UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

CONNECTIONS BETWEEN THE SPECTROSCOPY AND PHOTOMETRY OF TYPE IA SUPERNOVAE

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

ANTHONY R BURROW
Norman, Oklahoma
2024

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/340468

CONNECTIONS BETWEEN THE SPECTROSCOPY AND PHOTOMETRY OF TYPE IA SUPERNOVAE

A DISSERTATION APPROVED FOR THE
HOMER L. DODGE DEPARTMENT OF PHYSICS AND ASTRONOMY

BY THE COMMITTEE CONSISTING OF

Dr. Mukremin Kilic, Chair

Dr. Edward Baron

Dr. Sean Matt

Dr. Howard Baer

Dr. Henry Neeman

© Copyright by ANTHONY R BURROW 2024
All Rights Reserved.

Dedication

In loving memory of my Dad;

I am proud to be your son.

Acknowledgements

I am thankful for my family and their continuous love and support of my decisions throughout my life. I am lucky to have been raised in an environment that nurtured personal growth and open-mindedness, which was conducive to this personal and professional achievement.

I want to thank my partner, Bianca, for her unconditional love and for being such a significant and precious source of support in my life. I cannot express how much I value her advice and encouragement as I begin this next chapter of my life.

I would also like to thank my supervisor, Eddie Baron, who has given me the guidance and freedom needed to allow me to flourish as both a researcher and person in the professional world. His selfless support and encouragement has been paramount in my success as a graduate student.

I am grateful to the additional members of my committee — Mukremin Kilic, Sean Matt, Howard Baer, and Henry Neeman — for their combined guidance throughout my graduate career that led to this work. I would also like to thank former committee members Nate Kaib and John Wisniewski for their roles in facilitating this achievement. I would like to especially highlight the support from John Wisniewski throughout my undergraduate and graduate years at OU, whose impact in my professional life cannot be overstated.

I would like to thank members of our research group at OU, including James, Sarah, Zach, Vera, Sara, London, and many others. I am thankful for our strong teamwork and mutual support and respect of each other's work.

I would finally like to thank my collaborators whom I have had the privilege of meeting and working with over the past few years. I want to especially thank Chris Burns, Chris Ashall, and Jing Lu for their considerable support of my research and progress in my career. Thank you to all of the other members of the CSP and POISE groups, including Max, Mark, Eric, Melissa, Nidia, Lluís, Ben, Gastón, Kevin, Peter B., Peter H., Nick, Shuvo, and Tony, for welcoming me into this collaboration; without this group of individuals, this work would not be possible.

Contents

Abstract **ix**

1 Introduction **1**
1.1 Classification of Supernovae . . . 1
1.2 Type Ia Supernovae . . . 2
1.2.1 Single-Degenerate Scenario . . . 3
1.2.2 Double-Degenerate Scenario . . . 4
1.2.3 Core-Degenerate Scenario . . . 5
1.3 Core-Collapse Supernovae . . . 5
1.3.1 Type II Supernovae . . . 6
1.3.2 Type Ib/c Supernovae . . . 7
1.4 Dissertation Outline . . . 7

2 Observational Diversity of SNe Ia **10**
2.1 Light-Curve Properties . . . 10
2.2 Peculiar SNe Ia . . . 10
2.3 Branch Classification Scheme . . . 12
2.4 CSP I+II Data . . . 14

3 GMM Formulation of Branch Groups **15**
3.1 Introduction . . . 15
3.2 Data . . . 15
3.3 Methods . . . 24
3.3.1 Velocities and Pseudo-Equivalent Widths . . . 24
3.3.2 Cluster Analysis . . . 25
3.4 Results . . . 28
3.4.1 Branch Clustering . . . 28
3.4.2 Higher-Dimensional Clustering . . . 30
3.4.2.1 Inclusion of $v_{\mathrm{Si\,II}}$ . . . 30
3.4.2.2 Inclusion of M_B . . . 31
3.4.2.3 Inclusion of Both $v_{\mathrm{Si\,II}}$ and M_B . . . 32
3.4.3 M_B-vs-$v_{\mathrm{Si\,II}}$ Clustering . . . 34
3.4.3.1 Color of the Fast Group SNe Ia . . . 37
3.5 Discussion . . . 39
3.5.1 Constraining the Branch GMM . . . 39
3.5.2 M_B-vs-$v_{\mathrm{Si\,II}}$ Groups . . . 43
3.5.3 Branch Group Relation with s_{BV} . . . 45
3.6 Conclusion . . . 46

4 Variation of the NIR Secondary Peak **49**
4.1 Introduction . . . 49
4.2 Review of SNooPy . . . 50
4.3 SNooPy Residuals in the NIR . . . 50
4.4 Conclusion . . . 52

5 Extrapolation of Spectra into the NIR **53**
5.1 Introduction . . . 53
5.2 Data . . . 55
5.2.1 Data Source & Preprocessing . . . 55
5.2.2 Sample Selection . . . 56
5.3 Methods . . . 59
5.3.1 PCA Method . . . 59
5.3.2 PC Training . . . 60
5.3.3 Prediction Using PCA Eigenvectors . . . 61
5.3.4 Uncertainty Measurement . . . 61
5.4 Results . . . 62
5.4.1 Maximum-Light Eigenvectors . . . 62
5.4.2 Prediction Near Maximum Light . . . 64
5.4.3 Time-Dependent Predictions . . . 69
5.4.4 Prediction Using Si II . . . 70
5.5 Discussion . . . 72
5.5.1 Color-Matching with Broad-Band Photometry . . . 72
5.5.2 Comparison with Templates – Spectroscopy . . . 74
5.5.3 Comparison with Templates – Photometry . . . 76
5.6 Conclusion . . . 78

