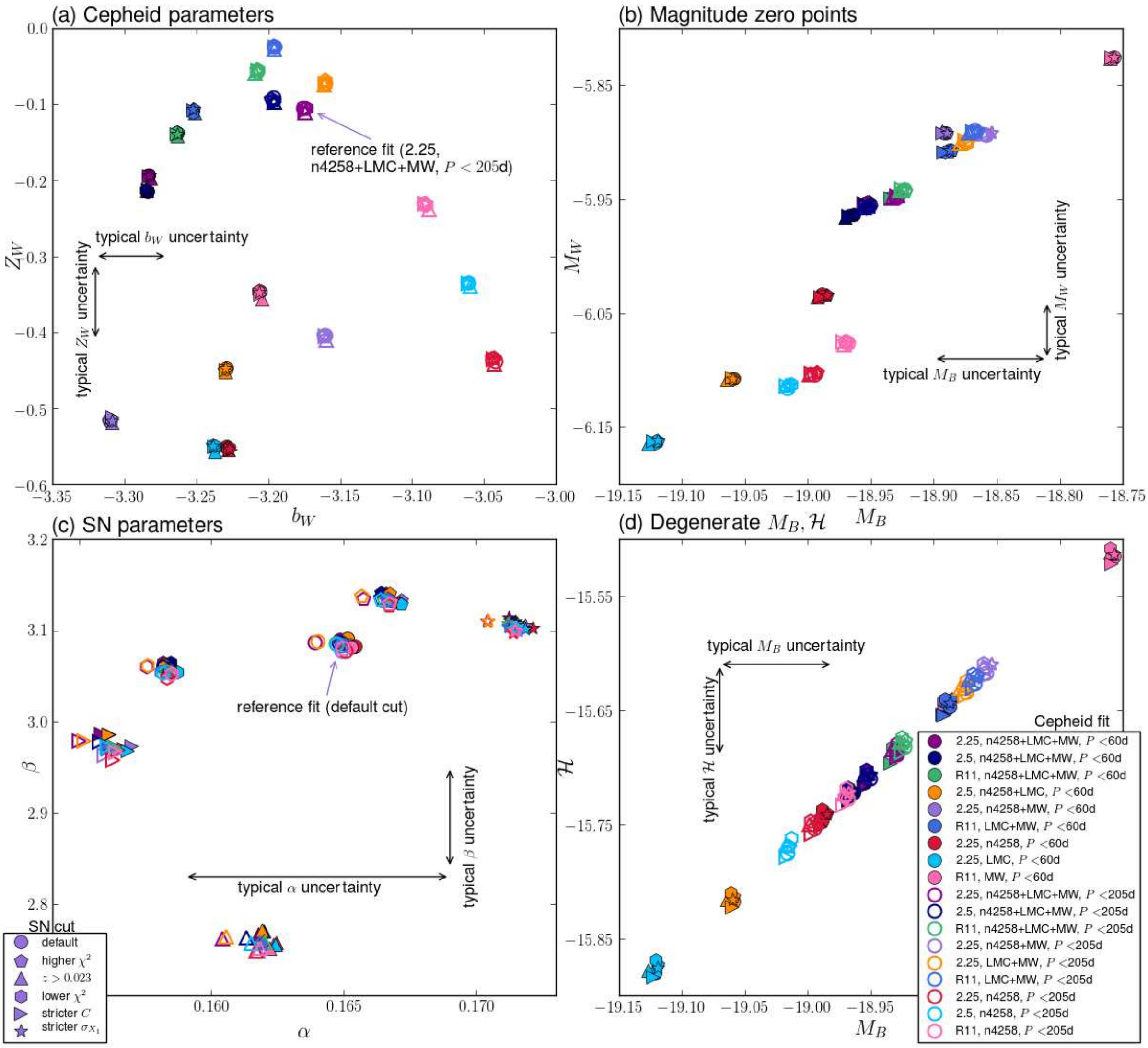


Figure 4.7 Results of all global fits to Equations 4.10–4.12 simultaneously, in the (a) $b_W, Z_W$- (b) $M_B, M_W$- (c) $\alpha, \beta$- and (d) $M_B, \mathcal{H}$-planes, when assuming various choices of SN cut and Cepheid fit. As shown in the legends the different combinations of colour and fill encapsulate information on the choice of Cepheid fit as described in Section 4.4.1, while the different shapes represent difference cuts on the supernovae from Section 4.5.5. The chosen reference fits (bolded in Table 4.7 and 4.8) are indicated by the violet arrows in (a) and (c). The overlap of points with the same colour and fill in (a) demonstrate that the Cepheid parameters $b_W, Z_W$ depend only on the Cepheid fit; similarly, the clusters of points with the same shape in (c) show that the SN Ia parameters $\alpha, \beta$ depend mostly on the SN cut. Subplot (b) shows that $M_W$ and $M_B$ both depend predominantly on the choice of Cepheid fit, with the effect more strong in $M_W$. A strong degeneracy between $M_B$ and $\mathcal{H}$ is evident in (d), indicating that $\mathcal{H}$ depends primarily on the Cepheid fit, and secondarily on the SN cut. There is no systematic difference in $M_W, M_B$, and $\mathcal{H}$ between fits with and without an upper limit on Cepheid period.

Table 4.7 Global fit results for supernova parameters $\alpha, \beta$ for each SN cut, averaged over Cepheid fits. The default rejection (bolded) detailed in Section 4.5.1 is chosen as our reference fit.

| SN cut | $\alpha$ | $\beta$ |
|---|---|---|
| **default** | **0.165** | **3.09** |
| higher $\chi^2$ | 0.167 | 3.134 |
| $z > 0.0233$ | 0.162 | 2.759 |
| lower $\chi^2$ | 0.158 | 3.057 |
| stricter $C$ | 0.156 | 2.974 |
| stricter $\sigma_{X_1}$ | 0.171 | 3.106 |

Table 4.8 Global fit results for Cepheid parameters $\{b_W, Z_W, M_W\}$ for each Cepheid fit, averaged over SN cuts. The bolded fit ($T = 2.25$ rejection, all three anchors, and no upper cut on the period) is chosen as our reference fit.

| Rejection | Distance anchor | $P < 60$d | $b_W$ | $Z_W$ | $M_W$ |
|---|---|---|---|---|---|
| 2.25 | all[a] | Y | −3.28 | −0.19 | −5.95 |
| 2.5 | all | Y | −3.28 | −0.21 | −5.96 |
| R11 | all | Y | −3.26 | −0.14 | −5.95 |
| 2.5 | n4258+LMC | Y | −3.23 | −0.45 | −6.11 |
| 2.25 | n4258+MW | Y | −3.31 | −0.52 | −5.89 |
| R11 | LMC+MW | Y | −3.25 | −0.11 | −5.91 |
| 2.25 | n4258 | Y | −3.23 | −0.55 | −6.03 |
| 2.25 | LMC | Y | −3.24 | −0.55 | −6.16 |
| R11 | MW | Y | −3.21 | −0.35 | −5.83 |
| **2.25** | **all** | **N** | **−3.17** | **−0.11** | **−5.95** |
| 2.5 | all | N | −3.20 | −0.10 | −5.96 |
| R11 | all | N | −3.21 | −0.06 | −5.94 |
| 2.25 | n4258+MW | N | −3.16 | −0.41 | −5.89 |
| 2.25 | LMC+MW | N | −3.16 | −0.07 | −5.90 |
| R11 | LMC+MW | N | −3.20 | −0.03 | −5.89 |
| 2.25 | n4258 | N | −3.04 | −0.44 | −6.10 |
| 2.5 | n4258 | N | −3.06 | −0.34 | −6.11 |
| R11 | n4258 | N | −3.09 | −0.23 | −6.08 |

[a] i.e. n4258+LMC+MW

subtle, and emerges from the simultaneous fit of the three data samples, most obvious in Figure 4.7(b). Even though the parameter $M_B$ only appears in the SN apparent magnitudes (Equations 4.11, 4.12), it is most strongly influenced by the Cepheid data via $M_W$, as the two parameters are tied to their respective data sets through the distance modulus offsets $\{\Delta\mu_i\}$. Furthermore, $M_B$ and $\mathcal{H}$ are degenerate with their difference determined by the low-$z$ SNe. Thus the resultant value of $\mathcal{H}$, hence $H_0$, is sensitive both to the choice of SN cut (via the $M_B - \mathcal{H}$ degeneracy) and to the choice of Cepheid fit (via the influence of $M_W$ on $M_B$). Unsurprisingly, the most extreme values of $M_B$ and $\mathcal{H}$ (both driven by $M_W$, as seen in Figure 4.7(b)) arise from Cepheid fits anchored on only the LMC or MW (most and least negative, represented by dark purple and pink symbols, respectively). It is clear from Figure 4.7(d) that the variation with Cepheid fits (anchor and rejection) is at least an order of magnitude larger than the variation with SN cuts, even when the fits anchored on the LMC or MW only are excluded.

### 4.6.4 Nuisance parameter results

Tables 4.7 and 4.8 contain results for the supernova and Cepheid nuisance parameters, averaged over the Cepheid fits and SN cuts respectively. We choose to average over these aspects of the fit that have minimal effect on the parameters, as shown in Figure 4.7: the SN parameters in (c) predominantly depend on shape (SN cut) and not on colour (Cepheid fit), while the Cepheid parameters in (a) depend

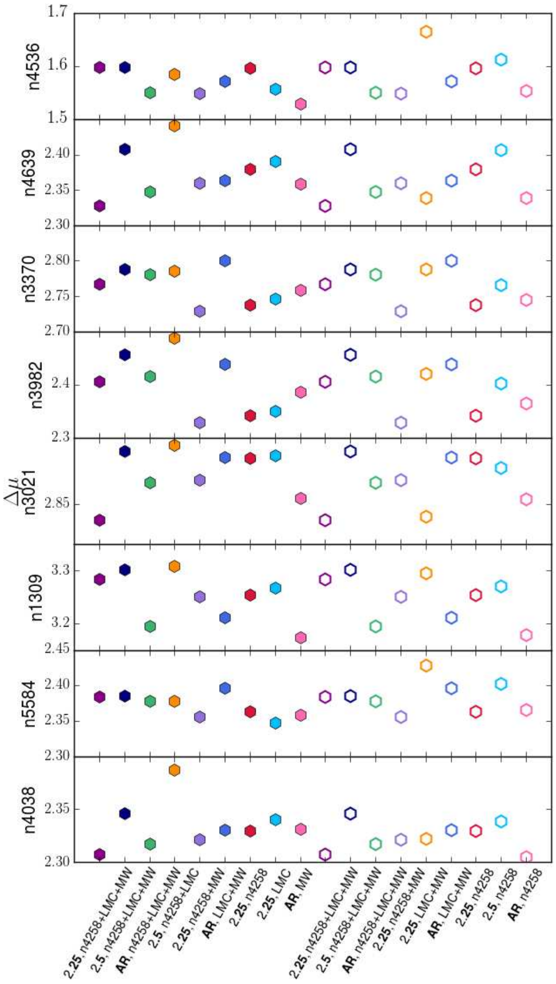


Figure 4.8 Visualisation of the best-fitting values for each $\Delta\mu_i$, which vary slightly with the different Cepheid fits in Section 4.4.1 (symbols shown in legends of Figures 4.7 and 4.10). Each horizontal subplot represents a different galaxy.

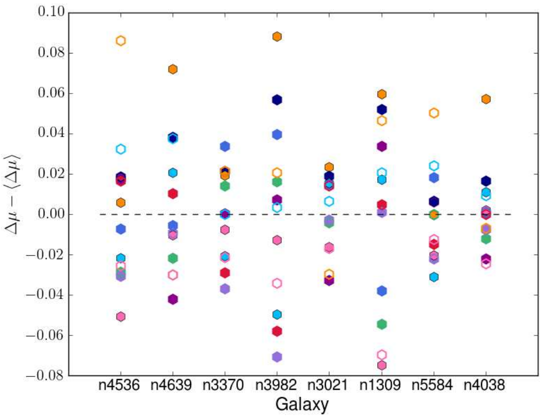


Figure 4.9 Visualisation of the relative value of the $\Delta\mu_i$ with respect to the mean over all Cepheid fits for each galaxy, marked by the black dashed line. The legends in Figures 4.7 and 4.10 show the Cepheid fit (colour and fill).

entirely on colour and not on shape. We omit statistical uncertainties of parameters in these tables as they can be obtained from the full set of results in table D3 of Z17. For the nuisance parameters we select the single best fit (bolded) and indicated in Figure 4.7). This is preferable to averaging over results in Tables 4.7 and 4.8, which are asymmetric, based on different premises (e.g. different distance anchors), and include more questionable fits (e.g. those SN cuts that reject a larger fraction of the total). Thus we use the maximal variation in these values to inform our systematic error budget, but not to influence the best fit.

Final values for the nuisance parameters are taken from the bolded reference fits, which have the default SN cut and the Cepheid fit with all three anchors, $T = 2.25$ rejection, and no cut on Cepheid period. We have chosen this fit because the results are representative and centred amongst the different choices. The Cepheid fit here also aligns with fits selected in E14 and R11. As in R11, we choose to not impose a cut on Cepheid period, and note the effects of including this cut on nuisance parameters described in Section 4.4.1: both $b_W$ and $Z_W$ are more negative with the $P < 60$ day cut, while there is no difference in $M_W$ when all three distance anchors are used. The global SN results differ slightly from the initial results in Table 4.6, and we again note that the most deviant (lowest) values of $\alpha$ or $\beta$ are where a large number of SNe have been rejected; the remaining cuts are in agreement with values derived from the default cut. In summary, the fit parameters and uncertainties from Tables 4.7 and 4.8 are:

$$\begin{aligned}
\alpha &= 0.165 \pm 0.010(\text{stat})^{+0.004}_{-0.005}(\text{sys}) \\
\beta &= 3.09 \pm 0.11(\text{stat})^{+0.04}_{-0.12}(\text{sys}) \\
b_W &= -3.17 \pm 0.04(\text{stat})^{+0.13}_{-0.11}(\text{sys}) \\
Z_W &= -0.11 \pm 0.09(\text{stat})^{+0.08}_{-0.10}(\text{sys}) \\
M_W &= -5.95 \pm 0.04(\text{stat})^{+0.06}_{-0.12}(\text{sys}).
\end{aligned} \tag{4.27}$$

These statistical uncertainties are found from table D3 of Z17. We generally take the maximal variation measured from the reference fits in Tables 4.7 and 4.8 as the systematic uncertainty, with the following exceptions. We disregard the higher low-redshift cut (associated with a large fraction of the SNe being discarded) in estimating the systematic uncertainty in $\beta$; as discussed in Section 4.5.5, the lower value of $\beta$ is impacted by a significantly reduced sample size, which does not show selection biases. Nevertheless, some evolution of $\beta$ with redshift has been noted in the literature (e.g. Conley et al., 2011)), and is worth investigating in future. For the uncertainty in $Z_W$, we only consider the variation over fits which include both the LMC and MW in the distance anchor: the constraints on the metallicity dependence provided by different distance anchors are inconsistent with each other, so we only consider these fits for estimating the uncertainty for the nuisance parameter $Z_W$ alone (i.e. the other anchors are considered for estimating uncertainties on $M_B$ and $\mathcal{H}$, in Section 4.6.5.) From Figure 4.7(c) it is clear that the statistical uncertainties in the SN parameters are around double the systematic uncertainty if we disregard the higher low-redshift SN cut. The opposite is true for the Cepheid parameters, where the statistical uncertainties are dwarfed by systematic variation with differing fits. If we restrict our analysis to only Cepheid fits anchored on all three galaxies, the statistical and systematic uncertainties are comparable in size.

The systematic errors are asymmetric for most parameters, especially for $\beta$ (due to the outlying $z > 0.0233$ cut) and $Z_W$. This can be observed in Figure 4.7(a)–(c), where it is evident our reference fits do not lie centrally within the parameter subspaces. Figure 4.7(b) shows that the MW as a distance anchor drives $M_W$ up, while the LMC (and to a lesser extent NGC 4258) drives $M_W$ down, an effect which propagates to $M_B$ and $\mathcal{H}$ (Figure 4.7(d)). Fits anchored on all three distance anchors lie centrally. Our Cepheid nuisance parameters remain consistent with R11 and E14 as initially found in Section 4.4.1.

We note that our best-fitting value for $\alpha$ is significantly higher than found in JLA (Betoule et al. (2014), table 10) and LOSS (Ganeshalingam et al., 2013), by $\sim$0.02 (around double the total uncertainty in $\alpha$). This difference occurs consistently over a range of SN cuts. While the JLA analysis always determines $\alpha$ from the low-$z$ sample in conjunction with a higher-redshift sample, Ganeshalingam et al. (2013) finds $\alpha = 0.146 \pm 0.007$ from the LOSS sample, which overlaps with ours considerably and is over a similar redshift range. Our results for $\beta$ are consistent with the literature with the exception of the $z > 0.0233$ SN cut, which results in a value $\sim 1\sigma$ below the other cuts (the triangles in Figure 4.7(c)). The impact of this cut on $\mathcal{H}$ can be seen in Figure 4.7(d): the triangles (higher low-$z$ cut) have higher $\mathcal{H}$ than the other shapes (cuts) for each colour/fill (Cepheid fit). This effect is much smaller than the differences from varying the Cepheid fit. Nevertheless, it is in agreement with the increase of $H_0$ with increasing low-$z$ observed in R16, figure 13.

The remaining nuisance parameters are the distance modulus offsets $\{\Delta\mu_i\}$, which, like all other nuisance parameters, are eventually marginalised over. Their values depend primarily on the Cepheid fits. The full table of fit values is left to table D2 in Z17. The $\Delta\mu_i$ are visualised in Figures 4.8 and 4.9 with different colour/fill representing Cepheid fit. Figure 4.8 gives some insight into the interplay and correlations between distance moduli of different galaxies, while Figure 4.9 shows the scatter and relative values of the $\Delta\mu_i$ from different fits. The statistical uncertainties in $\Delta\mu_i$ from individual fits range from 0.05 to 0.1, and is comparable to the scatter over different fits.

### 4.6.5 Results for $M_B$ and $\mathcal{H}$

We now consider the parameters $M_B$ and $\mathcal{H}$ which, together, directly reveal $H_0$. The degeneracy between them is apparent in Figure 4.7(d), which also shows that their primary dependence is on the Cepheid fits. Thus in Table 4.9 we present the global fit results averaged over the SN cuts.[12] Given that the fits in Table 4.9 anchored on all three galaxies are spread out, we average these fits rather than choose a best fit, and take the maximal variation in these fits as the systematic uncertainty. There is a slight systematic difference between fits in Table 4.9 with and without the upper period limit (on average, $\mathcal{H}$ is decreased by 0.015 mag where the $P < 60$ day cut is applied). From a theoretical standpoint, we have no reason to preference one cut over the other. Thus our best estimates for $M_B$ and $\mathcal{H}$ are averaged over all fits anchored on all three galaxies (including fits both with and without the upper period limit), represented by solid and empty navy, green and dark purple markers in Figures. 4.7–4.10.

Our **best estimates** are

$$\begin{aligned} M_B &= -18.943 \pm 0.088(\text{stat}) \pm 0.024(\text{sys}) \\ \mathcal{H} &= -15.698 \pm 0.093(\text{stat}) \pm 0.023(\text{sys}). \end{aligned} \tag{4.28}$$

Here the statistical uncertainties are found from relevant fits in table D3 of Z17, in which a representative fit is bolded (with default SN cut, $T = 2.25$ rejection, and no upper period limit). The above systematic uncertainties are given by the maximal variation in values with the combined NGC 4258+LMC+MW distance anchor. We impose this constraint on the anchor so that we can fairly assess the systematic uncertainty when all available distance information is used, and to allow better comparison with R11 and E14 who primarily report errors with all three anchors. In Section 4.7 we investigate and discuss uncertainties in $\mathcal{H}$, including converting from an absolute error in the logarithmic quantity $\mathcal{H}$ to a relative error in $H_0$.

We next consider fits anchored on NGC 4258 only, to estimate the uncertainty in this case, and for the sake of comparison with R11 and E14. These fits are represented by the empty turquoise and pink markers, and by all red markers in Figures. 4.10 and 4.7. We average these results from Table 4.9 to find Equation 4.29, as with Equation 4.28. The systematic uncertainties are given by the maximal variation in values derived from these fits, and the statistical uncertainties are found from

[12] We report $M_B$ and $\mathcal{H}$ to 3 decimal places, unlike most other parameters which have been truncated to 2 decimal places (but not rounded in the analysis). These two quantities are of particular interest, and it is desirable to retain precision in both their values and uncertainties throughout this section.

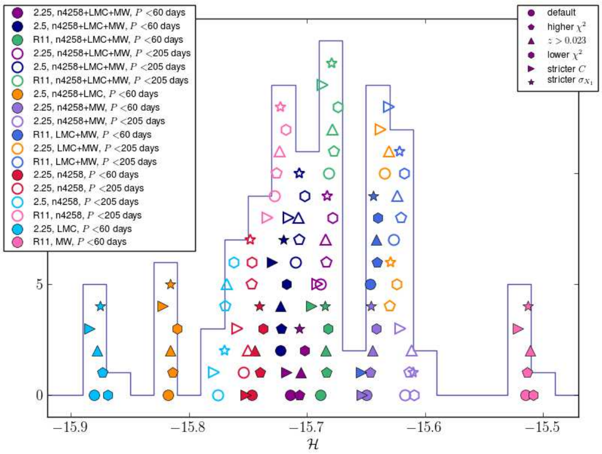


Figure 4.10 Histogram of best-fitting values for $\mathcal{H}$ from all global fits: essentially a histogram of Figure 4.7(d) projected on to its $y-$axis. The blue line shows the binned histogram, while the individual points are plotted with their true $\mathcal{H}$ values and to reflect the distribution (i.e. they are stacked vertically for each bin). The frequency reflects the fits we chose to include in the global fit, i.e. we deliberately included more fits with all three anchors (and to a lesser extent, anchored on NGC 4258), rather than an inherent distribution. The legends are the same as in Figure 4.7, and reflect the Cepheid fit (colour and fill, with solidity of markers reflecting the inclusion of a Cepheid period cut) and SN cut (shape). The fits anchored on NGC 4258 only have a much broader spread in $\mathcal{H}$, and are responsible for the lowest values. The range in values in the NGC 4258-anchored fits is much greater, extending from left-filled pink markers to top-filled navy and dark purple markers and spanning ∼0.11 mag. In contrast the fits anchored on all three galaxies extend from the solid navy markers to the empty green markers, spanning ∼0.04 mag.

