Title
Improvements to Type Ia Supernova Models

Author
Saunders, Clare Myers

Publication Date
2016

Peer reviewed|Thesis/dissertation

Improvements to Type Ia Supernova Models

by

Clare M. Saunders

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Saul Perlmutter, Chair
Professor Daniel Kasen
Professor Joshua Bloom

Summer 2016

Improvements to Type Ia Supernova Models

Abstract

Improvements to Type Ia Supernova Models

by

Clare M. Saunders

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Saul Perlmutter, Chair

Type Ia Supernovae provided the first strong evidence of dark energy and are still an important tool for measuring the accelerated expansion of the universe. However, future improvements will be limited by systematic uncertainties in our use of Type Ia supernovae as standard candles. Using Type Ia supernovae for cosmology relies on our ability to standardize their absolute magnitudes, but this relies on imperfect models of supernova spectra time series. This thesis is focused on using data from the Nearby Supernova Factory both to understand current sources of uncertainty in standardizing Type Ia supernovae and to develop techniques that can be used to limit uncertainty in future analyses.

Part I of this thesis is concerned with studying systematic errors that occur in the current generation of supernova lightcurve models. Lightcurve models are used to fit photometric supernova data, generating a small number of parameters that can then be used to 'correct' the supernova magnitude to a standard value. The analysis presented here estimates systematic errors due to K-corrections that occur when using such lightcurve models to fit high-redshift supernovae. These errors occur when the spectral template underlying the lightcurve fitter poorly matches the actual supernova spectral energy distribution and also if the model fit is dependent on the spectral coverage of the photometric filters used to observe the supernova.

In order to quantify this effect, synthetic photometry is performed on artificially redshifted spectrophotometric data from the Nearby Supernova Factory, simulating having photometric data for the same supernovae with a range of filter positions. The resulting lightcurves are fit with a conventional lightcurve fitter and the variation is measured in the resulting standardized magnitudes. We find significant variation in the measurements of the same supernovae placed at different redshifts regardless of filters used, which causes dispersion greater than ~ 0.05 mag for measurements of photometry using Sloan-like filters and a bias that corresponds to a 0.03 shift in w when applied to an outside data set. We also test the effect of population changes at high redshift and measure the resulting bias for the average of the supernova magnitudes. Lastly, methods are discussed for mitigating bias and dispersion due to K-corrections.

Part II presents an alternative to current lightcurve models. The supernovae from the Nearby Supernova Factory are used to develop an empirical model that will be able to more fully describe the variety of supernova behavior. The spectrophotometric time series for over 200 supernova are used to fit linear principal components that will be able to be used as a lightcurve model. This is done by first using Gaussian Processes to model the true spectral time series for each supernova and make a prediction for it on a regular grid in time and wavelength space. Once this has been done, principal components are calculated that describe the full set of supernovae using a method that incorporates the variance in the data. K-fold cross-validation is used to determine how many components best describe the full population without over-training on noise in the data. In the final version of the analysis, three different models are presented: one simple model that can be compared to the current generation of lightcurve models; one model that is the best for performing a linear standardization of supernova magnitudes following current standardization methods, at least when spectra for the supernova are available; and one complex model that provides the most complete model of spectral time series for the full population.

To my parents,

Carol and Richard

Contents

List of Figures

List of Tables

Acknowledgments

Many thanks are due to my advisor, Saul Perlmutter, whose enthusiasm for supernova cosmology drew me into the field. Alex Kim and Greg Aldering have also been valued advisors to me and I am very grateful for the time they have spent working with me. All three have set an example for me of indefatigable devotion to progress in our field.

None of the work presented here would be possible without the Nearby Supernova Factory collaboration. I have been very lucky to work with this international group of scientists and to benefit from the years of work that was done before I joined the group. Our collaboration meetings have let me see more of the world and taught me how much fun science can be.

My research group has been a much appreciated resource and has always provided an enjoyable work environment. Thanks especially to David Rubin, Brian Hayden, Rollin Thomas, and Kyle Barbary for always being willing to talk about research problems or answer my questions, and to Jakob Nordin, who has always been a sensible source of advice and encouragement.

Thanks is also due to the ARCS Foundation, which provided me with a fellowship in my first two years of graduate school. This funding allowed me to focus on my classwork and gave me flexibility in choosing a research group, along with freeing me from teaching responsibilities. Lawrence Berkeley National Laboratory has been a wonderful place to work for the last five years and has provided me with the best office that a graduate student could ever hope for.

I would like to thank my parents, Carol and Richard, who raised me to have the confidence that I could accomplish whatever I set out to do, and my sister Anne, whose looking up to me has given me something to live up to. Lastly, thanks to Alessandro Colavecchio, who did so much work planning our wedding so that I could write this thesis.

Chapter 1

An Overview of Supernova Cosmology

1.1 Cosmology with Standard Candles

Understanding the nature of dark energy and the history of our universe are some of the main goals of cosmology and among the greatest outstanding questions in physics today. Progress has been made in the past few decades through the use of a number of observational tools, among which Type Ia Supernovae, the Cosmic Microwave Background, and Baryon Acoustic Oscillations are the most powerful. These each provide complementary constraints on the nature of dark energy through different observational signatures.

We will be concerned here with the first of these, Type Ia Supernovae, which can be used as standard candles. A standard candle is arguably the most conceptually simple method for measuring the expansion of space: comparing the brightness of two objects known to have the same luminosity provides a measurement of their relative distances. This can be compared with the redshifts of the observed light, a proxy for how much space has expanded since the time the light was emitted, to calculate the expansion history of the universe.

The expansion history of the universe can be described by a single variable, $a(t)$, a scale factor that parametrizes the distance between any two points in space-time. The expansion rate of the universe, called the Hubble parameter H, can then be equated to the change in $a(t) = \dot{a}(t)$ in relation to $a(t)$, giving the Friedmann equation

$$H^2 = (\dot{a}/a)^2 = \frac{8\pi G\rho}{3} - \frac{k}{a^2} \tag{1.1}$$

Observations of standard candles are able to constrain H through measurements of their luminosity distance D_L, which describes their apparent brightness as a function of redshift:

$$D_L = \frac{c}{H_0}(1+z)\kappa_0^{-1/2} S\Big\{\kappa_0^{-1/2}\int_0^z dz'\Big[\sum_i \Omega_i(1+z')^{3+3w_i} - \kappa_0(1+z')^2\Big]^{-1/2}\Big\} \tag{1.2}$$

where $S(x) = \sin(x)$, x, or $\sinh(x)$ depending on whether the model of the universe is closed, flat, or open, respectively, and κ_0 is the curvature parameter defined as $\kappa_0 = \sum_i \Omega_i - 1$.

H_0 is the value of the Hubble parameter at the present epoch. The parameters Ω_i correspond to the current fractional densities of the components of the universe, namely radiation, matter, and dark energy, while $w_i = p_i/\rho_i$ are their corresponding equations of state relating the density and pressure. For supernova observations, D_L is then commonly converted to the distance modulus μ,

$$\mu \equiv m - M = 5\log_{10}(D_L) - 5 = 5\log_{10}\left(\frac{D_L}{10\text{pc}}\right) \tag{1.3}$$

where m and M are the (observed) apparent and absolute magnitudes of the object.

Thus, given a large ensemble of objects with known absolute magnitudes, in other words standard candles, distributed over a range of redshifts, it is possible to constrain the Hubble parameter and the model of dark energy, while measurements of D_L itself (by means of a distance ladder) allow us to measure H_0.

1.2 Overview of Type Ia Supernovae

In reality, Type Ia Supernovae are not perfect standard candles, and using them for cosmology requires a greater understanding of their nature. Supernovae and novae have been observed as "guest stars" since perhaps as early as AD 185 and possibly longer (Zhao, Strom, and Jiang, 2006). Baade and Zwicky (1934) separated these transient events into novae and supernovae, and Minkowski (1941) classified them further by dividing supernovae into two groups based on their spectra: Type I, which have no hydrogen in their spectra, and Type II, which do have hydrogen. For cosmology, the crucial division is between Type Ia supernovae, which are uniquely identified by strong SiII, except for some peculiar cases, and the rest of the subtypes (Elias et al., 1985). Although the progenitor has never been observed, Type Ia supernovae are believed to be the result of the thermonuclear disruption of a mass-accreting carbon-oxygen white dwarf, while the rest of the sub-types are due to the core-collapse of a massive star.

While Type II supernovae may theoretically be used for cosmology through the expanding photosphere method (first developed in Kirshner and Kwan, 1974 and explored more recently in Gall et al., 2016), Type Ia Supernovae are the best available standard candle because of their great brightness, which allows them to be observed at high enough redshifts to be useful for cosmology, and because their peak absolute magnitude in the B-band has a relatively small dispersion of about $\sigma(M_B) = 0.35$ (Sandage and Tammann, 1993). This can be significantly reduced by correcting magnitudes by finding correlations between the magnitudes and other parameters of the supernova, discussed further below.

Broad-band supernova behavior is driven by the radioactive decay of ^{56}Ni to ^{56}Co and then ^{56}Fe, products of the thermonuclear fusion. This is observed as a lightcurve rise time of about 20 days, followed by a long decline, with a second peak in the redder bands at about 15 days after the peak of the lightcurve. At maximum light, the majority ($\approx 85\%$) of the light from the supernova is in the optical, peaking at ≈ 4000Å (Howell et al., 2009). Spectral

features in the supernova are P-Cygni profiles with strong absorption and weak emission at early times and shifting to predominantly emission later on, which through the peak brightness phase are primarily from the intermediate mass elements (Si, S, Ca, and Mg), where burning from the thermonuclear fusion was incomplete, and C and O, where unburned material remains, then are dominated by iron group elements at later phases (Hillebrandt and Niemeyer, 2000).

Differences in Type Ia supernovae are driven by varying amounts of ^{56}Ni changing the shape of the lightcurve, and varying abundances and velocities of the elements causing the absorption and emission features.

1.3 Type Ia Supernova Standardization with Empirical Models

The variations among supernovae discussed above are what prevent them from being perfect standard candles. However, correlations between the supernova characteristics and absolute magnitudes enable us to 'correct' for their differences, thus allowing us to turn Type Ia Supernovae into standard candles. While a perfect understanding of Type Ia Supernovae would make it possible for their magnitudes to be fully standardized, we are limited computationally from making accurate theoretical models and limited observationally by the time and expense needed to acquire comprehensive data. At this time, theoretical models are far from accurately replicating supernova behavior at the level needed to precisely match the supernova luminosity, and so for cosmology all standardization is done using empirical methods based on data.

Early progress was made in standardizing magnitudes by correcting them for host-galaxy exctinction (Miller and Branch, 1990), providing motivation for further study of using Type Ia supernovae as standard candles. The first step towards standardization as it is currently performed was made possible by the discovery of a correlation between the peak B-band magnitude of a supernova and the decline rate of the B-band lightcurve, typically measured by the difference in the magnitudes at peak and 15 days after peak, $\Delta \mathrm{m}_{15}$ (Phillips, 1993). Fitting a linear relation between these two parameters and subtracting it from the peak magnitude results in a 'corrected' magnitude with a lower dispersion. A further correlation was found in Tripp (1998) between the peak magnitude and the difference in the B and V band magnitudes (called the $B-V$ color) at peak. The effectiveness of this standardization was proven when it was used in the first detections of dark energy in Riess et al. (1998) and Perlmutter et al. (1999).

This method of standardization is limited by both the time of observations and the redshift of the supernova. For cosmology, supernovae need to be observed at a range of redshifts, which means that photometric observations will cover different parts of a supernova's rest-frame spectrum. For standardization, however, a common rest-frame B-band is needed. To approximate this, a template spectrum is used to try to 'correct' the observer-frame photom-

etry to the supernova rest-frame. This is done by redshifting and warping the template to match the observed photometry, then using that spectrum to calculate what the photometry would be in the desired bandpass. This process is called a K-correction. Template spectral time series or template photometry is also needed to interpolate from the observed data points to the date of maximum and to find Δm_{15}. The templates are typically stretched in the time dimension to find the best match to the observed data. Both of these procedures lead to errors and subsequent uncertainty due to mismatches between the template spectrum or spectral time series and the true supernova behavior.

Current methods of supernova standardization ameliorate these issues by taking a more holistic approach to fitting supernova data. Supernova lighturve fitters such as SALT2 (Guy et al., 2007) are spectral time series models with parameters that can be tuned to fit the data. The standardization is done by fitting out the correlation between the model parameters and the model's peak magnitude, instead of using data plus a correction. Because the model is a spectral time series, no explicit K-corrections and only small interpolations are needed.

Nevertheless, these current lightcurve fitters are still based on a two-parameter model for supernovae, leaving a large 'intrinsic dispersion' left in the standardized magnitudes due to variations in the supernovae unexplained by the model. Additionally, recent studies have confirmed that supernovae have more diversity than can be modeled by two components (e.g. Kim et al., 2013, Sasdelli et al., 2015). Current and future supernova surveys will reduce statistical errors to less than 0.01 mag in some redshift regions, so it will become crucial to control for systematic errors related to supernova standardization in order to make further improvements in precision. This thesis is focused on analyzing the consequences for cosmology of using a two-parameter supernova model and on how the current model can be improved upon using data from the Nearby Supernova Factory (Aldering et al., 2002). Part I of this work describes an analysis of systematic errors found in calculating supernova corrected magnitudes with a two-parameter lightcurve fitter and was published in Saunders et al., 2015. Part II sets forth a method for creating a new supernova model that will capture more of supernova diversity and presents the results of that model, an analysis that will be the subject of an upcoming article.

Part I

Understanding Systematic Errors in Type Ia Supernova Lightcurve Fitting

Chapter 2

Introduction

The following chapters making up Part I of this work consist of the paper, "Type Ia Supernova Distance Modulus Bias and Dispersion from K-corrections Errors: A Direct Measurement Using Light Curve Fits to Observed Spectral Time Series," which was published in *The Astrophysical Journal* in 2015[1].

For several years now, Type Ia supernovae have no longer been limited by statistical uncertainty in their use as standardized candles. Therefore, in order to improve their use as cosmological tools, it has become necessary to further limit systematic errors to obtain the highest possible accuracy from supernova measurements. This becomes increasingly crucial as we find more supernovae to constrain the properties of dark energy. Modern supernova searches (such as the Supernova Legacy Survey, Astier et al., 2006, and the Dark Energy Survey, Bernstein et al., 2012) typically are designed to find supernovae in a certain redshift range using a fixed set of photometric filters, which means that the supernovae found will be observed with a range of supernova-frame filter configurations. To use these supernovae for cosmology, the observations generally are converted into peak standardized absolute magnitudes in a common band.

This conversion necessarily involves using an estimate of the spectral time series. Originally, photometric observations were converted to a common band by adding an explicit K-correction to the magnitudes, which was calculated with an example spectrum (a process originally developed for use on galaxies in Humason, Mayall, and Sandage, 1956 and Oke and Sandage, 1968). Kim, Goobar, and Perlmutter (1996) extended this method with the development of the cross-filter K-correction, which allowed for conversion between different filters. K-corrections were further improved by Nugent, Kim, and Perlmutter (2002) with a spectral time series template that could be stretched or warped to match a supernova's shape. Once the supernova observations have been K-corrected, they can be fit by a lightcurve template to get the standardized peak absolute magnitude, as done by MLCS2k2 or SNooPy (Jha, Riess, and Kirshner, 2007, Burns et al., 2011).

[1] Saunders et al., 2015.

Modern spectral time series-based lightcurve fitters combine the two-step process of K-correcting and then fitting photometric data. Lightcurve fitters like SALT2 and SiFTO (Guy et al., 2007, Conley et al., 2008) fit observed photometric data with integrals over a time-evolving spectral energy distribution model that can be adjusted by means of free parameters. The standardized magnitude is calculated from the parameters and the rest-frame magnitude of the best-fit model.

In either of these methods, errors in the supernova's magnitude will occur if the estimated spectrum is different from the true spectrum of the observed supernova, either in the spectral shape or the flux normalization with respect to the other phases. This can happen if the diversity of the supernova population is greater than can be encompassed by the model. Further, since the estimated spectrum is based on the available photometry, it can vary as a function of the filters used to observe the supernova, depending on which spectral features fall in each filter band and the relative weighting of different parts of the spectral time series model. While a separate K-correction is only explicitly done in the first type of lightcurve fitter, we will here refer more generally to errors due to inaccurate estimation of the supernova spectrum as K-correction errors.

Variation in the best-fit spectrum propagates into variation in the distance modulus. Uncertainty due to a random scatter in the errors will decrease in proportion to the square root of the number of supernovae observed. However, bias in the ensemble mean of the magnitudes will cause a systematic error in the fitting of the Hubble Diagram that will not be reduced with greater numbers of supernovae. Such errors are thought to be subdominant to other source of error like calibration uncertainty, but have not been exactly tested using a current-generation lightcurve fitter and a large supernova data set.

With photometric data alone, it would be extremely difficult to cleanly measure the size of this effect, but using supernova spectrophotometric time series data it is possible to exactly measure the effect on each supernova. Spectrophotometry can be integrated over any passband to synthesize photometry for the same supernova as if it were observed through different filters or at different redshifts; changes in the lightcurve fit for different filters due to K-correction error are then directly measurable. From the ensemble of supernovae in the data set, the bias and dispersion on that supernova population can be calculated. In this paper, we calculate the redshift-dependent distance modulus variation due to K-corrections by applying a state-of-the-art two-parameter lightcurve fitter to spectrophotometric time series data from the Nearby Supernova Factory (SNfactory, Aldering et al., 2002). The bias and distribution of distance modulus variation over the whole ensemble provide an estimate of the added uncertainty in distance modulus applicable to other supernovae without spectroscopic time series that will be used in cosmological measurements.

In Chapter 3, the SNfactory supernova data set is described. The questions we address with different filter sets are given in Section 4.1, with a description of the procedure used in Section 4.2. Chapter 5 compares the K-correction effects when using these different sets of photometric filters and discusses the effect of potential hidden calibration errors in the sample supernova set. In Chapter 6 we consider the contribution of K-correction effects to the total dispersion seen in supernova standardized magnitudes. We also consider the applications

of our results for future high-redshift programs, where different filter configurations are planned, and with the addition of possible shifts in the makeup of the supernova population in Chapter 7. Finally, in Chapter 8, we discuss the sources of the errors that are seen and we discuss possible methods for minimizing errors due to K-correction.

Chapter 3

The Supernova data set

The data set used in this study consists of spectrophotometric time series for a set of over 100 nearby supernovae, observed by the SNfactory between 2005 and 2009 with the SuperNova Integral Field Spectrograph (SNIFS, Lantz et al., 2004). Most of these supernovae have been previously presented, for example in Thomas et al. (2007, 2011), Bailey et al. (2009), and Chotard et al. (2011). SNIFS is a fully integrated instrument optimized for automated observation of point sources on a structured background over the full ground-based optical window at moderate spectral resolution ($R \sim 500$). It consists of a high-throughput wide-band pure-lenslet integral field spectrograph (IFS; based on the 'Tiger' principle, Bacon et al., 1995, 2000, 2001), a multi-filter photometric channel to image the field in the vicinity of the IFS for atmospheric transmission monitoring simultaneous with spectroscopy, and an acquisition/guiding channel. The IFS possesses a fully-filled $6\farcs4 \times 6\farcs4$ spectroscopic field of view subdivided into a grid of 15×15 spatial elements, a dual-channel spectrograph covering $3200 - 5200$ Å and $5100 - 10000$ Å simultaneously, and an internal calibration unit (continuum and arc lamps). SNIFS is continuously mounted on the South bent Cassegrain port of the University of Hawaii 2.2 m telescope on Mauna Kea and is operated remotely. A description of host-galaxy subtraction is given in Bongard et al. (2011).

The spectra are flux-calibrated following the procedure detailed in Buton et al. (2013) and the color calibration is trusted to within 1% based on observations of standard stars. Comparison between spectra from the same supernovae observed by SNIFS and by the Hubble Space Telescope (Maguire et al., 2012) confirms that any $B - V$ color offset is below this 1% level. In the analysis we have included a dispersion empirically measured in the lightcurve fits using SNfactory and other data sets as an error floor of 0.04 mag on the ability of the lightcurve model to match the photometric data points. The SNfactory supernova spectra are originally calibrated using a hybrid system of both CALSPEC (Bohlin, 2014 and references within) and Hamuy (Hamuy et al., 1992, 1994) standard stars, which are listed in Buton et al. (2013). However, in this analysis the supernovae have been recalibrated to the CALSPEC standards in order to match the calibration of the SALT2 training data. The effect of the calibration on the analysis is discussed further in Chapter 5.

The supernovae in our data set range in redshift from $z \approx 0.007$ to 0.11 and in epoch

from approximately 12 days before maximum to 55 days after, with an average of 14 spectra for each supernova. An example of a typical supernova spectral time series from our data set is shown in Figure 11.1. The supernovae spectra have been corrected for Milky Way dust extinction (Schlegel, Finkbeiner, and Davis, 1998, Cardelli, Clayton, and Mathis, 1989, where the here correction uses the values from the former) and placed in a common rest-frame at $z = 0$. In order to preserve any characteristics caused by circumstellar and host galaxy extinction, we do not correct for any further reddening due to these dust sources. We do not include in our sample a small number of supernovae that have been previously identified as Ia-CSM or super-Chandrasekhar mass Type Ia supernovae and that would be easily identified and removed from a photometric supernova survey (Aldering et al., 2006, Scalzo et al., 2010, Scalzo et al., 2012). After removing these, we are left with 119 supernovae in our set, which are listed in Childress et al. (2013).

Chapter 4

Method for Measuring Uncertainty due to K-corrections

4.1 Questions addressed and filters used

Different sets of photometric filters, shown in Figure 4.1 and described in Table 4.1, are used to answer a series of questions.

In order to test the best-case scenario where observer-frame and supernova-frame filters can be exactly matched, in Chapter 5 we use a set of top-hat shaped filters with logarithmically spaced edges, which extend to whatever wavelength range is needed to fully cover the supernova spectrum and so are called the "infinite" set (top panel of Figure 4.1 and Table 4.1). The filter spacing is chosen so that the LL4, LL5, and LL6 bands approximately match bands r, i, and z in the MegaCam-R09 filter set, allowing for a closer comparison between them. Because of the logarithmic spacing, these filters move in and out of alignment across redshift such that, at certain redshifts, filters in the observer frame exactly overlap with filters in the supernova frame. A similar technique was used in Schmidt et al. (1998), where filters meant to mimic B and V filters redshifted to $z = 0.35$ and $z = 0.45$ were used to minimize K-correction errors in the range $0.25 < z < 0.55$. The logarithmically-spaced filter set used here presents a more flexible variation on the redshifted B and V filters, since all of the filters in this set can line up with each other, extending the redshift ranges where close matches can be made between supernova-frame and observer-frame filters. Because of this characteristic, along with their on-off edges and full coverage of the wavelength range used by lightcurve fitters, the logarithmically-spaced filters are an idealized set. While narrower filters would provide more information, this would be balanced by a trade-off in increased observing time, meaning overall the filter configuration used here is approximately optimal.

From this we proceed to two degraded situations. To simulate the effect of a finite filter set losing coverage of the redder end of the supernovae's spectral energy distribution, we restrict the analysis to the logarithmically-spaced filters covering 3300 to 9800 Å, referred to as the "finite" set. Next, to simulate the effects of having filters that do not align with

Filter Set	Filter Name	Wavelength Range (Å)	Central Wavelength (Å)	Redshift Line-up[b]
Logarithmically-Spaced (for low-redhift SNe)	LL1	3345 – 4000	3672	—
	LL2	4000 – 4783	4391	$\mathrm{LL1}\cdot(1+0.20)$[c]
	LL3	4783 – 5719	5251	$\mathrm{LL1}\cdot(1+0.43)$
	LL4	5719 – 6839	6279	$\mathrm{LL1}\cdot(1+0.71)$
	LL5	6839 – 8179	7509	$\mathrm{LL1}\cdot(1+1.05)$
	LL6	8179 – 9780	8979	$\mathrm{LL1}\cdot(1+1.45)$
	LL7[a]	9780 – 11696	10738	$\mathrm{LL1}\cdot(1+1.92)$
	LL8[a]	11696 – 13987	12841	$\mathrm{LL1}\cdot(1+2.50)$
MegaCam-R09	u	3370 – 4110	3740	None
	g	4140 – 5590	4870	
	r	5640 – 6850	6250	
	i	6980 – 8430	7700	
	z	8230 – 10190	8900	
Logarithmically-Spaced (for high-redshift SNe)	LH1	10000 – 12600	11300	—
	LH2	12600 – 15873	14236	$\mathrm{LH1}\cdot(1.26) = \mathrm{LH1}\cdot\frac{(1+1.52)}{(1+1.0)}$[d]
	LH3	15873 – 20000	17936	$\mathrm{LH1}\cdot(1.59) = \mathrm{LH1}\cdot\frac{(1+2.17)}{(1+1.0)}$
Euclid-L11	Y	9200 – 11460	10330	None
	J	11460 – 13720	12590	
	H	13720 – 20000	16860	

Notes

[a]Not included in "finite" set of logarithmically-spaced filters.

[b]Determined by the redshifts at which $T_A(\lambda \cdot (1+z_1)) = T_B(\lambda \cdot (1+z_2))$, where T_A is the transmission for filter A as a function of λ.

[c]As an example, we show the redshifts at which the supernova-frame LL1 band lines up with each of the other bands at $z = 0$.

[d]Here we show the supernova-frame redshifts at which the LH1 band lines up with the other bands at the arbitrarily chosen supernova-frame redshift $z = 1$.

Table 4.1: Filter Set Specifications

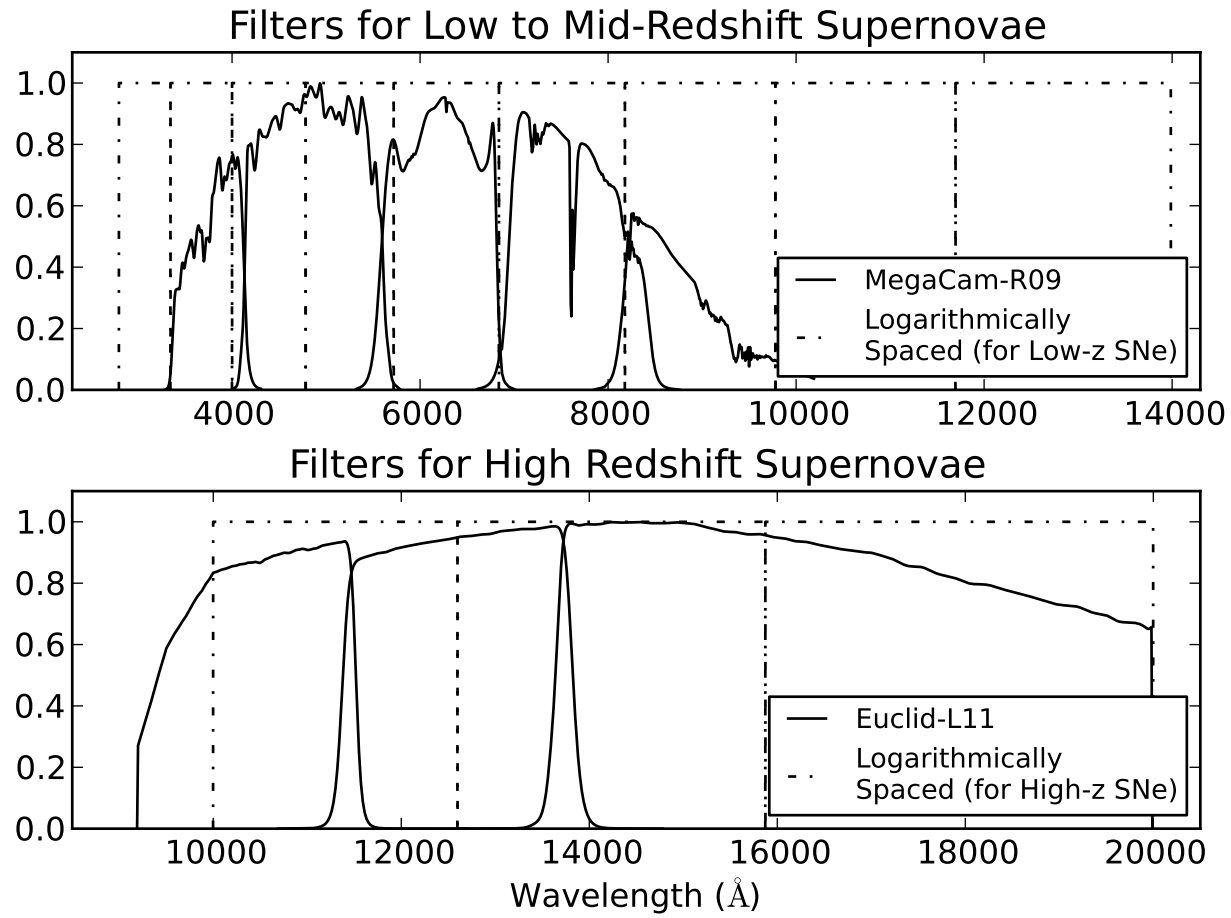

Figure 4.1: Transmission of filters sets used for synthetic photometry. The MegaCam-R09 (Sloan-type) filters and Euclid-L11 filters are shown with the logarithmically-shaped filters covering similar wavelength ranges.

each other at any redshift, we use the MegaCam (Sloan-like) filters (also shown in the top panel of Figure 4.1), which were used in the Supernova Legacy Survey (hereafter referred to as MegaCam-R09, Regnault et al., 2009). The MegaCam-R09 filter transmission curves are calculated at 14 cm from the center of the focal plane and include average atmospheric transmission at Mauna Kea, the CCD quantum efficiencies, the transmission along the optical path, and the mirror reflectivity. In Chapter 7, we address the issue of fitting high-redshift supernovae, where different filters need to be used for the measurements' low-redshift anchor. For higher-redshift supernovae, we use the filters planned for the original design of the Euclid Space Telescope (referred to as Euclid-L11 filters, Laureijs et al., 2011), the results of which are then contrasted with an artificial logarithmically-spaced filter set we have designed to have a similar wavelength coverage (both shown in the bottom panel of Figure 4.1).