6 Modeling Uncertainties of NIR Extrapolation **80**
6.1 Introduction . . . 80
6.2 The Logistic Model . . . 80
6.3 Fitting with Gaussian Processes . . . 82
6.4 Sampling . . . 84
6.4.1 Parameter Distributions . . . 84
6.4.2 Resulting Realizations . . . 87
6.5 Conclusion . . . 89

7 Correlations of PC Projections **90**
7.1 Introduction . . . 90
7.2 PC Relation to Branch Groups . . . 90
7.3 PC Relation to s_{BV} . . . 93
7.4 Conclusion . . . 97

8 Conclusion **98**

A Appendix - Software **110**
A.1 `Spextractor` 110
A.2 `SNIaDCA` 111
A.3 `SNEx` 111

B Appendix - Gaussian Process Regression **113**

Abstract

Type Ia supernovae (SNe Ia) play a vital role in the study of many topics of astrophysics. They act as cosmological probes through their well-established nature as precise standardizable candles. This nature can be described empirically by the systematic relationship between the width of their light-curve and their peak luminosity, otherwise known as the Phillips relation. As extragalactic distance indicators, SNe Ia aid in constraining the nature of dark energy. SNe Ia also enrich the universe with iron-group products, and provide insight into our understanding of stellar evolution as a whole.

To fully and precisely take advantage of the standardizable properties of SNe Ia, their time evolution must be fully understood, including the nature of their origin. Unfortunately, the exact progenitor systems of SNe Ia are still uncertain and are a popular topic in the study of SNe Ia. The different theorized progenitor scenarios and explosion mechanisms that SNe Ia undergo stem from the diversity of their observable properties. Understanding the causes of the variation seen between SNe Ia is imperative in minimizing the Hubble residual, which is used to deduce cosmological parameters such as H_0.

Due to the complexities of supernova physics, the variation observed in SNe Ia are often described using empirical models. Recent advances in statistical and machine-learning techniques have led to models that provide useful constraints on theories that link observables to the underlying progenitor system. Although the time-dependent nature of SNe Ia leads to difficulty in obtaining large samples in order to treat observations statistically, recent surveys of astronomical transients have led to sample sizes large enough such that modern machine-learning techniques can be applied. The goal of this thesis is to use this increase in data available to create empirical models that will aid in the classification and prediction of SN Ia properties, which will further constrain the progenitor problem. Specifically, these models will be created using a combination of both observed spectroscopy and photometry. In doing so, correlations between the two regimes are found which are useful for constraining spectroscopic models and improving light-curve models of SNe Ia.

This thesis is organized in the following manner: Chapter 1 reviews the fundamental nature of SNe Ia along with previously theorized progenitor systems and explosion mechanisms; Chapter 2 discusses the observational diversity of SNe Ia, including some of the relevant subtypes of SNe Ia and the observational data used in empirical models created here; Chapter 3 details a cluster analysis of a sample of SNe Ia that leads to a robust model for classifying SNe Ia; Chapter 4 reviews

the light-curve-fitter `SNooPy` and, using `SNooPy`, illustrates the limitations of the color-stretch parameter s_{BV} in the near-infrared; Chapter 5 discusses a technique using principal component analysis to extrapolate NIR spectra of SNe Ia using optical spectroscopy; Chapter 6 proposes a method of modeling residuals from aforementioned extrapolations to more thoughtfully provide correlated uncertainties for the application of light-curve calculations; Chapter 7 studies the principal components from the extrapolation model in Chapter 5 in more detail, and relates them back to the classification scheme defined in Chapter 3; finally, Chapter 8 provides a summary of the most notable conclusions in this work.

CHAPTER 1

Introduction

Supernovae (SNe) are some of the most important objects in the study of stellar astrophysics and cosmology. The ejected material from these events allow us to probe the interior structure and nature of stars, providing insight into stellar evolution as a whole, as SNe mark the end of the lifetime of many stars. The nucleosynthesis that occurs both before and during the explosion of stars enrich the material ejected from these events, which spreads further into the interstellar medium, driving the most significant source of chemical evolution in the Universe. Because of their exceptionally high luminosities and standardizable properties, SNe also serve as extragalactic distance indicators, which allows for a measurement of the expansion rate of the Universe.