NGC 4258-anchored fits in table D3 of Z17, with one representative fit bolded.

$$\begin{aligned} M_B &= -18.993 \pm 0.104(\text{stat}) \pm 0.023(\text{sys}) \\ \mathcal{H} &= -15.748 \pm 0.107(\text{stat}) \pm 0.023(\text{sys}). \end{aligned} \tag{4.29}$$

The resultant value of $\mathcal{H}$ in Equation 4.29 is 0.05 mag lower (corresponding to a 2.3% decrease in $H_0$) compared to where all three anchors are used (Equation 4.28). Moreover, $M_B$ (which is largely degenerate with $\mathcal{H}$) is also 0.05 mag lower (brighter). The systematic uncertainty (i.e. the spread in values between different fits) is the same, while the statistical uncertainties are larger, reflective of the fact that a distance scale is anchored on a smaller set of data.

Our best estimate of the peak SN Ia brightness $M_B$ (in Equation 4.28, from the three-galaxy anchor) appears mildly higher (dimmer) than values reported in JLA (assuming $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$), which are $M_B = -19.05 \pm 0.02$ from all SN Ia data, or $M_B = -19.02 \pm 0.03$ from a lower-redshift subsample consisting of low-$z$ and SDSS supernovae (table 10 of Betoule et al. (2014)). However, the supernova-only fit in JLA cannot constrain both $M_B$ and $H_0$, which are degenerate. As they have assumed a value for $H_0$ (while we have fitted separately using a distance ladder), our numerical values for $M_B$ are not directly comparable, but merely reflect the influence of different values of $H_0$.

Returning to $H_0$, Equation 4.28 corresponds to a value of $H_0 = 72.5 \pm 3.1(\text{stat}) \pm 0.77(\text{sys})$ km s$^{-1}$ Mpc$^{-1}$ (total uncertainty of 4.4%) from the combined NGC 4258+LMC+MW anchors. If we assume the older distance $\mu_{4258} = 29.31$ in R11 (Footnote 7), our best estimate increases to $H_0 = 73.8 \pm 3.2(\text{stat}) \pm 0.78(\text{sys})$. These central values agree with R11 ($H_0 = 73.8 \pm 2.4$) and E14 ($H_0 = 72.5 \pm 2.5$), which respectively assume $\mu_{4258} = 29.31$ and 29.404. Using only NGC 4258 as a distance anchor (and the new Humphreys et al. (2013) value of $\mu_{4258} = 29.404$) gives $H_0 = 70.9 \pm 3.5(\text{stat}) \pm 0.75(\text{sys})$ km s$^{-1}$ Mpc$^{-1}$,

Table 4.9 Global fit results for degenerate parameters $M_B$ and $\mathcal{H}$, averaged over SN cuts. The bolded fit ($T = 2.25$ rejection, all three anchors, and no upper cut on the period) is chosen as our reference fit.

| Rejection | Distance anchor | $P < 60$d | $M_B$ | $\mathcal{H}$ |
|---|---|---|---|---|
| 2.25 | all[a] | Y | −18.953 | −15.709 |
| 2.5 | all | Y | −18.967 | −15.722 |
| R11 | all | Y | −18.932 | −15.687 |
| 2.5 | n4258+LMC | Y | −19.061 | −15.816 |
| 2.25 | n4258+MW | Y | −18.892 | −15.647 |
| R11 | LMC+MW | Y | −18.889 | −15.644 |
| 2.25 | n4258 | Y | −18.988 | −15.743 |
| 2.25 | LMC | Y | −19.122 | −15.877 |
| R11 | MW | Y | −18.759 | −15.514 |
| 2.25 | all | N | −18.929 | −15.685 |
| 2.5 | all | N | −18.953 | −15.708 |
| R11 | all | N | −18.924 | −15.679 |
| 2.25 | n4258+MW | N | −18.859 | −15.614 |
| 2.25 | LMC+MW | N | −18.875 | −15.631 |
| R11 | LMC+MW | N | −18.868 | −15.623 |
| 2.25 | n4258 | N | −18.996 | −15.751 |
| 2.5 | n4258 | N | −19.015 | −15.770 |
| R11 | n4258 | N | −18.970 | −15.725 |

[a] i.e. n4258+LMC+MW

which is 2.3% lower than with the three anchors. The uncertainties in $\mathcal{H}$ are broken down in Section 4.7.1 and summarised in Table 4.10. We next discuss the uncertainties in $H_0$; their size informs the significance of the tension between values of the Hubble constant measured using different probes, so they are of equal interest to the values.

## 4.7 Uncertainties

We have emphasised the importance of quantifying and incorporating the scatter in parameters arising from varying aspects of the SN and Cepheid fits, and indeed we use this overall variation in results to gauge the systematic uncertainty in these parameters. However, we have also seen that the statistical uncertainty dominates for the supernova parameters $\alpha$ and $\beta$ (Figure 4.3 and 4.7(c)), as well as for $M_B$ and $\mathcal{H}$ when only considering the systematic variation between fits with all three anchors (Figure 4.7(d)). This dominance reflects the fact that the SN Ia samples, especially the nearby sample (i.e. in Cepheid hosts), are relatively small with large errors when compared to the Cepheids. Hence the SNe are statistically limited while the Cepheids are not.

For clarity we divide the contributions to the total uncertainty in the parameters into three classes:

1. The statistical (in the usual sense) portion of the uncertainty reported by MultiNest, which is dominated by noise in the nearby and low-$z$ supernovae.

2. The systematic elements of $\mathbf{C}_\eta$, which make up remainder of the uncertainty reported by MultiNest, listed in Table 4.11.

3. The systematic uncertainty associated with varying aspects of the fit between reasonable alternatives is dominated by the variation in the choice of Cepheid fit, as shown in Figures 4.7 and 4.10. For our final value and uncertainty of $H_0$ we focus on fits with all three anchors only (with some consideration of fits with only NGC 4258 as an anchor for the sake of comparison to R11 and E14). Then in effect we are only considering the variation with the rejection algorithm and the cut on Cepheid period.

Table 4.10 Summary of uncertainties in $\mathcal{H}$ from Section 4.6.5 (Equation 4.28 and 4.29), converted to relative errors in $H_0$ using Equation 4.30, and added in quadrature in line with R11. The statistical error below are those reported by MultiNest and include terms (i) and (ii) described at the start of Section 4.7. The systematic error is from the variation between fit results with different choices of Cepheid fits, and secondarily SN cuts, i.e. term (iii).

| Anchor | all | NGC 4258 only |
|---|---|---|
| $\mathcal{H}$ | −15.698 | −15.748 |
| $\sigma_{\mathcal{H}}$ | | |
| Statistical | 0.093 | 0.107 |
| Systematic | 0.023 | 0.023 |
| **Relative $H_0$ error (%)** | | |
| Statistical | 4.3 | 4.9 |
| Systematic | 1.1 | 1.1 |
| **Total** | **4.4** | **5.0** |

### 4.7.1 Uncertainties in $H_0$

We now address the uncertainty in the Hubble constant $H_0$ explicitly, using results in Section 4.6.5 (Equations 4.28 and 4.29). As the quantity $\mathcal{H}$ is related to the logarithm of $H_0$, its absolute error informs the relative error in $H_0$, via

$$\frac{\sigma_{H_0}}{H_0} = \frac{\log(10)}{5}\sigma_{\mathcal{H}}. \tag{4.30}$$

Table 4.10 summarises our calculations of the final uncertainty in $H_0$ from Equations 4.28 and 4.29. We find using Equation 4.30 relative errors in $H_0$ of 4.3% statistical and 1.1% systematic (corresponding to terms (i) and (ii) combined, and (iii) respectively as described at the start of Section 4.7) from all three distance anchors. From using only NGC 4258 as an anchor, these errors are 4.9% statistical, 1.1% systematic, 5.0% total. The final uncertainty in $H_0$ (the bottom row of Table 4.10) is the quadrature sum of the above statistical and systematic terms. Table 4.10 is comparable to the lower portion of table 5 of R11 (and subsequently table 7 of R16), which lists all systematic and statistical uncertainties contributing in quadrature to the uncertainty in $H_0$.

### 4.7.2 Relative size of SN Ia uncertainties

We now examine the breakdown of uncertainties contributing to the statistical error, which include the multiple statistical and systematic uncertainties in SN Ia parameters making up $\mathbf{C}_\eta$ as constructed in Sections 3.4 and 4.5.3.

To visually assess the impact on confidence contours we compare results from MultiNest with different covariance matrix inputs. For an example global fit (with Cepheid fit $T = 2.5$, NGC 4258 anchor, no priors, default SN cuts) we test each systematic, and compare their results from MultiNest. The full expression for the covariance matrix $\mathbf{C}_\eta$ for observed SN Ia quantities is described in Section 3.4. As entries of $\mathbf{C}_\eta$ in MultiNest we try the following: only the statistical contribution $\mathbf{C}_{\mathrm{stat}}$ (described in Section 3.4), each single systematic term added to $\mathbf{C}_{\mathrm{stat}}$, and all systematics added i.e. $\mathbf{C}_{\mathrm{stat}} + \mathbf{C}_{\mathrm{sys}}$ (the default for all global fits). The confidence contours with statistical uncertainties only and with all systematics are easily distinguishable in Figure 4.11, but the contours with individual systematics are not. Thus for clarity we only show in Figure 4.11 the systematic term from the uncertainty in host mass correction ($\mathbf{C}_{\mathrm{host}}$ in Equation 3.16, described in Section 3.4.3), in addition to $\mathbf{C}_{\mathrm{stat}}$ and $\mathbf{C}_{\mathrm{stat}} + \mathbf{C}_{\mathrm{sys}}$. The difference between the contours is slight, indicating that the uncertainties in the parameters only increase slightly when covariance matrices for different systematics are added to the statistical term $\mathbf{C}_{\mathrm{stat}}$.

Following the method in JLA (Betoule et al., 2014, section 6.2), we quantify the relative contribu-

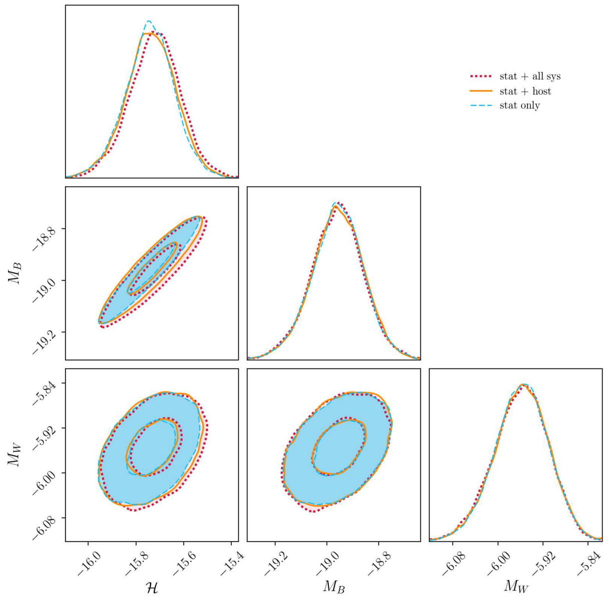


Figure 4.11 Constraints on parameters $\mathcal{H}$, $M_B$, and $M_W$ from an example global MultiNest fit (with Cepheid fit $T = 2.5$, NGC 4258 anchor, no priors, default SN cuts) with partial and full contributions to the full SN Ia covariance matrix. Confidence contours are shown with the statistical contribution $\mathbf{C}_{\text{stat}}$ only (turquoise filled), with one systematic term (the host mass correction) added i.e. $\mathbf{C}_{\text{stat}} + \mathbf{C}_{\text{host}}$ (orange solid line), and with all SN Ia systematics, i.e. $\mathbf{C}_{\text{stat}} + \mathbf{C}_{\text{sys}}$ (red dashed). The inclusion of systematic terms only increases the widths of contours marginally relative to the $\mathbf{C}_{\text{stat}}$-only (turquoise) contours, reflecting that the statistical contribution dominates the uncertainty in the parameters.

Table 4.11 Relative contributions to the uncertainty in $H_0$ (i.e. the variance) from individual statistical and systematic sources uncertainties, calculated as described in Betoule et al. (2014, Section 6.2).

| Source of uncertainty | Relative contribution to $\sigma^2(H_0)$ (%) | Described in Section |
|---|---|---|
| **Statistical:** | | |
| Lightcurves | 62.1 | 3.4.1 |
| SALT2 training | 1.2 | 3.4.1 |
| **Total statistical** | **63.3** | |
| **Systematic:** | | |
| Malmquist bias | 13.7 | 3.4.2 |
| Host galaxy | 13.0 | 3.4.3 |
| Lightcurve model | 6.8 | 3.4.7 |
| Calibration | 3.1 | 4.5.3 |
| Peculiar velocities | 0.04 | 3.4.4 |
| Milky Way extinction | 0.03 | 3.4.5 |
| **Total systematic** | **36.7** | |

tions, replacing the parameters $\{\Omega_m, w, \alpha, \beta, M_B, \Delta M\}$ with our parameters $\{H_0, M_B, \alpha, \beta\}$ (the only parameters in $\Theta$ which can be influenced by the low-$z$ SN Ia covariance matrices). The breakdown of relative contributions to the variance in $H_0$ from each term (the purely statistical term $\mathbf{C}_{\mathrm{stat}}$, and each systematic) are reported in Table 4.11. We emphasise that each of these numbers represents a proportion of the uncertainty (terms (i) and (ii) in Section 4.7 combined) from the systematic or statistical term in question alone, rather than reflecting an uncertainty in $H_0$.

From Table 4.11 and Figure 4.11 it is clear that $\mathbf{C}_{\mathrm{stat}}$ is the largest component of $\mathbf{C}_\eta$. Even though the contributions to $\mathbf{C}_\eta$ from SN systematics are included in the statistical uncertainty, all of these systematics together are smaller than the SN statistical uncertainties: the relative contributions to the variance in $H_0$ are dominated by $\mathbf{C}_{\mathrm{stat,diag}}$ (Table 4.11), and the contours with and without systematic covariance matrices added to $\mathbf{C}_{\mathrm{stat}}$ in Figure 4.11 are similar to each other. Of the systematic terms, the most significant in their impact on $H_0$ are from uncertainties in the Malmquist bias correction (including the selection function) and the host mass correction (Sections 3.4.2 and 3.4.3), followed by the uncertainty in lightcurve model. While JLA had found the photometric calibration (Section 3.1.1), especially from low-$z$ SNe, to be the dominant uncertainty for $\Omega_m$ and $w$, its effect on $H_0$ is decidedly smaller. It is interesting to note that despite conservative estimates of both the uncertainties in Milky Way extinction and peculiar velocity correction, their effects on the error in $H_0$ are negligible.

We comment on these results, particularly the increased error, in the next chapter.

Chapter 5

# Implications for the Hubble constant

In this chapter, we reflect on the findings in Chapter 4, and their place in the literature. We then review other efforts to elucidate the current tension in $H_0$ in the literature, including several other reanalyses of R11 and R16. Finally, we look toward future prospects for the Hubble constant.

## 5.1 Reflections on Chapter 4

First, we examine the results in Chapter 4, which had been published in Z17. As we will describe below, our reanalysis of R11 in Z17 confirmed the central value, a strong affirmation of the R11 result (noting we use the same data – just the analysis techniques differ), because our analysis used a different set of techniques, and was blinded. However, our uncertainty is substantially larger than in R11, and it is worthwhile to understand this discrepancy because it is central to the notion that there is tension between the CMB-related values of Planck, and SN Ia and Cepheid distance ladder derived values from the local Universe.

### 5.1.1 Summary of results in Z17

Our best estimate from R11 data is $H_0 = 72.5 \pm 3.1(\mathrm{stat}) \pm 0.77(\mathrm{sys})$ $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ using a three-galaxy (NGC 4258+LMC+MW) anchor. As stated in Footnote 7, this estimate uses the newer distance to NGC 4258 in Humphreys et al. (2013); adopting the lower value of $\mu_{4258}$ used in R11 increases our best estimate to $H_0 = 73.8 \pm 3.2(\mathrm{stat}) \pm 0.78(\mathrm{sys})$. Thus, our central value is in excellent agreement with both R11 and the E14 reanalysis. Our above value and uncertainty imply tension with the Planck value at $\sim 1.5\sigma$ significance, which is smaller than found in previous analyses of the R11 data, due to our increased uncertainties. While this discrepancy is less significant in our analysis than in the original analysis of R11 data, it has potential to be magnified by the improved data set in R16 (which has smaller statistical uncertainties compared to R11), and hence remains of interest. We also note a higher value of the stretch coefficient $\alpha = 0.165 \pm 0.010$, discussed in Section 4.6.4, a discrepancy at $\sim 2\sigma$ is surprising, which justifies further investigation.

We find a larger relative uncertainty in $H_0$ (4.4% total) compared to R11 and E14 analyses of the same data (3.3% and 3.4% total respectively), based on the NGC 4258+LMC+MW distance scale. The difference appears in the statistical error; our systematic term is similar to that in R11 (1.1% and 1.0% respectively), with the caveat that the separation of our total uncertainty into statistical and systematic components is not directly comparable to R11. These comparisons are fleshed out further in Section 5.1.2. Our larger error primarily arises from our use of covariance matrices to estimate correlated SN Ia systematic uncertainties, compared to the quadrature sum of individual errors in different parts of the analysis chain. Other significant differences in our analysis, which potentially contribute to the increased uncertainty, are our simultaneous fit of all three sets of data, allowing all parameters to interact, and our use of variation in results to quantify systematic error. These distinctions are in our view justified and desirable. Given the increase in uncertainty they produce compared to previous works, they are important to consider in future analyses.

As found in R11, our results are limited by statistics in the supernova samples. Steps to reduce this statistical uncertainty have been implemented in R16, increasing the number of nearby galaxies to 18 and improving the SN Ia photometry, to reduce the total uncertainty to 2.4%. We envisage a similar relative increase in precision when the techniques in this work are applied to the same data set. R16 also includes important changes to data analysis of the Cepheids. Other contributions to our error budget are the systematic uncertainty, which is dominated by the variation in the different Cepheid fits, and the SN Ia systematic terms in $\mathbf{C}_\eta$, the largest of which are $\mathbf{C}_{\rm bias}$ and $\mathbf{C}_{\rm host}$. As mentioned in Section 4.7.2, this contrasts with SNLS and JLA where calibration uncertainties dominate – however we note the difference there of combining supernova samples at different redshifts and considering uncertainties in $\Omega_m$ instead of $H_0$.

### 5.1.2 Increase in error compared to R11 and E14

Our final uncertainty in $H_0$ is 4.4% total (4.3% statistical and 1.1% systematic, with the statistical term inclusive of contributions from SN Ia covariance matrices) for the NGC 4258+LMC+MW distance anchor, which is significantly larger than previously found for the same data set (by 1% absolutely, or a ∼20% increase): R11 and E14[1] report total uncertainties of 3.3%[2] and 3.4% respectively. If NGC 4258 alone is used as a distance anchor, the above errors increase to 5.0% total (4.9% statistical and 1.1% systematic) for our fit, 4.1% (4.0% statistical, 1.0% systematic) from R11, and 4.7% total from E14. The difference between our errors and those found in E14 is smaller with the NGC 4258 anchor compared to when all three anchors are used – however, both are significantly larger than found in R11. For the remainder of Section 5.1.2, our discussion of errors pertains to fits with all three distance anchors.

Although it appears that the increase in our error lies in the statistical term (with the systematic term remaining the same), it is important to note the significant differences in how these terms are derived and defined in this work (given in points (i)–(iii) at the start of Section 4.7), compared to R11. Explicitly, the covariance matrices which quantify our SN Ia systematic terms directly contribute to the statistical errors in our global fits (i.e. increase the widths of PDFs) via the likelihood, while our systematic term contains the variation in parameters resulting from changing features of the fits. In comparison, the errors in each part of the calibration chain from the distance calibrators to the SNe Ia are separately given in R11, table 5. The total uncertainty is a quadrature sum of these individual terms, and the systematic variation described in R11, section 4.

The two major differences in our analysis that can potentially contribute to the increased error are the inclusion of the supernova covariance matrices, detailed in Sections 3.4 and 4.5.3, and the simultaneous fit to all parameters, described in Section 4.3.3. As outlined above, it is not possible to make a direct comparison between contributions to our error and errors given in R11, with the aim of isolating the source of the discrepancy. However, we attempt to separate the influences of the covariance matrices and simultaneous fit, replicating the quadrature sum in R11 as closely as we can below.

### 5.1.3 Analysis of increased error

First, we isolate the effect of the supernova covariance matrices alone on the size of uncertainties, by considering the error in the intercept of the SN Ia $m - z$ relation: this is $\mathcal{M}$ determined from our SN-only fit (Table 4.6 in Section 4.5.4), and is equivalent to $5a_V = 3.485 \pm 0.010$ in R11. Our error in $\mathcal{M}$ is ∼0.036, over three times larger than in R11. This error is roughly halved to 0.019 if we only consider the strictly statistical covariance matrix, i.e. $\mathbf{C}_{\rm stat}$ in Equation 3.16.[3] For the same supernova data, our statistical-only error exceeds the total error in R11. Including the SN Ia systematic covariance matrices doubles the error again. We infer that the increase in error in this analysis compared to R11 is

[1]E14 adopts the error in R11 for the SN Ia-side of the analysis.

[2]The errors reported in table 5 of R11 are: 2.9% statistical, 1.0% systematic, 3.1% total. However the final error given with all three distance anchors conservatively includes the larger statistical error associated with using two distance anchors instead of three, resulting in a total of 3.3%.

[3]Neglecting the uncertainty from the finiteness of the SALT2 training sample reduces the error slightly to 0.017, which reflects the statistical error in the observed SNe Ia only (i.e. the tridiagonal matrix $\mathbf{C}_{\rm stat,diag}$).

attributable to both the covariance matrix method of accounting for correlated SN Ia uncertainties, and to the individual systematic covariance matrices this method comprises.

Next, we attempt to replicate the quadrature summation of terms in R11, table 5 (most of which unfortunately do not have equivalent terms in our analysis) using projected uncertainties from our global fit. It is important to note that this comparison is not directly equivalent, because we are marginalising simultaneously over all nuisance parameters. With this caveat, we break down the uncertainty in the overall value of $H_0$ into three components: the uncertainties in the distance anchor, in the calibration of the SN Ia absolute magnitude $M_B$ via Cepheids, and in the measurement of the local expansion rate via SNe Ia (given in the intercept $\mathcal{M}$).[4] These can be determined separately from three disjoint data sets, as follows. The anchor distance is constrained by external data: the megamaser distance to NGC 4258 has a 0.066 mag uncertainty, corresponding to 3.0% in $H_0$. Only the low-$z$ SNe Ia are used to constrain $\mathcal{M}$ (or $5a_V$), with a 0.036 mag or 1.7% uncertainty. The calibration transfer from the Cepheids to the SNe Ia occurs in the simultaneous fit of the nearby supernova and Cepheid data[5] to Equations 4.10 and 4.11. The resultant uncertainty in $M_B$ is 0.103 mag (with only the NGC 4258 anchor) and incorporates both the uncertainty in the SN Ia-Cepheid calibration and the uncertainty in the distance anchor; thus the former is 0.079 mag or 3.6% in $H_0$.[6] In quadrature, these three terms sum to 5.0% in $H_0$ using the NGC 4258 anchor, and 4.3% with all three anchors. This decomposition, whilst approximate, indicates that a quadrature sum of uncertainties in independent components results in similar uncertainties to our simultaneous fit. Thus, the simultaneous fit does not by itself result in the increase in statistical error.