4.2 Procedure

In this analysis we use SALT2 (both version 2.2, Guy et al., 2010, and the recently released version 2.4, Betoule et al., 2014) as an exemplar of the capabilities of a state-of-the-art two-parameter lightcurve fitter. Other lightcurve fitters based on similarly simple templates are expected to give comparable results (SiFTO, for example, was shown in Conley et al., 2008 to be very similar to SALT2). Using the spectrophotometric data for each of the SNfactory supernovae, we construct mock supernovae at fixed redshift intervals, using a given set of filter functions to perform synthetic photometry. The supernova spectra are shifted in wavelength, but the flux normalization of the supernovae is not changed because we are interested in the differences in lightcurve fits due to filter coverage, not supernova distance.

We begin by using SALT2 to fit the multi-band lightcurves of a supernova at $z = 0$ and use this fit as a zero point to which other fits of the supernova are compared. SALT2 fits a spectral energy distribution model to the supernova data by using two free parameters that correspond to the color and lightcurve shape of the supernova. By combining the fitted peak B band magnitude with these color and x_1 (stretch-like) parameters, we get the standardized magnitude, M_{std}:

$$M_{std} = M_B + \alpha\, x_1 - \beta\, c \tag{4.1}$$

where α and β are redshift-independent terms that can be fit by minimizing the residuals in the Hubble diagram (in this analysis we use $\alpha = 0.137$ and $\beta = 3.07$, Rubin et al. (2013) to emulate cosomological analyses). We stress that this standardization method, including the α and β parameters, is entirely extrinsic to the SALT2 lightcurve fit; it is only the systematic variations on M_B, x_1 and c that concern us here. We then fit the lightcurve of the redshifted version of the same supernova. As discussed in Chapter 2, in order to fit this lightcurve and get the standardized peak magnitude of the redshifted supernova, a K-correction must be involved in fitting a template to the supernova, which necessarily uses an approximation (in the form of an imperfect spectral model) to fit the measured data. This introduces error into the calculation. Having set the standardized magnitude found for the rest-frame as our zero point, we can find the fit-dependent difference between the two distance estimators of the supernova:

$$\Delta\mu(z_i) \equiv M_{std|z=z_i} - M_{std|z=0} \tag{4.2}$$

This allows us to see the variation around the expected standardized magnitude (or equivalently the fit-dependent part of the variation in the distance modulus) found when using a two-parameter model to fit the same supernova placed at different redshifts. For simplicity, we call this difference the error due to K-correction.

At a given redshift, the above procedure is repeated for the 119 supernovae in our data set, and from the aggregate of the errors we assess the accuracy of the K-corrections. Examples of the range of errors across redshifts may be seen in the top panel of Figure 4.2, which shows histograms of the errors for each of the supernovae at redshift intervals. As a baseline we use robust statistics, meaning that the bias and dispersion are the median and normalized median

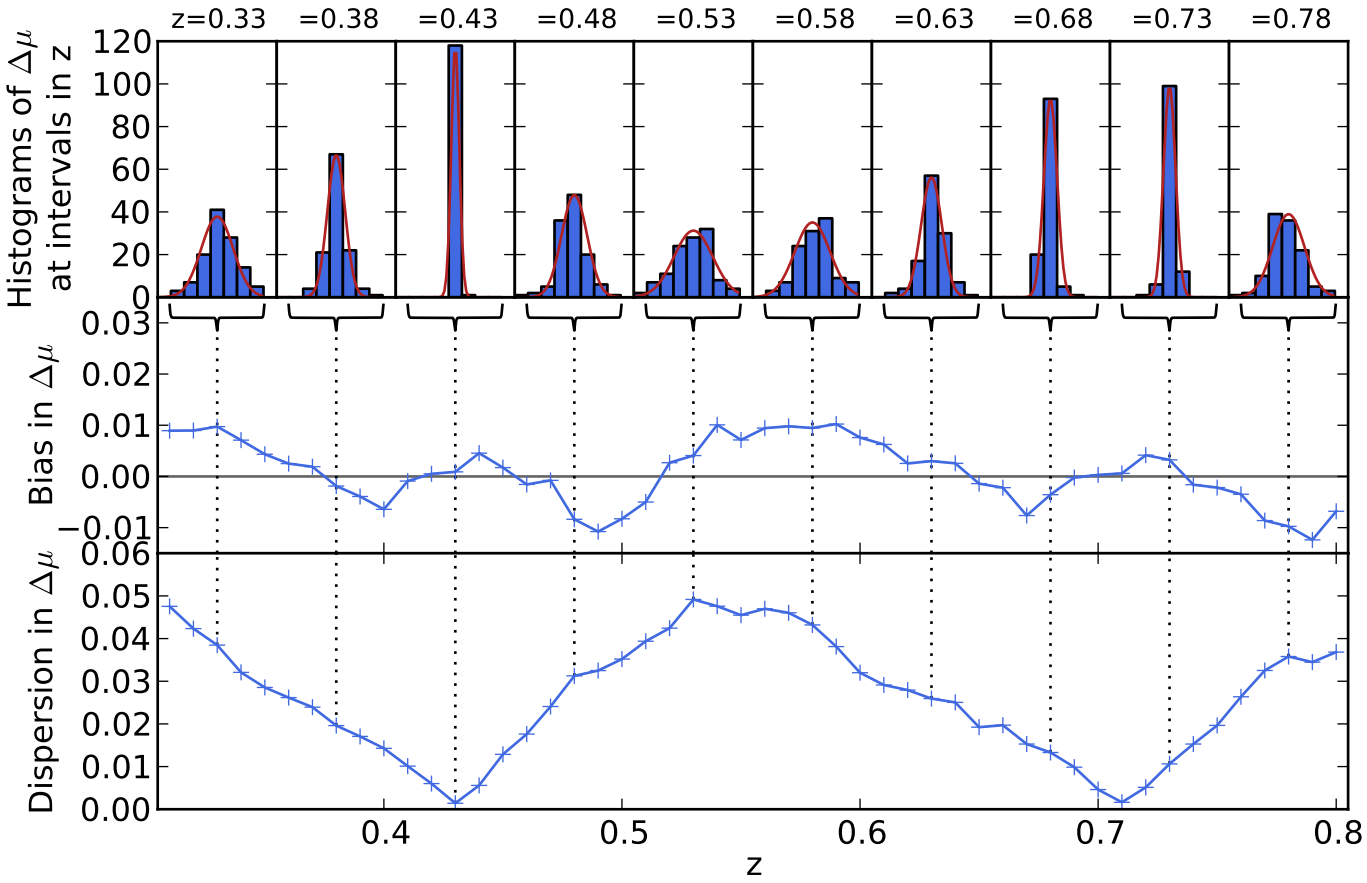

Figure 4.2: An example of the median bias and dispersion in the standardized magnitude variation due to K-correction errors found using SALT2.2 (with logarithmically-spaced filters) over a portion of the redshift range are shown in the middle and bottom panels. The error in the bias is approximately $\sqrt{\pi/(2N)} \approx 0.11$ times the dispersion, where $N = 119$. The top panel shows histograms of the data points at intervals of 0.05 in redshift, where each histogram has 9 bins spanning standardized magnitude errors from -0.15 to 0.15. A Gaussian curve fit to the histogram is overplotted in each panel. Low standard deviation in the histogram can be seen to correspond to low dispersion in the bottom panel, while a large standard deviation corresponds to high dispersion.

absolute deviation (nMAD) of the errors (i.e. the median and nMAD of the $\Delta\mu$ values for all the supernovae at a given redshift), ensuring results are not sensitive to outliers. However, the mean and standard deviation of a Gaussian function fit to the histogram of the errors at a given redshift give equivalent results, as may be seen by comparing the Gaussian curves in the top panel to the amount of dispersion in the bottom panel of Figure 4.2. A bootstrap test confirms that the error in the bias is approximately equal to $\sqrt{\pi/(2N)}\times$ the dispersion, as expected for bias calculated from the median, where $N = 119$, the number of supernovae used. This further shows that outliers are not driving the dispersion seen in the results.

Some supernovae are fit consistently by the template, such that their standardized magnitudes are reconstructed well; this means that even at redshifts where there is a large standard deviation, these supernovae are still in the peak of the histogram. Less well fit supernovae have standardized magnitudes which vary more across redshifts and make up the wings of the histograms at redshifts where there is a high standard deviation. The errors assembled across redshifts allow us to gauge general trends in the standardized magnitudes, in particular the amount of dispersion at different redshifts, or consistent positive or negative biases.

As an example of the possible variation of SALT2 lightcurve fits, in Figure 4.3 we show the rest-frame spectra nearest maximum of three supernovae, along with the SALT2 model spectra constructed from the parameters fit to each supernova at a range of redshifts. It can be seen that for the supernova in the top panel, the best-fit SALT2 spectral reconstructions vary noticeably across redshift, while for the supernova in the bottom panel, the SALT2 reconstructions are more difficult to distinguish from one another. This is most likely driven by how well the SALT2 spectral model can fit the given supernova spectra.

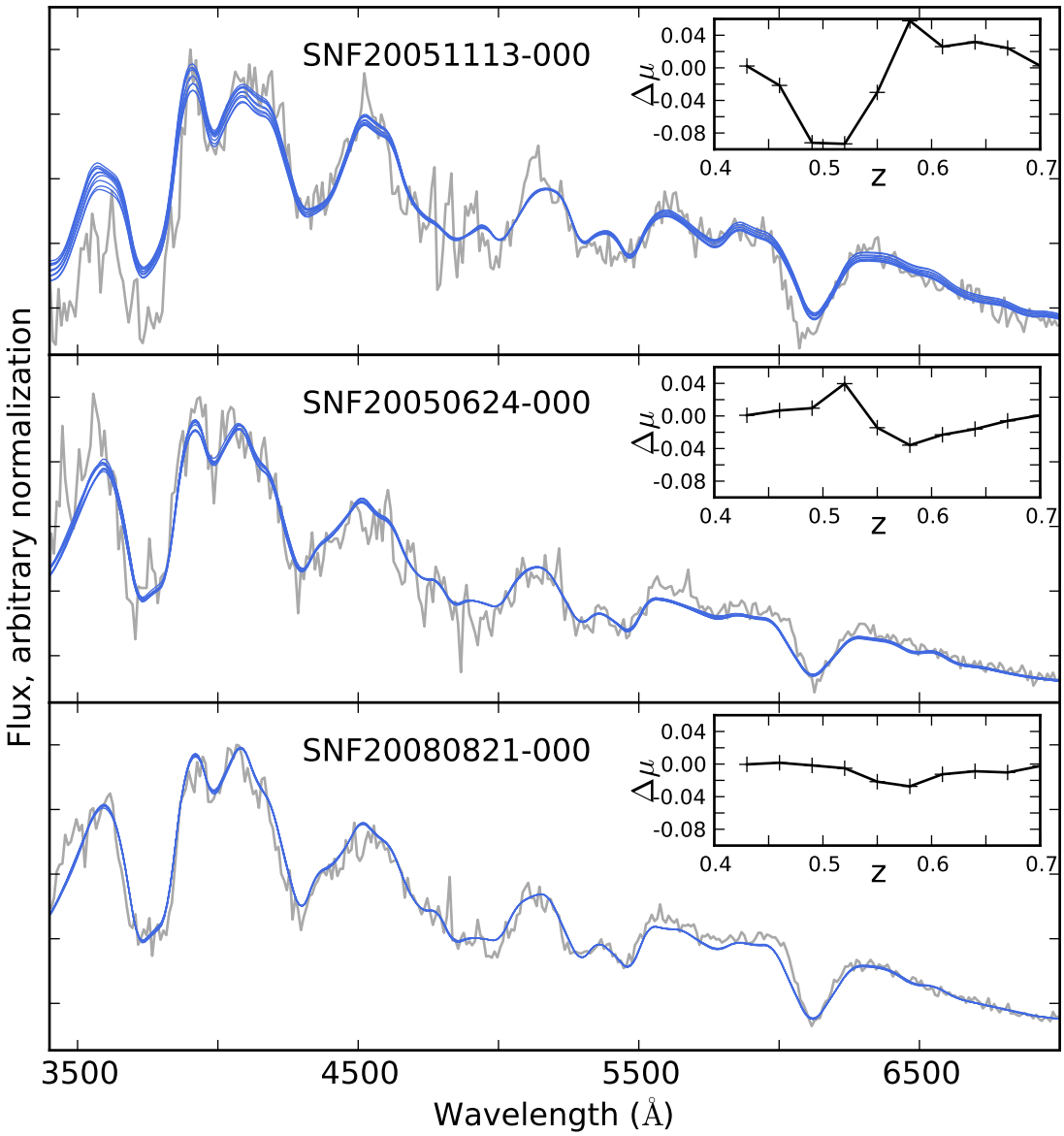

Figure 4.3: Examples of spectral reconstructions of different qualities of SALT2.2 lightcurve fits calculated when using logarithmically-spaced filters. Spectra near maximum for three example supernovae are shown in gray. Overplotted are the spectra constructed from the SALT2 parameters fit to the supernova at a range of redshifts. The inset plots show the values for the variation in the standardized magnitudes at the corresponding redshifts. The top panel shows a supernova for which the fits vary widely, the middle panel shows a supernova for which there is moderate variation, and the bottom panel shows a supernova which SALT2 fits very similarly at any redshift.

Chapter 5

Sources of K-correction Errors

We consider first the results of our analysis when using version 2.2 of SALT2: For all supernovae, when an object is measured with logarithmically-spaced filters, shown as dotted lines in Figure 4.1, which fully cover the wavelength range used by SALT2, there are some redshifts where the filters in the frame of the supernova and the filters in the frame of the observer line up exactly. When a supernova is measured at one of these redshifts, with the same number of filters as were used for the rest-frame measurement, the K-correction is done in exactly the same way for the $z = 0$ and the redshifted supernova, which means that the calculated standardized magnitudes are the same.

When we move away from this ideal situation, errors start to appear. The most easily identifiable source of error is when filters are out of alignment. As can be seen in Figures 4.2 and 5.1, the dispersion in the errors increases and the bias oscillates as the filters fall out of alignment, and then both go to zero as the filters become more aligned. When the filters are maximally misaligned, the dispersion peaks at about 0.04 to 0.05 magnitudes, while the bias varies between -0.01 and 0.01 mag. In both the bias and the dispersion we see a periodically repeating pattern, the size of which does not increase as we go up to higher redshifts. It should be noted that in a cosmological analysis, a periodic bias such as this can in principle be fit out, whether or not the size of the bias is known, though in practice this has never been done.

For the subset of supernovae that are well fit by the SALT2 templates, the SALT2 fit is almost the same at every redshift (the supernova in the bottom panel of Figure 4.3 is an example of such a supernova). This source of error is thus the result of supernova spectral diversity not captured by the lightcurve fitter parametrization of the underlying spectrum. The dispersion is set by the degree of diversity, while the bias reflects a shift between the demographics of the SNfactory data set and that of the supernova set used to train the SALT2.2 spectral templates.

A further source of error occurs if we no longer have enough filters to always cover the wavelength range of the supernova at a given redshift. This means that as we get to higher redshifts, we run out of filters and measure a supernova with fewer filters than we had for the rest-frame measurement. To test this situation, the filters used are limited to those

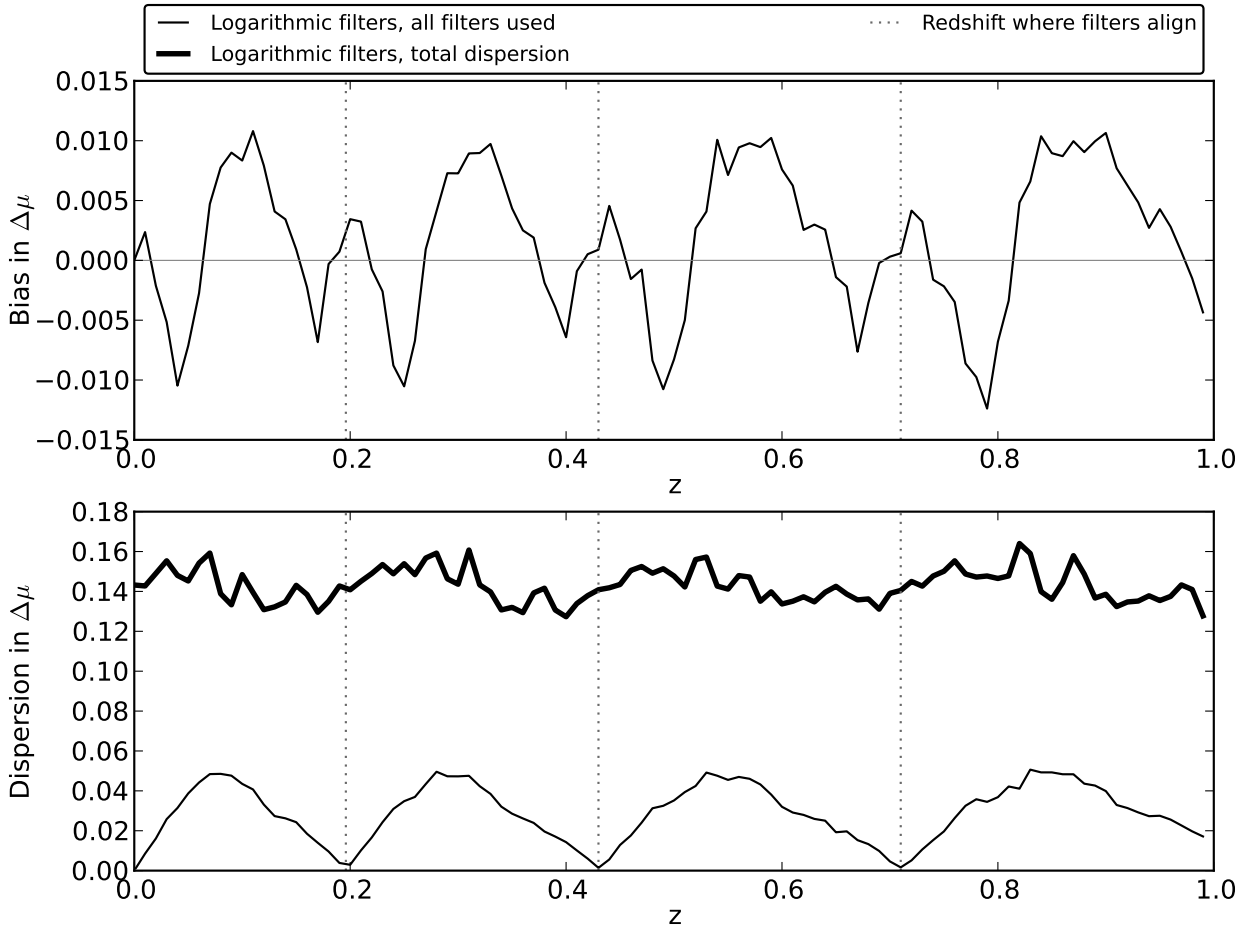

Figure 5.1: Bias and dispersion in the standardized magnitude variation due to K-correction errors when an "infinite" set of logarithmic filters are used with SALT2.2. Note that these are the same results as in Figure 4.2, shown over a greater redshift range. As a comparison, the lower panel shows, the total dispersion in the standardized magnitudes, i.e. the dispersion in μ, not $\Delta\mu$, which is what would actually be observed (this is equivalent to the dispersion in the distance moduli on a Hubble diagram). The contribution of the K-correction errors to the total dispersion is discussed further in Chapter 6.

covering observer-frame 3345 to 9780 Å, the "finite" set. We then calculate the standardized magnitudes that would be found with the limited filters at a given redshift, compare them to the rest-frame magnitudes, and find the bias and dispersion in the aggregate of the errors (Figure 5.2). Filter coverage of the redder wavelengths is lost starting at $z = 0.53$ since the LL7 filter which would otherwise cover this range is not included in the "finite" set. As the lack of overlapping wavelength coverage becomes more severe around $z = 0.7$, a large amount of dispersion appears, increasing to more than 0.15 magnitudes, and the bias oscillates over a much larger range. The results suggest that some loss of filter coverage can be tolerated, but that it will become a major source of systematic errors when the filter coverage is limited and covers bluer supernova-frame wavelengths. The possible causes of this effect are discussed below.

A starker difference in the standardized magnitudes is seen if filters are used which are not logarithmic in their shape and relative spacing, as is shown in Figure 5.2. As a comparison to the logarithmically-spaced filter set, we look at the filters used in the MegaCam wide-field imager, a set of Sloan-like filters which are shown in Figure 4.1 and have wavelength coverage similar to that of the "finite" set of logarithmic filters. Here there is no longer any redshift where the supernova-frame filters line up with the observer-frame filters and thus there are no redshifts at which a K-correction can be done perfectly. Instead, there is a consistent amount of dispersion at around 0.05 magnitudes, increasing to about 0.20 mag as wavelength coverage is lost. We find a small bias in the data, varying across redshift within about 0.02 mag between $z = 0$ and 0.7, but by an increasing amount as redder wavelength coverage is lost at higher redshifts.

We next consider the results of our analysis when using the recently released version 2.4 of SALT2 on synthetic photometry from the same three filter sets. SALT2.4 uses the same model as SALT2.2, but the templates and color law have been retrained on a larger data set, which changes both the individual fits and the ensemble average, and the error bands have been improved. As can be seen in Figure 5.3, at redshifts below $z = 0.6$, the results are very similar to those found with version 2.2. However, at higher redshifts, the bias is much lower than was found when using 2.2. The dispersion, on the other hand, is very similar to that found using SALT2.2. Reasons for this are discussed below.

5.1 Isolating the Cause of the Results at Higher Redshifts

The high bias and dispersion seen above $z \approx 0.7$ in Figure 5.2 raise questions about lightcurve fitting when coverage of the red side of supernova's spectra is lost at higher redshifts. Here we examine possible causes for this breakdown in K-correction accuracy, such as an overall offset between the calibration of SALT2 and the SNfactory data, which would lead to biased results, or dispersion in the calibration, which would lead to dispersion in the results.

First, there is uncertainty in the calibration of the data used to train SALT2. We perform

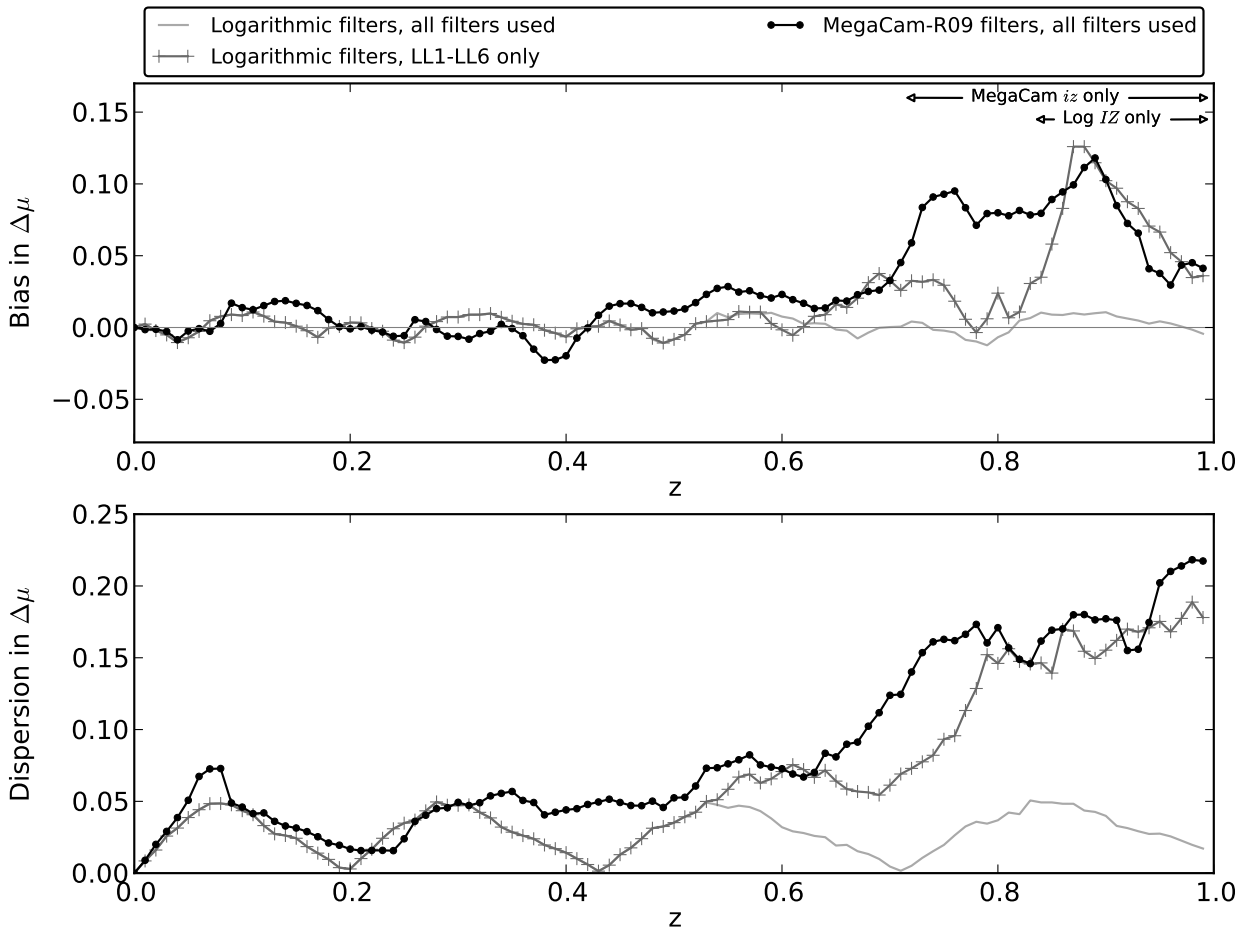

Figure 5.2: Bias and dispersion in standardized magnitude variation found using the MegaCam-R09 filters with SALT2.2. Results from previous iterations using an "infinite" and a finite number of logarithmically-shaped filters are shown in light and dark gray. The redshift ranges in which only two MegaCam-R09 or Logarithmic filters are available with the SNfactory data are indicated in the top panel.

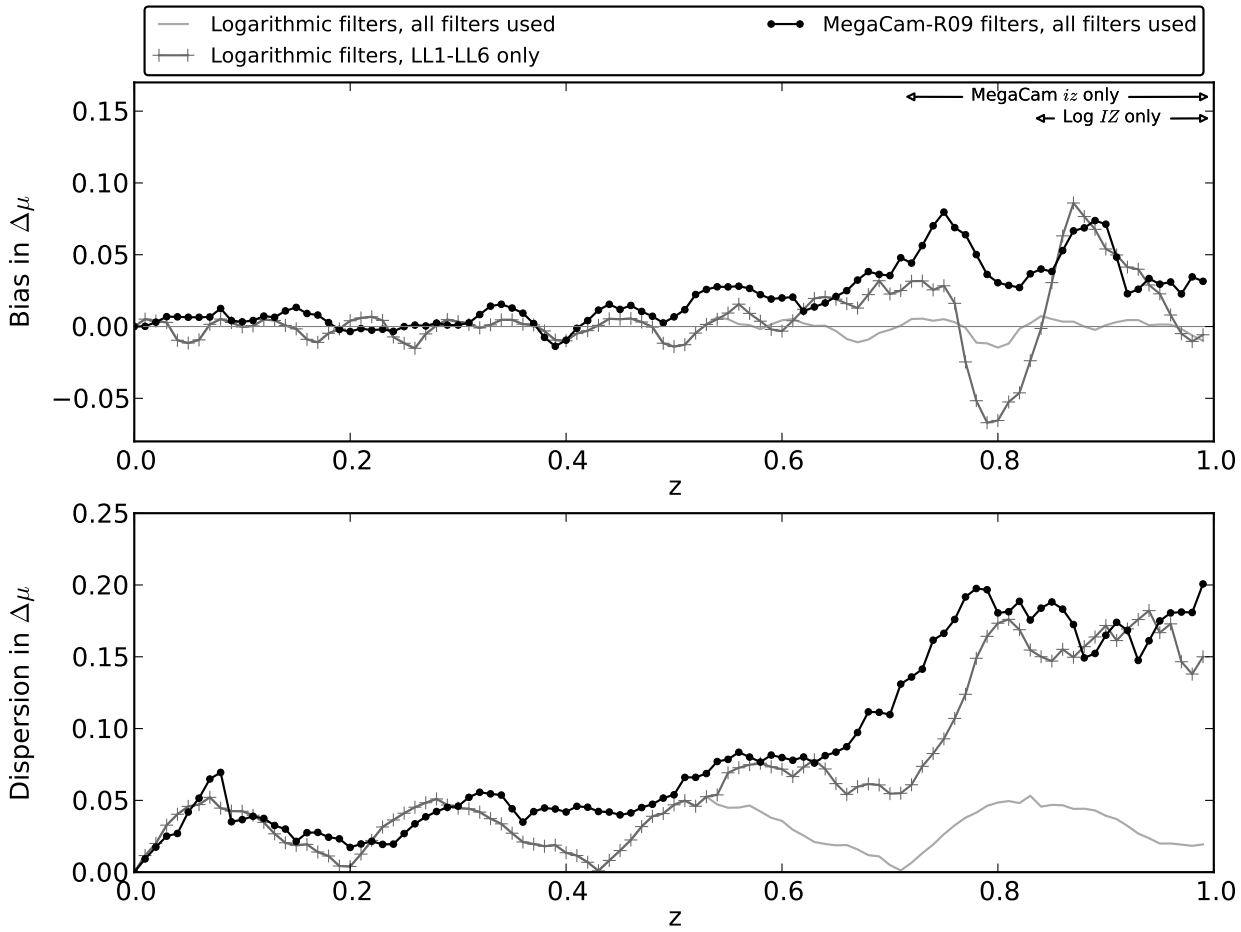

Figure 5.3: Same as Figure 5.2 but using SALT2.4 for the lightcurve fits.

a Monte Carlo simulation to test the impact of SALT2 input calibration errors following the error distribution shown in Appendix A of Guy et al. (2010). These errors are used to produce variations in the SALT2 spectral template, which are then propagated through the K-corrections analysis by fitting the template variations with SALT2. This gives an error on the bias equal to about 0.01 at low redshifts, increasing to about 0.04 at $z = 1$. This uncertainty is shown for both SALT2.2 and 2.4 in gray in Figure 5.4. A similar effect is shown in Mosher et al. (2014), where varying the SALT2 training data and scatter model is shown to result in biased Hubble diagrams.