1.1. Classification of Supernovae

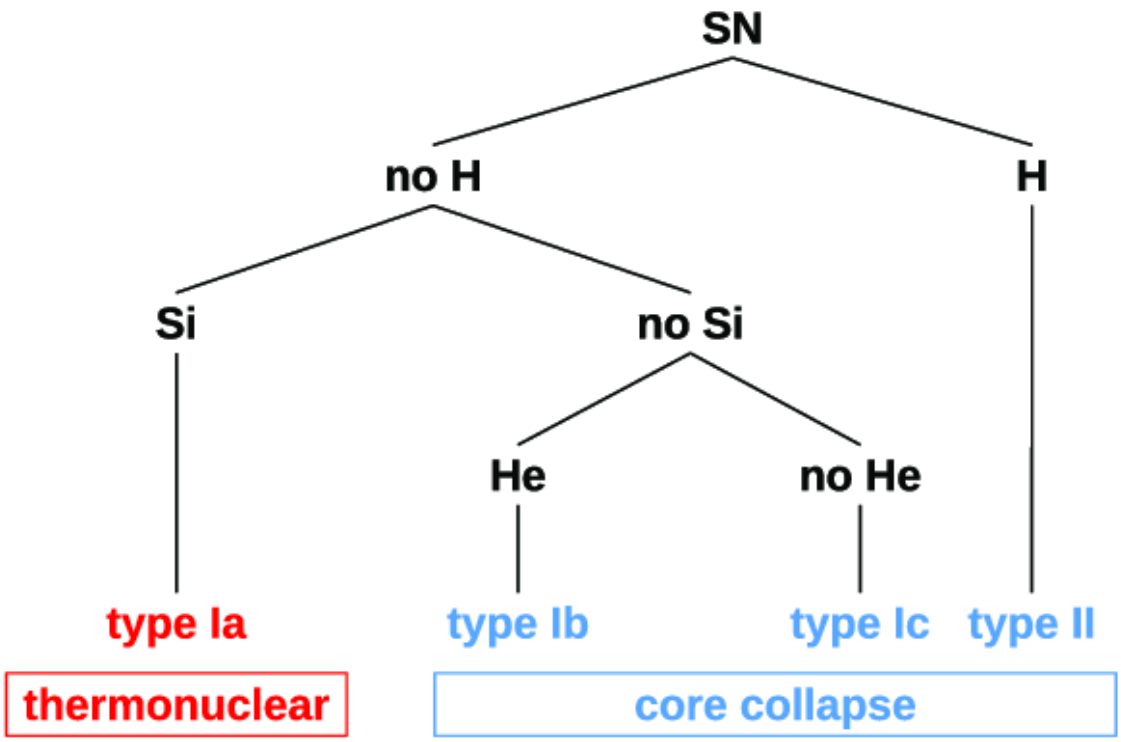


Figure 1.1: A basic decision tree on the categorization of supernovae, taken from Figure 3.3 by Paz Arribas (2017).

The classification of astronomical transients is carried out by examining their optical spectra. Through line identifications of the strongest features in the spectrum, SNe have been categorized into a handful of different classes and subclasses. At the highest level, SNe have been distinguished by either an indication of hydrogen in their spectrum or a lack thereof. Spectra containing hydrogen features belong to the class of SNe known as Type II supernovae (SNe II). Those without hydrogen features are deemed Type I supernovae. Figure 1.1 demonstrates the basic deciding process of categorizing supernovae.

Type I supernovae can be further subcategorized based on the presence of silicon in the spectrum. If the spectrum contains strong silicon features, then the corresponding object is called a Type Ia supernova (SN Ia), which originates from a thermonuclear explosion (see Section 1.2). The remainder of SNe are core-collapse supernovae (CCSNe; see Section 1.3), including SNe II and Type Ib and Ic supernovae (SNe Ib and SNe Ic, respectively). SNe Ib/c do not show strong silicon (or hydrogen) features in the ejecta. The two differ based on their helium content; those showing strong helium features in their spectra are SNe Ib, and those without are SNe Ic. These classification criteria are summarized visually looking at the line identifications shown in Figure 1.2 (taken from D. Kasen[1] and Branch & Wheeler (2017) Figure 1.2).

The topic of this thesis will be focused primarily on SNe Ia and their subtypes, which are fundamentally different objects from CCSNe. The remainder of this chapter will provide a brief overview of each aforementioned subclass of SNe, with particular emphasis on SNe Ia. For a more detailed summary, Branch & Wheeler (2017) provide a broad review of modern research on all types of supernovae.

1.2. Type Ia Supernovae

SNe Ia are believed to be the result of a thermonuclear explosion of C/O white dwarfs (WDs) that exist in binary systems (Hoyle & Fowler, 1960). The interaction with and accretion of mass from its companion is thought to be what eventually leads to a thermonuclear runaway process that causes a subsequent explosion. However, because WDs are such faint objects, observation of the progenitors that lead to SNe Ia is incredibly unlikely (Rebassa-Mansergas et al., 2019), and even impossible for SNe Ia hosted by galaxies other than the Milky Way galaxy, as we can only observe WDs on the order of hundreds to thousands of parsecs away (Debes et al., 2011; Werner & Rauch, 2015). This leads to the so-called progenitor problem: We do not know the progenitor systems of the SNe Ia that we observe. In this section, we briefly discuss some of the more accepted theoretical models of both the progenitors of SNe Ia (single-, double-, and core-degenerate models) and their underlying explosion mechanisms.

[1] https://supernova.lbl.gov/~dnkasen/tutorial/

1.2.1. Single-Degenerate Scenario

The canonical single-degenerate (SD) channel of SNe Ia (Whelan & Iben, 1973; Nomoto, 1982) puts focus on the mass-transfer from the secondary binary companion onto the surface of the primary WD that is expected to explode. This mass-transfer occurs either through the expansion of the secondary which leads to a Roche-lobe overflow, or from stellar winds of the secondary; both of which lead to the accretion of material onto the WD. As the C/O core of the WD has already reached degeneracy, it begins this process with a mass under the Chandrasekhar limit (1.4 $M_{\odot}$). As mass-transfer occurs, the mass of the WD approaches the Chandrasekhar mass, leading to a compression near the center of the WD which ignites carbon in the core and is followed by explosive thermonuclear runaway.

It should be noted that models show that achieving an explosion is sensitive to the accretion rate onto the WD to allow for the stable burning of hydrogen as the WD mass increases (Paczynski & Zytkow, 1978; Sion et al., 1979; Sienkiewicz, 1980; Nomoto, 1982). If mass-transfer is too fast, the WD would be completely immersed in hydrogen, and this hydrogen would be present in the spectrum — which by definition is indicative of SNe II. If the mass-transfer is slow enough, the accreted material becomes degenerate, producing strong hydrogen flashes that lead to classical nova explosions.

The detection of SNe Ia within the SD channel with nondegenerate companions has been thought to be indicated by narrow hydrogen features such as Paschen-β in the near-infrared (NIR) found after maximum light or other spectroscopic features, but they are often subject to constraints on the viewing angle (Marietta et al., 2000; Kasen, 2010; Maeda et al., 2014; Kutsuna & Shigeyama, 2015). These hydrogen lines occurring postmaximum would originate from the interaction of the ejecta with material from the companion that is stripped away, which can be seen as the photosphere recedes. Models have also shown that interaction with the ejecta and a companion star may lead to a rise in the early light-curves of SNe Ia that depend on the viewing angle (Kasen, 2010); however, some also find that overall photometric properties are not very sensitive to the existence of companion stars (Maeda et al., 2014).