### 5.1.4 Conclusions from Z17

Our independent analysis of the Riess et al. (2011) data complements the R11 and E14 analyses in understanding the local measurement of the Hubble constant from type Ia supernovae. This work occupies the unique position of combining the precise Cepheid calibration of nearby SNe Ia (R11) with the sophisticated, thorough treatment of supernova lightcurves and systematics within a SALT2 framework (Betoule et al., 2014). In the context of the present tension in $H_0$, we presented the first blinded SN Ia-based determination of $H_0$, eliminating confirmation and other biases.

Foremost, we find that both the use of covariance matrices and the simultaneous fit of data from different rungs of a distance ladder will be important in future analyses in order to wholly account for uncertainties. Furthermore, our findings recommend more sophisticated techniques for quantifying host galaxy dependence of SN Ia magnitudes, and modelling of Malmquist bias – both of which have the potential to diminish the systematic error in $H_0$. These techniques are continually improved in supernova analyses, particularly in the pursuit of more precise measurements of dark energy, for example in the Dark Energy Survey (DES; Dark Energy Survey Collaboration et al., 2016). Meanwhile a uniform, non-targeted low-$z$ sample (e.g. the SkyMapper Transient Survey (Scalzo et al., 2017), or the Pan-STARRS Survey (Rest et al., 2014)) will simplify photometric calibration and the selection function, reducing associated uncertainties, and will avoid peculiar velocity biases from coherent flows. Adopting these changes will benefit future SN Ia-based $H_0$ measurements.

Our work in Z17, like many of the reanalyses of SH0ES (detailed in Section 5.2.2), largely supports the results found in R11 and R16. The lower central value that had resulted from the updated megamaser distance to NGC 4258 from Humphreys et al. (2013) slightly reduces the tension relative to Planck, as does E14. Meanwhile, the decreases in total (statistical and systematic) uncertainty achieved by both Planck in Planck Collaboration et al. (2016) and the supernova distance ladder in R16 have only served to magnify the disagreement, whether perceived or real, between the values. The work we have done in Z17 has been valuable for independent verification of the results found in R11 and R16. The increased errors notwithstanding, the precise agreement of the central values in R11, E14, and Z17 of the data set signifies that the discrepancy between SN Ia- and CMB-derived values continues to exist.

[4] This decomposition essentially follows equation 4 of R11.

[5] For the uncertainty in $M_B$ to be independent of the error in $\mathcal{M}$, only these data can be included.

[6] The same calculation with all three anchors results in the same number. In the setup of R11, equation 4, this SN Ia-Cepheid calibration uncertainty is the error in $m^0_{v,4258}$, which is equivalent to $M_B + \mu_{4258}$.

Our blind method for doing so, without a view of the numerical result of our analysis during the process nor the effects of our changes and decisions throughout, bolsters the extent to which our results supports the lingering tension.

## 5.2 Subsequent developments in $H_0$

The work in Z17 was intentionally applied to the well understood historical work of R11 and E14, as a proof of concept. It was our aim to extend this analysis to the updated and improved sample in Riess et al. (2016) (hereafter R16); however, it has not yet been possible to access the photometry for the nearby supernova sample required for such an analysis. In this section, we briefly describe the improvements in R16 compared to R11 which had led to the improved uncertainty, and review the developments in the literature since.

### 5.2.1 R16

The analysis in MegaSH0ES (R16) benefits from improvements in data in two obvious ways: an increase in number of calibrator galaxies containing both SNe Ia and Cepheids, and a more uniform, well-calibrated Hubble flow sample. The number of SN Ia hosts containing Cepheids increased from eight to 19, a result of of improved and more streamlined supernova searches in the late 2000s and 2010s, leading to earlier detection and follow-up, and yielding more supernovae falling into the criteria for SN Ia hosts to be targeted for Cepheids with the HST. In these 19 galaxies, the total number of Cepheids has increased from 570 to 1486 (Hoffmann et al., 2016). Given that the error in $H_0$ in R11 was limited by statistical uncertainties from the sample size, this enlargement of the nearby SN hosts has been crucial for reducing the error.

Additionally, R16 includes an additional geometric anchor of M31, with 375 Cepheids, calibrated by detached eclipsing binaries, as well as two more Cepheids in the Milky Way (Casertano et al., 2016). Although the addition of M31 anchor was tested, the best estimates of $H_0$ remained those derived from the other three anchors in combination, as in R11. Meanwhile, the Hubble flow supernovae sample, which consisted of 280 SNe Ia combined from CfA3 and LOSS in R11, had been made more uniform by replacing the LOSS sample with newer CfA4 supernovae (Hicken et al., 2012). The SN Ia photometry in R11, in the natural systems of CfA3 and LOSS instruments, had been recalibrated in Supercal (Scolnic et al., 2015), a Pan-STARRS-based effort to reduce calibration uncertainties across a wide area covered by low-redshift surveys. The Supercal photometry and replacement of LOSS (with measurements from five instruments, KAIT1–4 and Nickel) with CfA4 has improved and homogenised the photometry of the low-z SN Ia sample. We discuss Supercal, and a caveat of using photometry thus derived, further in Section 6.5.2.

### 5.2.2 Reanalyses of SH0ES

The SN Ia distance ladder approach to measuring the Hubble constant is interesting, relevant, and of value. Especially in the face of the tension in $H_0$, many parties have been compelled to re-examine the result. Here we review some of the approaches, including those on the Cepheid side, on the supernova side, or both.

Numerous reanalyses have focused on the Cepheids which form the basis of the distance ladder calibration. As we discussed in Section 4.2.1, there are many subtleties and intricacies to the measurement, outlier correction, and extinction correction for Cepheids. Their luminosity-metallicity dependence remains controversial. We have measured the effects of varying choices, such as inclusion of different geometric anchors (with different derived constraints on the metallicity dependence $Z_W$), choosing whether to include an upper limit on the periods of Cepheids, and different rejection algorithms. As discussed in the previous chapter, many of our analysis methods and choices there were informed by Efstathiou (2014), who repeated the Cepheid analysis in R11 assuming the SN Ia Hubble diagram intercept from R11. Largely motivated by concerns that the outlier rejection in R11 was exaggerating the metallicity dependence by selectively rejecting a low-luminosity subsample (Section 4.4.2), they implemented a new iterative outlier rejection algorithm and tested the effect of priors on the slope and metallicity dependence.

The analysis of Cepheids in R16 has been reanalysed by Cardona et al. (2017) and Follin & Knox (2018), who further probed the dependence of the result on the rejection of outliers and on calibration choices, respectively. The former use BHM methods (Section 3.2.3) to re-examine the Cepheid data in both the R11 and R16 analyses using Bayesian hyperparameters to down-weight portions of the data. This was motivated by (section 2.2, Cardona et al., 2017) the fact that potentially poor or arbitrary choices in outlier rejection could affect or bias results, or result in inaccurate error bars. The hyperparameters were introduced to offer an alternative to outlier rejection, and to weight data points accordingly. This approach made a difference for the R11 data set (where portions of the data were downweighted), but not for the larger R16 sample which mostly obeyed expectations for a gaussian distribution. In line with our results in Z17/Chapter 4, they found substantial dependence of results on the choice of anchor used (Section 4.4.2).

In a similar vein, Follin & Knox (2018) assesses the possibility that choices made in the calibration of Cepheids could affect the result, including a novel model-free approach to determining the distances to the SN Ia hosts. They find both insensitivity of the R16 result to calibration choices, and remarkable consistency with R16.

The supernova side of the distance ladder $H_0$ measurement has been revisited less often. In Section 4.1, we had discussed the works in Rigault et al. (2013, 2015); Jones et al. (2015)[7], investigating of the potential differences in host galaxy properties between the nearby and Hubble flow supernovae. More recently, Dhawan et al. (2018) made use of the smaller dispersion of SNe Ia in the near-infrared (NIR), referred to in Section 2.2.2, to reanalysis R16. They compiled a Hubble-flow sample from CSP and CfA supernovae with NIR photometry, together with a subset of nine out of 19 nearby SNe in R16 with sufficient NIR photometry. Using only half the number of calibrator galaxies as R16, they found $H_0 = 72.8 \pm 1.6(\text{stat}) \pm 2.7(\text{sys})\text{km}\,\text{s}^{-1}\,\text{Mpc}^{-1}$.

To our knowledge, Z17 and Feeney et al. (2018a) have been unique out of the recent reanalyses of the distance ladder measurements of $H_0$ (as presented in R11 and R16) in reanalysing both the supernovae and Cepheid data, and moreover, doing so simultaneously (Section 4.3.3). Feeney et al. (2018a) extends the BHM methodology in Cardona et al. (2017) of downweighting less trustworthy portions of data rather than outright excluding them, to the local distance ladder. Their hierarchical model includes underlying parameter distributions for the combined Cepheid and supernova data sets, in both the geometric anchors, calibrator galaxies, and Hubble flow. They interpret the difference between values in $H_0$ derived from the distance ladder and the CMB as Bayesian belief, rather than $n-\sigma$ separations between parameters distributions assumed to be gaussian. In this reframed view, they find odds of between 10:1 and 7:1 in the face of observed data from supernovae and Planck.

Other studies looked beyond Cepheids and supernovae, at the geometric distances anchoring the ladder and the set of observations as a whole. New distances to a much larger sample of Milky Way Cepheids (Casertano et al., 2017) have supported the R16 result for $H_0$, indicating that a systematic bias in one or more of the geometric anchors is unlikely. Meanwhile, Wu & Huterer (2017) modelled observations in R16 to examine the effects of sample variance using $N$-body simulations, and concluded that within a ΛCDM Universe, sample variance could not adequately explain the difference in values.

### 5.2.3 Other probes

Other modes of distance or scale measurement have been applied to the tension in the Hubble constant. The existing measurements from strong lensing (Bonvin et al., 2017) and standard siren GW170817 (Abbott et al., 2017) have beeen mentioned. In addition, Guidorzi et al. (2017) attempt to better

[7]Becker et al. 2015 is often cited on the same topic, including in R16; however, the manuscript has been withdrawn from the arXiv, citing an overlooked step in finalising the analysis.

constrain the inclination of the binary neutron star source of GW170817 using later radio and X-ray data, resulting in a slightly increased estimate of $H_0 = 74^{+12}_{-8}$ km s$^{-1}$ Mpc$^{-1}$, still consistent with both SN Ia and Planck results. Looking forward, Feeney et al. (2018b) forecasts prospects using standard sirens from binary NS mergers to resolve the tension in $H_0$, conluding that arund 50 will be necessary. Like Wu & Huterer (2017), Feeney et al. (2018b) examine sample variance in $H_0$; they also examine the inverse distance ladder (below). In the strong lensing realm, Vega-Ferrero et al. (2018) have presented the first $H_0$ measurement from a lensed supernova, SN Refsdal (Kelly et al., 2016) of $H_0 = 64^{+9}_{-11}$ km s$^{-1}$ Mpc$^{-1}$; Courbin et al. (2018) present time-delay measurements for a new quadruply-lensed quasar in DES which could be used to measure $H_0$ alongside the systems in H0LiCOW.

Looking at the early Universe, Spergel et al. (2015) had revisited earlier Planck data and found a similar value of $H_0$ to Planck Collaboration et al. (2014). Furthermore, Addison et al. (2018) demonstrate using WMAP9, South Pole Telescope (SPT) and Atacama Cosmology Telescope (ACT) data that the tension in $H_0$ remains even without Planck data (similarly to Bennett et al. (2014), as discussed in Section 4.1). However in both these analyses, the low- and high-multipole regions of the CMB power spectrum provide discordant views of $H_0$, with $\ell > 800$ showing significant tension with the R16 results, and $\ell < 800$ results much less so.

The 'inverse distance ladder' is a novel approach for calibrating SNe Ia using a distance scale from the early Universe: the BAO scale, with physical parameters set by the CMB, rather than using Cepheids. Aubourg et al. (2015) apply this method to calibrate JLA supernovae using BAO data from BOSS and CMB measurements from Planck, and measure $H_0 = 67.3 \pm 1.1$ km s$^{-1}$ Mpc$^{-1}$. This method is extended to the DES3YS data (presented in Chapter 6) in a blind analysis in Macaulay et al., (in prep.).

### 5.2.4 Theoretical approaches

Possible theoretical modifications to standard ΛCDM to reconcile the tension in $H_0$ include an increased neutrino effective number (the existence of dark radiation), and/or a more negative dark energy equation-of-state parameter $w$ at late times. Di Valentino et al. (2016) explore these scenarios in a higher-dimensional parameter space, with their findings supporting phantom dark energy with $w \sim -1.3$, while Wyman et al. (2014); Dvorkin et al. (2014); Leistedt et al. (2014) focus on the implications of an additional massive sterile neutrino species. Meanwhile, Bernal et al. (2016) examine the model dependence of the Universe's distance scale (anchored by $H_0$ and by the scale $r_S$ of the sound horizon at radiation drag, at late and early times respectively) by reconstructing its expansion history with minimal cosmological assumptions.[8] They conclude that the tension in $H_0$ translates to a mismatch in the normalisations provided by $H_0$ and $r_s$ at two opposite ends of the distance ladder.

## 5.3 Future directions

In the past five years, the discrepancies between local and early value of $H_0$ have attracted much interest and attention since the Planck 2013 (Planck Collaboration et al., 2014) data release revealed a substantially lower value than SH0ES (R11). Since then, the SH0ES data set has grown substantially, and numerous reanalyses of both measurements (Section 5.2.2) have taken place. While small oscillations have occurred at both ends, these subsequent works have for the most part reinforced the two clusters of values: the local distance ladder based measurements at ∼73 km s$^{-1}$ Mpc$^{-1}$, and the CMB measurements at ∼68 km s$^{-1}$ Mpc$^{-1}$. Of particular interest are the measurements from modes other than SNe or the CMB. The strong lensing measurement from H0LiCOW, like SNe Ia on a distance ladder, is based on distance measurements in the local Universe and is in good agreement with supernovae at $H_0 = 71.9^{+2.4}_{-3.0}$ km s$^{-1}$ Mpc$^{-1}$ (Bonvin et al., 2017). The inverse distance ladder method in Aubourg et al. (2015), calibrated on the absolute distance scale in the early Universe (via $r_s$, the sound horizon at radiation drag) agrees with measurements from the CMB. These probes have played a

[8]This is possible as the combination of SNe Ia and BAO as probes constrains the product $r_S H_0$ in a model-independent way.

signifcant role in corroborating measurements at late and early times.

As we have seen in Section 5.2.2, the numerous reanalyses of the distance ladder Hubble constant measurements in R11 and R16, including ours in Z17, have steadfastly supported their central numerical results. Moreover, other than numerical agreement, the analyses all show reasonable robustness to choices in the analysis. These results, as well as the new geometric distances to Galactic Cepheids, suggest that a systematic uncertainty in some step of the distance ladder is unlikely. The tension in the numerical value of $H_0$ has persisted over five years. The agreement from external probes and consistencies in reanalyses have lent support to a real physical or cosmological origin of the tension in $H_0$.

Whichever of these scenarios we are inclined to believe is true, whether there is something odd in cosmology, whether there is some undetected systematic in one of the measurements, or the two measurements are both in some ways consistent with $\Lambda$CDM, or something in between, it is (at least) as clear now as it had been in the past several years that it is crucial that our unconscious biases, or other human factors, do not impact our numerical results, which continue to have profound consequences. The growing trend of blinding cosmological analyses in recent years (Section 1.5) remains as important now as ever before.

Chapter 6

# Dark Energy Survey

This chapter concerns the supernova cosmology analysis with the Dark Energy Survey 3-year Spectroscopic Supernova (DES3YS) sample, focusing on the established maximal likelihood 'JLA-like' approach as part of the Dark Energy Survey Supernova working group (DES SNWG). For an overview of DES, see Section 6.1. Following an introduction to context for the analysis in Section 6.1, I describe the JLA-like analysis in Section 6.1.2 and present the equations for the physical Universe's expansion in Section 6.2. Then I present an overview of the data (Section 6.3), followed by equations and methods for preparing the data and cosmology fitting in Sections 6.4 and 6.5, respectively. My main contributions to the DES SNWG-wide analysis have been developing the computation of systematics (Section 6.5.2) and the cosmology fitting code (Section 6.5.1) for the JLA-like approach. The work described in this section encompasses contributions from several members of the DES SNWG. In particular, I have developed the code for this analysis in collaboration with Chris Lidman; Tamara Davis will continue to run this analysis on updated DES3YS data. I reference papers in preparation for parts of the project where relevant, and otherwise identify the DES SNWG member who was responsible for doing the work for a specific part of the analysis. Section 6.6 contains results, including comparisons of results to other analysis methods on the same data and a breakdown of sources of errors, and discussions. The work in this chapter forms portions of the wider alphabetically-authored DES3YS cosmological results paper Dark Energy Survey Collaboration, in prep..

## 6.1 Introduction

Type Ia supernovae were pivotal in discovering cosmic acceleration, and remain a leading method in the Dark Energy Survey. Earlier, Section 1.2.1 detailed their significance in modern cosmology, while Section 2.4.5 pinpointed the significance of DES SNe in the context of wider supernova surveys. We refer the reader to the above sections for introductions to those areas. In this section, we will examine DES supernovae in more detail and motivate the JLA-like analysis.

### 6.1.1 DES supernovae

The final DES SN Ia sample is expected to number ∼2500, and will consist predominantly of photometrically classified supernovae. A smaller proportion of supernovae will be spectroscopically classified. Current analyses within the DES SNWG are focused on the DES 3-year spectroscopic (DES3YS) SN Ia sample. The analyses of this sample, including the JLA-like method presented in this chapter and those in Brout et al., (in prep.) and Hinton et al., (in prep.), are intended as an intermediate stepping stone on the way to the final 5-year photometric sample, allowing the development and verification of techniques that will be used then. Constraints from the DES3YS sample were expected to be comparable with constraints in the Joint Lightcurve Analysis (JLA) of SNLS-SDSS supernovae (Betoule et al., 2014, hereafter B14) and have been largely so (see preliminary results in Sections 6.6 and 6.7). JLA follows a series of developments focusing on improving and enlarging collections of SNe Ia for cosmological studies, including notably the Supernova Legacy Survey (SNLS) at high redshift, the SDSS supernova survey at intermediate redshifts, and a collection of low-redshift surveys including the Harvard-Smithsonian Center for Astrophysics (CfA) supernova survey, Carnegie Supernova Project (CSP), and Lick Observatory Supernova Survey (LOSS). A major lesson from Conley et al. (SNLS; 2011); Guy et al. (SNLS; 2010); Sullivan et al. (SNLS; 2011) has been that uncertainties in photometric calibration significantly dominate the systematic error budget. JLA originated as an effort to reduce the impact of calibration (Betoule et al., 2013), while combining the aforementioned notable surveys over different redshift ranges. This chapter focuses on the JLA-like method, based on methods and in particular techniques for quantifying systematic errors developed in SNLS and JLA.

While DECam will obtain photometry for all supernovae to determine distances for the SN Ia Hubble diagram, and photometric redshifts are possible, (e.g. Bolzonella et al., 2000; Salvato et al., 2009; Benítez, 2000; Laurino et al., 2011) spectroscopic redshifts for DES supernovae and their host galaxies will come from a variety of external spectroscopic facilities, the primary one being the Australian Dark Energy Survey (OzDES; Yuan et al., 2015; Childress et al., 2017). OzDES and the other sources of spectroscopy are described in Sections 2.4.5 and 6.3.2. In the first three years of DES, $\sim$250 SNe Ia were spectroscopically confirmed by OzDES and other sources of spectroscopy D'Andrea et al., (in prep.); around 2500 host galaxy redshifts have been observed by OzDES.

As discussed in Sections 1.2.1 and 2.4.5, DES is necessary for improving the statistical constraining power at high redshift, and will enable measurements of cosmological parameters to a new degree of accuracy. These parameters and SN Ia-specific parameters are presented in Section 6.2, and notably include the densities[1] $\Omega_m, \Omega_{\mathrm{DE}}$ of matter (both baryonic and dark) and dark energy,[2] the equation-of-state parameter $w$ for dark energy, and $H_0$. The Hubble constant $H_0$ – the focus of Chapters 4 and 5 and discussed in more detail there – represents the Universe's current expansion rate. This is degenerate with the absolute SN Ia peak magnitude, and is often written as the dimensionless Hubble parameter $h := H_0/100\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$.

In particular, DES sets out to affirm or refute, with unprecedented precision, the default scenario where dark energy takes the form of a cosmological constant with equation-of-state parameter $w = -1$. Finally, DES will pave the way for the next generation of large-scale cosmology experiments starting with the Large Synoptic Survey Telescope (LSST).

### 6.1.2 JLA-like cosmological analysis

There are three analysis methods within the DES SNWG for performing cosmological analyses of the DES3YS sample: the BEAMS with Bias Corrections Brout et al., (in prep.), the Bayesian Hierarchical Model (BHM) method Hinton et al., (in prep.), and the JLA-like analysis, which we focus on in this chapter.

We present a maximal-likelihood (or $\chi^2$-minimising) fit of the DES3YS sample supplemented by a hybrid sample of low-redshift supernovae from CfA3/4 and CSP. We closely follow methods in JLA, which estimate and correct for systematics, namely Malmquist bias, host galaxy mass dependence, and peculiar velocities (discussed in Section 6.4.3), and use covariance matrices to account for correlated uncertainties: in these corrections, in statistical uncertainties, and in other systematics (specifically photometric calibration, scatter model dependence, non-Ia contamination, Milky Way extinctions; see Section 6.5.2). The corrected SN Ia distance moduli and redshifts are fitted for parameters of the $w$CDM or $\Lambda$CDM cosmological models according to equations in Section 6.2, using standard Markov Chain Monte Carlo methods for parameter estimation, described in Section 6.5.1.

In short, the analysis of the DES3YS data set using the JLA-like method is a proof of concept, which demonstrates what is possible using currently available DES data and existing techniques. It allows comparisons (i) between established JLA-like techniques and the newer methods of BHM and BBC, and (ii) between the present DES sample and the existing JLA data set. SNLS, and by extension JLA and subsequently Pantheon, remains the benchmark for high-redshift SN Ia surveys until DES is

[1] Written as fractions of the Universe's critical density (Equation 6.5).