Next, we evaluate whether undiagnosed calibration errors might play a role in the effects seen. Dispersion in the calibration of the U band (averaging to zero) would be a plausible cause of the dispersion in the standardized magnitudes, since the lightcurve fit becomes more sensitive to the U band photometry when redder bands are lost. To test whether a calibration error in the SNfactory data could cause this effect, a Monte Carlo simulation was performed in which a random, Gaussian distributed calibration error with $\sigma = 0.06$ mags was

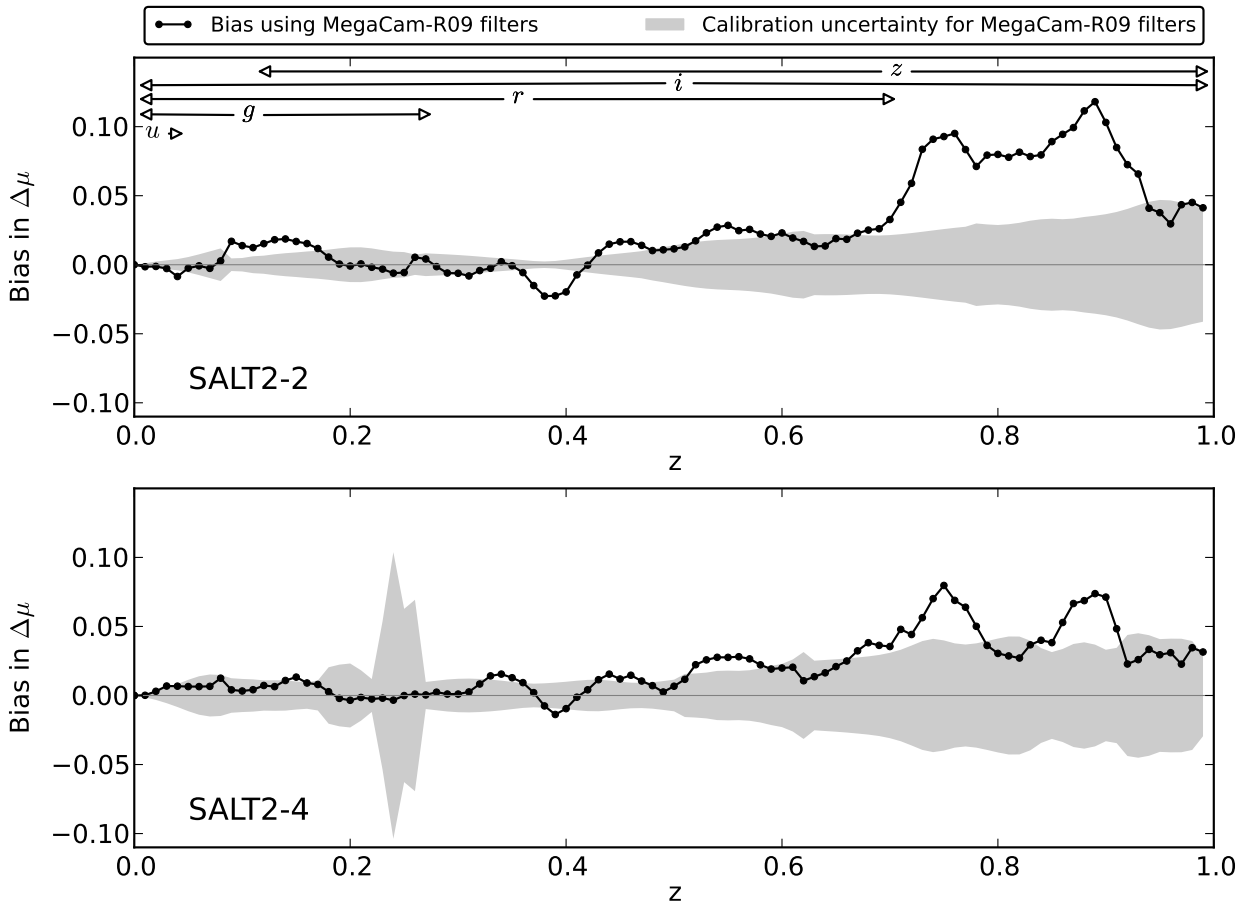

Figure 5.4: Calibration uncertainty is shown as an error band on the bias zeropoint, overlaid on the bias calculated using SNfactory spectra. The top panel shows the bias and calibration uncertainty calculated using SALT2.2 and the MegaCam-R09 filters, while the bottom panel shows the same results calculated using SALT2.4. The spike in the uncertainty for SALT2.4 at around $z = 0.25$ reflects the fact that at this redshift the g band covers the far UV in the supernova-frame where there is very large calibration uncertainty in the training data. The redshift range where each of the MegaCam-R09 filters can be used is shown in the top panel; the redshift ranges for the results in the bottom panel are the same.

added to the U band of fake supernovae made using the SALT2 template. These supernovae were then placed at a range of redshifts and the above analysis was repeated to synthesize photometry and fit the standardized magnitude at each redshift. While SALT2 does not find exactly the same standardized magnitude at every redshift for each of these simulated supernovae, the variation is negligible in comparison to the results seen in Figure 5.2, with dispersion at about 0.005 mags. At higher redshifts, as redder filters are lost and the U band becomes more important when the MegaCam or finite logarithmic filters are used, the dispersion does increase to about 0.04 mags at $z = 1.0$, but this is still only a small fraction of the dispersion that we see from variation due to K-corrections (about 0.2 mags).

An additional calibration test was done in which a grey calibration error was added to the SALT2 template. In this scenario, a random, Gaussian-distributed wavelength-independent error with $\sigma = 0.05$ mag was added to the spectrum. As in the previous test, SALT2 did not find exactly the same standardized magnitude at every redshift for the simulated supernovae. However, the bias and dispersion in these standardized magnitudes is again small compared to the bias and dispersion that we see in the fits to the real supernova data. Even with this conservative estimate of a possible grey calibration error, the dispersion is at most about 20% of what is seen with the real data, and the bias is negligible until filters are lost at high redshifts, where the bias increases to about 0.004, again small compared to the bias in the real data (compare to bias in top panel of Figure 5.2).

This suggests that more deep-seated differences between the U bands of the SNfactory data and the SALT2 template must be the source of the large biases seen. The decrease in the biases seen between SALT2.2 and SALT2.4 in particular shows that the additional training material in the new version has brought the blue end of the SALT2 model closer to that of the average SNfactory supernova spectrum. Even with the new model the dispersion at higher redshifts has stayed high however. This implies that the UV spectra of Type Ia supernovae are inherently more diverse than can be accounted for by a two-parameter lightcurve fitter, or that there are UV subpopulations that are not represented by the spectral templates (Figure 4.3 shows examples in which the SALT2 reconstructions do not match the spectra in the UV). Hints of the importance of subpopulations in the UV spectra have been seen already in the literature, but this will require further study (Jha, Riess, and Kirshner, 2007, Ellis et al., 2008, Nordin et al., in preparation).

The U band and grey calibration issues have been ruled out as causes of the dispersion seen at higher redshifts. Systematic errors in the standard star calibration of both the SALT2 training data and the SNfactory data cause uncertainty which may explain some of the bias that is seen. The remaining bias and the dispersion must then be due to either spectral feature differences or broad band differences in the SNfactory data set and the SALT2 training set. Visual inspection suggests that both of these are in play, but further work will be required to decide quantitatively which has the greater effect.

Chapter 6

Contribution of K-corrections to the Total Dispersion in Magnitude Measurements

The effects of K-corrections have not been seen in the data from past photometric supernova surveys because their effects are lost in what we actually observe, which is the total dispersion in supernova standardized magnitudes, or equivalently the dispersion in Hubble diagram residuals. As a comparison, the dispersion due to K-correction errors is shown with the total dispersion in the supernova standardized magnitudes in Figure 5.1. For both logarithmically and non-logarithmically-spaced filters, the dispersion in the standardized magnitudes (as opposed to the dispersion only in the K-correction error) is close to a constant value of about 0.15 magnitudes across redshifts, until filter coverage is lost at higher redshifts. (This is comparable to the RMS of the Hubble residuals, 0.161 mag, found in SALT2, Guy et al., 2007). Thus, while K-corrections have the effect of moving individual magnitudes away from their standardized magnitude as measured in the rest-frame, these changes are subdominant to the overall spread in the data.

To test whether the results using synthetic photometry of SNfactory data agree with recent supernova searches, the total dispersion in standardized magnitudes of the redshifted SNfactory data set used here is compared to the results of the Sloan Digital Sky Survey (SDSS-II, Holtzman et al., 2008, Frieman et al., 2008, Rubin et al, in preparation), shown in Figure 6.1. The SDSS-II supernova set consists of 129 photometrically measured supernovae spread across redshifts 0.06 to 0.42. To compare the SDSS-II results with our own, the Hubble residuals of the standardized magnitudes fit by SALT2.2 (calculated with $\Omega_m = 0.27$) were binned and then the bias and dispersion of the Hubble residuals in each bin were calculated. The variation of the bias in the SDSS-II data is on a comparable scale with the bias variation predicted from the synthetic SNfactory distance moduli. While the dispersion in this data set is not as constant, the uncertainties on the value in each bin are consistent with a constant dispersion.

Again, the amount of dispersion in the data means that the effect of K-corrections by

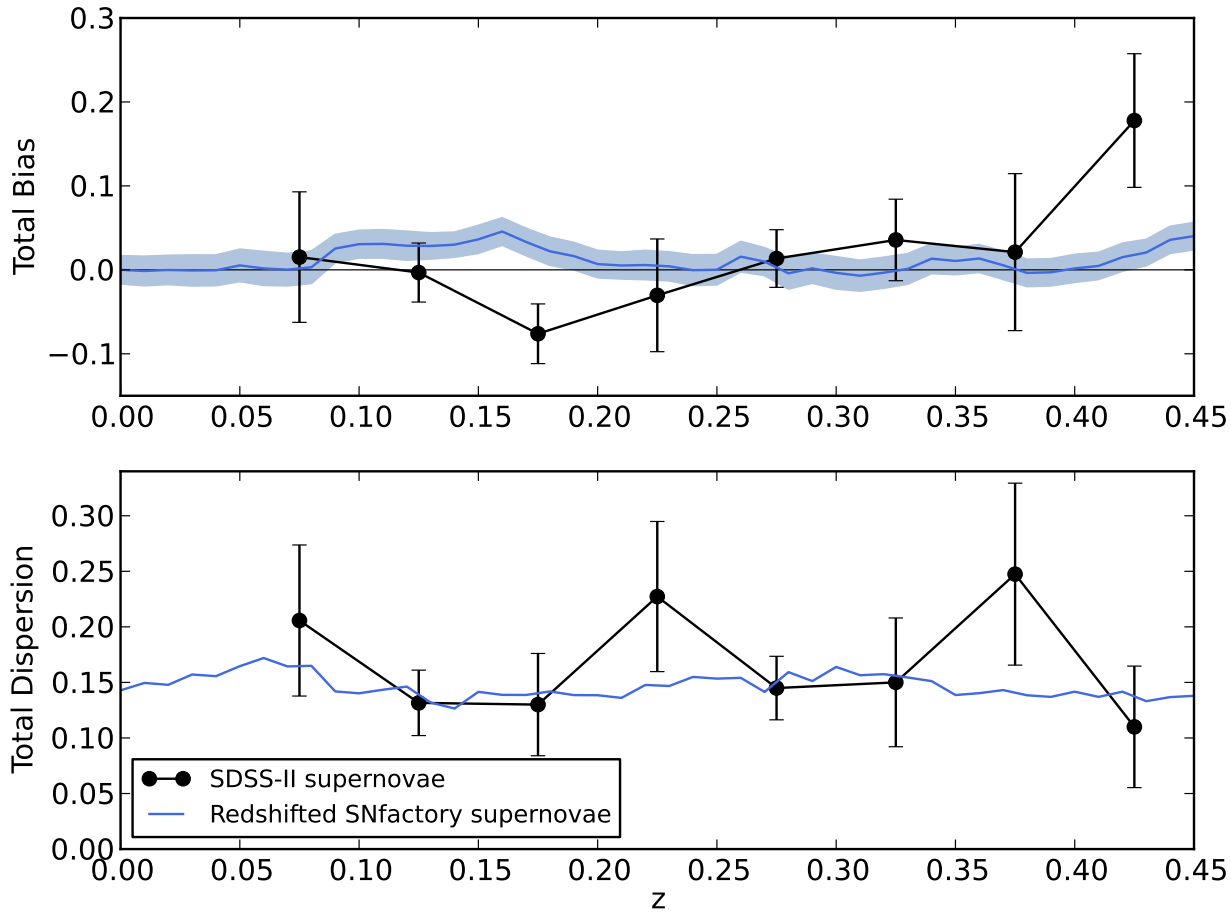

Figure 6.1: Bias and dispersion in the Hubble residuals of supernovae from SDSS-II are shown by black points with error bars, binned at intervals of 0.05 in redshift. The bias and dispersion in the standardized magnitudes of our mock supernova observations (where the median of the standardized magnitudes at $z = 0$ is used as the zero point for the bias) are shown as a comparison by the blue line and error band. Both data sets are fit using SALT2.2.

themselves cannot be seen in the total dispersion of the data set, where they are subdominant to measurement uncertainty and intrinsic dispersion. This does not, however, mean that it is not important to isolate the K-correction errors: future advances may reduce sources of dispersion that are currently dominating what is observed. Additionally, future supernova searches will use hundreds or thousands of supernovae to narrow the statistical dispersion of binned distance moduli (Bernstein et al., 2012, LSST Science Collaboration et al., 2009). Thus, while K-corrections are currently subdominant, due to the fact that they are a systematic source of bias they will become a limiting factor in precision if future searches do not address these issues, as discussed below.

Chapter 7

Implications for Supernova Cosmology

An important application of this study will be for high redshift supernova analyses, requiring that our previous results be extended to higher redshifts. Doing so for an internally-consistent set of logarithmically-spaced filters would simply extend to higher redshifts the oscillatory behavior seen in Figure 5.1. A more interesting and relevant case is to address the effect of the jump from ground-based to space-based experiments when going to higher redshift, and the attendant filter choices. In this case, high-redshift supernovae observations will need to be compared to low-redshift supernovae that have been observed under different conditions. To calculate the K-correction errors in this situation, for the $z = 0$ zero-point measurements we will use the MegaCam-R09 filters, which are a representative of ground-based experiments like SNLS, DES and LSST. For the high-redshift measurements, we use two alternative filter sets that would be realistic for a space-based mission: first, a set of three filters once planned for the Euclid mission, and second, three logarithmically-spaced filters that Euclid or other space missions could use instead (bottom panel of Figure 4.1).

Such externally-imposed changes in filter spacing mean that the ground-based and space-based rest-frame wavelength coverages never line up exactly. This means that the K-corrections cannot be done perfectly at any redshift, but we do still see that as filters move closer and farther from approximate alignment the dispersion rises and falls (Figure 7.1). Though neither filter set appears particularly favored in the comparison with a MegaCam-R09 baseline, the logarithmically-spaced filter set has the advantage that internally measurements can be compared exactly: a supernova will have the same supernova-frame standardized magnitude if it is, for example, observed with the two bluer bands at $z = 1.2$ or the two redder bands at $z = 1.77$, because the LH1 and LH2 bands cover the same supernova-frame wavelength range at $z = 1.2$ as the LH2 and LH3 bands at $z = 1.77$. Also, as with the lower-redshift results using logarithmically-spaced filters, a periodic bias found in the high-redshift results using logarithmically-spaced filters can in principle be fit out in a cosmological analysis.

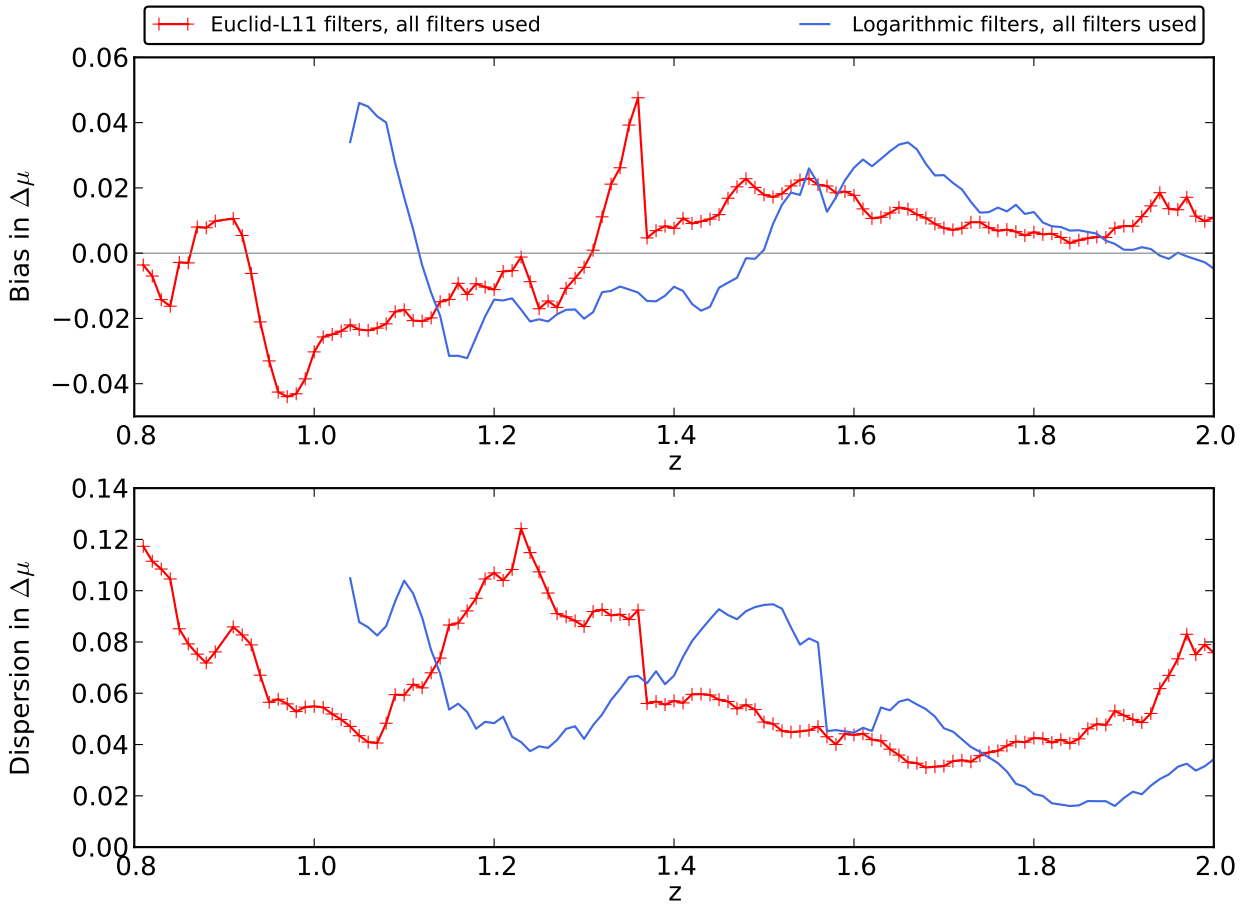

Figure 7.1: Bias and dispersion of the standardized magnitude variation found using SALT2.2 with the Euclid-L11 filters and a set of logarithmically-shaped filters covering a corresponding wavelength range. Since these filters cannot measure low-redshift supernovae, the results from the MegaCam-R09 filters are used as the zero redshift fiducial points. The Euclid-L11 filters approximately align with the MegaCam-R09 filters in the regions around $z = 1.05$ and 1.73 and the logarithmically-spaced filters approximately align with the MegaCam-R09 filters around $z = 1.27$ and 1.9. The importance of wavelength coverage can be seen in the Euclid-L11 results, where the inclusion of the H band at redshifts above 1.37 causes a large drop in the bias and dispersion.

7.1 Impact on w and the Effect of a Possible Population Shift

For applications at higher redshifts, a trend in the biases of the K-correction errors that is monotonic with redshift, such as was seen at lower redshifts in the results in using the MegaCam-R09 filters (Chapter 5), will be a concern when supernova searches are applied to questions in cosmology, in particular the calculation of the time variation of the dark energy equation of state ratio w. Subtracting out the bias found using the MegaCam-R09 filters from data in the range $0 < z < 0.8$ causes w to shift by about -0.03 (using as a representative sample supernovae from the Union Supernova Ia Compilation, Rubin et al., in preparation). Since the bias has been calculated using the supernovae that we artificially redshifted, we could simply correct for it in future searches, but that assumes that the supernova population at higher redshifts is the same as the low redshift population.

To measure the effects of a change in population, we repeated the calculation of the bias and dispersion on supernova sets where the population shifts towards a certain subset of supernovae as the redshift increases. As a simple test, the supernovae were divided into two sets, one containing supernovae with average standardized magnitude variation (using the magnitudes calculated at redshifts between $z = 0.3$ and 0.7) above the median, called the high-bias set, and one with average standardized magnitude variation below the median, called the low-bias set. We then simulated supernova sets, which at $z = 0$ would have the original population of 50% high-bias supernovae and 50% low-bias supernovae, and would then shift quadratically with redshift so that at $z = 2.0$ they would have 95% high-bias supernovae and 5% low-bias supernovae. The bias and dispersion of the supernova sets were calculated with the MegaCam-R09 filters up to $z = 0.8$ and with the Euclid-L11 filters from there to $z = 2.0$. This was then repeated with supernova sets that shift toward having more low bias supernovae at high redshift.

The results, shown in Figure 7.2, indicate that even at redshifts like $z = 0.6$ where the population is only shifted from 50% to about 70% towards one population, the difference between the biases is near 0.05 mag. At higher redshifts, the effect is more consistently large, though particularly large at the redshifts around $z = 1.25$ where the filter configuration gives more weight to regions of the spectrum where the two populations differ more. The difference between the bias calculated with the original population and the bias with a population shift would correspond to a change in w of ~ 0.03 when calculated with data up to $z = 0.8$ (again using supernovae from the Union Supernova Ia Compilation). This indicates that without some knowledge of the supernova population, attempts to correct for a bias in magnitudes due to K-correction errors will introduce a significant amount of uncertainty. It is important to note that these biases, too, always become small when there are redshifts at which the low- and high-redshift observer's filters align (e.g. around $z = 1$ in Figure 7.2). While an accurate measurement could be made at these points, this would only be possible for a limited range of redshifts for each filter set, an expensive proposition as a way to study the variation of w with time.

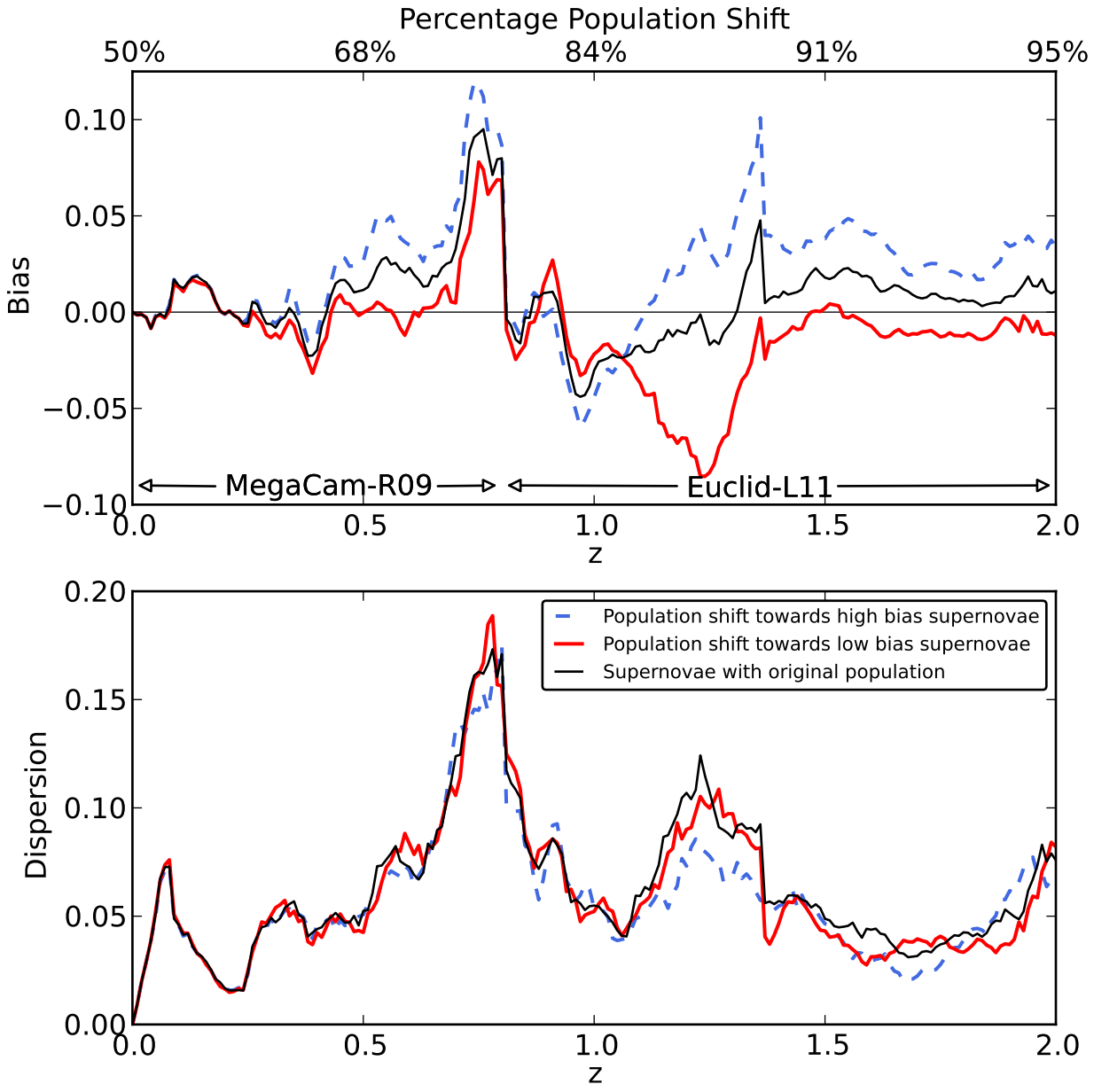

Figure 7.2: Bias and dispersion in the standardized magnitude variation due to K-correction errors in supernova populations shifted towards high-bias supernovae (in blue) or low-bias supernovae (in red) when using SALT2.2. The results from the MegaCam-R09 filters are used from $z = 0$ to $z = 0.8$ and the results from Euclid-L11 are used from there to $z = 2.0$. The population consists of 50% high-bias and 50% low-bias supernovae at $z = 0$ and shifts quadratically such that in the region of $z = 0.6$ to 0.8 about 70% of the supernovae are from one population, and at $z = 2.0$, 95% of the supernovae are from one and 5% are from the other. The results with the original population are shown in black.

Chapter 8

Discussion

We have isolated the component of the standardized magnitude variation due to K-corrections across a range of redshifts. These inconsistencies in the magnitudes measured are a direct consequence of the imperfect way in which the lightcurve fitting template is able to match the underlying spectrum of each supernova at each phase. The best fit of an individual supernova varies depending on the filter configuration with which it was observed: examination of the actual spectral time series of a supernova juxtaposed with the spectra constructed from the SALT2 fit of the supernova shows that there are differences between the data and the template fit to it at every redshift. When a fixed filter system is used over a range of redshifts, these differences move across the photometric filters, causing changes in the best fit of the parameters to the lightcurve. As a result we see a variation in the standardized magnitudes fit to the supernova. Biases appear in the average fitted standardized magnitudes when there is an offset between the population of the observed supernovae and the supernovae in the lightcurve fitter's training set. In the current generation of lightcurve fitters, this may be due to the fact that the models are trained on photometry from limited redshift ranges with no spectra or with spectra from a relatively small number of supernovae. The dispersion that is seen in the standardized magnitudes, however, is an inherent effect of the fact that the flexibility of a model built with two linear components and a color law is unable to capture the diversity of Type Ia supernovae. This is reinforced by the finding in Kim et al. (2013) that a model incorporating more parameters produced a significant reduction in the dispersion.