Models of SN Ia explosions began with the W7 model (Nomoto et al., 1984), which is a 1-D deflagration model — deflagration meaning that a slow burning occurs below local sound speed, as opposed to detonation, which describes burning that occurs supersonically, where a shock wave forms ahead of the reaction front. For pure detonation models, the supersonic burning front would propagate through the entirety of the WD before reactions leading to iron-group elements can occur. On the opposite end, pure deflagration models lead to significant amounts of carbon and oxygen left unburned, which contradicts spectroscopic observations of SNe Ia. In addition, pure deflagration does not provide the system with enough energy to match observations of most spectroscopically

normal SNe Ia. Instead, a "delayed detonation" model is often used by implementing a deflagration-to-detonation transition (DDT) that occurs at a critical density (Khokhlov, 1991; Hoeflich et al., 1995).

Other central-detonation SD models also exist, such as the double-detonation model of a sub-Chandrasekhar C/O WD (Woosley & Weaver, 1994; Shen & Bildsten, 2014). In this scenario, a thin layer of helium accretes onto the surface of the WD, which can lead to an ignition of helium once the density reaches $\sim 10^6$ g cm^{-3} (Woosley & Weaver, 1994). This ignition causes an inward-directed shock wave to propagate through the star, and because it is a nearly spherically symmetric shock front, the point of convergence is near the center of the core. This convergence disrupts the C/O core, leading to densities allowed for a second detonation to occur, when the mass of the WD is still below the Chandrasekhar limit.

1.2.2. Double-Degenerate Scenario

In the canonical double-degenerate (DD) channel, the merger of two degenerate WDs spins down and, over time, the central densities become high enough to ignite carbon in the core (Iben & Tutukov, 1984; Webbink, 1984; Paczynski, 1985). This spin-down is achieved through dissipation of orbital energy via the emission of gravitational radiation (gravitational waves). The total mass of the system exceeds the Chandrasekhar limit for the canonical DD model, and observations of such an event should resemble those of the SD super-Chandrasekhar case such as those allowed by rotation (see, e.g., Domínguez et al., 2006).

Aside from the canonical model involving spun-down merger remnants, violent mergers have been modeled to detonate in one WD first while still undergoing a merger with the other WD (Pakmor et al., 2011, 2012; Kromer et al., 2013). In such models, the secondary WD explodes around a second later than the first, leading to asymmetries in the ejecta. Pakmor et al. (2013) find that simulations of the violent merger of two C/O WDs allow for thinner helium shells around the WD, and thus the products of burning helium would be less likely to show up in spectra, which would better match observations. In addition, they find that violent merger scenarios may be able to explain the observed colors and spectra of subluminous SNe Ia such as 91bg-likes (see also Pakmor et al., 2010); however, the fast-declining light-curves exhibited by 91bg-likes are proposed to be due to binaries of helium and C/O WDs, which their models also allow (Pakmor et al., 2013).

The detonation of the primary WD may sometimes be due to the ignition of a helium shell as the primary accretes material from a helium-rich WD companion (Guillochon et al., 2010; Dan et al., 2014), causing a double-detonation case for the DD scenario. As discussed in Section 1.2.1, the double-detonation scenario can occur for sub-Chandrasekhar-mass explosions. Modern models of DD double-detonation explosions include the so-called "dynamically driven double-degenerate double-detonation" (D6) model (Shen et al., 2018; Polin et al., 2019). In this model, a small but

rapid transfer of helium from a WD donor is sufficient in igniting the helium shell of the other WD, leading to carbon detonation in its core via double-detonation. Shen et al. (2018) find that the D6 model provides both synthetic light-curves and spectra that are roughly consistent with observations. In previous DD models that employ double-detonation on a relatively large (~ 0.1 $M_\odot$) helium shell (Woosley & Weaver, 1994), the resulting spectra have been shown to be too blue (Hoeflich & Khokhlov, 1996; Nugent et al., 1997). Instead, the D6 model replaces helium burning via the triple-alpha process with $^{12}C(\alpha,\gamma)^{16}O$, which is a faster burning process, allowing for much thinner (~ 0.01 $M_\odot$) helium shells.

1.2.3. Core-Degenerate Scenario

The core-degenerate (CD) scenario (Kashi & Soker, 2011; Ilkov & Soker, 2012) has been somewhat recently introduced in order to address some of the shortcomings of the DD scenario. The DD channel requires that both WDs survive after the common-envelope (CE) phase, where the primary WD shares an envelope with its pre-WD giant companion that has vastly expanded as it has evolved. In the CD channel, the companion of the WD becomes an asymptotic-giant-branch (AGB) star, and a merger occurs in the CE phase between the primary white dwarf and the hot C/O core of the AGB star (see also Sparks & Stecher, 1974). This hot core is usually larger and more dense than the primary WD, so the primary is disrupted and is accreted onto the larger core. The merger remnant then becomes a rapidly rotating WD, which then spins down over time, allowing for an explosion to occur.

There have currently not been many detailed investigations of or attempts at modeling CD-channel SNe Ia. In addition, the rate at which SNe Ia are produced via the CD channel is also not well known, as there are estimates in favor of the CD scenario playing a major role in SN production (Ilkov & Soker, 2013), and some that show that the CD scenario may only produce a small fraction of SNe Ia (Zhou et al., 2015).

1.3. Core-Collapse Supernovae

CCSNe occur in massive stars ($M \gtrsim 8$ $M_\odot$), as opposed to lower mass C/O WDs in binary systems which lead to thermonuclear explosions (SNe Ia). They are caused by an instability in the cores of massive stars, to such extent that the core cannot fully support itself against an immense gravitational collapse. Instability in cores of massive stars is primarily due to a combination of electron capture (Rakavy et al., 1967; Finzi & Wolf, 1967; Nomoto, 1984) and photodisintegration (Hoyle, 1946; Clifford & Tayler, 1965).