[2] The latter is often written $\Omega_\Lambda$ for a cosmological constant, but in a generalised situation where we do not assume that dark energy takes this form, we write it as a more general $\Omega_{\mathrm{DE}}$.

complete; using the method therein allows a faithful comparison of the intermediate DES data, which is now producing results which are close to competitive with JLA.

The sizes of the SN Ia samples, relative to JLA, are shown in Table 6.1. The DES spectrosopic sample is comparable in size to the SNLS data set; however it is more similar in redshift to the combined SDSS-SNLS high-$z$ sample in JLA. Meanwhile the addition of CfA4 and CSP-II supernovae results in a low-$z$ sample that is mostly independent of JLA (with only 68 CfA3 supernovae in common). The new CfA-CSP combined low-$z$ sample is larger than in JLA and more homogeneous, originating from only three telescopes and (broadly classed) magnitude systems: 4Shooter2, KeplerCam, and Swope.

This method uses covariance matrices to quantify correlated statistical and systematic error terms between supernovae, which are included in the calculation of the $\chi^2$ figure-of-merit and likelihood. As shown in Zhang et al. (2017) (Chapter 4), these are powerful in accounting for correlated systematic errors. The JLA-like analysis sets a baseline to be improved upon with greater numbers (by up to an order of magnitude in the DES 5-year photometric sample) and hence statistics enabled by photometric classification and with improved photometry and calibration, and possible new statistical analysis techniques.

Furthermore, being able to compare different analysis methods − including the JLA-like − on the DES3YS data gives us an additional comparison of potential methods to be applied to future DES SN Ia samples. Thus, we place an emphasis on using an identical sample (with identical cuts, Section 6.3.4) to the BBC method. Final cosmological constraints from DES will necessarily improve on those from the DES3YS sample, and will likely use some of the same analysis methods, as well as some new ones.

## 6.2 Measuring dark energy

In this section, we outline the framework and equations for describing the physical Universe, with the intent of demonstrating where dark energy and cosmological parameters fit in.

We are able to measure cosmology from SN Ia distances (as a function of redshift) because these distances are reflective of the underlying cosmology. This is Bayes' theorem paraphrased: we can arrive at the posterior beliefs about cosmological parameters given the data we can observe, by using the data and the likelihoods we compute from the theory. Here, we write out the theory, starting from Friedmann's equations for an isotropic and homogeneous Universe obeying General Relativity, and finishing at the theoretical distance modulus (a function of redshift and cosmology) to be compared to the observed distance modulus from the data.

The generalised isotropic homogeneous Universe is described by Friedmann's equations:

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} - \frac{kc^2}{a^2} \tag{6.1}$$

$$\frac{\ddot{a}}{a} = -4\pi G(\rho^2 + 3p^2), \tag{6.2}$$

along with the Friedmann-Lemaître-Robertson-Walker (FLRW) metric. In the above, $a$ is a dimensionless scale factor and $k$ is the spatial curvature parameter, which takes values $+1$, 0, or $-1$ for a closed, flat, or open Universe respectively. Overdots and double overdots represent derivatives with respect to the time parameter $t$. Written in hyperspherical coordinates with the dimensionless comoving coordinate $\chi$, the FLRW metric is

$$ds^2 = dt^2 - R^2(t)\left(d\chi^2 + S_k^2(\chi)(d\theta^2 + \sin^2\theta\ d\phi)\right) \tag{6.3}$$

where $R$ is the scale factor with units of length (with $a = \frac{R}{R_0}$ with $R_0$ representing the scale factor at present time) and

$$S_k(x) = \begin{cases} \sin x & k = 1, \\ x & k = 0, \\ \sinh x & k = -1, \end{cases} \tag{6.4}$$

The density $\rho$ in Friedmann's equations has contributions from matter, radiation, and dark energy components, which evolve with time as the Universe expands. These densities are more commonly written as relative densities, compared to the critical density at which the Universe is spatially flat. The value of the critical density at the present time can be derived using elementary mechanics as

$$\rho_{\text{crit}} = \frac{3H_0^2}{8\pi G}. \tag{6.5}$$

At late times the energy density due to radiation is effectively zero, so we can assume $\Omega_m$ and $\Omega_{\text{DE}}$ are the only contributions.

Although we only allow a spatially flat Universe (with curvature $k = 0$) in our cosmology fits, the following equations allow a non-flat Universe with $k = \pm 1$. We use $\kappa_0$ to denote scalar curvature in the units of density, with $\kappa_0 \equiv \left(\frac{c}{aH_0}\right)^2$. The present length scale is related to the curvature by:

$$R_0 = \frac{c}{H_0 |\kappa_0|^{\frac{1}{2}}} \tag{6.6}$$

Combining Equations 6.1 and 6.5 gives

$$\kappa_0 = \frac{1}{H_0^2}\left(\frac{8\pi G \sum_i \rho_i}{3} - H_0^2\right) \tag{6.7}$$

$$= \sum_i \Omega_i - 1 \tag{6.8}$$

which is sometimes written as $-\Omega_k$, or the (negative) relative energy density attributable to spatial curvature. In a flat Universe, the energy densities from to all contributions sum to 1, and so $k$ and $\kappa$ are both zero.

The energy density of each term which contributes to the total energy evolves differently with time. For example, normal matter has density which scales as inverse volume or inverse length cubed, whilst radiation density scales as the inverse fourth power of length. Meanwhile the energy density of a cosmological constant does not change, hence its name. More generally this scaling of the density $\rho_i$ is parametrised by the equation-of-state parameters $w_i$, with:

$$\rho_i \propto V^{-(1+w_i)} \propto a^{-3(1+w_i)} \propto (1+z)^{3(1+w_i)}, \tag{6.9}$$

where $w_i$ takes values 0 for matter, $\frac{1}{3}$ for radiation, and $-1$ for a cosmological constant, respectively.

To compute the distance corresponding to a given redshift $z$, we integrate over redshifts from zero (the present time) to $z$. The dimensionless comoving coordinate is given by the integral

$$\chi(z) = \frac{c}{R_0}\int_0^z \frac{dz'}{H(z')} \tag{6.10}$$

$$\tag{6.11}$$

where the Hubble parameter $H := \frac{\dot{a}}{a}$ is the evolving expansion rate. The comoving and luminosity distances are related to $\chi$ by:

$$D(z) = R_0\chi(z) \tag{6.12}$$
$$D_L(z, z_h) = (1 + z_h)R_0S_k(\chi(z)) \tag{6.13}$$

Unless otherwise specified, $z$ here and hereafter refers to the cosmological redshift in the CMB frame ($\bar{z}$ in Sections 3.1.3 and 4.3.1); for low-$z$ supernovae in the DES3YS sample, this has been corrected for peculiar velocities according to Section 3.1.3. Note that the luminosity distance contains both the cosmological redshift $z$ and the heliocentric redshift $z_h$. This is because the factors affecting the $(1 + z_h)$ prefactor (redshifting and beaming) depend on the total relative velocities, whereas the cosmological distance only depends on $z$, the redshift due to expansion (Calcino & Davis, 2017, Section A).

Using the common substitution $E(z) \equiv H(z)/H_0$, Equation 6.1 gives

$$E(z) = \left[\sum_i \Omega_i(1 + z)^{3(1+w_i)} - \kappa_0(1 + z)^2\right]^{1/2}, \tag{6.14}$$

with terms for each density contribution (matter, radiation, and dark energy), and spatial curvature.

For a flat Universe with $k = 0$, where dark energy (with constant equation-of-state $w$) has energy density $\Omega_{\mathrm{DE}} = 1 - \Omega_m$, the luminosity distance reduces to (using Equation 6.14)

$$D_L(z, z_h) = \frac{c(1 + z_h)}{H_0}\int_0^z \frac{dz'}{\sqrt{(\Omega_m(1 + z')^3 + \Omega_{\mathrm{DE}})}}. \tag{6.15}$$

Then, the predicted distance modulus, as a function of redshift and cosmology, is

$$\mu_{\mathrm{mod}}(z, z_h, \Omega_m, w) = 5\log_{10} D_L(z, z_h, \Omega_m, \Omega_{\mathrm{DE}}, w)/(\mathrm{Mpc}) + 25 \tag{6.16}$$

## 6.3 Data

This section describes the data used in the DES3YS analyses, including correction terms in the JLA-like method. This includes details of the photometry pipeline (including observations and calibration) and collection of redshifts. Brout et al., (in prep.) will present the Scene Modelling Photometry (SMP) method used for obtaining lightcurves, and Lasker et al., (in prep.) will describe the DES Forward Global Calibration Method (FGCM) applied to the DES SNe Ia. These important parts of the analysis chain will be summarised in Sections 6.4.1 and 6.4.2 below. The spectroscopic redshifts follow, then correction terms for SN Ia distances and redshifts following JLA.

### 6.3.1 Observations

The DES overview paper (Dark Energy Survey Collaboration et al., 2016) describes observations for the survey as a whole, which use the Dark Energy Camera (DECam) frontend on the CTIO Blanco telescope, as part of the wider DES campaign, in B semesters from August 2013 to February 2019. The image processing pipeline is presented in Morganson et al. (2018).

The DES3YS sample, which this analysis and chapter focus on, will be outlined in a forthcoming paper D'Andrea et al., (in prep.). The 249 SNe Ia in the DES3YS sample, at redshift $0.017 < z < 0.85$, are supplemented by a hybrid low-redshift sample of 312 SNe Ia, described in Section 6.3.3. All of these supernovae have been spectroscopically classified as type Ia. There are two levels of confidence for the spectroscopic classification of SNe Ia in DES: supernova are labelled 'Ia' (223 SNe), and 'Ia*' (or probable Ias, numbering 26 SNe), based on spectroscopic features (specifically, silicon, iron, and sulphur absorption features). Both sets are included in the cosmology fit, and the systematic error associated with uncertain classification or possible misclassification is discussed in Section 3.4.7.

Table 6.1 Contributions from various surveys to the SN sample in this work, compared to in JLA (table 8, Betoule et al., 2014).

| source | JLA | this work |
|---|---|---|
| CSP | 13 | 85[a] |
| CfA3 | 55 | 185 (43) |
| CfA4 | - | 94 (9) |
| other low-$z$ | 28 | - |
| SDSS | 374 | - |
| high-$z$ | 248[b] | 249[c] |
| **Total** | 740 | 562 |

[a] Including 52 SNe Ia which are also in CfA3/4, in brackets below. See Section 6.3.3 for treatment of these supernovae.
[b] Comprising 239 SNe from SNLS and 9 from HST (table 8, Betoule et al., 2014).
[c] DES 3-year spectroscopic sample.

The DES3YS and low-redshift supernova samples together go through a set of cuts (Section 6.3.4) which are applied consistently to both sets. The 338 supernovae (210 DES and 128 low-redshift) that pass cuts are placed on the same Hubble diagram and fit together for cosmological parameters (Section 6.5). Table 6.1 lists the contributions to the total SN Ia sample studied in this work as compared to JLA, from the DES3YS sample and three low-$z$ surveys making up the low-$z$ sample used in this analysis.

### 6.3.2 Redshifts

Redshifts for placing SNe Ia on a Hubble diagram can come from spectra of the supernova host galaxy, or the supernova itself. If a supernova has an available spectrum, the spectrum can be used to spectroscopically classify the supernova as well as to determine its redshift. Several techniques were used to classify and obtain redshifts for the transients making up DES3YS, including: `Superfit`, SNID, and `DASH`. These techniques were described in Section 2.1.2.

Practical considerations mean that for most DES supernovae, only spectra of their host galaxies are obtained: this removes the logistical constraint of only having the supernova bright enough for a short window of opportunity. The redshift of the supernova is inferred from the spectrum of the host, but the supernova is classified photometrically (see Section 3.3.3). This will be true for most supernovae in the final DES 5-year sample. In contrast, the DES3YS supernovae are all 'spectroscopic' in the sense that a spectrum is of the supernovae itself. Moreover, the later DES 'photometric' sample is not to be confused with the technique of photometric redshifts, where relative brightnesses of objects in different passbands are used to estimate the redshift. This is currently not considered for DES supernovae, because the redshifts obtained lack the precision required for Hubble diagram cosmology, but the technique is used for other DES science (with validation on OzDES spectra).

For DES objects with spectra of the supernova and of only the host galaxy, the inferred redshifts are compared to ensure agreement. This additional check will not be possible for future photometric samples; however agreement at this stage allows validation of the same methods used later for host galaxy redshifts, and allows the estimation of the probability and frequency of host galaxy mismatch, and the associated systematic uncertainty. This can be caused by supernovae having high velocity (and are displaced from the host and/or assigned the wrong host), or are located along the same line of sight of another galaxy attributed as the host, as discussed in e.g. Gupta et al. (2016).

The measurement errors in redshifts from supernova spectra are around $\sigma_z \sim 0.01$, which is larger than uncertainties in redshifts from host galaxy spectral features, and around the same order of magnitude as uncertainties in magnitude. We limit our treatment of errors to the so-called dependent variable, the distance modulus, only (see Section 6.5.1 for further discussion of errors in both variables) so we propagate redshift errors to magnitude or distance modulus as in Equation 6.22. Effects of potential systematic errors in redshift, and in particular biases these will induce in cosmological results, are studied in Andersen (2018). Uncertainties in redshift due to sources other than measurement error include peculiar velocity uncertainties, included in Section 3.4.

**OzDES**

While all photometry for the DES3YS come from DECam, there are several sources for spectroscopy of supernovae, or more often, their host galaxies. The primary source is OzDES (Yuan et al., 2015; Childress et al., 2017) (introduced in Section 2.4.5, an Australian-led spectroscopic programmme to shadow DES fields on the 4-m Anglo-Australian Telescope (AAT). Other telescopes which contribute to the DES spectroscopic task force are Southern African Large Telescope (SALT), the Magellan Telescopes, Gran Telescopio Canarias (GTC), Gemini South, and the Very Large Telescope (VLT). These telescopes are larger than the AAT, with mirror diameters from 6 to 10 metres, and are significantly represented in the DES3YS sample where the presence of spectra of supernova (rather than hosts) necessarily requires (i) greater light-collecting area, and (ii) access to Target of Opportunity (ToO) time on other telescopes during periods outside of the pre-scheduled OzDES observing runs. Depending on timing, spectra of supernovae may be taken after significant decline in brightness, which also explains the greater representation of larger telescopes in the DES3YS sample. In particular, VLT contributions to redshift are notable at $z > 0.5$, while OzDES dominates below this.

OzDES targets host galaxies of DES transients, removing the strict constraints (in terms of time, scheduling, and logistics) of targeting live transients from the bulk of the supernovae which will be used for science. However, in some cases (where a supernova candidate is thought to be bright enough to be targeted at the AAT at time of observation, it is possible for OzDES to take spectra of live transients (including Type Ia supernovae). We emphasise again that all supernovae in the DES3YS sample have spectra (from OzDES or otherwise) of the supernova itself. Unlike the DES3YS, for which we have a spectrum of each supernova, the final DES photometric sample will predominantly consist of supernovae for which only host-galaxy spectra are available.

### 6.3.3 Low-redshift samples

The DES3YS supernovae, with redshift range $0.017 < z < 0.85$, are supplemented with a collection of 312 low-redshift supernovae from CfA3, CfA4 and CSP, at redshifts $z < 0.09$. The Harvard-Smithsonian Center for Astrophysics supernova survey operated on the 1.2-m telescope at the Fred Lawrence Whipple Observatory (FLWO). During 2001–2004 the 4Shooter2 imager with $UBVri$ filters was used, and from 2004–2011 Keplercam with $UBVRI$ filters was used. The third data release (CfA3; Hicken et al., 2009a) comprised 185 SNe Ia, while the fourth (CfA4; Hicken et al., 2012) added 94 SNe Ia. The later stages of the CfA supernova survey continued on work presented in Riess et al. (1999) and Jha et al. (2006, CfA2), concentrating on following up SNe Ia at low redshift discovered by various sources and obtaining multi-colour lightcurves. These sources of discovery included the SDSS supernova survey, Li et al. (LOSS; 2000) (until 2003, this was LOTOSS, the Lick Observatory and Tenagra Observatory Supernova Survey), which built on the Beijing Astronomical Observatory Supernova Survey (BAOSS; Qiu et al., 1999; Li et al., 1999), and also from amateur astronomers. The Carnegie Supernova Project (CSP; Contreras et al., 2010; Stritzinger et al., 2011) added near-infrared (NIR) photometry. CSP operated from 2004–2009 at Las Campanas Observatory. Optical photometry came from the Swope 1-m telescope in $ugriBV$ bands, with the SITe3 camera. Near-infrared photometry came from the Wide-Field Infrared Camera on the du Pont 2.5-m telescope at first, and later was replaced by the NIR imager RetroCam on the Swope 1-m. Further details of observations in CfA3/4 and CSP were described in the above references.

There are 52 supernovae in both CSP and either CfA3 (43) or CfA4 (9). For these supernovae, we concatenate the lightcurves of both sources and fit using SALT2. Tests performed in Zhang et al. (2017) have shown that resultant lightcurve fit results generally are an average of the two sources of photometry, weighted by number and quality (in terms of sizes of errors) of data points; given that there is no reason to prefer one photometry source over the other and they complement each other, we use

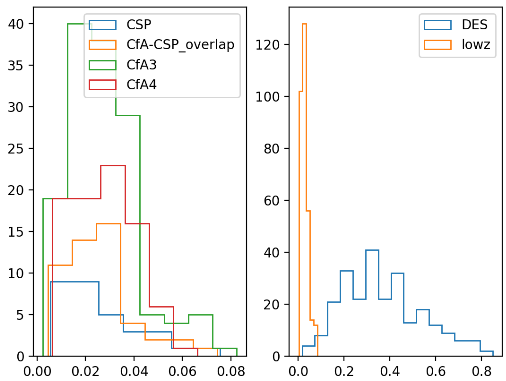


Figure 6.1 Redshift histograms for supernovae before cuts. Top: DES3YS with all low-redshift together. Bottom: Individual low-redshift subsamples.

both sources where available.

We use natural photometry in the systems used for observations, along with their filter transmissions, for lightcurve fitting with Guy et al. (SALT2; 2007). The version of photometry we use is the most recent in PS1 (and later in Pantheon) with the revisited calibration there (described in Scolnic et al., 2015, 2018). These have improved on the natural system photometry in the original data releases (Hicken et al., 2009a, 2012; Contreras et al., 2010), as the offsets between different survey calibrations has been minimised to reduce estimated systematic errors, using Supercal (discussed in Section 6.5.2).

The low-redshift anchor of the Hubble diagram is important for cosmology because it sets the zero point of the Hubble diagram. As found in JLA, the systematic in the calibration of the low-redshift supernova samples is at present a dominant contribution to the error budget. This can largely be put down to the heterogeneity of the data sets. Since SNLS and JLA, the expanded CfA and CSP sets have allowed more homogeneous low-$z$ samples of over 100 supernovae, which was not possible for JLA (table 8 Betoule et al., 2014). Although our low-$z$ sample is at most comparable in size to JLA, it is more homogeneous. For the later 5-year DES photometric analysis, future surveys (including low-$z$ SNe Ia) such as PS1 (Rest et al., 2014) and Foundation (Foley et al., 2018) will offer improved uniformity hence better understood selection functions and more precise calibration.

### 6.3.4 Cuts

The duration of the 249 SNe Ia in the DES3YS sample will be described in detail in D'Andrea et al., (in prep.). We apply further cuts to this sample, based on principles of removing supernovae which are judged as too adversely affected by systematics, or outside the bounds of spectroscopically typical Ias used to train SALT2. We follow these principals with cuts based on the following criteria, applied consistently to the DES and low-redshift samples. These cuts are largely informed by standards set in JLA (based on SNLS), and are echoed in PS1 as well in numerous low-redshift surveys (CfA, LOSS, CSP).

The cuts are used to exclude outliers in stretch and colour, only retaining SNe Ia with $|X_1| < 3$ and $|C| < 0.3$ respectively. We remove SNe Ia with lightcurves that are poorly constrained in the time domain, with the constraints $\sigma_{X_1} < 0.5$ and with the day of maximum constrained to within half a day. In the low-redshift sample, we also eliminate supernovae in locations significantly affected by Milky Way extinction and with redshifts which are too low, i.e. are dominated by peculiar velocities rather than the Hubble flow: thus we restrict the low-redshift sample to supernovae with extinction $E(B-V) < 0.2$ and redshift $z > 0.01$.

Our cuts described above result in 238 and 207 DES and low-$z$ supernovae, respectively. However, the other cosmology fitting methods (BHM and BBC for the same DES3YS data set) implement additional cuts which are specific to the Kessler et al. (SNANA; 2009b) analysis package, such as the lightcurve goodness-of-fit metric. These additional cuts leave 210 DES and 128 low-$z$ supernovae remaining, notably cutting a larger fraction of the low-redshift supernovae. We identically adopt these cuts, which were essentially frozen in place by the unblinding event in December 2017, in order to make a meaningful comparison to the other fitting methods. Thus, there will be minor differences between preliminary results presented here, and later published DES 3-year results.

## 6.4 Methods

This section contains methods to prepare the data for cosmology fitting (Section 6.5): photometric calibration, photometry, and accounting for correction terms (previously introduced in Section 2.3.1) including selection effects.

### 6.4.1 Forward Global Calibration Method

From its beginning DES was envisioned with SN Ia acting as a major probe for dark energy, with the knowledge that accurate calibration was paramount to making precise cosmological measurements. Thus, photometric calibration had always been a focus. The current method of calibration, the Forward Global Calibration Method (FGCM) (Burke et al., 2018), sets out to improve on existing calibration methods.

In recent years, photometric calibration of wide-field surveys has depended on calibrating a network of standard stars based on the CALSPEC network, consisting of a chain non-variable stars resting on the three HST DA white dwarfs, as described in Section 3.1.1.

Surveys which have used this approach included SDSS (Padmanabhan et al., 2008), SNLS (Regnault et al., 2009), and PS1 (Rest et al., 2014). The first iteration of the DES calibration, Drlica-Wagner et al. (Y1A1 Gold; 2018), followed this method; in the three-year data release (Y3A1) the PGCM (Photometric Global Calibrations Model) was used as a cross-check for FGCM.

The model-based approach in FGCM addresses limitations in the above calibration method, primarily the approximate nature of the spectral energy distributions (SEDs) of objects used to calculate them. A magnitude system is typically calibrated by a set of non-variable stars and extended to other objects. Then extrapolating thus calibrated magnitude systems to supernovae, which have significantly different SEDs, carries some uncertainty which must be corrected for or taken into account. In particular, FGCM eliminates the uncertainty associated with varying SEDs by including functions called 'chromatic corrections', which correct broadband magnitudes given an object's SED.