We see then that even when using a fairly sophisticated two-parameter lightcurve fitter systematic errors remain. These produce significant, though currently subdominant, amounts of dispersion in supernovae magnitudes, along with bias if the fitter training set does not match the distribution of the fitted supernovae. A more precise model cannot be trained with data insufficient to capture the full range of Type Ia supernova behaviors. However, the existence of the SNfactory spectrophotometric time series suggests possible methods for improving results. As described in Chapter 7, the SNfactory data could be used to predict bias and subtract it out of high-redshift results, but this is limited by our knowledge of the high-redshift population. A more sophisticated approach would be using

the SNfactory data as the low-redshift anchor for the Hubble diagram and performing direct matching between the SNfactory supernovae and high redshift supernovae, eliminating the step of using a template-based lightcurve fitter (Fakhouri et al., 2015). Another possibility would be to use SNfactory data to train a more complex spectral model with the flexibility to capture more of Type Ia supernova diversity, which is performed in Part II of this thesis, though lightcurves alone may not have sufficient data on which to fit such a model.

This then highlights the primary remaining issues: that we cannot currently recognize which model or observed low-redshift spectrum corresponds to a given high-redshift supernova from photometric light curves alone, and that we cannot assume that the overall population stays the same across redshifts. Future work must determine whether spectroscopy of high-redshift supernovae is necessary or if there are unique photometric indicators beyond the currently employed lightcurve width and color that can be used to identify a given supernova among the diverse range of possible Type Ia supernovae.

Conclusion

We have calculated the variation in distance moduli due to K-corrections for filter sets of various configurations and with wavelength coverages selected for low to mid-range and high redshift supernovae. We find that the dispersion and bias in the distance moduli are the most predictable and the most constrained when we use logarithmically-shaped filters that are sufficiently numerous to ensure that the full wavelength range of the supernovae is covered. However, this variation in the distance moduli is a symptom of using a template that lacks the flexibility to span the diversity of Type Ia supernovae, meaning it will occur at most redshifts regardless of the choice of filters. While we have used the two-parameter lightcurve fitter SALT2 in this analysis, no other current fitters provide different methods that would resolve the issues seen here. Thus, the results show that if we use photometry and simple template-based lightcurve fitters, we must expect that errors in our magnitude measurements will be propagated into cosmology measurements.

A supernova spectral-time-series template trained on low-redshift spectrophotometry from the SNfactory should be capable of making an improvement in the bias of distance moduli due to K-corrections, but would be handicapped by lack of information concerning shifts in the supernova population. In this case then, we would need to spot check our high-redshift photometry for changes in supernova population or environment, which would be expensive. Another possibility for improving results would be redundant, overlapping filters, but this is also expensive. Nevertheless, in order to accurately measure high-redshift supernovae, future searches must have a way to identify and account for population drift, and this may in the end require partial or complete spectrophotometry.

Part II

Building a Better Empirical Model for Type Ia Supernovae

Chapter 9

Introduction

As demonstrated in Part I, insufficiently detailed models for Type Ia supernovae are a limiting factor in improving the precision of magnitude standardization and the use of Type Ia supernovae for cosmology. While current and upcoming supernova surveys such as DES and LSST (Bernstein et al. 2012, LSST Science Collaboration et al. 2009) will eliminate any significant source of statistical uncertainty, calibration errors and intrinsic dispersion still lead to a large amount of uncertainty and potential bias in standardized magnitudes.

On the question of standardized magnitudes, progress should be possible. As discussed in the introduction, current techniques for standardization still rely on simple supernova models that have only one or two degrees of freedom (SALT2, MLCS2k2, SNooPy, Guy et al. 2007, Jha, Riess, and Kirshner 2007, Burns et al. 2011). After standardization, there is a remaining dispersion in supernova magnitudes of about 0.14 (Betoule et al. 2014), which is a combination of scatter due to calibration errors and intrinsic and extrinsic sources of remaining dispersion in supernova magnitudes. Recent work has shown that Type Ia Supernovae have more diversity than can be described by two components (Saunders et al. 2015, which is Part I of this work, Kim et al. 2013, Sasdelli et al. 2015, etc.), suggesting that the supernova-dependent sources of dispersion in the corrected magnitudes can be reduced by using an improved model for Type Ia supernovae.

The complexity of the existing models was constrained by the availability of supernova observations, in particular small numbers of spectra, which limited how much information a model could contain. With spectrophotometric time series data from the Nearby Supernova Factory (SNfactory, Aldering et al. 2002), a much more complex empirical model can be built. The SNfactory dataset includes hundreds of flux calibrated spectral time series which can be used to build a more complete picture of SNIa because there is so much more information per supernova. To equivalently constrain a model with photometry, we would have to have a photometric data set orders of magnitude larger and at a wide range of redshifts in order to deduce what spectral behavior drives variation in the observed broad band measurements and to break degeneracies between different spectral behavior that results in identical broad band behavior in certain wavelength regions.

Given the nature of variations among supernova spectra, such as the strengths and ve-

locities of emission and absorption features, it may be that supernova spectra can only be perfectly described by complex nonlinear models or by simulations. However, in this analysis, we aim to make the next logical extension from the current generation of empirical models. Models such as SALT2 are essentially performing Principal Component Analysis (PCA), finding the first two eigenvectors that describe the space of SNIa lightcurves or spectral time series. The simplest improvement on this is to find more eigenvectors, and thus build a model that describes a larger fraction of the variation in SNIa. The rich data from SNfactory allows us to do this by performing PCA on a matrix made up of our collection of spectral time series. The resulting principal components, which are ranked eigenvectors of the matrix, become the linear components of a new spectral time series model. Using the model to fit new SNIa provides a method for standardizing the SNIa magnitudes, since correlations between the peak magnitude and fit coefficients can be calculated and corrected for.

We review the SNfactory data in Chapter 10. The method used to calculate the model components is described in Chapter 11. The results of our cross-validation procedure are discussed in Chapter 12 and the final results are shown in Chapter 13.

Chapter 10

The Type Ia Supernova Data Set

The supernova data used in this analysis are spectrophotometric time series observed by the SNfactory between 2004 and 2014 with the SuperNova Integral Field Spectrograph (SNIFS, Lantz et al. 2004). This set includes all of the supernovae used in Part I of this work, with the addition of supernovae that have been observed more recently or that have only recently been fully processed and calibrated.

The supernovae are flux-calibrated following the procedure described in Buton et al. (2013), using a system of both CALSPEC (Bohlin 2014 and references within) and Hamuy (Hamuy et al. 1992, 1994) standard stars, which are listed in Buton et al. (2013). The supernovae have been corrected for Milky Way dust extinction (Schlegel, Finkbeiner, and Davis 1998, Cardelli, Clayton, and Mathis 1989). The supernova fluxes have been normalized to a reference redshift of $z = 0.05$ and the wavelengths have been shifted to a common restframe at $z = 0$. All references hereafter to the phase or wavelength of the supernova data will refer to the values in this restframe.

The supernovae used in this analysis range in redshift from $z = 0.01$ to 0.08 and are required to have observations on at least 5 epochs. There must be at least one epoch between 10 days before and 7 days after maximum light, as determined by fitting the data with SALT2, and the supernova must have at least 4 epochs between 10 days before and 35 days after maximum light. Stricter phase-based cuts on the data set are also discussed below. SNfactory judges a supernova spectral time series to be 'bad' if the SALT2 fit to the data does not meet a certain criterion. This procedure identifies supernovae with fatal observational errors, but also can exclude some more peculiar supernovae. In this analysis, we include all peculiar SNe, such as SN1991T and SN1991bg-like objects, plus other SNe that might not be currently included in a cosmology analysis because they cannot be well fit by the SALT2 model. Whether and how to exclude such supernovae from the set of supernovae used for cosmology is discussed below. This analysis does not include Super-Chandrasekhar supernovae, which excludes two supernovae that would otherwise meet the cuts to be used in the analysis. After excluding these and making the above data cuts, the SNfactory data set used here includes 241 supernovae.

Chapter 11

Procedure

The goal of the analysis is to build a model for Type Ia supernovae that will consist of spectral time series components that can be added linearly. Two reasonable choices for the model would be to construct this model for the supernova fluxes, or for the logarithm of the fluxes, which would correspond to the supernova magnitudes. Early testing showed that a linear model in flux space was a better description for the data and we have proceeded with that option. While the logarithm of the fluxes may make it easier for a linear model to fit the color variation in supernovae, a linear model of the fluxes themselves may be a better approximation of the radiative processes underlying supernova behavior. Also, as discussed in Section 11.2, it is possible to separately add a color term to a model in flux-space.

If we had identically sampled, noiseless data, Principal Component Analysis (PCA) would be the clear choice for doing finding a linear model. However, the supernova data available is determined by the time of the supernova discovery, observing cadence and conditions, and the redshift of the supernova, all of which also lead to varying amounts of noise in the data. In this situation it is necessary to devise a procedure for calculating principal components that can handle missing and noisy data.

11.1 Methods for Calculating Principal Components

In the course of this analysis, a number of methods for finding a linear model were tried. The first method was Expectation Maximization Principal Component Analysis (see Bailey, 2012 for a description of the algorithm), which is one method for adding weighting to PCA. This method is supposed to calculate the first N eigenvectors for a given data set. This is done by starting from a set of random eigenvectors, and then alternating between fitting those eigenvectors to the data set and fitting new eigenvectors to the fit coefficients. This method was found to give very inconsistent results depending on the training set and the number of eigenvectors calculated and to not deal very well with noise.

The second method tried was an extension of the method used in Kim et al. (2013) on photometric supernova lightcurves. This method used Gaussian Processes (discussed further

below) to model the true spectral time series of the supernova. Briefly, Gaussian Processes models the true function underlying observations by using a parametrized model of the covariance of the function. For a supernova, this means modeling the correlation between different phases or different wavelengths. The Gaussian Process results in a conditional probability distribution of function values for each supernova. From this distribution, 10 spectral time series realizations were drawn, intended to sample the space of possible true spectral time series for a given supernova. Once this was done, all of the realizations were put into a matrix on which standard PCA could be done. The main problem with this method was that, while the best prediction and covariance for a given supernova were reasonable, individual realizations had characteristics not seen in supernovae. It was not possible to find a wavelength correlation scale short enough to capture all the spectral lines in supernova spectra without allowing the realizations to oscillate too much. Finding the appropriate amount of correlation in the phase dimension might have been achievable, but this required an undesirable amount of tuning of the correlation parameters.

The most successful method, and the one which we will use in this analysis, is in some ways a combination of the above two. First, Gaussian Processes is used to model the true supernova spectral time series. However, in this method we use the expected mean and covariance of the Gaussian Process distribution instead of realizations of the Gaussian Process. Once this is done, Expectation Maximization Factor Analysis (EMFA, discussed more below) is used to calculate principal components. EMFA can incorporate data covariance and has been seen to be more robust to large errors than EMPCA. The resulting output is a set of ranked orthogonal eigenvectors with can then be used as a supernova spectral time series model. The steps in this Gaussian Processes plus EMFA procedure are described in more detail below.

11.2 Data Pre-processing (Gaussian Processes)

Gaussian Processes (Rasmussen and Williams, 2006) models an observation as a sample drawn from some underlying function with an error model:

$$y = f(x) + N(0, \sigma_n^2) \tag{11.1}$$

where $N(0, \sigma_n^2)$ is a normal distribution with a standard deviation of σ_n that accounts for the error in the data. The underlying function is estimated as

$$f \approx N(a\, m(X_\star), K(X_\star, X_\star)) \tag{11.2}$$

where $X_\star$ are the input data points, $m(X_\star)$ is an estimate of the mean function, a is a fit normalization hyperparameter, and K is a covariance matrix made up of elements

$$k(x, x') = \sigma_f^2 \exp\left[\frac{-(x - x')^2}{2l^2}\right] + \sigma_n^2 \delta(x, x') \tag{11.3}$$

where $k(x, x')$ is called the kernel and x is an individual input point. The hyperparameter l determines the length scale of correlation between data points, while σ_f determines how much variation away from the mean is allowed and σ_n models unaccounted for error in the input data, which is diagonal and so is multiplied by the Dirac delta function $\delta(x, x')$. The hyperparameters can be fixed or can be fit by, for example, maximizing the log likelihood of the model. Given input data at $X_\star$ with variance V, the prediction for the underlying supernova spectrum at points X is then the mean of the conditional distribution of function values drawn from the normal distribution:

$$\mathbf{f} = a\ m(X) + K(X_*, X)[K(X, X) + V]^{-1}(\mathbf{y} - a\ m(X)) \tag{11.4}$$

with covariance

$$cov(\mathbf{f}) = K(X_*, X_* - K(X_*, X)[K(X, X)]^{-1}K(X, X_*) \tag{11.5}$$

In the application at hand, the relevant hyperparameters are the correlation length scales of the supernova data points in phase and wavelength, the normalization of the mean function, the kernel amplitude, i.e. the amount by which the prediction can vary from the mean function, and the error model. As discussed above, correlation between points at different wavelengths tends to lead to over-smoothing of spectral features, and it is not particularly necessary here, since the data is already at the spectral resolution desired. For these reasons, we do not include a wavelength correlation parameter in the kernel.

The phase correlation length scale should be similar for all supernovae. However, if it is fit by maximizing the log likelihood of the model, the value will be highly sensitive to the average separation in time of the data points. Instead we choose the value that maximizes the total log likelihood for all of the supernovae. This results in a phase correlation length of 10 days. The kernel amplitude and mean normalization can be reasonably assumed to be different for individual supernovae and are fit for each by maximizing the log likelihood of the model. The mean function itself is calculated by fitting a normalized mean to all the raw supernova data, with the Hsiao supernova template (Hsiao et al., 2007) as a prior, and with Savitzky-Golay smoothing (Savitzky and Golay, 1964) in the time axis.

The error model presents more complications. There are two options available for accounting for error. Assuming that the covariance matrix of the observations is known, it can be fixed in the Gaussian Processes model and accounted for in fitting the other hyperparameters. The measured supernova variance in the SNfactory data set includes the photon noise, but does not include any uncertainties from the data processing pipeline. Observations of standard stars indicate a grey scatter in the SNfactory data set at about $2 - 3\%$ of the flux, depending on the observing conditions. This can be explicitly included in the error model, but here we let the error be fit as described below.

The alternative to using known errors is to assume some model for the covariance and fit for exact values at the same time as fitting the other hyperparameters. In this analysis, we assume that there is covariance between data points with the same phase, corresponding to adding a grey scatter term, and an uncorrelated scatter term, corresponding in turn to

any noise in the data that is not accounted for by the photon noise. Testing revealed that if the grey scatter term is fit freely, it will become very large and actually result in the gaussian process giving a very biased prediction for the supernova. Putting a strong prior on the size of the error prevents this from happening, but relies on knowing the expected value. In the end, we decided to fix the grey scatter term to zero and let the diagonal noise term be fit along with the other hyperparameters. This gave the results that were the least biased with respect to and had error bars most consistent with the data, and resulted in predictions that evolve on the time scale of known Type-Ia supernova behavior without relying on hand-tuning the hyperparameters.

With the hyperparameters determined thusly, the best fit for the true spectral time series is found for each supernova in the data set. The predicted spectral time series for each is calculated at the same wavelengths as the data, at two day intervals from 10 days before maximum to 46 days after maximum. Now the supernovae are evenly sampled with a comprehensive covariance matrix and are ready to be used to calculate principal components. Figure 11.1 shows an example of the Gaussian Process prediction for a moderately well-sampled supernova with some grey scatter in the data. Figure 11.2 shows the performance of the predictions over the whole supernova set, demonstrating that the predictions are unbiased and have residuals consistent with error on the data.

Inclusion of a color relation

Supernova color difference has been shown to be a major contributor to variability in supernova spectra and magnitudes (Tripp 1998). This color is some combination of intrinsic color and reddening due to dust, but a definitive method for separating these components has not yet been discovered. The portion due to dust reddening is thought to obey the Cardelli dust law (Cardelli, Clayton, and Mathis 1989), or another similar model such as Fitzpatrick and Massa (2007). Dust reddening is modeled as an exponential multiplier to the flux and thus cannot be described by a single linear principal component. Dereddening the supernovae using the Cardelli law before performing PCA may better match supernova behavior and mean that a smaller basis of linear components is needed to describe supernova variability. Thus, in addition to calculating a purely linear supernova model, we will also calculate models using supernova data dereddened with the Cardelli and Fitzpatrick dust reddening models.

The color relation is applied by fitting the color difference between each supernova and a fiducial supernova and dereddening the supernova by that amount. This is done by minimizing the quantity

$$\sum_{SNe}\sum_{p,\lambda}\frac{(f_{SN_i}(p,\lambda) - a\ 10^{-0.4\ E(B-V)\ c(\lambda)}\ f_{fid}(p,\lambda))^2}{\sigma^2_{SN_i}} \tag{11.6}$$

as a function of a, $E(B-V)$, and $f_{fid}(p,\lambda)$, where $E(B-V)$ is the color difference, a is a normalization factor, and $f_{fid}(p,\lambda) = f_{fid}(phase, wavelength)$ is a freely fit spectral time

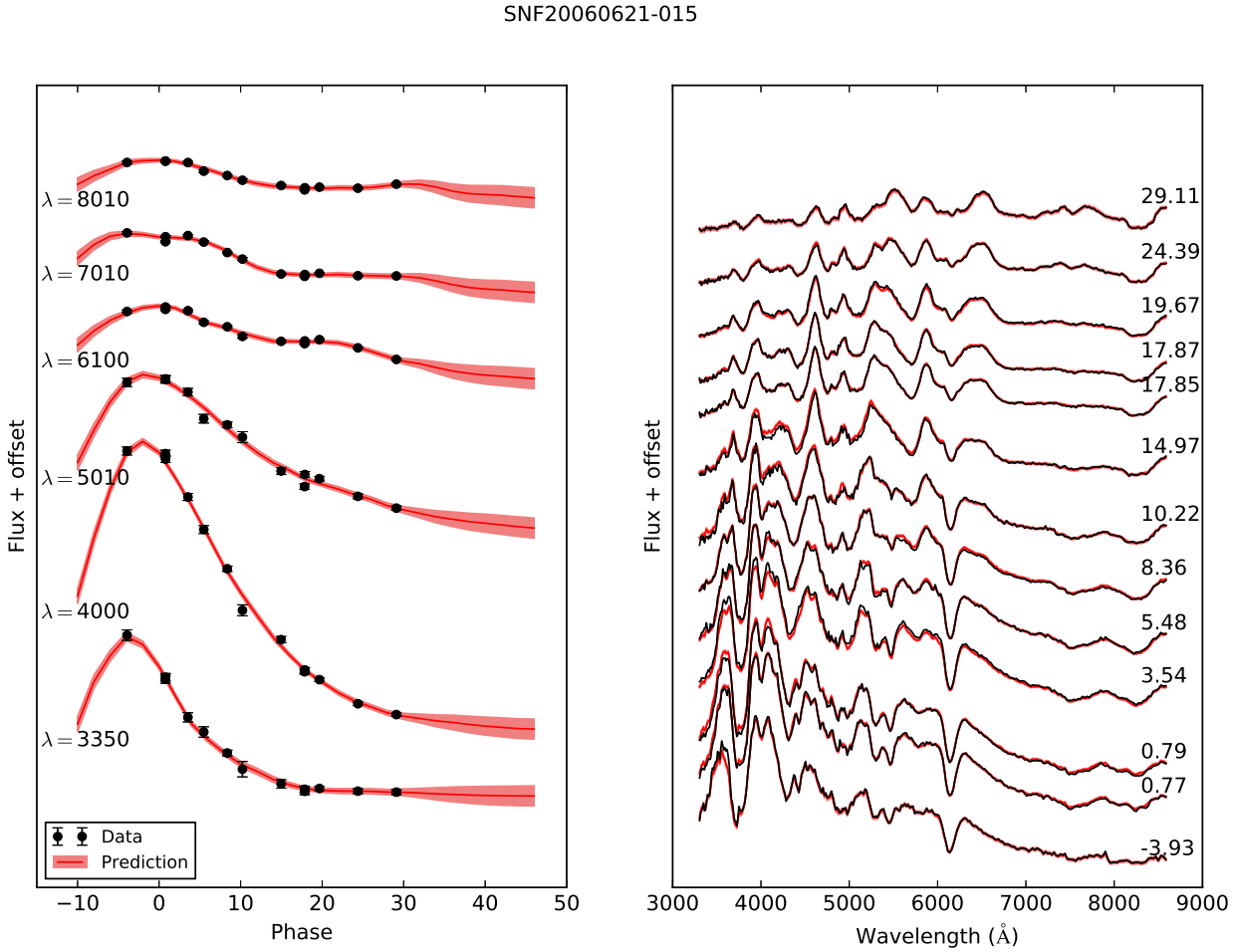

Figure 11.1: A demonstration of the supernova model predicted by Gaussian Processes compared with the supernova data. The left panel shows monochromatic lightcurves while the right panel shows the spectra predicted by the gaussian process at the phases where there is supernova data.

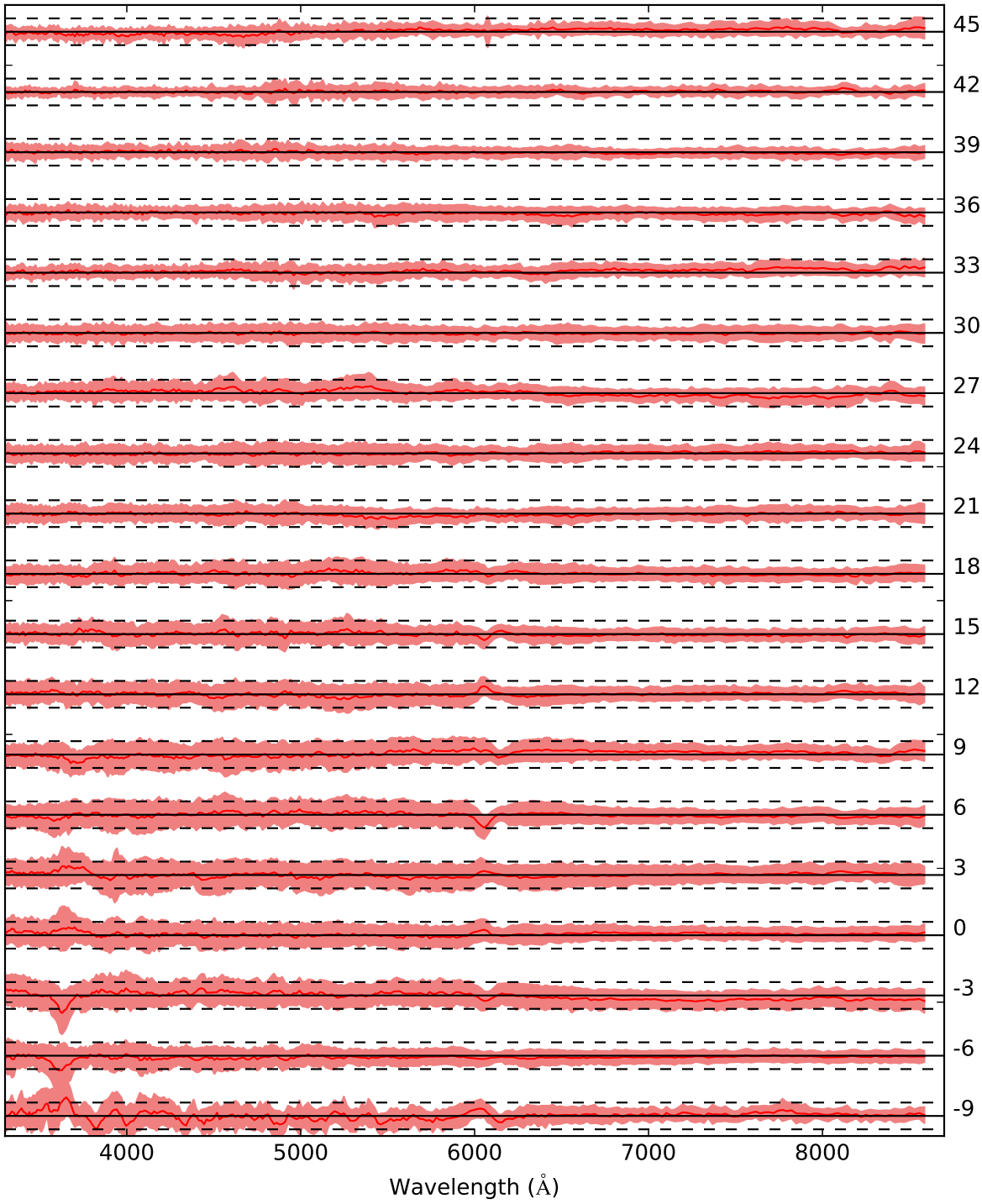

Figure 11.2: Average and standard deviation of the Gaussian Process pulls. The red lines show the average pull, which is the residual divided by the error, of the Gaussian Process prediction to the data at the phase indicated to the right of the figure, while the shaded region shows the standard deviation in the pulls. The solid black lines are the zeropoint for the averages, where a level of zero indicates that there is no bias. The dashed lines show the level of plus and minus one pulls.

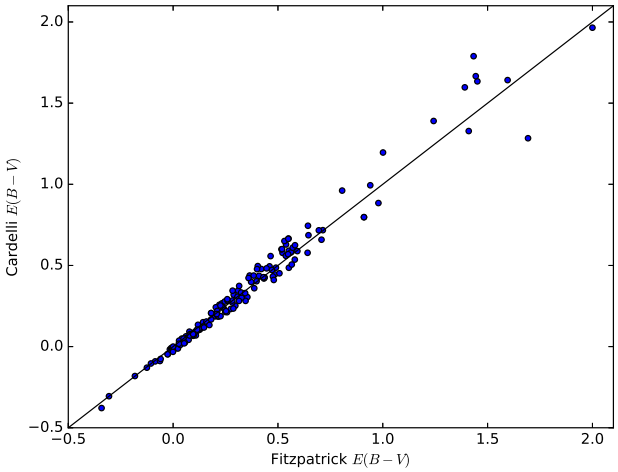

Figure 11.3: Relation between $E(B-V)$ values found for Cardelli and Fitzpatrick dereddening relations. A one-to-one line is shown for comparison.

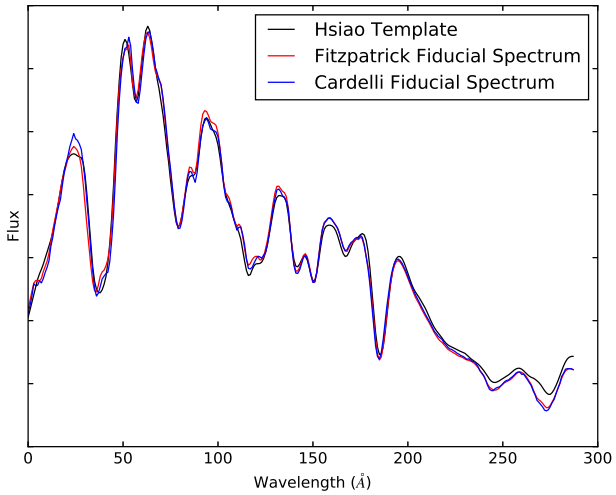

Figure 11.4: Fiducial spectra at maximum found for Cardelli and Fitzpatrick dereddening relations, dereddened and normalized to match the Hsiao template (Hsiao et al. 2007).

series. $c(\lambda)$ is either the Cardelli or Fitzpatrick dereddening model, both using $R_V = 3.1$. The results for the $E(B-V)$ distributions are shown in Figure 11.3 and the fiducial spectra for the Cardelli and Fitzpatrick dereddening models are shown in Figures 11.4. The fiducial spectra, fit by minimizing Equation 11.6, are essentially the spectral time series that is on average the best match to all the normalized and dereddened supernovae.

The supernovae dereddened using the fit $E(B-V)$ values (but not including the normalization factor a) are then used as input to the EMFA. These results are compared to the

EMFA on the original spectra in the cross-validation process discussed below.

11.3 Linear Decomposition (Expectation Maximization Factor Analysis)

Expectation Maximization Factor Analysis (EMFA) is a method for performing dimensionality reduction presented in Ghahramani, Hinton, et al. (1996). The generative model is given by

$$\mathbf{x} = \Lambda \mathbf{z} + \mathbf{u} \tag{11.7}$$

where x is the $p \times n$ dimensional data vector and Λ is a $n \times k$ dimensional matrix of model components, where k is generally much smaller than p, with factors z, corresponding to the coefficients of the model. u is a p-dimensional random variable with distribution $N(0, \Psi)$, corresponding to the noise in the data. The method has been adapted here to accommodate correlated noise in the data. EMFA is preferable over tradition PCA here because it is able to define a proper density matrix for the data. The method consists of an iterative process in which the expected value of the factors z is calculated, given Λ, and then new values for Λ and Ψ are calculated given the data and the new values for z. This is repeated until the log likelihood converges. Once convergence is reached, the components in the matrix $\mathbf{\Lambda}$ can be orthogonalized and ranked.