In higher-mass stars with iron cores, because iron is so stable and is unable to be fused into heavier elements, the core is unable to produce enough energy to prevent gravitational collapse as it exceeds the Chandrasekhar mass limit for iron. Photodisintegration of iron occurs as the core

heats up, and in the process it produces alpha particles and free neutrons. As the core becomes more dense, electron capture occurs, merging protons and electrons to produce neutrons and neutrinos; energy is lost as these neutrinos escape the system. Collapse continues until it is suddenly halted by the nuclear force between neutrons and protons. The sudden halt of infalling material forms a buildup of sound waves at the shock point, building into a shock wave that eventually propagates outward, creating what we observe as a CCSN.

Electron capture has been also suggested to occur for stars with a zero-age main-sequence (ZAMS) mass of 8–13 $M_\odot$ with degenerate O/Ne/Mg cores, where Ne and Mg are susceptible to electron capture (Miyaji et al., 1980; Nomoto, 1980, 1984, 1987). However, most studies of CCSNe are focused on the cases of stars massive enough to fuse up to iron in their cores.

1.3.1. Type II Supernovae

SNe II are a subset of CCSNe, which arise from the catastrophic collapse of massive stars. Observationally, they are classified by the presence of strong hydrogen (Figure 1.1) features in their spectra, which are due to hydrogen still present in the envelopes of these massive stars. The first SN II studied in great detail was SN 1987A, which occurred in the Large Magellanic Cloud. SN 1987A sparked a revolution in the observation of SNe, as it was the first SN to occur in the Local Group since the invention of the telescope. Its progenitor star was identified as Sanduleak -69 202 — a B3 supergiant — and neutrinos from its core-collapse were detected in terrestrial neutrino detectors: 12 from Kamiokande II, and 8 from the Irvine-Michigan-Brookhaven (IMB) detector (Arnett et al., 1989). The observation of this SN II provided much insight into core-collapse physics and the progenitors of CCSNe (see reviews by Arnett et al., 1989; Hillebrandt & Hoflich, 1989; Imshennik & Nadezhin, 1989; McCray, 1993).

SNe II have been classified into the two subclasses Type IIP and Type IIL SNe (SNe IIP and SNe IIL, respectively). SNe IIP are defined by a plateau phase in their light-curve that lasts around 100 days. In contrast, SNe IIL show a nearly linear decline (in magnitudes) in their light-curves (Barbon et al., 1979). More recent studies of SNe II have suggested that these two subclasses are the extremes of a continuum, and that there is no statistical distinction between two individual groups of SNe II (Anderson et al., 2014; Valenti et al., 2016). Instead, the slope of the light-curve is a function of the mass of the progenitor envelope at the time of explosion. Anderson et al. (2014) find that SNe II with faster-declining light-curves have lower ejecta mass.

In addition to photometric classifications, SNe II can also be subcategorized by their spectroscopic qualities. Type IIn supernovae (SNe IIn, Schlegel, 1990) show narrow hydrogen emission lines in their optical spectra on top of broad lines seen in typical Type II spectra. These emission features originate from interactions with substantial amounts of circumstellar material. Another spectroscopic subclass of SNe II are Type IIb SNe (SNe IIb). The “b” in SNe IIb follows the

convention of SNe Ib (Section 1.3.2), in that they exhibit helium in their ejecta (resembling spectra of SNe Ib) following an earlier hydrogen-dominated spectral phase (Filippenko, 1988). The progenitors of SNe IIb have retained a thin layer of hydrogen in their outer envelope prior to explosion, and after explosion, the photosphere recedes, quickly uncovering the helium-rich layers of the ejecta, forming a spectrum similar to SNe Ib.

1.3.2. Type Ib/c Supernovae

SNe Ib/c are physically similar to SNe II, as they originate from core-collapse events instead of a thermonuclear explosion of a C/O WD. They are spectroscopically defined as SNe I because they do not show hydrogen in their optical spectra, which is because the hydrogen and sometimes helium in the progenitor envelope have been stripped away (Smartt, 2009; Smith et al., 2011; Smith, 2014; Bersten & Mazzali, 2017). The cause of SNe Ib/c having lost their envelopes is uncertain; however, popular theories suggest that strong stellar winds similar to those in Wolf-Rayet stars and binary Roche lobe overflow are the source of mass-loss of these progenitors (Gaskell et al., 1986; Smith et al., 2011; Smith, 2014; Smartt, 2015).

As was hinted in the discussion of SNe IIb in Section 1.3.1, SNe Ib are CCSNe whose progenitors have had their hydrogen envelopes stripped away, however retaining their underlying helium envelopes, leaving helium features in their spectra. On the other hand, SNe Ic are those that have had both their hydrogen and helium envelopes removed, leaving behind spectra dominated by intermediate mass elements such as Mg, Na, O, and Ca. Although there may be evidence in suggesting that SNe Ic undergo different physical mechanisms prior to and during explosion than those of SNe Ib, these two categories may be two extremes of a continuum of helium abundance (Filippenko et al., 1990; Wheeler et al., 1994; Hamuy et al., 2002). For this reason, the two subclasses are often jointly studied, referring to them in conjunction as SNe Ib/c or stripped-envelope SNe.

1.4. Dissertation Outline

This thesis focuses on studies of different subtypes of SNe Ia. In Chapter 2, a summary of some of the major classification schemes of SNe Ia is provided. Historically, studies have found that SNe Ia exhibit a wide variety of both photometric and spectroscopic properties. This work puts particular emphasis on using statistical analyses to arrive at conclusions that connect SN Ia photometry and spectroscopy. In Chapter 3, the classification scheme of Branch et al. (2006) is extended to include photometric properties, and they are shown to be statistically robust. This classification scheme is also used to discuss the possibility suggested by Polin et al. (2019), that there is a dichotomy when looking at the maximum B-band magnitude of SNe Ia and the blueshift velocity of the Si II $\lambda6355$ feature at the same phase with respect to B-band maximum, corresponding to a distinction between Chandrasekhar- and sub-Chandrasekhar-mass explosion mechanisms.