In summary, the FGCM method fits a set of parameters nightly to predict zero points in the DES standard system, and additionally computes colour correction terms for an estimated or exact SED. For each exposure and CCD, the model forward computes (for a set of observed stars) the number of photons at top of the atmosphere, compared to photons predicted to be detected in the camera. The model parameters, consisting of eight atmospheric terms (for each night of observation) and two optical terms (for each time the mirrors are washed), are solved for to model the responses of the DECam passbands. Repeated observations of non-variable stars in the DES footprint are used to predict the number of photons at the top of the atmosphere are predicted iteratively, to determine these model parameters.

Thus, FGCM establishes a typical and representative DES standard (referred to as just 'standard' for the remainder of this section), which comprises a set of filter passbands from the top of atmosphere to the detector, as well as a network of stars (acting as tertiary standards) which cover the DES footprint and are used to calibrate the science fields, determining observing conditions for individual exposures. The FGCM calibration is tied to the HST system (for absolute magnitudes) via the CALSPEC standard star C26202. This process occurs outside of FGCM and has its own error contributions (Section 6.5.2). For each exposure and CCD the FGCM computes a zero point. An additional AB offset (determined from observations of C26202) is added to the FGCM zero point, to place magnitudes on an AB scale (Burke et al., 2018, equation 40).

Chromatic corrections are applied to give differential magnitudes in the DECam system, given either an object's broadband colours (an approximation) or precise SED: the 'full integrated chromatic correction', necessary for supernovae. To first order, the chromatic correction integral (Burke et al., 2018, equations 6 and 14)is computed for each CCD and exposure.

Actual passband throughputs evolve with time and position, and the difference from the standard system are measured. The calibration is adjusted in real time to account for deviations from DES standard for each exposure and CCD, comprising both the FGCM fit for the atmosphere and mirror model, and DECal: in situ monitoring, which measures the variation of optical properties of the survey system and atmospheric conditions in real time including both spatial and temporal variations of the instrument throughput, similar to other in situ systems such as SNDICE (Barrelet & Juramy, 2008; Regnault et al., 2016; Barrelet, 2016).

The release products of FGCM are the DES standard system, including instrument and atmosphere, in the form of transmission curves and zero points. The end goal of photometric calibration with FGCM is the ability to determine the magnitude zero points of any observation in standard passbands, given the SED (or an approximation of it in broadband colours), making use of atmospheric models and DECal data, as well as the FGCM model itself. This differs from previous broadband photometry methods, which have end products of magnitude zero points for the instrument passbands, which along with instrument transmissions allow transformations into other magnitude systems.

The most notable improvement from previous methods is the elimination (to a great extent) of the ambiguity associated with broadband photometry (from matrix transformations), by greatly improving SED estimates. Current performance is to a level of random residuals around 6 mmag and uniformity to within $\sigma = 7$ mmag (Burke et al., 2018). The computation of associated systematic uncertainties is described in Section 6.5.2. Further details of FGCM are given in Burke et al. (2018), and details of application to differential photometry for SN Ia observations are presented in Lasker et al., (in prep.). Next, we detail uncertainties in both FGCM and the calibration transfer from the HST CALSPEC scale to FGCM, which go into our computations of $\mathbf{C}_{\rm cal}$.

#### FGCM calibration uncertainties

We describe the following sources of uncertainty in the FGCM calibration, and their contributions to $\mathbf{C}_{\kappa}$. The two parts of this are the determination of the FGCM zero points, and the AB offsets. In addition, we consider the large scale spatial variations of FGCM. So far, these estimates neglect an uncertainty in the chromatic correction, which is not yet modelled well enough.

Unless otherwise specified, the following terms are added to diagonal elements of $\mathbf{C}_{\kappa}$ which correspond to DECam zero points.

- **Zero point uncertainty per exposure:**
  The accuracy of FGCM zero points (for a given CCD and exposure) depends on the number of stars used to measure the ZP in Scene Modelling Photometry (Section 6.4.2; used to then estimate SN Ia magnitudes), and the ZP accuracy of a single star's measurement. The latter is informed by the scatter of tertiary stars in FGCM, 5–6 mmag (figure 11, Burke et al., 2018). There are sufficiently many stars used to determine the ZP in SMP that we can take the zero point uncertainty to be 1 mmag.

- **AB offset uncertainty from filter curve:**
  The CALSPEC standard C26202 is used to determine the AB offsets for the DECam filters using

DES synthetic magnitudes of C26202 (using DES filters) and the DES standard magnitude of C26202 to transform. For DES instruments, these have errors of $5\mathring{A}$. These uncertainties are also correlated with zero points, resulting in off-diagonal terms in $\mathbf{C}_\kappa$.

- **AB offset uncertainty from C26202 magnitude:**
The uncertainty in a single measurement of C26202 in the DECam standard system is given by the residual scatter of 5 mmag in FGCM (Burke et al., 2018), divided by the square root of the number of observations of C26202 in that passband.

- **Large scale spatial variations of FGCM:**
The uniformity of the FGCM calibration over the sky is assessed by comparison to GAIA observations above the atmosphere, in Burke et al. (2018, section 5.3). In particular, correlated measurement error in observations that were made close in time can imprint on the large scale uniformity. The difference in observed GAIA DR1 $G$ magnitudes and those predicted by transforming DES $gri$ bands are binned in HEALPIX with NSIDE= 256.[3] The binned residuals are distributed as found to resemble a gaussian distribution with $\sigma_{\rm res} = 6.6$ mmag. This mean uncertainty applied to the calibration of DES SNe Ia is reduced by a factor $\sqrt{2}$ by the (assumed) equal contribution of GAIA and FGCM scatter to the residual, and by a further factor of $\sqrt{2}$ when considering that spatial uniformity applied to the calibration of DES SN fields is applied to the difference between two positions, of the SN Ia and the standard stars. Thus we add $\frac{1}{2}\sigma_{\rm res}$ to diagonal DECam zero point errors in $\mathbf{C}_\kappa$, assuming that this scatter is uncorrelated between filters.

**Calibration transfer uncertainties**

The following uncertainties are associated with the transfer of the HST CALSPEC scale to the FGCM natural system, placing the FGCM standard stars on an absolute scale. These largely follow JLA, and are as described in Betoule et al. (2014, section 3.4). We compute these contributions to $\mathbf{C}_\kappa$ following descriptions therein.

- **Uncertainty in white dwarf colour:** following Betoule et al. (2014, section 3.4.1), we estimate the uncertainty in the colour of the three DA white dwarfs that anchor the HST scale (are primary standards) as a 0.5% uncertainty in the slope of their spectra over the 3000–10000Å range.

- **Uncertainty of calibration transfer to CALSPEC standard C26202:** the FGCM calibration (viewed as magnitudes of standard stars in science fields) is tied to the CALSPEC network via the standard star C26202, which is calibrated by the primary HST stars. In JLA, the uncertainty in this transfer was estimated by modelling the measurement error in individual spectra by monitoring the repeatability of the spectrum of AGK +81 266. For DES, we assume 0.3% uncertainties based on Bohlin (2000) in the absence of such spectral monitoring. In both cases, the uncertainty falls with the square root of number of observations of the CALSPEC standard with STIS; however C26202 only has a single STIS observation.

## 6.4.2 Scene Modelling Photometry

Photometry for the DES3YS is determined using the Scene Modelling Photometry (Brout et al., in prep.) method, which improves on the preliminary difference imaging (DIFFIMG) pipeline described in Kessler et al. (2015). The value of cosmology results from DES supernovae relies on the the best photometric calibration possible. FGCM sets the instrumental zero points and models the DECam transmission curves for any given exposure, but for supernova exposures, a further step is needed to obtain photometry of the transient without any galaxy light. Just as the precise calibration from FGCM is crucial for the survey-wide calibration, the care and attention to detail in the SMP process is necessary for the accurate measurement of SN Ia-only light. Brout et al., (in prep.) will give a full description of the SMP pipeline which produces the DES3YS lightcurves.

[3]HEALPIX is a system for the subdivision of the surface area of a sphere into pixels of equal area, used in GAIA. NSIDE is an index for size of pixels.

The SMP method builds on techniques initially developed in Astier et al. (SNLS; 2006) and Holtzman et al. (SDSS; 2008). The principal is to calculate magnitudes of individual lightcurve points without subtracting galaxy templates. Traditionally, difference imaging (e.g. Skymapper (Scalzo et al., 2017) and DIFFIMG for DES (Kessler et al., 2015); described earlier in Chapter 2) have involved convolving galaxy templates to the same point spread function (PSF) as the supernova image, which almost always has poorer resolution. This resampling process invariably leads to a degradation in PSF. In particular, noise that this subtraction introduces is often correlated between pixels.

Instead, like with the FGCM method, SMP is a model-based approach which fits detection images simultaneously as the sum of a galaxy background and supernova point source. The framework fits Equations 3.8 where the data consists of the set of all individual pixels, and the model is the prediction from galaxy and supernova templates.

In SNLS (Astier et al., 2006), each image is modelled as a sum of a time-independent galaxy background plus a time-dependent supernova (modelled as a point source), convolved with the PSF of each image. In SDSS (Holtzman et al., 2008) the contributions to the image also include a sky background and stars (also point sources) in the extensive SDSS calibration network. All images resampled to a common pixel grid; Holtzman et al. (2008) avoids this resampling, to avoid correlated uncertainties between pixels that are close to each other, which can lead to underestimated parameter errors.

For the DES SMP, Brout et al., (in prep.) builds on the photometry methods in Astier et al. (2006); Holtzman et al. (2008) with the addition of extensive simulations. In this framework, each image (a neighbourhood of the SN) is modelled as the sum of a point source, a galaxy sampled on a pixel scale, and the sky background. The parameters of the model are fitted by minimising $\chi^2$ (defined above) simultaneously over all epochs. Simulations of thousands of fake supernovae lightcurves are generated, with a fake image at each epoch composed of the sum of a galaxy template and a two-dimensional delta function for the SN Ia, which then have a sky background added, and are convolved with a PSF for each image. These simulated lightcurves are fitted with SMP alongside the real lightcurves, to validate the model and method, and to estimate the residual flux.

### 6.4.3 Correction terms

In this section we outline the corrections applied to SN Ia observations, introduced in Section 2.3.1, which are applied to observed quantities $m_B, X_1, C, z_{\mathrm{CMB}}$ of each supernova, yielding its corrected model and observed distance moduli $\mu_{\mathrm{mod}}$ and $\mu_{\mathrm{obs}}$, respectively.

From observed DES3YS SN Ia lightcurves, the observed distance modulus is (following Tripp corrections in Section 2.2.2):

$$\mu_{\mathrm{obs}}(\alpha, \beta, M_B) = m_B - M_B + \alpha X_1 - \beta C \tag{6.17}$$

Beyond the Tripp parameters, Equation 6.18 is corrected for the host galaxy mass dependence of the intrinsic SN Ia absolute magnitude as described in Section 2.3.2. Further, the observed distance modulus is corrected for Malmquist bias, a term estimated using the spectroscopic selection efficiency, described in Section 6.4.4. Similarly, the redshift $z$ in the model or predicted distance modulus $\mu_{\mathrm{mod}}$ in Equation 6.16 have been corrected for peculiar velocities where appropriate (i.e. for low-$z$ SNe Ia) as described in Section 3.1.3.

For the DES supernovae, host galaxy masses are estimated from broadband galaxy photometry Smith et al., (in prep.) using methods described in Smith et al. (2012); Sullivan et al. (2006). Following Section 2.3.2, we correct the absolute magnitude using the quantity $M_B^*$ (Equation 2.4:

$$\mu_{\mathrm{obs}}(\alpha, \beta, M_B, \Delta M_B) = m_B - M_B^* + \alpha X_1 - \beta C \tag{6.18}$$

The mass step $\Delta M_B$ is fitted for as one of the nuisance parameters; we find (Section 6.6) $\Delta M_B = -0.08 \pm 0.04$.

The observed distance modulus requires correction for Malmquist or other selection biases also, according to

$$\mu_{\rm cor}(\alpha, \beta, M_B, \Delta M_B, z) = \mu_{\rm obs} - \delta\mu(z) \tag{6.19}$$

where the bias correction function $\delta\mu$ is estimated below in Section 6.4.4.

On the model side, the CMB-frame redshift $z$ in $\mu_{\rm mod}$ in Equation 6.16 below is corrected for peculiar velocities (Section 3.1.3) for the low-$z$ supernovae in the DES3YS sample. Equation 3.3 allows the cosmological redshift (due to expansion alone) – denoted $\bar{z}$ in Section 3.1.3, and $z$ in the remainder of this chapter – to be isolated from other contributions to redshift from peculiar velocities, which include any motion other than from cosmological expansion, including the motion of the Solar system relative to the CMB, and the peculiar velocity of the supernova and its host galaxy relative to the CMB. In the DES3YS analysis, peculiar velocity corrections are only applied to the low-$z$ supernovae, where the $\boldsymbol{v}_{\rm SN}^{\rm pec}$ term is significant compared to $cz$, and where 2M++ provides density field information. At higher redshifts, velocity estimates are mostly not available; even if they were they would have a much smaller impact despite their larger uncertainties.

### 6.4.4 Malmquist bias from selection efficiency

In Section 3.1.2 we had outlined the importance of modelling and correcting for selection biases, which are especially potent in supernova surveys. In magnitude limited surveys, intrinsically brighter objects are preferentially detected, leading to Malmquist bias: a skewed estimate of the absolute magnitude distribution. The Malmquist bias can be estimated by modelling the selection efficiency (i.e. the rate of successful detection as a function of magnitude) to match observed distributions (in redshift, stretch, and colour), by simulating the survey with SNANA (Kessler et al., 2009a) to obtain the bias $\delta\mu$ in distance modulus.

In DES3YS, two parallel approaches are employed to estimate the spectroscopic selection efficiency: via data/MC simulations, from ratios of Ias to all transients in the data. The methods are in agreement and are propagated through to a bias in distance modulus as a function of binned redshift via realistic simulations; this bias correction can be applied directly to correct distance moduli prior to cosmology fitting. In other simulation-based analysis methods, the selection efficiency is input into the simulations directly.

For the JLA-like analysis of the DES3YS sample, we apply a spectroscopic selection function to a simulated DES3YS-like data set to recover the Malmquist or selection bias $\delta\mu := \mu_{\rm out} - \mu_{\rm in}$ as a function of binned redshift. We interpolate this function to find $\delta\mu(z)$ for any redshift in the range and correct the observed distance modulus, according to Equation 6.19.

The simulation-based analysis methods for the DES3YS sample, BBC (Brout et al., in prep.) and BHM (Hinton et al., in prep.), instead forward-model the selection function through numerous simulations, to retrieve bias-corrected distances as discussed in Section 3.1.2. The bias corrections are necessary for simulations which form part of the cosmology fitting process in these analyses, whereas in the JLA-like analysis they are separable. More details on the bias corrections are given in the above references.

An earlier iteration of the spectroscopic selection function for the DES3YS sample is shown in Figure 6.2. The selection function is computed as a function of peak $i$-band magnitude in two ways. The first, led by Anais Möller, compares data and simulation distributions: the observed data sample, after cuts, is compared to simulated parent SN Ia populations using SNANA, using survey simulation parameters representative of the DES3YS sample. Two models of intrinsic colour variation (**G10**

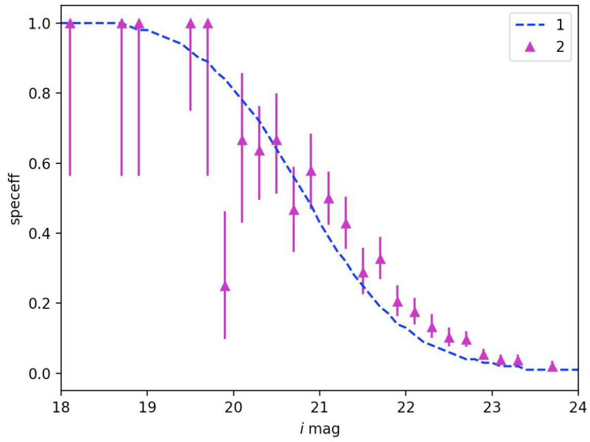


Figure 6.2 An example spectroscopic selection function for the DES3YS sample determined from sigmoid fit to data-to-simulation comparisons (blue dashed line) and ratios of Ias to all detected transients (magenta triangles). Error bars for the latter span the blue line.

and **C11**, explained in Section 3.4.7) are tested, and the **G10** model is used. A sigmoid function is fitted to the binned points, shown in the blue dashed line in Figure 6.2. The second approach, led by Chris D'Andrea and Mathew Smith (magenta points in Figure 6.2) compared the fraction of detected spectroscopically confirmed Ias, before any cuts, to all detected transients – see D'Andrea et al., (in prep.) for an explanation of this method. These points are binned in intervals of 0.2 mag. The physical motivations for denominators in both methods are different, but the derived selection functions (Figure 6.2) and Malmquist biases (Figure 6.3) agree well.

We model the Malmquist bias as a Gaussian process, using methods (Rasmussen & Williams, 2006) described in Section 3.3.4. Both of the selection functions and their uncertainties are treated as input data points, and (given a prior distribution) return a posterior distribution of smooth functions, where each function in the posterior is defined by its behaviour at all test points (in this case, linearly spaced points in the redshift range $0.05 < z < 0.8$ – see Figure 6.3). A multivariate gaussian distribution is chosen as a prior, parametrised by a kernel width and scaled by the variance of the data. The kernel is set to $k = 0.05$ for the input data to ensure that the scale of features is appropriate for the errors in the input data. We take the mean (the red line in Figure 6.3) of the posterior distribution of functions as the combined Malmquist bias correction $\delta\mu(z)$, which is interpolated to give $\delta\mu(z)$ in Equation 6.19. We conservatively choose to use five times the standard deviation in the posterior as the uncertainty in the bias correction to reflect the fact that the difference between the corrections derived from the two methods is greater than the uncertainty carried from either method (similarly to the approach in Section 3.4.4). This error in the posterior shown in the grey shaded area in Figure 6.3). The propagation of this uncertainty to cosmology results is discussed in Section 6.5.2.

This method, of combining results of two approaches using Gaussian processes, differs from JLA, where the few binned points are interpolated using a spline polynomial fit and errors are propagated to the distance modulus using uncertainties in the polynomial coefficients. Our method is justified because it applies the same principal (fitting a smooth function to the discrete points Betoule et al. (figure 5, 2014) and estimating the error), in a more natural way: rather than forcing behaviour (specific functional forms), Gaussian processes only influence resultant functions toward smoothness. Thus, we can combine the two congruous Malmquist bias correction functions from the selection functions in Figure 6.2 into a smooth function (the red line) with uncertainties (the grey shaded area), from the function values and their uncertainties. In short, our method results in an output function and errors which reflect both bias estimation methods and their errors.

The above methods, for the DES3YS sample, assume a magnitude-limited scenario. For the low-

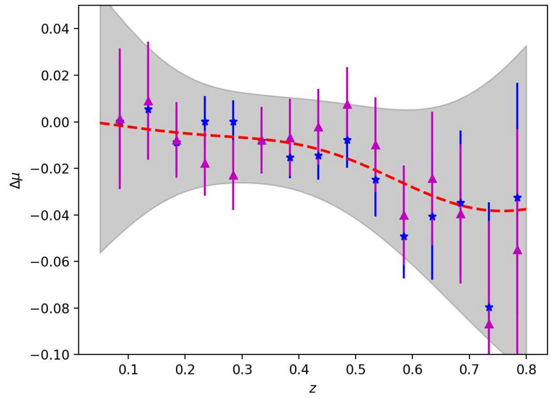


Figure 6.3 Malmquist bias functions determined from the spectroscopic selection functions in Fig 6.2, which are combined using Gaussian processes. the blue stars are from the data-to-simulation comparisons, whilst the magenta triangles are from the ratios of Ias to all detected transients. The red line and grey area are respectively the mean and $5\sigma$ uncertainty region for the posterior distribution of smooth functions.

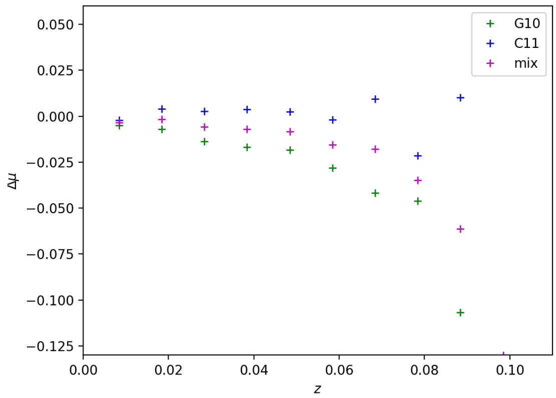


Figure 6.4 Malmquist bias for low-redshift surveys, using SNANA simulations for the CSP and CfA3/4 surveys assuming a magnitude-limited survey. The magenta points are the average of the two scatter models.

$z$ samples which supplement the DES3YS sample, we use SNANA simulations from Pantheon (private communication from Dan Scolnic) restricted to the CfA3/4 and CSP surveys. These are generated for both magnitude-limited and volume-limited surveys, with **G10** and **C11** scatter models (discussed in Section 3.1.2). We use the same approach as in Scolnic et al. (2018), which is to assume the magnitude-limited simulations and average the **G10** and **C11** scatter models, using the difference between them as the systematic error in the bias correction function. Given that the biases in a volume-limited scenario fall within these limits, this method is consistent with JLA. The average function (purple points in Figure 6.4) is interpolated to give $\delta\mu(z)$ for the low-redshift supernovae.

## 6.5 Cosmology fitting

This section contains methods for fitting for our cosmological parameters introduced in Section 6.2. Following on from Section 6.2, we describe the Bayesian parameter estimation methods which fit these equations to yield posterior distributions.

The parametrisation in Equation 6.16 is for a flat $w$CDM or $\Lambda$CDM Universe with non-evolving dark energy, but can be easily extended to other cosmologies: for evolving dark energy (e.g. with $w = w_0 + w_a(1-a)$, for scale factor $a$) the single parameter $w$ can be replaced by $(w_0, w_a)$, whereas for the more specific ΛCDM it is fixed at $w = -1$. If the flatness constraint is removed, both $\Omega_m$ and $\Omega_{\rm DE}$ are fitted for, as they do not necessarily add to 1. In this analysis, we do not allow for time-evolving dark energy, nor non-flat geometries.

We focus on flat ΛCDM and $w$CDM cosmologies. For the former, SNe Ia alone can constrain $\Omega_m$ well, but if fitting for $w$ also, the $\Omega_m$–$w$ posterior inhabits a non-elliptical contour ('banana plot') in parameter space. External orthogonal priors from BAO and the CMB are required to constrain $w$.