The resulting model for the data takes the form

$$f_{SN}(p, \lambda) = \sum_{i=0}^{N} c_{SN,i} e_i(p, \lambda) \tag{11.8}$$

where f_{SN} is the flux of the supernova at the phase p and wavelength λ, N is the number of model components, e_i are the model components, and $c_{SN,i}$ are the model coefficients fit to the individual supernova. Since the model is expected to be used for supernovae at a range of redshifts, we rewrite the model in the form

$$f_{SN}(p, \lambda) = c_{SN,0} \left(e_0(p, \lambda) + \sum_{i=1}^{N} c'_{SN,i} e_i(p, \lambda) \right) \tag{11.9}$$

so that $c_{SN,0}$ is the only model coefficient that is dependent on the flux normalization and is thus the only coefficient that is a function of the luminosity distance.

11.4 Model Testing

Given a sufficient number of components, the results of the EMFA a priori will perfectly describe the data on which they were trained. Given the noise in the data and the incomplete nature of a finite training set, only some of the principal components will describe true

supernova behavior while the rest will be overtrained on the noise. To determine how many components are describing real supernova behavior, and to make any other decisions about the model, it is necessary to use cross-validation, in which a model is trained on one set of observations and tested on another independent set of observations.

In this analysis, to allow for testing different techniques and to ensure especially robust results, we use the following procedure: from the original set of 240 supernovae, 60 supernovae, or 25% of the data, are removed in order to be used as a final demonstrative test set, then K-fold cross-validation is used on the remaining 180 supernovae.

In K-fold cross-validation, the data set is split into K subsets, where here K equals four, then the model is trained on all but one of the subsets and tested on the remaining subset. This is repeated K times, using each subset as the test set once, and the results are combined. This method uses more of the data than having a single test set, and guarantees that the results will not be skewed by an anomalous distribution of supernovae that could occur in a single test set, while being less expensive than full leave-one-out cross-validation. Once all decisions have been made about the best training set and number of components to use, the model is retrained on all 180 supernovae used in the K-fold cross validation. The final results are applied to the 60 remaining supernovae in the hold-out set as a final test and a demonstration how well the results work with outside data.

In this analysis, there are two important metrics upon which to judge the models. The first is simply how well the model fits the data, which can be quantified with a number of metrics, which here include the χ^2 metric, Bayesian Information Criterion, and Akaike Information Criterion. The second is how well the model can be used to standardize supernova magnitudes. Both are discussed further in Chapter 12.

Variations on the General Data Set

Cross-validation is needed to choose the best option among three types of variations on the model. The first variation is the number of components that can be used to model the full data set, as discussed above. The second variation is whether to deredden the spectral time series prior to calculating the principal components and, if so, whether to use the Cardelli or Fitzpatrick reddening relation.

The third variation is whether to make any cuts to the data set based on the phase coverage of the original observations (in addition to the selection criteria described in Chapter 10). In training a model, it is preferable to have as many supernovae as possible. However, if a given supernova has no observations at early or late times, a large amount of extrapolation will be done by the Gaussian Processes model, described in Chapter 11, making the model of that supernova less reliable. Since significant information about the supernova is believed to be contained in the early and late time spectra (Scalzo et al. 2014, Nordin et al. in preparation), using such supernovae in the PCA training set may dilute the strength of important early and late time indicators in the principal components, weakening their ability to replicate the true variability in supernovae. Additionally, lack of observations before the

supernova's peak weakens the constraint on the supernova phase, which will contaminate the training data.

For this reason, in addition to using all of the supernovae for training the model, we test making two cuts on the supernovae used in the analysis based on the phases of the data as determined from a SALT2 fit to their lightcurves: first, we cut all supernovae that do not have an observation before maximum; second, we cut all supernovae that do not have observations both before maximum and after 30 days past maximum.

In sum, we must decide how many components to use, whether and how to deredden the data, and how well observed the training data must be. The K-fold cross-validation process is used to decide which of these options produces the principal components that best describe the data.

Chapter 12

Results of Cross-Validation

As described in Chapter 11, Gaussian Process predictions are made for all the supernovae, and then used to calculate principal components with EMFA. We calculate the EMFA with up to 16 principal components. For each data set configuration and each dereddening option (Cardelli, Fitzpatrick, or none), we have three metrics to judge how well the models fit the data: the χ^2 values, the Bayesian Information Criterion, and the Akaike Information Criterion. Along with these, we will consider how well each models can standardize the supernova magnitudes. All of these metrics are considered as a function of the number of eigenvectors used in the model. Each is calculated separately for each of the k-fold training and test sets and then averaged to get combined training set and test set results.

12.1 Model Reconstruction of the Data

For judging how well the model matches supernovae, we want to consider the improvement in the χ^2 value as a function of the number of components in the model. The χ^2 will improve (i.e. decrease) continuously for the training data set. However, if some of the components are overtrained, the χ^2 for the test set will not show the same improvement when those components are added to the model. If the the test set χ^2 stops improving as much as the training set χ^2 at a certain number of components, that will show how many components are describing true supernova behavior and how many are overtrained to the training set. The relative values for the χ^2 metrics also give us a way to compare the models resulting from different data selections and dereddening options.

The χ^2 results as a function of the number of eigenvectors, shown in the upper panels of Figures 12.1, 12.2, and 12.3, are largely the same regardless of the data selection. Additionally, applying the eigenvectors trained on the restricted data sets to the full test sets removes most of the remaining differences. Comparing the χ^2 values between the models with and without dereddening shows a slight preference for the Fitzpatrick reddening relation.

The Bayesian and Akaike Information Criteria, shown in the lower panels of Figures 12.1, 12.2, and 12.3 provide methods for choosing the number of model components to use. Both of

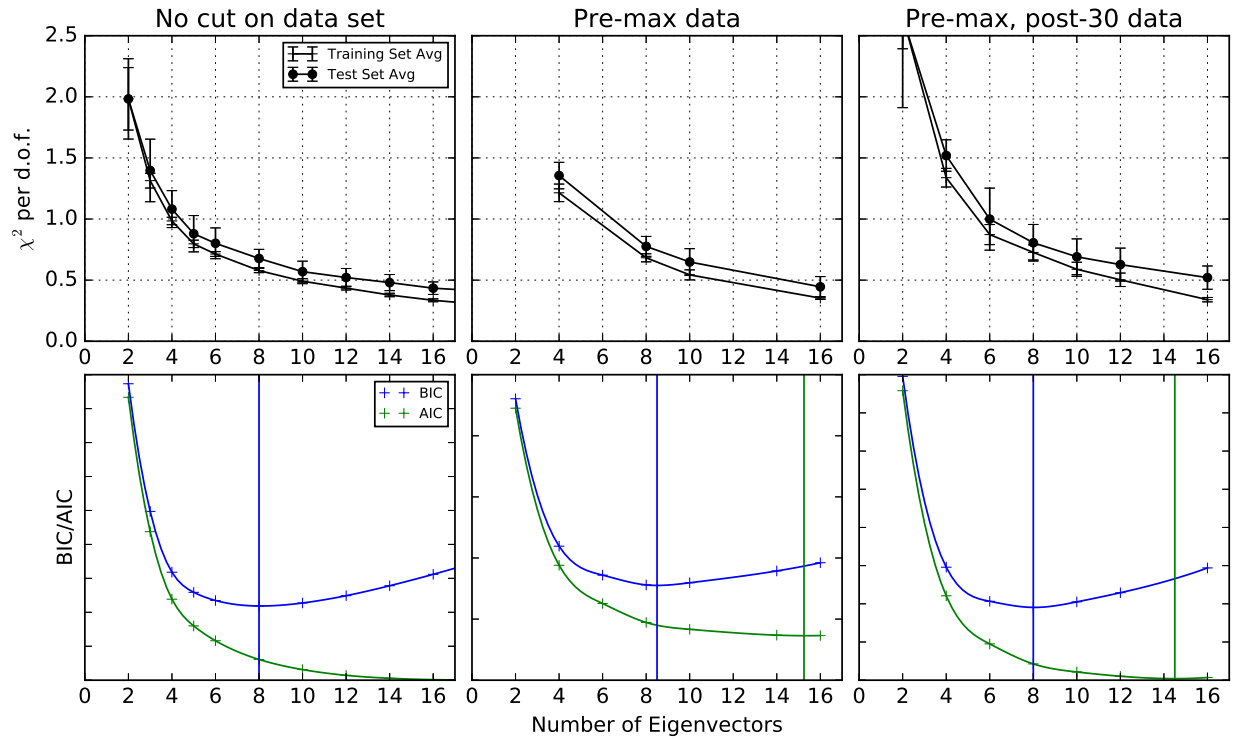

Figure 12.1: Cross Validation results for data with no dereddening with the metrics calculated as a function of the number of components in the model. Error bars indicate the standard deviation among the four training or test sets. The blue vertical lines in the lower panel show the minima for the BIC metrics, while the green vertical lines show the minima for the AIC metrics

these are based on weighing the likelihood of the model versus the number of free components in the model and the number of data points fit.The BIC puts a heavier penalty on adding model components than the AIC, and gives a minimum value at 7 eigenvectors. The AIC has a relatively light penalty on adding model components and does not reach a minimum until about 15 eigenvectors.

12.2 Magnitude Standardization

The second way to judge the models is to look at how well they can be used to standardize the supernova magnitudes. As in other supernova lightcurve models, the correlation between the peak B-band magnitude (or other bands' magnitudes) and the supernova model coefficients can be calculated and used to standardize the supernova magnitudes. While it may turn out not to be sufficient, the simplest and currently most common way to do this is by fitting a linear relation between the coefficients and the B-band peak magnitude, plus the fit value

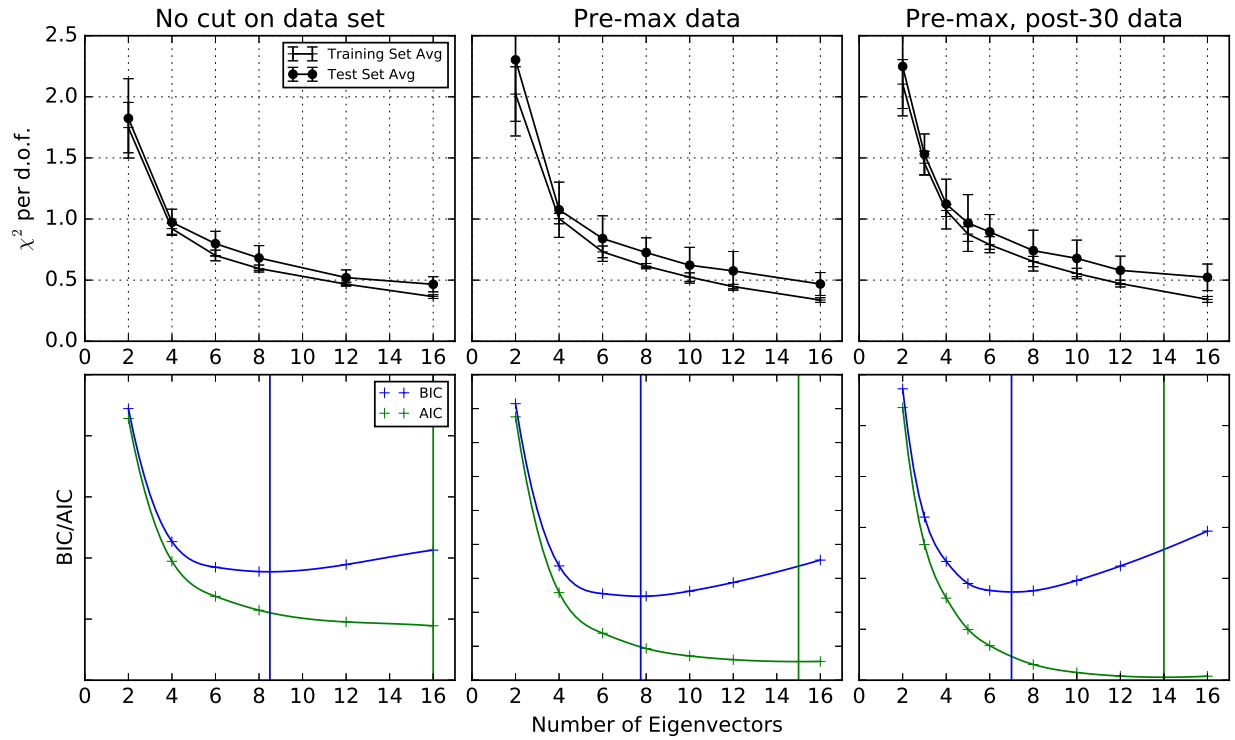

Figure 12.2: Same as Figure 12.1 but for data with Cardelli dereddening.

for E(B-V) in the case where a color law is used. For example,

$$\mu_B = m_B^{corr} - M_B = m_B + \sum_i \alpha_i \, c_i \;\; (+ \beta \times E(B-V)) - M_B \tag{12.1}$$

where M_B is the standardized absolute magnitude of the supernova and μ is the distance modulus. Following the method of the 'Union' compilation of supernovae (Amanullah et al., 2010), the $\boldsymbol{\alpha}$ and M components are fit by iterating between maximizing the log-likelihood function (neglecting the constant π term)

$$\mathcal{L} = -\ln\det(\sigma_{pv}^2 + \sigma_{int}^2 + \sigma_m^2) - \sum_{SNe} \frac{[\mu_B(\boldsymbol{\alpha}, \beta, M_B) - \mu(z,\theta)]^2}{\sigma_{pv}^2 + \sigma_{int}^2 + \sigma_m^2} \tag{12.2}$$

and fitting the value of the intrinsic dispersion σ_{int} so that the reduced χ^2 is equal to unity. Here $\mu(z,\theta)$, where θ stands for the parameters of the fitted cosmological model, is zero since the supernovae in the data set are in the restframe, σ_{pv} accounts for uncertainty due to host galaxy peculiar velocities of 300 km s^{-1} (changing this value by 100 km s^{-1} results in $< .01$ mag changes in the intrinsic dispersion), and σ_{lc} is the uncertainty in the supernova model:

$$\sigma_m = \mathbf{V}^T \mathbf{C} \mathbf{V} \qquad \text{where} \qquad \mathbf{V} = \begin{pmatrix} 1 \\ \beta \\ \boldsymbol{\alpha} \end{pmatrix} \tag{12.3}$$

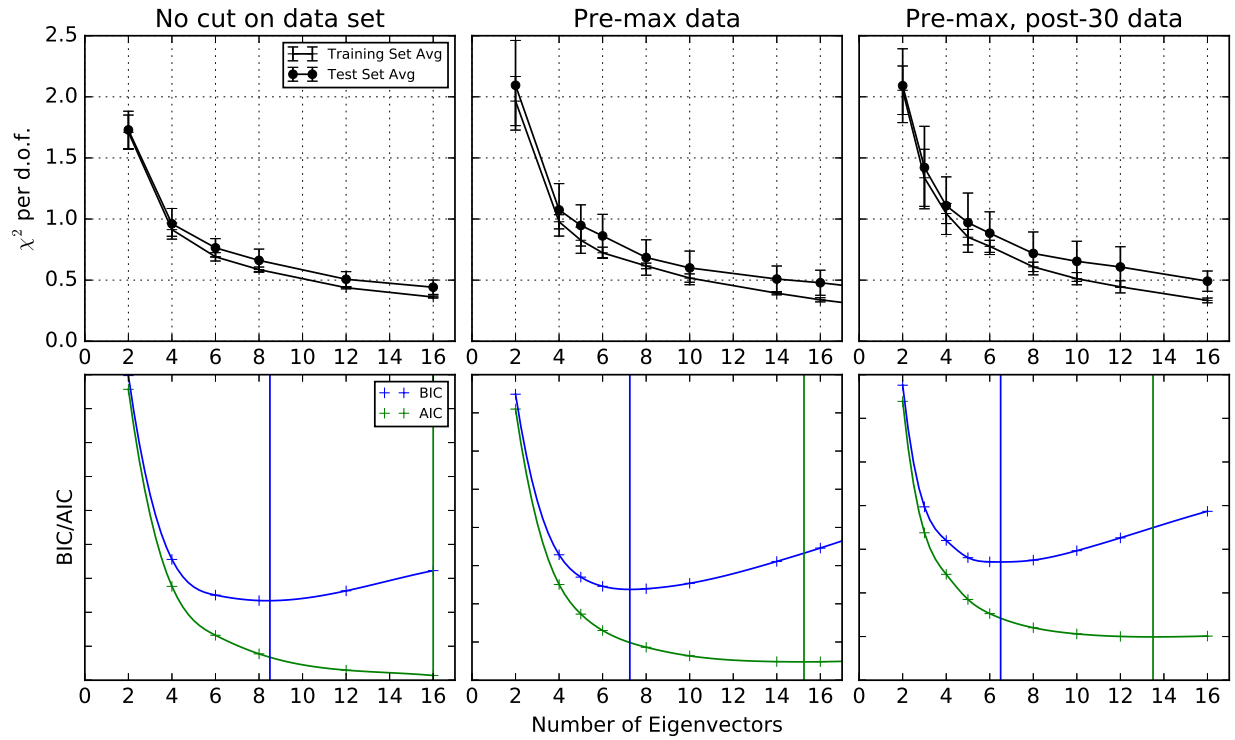

Figure 12.3: Same as Figure 12.1 but for data with Fitzpatrick dereddening.

and $\mathbf{C}$ is the covariance in the model fits to the supernova realizations.

To judge how many of the model components c_i are truly correlated with the magnitude, and thus useful for cosmology, we can apply a method similar to the one used above with the χ^2 metric: the global correlations between the magnitudes and model parameters can be fit to the training data and then applied to the test data. This is done using only the first model component, then the first two model components, etc. As more components are added, the dispersion in the standardized magnitudes will get lower and lower for the training data set. However, after a certain number of components, the correlations fit will only be noise. At this point, the dispersion in the corrected magnitudes for the training set may keep decreasing, but the correlations will not fit the test set, so the dispersion in the corrected magnitudes for the test set will stop improving. This will tell us how many model components will be useful for cosmology, which will necessarily be equal or less than the number of components that describe real features of the data set, identified by the χ^2 in the method above.

For calculating the correction function and the resulting dispersion in the corrected magnitude we first exclude extreme outlier supernovae. To do this we calculate the Mahalanobis distance, or normalized metric distance, of each of the supernovae in the space of the model coefficients. The coefficient corresponding to the normalization of the model, which is degenerate with the supernova magnitude, is not included in this calculation. Supernovae are

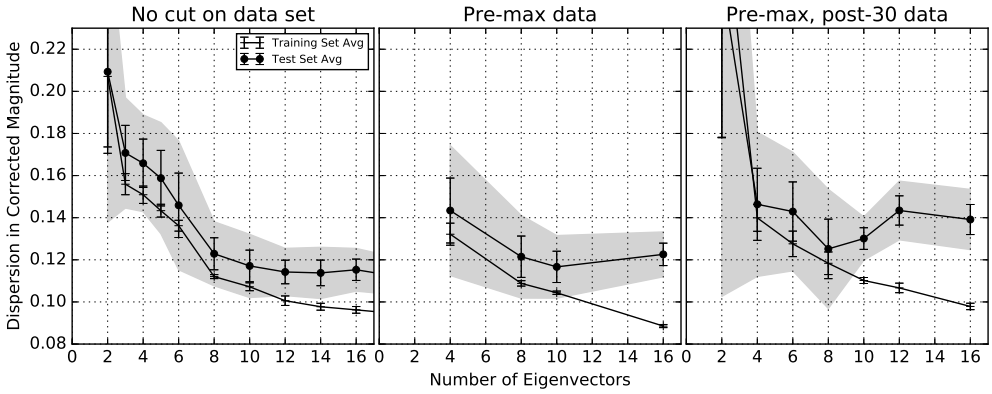

Figure 12.4: Dispersion in standardized magnitudes for data with no dereddening. The solid lines show the result averaged over all of the K-fold sets, while the shaded region shows the standard deviation among the sets and the error bars shows the error on the mean.

excluded that have a probability of less than .001% according to the expected Chi-squared distribution of the Mahalanobis distances. This procedure is then repeated once, since outliers affect the calculation of the Mahalanobis distances. This removed six supernovae from the standardization analysis, which were identified after the fact as either SN1991T or SN1991bg-like. Such supernovae would generally be excluded in current cosmological analyses, assuming they could be correctly identified, though future non-linear methods for magnitude standardization may allow these types of supernovae to be brought back into the set used for cosmology. In Appendix B we show the results of the EMFA analysis with theses outlier supernova removed from the training data.

While it is not a direct measurement of how well the eigenvectors model the data, the dispersion in the corrected magnitudes (shown in Figures 12.4, 12.5, and 12.6) is an important metric for using the eigenvectors for supernova cosmology. The dispersion reaches a minimum or a rough plateau for the test sets at between 4 and 6 eigenvectors. Like the χ^2 method, the dispersion shows a slight preference for the eigenvectors using the Fitzpatrick dereddening relation. The set where we only use supernovae with pre-max data gives the lowest dispersion for the smallest number of eigenvectors. Again, this is determined by this particular standardization method, and more advanced standardization techniques may favor a different number of eigenvectors.

12.3 Conclusions of K-fold Cross-Validation

Between the χ^2, BIC, and AIC per number of eigenvectors, there is not a strong preference for one number of eigenvectors. For this reason, we decide to provide three options for the final analysis: a set of seven eigenvectors, corresponding to the model that is preferred for the BIC and is best for making a linear correction to the supernova magnitude; a set of 15

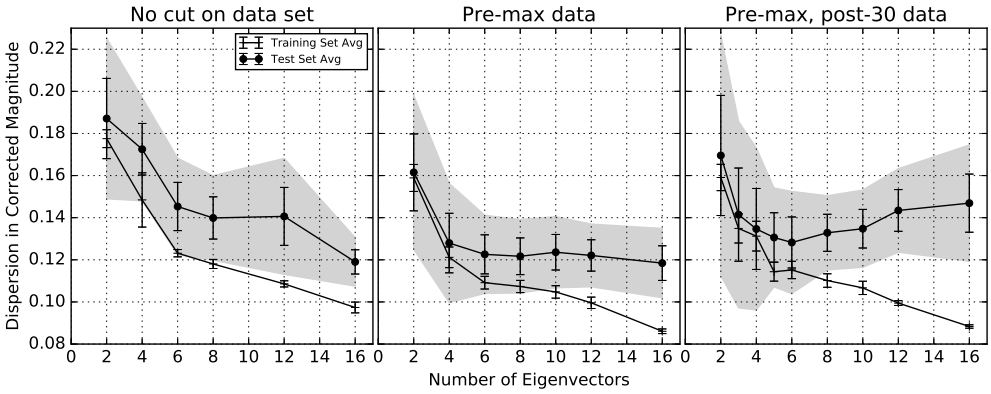

Figure 12.5: Same as Figure 12.4 but for data with Cardelli dereddening.

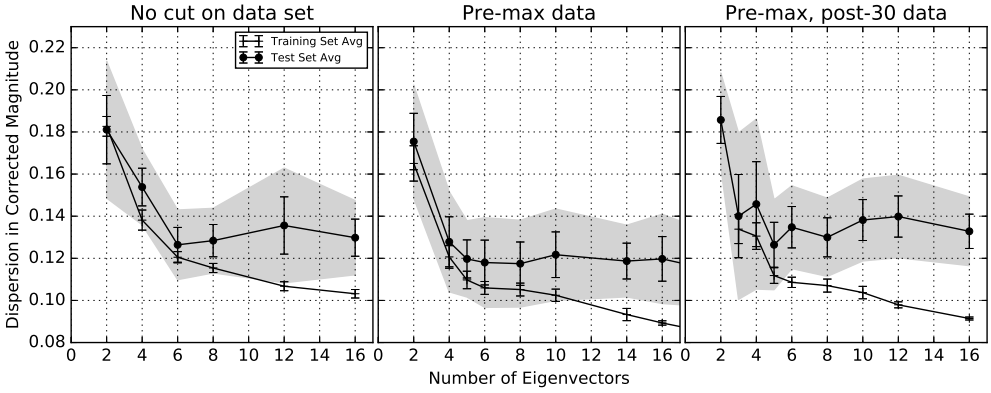

Figure 12.6: Same as Figure 12.4 but for data with Fitzpatrick dereddening.

eigenvectors, which is preferred by the AIC and corresponds to a more complete model of the supernova spectral time series; and a set of two eigenvectors that, together with a redenning relation, will provide an analogue to the SALT2 model for comparison purposes. Based on the dispersion results from the cross-validation, we will only use the first six components for standardizing magnitudes. These will be trained on the set of 121 supernovae with data before maximum, dereddened with the Fitzpatrick relation, which is the set that resulted in the lowest dispersion in standardized magnitudes at the number of eigenvectors preferred by the BIC metric. This choice is also conceptually preferable based on our previous knowledge of the impact of dust reddening on supernova spectral time series and the fact that having data before maximum is important for constraining the date of maximum and correctly aligning the supernova data used in our calculation of the eigenvectors.

12.4 Comments on Unblinding the Model

In performing the cross-validation here, we have found that the the K-fold training and test sets have been largely consistent with each other. This can be seen foremost from the fact that the model coefficients of the training and test sets have consistent distributions, which indicates that the populations are largely similar. It is possible that by chance the randomly selected supernovae in the final hold-out set will have a population that is significantly different from the full training set. Should we find upon retraining the model that the final training and hold-out sets have significantly differently distributed model coefficients, we will have reason to believe that we will not see the same results. The average χ^2 and the dispersion in the corrected magnitudes are expected to be worse if the hold-out set coefficient distributions are larger or biased with respect to the training set coefficients, but the average χ^2 and dispersion for the hold-out set could also be lower than in the training set if by chance there are fewer peculiar supernovae in the validation set.

For calculating the corrected magnitudes, in the cross-validation procedure we fit the correction parameters to the training sets and applied them to the test sets. Once we have retrained and applied the model to the final hold-out set, the correction parameters will be calculated on the final validation set itself. This replicates what would be done in a cosmology analysis, where the correction parameters are fit at the same time as the parameters of the models for cosmological expansion. As in the cross-validation, outlier supernovae will be removed before calculating the correction parameters and the final dispersion. As a final check, we can determine whether calculating the correction parameters on the full training set and the final validation set give consistent results. In this iteration, we will be fixed to correcting the magnitudes with the number of components determined in the cross-validation step. In a true cosmological analysis with a large enough data set we would not be concerned about over-training the magnitude corrections and could hypothetically use more of the model coefficients for standardization, assuming that the number of supernovae would be many times the number of possible correction components. However, based on the cross-validation results, we do not expect to find lower dispersion using a linear correction on more than six components. Using more complicated functions for standardizing the magnitudes may allow more components to be utilized, which will be tested in future work.

Chapter 13

Final Results

The models with 2, 7, and 15 components are recalculated with the full training set (restricted to supernovae with observations before maximum) of 121 supernovae. The components of the models are shown in Figures 13.1, 13.3, and 13.4. For demonstrative purposes, we only show spectra at a few intervals in phase and monochromatic lightcurves at illustrative wavelengths. For the 7 and 15 component models, only the first four eigenvectors are shown, with the rest shown in Appendix A. The SALT2 model is also shown for comparison with the two component model in Figure 13.2. Another version of the models is shown in Appendix B in which no outlier (which here are all SN1991T or SN1991bg-like) supernovae are included in the analysis.

A number of characteristics are immediately apparent in the models. As expected, Eigenvector 0 for each of the models looks like a very standard Type Ia supernova, though there is some difference in the lightcurve shapes for the 7 and 15 component models, with blue-bands that appear less bright at peak with respect to the other bands than is seen in the 2-component and SALT2 models. It is also clear that the lightcurve shapes of the next few components indicates that they will have a strong impact on the peak brightness of the supernova at peak and at 15 days past maximum, the approximate time of the second maximum in the redder bands. Lastly, there is a lot of broad-band variation around 30 days past maximum, which may correlate with the variation in the Nickel mass, found in Scalzo et al. (2014) to have a correlation with the behavior of the lightcurve at later phases.

Spectroscopically, it can be seen that the components have the strongest impact on the model spectrum in areas in which we know there is a lot of supernova variability. In the model components shown, most have a strong feature at the wavelength of the Si II '6150' absorption feature. Also very noticeable are strong features in the near-UV section of the spectra before and at maximum light.