Chapter 4 transitions into examining NIR spectroscopy and photometry by illustrating a limitation of empirical models of SN Ia photometric light-curves. Chapter 5 presents a method of estimating NIR spectral energy distributions (SEDs) using observed spectra that addresses this limitation of photometry. In Chapter 6, the uncertainties in predictions made in Chapter 5 are further explored to be more suitable and accurate for light-curve fitters. Chapter 7 ties together the results and classification scheme from Chapter 3 with the models created in Chapter 5 to demonstrate correlations seen (or unseen) between spectroscopic and photometric properties. Finally, in Chapter 8, the results shown in this dissertation are summarized.

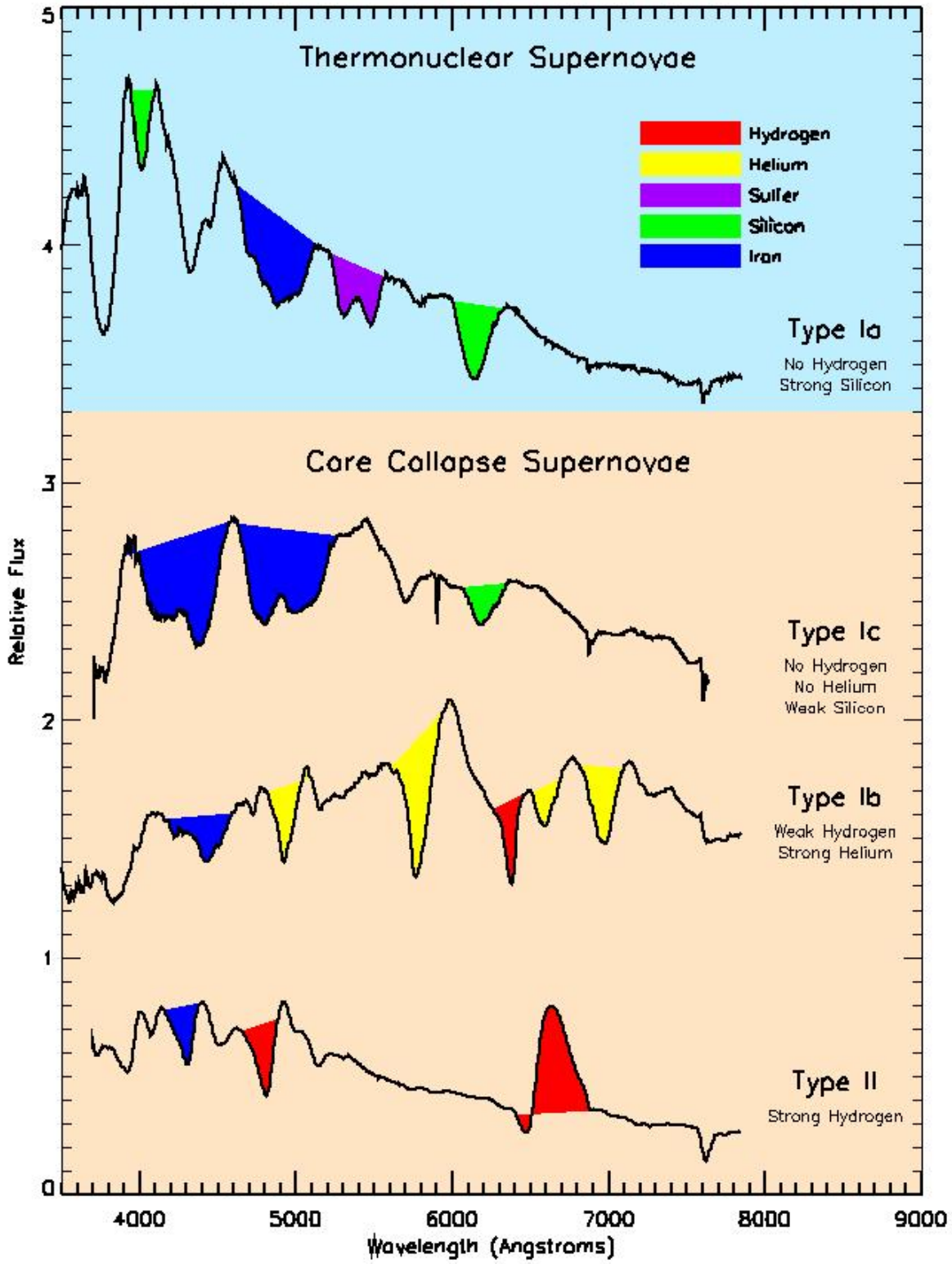


Figure 1.2: Example spectra of SN subclasses, showing identifications of lines that are defining or typical of the major subclasses of SNe. Figure was created by D. Kasen.

CHAPTER 2

Observational Diversity of SNe Ia

2.1. Light-Curve Properties

SNe Ia have been made standardizable, correctable candles based on their light-curve shape. This property of SNe Ia is the reason that they are so useful as cosmological distance indicators. The shape of their light-curves can be parameterized by the quantity $\Delta m_{15}(B)$, which describes the width of the light-curve by the decline (in magnitudes) of the B-band light-curve from the primary maximum to $+15$ days postmaximum. Phillips (1993) found that intrinsically brighter and bluer SNe Ia have wider light-curves with smaller $\Delta m_{15}(B)$ values; wider light-curves mean that the SN Ia is a "slow-decliner". Conversely, SNe Ia that are intrinsically dimmer and redder are "fast-decliners" with higher decline rates. This is commonly referred to as the Phillips relation (Phillips, 1993; Phillips et al., 1999). However, the simplicity of $\Delta m_{15}(B)$ leads to limitations in its effectiveness in describing SN Ia light-curves. First, the shape of the B-band light-curve changes with reddening, and $\Delta m_{15}(B)$ cannot account for this change. Various other technical problems arise as well, such as the need to interpolate the light-curve at two specified epochs.