### 6.5.1 Maximising likelihood with MCMC

The likelihood is a function of the observed data and the underlying cosmological model, and is the workhorse of Bayesian methods for making cosmological inferences (Section 3.2). In this section we discuss computing the likelihood from the theory in Section 6.2, and Markov Chain Monte Carlo (MCMC; Section 3.2.2) methods for maximising the likelihood to do cosmology fitting.

In particular, the distance modulus in Equation 6.19 from the data, and the theoretical or model distance modulus Equation 6.16 from the theory are combined in the likelihood. Simplistically, the closer these are for a given set of cosmological parameters Θ, the higher the likelihood. This is scaled by correlated errors in the data (contained in only the dependent variable, typically referred to as the 'Y' variable, in this case the distance modulus $\mu$) contained in a covariance matrix $\mathbf{C}_\mu$. A limitation of maximal likelihood methods is that they only allow for errors in the 'Y' or dependent data (in this analysis, $\eta$ or the distance modulus $\mu$) to be taken into account (via $\mathbf{C}_\eta$ or $\mathbf{C}_\mu$), and not errors in the 'X' or independent data (the redshift $z$ in this analysis). This limitation was one motivation (Gull, 1989) for Bayesian hierarchical methods (discussed in Section 3.2.3), which will be independently applied to this sample and in future DES analyses.

The equations in Section 6.2 go into a maximal likelihood, or minimal $\chi^2$, method for fitting cosmological parameters. The foundations of Bayesian methods for cosmology were presented in Section 3.2. In summary, a likelihood term $\mathcal{L}$, a function of a set of cosmological and supernova parameters Θ, is calculated analytically over a parameter space to determine parameters (specifically, posterior distributions for them) that maximise the likelihood. Implementation of this is typically through sampling techniques such as MCMC (Section 3.2.2).

The likelihood is determined from the $\chi^2$statistic, both functions of the parameters of fit Θ. For a supernova Hubble diagram with distance moduli from observational data ($\mu_{\rm cor}$ from Equation 6.19) and from a parametrised model as function of redshift ($\mu_{\rm mod}(z, \Theta)$), are

$$\chi^2(\Theta) = (\boldsymbol{\mu}_{\mathbf{cor}} - \mu_{\rm mod}(\boldsymbol{z}, \Theta, \delta\mu))\mathbf{C}_\mu{}^{-1}(\boldsymbol{\mu}_{\mathbf{cor}} - \mu_{\rm mod}(\boldsymbol{z}, \Theta, \delta\mu))^T \tag{6.20}$$

$$\mathcal{L}(\Theta) = \exp\left(-\frac{1}{2}\chi^2(\Theta)\right) \tag{6.21}$$

which is Equation 3.12 where the observable is the bias-corrected distance modulus $\mu_{\rm cor}$. We assume a flat Universe with $k = 0$ and $\Omega_{\rm DE} = 1 - \Omega_m$, for which the parameter vector Θ to be determined is $\{\Omega_m, \alpha, \beta, M_B, \Delta M_B\}$ for a ΛCDM cosmology or $\{\Omega_m, w, \alpha, \beta, M_B, \Delta M_B\}$ for $w$CDM. If the flatness condition is removed Θ will also include a $\Omega_{\rm DE}$ term. If time-evolving dark energy is allowed, $w$ is replaced by $w_0$ and $w_a$.

The matrix $\mathbf{C}_\mu{}^{-1}$ is the covariance matrix of correlated uncertainties in the observed distance moduli $\mu_{\rm obs}$, a generalisation of the diagonal uncertainties $\sigma_{\mu_i}$ in each distance modulus (see Section 3.2). Descriptions of the computation of the individual contributions to $\mathbf{C}_\mu{}^{-1}$ follow in Section 6.5.2.

Markov Chain Monte Carlo (described in Section 3.2.2) is a probabilistic Bayesian method that is useful for parameter estimation and model selection. We use MCMC, with the `emcee` (Foreman-Mackey et al.,

2013) implementation, to constrain values, i.e. the posterior distributions of the set of parameters parameters $\Theta$ which maximise the likelihood $\mathcal{L}$ in Equation 6.21. We run MCMC fits with 100 walkers and 500 steps, and burn in 100 steps (10000 points). We ensure chains have properly converged at this burn-in by examining the evolution of the likelihood of the walks, and affirming that resultant values remain the same if further burnt in. Results are shown in Section 6.6.

### 6.5.2 Systematics

Chapters 2 and 3 described numerous systematics which affect supernova surveys, which include those specific to supernovae, and more general effects. It is especially important to account for these accurately in cosmological studies of supernovae in the near future, as sample sizes increase and the absolute size of the statistical uncertainties decrease. In current analyses of the DES3YS sample, there are two broad approaches to estimating systematic uncertainties: using covariance matrices, and via simulation. We first briefly describe the simulation method, then explain the covariance matrices which underpin the treatment of systematics in the JLA-like analysis.

In the simulation approach, hundreds of realisations of a DES-like survey are generated to closely resemble the DES and low-$z$ combined sample. A number of systematic terms, including Milky Way extinction, instrument zero points and wavelengths, are individually perturbed and the simulation repeated, and the shift in the resultant cosmological and supernova parameters are taken as the uncertainty due to that particular systematic. These shifts are added in quadrature for all systematics terms. This is a similar principle to Equation 3.17, but interpreted via simulation.

The covariance matrix method described in Section 3.4 is a natural way to estimate uncertainties in this framework, and combined with using SALT2 for lightcurve fitting, enable us to follow JLA. An alternative (used in the BBC method) is to bin by redshift and use a 'compressed' covariance matrix (Brout et al., in prep.). Through Equation 3.15, $\mathbf{C}_\eta$ is transformed into $\mathbf{C}_\mu$, which has size $N \times N$ matrix with correlated errors for the distance modulus of each supernova, and fits into Equations 6.20 and 6.21.

The host mass, Malmquist bias, peculiar velocity systematics are corrected for peak absolute SN Ia magnitude, $z_{\rm cor}$, and distance modulus respectively (Section 6.4.3). The systematic is the error in these correction terms. The other systematics are errors in an input of the analysis such as lightcurve fitting, then computing the uncertainty to propagate through to final cosmology fit results.

The remainder of this section details calculations of covariance matrices descriptions of $\mathbf{C}_\eta$ and $\mathbf{C}_\mu$ for the DES3YS sample. We start by considering the diagonal terms (which go directly into $\mathbf{C}_\mu$), then each matrix (statistical and various systematic terms) comprising $\mathbf{C}_\eta$. For the computation of uncertainties from corrections for selection bias and peculiar velocities, as well as diagonal uncertainties and systematics associated with calibration and non-Ia contamination, we describe specifics of the DES3YS survey, with additions or deviations from general descriptions in Chapter 3, Section 3.4. For the remaining covariance matrices (statistical, host mass corrections, dust extinction, lightcurve model) we refer to the descriptions in Section 3.4 directly.

#### Differences from Chapter 3

Several covariance matrices, including $\mathbf{C}_{\rm stat}$ and the systematic terms $\mathbf{C}_{\rm host}$, $\mathbf{C}_{\rm dust}$, $\mathbf{C}_{\rm pecvel}$, and $\mathbf{C}_{\rm model}$, computations are virtually unchanged from the generalised methods in Chapter 3 (Sections 3.4.3, 3.4.7, and 3.4.5 respectively). We make the following comments or additions for these matrices, as follows.

When computing the diagonal terms (described more generally in Section 3.4), the size of the uncertainty in peculiar velocities is estimated to be[4] $c\sigma_{z,\rm pec} = 220 \mathrm{km\,s}^{-1}$, and the corresponding uncertainty in redshift is related via $\sigma_z = (1+z)\sigma_{z,\rm pec}$. All redshifts used in computing these diagonal

[4]From Michael Hudson via private communication; see Andersen (2018) for further discussion.

terms have been corrected for peculiar velocities (i.e. $z_{\rm cor}$ in Section 3.4.4). We use estimates in JLA (following SNLS) for the sizes of the uncertainties in SN Ia peak magnitudes from gravitational lensing and intrinsic scatter, being $\sigma_{\rm lens} = 0.055z$, and constant $\sigma_{\rm int,SN}$ for each DES3YS and low-$z$ samples, of 0.08 and 0.134 respectively. Subtleties of the latter are discussed in Section 3.4.7.

For $\mathbf{C}_{\rm host}$ we fit for the size of the mass step $\Delta M_B$, rather than fixing it as in Chapter 4.

We update the relative uncertainty in extinction due to Galactic dust from 20% (in JLA) to 5%, for estimating $\mathbf{C}_{\rm dust}$. We also adopt the newer (Schlafly & Finkbeiner, 2011) dust maps instead of SFD, which scales the extinction by a factor of 0.86 in effect. These changes follows Pantheon, where Scolnic et al. (2014) note that some small patches of sky have ∼10% coherent error, despite the average < 1% error in dust maps, thereby using 5% of the dust extinction as an all-sky estimate. As noted at the end of this section, SALT2 has not been retrained with these smaller uncertainties in extinction.

In computing $\mathbf{C}_{\rm pecvel}$, we improve on Equation 4.21 when propagating the uncertainty in redshift through to apparent magnitude, using

$$\sigma_{\mu,\rm pecvel} = \frac{5\sigma_z}{\log(10)}\left(\frac{1+z}{z(1+\frac{z}{2})}\right). \tag{6.22}$$

which follows (Davis et al., 2011, eq A1, A4). This differs from Equation 4.21, which was used in JLA in Betoule et al. (2014, eq 13), by a factor of $\frac{1+z}{1+\frac{z}{2}}$ (only marginally different from unity at low redshifts), and is more precise, as explained in Andersen (2018).

The computation of $\mathbf{C}_{\rm model}$ follows JLA exactly, in the method described in Section 3.4.7. Since JLA, Scolnic & Kessler (2016) have performed simulations to further model realistic SN Ia parent populations, including the colour-dependence of intrinsic scatter; these are incorporated in more recent analysis such as Pantheon (Scolnic et al., 2018), and in the BBC analysis of the DES3YS sample.

We now describe in full methods of computing the matrices $\mathbf{C}_{\rm bias}$, $\mathbf{C}_{\rm cal}$, and $\mathbf{C}_{\rm nonIa}$ which are specific to the DES sample.

### Malmquist bias correction

The uncertainty due to selection or Malmquist biases is computed according to Section 3.4.2, using the selection functions in Section 6.4.4. This differs from JLA, in using new simulations to model the selection function, which is propagated to a $\delta\mu$ function. We use this combined bias function, with errors inferred from using Gaussian processes to combine multiple bias functions, as described in Section 6.4.4. These errors are then to estimate uncertainties in $\eta$ due to selection bias, as in Section 3.4.2.

For the low-$z$ sample, the uncertainty in bias correction is derived from differences in bias functions associated with different scatter models, as modelled in Pantheon and Scolnic & Kessler (2016); this is substantially lower than in JLA, where the uncertainty was estimated as the same size as the low-$z$ bias correction itself (i.e. the errors allowed anywhere between zero and double the bias correction).

### Calibration

Calibration via FGCM is described above in Section 6.4.1, with a more general view of photometric calibration in Section 3.1.1. Recent studies have emphasised the significant contribution of calibration uncertainties to the error budget in cosmological parameters. We base our methods for estimating these on JLA (Betoule et al., 2013, 2014) to reproduce the calibration covariance matrix for our SN sample and the telescopes used to observe them.

The general principle for computing $\mathbf{C}_{\rm cal} = \mathbf{J}\mathbf{C}_\kappa\mathbf{J}^T$ follow descriptions in Section 3.4.6. The instruments in $\kappa$ are listed in Table 6.2, and include the DECam passbands used for SN Ia observations

Table 6.2 Systematics in $\kappa$

| Instrument | Filters | ZP (mmag) | $\lambda_{\rm eff}(\AA)$ |
|---|---|---|---|
| DECam | *griz* | $7^a$ | 5 |
| KeplerCam | *Us* | 31 | 25 |
| KeplerCam | $B^b Vri$ | 6,6,4,3 | 7 |
| 4Shooter2 | *Us* | 70 | 25 |
| 4Shooter2 | *B* | 6,4,3,5 | 7 |
| Swope | $ugriBV^c$ | 23,4,3,5,5,5 | 7,8,4,2,7,3 |

[a] from FGCM residuals
[b] including *Bc*, the KeplerCam B filter changed after some date
[c] including the three different *V* filters used: *V*-3014, *V*-3009, *V*-9844 used for Swope (Contreras et al., 2010)

(*griz*), and instruments which made CSP and CfA3/4 observations (Swope, and Keplercam and 4Shooter2, respectively).

The Jacobian $\mathbf{J}$ is computed as described in Section 3.4.6, while $\mathbf{C}_\kappa$ is reevaluated. A subset of the terms which contribute to $\mathbf{C}_\kappa$ are those which are specific to the FGCM calibration method, described in Section 6.4.1; these are analogous to the methods described in Betoule et al. (section 3.4.2, 2014) but not directly comparable. The remaining terms are tied to the absolute CALSPEC-based magnitude scale that the FGCM calibration is anchored to, described in Section 6.4.1 directly following methods in Betoule et al. (2014, section 3.4.1).

**Low-redshift uncertainties**

For the low-redshift supernovae, we use numerical uncertainties reported in Pantheon (table 5, Scolnic et al., 2018), following the method in JLA. These have decreased substantially since the JLA (errors summarised in Betoule et al. (section 3.4.3, 2014), and described earlier in Conley et al. (2011) for SNLS) through the Supercal calibration of PS1 (Scolnic et al., 2015). The photometry and calibration uncertainties of these low-$z$ SNe Ia are taken from PS1 directly (as in Pantheon), and described in Scolnic et al. (2018). In summary, Supercal compares PS1 photometry of standard stars over a wide area of the sky, along with photometry in natural systems of multiple surveys, to reduce discrepancies in the calibrations of different low-$z$ supernova surveys to the milli-mag level.

Calibration offsets and uncertainties determined in PS1 are applied to low-$z$ photometry used for this analysis, and summarised in Table 6.2. A further limitation of the present low-$z$ calibration used is that the correlated uncertainties in tying together calibrations of the different low-$z$ surveys in PS1 are not fully understood, and thus not fully taken into account in $\mathbf{C}_\kappa$.

While Supercal has been very successful in reducing estimated uncertainties in the low-redshift sample, a limitation is that the SALT2 model in JLA, SALT2.4, has not been retrained using the updated uncertainties in Supercal. There is hence an inconsistency, likely slight but not yet measured, between the natural system for the low-$z$ photometry in the DES3YS analysis (and R16, discussed in Section 5.2.1), and the SALT2 model for fitting lightcurves and computing calibration uncertainties in $\mathbf{C}_\kappa$. Similarly, the updated dust extinction, and its reduced uncertainty from using (Schlafly & Finkbeiner, 2011), have not been incorporated into the SALT2 model training.

## 6.6 Constraints on ΛCDM from supernovae only

The next two sections contain preliminary results of our MCMC fits for cosmological parameters and SN Ia nuisance parameters, according to methods discussed in Section 6.5.1. Corrections to data were applied as described in Section 6.4.3. Cuts on SN Ia samples are the same as those used in the BBC analysis (Section 6.3.4).

We emphasise the preliminary nature of the results in this section, which have been derived from the frozen version of DES3YS data from the time of the December 2017 unblinding. Since then, minor changes to aspects of the analysis such as photometry and bias corrections have occurred which have not been reflected in this chapter. Instead, the methods in this chapter will be applied by Tamara Davis to the DES3YS data alongside the other methods, as part of the key Dark Energy Survey Collaboration, (in prep.) paper.

We fit for two cosmological models, both assuming an energy composition of mostly dark energy and cold dark matter and spatial flatness (with $k = 0$, $\Omega_m + \Omega_{\rm DE} = 1$). This section examines results in a $\Lambda$CDM Universe, where dark energy takes the form of a cosmological constant $\Lambda$. Using SNe Ia alone, we can constrain the matter density $\Omega_m$ as well as SN Ia nuisance parameters. Subsequently, Section 6.7 visits the more general $w$CDM case. We also present an error breakdown in Section 6.6.3 as well as a discussion of our results compared to JLA (the benchmark data set with the same method).

### 6.6.1 Validation on JLA data

Our methods are validated on JLA data, affirming that we do retrieve the same values as B14 using our method which is intended to closely reproduce theirs. We can then compare our results from the DES3YS to those we derive from JLA data.

We apply the same methods described in Sections 6.2 and 6.5 to the JLA lightcurve parameters,[5] with the only difference being that Malmquist bias corrections (Section 6.4.4) are only included in the apparent magnitudes released by JLA, so we do not apply corrections again. The results for our fits on JLA data, compared to those in Betoule et al. (2014, table 10), are given in Table 6.3. Comparing the first two rows of Table 6.3 reveals that our central values for $\{\Omega_m, \alpha, \beta, M_B, \Delta_M\}$ agree very well, while our errors are consistently larger than in JLA by approximately 30%. Our fits to JLA data are displayed as the red contours in Figure 6.5 and later Figure 6.6, alongside our fits to the DES3YS+low-$z$ combined sample.

Table 6.3 Results from $\Lambda$CDM fits for $\Omega_m$ and supernova parameters $\alpha, \beta, M_B, \Delta$, compared to results from JLA (both our fits, and as published in Betoule et al. (2014)

| | $\Omega_m$ | $\alpha$ | $\beta$ | $M_B$ | $\Delta M_B$ |
|---|---|---|---|---|---|
| JLA (B14) | $0.295 \pm 0.034$ | $0.141 \pm 0.006$ | $3.101 \pm 0.075$ | $-19.05 \pm 0.02$ | $-0.070 \pm 0.023$ |
| JLA | $0.293^{+0.044}_{-0.037}$ | $0.142 \pm 0.008$ | $3.10 \pm 0.10$ | $-19.05 \pm 0.028$ | $-0.072 \pm 0.03$ |
| DES3YS | $0.268^{0.053}_{-0.046}$ | $0.144^{+0.009}_{-0.012}$ | $3.14 \pm 0.11$ | $-19.09 \pm 0.037$ | $-0.076 \pm 0.041$ |

### 6.6.2 DES3YS sample

Figure 6.5 shows the results of the JLA-like fit to the DES3YS sample combined with the low-redshift sample (blue), as well as the SN Ia-only $\Lambda$CDM fit from JLA (red). With the condition that dark energy is a cosmological constant, in the form of the constraint $w = -1$, these results represent the best cosmological constraints available from supernovae alone, without external priors. Our best estimate for the relative density of matter $\Omega_m = 0.268^{+0.053}_{-0.046}$. This is slightly lower than in JLA ($\Omega_m = 0.295 \pm 0.034$), and larger errors. The larger errors are in line with expectations, as the DES3YS+ sample has fewer than half the number of supernovae as JLA (338 compared to 740 SNe Ia); we also noted in Section 6.6.1 that our method results in a larger error for the same data set as JLA.

### 6.6.3 Error contributions

We estimate the breakdown of contributions of different error sources (terms in $\mathbf{C}_\eta$ in Equation 3.16) following JLA, according to the method described in (section 6.2, Betoule et al., 2014). This involves

[5]Data downloaded from the JLA release page at http://supernovae.in2p3.fr/sdss_snls_jla/ReadMe.html.

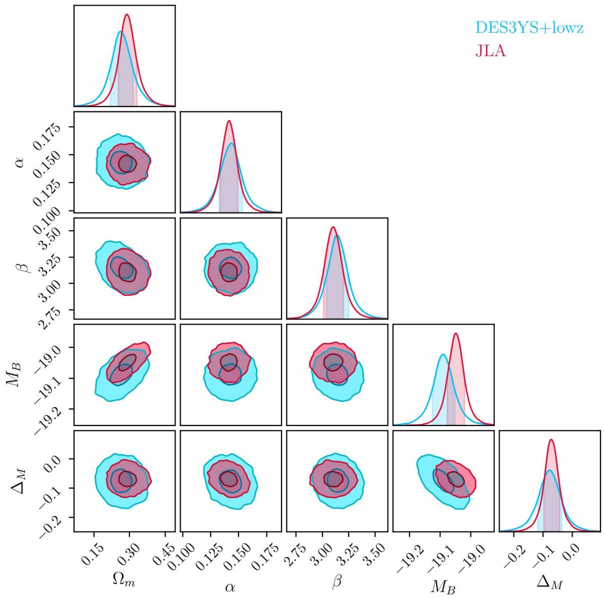


Figure 6.5 Contours for $\Theta = \{\Omega_m, \alpha, \beta, M_B, \Delta M_B\}$ for a ΛCDM fit of the combined DES3YS+low-$z$ sample (blue), plotted over JLA ΛCDM results (Betoule et al. (table 10, 2014)).

Table 6.4 Relative contributions to the uncertainty in $\Omega_m$ (i.e. the variance) from individual statistical and systematic sources uncertainties, calculated as described in Betoule et al. (2014, Section 6.2).

| Source of uncertainty | Variance $\sigma^2(\Omega)$ | Contribution to $\sigma^2(\Omega_m)$ (%) | Described in Section |
|---|---|---|---|
| **Statistical** | 0.0297 | **63.8** | 3.4.1 |
| **Systematic:** | | | |
| Calibration | 0.0191 | 26.4 | 6.5.2 |
| Host galaxy | 0.0081 | 4.8 | 3.4.3 |
| Lightcurve model | 0.0074 | 4.0 | 3.4.7 |
| Contamination | 0.0030 | 0.7 | 3.4.7 |
| Malmquist bias | 0.0014 | 0.1 | 6.5.2 |
| MW extinction | 0.0013 | 0.1 | 3.4.5 |
| Peculiar velocities | 0.0011 | 0.1 | 3.4.4 |
| **Total systematic** | 0.0223 | **36.2** | |

decomposing the variance of each covariance matrix for the parameter $\Omega_m$, chosen because it is a parameter of interest that can be constrained by only supernovae, unlike $w$. Table 6.4 shows the individual contributions as a fraction of the total variance in $\Omega_m$ from each systematic term in Section 6.5.2.

As indicated in B14, this method of assessing relative errors serves to provide qualitative insight into which systematics affect cosmological parameters the most, rather than providing a rigorous decomposition of sources of uncertainties in $w$. It is not directly comparable to commonly used quadrature methods, which estimate the total uncertainty as a quadrature sum of uncertainties from numerous sources, where the size of the uncertainty form each source is directly reportable. Meanwhile, an alternative method used in the other analyses on the DES3YS data is to perform many simulations with individual systematic terms perturbed (described briefly at the start of Section 6.5.2) and measuring the relative change in contour area (e.g. table 10 in Conley et al. (2011) for SNLS), and using these test results to validate the sizes of uncertainties; these tests will be described in the upcoming DES3YS cosmology paper.