Figure 13.5 shows an example of fitting a supernova with each of the three models. It can be seen from the residuals between the data and the models that there are major improvements made moving from the two component model to the seven component model. In the example shown, the two component model misses both broad-band behavior and spectral features of the supernova. Most of these are matched by the seven component

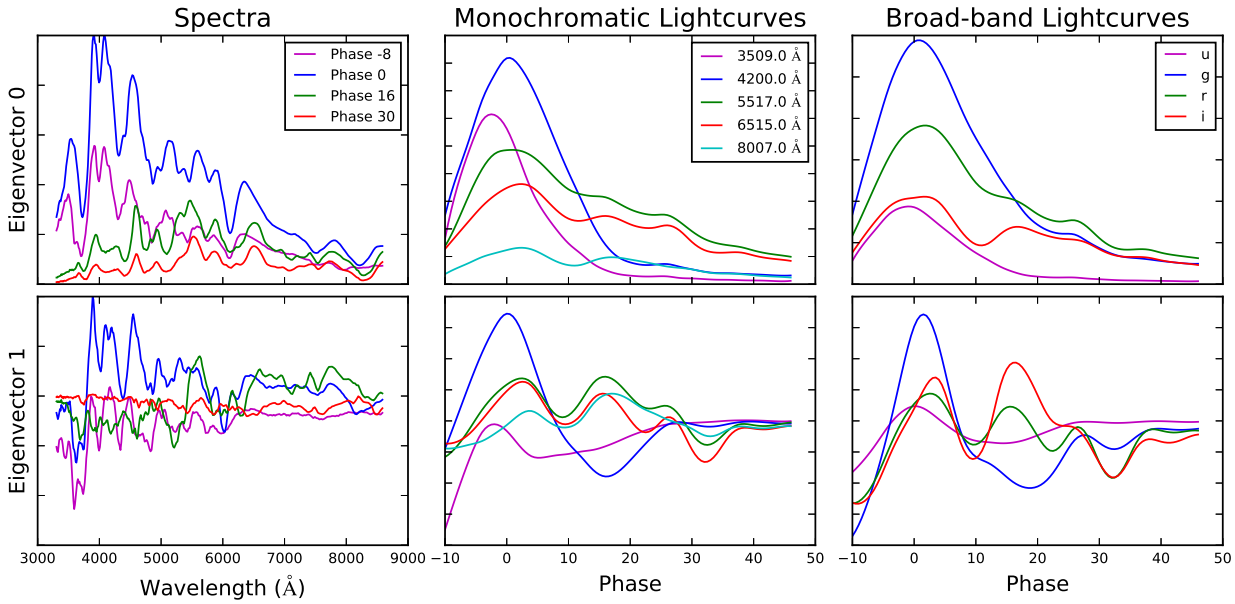

Figure 13.1: Demonstrative spectra and lightcurves for the two component model

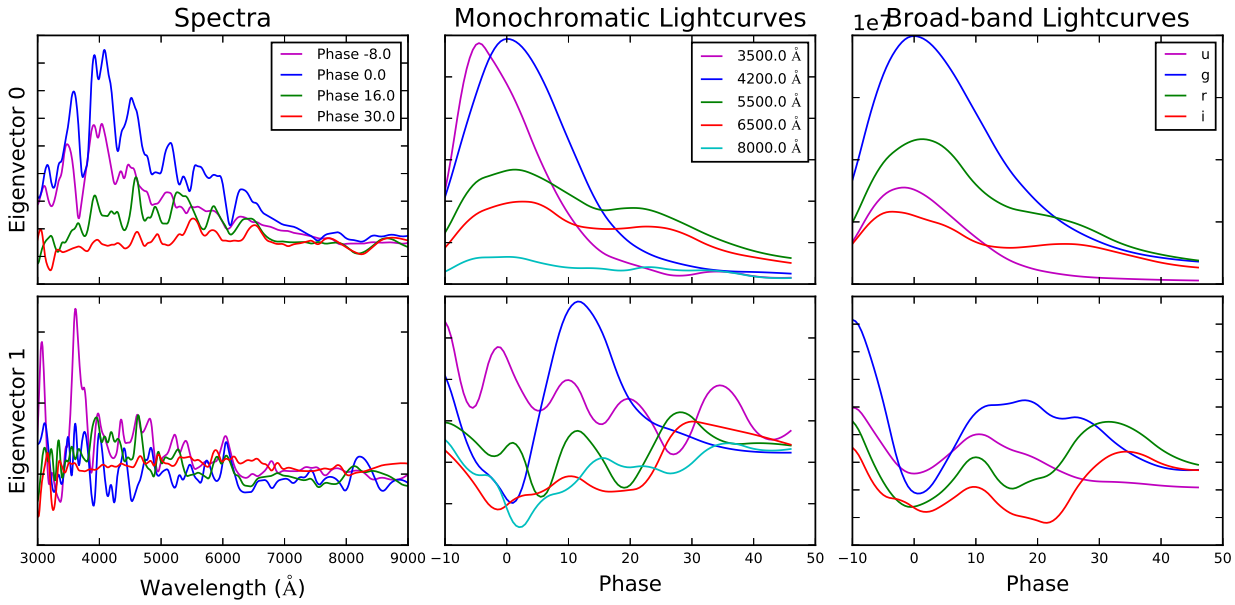

Figure 13.2: Demonstrative spectra and lightcurves for the SALT2 Model

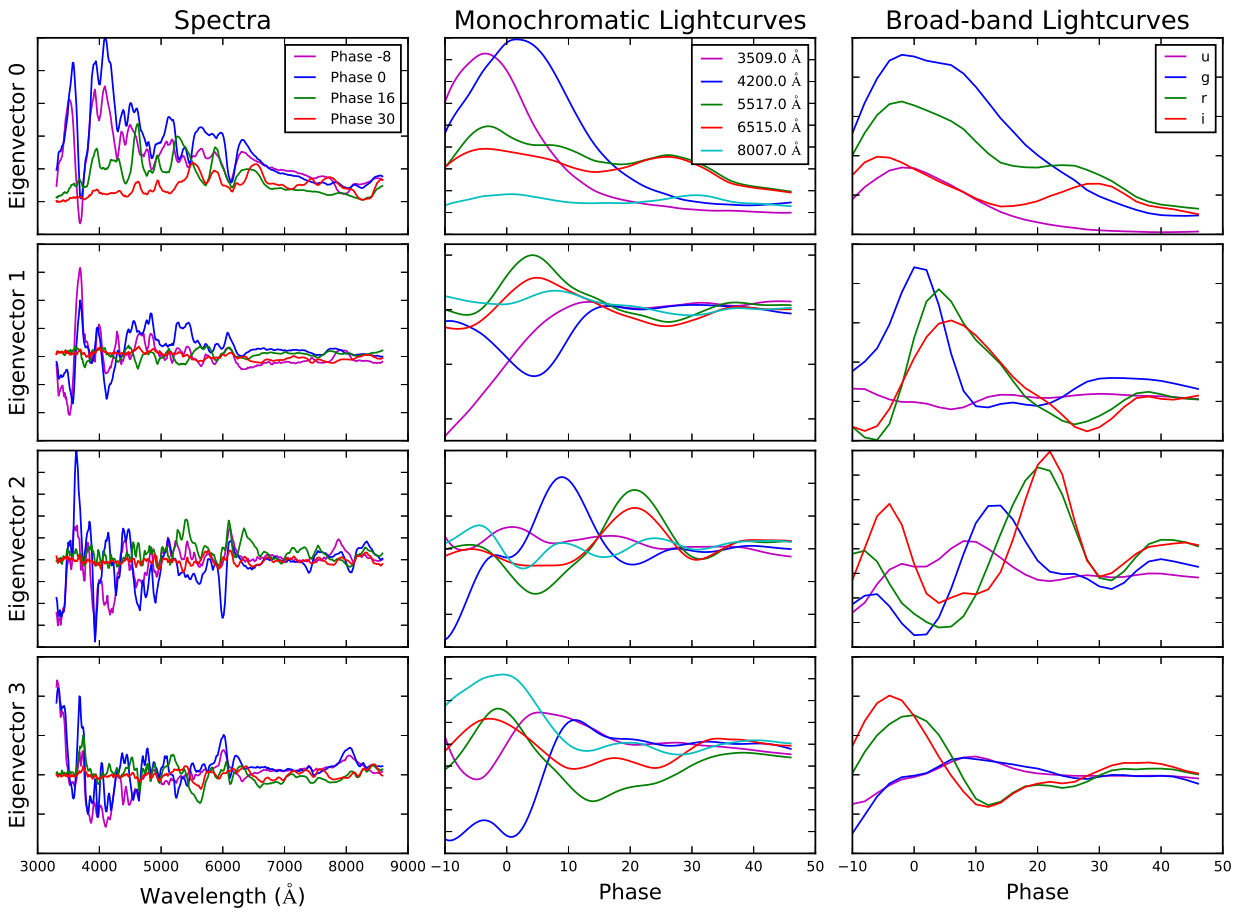

Figure 13.3: Demonstrative spectra and lightcurves for the first four components of the seven component model

model, while the fifteen component model makes some additional improvements.

A more comprehensive view is found looking at the standard deviation of the residuals for all of the supernovae. This is shown at intervals in phase in Figure 13.6. A large standard deviation in the residuals shows places where there is a lot of diversity in the supernova spectral time series that is not fit by the model. In this case, we see that for the two-component model, there are large residuals at the Si II 6150 feature, with somewhat smaller residuals at the Si II features near 4000Å. Large residuals are also apparent in the near-UV and Ca H&K region, particularly around maximum light. At and after maximum light, we also see a lot of spectral features in the residuals corresponding to the Fe II blend around 5000Å. Aside from spectral features in the residuals, broad-band residuals are also apparent at and after maximum light.

For the seven component model, the spectral and broad-band residuals are greatly reduced. Some unexplained supernova diversity is still apparent at Si II 6150 and in the

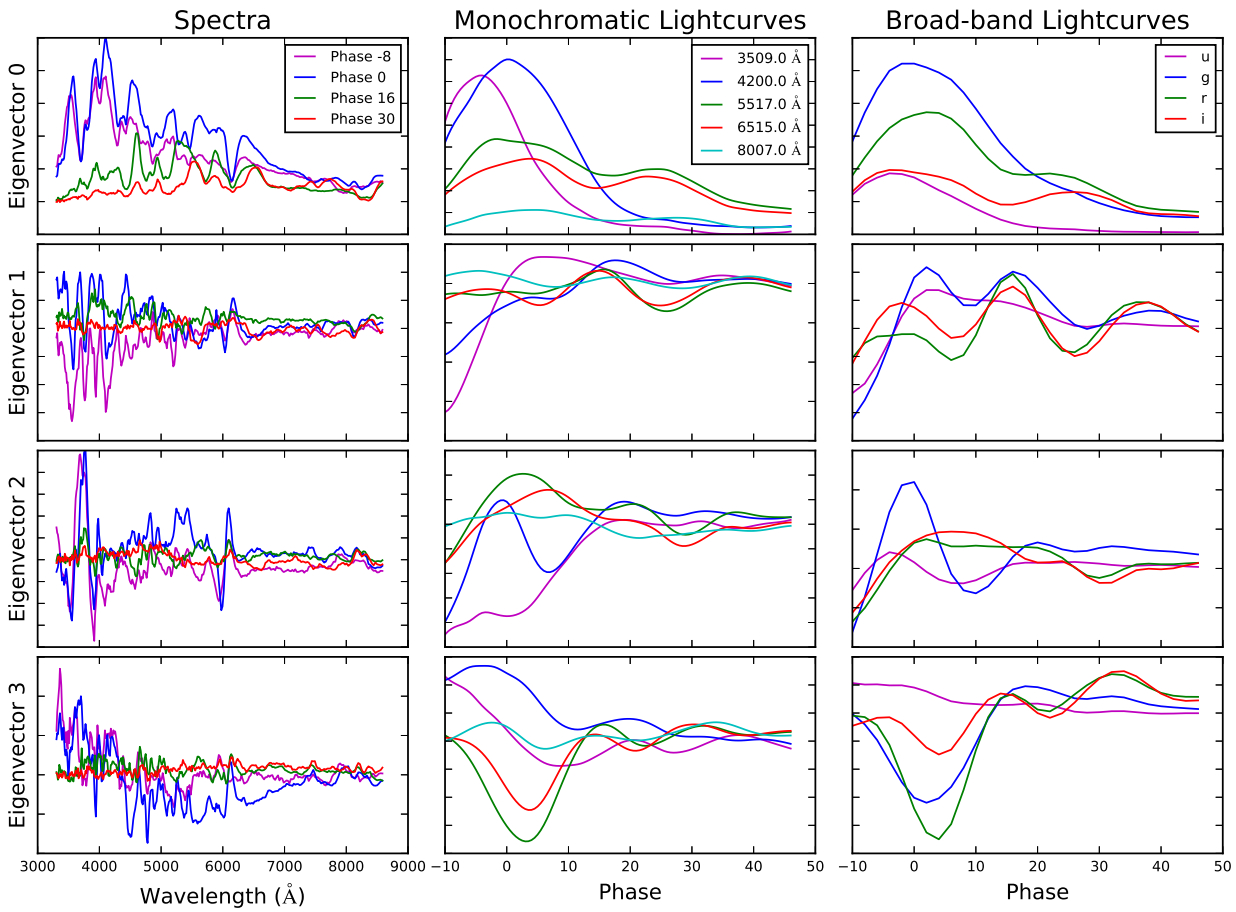

Figure 13.4: Demonstrative spectra and lightcurves for the first four components of the fifteen component model

near-UV near maximum light. Moving further to the fifteen component model, almost all of these are gone and the residuals are at or below the scale of the error in the data.

13.1 Comparison with the SALT2 model components

Comparing the models presented here with the SALT2 model can give some context for understanding the new model components. The SALT2 auxiliary template is known to be highly correlated with the stretch of the supernova, so comparing the new model components with the SALT2 auxiliary template gives us a proxy for how correlated our model components are with stretch.

The most direct comparison is with the two-component model, which has the same number of degrees of freedom as the SALT2 model and so should capture about the same amount

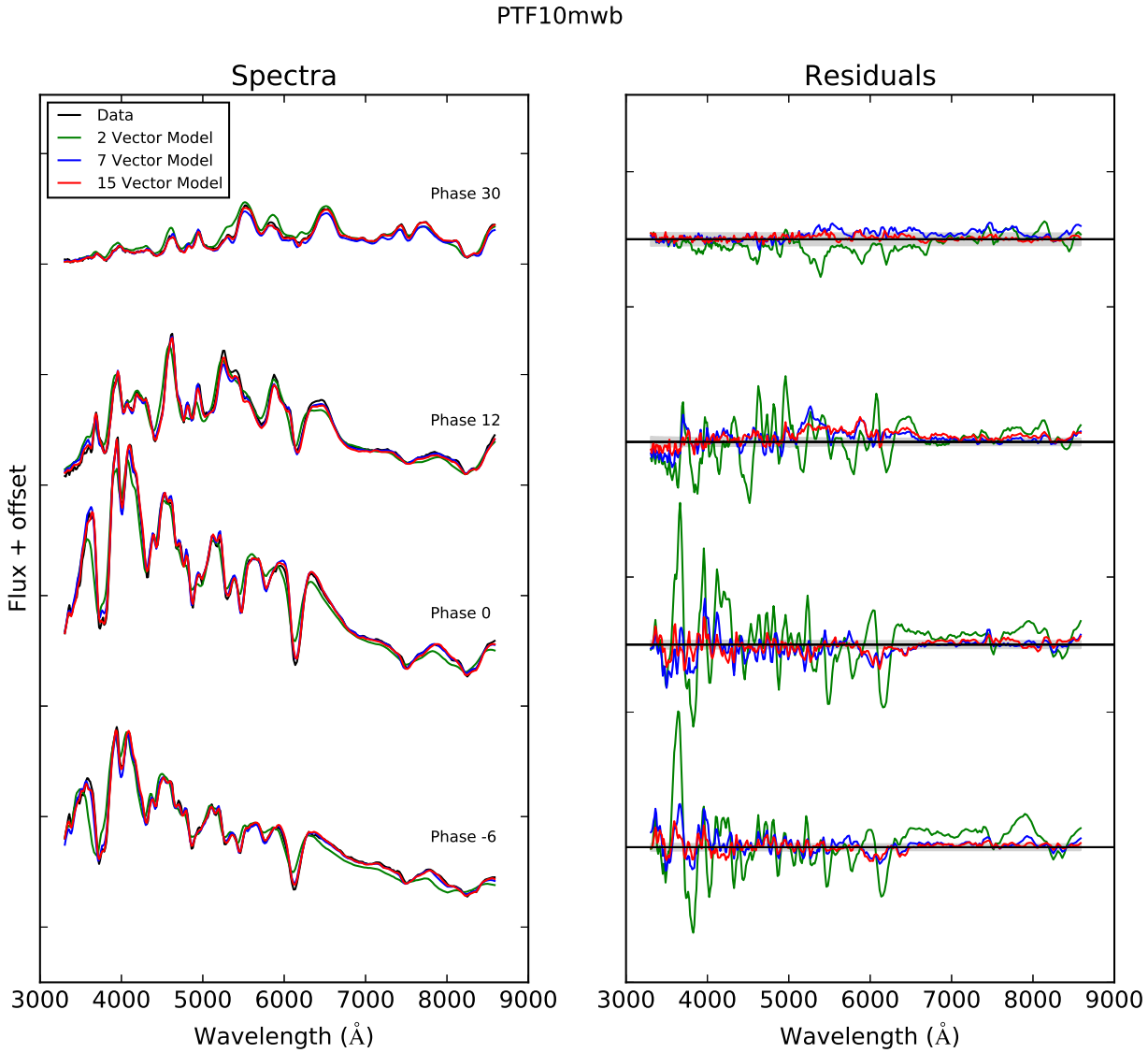

Figure 13.5: Model reconstructions of a sample supernova. The data and models for four demonstrative phases are shown in the left panel, while the right panel shows the residuals of the models, where the solid lines show the zero residual level and the gray shaded region shows the level of the supernova error.

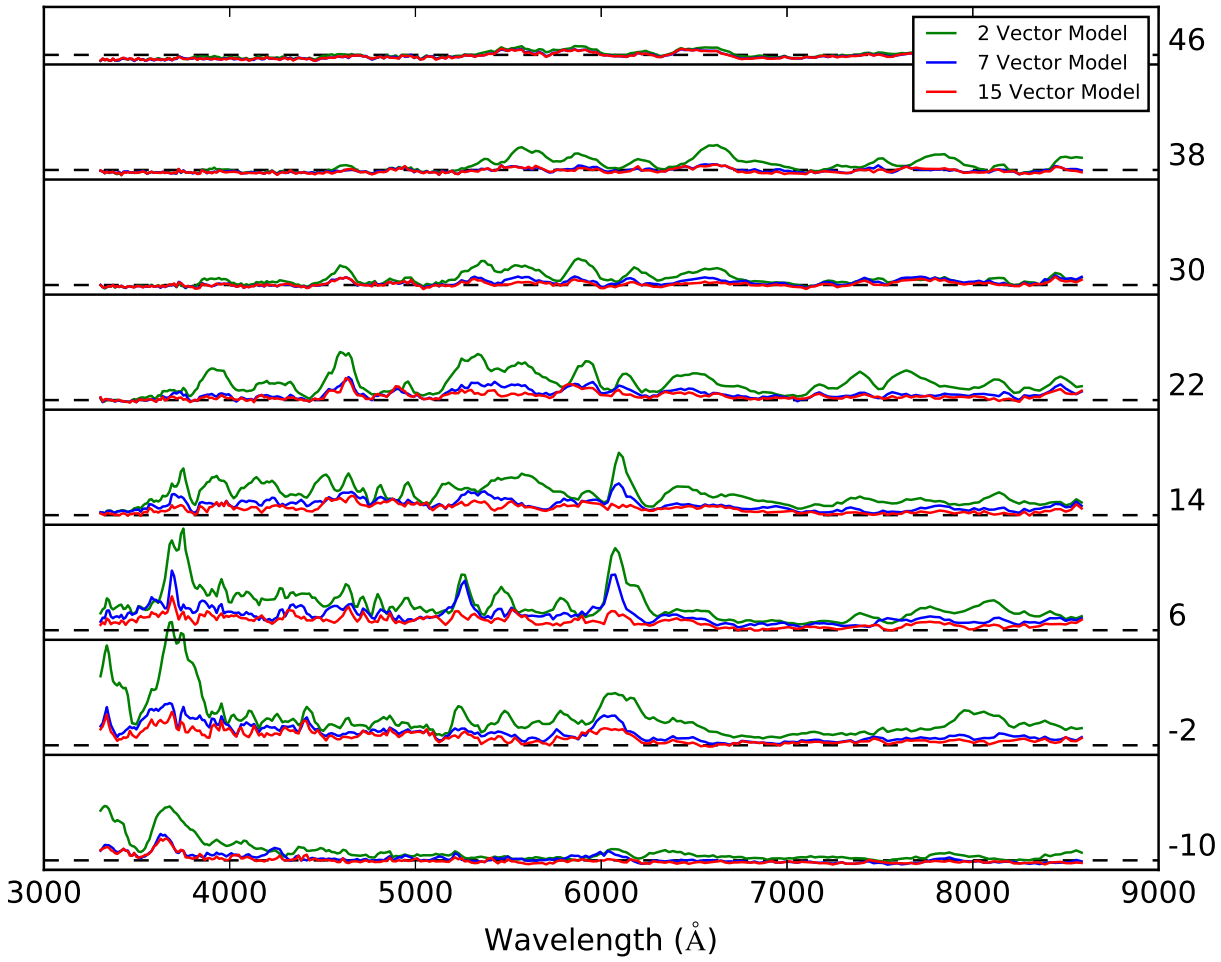

Figure 13.6: Standard deviation of the pulls (residuals divided by the error) between the models and the data from the hold-out set at demonstrative phases, indicated at the right of the figure. Solid and dashed lines show zero and plus-one in units of the residuals. The plus-one level shows where the residuals are on the same scale as the error in the data.

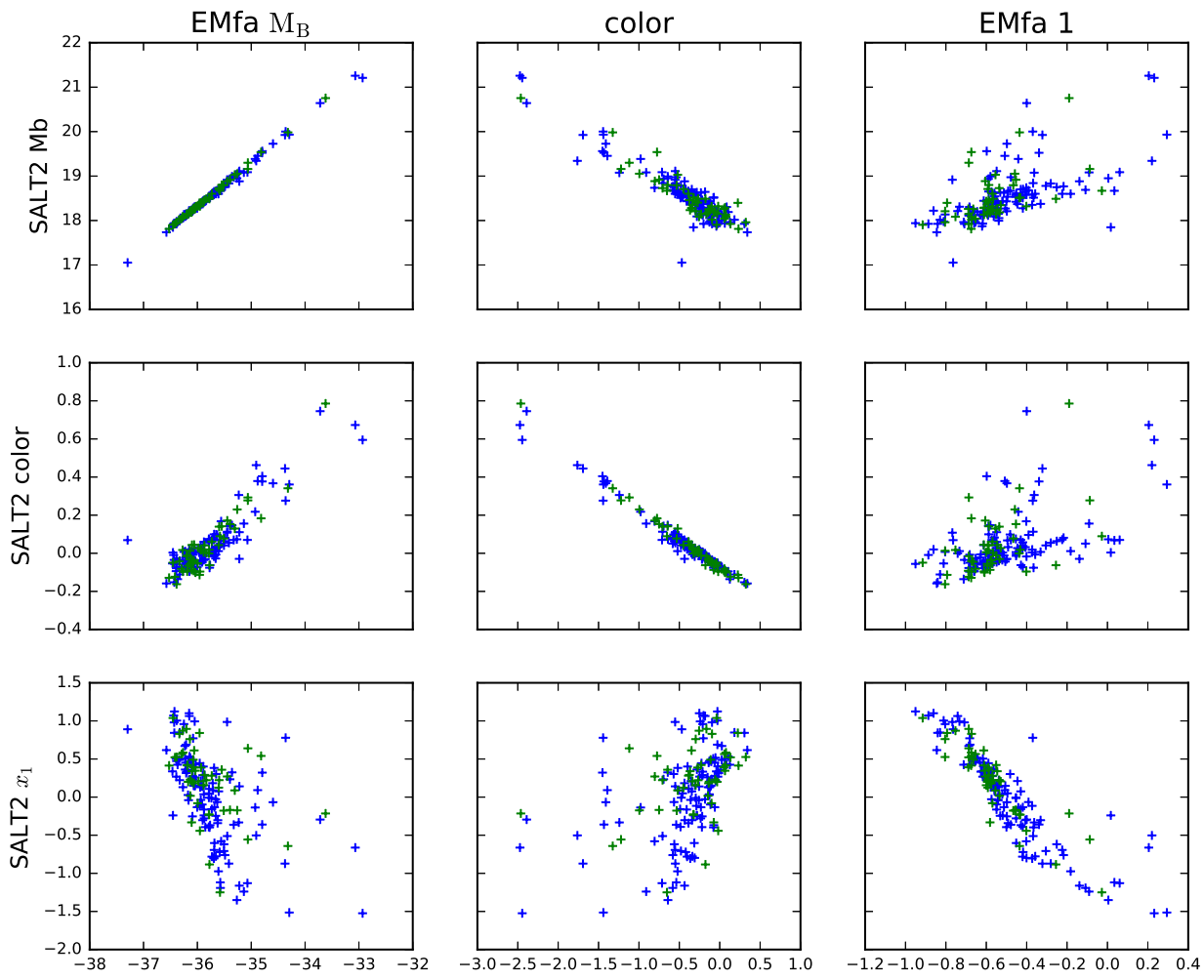

Figure 13.7: Model coefficients for the two component model versus the model coefficients for the SALT2 model. Blue points show training set supernovae and green points show test set supernovae.

of supernova diversity. The supernova fit coefficients are compared in Figure 13.7. While the second components of each (Eigenvector 1) do not appear very similar, it is clear that the coefficients are highly correlated (middle right panel of the figure). Thus, while they may be capturing different behavior, the two models find roughly the same way to discriminate among the supernovae. Some difference may be attributable to the fact that the SALT2 x_1 and c components are intended to be uncorrelated, while this is not the case with the models shown here. Comparing the seven and fifteen component models (not shown) finds that each also has one or two components that have a strong correlation with the SALT2 x_1 coefficient, though not so clearly as in the two-component model.

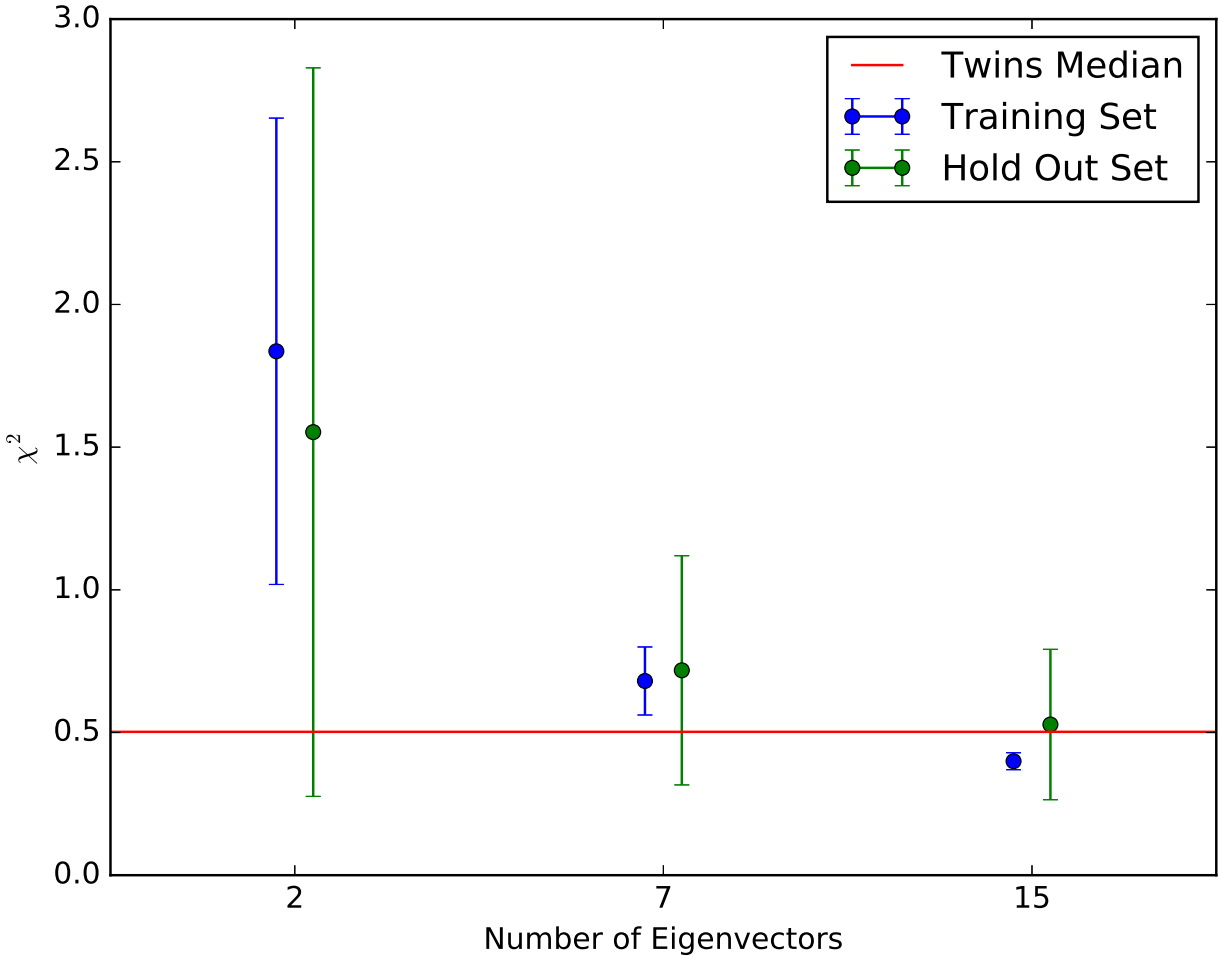

Figure 13.8: Reduced χ^2 of the three models in comparison to the χ^2 of the best supernova twins. The blue points correspond to the median result for the training set and the green points correspond to the median of the hold-out set supernovae, while the error bars show the error on the median. The red line is the median χ^2 between the best supernova twins.