There have been other parameters since the introduction of the Phillips relation that have been implemented to quantify the light-curve shape of SNe Ia. The light-curve stretch parameter s is essentially a time-stretching factor that corrects the rest-frame light-curve into a template (Perlmutter et al., 1999). However, use of this stretch parameter has evolved into use on an SED model in place of light-curves in the light-curve-fitter `SALT` (Guy et al., 2005), and it has been more generalized in updated versions `SALT2` and `SALT3` (Guy et al., 2007; Kenworthy et al., 2021). To address many problems with the decline rate $\Delta m_{15}(B)$, Burns et al. (2014) defined the color-stretch parameter $s_{\rm BV}$ as the time of maximum $B-V$ relative to B-band maximum measured in the rest frame, divided by 30 days. In general, the photometric diversity of SNe Ia is described using templates based on empirical models that are parameterized by various quantities that describe the stretch of their light-curves. For example, see Burns et al. (2011), in particular Figures 4 and 5.

2.2. Peculiar SNe Ia

By definition, spectroscopically normal SNe Ia are those that exhibit the most homogeneity in their strongest spectral features. They are, therefore, the most populous category of SNe Ia, and they serve as a good control group in statistical analyses of SNe Ia that show exceptional distinctions in their spectra. SNe Ia that display uncharacteristic features in their spectra are commonly referred to

as peculiar SNe Ia. Understanding the diversity of peculiar SNe Ia and the fundamental reasons that make these observables distinct is key to improving upon the effectiveness of SNe Ia as cosmological probes and to understanding stellar evolution as a whole. It is possible that peculiar SNe Ia experience completely different physical scenarios leading up to explosion and thereafter, which is what the supernova community aims to deduce. In this section, we describe the observable properties of some of the most distinguishable subtypes of peculiar SNe Ia.

91T-like SNe Ia are overluminous at B-maximum and slow-declining; they are expected to follow the same behavior as the SN 1991T (Filippenko et al., 1992a; Phillips et al., 1992). They have comparably weak Si II and Ca II features near maximum light (Phillips et al., 2024), which is discussed further in Section 2.3 regarding shallow-silicons. 91T-like SNe Ia are also notable for their iron-rich composition, with spectra dominated by Fe III lines at early times (Filippenko et al., 1992a; Ruiz-Lapuente et al., 1992; Jeffery et al., 1992; Mazzali et al., 1995; Fisher et al., 1999).

91bg-likes are subluminous, fast-declining SNe Ia, following the trends found in the SN 1991bg (Filippenko et al., 1992b; Leibundgut et al., 1993). 91bg-like SNe Ia are cooler in temperature and exhibit redder colors around maximum light. Spectroscopically, these SNe Ia have a characteristic blue trough around 4200 Å, which has been identified to be produced by Ti II (Filippenko et al., 1992b; Nugent et al., 1995, 1997). Filippenko et al. (1992b) find that SN 1991bg exhibits properties indicative of a slow deflagration of the white dwarf progenitor; however, a sub-Chandrasekhar-mass explosion is required to achieve high enough expansion velocities that have been observed. This seems to support the argument by Polin et al. (2019), that dimmer SNe Ia at maximum light (in the B band) follow sub-Chandrasekhar-mass explosion models.

There are also many other peculiar SNe Ia that have shown to have distinct characteristics: 02cx-likes, otherwise known as SNe Iax, have lower luminosities but with early spectra resembling 91T-likes (Li et al., 2003; Foley et al., 2013; Jha, 2017); 03fg-like SNe Ia are overluminous objects which are also referred to as super-Chandrasekhar-mass SNe Ia because their spectra show low velocities and strong carbon features, which has been interpreted to imply that the explosion stems from a super-Chandrasekhar-mass progenitor (Howell et al., 2006; Hicken et al., 2007; Scalzo et al., 2010); 02ic-likes have a strong presence of hydrogen emission in their spectra, which for SNe Ia imply an interaction of the ejecta with circumstellar material (Hamuy et al., 2003); and 06bt-like SNe Ia are slow-decliners (indicative of high luminosity) without a secondary maximum in their light-curve, similar to SNe Ia with lower luminosities (Foley et al., 2010, see Chapter 4 for more information on the secondary maximum).

The recent discovery and study of the SN 2019ein has been another particularly interesting case. SN 2019ein was classified as a broad-line (see Section 2.3) due to it's broad Si II features and high velocities (Kawabata et al., 2020; Pellegrino et al., 2020). These velocities were some of the highest ever observed for SNe Ia, with Si II $\lambda 6355$ absorption blueshifted as much as 24,000 km s^{-1}

14 days after maximum light. Interestingly, this blueshift did not occur for only the absorption minimum; the emission peak of the P-Cygni profile was also blueshifted to some extent, which is not expected for spherical ejecta. Xi et al. (2022) suggest that this behavior may arise from asymmetries in the ejecta; however, using spectropolarimetry, Patra et al. (2022) found no significant evidence of polarization due to asymmetric geometry of the ejecta. Yarbrough et al. (2023) have examined the time-dependence of the spectra of SN 2019ein to find that its spectra resemble those of a core-normal SN Ia (see Section 2.3) that has been Doppler-shifted.