As expected, errors in cosmological parameters are statistically dominated for the DES3YS sam-

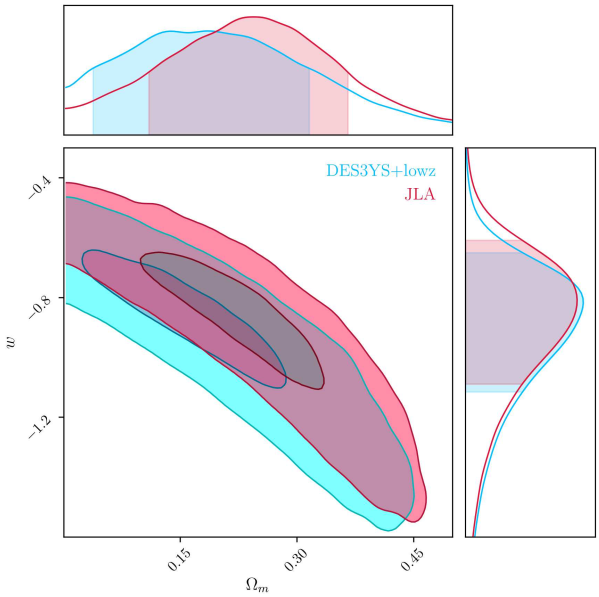


Figure 6.6 The $\Omega_m$–$w$ contour in Figure 6.7 (turquoise) compared to the same contour from JLA supernova data. Both statistical and systematic errors are included.

ple. The systematics budget is dominated significantly by calibration, like in JLA. Notably, the total variance in $\Omega_m$ from $\mathbf{C}_{\text{cal}}$ is similar to JLA (0.0191 compared 0.0203 in Betoule et al. (table 11, 2014)) but makes up a smaller proportion of the error budget due to our larger total error. Our statistical uncertainties are unsurprisingly larger than in JLA absolutely, and also relatively (63.8% compared to 51.6%). We find a larger contribution from the host galaxy mass correction uncertainty (4.8% compared to 1.3%), conceivably from more of our sample having uncertain host mass bin. Also unsurprisingly our uncertainty from dust is much smaller (0.1% compared to 4.6%) by virtue of our smaller estimation of errors in extinction maps (Section 3.4.5). To a lesser extent, the same is true for the Malmquist bias correction uncertainty, down to 0.1% from 1.4% due to our smaller (and likely more realistic) estimate of the uncertainty in the low-redshift sample (Section 6.5.2).

## 6.7 Constraints on dark energy

In this section, we examine fits to our DES3YS data for a $w$CDM Universe. We assume the dark energy component is constant in time, and parametrised by the equation-of-state $w$. Fit results for $w$CDM from SNe Ia alone do not adequately constrain both $w$ and $\Omega_m$. In Section 6.7.2, we combine chains from DES3YS supernovae with external chains from the CMB (Planck Collaboration et al., 2016) and BAO (Alam et al., 2017), to place constraints on $w$.

### 6.7.1 $w$CDM SN Ia-only fits

First, we display the posterior probability distributions from SNe Ia alone. This contour in $\Omega_m$–$w$ space is shown in Figure 6.7 and enlarged in Figure 6.6. Given this good agreement in Table 6.5 (i.e. the SN Ia parameters are the same for the DES3YS+low-$z$ data from $\Lambda$CDM and $w$CDM fits), we only show the enlarged $\Omega_m$–$w$ contour for this fit. The supernovae by themselves do not inform the value of $w$.

Figure 6.7 shows the full 6-dimensional $w$CDM fit from the combined DES3YS and low-redshift sample in turquoise, with the additional purple contours showing results with only statistical errors (i.e. using $\mathbf{C}_{\text{stat}}$ in place of $\mathbf{C}_\eta$, or equivalently neglecting the $\mathbf{C}_{\text{sys}}$ term in Equation 3.16). This plot compares the size of posterior contours from systematic and statistical sources, alongside the discussion in Section 6.6.3. Table 6.5 shows the results for the nuisance parameters, in excellent agreement with the

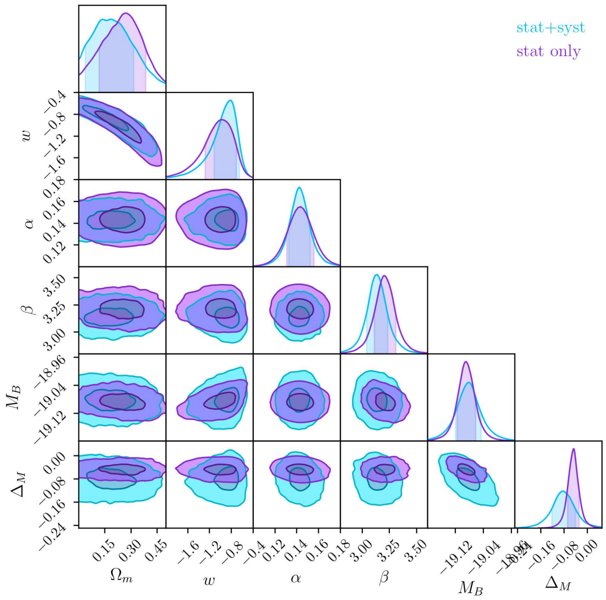


Figure 6.7 Contours for $\Theta = \{\Omega_m, w, \alpha, \beta, M_B, \Delta M_B\}$ for a $w$CDM fit of the combined DES3YS+low-$z$ sample. The canonical results with both statistical and systematic errors contained in $\mathbf{C}_\eta$ are in turquoise, and the statistical error only results are shown in purple.

$\Lambda$CDM fit, as well as the poorly constrained cosmological parameters $\Omega_m$ and $w$. The $\Omega_m$–$w$ contour, or banana plot, in Figure 6.7 from the SN Ia-only $w$CDM fit is replicated in Figure 6.6 along with the same fit from JLA Betoule et al. (table 14/figure 16, 2014) in red.

The results for SN Ia parameters $\{\alpha, \beta, M_B, \Delta M_B\}$ for both models, and relevant cosmological parameters for each model, are shown in Table 6.5. The nuisance parameter values are largely independent on the model, as expected. The largest difference is in the host mass step $\Delta M_B$, which is 0.007 lower for the $w$CDM fit. However, this difference is at $\sim$0.2$\sigma$, and insignificant. Our larger errors in the posterior parameter distributions are visible in Figure 6.5, as well as our slightly smaller value for $\Omega_m$. Additionally, our value for the nuisance parameter $M_B$ (the absolute SN Ia magnitude at peak) is smaller than in JLA by $\sim$0.04 mag, or 1$\sigma$. Otherwise, the our $\Lambda$CDM results agree well with JLA (table 10, Betoule et al., 2014).

Table 6.5 Results from $w$CDM fits for cosmological parameters and supernova parameters $\alpha, \beta, M_B, \Delta$

| model | $\Omega_m$ | $w$ | $\alpha$ | $\beta$ | $M_B$ | $\Delta M_B$ |
|---|---|---|---|---|---|---|
| $\Lambda$CDM | $0.268^{0.053}_{-0.046}$ | - | $0.144^{+0.009}_{-0.012}$ | $3.14 \pm 0.11$ | $-19.09 \pm 0.037$ | $-0.076 \pm 0.041$ |
| $w$CDM | $0.18^{+0.13,a}_{-0.15}$ | $-0.82^{+0.17,a}_{-0.30}$ | $0.143 \pm 0.010$ | $3.14 \pm 0.10$ | $-19.08 \pm 0.037$ | $-0.083^{0.043}_{-0.039}$ |

[a]: poorly constrained by the banana plot alone

Our $\Omega_m$–$w$ contour (Figure 6.6) agrees well with JLA, with the most notable difference being that our $\Omega_m$ is smaller than in JLA, as noted in the $\Lambda$CDM fit in Section 6.6. The marginalised posteriors for $w$ agree very well. The difference from JLA for the nuisance parameters can be seen in Figure 6.5, as they are the same for our $\Lambda$CDM and $w$CDM fits (Table 6.5).

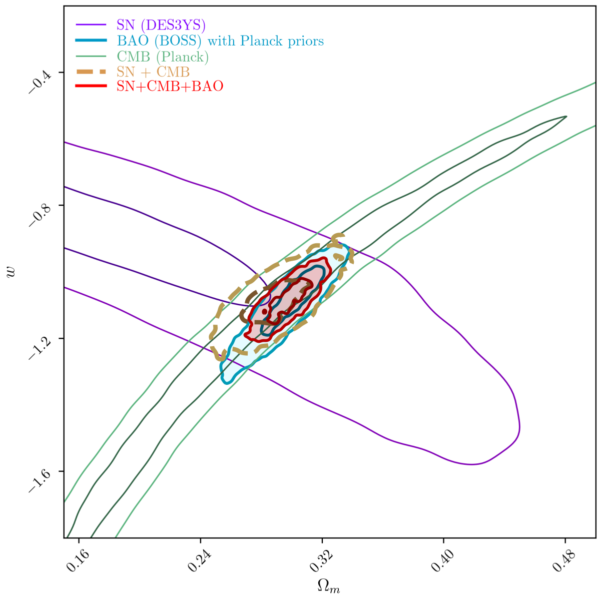


Figure 6.8 Contours in the $\Omega_m$–$w$ plane combined with other probes.

### 6.7.2 Combining with external probes to constrain $w$

To determine $w$, additional data are needed, in the form of a prior on (for example) $\Omega_m$, or contours from external data – usually from BAO in combination with CMB constraints on matter densities, which give orthogonal contours to supernovae. We now consider these additional data combined with the DES3YS and low-redshift SNe Ia.

Table 6.6 Results for cosmological parameters $\Omega_m$ and $w$ from our $w$CDM fit, combined with external datasets

| | $\Omega_m$ | $w$ |
|---|---|---|
| SN[a] | $0.19^{+0.13}_{-0.15}$ | $-0.81^{+0.16}_{-0.30}$ |
| BAO (BOSS)[b] | $0.30 \pm 0.02$ | $-1.064^{+0.077}_{-0.13}$ |
| CMB (Planck) | $0.17^{+0.24}_{-0.03}$ | - |
| SN+BAO | $0.297 \pm 0.014$ | $-1.10^{+0.073}_{-0.057}$ |
| SN+CMB | $0.282^{+0.029}_{-0.014}$ | $-1.099^{0.088}_{-0.064}$ |
| SN+CMB+BAO | $0.297 \pm 0.014$ | $-1.094^{+0.067}_{-0.054}$ |

[a] both parameters are poorly constrained by supernovae alone
[b] with Planck priors on $\Omega_m$ and $\Omega_b h^2$

### 6.7.3 Comparing with other methods

There is some overlap of our 1-$\sigma$ $\Omega_m$–$w$ contour with BBC, however it is shifted in the negative $\Omega_m$ direction. Compared to BBC Brout et al., (in prep.), our posterior contours in $\Omega_m$–$w$ space are larger are less oblate or compressed; although the $2\sigma$ boundaries line up in the positive $\Omega_m$ positive $w$ direction, they extend much further in the negative direction (Figure 6.9). This comparison is intended as preliminary and illustrative, and will be expanded in the full DES3YS cosmology paper.

## 6.8 Summary and next steps

This chapter covers my work in the JLA-like analysis, one of several parallel analyses, of the intermediate DES 3-year spectroscopic supernova sample. These results stand on their own to demonstrate the

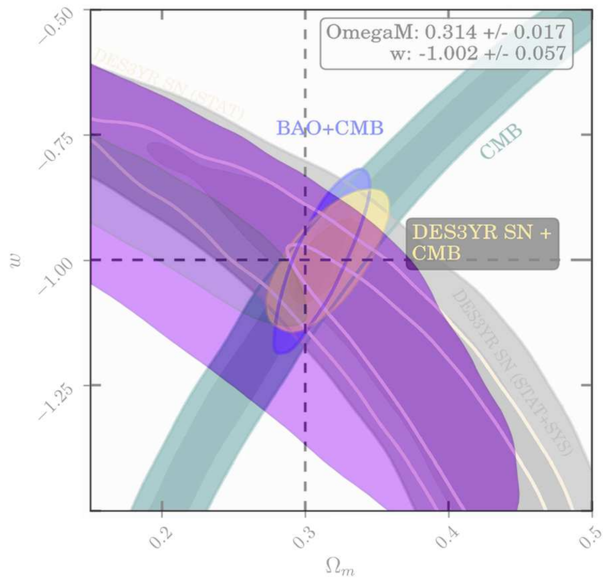


Figure 6.9 The $\Omega_m$–$w$ contour overlaid on BBC results and external priors.

constraining power of the current data set combined with established methods. Additionally, they faciliate two comparisons: of the DES3YS data set compared to JLA (using the same methods), and of the multiple methods applied to this and potentially future DES data sets (using the same data).

We have adopted methods in JLA in particular using full covariance matrices of supernova observables to account for correlated systematic and statistical uncertainties. These covariance matrices go into the calculation of the likelihood term to be minimised in the MCMC fit for cosmological parameters. We have performed fits for flat $\Lambda$CDM and $w$CDM cosmologies, for the DES3YS sample compared with a composite low-redshift SN Ia sample.

From supernovae alone (Section 6.6), we find $\Omega_m = 0.268^{+0.053}_{-0.046}$, slightly (less than $1\sigma$) lower than in JLA and with larger uncertainties, commensurate with the decreased sample size in the DES3YS+low-$z$ which is less than half that in JLA after cuts. In combination with external BAO and CMB priors (Section 6.7), our SN Ia sample constrains $w = -1.094^{+0.067}_{-0.054}$ and $\Omega_m = 0.297 \pm 0.014$. This value of $w$ is slightly lower than $-1$, but not by any means conclusively inconsistent with a cosmological constant.

We emphasise that aspects of our fit which include the choice of cuts, selection function, photometry, choice of dust map and size of uncertainty, had been frozen to match the state of the BBC analysis before results of that analysis were unblinded in December 2017. Additionally, our results here are preliminary – see discussion at the start of Section 6.6 – and will be superseded by those in the relevant section of the alphabetically-authored Dark Energy Survey Collaboration, (in prep.) paper.

Before the December 2017 unblinding, modifications to any of the DES3YS analyses were made blind, without being able to see the impact of changes on cosmological parameters $\Omega_m$ and $w$. Freezing the state of the JLA-analysis has two effects: to prevent seeing the results affecting how we perform the analysis, and to ensure that our comparison to the BBC method is made on the same data set. We note that the process of publically unblinding in late 2017 to reveal results, and subsequently continually making adjustments to the data and analysis, leads to results which cannot be truly described as having been reached blind.

The constraining power of the DES sample will considerably improve in the next few years, with the sample size increasing by an order of magnitude. The most significant change will be the implementation of methods to validate the photometrically classification of supernova and estimate the associated systematic uncertainty. Newer Bayesian methods intended to address complexities of bias and selection effects, including BHM and BBC, will likely be used. They will address, for example, the limitation in the JLA-like method of only including errors in the distance moduli and not the redshifts.

Chapter 7

# Conclusions and future directions

This thesis introduces and describes a collection of work in supernova cosmology, both near and far. I have provided an independent and blinded validation of the distance ladder based Hubble constant measurement in Chapter 4, and developed an analysis framework for the Dark Energy Survey using techniques in JLA in Chapter 6. My work has been important for cementing the persistence of the tension in $H_0$, and for bridging existing and newer methods for studying dark energy. These works and the broader results they support, are representative of the broader shift in paradigm within cosmology, where the focus has shifted from querying dark energy to examining discrepancies in the Hubble constant.

Future directions for both the Hubble constant and dark energy were discussed in Sections 5.3 and 6.8. To summarise, the discrepancy remains between local direct measurements – from supernovae and strong lensing, $H_0 \sim 73\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$ – and values inferred from the early Universe assuming a cosmological model – from CMB anisotropies and the BAO-calibrated inverse distance ladder, $H_0 \sim 68\ \mathrm{km\,s^{-1}\,Mpc^{-1}}$. In the five years since the first Planck results hinted at this tension, numerous studies (Chapter 5) have only served to reinforce it. In contrast, the combination of probes of dark energy appear to be converging on a flat Universe, made up of two components. The first, normal baryonic and dark matter, has relative density $\Omega_m \sim 0.3$, while the remainder ($\Omega_\Lambda \sim 0.7$) has equation-of-state parameter close to $w = -1$, indicating consistency with a cosmological constant.

In the pursuit to understand both the nature of dark energy and the tension in measurements of $H_0$, the data and methods continue to improve. Dark energy data sets are growing larger and more automated, while nearby SN Ia distance measurements are becoming more uniform and better-calibrated. Newer Bayesian analysis methods have proven powerful for representing and handling the complexities and incompleteness of cosmological data. In addition, the first standard siren measurement in GW170817 holds hope that this additional method for determining $H_0$ can bring independent and more precise measurements in the near future. Within all of these developments, it remains paramount to keep human convictions out of analyses, whether we believe that our Universe is adequately described by the ΛCDM concordance model, or that the discrepancy in $H_0$ measurements reveals further unknowns in physics, or some combination of the two. Only the fullness of time, and new data and analyses, will reveal the answer.

# Bibliography


Abadi M., et al., 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, `https://www.tensorflow.org/`

Abazajian K., et al., 2003, AJ, 126, 2081

Abazajian K., et al., 2004, AJ, 128, 502

Abbott B. P., et al., 2016, Physical Review Letters, 116, 061102

Abbott B. P., et al., 2017, Nature, 551, 85

Abbott T. M. C., et al., 2018, preprint, (`arXiv:1801.03181`)

Addison G. E., Watts D. J., Bennett C. L., Halpern M., Hinshaw G., Weiland J. L., 2018, ApJ, 853, 119

Alam S., et al., 2017, MNRAS, 470, 2617

Alard C., Lupton R. H., 1998, ApJ, 503, 325

Alcock C., et al., 2000, ApJ, 542, 281

Andersen P., 2018, PhD thesis

Arcavi I., et al., 2016, ApJ, 819, 35

Astier P., et al., 2006, A&A, 447, 31

Aubourg É., et al., 2015, PhysRevD, 92, 123516

Baade W., 1938, ApJ, 88, 285

Bacon D. J., Refregier A. R., Ellis R. S., 2000, MNRAS, 318, 625

Baker T., Ferreira P., Skordis C., 2014, PhysRevD, 89, 024026

Barbary K., et al., 2016, SNCosmo: Python library for supernova cosmology, Astrophysics Source Code Library (ascl:1611.017)

Barrelet E., 2016, A&A, 594, A38

Barrelet E., Juramy C., 2008, Nuclear Instruments and Methods in Physics Research A, 585, 93

Bassett B., Hlozek R., 2010, Baryon acoustic oscillations. p. 246

Baumann D., Peiris H. V., 2008, preprint, (`arXiv:0810.3022`)

Bazin G., et al., 2009, A&A, 499, 653

Bazin G., et al., 2011, A&A, 534, A43

Becker A., 2015, HOTPANTS: High Order Transform of PSF ANd Template Subtraction, Astrophysics Source Code Library (ascl:1504.004)

Bellm E., 2014, in Wozniak P. R., Graham M. J., Mahabal A. A., Seaman R., eds, The Third Hot-wiring the Transient Universe Workshop. pp 27–33 (`arXiv:1410.8185`)

Benítez N., 2000, ApJ, 536, 571

Bennett C. L., Larson D., Weiland J. L., Hinshaw G., 2014, ApJ, 794, 135

Bernal J. L., Verde L., Riess A. G., 2016, JCAP, 10, 019

Bertone G., Hooper D., 2016, preprint, (arXiv:1605.04909)

Bertone G., Hooper D., Silk J., 2005, Phys. Rep., 405, 279

Bessell M. S., 1990, PASP, 102, 1181

Bessell M. S., 2005, ARA&A, 43, 293

Bessell M., Bloxham G., Schmidt B., Keller S., Tisserand P., Francis P., 2011, PASP, 123, 789

Betoule M., et al., 2013, A&A, 552, A124

Betoule M., et al., 2014, A&A, 568, A22

Beutler F., et al., 2011, MNRAS, 416, 3017

Beutler F., et al., 2017, MNRAS, 466, 2242

Bildsten L., Shen K. J., Weinberg N. N., Nelemans G., 2007, ApJL, 662, L95

Bird J. C., Stanek K. Z., Prieto J. L., 2009, ApJ, 695, 874

Blake C., Glazebrook K., 2003, ApJ, 594, 665

Blake C., et al., 2011, MNRAS, 418, 1707

Blondin S., Tonry J. L., 2007, ApJ, 666, 1024

Bloom J. S., et al., 2012, PASP, 124, 1175

Blumenthal G. R., Faber S. M., Primack J. R., Rees M. J., 1984, Nature, 311, 517

Bohlin R. C., 2000, AJ, 120, 437

Bolzonella M., Miralles J.-M., Pelló R., 2000, A&A, 363, 476

Bonvin V., et al., 2017, MNRAS, 465, 4914

Branch D., Tammann G. A., 1992, ARA&A, 30, 359

Branch D., Fisher A., Nugent P., 1993, AJ, 106, 2383

Branch D., Livio M., Yungelson L. R., Boffi F. R., Baron E., 1995, PASP, 107, 1019

Brandt T. D., 2016, ApJL, 824, L31

Brink H., Richards J. W., Poznanski D., Bloom J. S., Rice J., Negahban S., Wainwright M., 2013, MNRAS, 435, 1047

Brout D., Coauthors M., Dark Energy Survey Collaboration in prep.a, DES Supernova Analysis, Systematics, and Validation

Brout D., Coauthors M., Dark Energy Survey Collaboration in prep.b, DES Supernova Photometry and Lightcurves

Buchner J., et al., 2014, A&A, 564, A125

Burke D. L., et al., 2018, AJ, 155, 41

Buta R. J., Turner A., 1983, PASP, 95, 72

Calcino J., Davis T., 2017, JCAP, 1, 038

Campbell H., et al., 2013, ApJ, 763, 88

Cardona W., Kunz M., Pettorino V., 2017, JCAP, 3, 056

Carrick J., Turnbull S. J., Lavaux G., Hudson M. J., 2015, MNRAS, 450, 317

Casertano S., et al., 2016, ApJ, 825, 11

Casertano S., Riess A. G., Bucciarelli B., Lattanzi M. G., 2017, A&A, 599, A67

Chandrasekhar S., 1931, ApJ, 74, 81

Childress M., et al., 2013, ApJ, 770, 108

Childress M. J., et al., 2017, MNRAS, 472, 273

Chotard N., et al., 2011, A&A, 529, L4

Conley A., et al., 2011, ApJS, 192, 1

Contreras C., et al., 2010, AJ, 139, 519

Courbin F., et al., 2018, A&A, 609, A71

Cousins A. W. J., 1976, Memoirs of the Royal Astronomical Society, 81, 25

Croft R. A. C., Dailey M., 2011, preprint, (arXiv:1112.3108)