13.2 Comparisons of the Model Reconstructions of the Data

Figure 13.8 shows the median χ^2 values for the two, seven, and fifteen component models. The size of the error is determined by the sample size and is also driven by outliers, which can be seen in the histograms in Figure 13.9. The values for the training and hold-out sets are consistent with the results found in the cross-validation process. The two component model yields χ^2 values consistent with the results seen in cross-validation; the value is somewhat lower, but the error on the median is large.

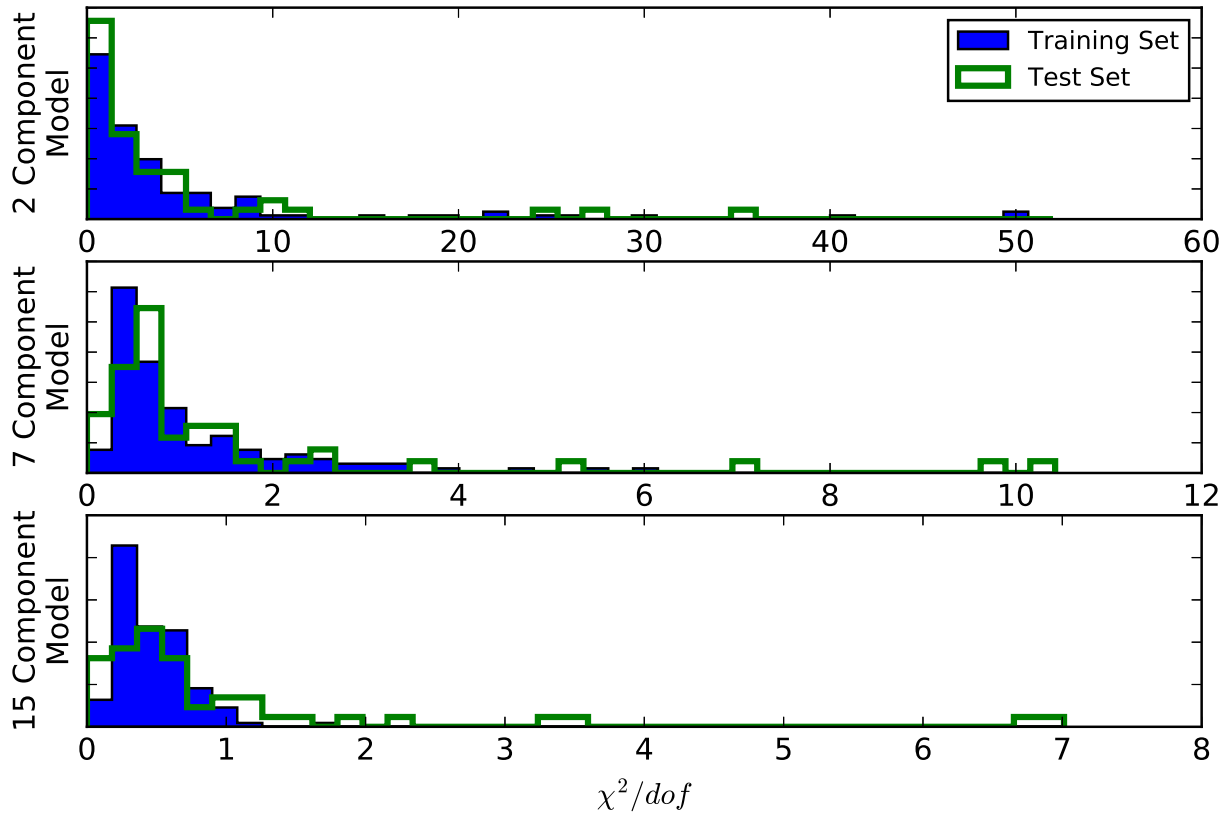

Figure 13.9: The three panels show histograms of the reduced χ^2 of the supernovae for the three models in comparison. The histograms are normalized to more easily compare the training and hold-out set results.

Examination of the supernovae from the hold-out set that still have somewhat large χ^2 values in Figure 13.9 shows that most are peculiar supernovae. It may be that from the small number of 1991T-like objects in the training set, we do not have enough information to fully capture the behavior for this type of supernova. One outlier is PTF09dnp, an otherwise normal object that is a rare case of a supernova with very narrow absorption lines where the CaII H&K and Si lines can be separately resolved. It can be seen from the residual between the model and supernova spectra that this feature is driving the relatively large χ^2, which is not unreasonable since this is the only supernova in our data set to have this feature. The quantity of such outlier χ^2 values gives an indication of the scale at which we have captured the full diversity of Type Ia supernova spectral time series.

These values can be compared with the analysis of 'twin' supernovae from Fakhouri et al. (2015). In this analysis, supernovae were identified in the SNfactory data set that, after normalizing and dereddening, had very similar spectral time series. The concept of the analysis was that such supernovae would be perfect standard candles requiring no magnitude

correction. These supernovae were called 'twins,' and it was found that there was a dispersion of 0.072 mag between the B-band magnitudes of twins, which is the best that has been found for Type Ia supernovae.

Here, we took Gaussian Process predictions for the best matches from the supernova twins analysis and calculated the χ^2 values, showing how closely the spectral time series for a pair of twins match each other (after normalizing and dereddening the spectra, as in the twins analysis). The median value for these is shown in Figure 13.8, where it is at about the level of the median χ^2 for the fifteen vector model fit to the hold-out set.

This shows that at fifteen components, our model is capturing the same amount of supernova behavior as the best twin supernovae match each other, which implies that with the fifteen component model we should be able to achieve the same level of magnitude standardization as with twins. This could be done in two ways: the simple way would be to fit supernovae with the model and identify ones that are very close to one another in the space of model coefficients, effectively finding twins through the model instead of looking directly at the spectra. As with twins, one would then use the 'model twins' as perfect standard candles. The other possibility is to take a more holistic approach and find a mapping from the space of model coefficients onto the possible range of B-band magnitudes. If the coefficients are a unique identifier of the supernova, then there must be some holomorphic map from them to the magnitudes. This should achieve the same level of magnitude dispersion that the twins method finds.

13.3 Dispersion in Standardized Magnitudes with the new model

The seven component model is used to demonstrate the improvements made in linear standardization of magnitudes. As described above, we find a standardized magnitude by calculating the linear correlations between the model coefficients and the B-band magnitude. The results are shown in Table 13.1. As discussed in 12.4, the corrections are calculated using the final hold-out set, since in a cosmological analysis the corrections would be calculated on the data set used. For comparison, we also show the dispersion found when calculating the corrections using the training set, which gives us a larger number of supernovae to fit.

As in the 'Union' compilation (Amanullah et al., 2010), we remove supernovae with corrected magnitude pulls (the residual divided by the error) of more than 3σ. This removes one supernova from the training set and one supernova from the hold-out set. Without removing the outlier supernovae, the results fit to the training set give a higher dispersion when applied to the hold-out set. However, without the outlier supernovae, the dispersion levels are within 1σ of one another. The results actually fit to the hold-out set are also consistent with the results fit to the training set. The dispersion found matches what is seen in the χ^2 of the seven component model: while it is lower than the SALT2 values, it is not quite at a level matching 'twin' supernovae. Again, moving from linear standardization to

	Intrinsic Dispersion	
Data Set	Corrections fit to Training Set	Corrections fit to Hold-out Set
Training Set	$0.114^{+0.011}_{-0.009}$	—
Hold-out Set	$0.146^{+0.023}_{-0.016}$	$0.130^{+0.021}_{-0.015}$
Training Set with 3σ cuts	$0.108^{+0.011}_{-0.009}$	—
Hold-out Set with 3σ cuts	$0.115^{+0.018}_{-0.013}$	$0.108^{+0.017}_{-0.012}$

Table 13.1: Final Results for the Intrinsic Dispersion in B-band Standardized Magnitudes using the Seven Component Model

a more sophisticated method may be necessary for maximizing the possible improvement in standardized magnitudes.

13.4 Final Discussion of the Models

In sum, we have calculated three models for Type Ia supernovae with a range of applications. The two component model serves mainly as a tool for comparison with other currently used two-component empirical models for supernova spectral time series. We find that while the components of the model appear quite different from the components in the SALT2 model, they describe similar effects in supernova behavior, which can be seen from the correlation in the model coefficients.

The seven component model has been shown to describe much more of supernova variability and is also the model best among the models presented here for making linear corrections to the B-band magnitude and minimizing dispersion among supernovae. The number of model components is small enough that it should be possible to fit the model to supernova data with only photometric lightcurves, though testing of this is left for future study.

Lastly, the fifteen component model has been shown to match spectral time series for all the supernovae in our data set on the same level that the best 'twin' supernovae match each other. Given the number of components, it will probably be difficult to constrain such a model without at least one spectrum from the fit supernova. However, if the data can sufficiently constrain the model, it should also be possible to achieve magnitude standardization at the level of supernova twins, around 0.07 mags. A lightcurve fitter that could match supernova spectra so well would also remove the systematic errors discussed in the Part I of this thesis, which may not be the case with the seven component model. This would be possible if the model parameters can be fit accurately enough to minimize differences between the spectral time series model and the true supernova spectrum. If this could be done, the same standardized magnitude would be fit independent of the filters used, unlike what was found using current two-parameter lightcurve fitters. The fifteen component model will also be particularly useful for those trying to accurately measure spectral features and for theoretical modelers who are trying to replicate the full range of Type Ia supernova behavior.

Chapter 14

Future Steps

There are two major next steps to be taken in followup to this analysis. These will determine how this model can best be used to improve the use of Type Ia supernovae for cosmology, since much of the utility of the model will depend on how it can be applied to new supernova data and how well we can interpret the model parameters in the context of supernova analyses over a wide range of redshifts and with varying conditions of data.

The first priority will be testing the Type Ia supernova model on photometric supernova data. While the model describes more of Type Ia supernova behavior than current empirical models, it remains to be seen whether this greater amount of information can be detected in lower resolution photometric lightcurves. Given that high-redshift supernovae only have observations at a few phases in two or three photometric bands, there is reason to be concerned that it will not be possible to constrain the parameters of the high-precision fifteen component model. This can be tested with a large set of photometric data, such as is available from SDSS or SNLS. The primary question here will be whether the magnitudes can be standardized with greater precision even without spectral information, meaning that there will be less scatter and potential for bias on the Hubble diagram, and hence lower uncertainty in determining cosmological parameters. It will also be possible to test whether the addition of spectroscopy improves the model fit enough that it would justify the additional time investment necessary to obtain it.

Another big question will be whether issues seen in the current lightcurve fitters, such as filter-dependent bias (as addressed in the first part of this thesis) and bias depending on the host galaxy, persist with this supernova model. The models can also be used to prevent data set contamination. For example, a model parameter that has significantly different distributions for Ia and non-Ia supernovae can be used to remove non-Ia supernovae from the data set. Answering questions such as these will be crucial for current and future surveys such as the Dark Energy Survey and the Large Synoptic Sky Survey observing program, in which very large numbers of supernovae will be observed with spectroscopic followup of only a small percentage of the objects. Meanwhile, the fifteen component model will be useful for fitting supernova spectrophotometry observed with the planned Wide Field Infrared Survey Telescope, which will include an integral field unit spectrograph.

The second priority will be to test more advanced magnitude standardization techniques. As discussed briefly above, with a more complex model it should be possible to make further improvements in supernova standardization. While the model is free to fit supernovae with identical spectral features and lightcurve shapes but different absolute magnitudes, it is difficult to find a physically valid description of supernovae that could explain such behavior: the spectra and lightcurve shapes are a window onto the explosion physics and it would not seem possible that they could be identical while corresponding to supernovae with different intrinsic brightnesses. Having the same spectra and lightcurve shape but different normalizations would require some very contrived scenario, or possibly the discovery of heretofore unknown grey dust between the supernova and the observer. Thus, given a perfect match to the supernova spectrum and lightcurve, which we are approaching with the 15 component model, some mapping must be possible which allows us to standardize supernova magnitudes at a level equivalent to finding twin supernovae. Of course, this is limited by other sources of magnitude dispersion, in particular measurement errors such as data calibration.

This improved standardization could be done a number of ways. The linear standardization is essentially fitting a hyper-plane to the space of model coefficients and magnitudes. A simple step up from linear standardization would be to fit some non-linear hyper-surface between the coefficients and magnitudes. Alternatively, magnitude standardization is a good occasion for using machine learning. Machine learning can find a non-analytic function that maps between model coefficients and magnitudes in a relatively agnostic manner.

Lastly, there are supplementary studies that will be important from the standpoint of better understanding supernovae, such as further interpreting the components of the new Type Ia supernova models. Correlations between the model components and other supernova features such as line strengths and velocities or host properties will help to explain the diversity of Type Ia supernovae. Possible clustering in the model parameters would help determine the question of whether there are discrete supernova populations or different types of progenitor systems.

The models presented here will be released in an upcoming paper and will also be added to the supernova lightcurve fitting Python package *sncosmo*, meaning that those wishing to apply a model to their data will not have to develop their own implementation. The coefficient distributions of the supernovae will also be released with the paper so that the model can be used to create simulated data sets.

Supernova cosmology has relied on two-component models for almost twenty years. For most of this time, models were limited by the lack of sufficient data to make any improvements upon this. We now have the capacity to move past this barrier, with advances possible both for our understanding of Type Ia supernovae and our capacity to measure dark energy.

Appendix A

Full EMFA Results

The following Figures A.1 through A.3 show demonstrative plots for all of the components of the seven and fifteen component models presented in Chapter 13.

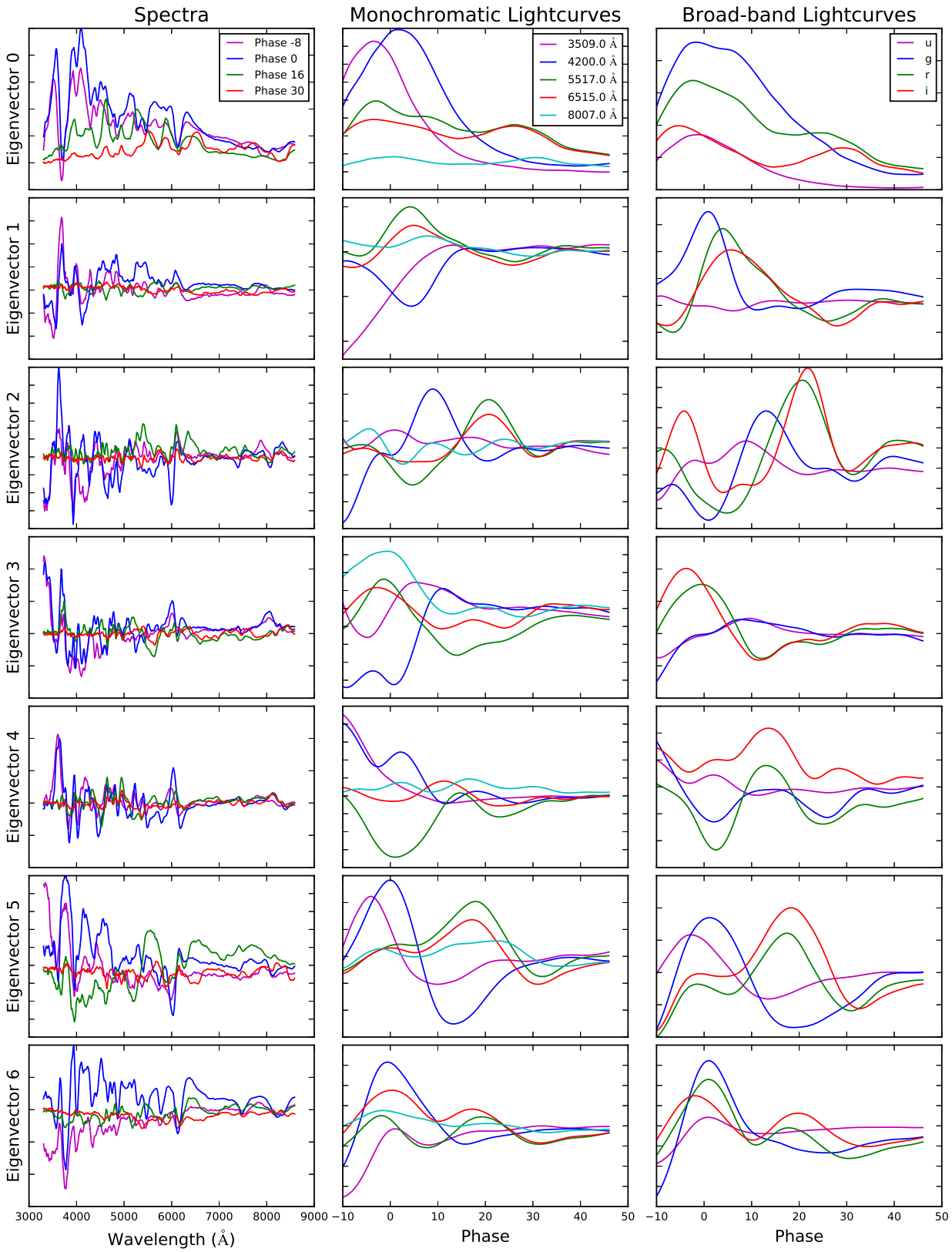

Figure A.1: Demonstrative spectra and lightcurves for all eigenvectors of the seven component model.

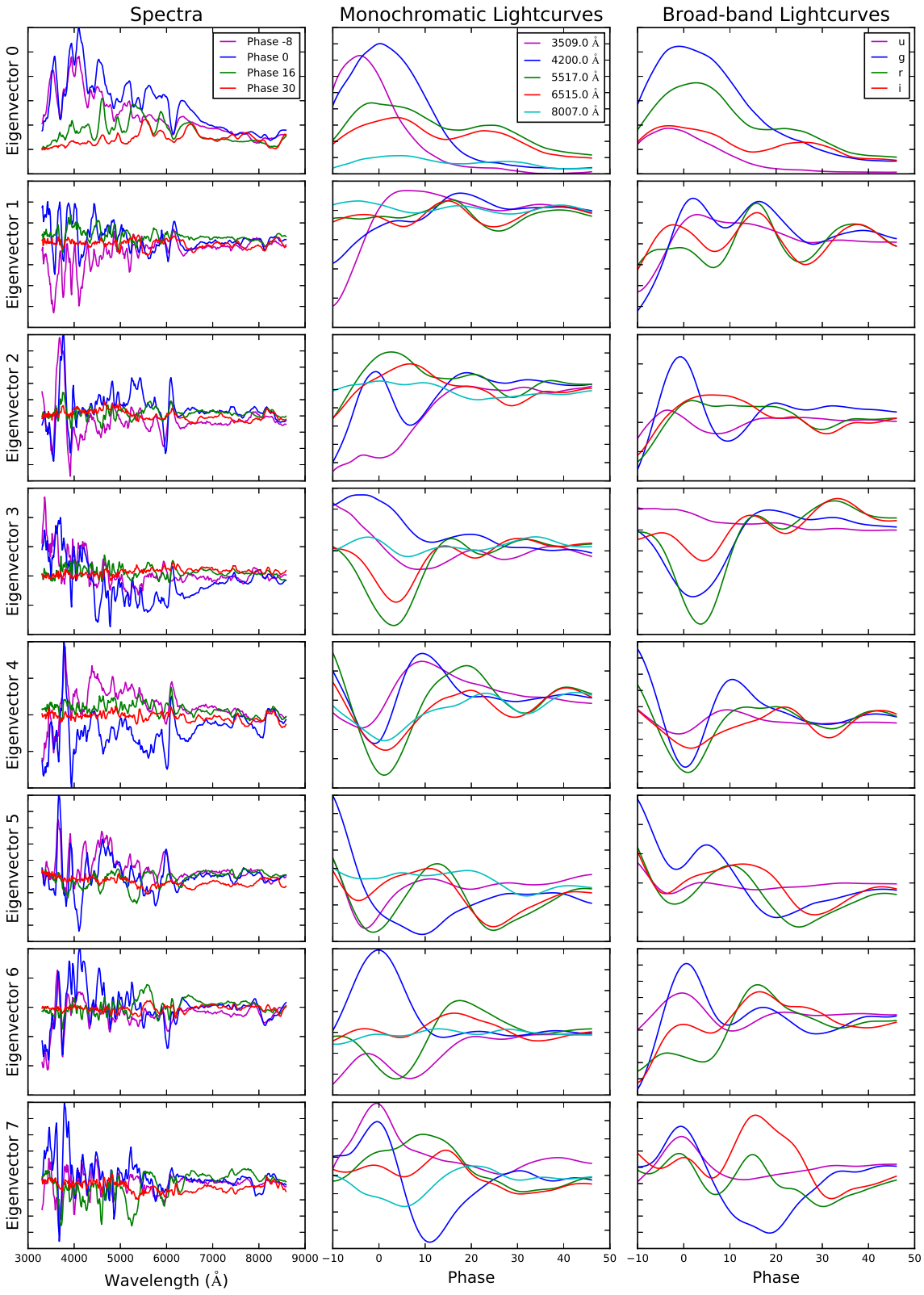

Figure A.2: Demonstrative spectra and lightcurves for the first eight eigenvectors of the fifteen component model.

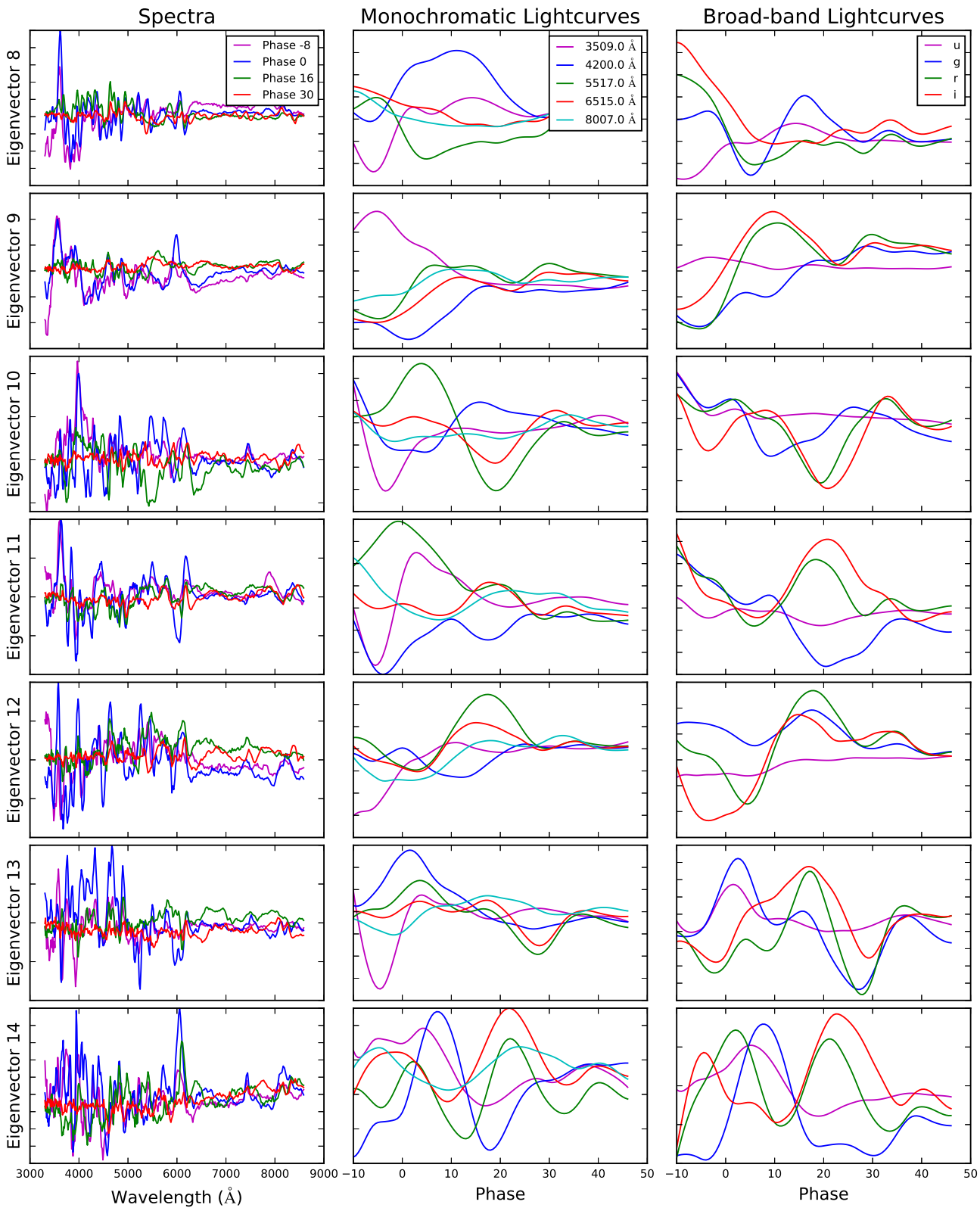

Figure A.3: Demonstrative spectra and lightcurves for the last seven eigenvectors of the fifteen component model.

Appendix B

Supernova Models with no Outlier Supernovae

As discussed in Section 12.2, six supernovae in the training set are identified as outliers in model coefficient space. These supernovae are not included in calculating the magnitude standardization of the supernovae, but are not taken out of the EMFA training set. After being identified as coefficient outliers, these supernovae were also spectroscopically identified by eye as being peculiar supernovae. Three are 1991T-like supernovae, identified by the lack of SiII and CaII H&K absorption before maximum (Filippenko et al., 1992a) and three are 1991bg-like supernovae, mainly identified by broad TiII absorption around 4200Å (Filippenko et al., 1992b).

These supernovae would generally be excluded from cosmological analyses because and they are not expected to be well fit by current lightcurve fitters like SALT2 (Amanullah et al., 2010). Since these are not currently useful for cosmology, there is some motivation to exclude them from the entire analysis: the model might be able to describe 'normal' supernovae better or with fewer components if it was not required to also fit these peculiar supernovae. Testing showed that removing these supernovae does change the individual components of the models, but does not lead to a significant improvement in the χ^2 of the model or the resulting dispersion in the standardized magnitudes, except in the two component model, where it did improve the χ^2 metric. K-fold cross validation showed that the favored model for standardizing magnitudes and the best spectral model were at six and fourteen components respectively, as opposed to seven and fifteen when the peculiar supernovae are included. Thus, there is the benefit of needing one less component to describe the more normal supernovae. We present the resulting models using two, six, and fourteen components in Figures B.1 through B.4.

Ultimately it was decided that excluding the peculiar supernovae did not provide enough benefits to outweigh the somewhat subjective decision to exclude them from the model training set. There is a continuum in the behavior that allows these supernovae to be identified as 1991T or 1991bg-like, and supernovae that share some of these features are seen in our data set that are not outliers in the model coefficient space. This means that to have

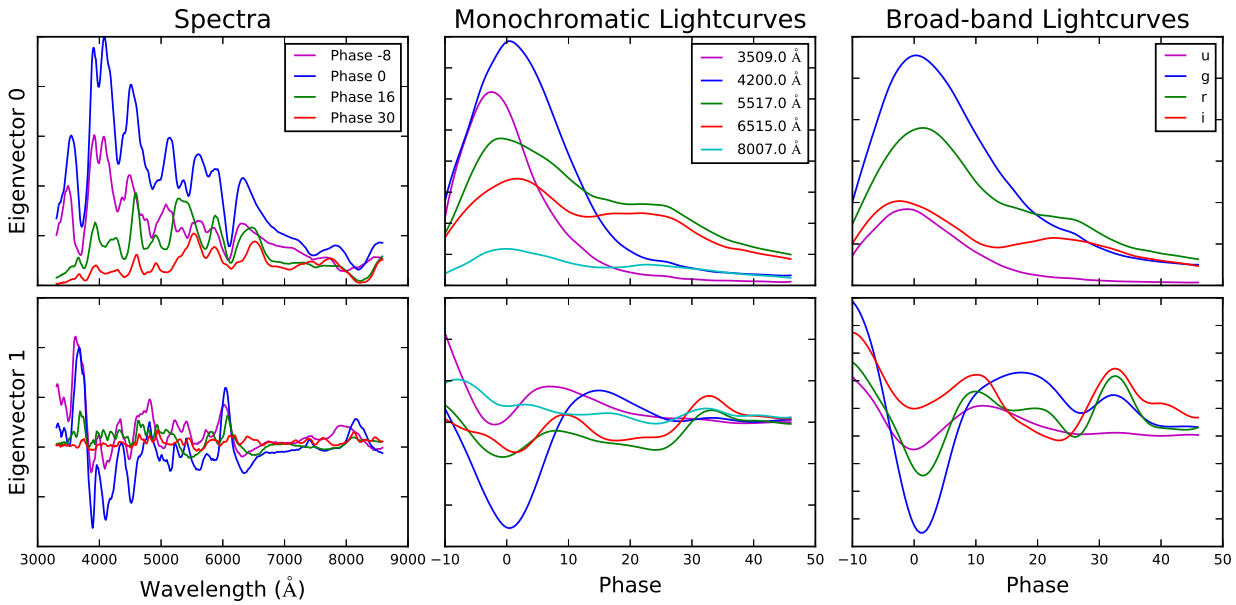

Figure B.1: Demonstrative spectra and lightcurves for the eigenvectors of the two component model trained with no outlier supernovae

a model trained only on 'normal' supernovae, all of the 1991T or 1991bg-like supernovae would have to be identified by eye in order to remove them, which would be dependent on some chosen limits and would probably be biased by the phase coverage of the supernovae, since 1991T-like supernovae appear normal at later phases. On balance, it is preferable to avoid subjective choices in training the model, with the benefit that since the model has been trained on peculiar supernovae it will be easier to identify these supernovae when the model is applied to new supernova set, and to remove them if necessary for cosmological analyses.