D'Andrea C. B., Coauthors M., Dark Energy Survey Collaboration in prep., DES Supernova Spectroscopy

Dark Energy Survey Collaboration in prep., Cosmology Results from Spectroscopically Confirmed Type Ia Supernovae in the First Three Years of the Dark Energy Survey

Dark Energy Survey Collaboration et al., 2016, MNRAS, 460, 1270

Dark Energy Survey Collaboration et al., 2017, preprint, (arXiv:1708.01530)

Davies R. L., Efstathiou G., Fall S. M., Illingworth G., Schechter P. L., 1983, ApJ, 266, 41

Davis T. M., et al., 2011, ApJ, 741, 67

Dhawan S., Jha S. W., Leibundgut B., 2018, A&A, 609, A72

Di Stefano R., Kilic M., 2012, ApJ, 759, 56

Di Valentino E., Melchiorri A., Silk J., 2016, Physics Letters B, 761, 242

Dilday B., et al., 2008, ApJ, 682, 262

Drlica-Wagner A., et al., 2018, ApJS, 235, 33

Dvorkin C., Wyman M., Rudd D. H., Hu W., 2014, PhysRevD, 90, 083503

Eddington A. S., 1917, The Observatory, 40, 290

Eddington A. S., 1919, The Observatory, 42, 119

Efstathiou G., 2014, MNRAS, 440, 1138

Efstathiou G., Sutherland W. J., Maddox S. J., 1990, Nature, 348, 705

Einasto J., 2001, in Martínez V. J., Trimble V., Pons-Bordería M. J., eds, Astronomical Society of the Pacific Conference Series Vol. 252, Historical Development of Modern Cosmology. p. 85 (arXiv:astro-ph/0012161)

Einasto J., Saar E., Kaasik A., Chernin A. D., 1974, Nature, 252, 111

Eisenstein D. J., Hu W., 1998, ApJ, 496, 605

Eisenstein D. J., et al., 2005, ApJ, 633, 560

Faber S. M., Jackson R. E., 1976, ApJ, 204, 668

Feeney S. M., J Mortlock D., Dalmasso N., 2018a, MNRAS,

Feeney S. M., Peiris H. V., Williamson A. R., Nissanke S. M., Mortlock D. J., Alsing J., Scolnic D., 2018b, preprint, (arXiv:1802.03404)

Feroz F., Hobson M. P., 2008, MNRAS, 384, 449

Feroz F., Hobson M. P., Bridges M., 2009, MNRAS, 398, 1601

Feroz F., Hobson M. P., Cameron E., Pettitt A. N., 2013, preprint, (arXiv:1306.2144)

Filippenko A. V., 1997, ARA&A, 35, 309

Filippenko A. V., et al., 1992a, AJ, 104, 1543

Filippenko A. V., et al., 1992b, ApJL, 384, L15

Filippenko A. V., Li W. D., Treffers R. R., Modjaz M., 2001, in Paczynski B., Chen W.-P., Lemme C., eds, Astronomical Society of the Pacific Conference Series Vol. 246, IAU Colloq. 183: Small Telescope Astronomy on Global Scales. p. 121

Fink M., Röpke F. K., Hillebrandt W., Seitenzahl I. R., Sim S. A., Kromer M., 2010, A&A, 514, A53

Foley R. J., et al., 2013, ApJ, 767, 57

Foley R. J., et al., 2018, MNRAS, 475, 193

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., Madore B. F., 2010, ARA&A, 48, 673

Freedman W. L., Madore B. F., 2011, ApJ, 734, 46

Freedman W. L., et al., 2001, ApJ, 553, 47

Freedman W. L., et al., 2011, AJ, 142, 192

Freeman K. C., 1970, ApJ, 160, 811

Friedmann A., 1922, Zeitschrift fur Physik, 10, 377

Frieman J. A., Turner M. S., Huterer D., 2008, ARA&A, 46, 385

Gal-Yam A., 2012, Science, 337, 927

Gamow G., 1946, Physical Review, 70, 572

Ganeshalingam M., et al., 2010, ApJS, 190, 418

Ganeshalingam M., Li W., Filippenko A. V., 2013, MNRAS, 433, 2240

Goldstein D. A., et al., 2015, AJ, 150, 82

Goobar A., et al., 2014, ApJL, 784, L12

Guidorzi C., et al., 2017, ApJL, 851, L36

Gull S., 1989, in Skilling J., ed., , Vol. 36, Maximum Entropy and Bayesian Methods. Springer, Dordrecht, pp 511–518, doi:10.1007/978-94-015-7860-8˙55

Gupta R. R., et al., 2016, AJ, 152, 154

Guth A. H., 1981, PhysRevD, 23, 347

Guy J., Astier P., Nobili S., Regnault N., Pain R., 2005, A&A, 443, 781

Guy J., et al., 2007, A&A, 466, 11

Guy J., et al., 2010, A&A, 523, A7

Hamacher D. W., 2014, Journal of Astronomical History and Heritage, 17, 161

Hamuy M., et al., 1993, AJ, 106, 2392

Hamuy M., Phillips M. M., Suntzeff N. B., Schommer R. A., Maza J., Aviles R., 1996, AJ, 112, 2391

Hanany S., et al., 2000, ApJL, 545, L5

Hicken M., et al., 2009a, ApJ, 700, 331

Hicken M., Wood-Vasey W. M., Blondin S., Challis P., Jha S., Kelly P. L., Rest A., Kirshner R. P., 2009b, ApJ, 700, 1097

Hicken M., et al., 2012, ApJS, 200, 12

Hillebrandt W., Kromer M., Röpke F. K., Ruiter A. J., 2013, Frontiers of Physics, 8, 116

Hinshaw G., et al., 2003, ApJS, 148, 135

Hinton S. R., 2016, The Journal of Open Source Software, 2016, 00045

Hinton S. R., Kim A. G., Davis T. M., in prep., Steve-o: A Hierarchical Bayesian Model for Supernova Cosmology

Hobson M. P., Jaffe A. H., Liddle A. R., Mukherjee P., Parkinson D., 2014, Bayesian Methods in Cosmology

Hoffmann S. L., et al., 2016, ApJ, 830, 10

Holtzman J. A., et al., 2008, AJ, 136, 2306

Holz D. E., Hughes S. A., 2005, ApJ, 629, 15

Howell D. A., 2011, Nature Communications, 2, 350

Hoyle F., 1948, MNRAS, 108, 372

Hsiao E. Y., Conley A., Howell D. A., Sullivan M., Pritchet C. J., Carlberg R. G., Nugent P. E., Phillips M. M., 2007, ApJ, 663, 1187

Hubble E., 1929, Proceedings of the National Academy of Science, 15, 168

Hui L., Greene P. B., 2006, PhysRevD, 73, 123526

Humphreys E. M. L., Reid M. J., Moran J. M., Greenhill L. J., Argon A. L., 2013, ApJ, 775, 13

Huterer D., 2010, General Relativity and Gravitation, 42, 2177

Inserra C., et al., 2016, MNRAS, 459, 2721

Jennings E., Madigan M., 2017, Astronomy and Computing, 19, 16

Jennings E., Wolf R., Sako M., 2016, preprint, (`arXiv:1611.03087`)

Jha S., et al., 2006, AJ, 131, 527

Jha S., Riess A. G., Kirshner R. P., 2007, ApJ, 659, 122

Johnson H. L., 1966, ARA&A, 4, 193

Jones D. O., Riess A. G., Scolnic D. M., 2015, ApJ, 812, 31

Jones D. O., et al., 2017, ApJ, 843, 6

Jones D. O., et al., 2018, ApJ, 857, 51

Joyce A., Lombriser L., Schmidt F., 2016, Annual Review of Nuclear and Particle Science, 66, 95

Justham S., 2011, ApJL, 730, L34

Kaiser N., 1992, ApJ, 388, 272

Kaiser N., Wilson G., Luppino G. A., 2000, ArXiv Astrophysics e-prints,

Kapteyn J. C., 1922, ApJ, 55, 302

Karpenka N. V., Feroz F., Hobson M. P., 2013, MNRAS, 429, 1278

Kasliwal M. M., 2012, PASA, 29, 482

Kasliwal M. M., et al., 2012, ApJ, 755, 161

Keller S. C., et al., 2007, PASA, 24, 1

Kelly P. L., Hicken M., Burke D. L., Mandel K. S., Kirshner R. P., 2010, ApJ, 715, 743

Kelly P. L., et al., 2016, ApJ, 831, 205

Kennicutt Jr. R. C., et al., 1998, ApJ, 498, 181

Kerzendorf W. E., et al., 2013, ApJ, 774, 99

Kessler R., Scolnic D., 2017, ApJ, 836, 56

Kessler R., et al., 2009a, PASP, 121, 1028

Kessler R., et al., 2009b, ApJS, 185, 32

Kessler R., et al., 2015, AJ, 150, 172

Kilbinger M., 2015, Reports on Progress in Physics, 78, 086901

Kim A. G., 2004, LBNL Report LBNL-56164

Kim A., Goobar A., Perlmutter S., 1996, PASP, 108, 190

Kirshner R. P., Kwan J., 1974, ApJ, 193, 27

Komatsu E., et al., 2009, ApJS, 180, 330

Kowal C. T., 1968, AJ, 73, 1021

Kowal C. T., Sargent W. L. W., Zwicky F., 1971, PASP, 83, 307

Kuchner M. J., Kirshner R. P., Pinto P. A., Leibundgut B., 1994, ApJL, 426, 89

Kuijken K., et al., 2015, MNRAS, 454, 3500

Kunz M., Bassett B. A., Hlozek R. A., 2007, PhysRevD, 75, 103508

Lampeitl H., et al., 2010, ApJ, 722, 566

Landolt A. U., 1983, AJ, 88, 439

Landolt A. U., 1992, AJ, 104, 340

Lasker J., Coauthors M., Dark Energy Survey Collaboration in prep., DES FGCM Supernova Chromatic Corrections

Laurino O., D'Abrusco R., Longo G., Riccio G., 2011, MNRAS, 418, 2165

Law N. M., et al., 2009, PASP, 121, 1395

Le Borgne D., Rocca-Volmerange B., 2002, A&A, 386, 446

Le Borgne D., Rocca-Volmerange B., 2010, ZPEG: An Extension of the Galaxy Evolution Model PEGASE.2, Astrophysics Source Code Library (ascl:1011.003)

Leavitt H. S., 1908, Annals of Harvard College Observatory, 60, 87

Leavitt H. S., Pickering E. C., 1912, Harvard College Observatory Circular, 173, 1

Leibundgut B., Tammann G. A., Cadonau R., Cerrito D., 1991, A&AS, 89, 537

Leibundgut B., et al., 1993, AJ, 105, 301

Leistedt B., Peiris H. V., Verde L., 2014, Physical Review Letters, 113, 041301

Lemaître G., 1927, Annales de la Société Scientifique de Bruxelles, 47, 49

Lemaitre G., 1934, Proceedings of the National Academy of Science, 20, 12

Lewis A., Bridle S., 2002, PhysRevD, 66, 103511

Li W., Qiu Y., Qiao Q., Hu J., Li Q., 1999, Science in China A: Mathematics, 42, 1075

Li W. D., et al., 2000, in Holt S. S., Zhang W. W., eds, American Institute of Physics Conference Series Vol. 522, American Institute of Physics Conference Series. pp 103–106 (`arXiv:astro-ph/9912336`), doi:10.1063/1.1291702

Li W., et al., 2011a, MNRAS, 412, 1441

Li W., et al., 2011b, Nature, 480, 348

Linde A., 2014, preprint, (`arXiv:1402.0526`)

Linder E. V., 2005, PhysRevD, 72, 043529

Lira P., et al., 1998, AJ, 115, 234

Macaulay E., Nichol R. C., Bacon D., Collett T. E., Dark Energy Survey Collaboration in prep., First Cosmological Analysis using Type Ia Supernovae from the Dark Energy Survey: Measurement of the Hubble Constant

Maccoun R., Perlmutter S., 2015, Nature, 526, 187

Macri L. M., Stanek K. Z., Bersier D., Greenhill L. J., Reid M. J., 2006, ApJ, 652, 1133

Madore B. F., 1982, ApJ, 253, 575

Maeda K., Terada Y., 2016, International Journal of Modern Physics D, 25, 1630024

Mandel K. S., Wood-Vasey W. M., Friedman A. S., Kirshner R. P., 2009, ApJ, 704, 629

Mandel K. S., Scolnic D. M., Shariff H., Foley R. J., Kirshner R. P., 2017, ApJ, 842, 93

Mandelbaum R., 2017, preprint, (`arXiv:1710.03235`)

March M. C., Trotta R., Berkes P., Starkman G. D., Vaudrevange P. M., 2011, MNRAS, 418, 2308

Mather J. C., et al., 1994, ApJ, 420, 439

Mazzali P. A., Röpke F. K., Benetti S., Hillebrandt W., 2007, Science, 315, 825

Minkowski R., 1941, PASP, 53, 224

Miralda-Escude J., 1991, ApJ, 380, 1

Möller A., et al., 2016, JCAP, 12, 008

Morganson E., et al., 2018, PASP, 130, 074501

Mosher J., et al., 2014, ApJ, 793, 16

Muthukrishna D., Parkinson D., Tucker B. E., in prep., DASH: Deep Learning for the Spectral Classificationo of Supernovae

Neill J. D., et al., 2009, ApJ, 707, 1449

Norgaard-Nielsen H. U., Hansen L., Jorgensen H. E., Aragon Salamanca A., Ellis R. S., 1989, Nature, 339, 523

Nugent P., Kim A., Perlmutter S., 2002, PASP, 114, 803

Nugent P. E., et al., 2011, Nature, 480, 344

Oke J. B., 1965, ARA&A, 3, 23

Oort J. H., 1932, Bulletin of the Astronomical Institutes of the Netherlands, 6, 289

Ostriker J. P., Steinhardt P. J., 1995, Nature, 377, 600

Padmanabhan N., et al., 2008, ApJ, 674, 1217

Pakmor R., Kromer M., Taubenberger S., Sim S. A., Röpke F. K., Hillebrandt W., 2012, ApJL, 747, L10

Peebles P. J. E., 1993, Principles of Physical Cosmology

Penzias A. A., Wilson R. W., 1965, ApJ, 142, 419

Perlmutter S., Muller R. A., Newberg H. J. M., Pennypacker C. R., Sasseen T. P., Smith C. K., 1992, in Filippenko A. V., ed., Astronomical Society of the Pacific Conference Series Vol. 34, Robotic Telescopes in the 1990s. pp 67–71

Perlmutter S., et al., 1999, ApJ, 517, 565

Perrett K., et al., 2010, AJ, 140, 518

Persson S. E., Madore B. F., Krzemiński W., Freedman W. L., Roth M., Murphy D. C., 2004, AJ, 128, 2239

Phillips M. M., 1993, ApJL, 413, L105

Phillips M. M., Wells L. A., Suntzeff N. B., Hamuy M., Leibundgut B., Kirshner R. P., Foltz C. B., 1992, AJ, 103, 1632

Pietrzyński G., et al., 2013, Nature, 495, 76

Pike R. W., Hudson M. J., 2005, ApJ, 635, 11

Piro A. L., Nakar E., 2014, ApJ, 784, 85

Planck Collaboration et al., 2014, A&A, 571, A16

Planck Collaboration et al., 2016, A&A, 594, A13

Primack J. R., 1997, ArXiv Astrophysics e-prints,

Qiu Y., Li W., Zhao Z., Qiao Q., Rao Y., Hu J., Li Q., 1999, Science in China A: Mathematics, 42, 220

Rasmussen C. E., Williams C. K. I., 2006, Gaussian Processes for Machine Learning

Rau A., et al., 2009, PASP, 121, 1334

Regnault N., et al., 2009, A&A, 506, 999

Regnault N., Barrelet E., Guyonnet A., Juramy C., Rocci P.-F., Le Guillou L., Schahmanèche K., Villa F., 2016, in Deustua S., Allam S., Tucker D., Smith J. A., eds, Astronomical Society of the Pacific Conference Series Vol. 503, The Science of Calibration. p. 165 (`arXiv:1208.6301`)

Rest A., et al., 2014, ApJ, 795, 44

Richardson D., Branch D., Casebeer D., Millard J., Thomas R. C., Baron E., 2002, AJ, 123, 745

Riess A. G., Press W. H., Kirshner R. P., 1996, ApJ, 473, 88

Riess A. G., et al., 1998, AJ, 116, 1009

Riess A. G., et al., 1999, AJ, 117, 707

Riess A. G., et al., 2005, ApJ, 627, 579

Riess A. G., et al., 2009a, ApJS, 183, 109

Riess A. G., et al., 2009b, ApJ, 699, 539

Riess A. G., et al., 2011, ApJ, 730, 119

Riess A. G., et al., 2012, ApJ, 752, 76

Riess A. G., et al., 2016, ApJ, 826, 56

Rigault M., et al., 2013, A&A, 560, A66

Rigault M., et al., 2015, ApJ, 802, 20

Roodman A., 2003, in Lyons L., Mount R., Reitmeyer R., eds, Statistical Problems in Particle Physics, Astrophysics, and Cosmology. p. 166 (`arXiv:physics/0312102`)

Rubin V. C., Ford Jr. W. K., 1970, ApJ, 159, 379

Rubin D., et al., 2015, ApJ, 813, 137

Ruiz-Lapuente P., 2014, New Astronomy Reviews, 62, 15

Sakai S., Ferrarese L., Kennicutt Jr. R. C., Saha A., 2004, ApJ, 608, 42

Sako M., et al., 2014, preprint, (`arXiv:1401.3317`)

Salvato M., et al., 2009, ApJ, 690, 1250

Scalzo R. A., Ruiter A. J., Sim S. A., 2014, MNRAS, 445, 2535

Scalzo R. A., et al., 2017, PASA, 34, e030

Schlafly E. F., Finkbeiner D. P., 2011, ApJ, 737, 103

Schlegel D. J., Finkbeiner D. P., Davis M., 1998, ApJ, 500, 525

Schmidt B. P., Kirshner R. P., Eastman R. G., 1992, ApJ, 395, 366

Schmidt B. P., et al., 1998, ApJ, 507, 46

Schutz B. F., 1986, Nature, 323, 310

Scolnic D., Kessler R., 2016, ApJL, 822, L35

Scolnic D., et al., 2014, ApJ, 795, 45

Scolnic D., et al., 2015, ApJ, 815, 117

Scolnic D. M., et al., 2018, ApJ, 859, 101

Scowcroft V., Bersier D., Mould J. R., Wood P. R., 2009, MNRAS, 396, 1287

Scowcroft V., Freedman W. L., Madore B. F., Monson A. J., Persson S. E., Seibert M., Rigby J. R., Sturch L., 2011, ApJ, 743, 76

Scrimgeour M. I., et al., 2016, MNRAS, 455, 386

Sebo K. M., et al., 2002, ApJS, 142, 71

Shariff H., Jiao X., Trotta R., van Dyk D. A., 2016, ApJ, 827, 1

Shen K. J., Kasen D., Weinberg N. N., Bildsten L., Scannapieco E., 2010, ApJ, 715, 767

Silverman J. M., et al., 2012, MNRAS, 425, 1789

Silverman J. M., et al., 2013, ApJS, 207, 3

Skilling J., 2004, in Fischer R., Preuss R., Toussaint U. V., eds, American Institute of Physics Conference Series Vol. 735, American Institute of Physics Conference Series. pp 395–405, doi:10.1063/1.1835238

Slipher V. M., 1915, Popular Astronomy, 23, 21

Smith S., 1936, ApJ, 83, 23

Smith A. M., et al., 2011, MNRAS, 412, 1309

Smith M., et al., 2012, ApJ, 755, 61

Smith M., Coauthors M., Dark Energy Survey Collaboration in prep., DES Supernova Host Galaxies

Smoot G. F., et al., 1992, ApJL, 396, L1

Spergel D. N., Flauger R., Hložek R., 2015, PhysRevD, 91, 023518

Springel V., et al., 2005, Nature, 435, 629

Springob C. M., et al., 2014, MNRAS, 445, 2677

Stehle M., Mazzali P. A., Benetti S., Hillebrandt W., 2005, MNRAS, 360, 1231

Stritzinger M. D., et al., 2011, AJ, 142, 156

Sullivan M., et al., 2006, ApJ, 648, 868

Sullivan M., et al., 2010, MNRAS, 406, 782

Sullivan M., et al., 2011, ApJ, 737, 102

Suyu S. H., et al., 2017, MNRAS, 468, 2590

The Dark Energy Survey Collaboration et al., 2017, preprint, (`arXiv:1712.06209`)

Tonry J., Davis M., 1979, AJ, 84, 1511

Tripp R., 1998, A&A, 331, 815

Tully R. B., Fisher J. R., 1977, A&A, 54, 661

Van Waerbeke L., et al., 2000, A&A, 358, 30

Vega-Ferrero J., Diego J. M., Miranda V., Bernstein G. M., 2018, ApJL, 853, L31

Walker E. S., et al., 2015, ApJS, 219, 13

Wang L., Steinhardt P. J., 1998, ApJ, 508, 483

Weinberg D. H., Mortonson M. J., Eisenstein D. J., Hirata C., Riess A. G., Rozo E., 2013, Phys. Rep., 530, 87

Wittman D. M., Tyson J. A., Kirkman D., Dell'Antonio I., Bernstein G., 2000, Nature, 405, 143

Wolf R. C., Jennings E., Sako M., Kovacs E., Coauthors M., in prep., BAMBIS: Bayesian Hierarchical Modeling with Biased Simulations for SN Ia Cosmology Including Selection Effects

Wood-Vasey W. M., et al., 2007, ApJ, 666, 694

Woosley S. E., Weaver T. A., 1994, ApJ, 423, 371

Wu H.-Y., Huterer D., 2017, MNRAS, 471, 4946

Wyman M., Rudd D. H., Vanderveld R. A., Hu W., 2014, Physical Review Letters, 112, 051302

Yuan F., et al., 2015, MNRAS, 452, 3047

Zhang B. R., Childress M. J., Davis T. M., Karpenka N. V., Lidman C., Schmidt B. P., Smith M., 2017, MNRAS, 471, 2254

Zwicky F., 1933, Helvetica Physica Acta, 6, 110

Zwicky F., 1937, ApJ, 86, 217

Zwicky F., 1964, Annales d'Astrophysique, 27, 300

de Bernardis P., et al., 2000, Nature, 404, 955

van Leeuwen F., Feast M. W., Whitelock P. A., Laney C. D., 2007, MNRAS, 379, 723

van den Bergh S., 1975, Physics Bulletin, p. 509