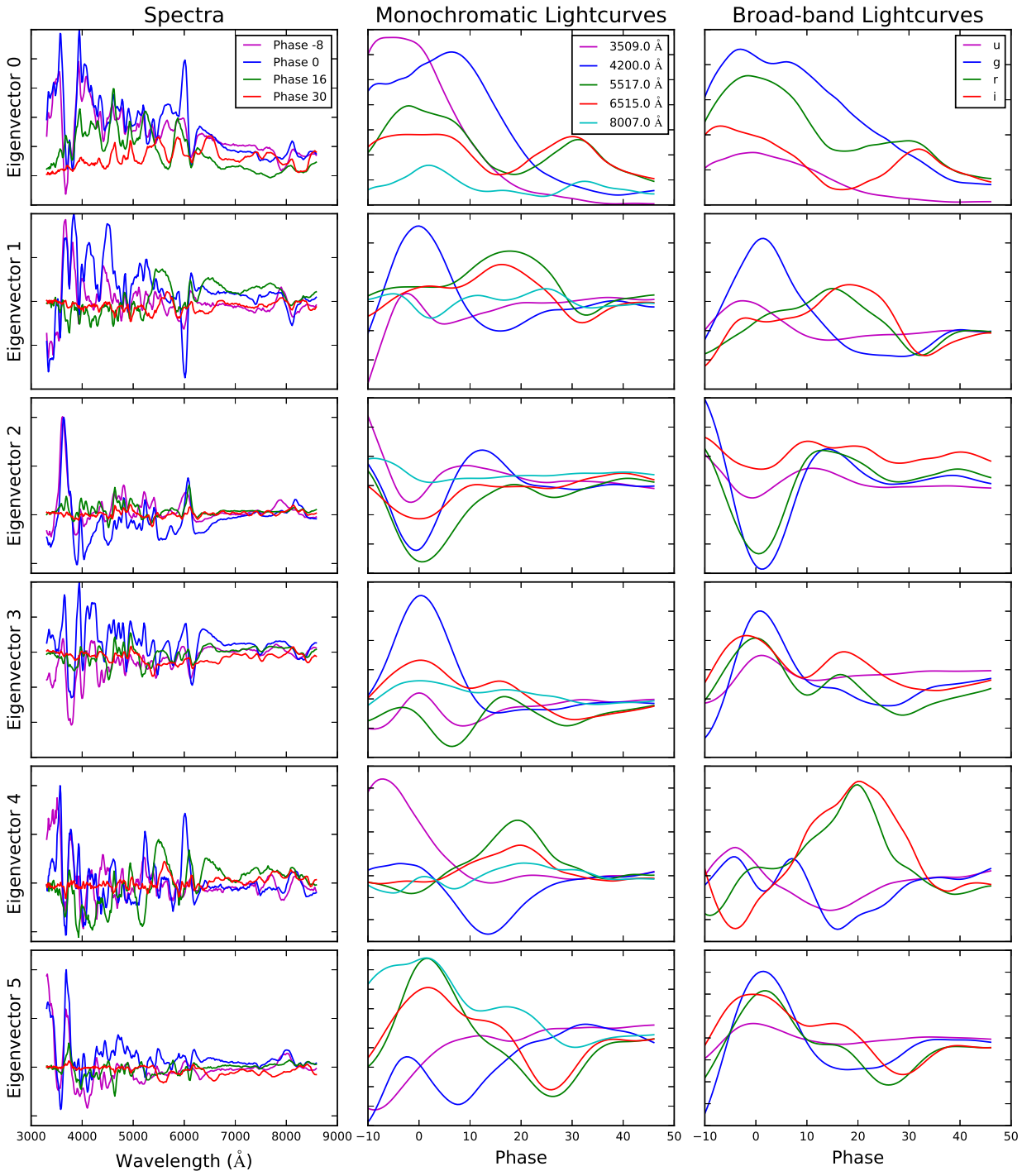

Figure B.2: Demonstrative spectra and lightcurves for all eigenvectors of the six component model trained with no outlier supernovae

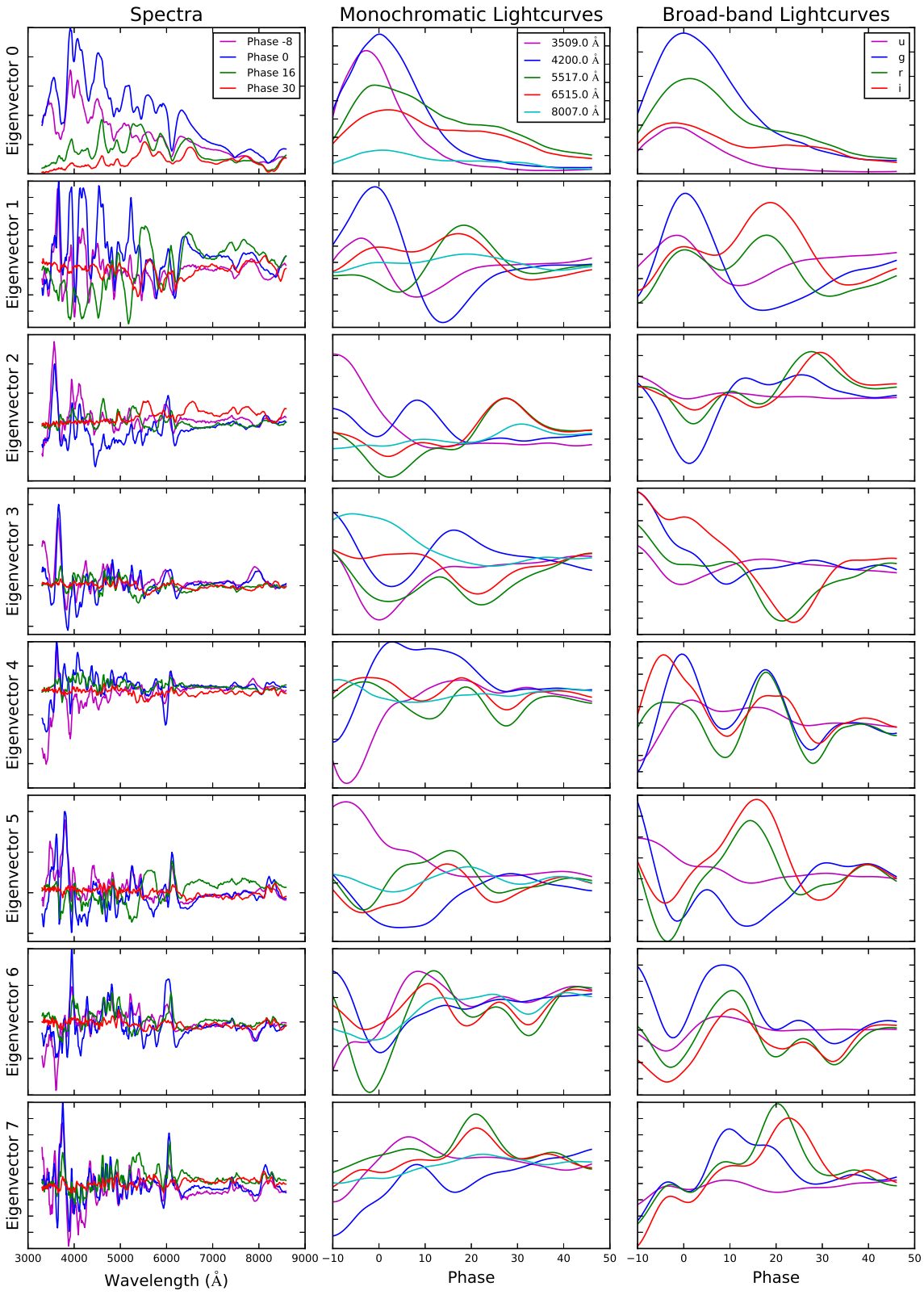

Figure B.3: Demonstrative spectra and lightcurves for the first eight eigenvectors of the fourteen component model trained with no outlier supernovae.

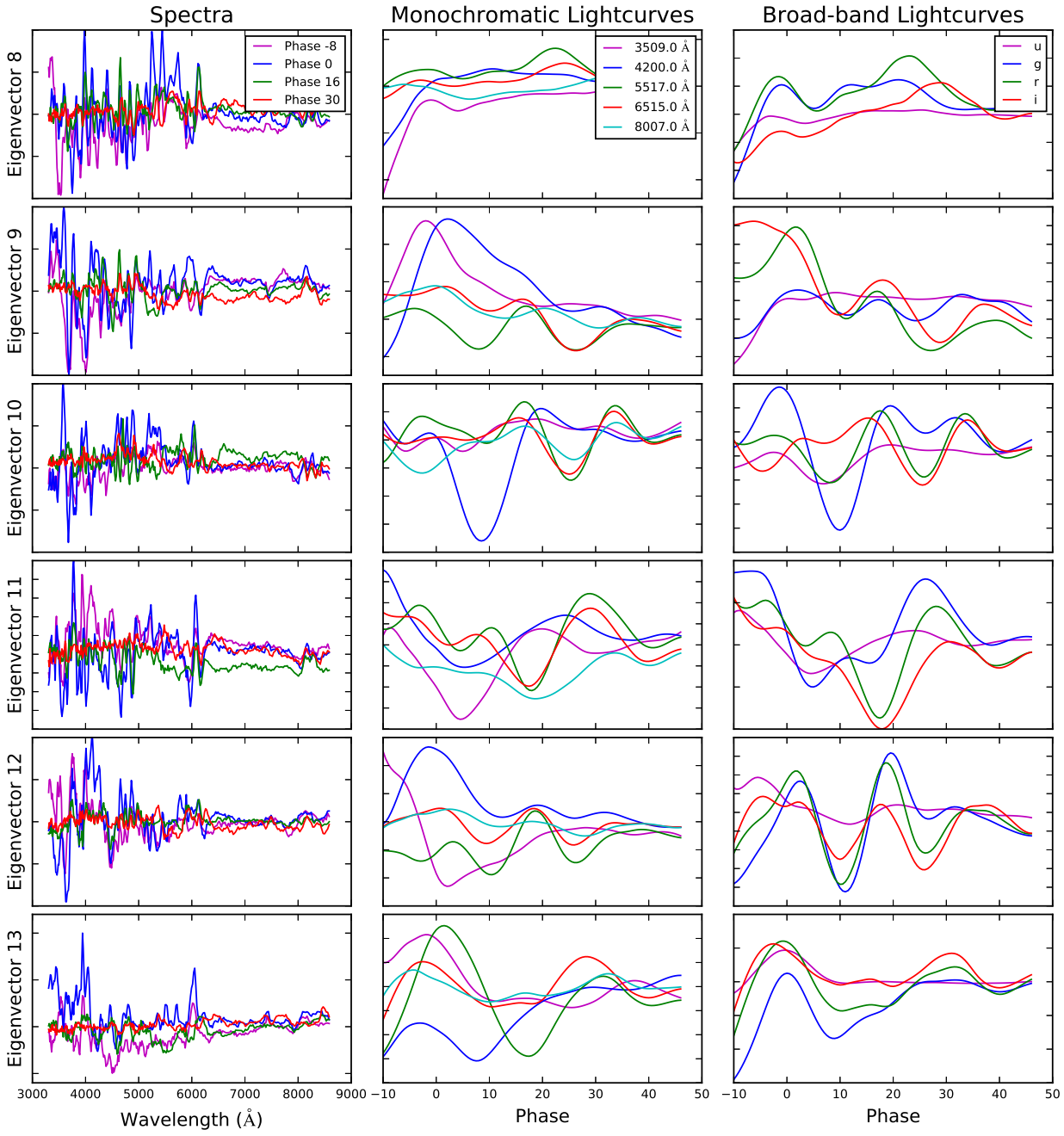

Figure B.4: Demonstrative spectra and lightcurves for the last six eigenvectors of the fourteen component model trained with no outlier supernovae.

Bibliography

Aldering, G. et al. (2002). "Survey and Other Telescope Technologies and Discoveries". In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series* 4836, pp. 61–72.

Aldering, G. et al. (Oct. 2006). "Nearby Supernova Factory Observations of SN 2005gj: Another Type Ia Supernova in a Massive Circumstellar Envelope". In: *ApJ* 650, pp. 510–527. DOI: 10.1086/507020. eprint: astro-ph/0606499.

Amanullah, R. et al. (June 2010). "Spectra and Hubble Space Telescope Light Curves of Six Type Ia Supernovae at 0.511 < z < 1.12 and the Union2 Compilation". In: *ApJ* 716, pp. 712–738. DOI: 10.1088/0004-637X/716/1/712. arXiv:1004.1711 [astro-ph.CO].

Astier, P. et al. (Feb. 2006). "The Supernova Legacy Survey: measurement of Ω_M, Ω_Λ and w from the first year data set". In: *A&A* 447, pp. 31–48. DOI: 10.1051/0004-6361:20054185. eprint: astro-ph/0510447.

Baade, W. and F. Zwicky (July 1934). "Remarks on Super-Novae and Cosmic Rays". In: *Physical Review* 46, pp. 76–77. DOI: 10.1103/PhysRev.46.76.2.

Bacon, R. et al. (Oct. 1995). "3D spectrography at high spatial resolution. I. Concept and realization of the integral field spectrograph TIGER." In: *A&AS* 113, p. 347.

Bacon, R. et al. (2000). "Supermassive Black Hole Searches with 3-D Spectroscopy". In: *Imaging the Universe in Three Dimensions*. Ed. by W. van Breugel and J. Bland-Hawthorn. Vol. 195. Astronomical Society of the Pacific Conference Series, p. 173.

Bacon, R. et al. (Sept. 2001). "The SAURON project - I. The panoramic integral-field spectrograph". In: *MNRAS* 326, pp. 23–35. DOI: 10.1046/j.1365-8711.2001.04612.x. eprint: arXiv:astro-ph/0103451.

Bailey, S. (Sept. 2012). "Principal Component Analysis with Noisy and/or Missing Data". In: *PASP* 124, pp. 1015–1023. DOI: 10.1086/668105. arXiv:1208.4122 [astro-ph.IM].

Bailey, S. et al. (June 2009). "Using spectral flux ratios to standardize SN Ia luminosities". In: *A&A* 500, pp. L17–L20. DOI: 10.1051/0004-6361/200911973. arXiv:0905.0340 [astro-ph.CO].

Bernstein, J. P. et al. (July 2012). "Supernova Simulations and Strategies for the Dark Energy Survey". In: *ApJ* 753, 152, p. 152. DOI: 10.1088/0004-637X/753/2/152. arXiv:1111.1969 [astro-ph.CO].

Betoule, M. et al. (Aug. 2014). "Improved cosmological constraints from a joint analysis of the SDSS-II and SNLS supernova samples". In: *A&A* 568, A22, A22. DOI: 10.1051/0004-6361/201423413. arXiv:1401.4064.

Bohlin, R. C. (June 2014). "Hubble Space Telescope CALSPEC Flux Standards: Sirius (and Vega)". In: *AJ* 147, 127, p. 127. DOI: 10.1088/0004-6256/147/6/127.

Bongard, S. et al. (Nov. 2011). "3D deconvolution of hyper-spectral astronomical data". In: *MNRAS* 418, pp. 258–270. DOI: 10.1111/j.1365-2966.2011.19480.x. arXiv:1107.4049 [astro-ph.IM].

Burns, C. R. et al. (Jan. 2011). "The Carnegie Supernova Project: Light-curve Fitting with SNooPy". In: *AJ* 141, 19, p. 19. DOI: 10.1088/0004-6256/141/1/19. arXiv:1010.4040 [astro-ph.CO].

Buton, C. et al. (Jan. 2013). "Atmospheric extinction properties above Mauna Kea from the Nearby SuperNova Factory spectro-photometric data set". In: *A&A* 549, A8, A8. DOI: 10.1051/0004-6361/201219834. arXiv:1210.2619 [astro-ph.IM].

Cardelli, J. A., G. C. Clayton, and J. S. Mathis (Oct. 1989). "The relationship between infrared, optical, and ultraviolet extinction". In: *ApJ* 345, pp. 245–256. DOI: 10.1086/167900.

Childress, M. et al. (June 2013). "Host Galaxies of Type Ia Supernovae from the Nearby Supernova Factory". In: *ApJ* 770, 107, p. 107. DOI: 10.1088/0004-637X/770/2/107. arXiv:1304.4719 [astro-ph.CO].

Chotard, N. et al. (May 2011). "The reddening law of type Ia supernovae: separating intrinsic variability from dust using equivalent widths". In: *A&A* 529, L4, p. L4. DOI: 10.1051/0004-6361/201116723. arXiv:1103.5300 [astro-ph.CO].

Conley, A. et al. (July 2008). "SiFTO: An Empirical Method for Fitting SN Ia Light Curves". In: *ApJ* 681, pp. 482–498. DOI: 10.1086/588518. arXiv:0803.3441.

Elias, J. H. et al. (Sept. 1985). "Type I supernovae in the infrared and their use as distance indicators". In: *ApJ* 296, pp. 379–389. DOI: 10.1086/163456.

Ellis, R. S. et al. (Feb. 2008). "Verifying the Cosmological Utility of Type Ia Supernovae: Implications of a Dispersion in the Ultraviolet Spectra". In: *ApJ* 674, pp. 51–69. DOI: 10.1086/524981. arXiv:0710.3896.

Fakhouri, H. K. et al. (Dec. 2015). "Improving Cosmological Distance Measurements Using Twin Type Ia Supernovae". In: *ApJ* 815, 58, p. 58. DOI: 10.1088/0004-637X/815/1/58. arXiv:1511.01102.

Filippenko, A. V. et al. (Jan. 1992a). "The peculiar Type IA SN 1991T - Detonation of a white dwarf?" In: *ApJ* 384, pp. L15–L18. DOI: 10.1086/186252.

Filippenko, A. V. et al. (Oct. 1992b). "The subluminous, spectroscopically peculiar type IA supernova 1991bg in the elliptical galaxy NGC 4374". In: *AJ* 104, pp. 1543–1556. DOI: 10.1086/116339.

Fitzpatrick, E. L. and D. Massa (July 2007). "An Analysis of the Shapes of Interstellar Extinction Curves. V. The IR-through-UV Curve Morphology". In: *ApJ* 663, pp. 320–341. DOI: 10.1086/518158. arXiv:0705.0154.

Frieman, J. A. et al. (Jan. 2008). "The Sloan Digital Sky Survey-II Supernova Survey: Technical Summary". In: *AJ* 135, pp. 338–347. DOI: 10.1088/0004-6256/135/1/338. arXiv:0708.2749.

Gall, E. E. E. et al. (Mar. 2016). "Applying the expanding photosphere and standardized candle methods to Type II-Plateau supernovae at cosmologically significant redshifts: the distance to SN 2013eq". In: *ArXiv e-prints*. arXiv:1603.04730.

Ghahramani, Zoubin, Geoffrey E Hinton, et al. (1996). *The EM algorithm for mixtures of factor analyzers*. Tech. rep. Technical Report CRG-TR-96-1, University of Toronto.

Guy, J. et al. (2007). "SALT2: using distant supernovae to improve the use of Type Ia supernovae as distance indicators". In: *Astron.Astrophys.* 466, pp. 11–21. eprint: astro-ph/0701828. URL: http://arxiv.org/abs/astro-ph/0701828.

Guy, J. et al. (Nov. 2010). "The Supernova Legacy Survey 3-year sample: Type Ia supernovae photometric distances and cosmological constraints". In: *A&A* 523, A7, A7. DOI: 10.1051/0004-6361/201014468. arXiv:1010.4743 [astro-ph.CO].

Hamuy, M. et al. (July 1992). "Southern spectrophotometric standards." In: *PASP* 104, pp. 533–552. DOI: 10.1086/133028.

Hamuy, M. et al. (June 1994). "Southern spectrophotometric standards, 2". In: *PASP* 106, pp. 566–589. DOI: 10.1086/133417.

Hillebrandt, W. and J. C. Niemeyer (2000). "Type IA Supernova Explosion Models". In: *ARA&A* 38, pp. 191–230. DOI: 10.1146/annurev.astro.38.1.191. eprint: astro-ph/0006305.

Holtzman, J. A. et al. (Dec. 2008). "The Sloan Digital Sky Survey-II: Photometry and Supernova IA Light Curves from the 2005 Data". In: *AJ* 136, pp. 2306–2320. DOI: 10.1088/0004-6256/136/6/2306. arXiv:0908.4277 [astro-ph.CO].

Howell, D. A. et al. (Jan. 2009). "The Effect of Progenitor Age and Metallicity on Luminosity and ^{56}Ni Yield in Type Ia Supernovae". In: *ApJ* 691, pp. 661–671. DOI: 10.1088/0004-637X/691/1/661. arXiv:0810.0031.

Hsiao, E. Y. et al. (2007). "K-corrections and spectral templates of Type Ia supernovae". In: *Astrophys.J.* 663, pp. 1187–1200. eprint: astro-ph/0703529. URL: http://arxiv.org/abs/astro-ph/0703529.

Humason, M. L., N. U. Mayall, and A. R. Sandage (1956). "Redshifts and magnitudes of extragalactic nebulae." In: *AJ* 61, pp. 97–162. DOI: 10.1086/107297.

Jha, S., A. G. Riess, and R. P. Kirshner (Apr. 2007). "Improved Distances to Type Ia Supernovae with Multicolor Light-Curve Shapes: MLCS2k2". In: *ApJ* 659, pp. 122–148. DOI: 10.1086/512054. eprint: astro-ph/0612666.

Jha, Saurabh, Adam G. Riess, and Robert P. Kirshner (2007). "Improved Distances to Type Ia Supernovae with Multicolor Light Curve Shapes: MLCS2k2". In: *Astrophys.J.* 659, pp. 122–148. eprint: astro-ph/0612666. URL: http://arxiv.org/abs/astro-ph/0612666.

Kim, A., A. Goobar, and S. Perlmutter (Feb. 1996). "A Generalized K Correction for Type IA Supernovae: Comparing R-band Photometry beyond z=0.2 with B, V, and R-band

Nearby Photometry". In: *PASP* 108, p. 190. DOI: 10.1086/133709. eprint: arXiv:astro-ph/9505024.

Kim, A. G. et al. (Apr. 2013). "Standardizing Type Ia Supernova Absolute Magnitudes Using Gaussian Process Data Regression". In: *ApJ* 766, 84, p. 84. DOI: 10.1088/0004-637X/766/2/84. arXiv:1302.2925 [astro-ph.CO].

Kirshner, R. P. and J. Kwan (Oct. 1974). "Distances to extragalactic supernovae". In: *ApJ* 193, pp. 27–36. DOI: 10.1086/153123.

Lantz, B. et al. (Feb. 2004). "SNIFS: a wideband integral field spectrograph with microlens arrays". In: *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*. Ed. by L. Mazuray, P. J. Rogers, and R. Wartmann. Vol. 5249. Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, pp. 146–155. DOI: 10.1117/12.512493.

Laureijs, R. et al. (Oct. 2011). "Euclid Definition Study Report". In: *ArXiv e-prints*. arXiv:1110.3193 [astro-ph.CO].

LSST Science Collaboration et al. (Dec. 2009). "LSST Science Book, Version 2.0". In: *ArXiv e-prints*. arXiv:0912.0201 [astro-ph.IM].

Maguire, K. et al. (Nov. 2012). "Hubble Space Telescope studies of low-redshift Type Ia supernovae: evolution with redshift and ultraviolet spectral trends". In: *MNRAS* 426, pp. 2359–2379. DOI: 10.1111/j.1365-2966.2012.21909.x. arXiv:1205.7040 [astro-ph.CO].

Miller, D. L. and D. Branch (Aug. 1990). "Supernova absolute-magnitude distributions". In: *AJ* 100, pp. 530–539. DOI: 10.1086/115534.

Minkowski, R. (Aug. 1941). "Spectra of Supernovae". In: *PASP* 53, p. 224. DOI: 10.1086/125315.

Mosher, J. et al. (Sept. 2014). "Cosmological Parameter Uncertainties from SALT-II Type Ia Supernova Light Curve Models". In: *ApJ* 793, 16, p. 16. DOI: 10.1088/0004-637X/793/1/16. arXiv:1401.4065.

Nugent, Peter, Alex Kim, and Saul Perlmutter (2002). "K-corrections and Extinction Corrections for Type Ia Supernovae". In: *Publ.Astron.Soc.Pac.* 114, pp. 803–819. eprint: astro-ph/0205351. URL: http://arxiv.org/abs/astro-ph/0205351.

Oke, J. B. and A. Sandage (Oct. 1968). "Energy Distributions, K Corrections, and the Stebbins-Whitford Effect for Giant Elliptical Galaxies". In: *ApJ* 154, p. 21. DOI: 10.1086/149737.

Perlmutter, S. et al. (June 1999). "Measurements of Ω and Λ from 42 High-Redshift Supernovae". In: *ApJ* 517, pp. 565–586. DOI: 10.1086/307221. eprint: astro-ph/9812133.

Phillips, M. M. (Aug. 1993). "The absolute magnitudes of Type IA supernovae". In: *ApJ* 413, pp. L105–L108. DOI: 10.1086/186970.

Regnault, N. et al. (Nov. 2009). "Photometric calibration of the Supernova Legacy Survey fields". In: *A&A* 506, pp. 999–1042. DOI: 10.1051/0004-6361/200912446. arXiv:0908.3808 [astro-ph.IM].

Riess, A. G. et al. (Sept. 1998). "Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant". In: *AJ* 116, pp. 1009–1038. DOI: 10.1086/300499. eprint: astro-ph/9805201.

Rubin, D. et al. (Jan. 2013). "Precision Measurement of The Most Distant Spectroscopically Confirmed Supernova Ia with the Hubble Space Telescope". In: *ApJ* 763, 35, p. 35. DOI: 10.1088/0004-637X/763/1/35. arXiv:1205.3494 [astro-ph.CO].

Sandage, A. and G. A. Tammann (Sept. 1993). "The Hubble diagram in V for supernovae of Type IA and the value of H(0) therefrom". In: *ApJ* 415, pp. 1–9. DOI: 10.1086/173137.

Sasdelli, M. et al. (Feb. 2015). "A metric space for Type Ia supernova spectra". In: *MNRAS* 447, pp. 1247–1266. DOI: 10.1093/mnras/stu2416. arXiv:1411.4424 [astro-ph.SR].

Saunders, C. et al. (Feb. 2015). "Type Ia Supernova Distance Modulus Bias and Dispersion from K-correction Errors: A Direct Measurement Using Light Curve Fits to Observed Spectral Time Series". In: *ApJ* 800, 57, p. 57. DOI: 10.1088/0004-637X/800/1/57. arXiv:1412.5533.

Savitzky, A. and M. J. E. Golay (1964). "Smoothing and differentiation of data by simplified least squares procedures". In: *Analytical Chemistry* 36, pp. 1627–1639.

Scalzo, R. et al. (Sept. 2012). "A Search for New Candidate Super-Chandrasekhar-mass Type Ia Supernovae in the Nearby Supernova Factory Data Set". In: *ApJ* 757, 12, p. 12. DOI: 10.1088/0004-637X/757/1/12.

Scalzo, R. et al. (May 2014). "Type Ia supernova bolometric light curves and ejected mass estimates from the Nearby Supernova Factory". In: *MNRAS* 440, pp. 1498–1518. DOI: 10.1093/mnras/stu350. arXiv:1402.6842.

Scalzo, R. A. et al. (Apr. 2010). "Nearby Supernova Factory Observations of SN 2007if: First Total Mass Measurement of a Super-Chandrasekhar-Mass Progenitor". In: *ApJ* 713, pp. 1073–1094. DOI: 10.1088/0004-637X/713/2/1073. arXiv:1003.2217 [astro-ph.CO].

Schlegel, D. J., D. P. Finkbeiner, and M. Davis (June 1998). "Maps of Dust Infrared Emission for Use in Estimation of Reddening and Cosmic Microwave Background Radiation Foregrounds". In: *ApJ* 500, p. 525. DOI: 10.1086/305772. eprint: arXiv:astro-ph/9710327.

Schmidt, B. P. et al. (Nov. 1998). "The High-Z Supernova Search: Measuring Cosmic Deceleration and Global Curvature of the Universe Using Type IA Supernovae". In: *ApJ* 507, pp. 46–63. DOI: 10.1086/306308. eprint: astro-ph/9805200.

Thomas, R. C. et al. (Jan. 2007). "Nearby Supernova Factory Observations of SN 2006D: On Sporadic Carbon Signatures in Early Type Ia Supernova Spectra". In: *ApJ* 654, pp. L53–L56. DOI: 10.1086/510780. eprint: arXiv:astro-ph/0611356.

Thomas, R. C. et al. (Dec. 2011). "Type Ia Supernova Carbon Footprints". In: *ApJ* 743, 27, p. 27. DOI: 10.1088/0004-637X/743/1/27. arXiv:1109.1312 [astro-ph.CO].

Tripp, R. (Mar. 1998). "A two-parameter luminosity correction for Type IA supernovae". In: *A&A* 331, pp. 815–820.

Zhao, F.-Y., R. G. Strom, and S.-Y. Jiang (Oct. 2006). "The Guest Star of AD185 must have been a Supernova". In: *Chinese J. Astron. Astrophys.* 6, pp. 635–640. DOI: 10.1088/1009-9271/6/5/17.