SNe Ia have a wide variety of observables and therefore, to distinguish them, a classification scheme is necessary for consistent analyses. For example, Branch et al. (2006) classified SNe Ia into four groups using Si II properties in their spectra; more detail on this is given in Section 2.3. Wang et al. (2009) classified SNe Ia into another four groups — Normal, High-Velocity, 91T-like, and 91bg-like — which are similar to those by Branch et al. (2006), with instead a larger focus on photometric quantities and velocity of the ejecta. Blondin et al. (2012) provides a comparison between the two with a larger data set, finding that the groups are continuous with ill-defined boundaries, and that a single parameter is unlikely to describe the variation between SNe Ia of different subclasses. Ashall et al. (2020) illustrate a method of distinguishing different subtypes of SNe Ia using photometric quantities $s_{\rm BV}$ and the time of i-band maximum relative to B-band maximum or g-band maximum ($t_{\rm max}^{i-B}$ or $t_{\rm max}^{i-g}$, respectively). This method is able to delineate between normal SNe Ia and 91bg-likes, 02cx-likes, and 03fg-likes, however not 91T-likes.

2.3. Branch Classification Scheme

One spectroscopic classification scheme that has been successful in explaining much of the fundamental diversity of SNe Ia is that formulated by Branch et al. (2006). This scheme compares the pseudo-equivalent widths (pEWs) of two Si II features at 5972 Å and 6355 Å from spectra observed around B-band maximum to find four visually distinct subclasses of SNe Ia. These groups are referred to as core-normals (CNs), shallow-silicons (SSs), broad-lines (BLs) and cools (CLs). Throughout this work, these groups are referred to as the Branch groups. The typical visualization of the Branch groups, which is adopted in this work, is what we refer to henceforth as the Branch diagram, where the measured pEW(Si II $\lambda5972$) is plotted on the y-axis against the corresponding pEW(Si II $\lambda6355$) on the x-axis (see Figure 3.3 for an example). In the remainder of this section, we provide a description of these four Branch groups, as Chapter 3 focuses on the creation of a robust definition for the Branch groups.

CNs are the most spectroscopically standard and populous group of SNe Ia, lying in the middle of the Branch diagram. Branch et al. (2006) find strong homogeneity in the spectra of CNs except for a relatively prominent high-velocity Ca II feature in one SN Ia in their sample. CNs typically exhibit lower velocities ($v_{\rm Si\,II} \leq 12{,}000$ km s^{-1}, where $v_{\rm Si\,II}$ is the measured line velocity of Si II).

However, we will show in Chapter 3 that, according to a cluster analysis, this may not always be the case.

SSs show shallower absorption widths for both Si II $\lambda5972$ and Si II $\lambda6355$, and as such they tend to lie to the lower left of CNs on the Branch diagram. They typically show comparable ejecta velocities to CNs ($v_{\text{Si II}} \leq 12{,}000$ km s^{-1}). Branch et al. (2006) find that most differences between SSs and CNs stem from SSs having higher temperatures than CNs. Branch et al. (2006) also find that SN 1991T lies among the SSs; 91T-like SNe Ia have been recently shown to be extreme SS SNe Ia, showing the weakest Si II and Ca II features of all SNe Ia around B-band maximum (Phillips et al., 2024). Phillips et al. (2024) also find that 99aa-like SNe Ia are also members of the SS group that display less extreme SS qualities than 91T-likes, and they are distinguishable by having little time-evolution in their $v_{\text{Si II}}$ over the range of -10 to $+10$ days relative to B-band maximum.

BLs are named as such because they are defined by exhibiting a broader, deeper Si II $\lambda6355$ feature than CNs, and because of this, they lie to the right of CNs on the Branch diagram. Because the breadth of Si II $\lambda6355$ is tied to its velocity for spherical ejecta (i.e., assuming the feature has a P-Cygni profile), these BLs also exhibit higher velocities in the ejecta ($v_{\text{Si II}} \geq 12{,}000$ km s^{-1}). BLs indicated by higher $v_{\text{Si II}}$ are more scarcely observed, and therefore the study of BLs is much more limited.

CLs exhibit the highest values of pEW(Si II $\lambda5972$) and therefore lie above CNs on the Branch diagram. They are less luminous than SNe Ia of other Branch subtypes, and they have lower photospheric temperatures. Most of the CL sample used by Branch et al. (2006) show an absorption trough around 4000–4400 Å, which is due partially to Ti II at photospheric velocity. CLs also display a strong, narrow absorption feature around 5675 Å, which seems due to Na I D (Filippenko et al., 1992b; Garnavich et al., 2004). The fast-declining 91bg-like SNe Ia are a subset of CLs, and the CL group has been shown to be a fairly homogeneous group of SNe Ia (Doull & Baron, 2011).

Prior to work done by Burrow et al. (2020), these four subtypes of SNe Ia had no strict boundaries that asserted group membership for various SNe Ia. Because of this, consistency between different analyses was lacking. In Chapter 3, we review the work done by Burrow et al. (2020), where a cluster analysis shows that these groups are statistically significant, and the groups are not only more vigorously defined with probabilistic distributions, but they are also extended to contain photometric and velocity information. In addition, note that, since SNe Ia spectra lack a well-defined continuum, the calculation of pEW for different features is not consistent either. In Chapter 7, we suggest an alternative measurement that shows potential in yielding a consistent classification scheme between analyses; however, future work would be necessary to reinforce this suggestion.

2.4. CSP I+II Data

To train empirical models described in Chapter 3 and Chapter 5, we primarily make use of optical+NIR spectroscopy and photometry provided by the Carnegie Supernova Project (CSP) I and II data sets. We include this section to summarize the relevant technical details of the observations within these data sets that are used to model the diversity of SNe Ia.

The CSP I project (Hamuy et al., 2006; Folatelli et al., 2010; Krisciunas et al., 2017; Morrell et al., 2024) was a targeted survey of low-redshift SNe Ia with the goal of establishing a homogeneous set of optical and NIR light-curves in a natural photometric system paired with corresponding spectrophotometry. Observations were carried out at the Las Campanas Observatory (LCO) Henrietta Swope $1\ m$ telescope ($ugriBV$;YJH bands), and NIR photometry is obtained using the Wide-field Infrared Camera (WIRC) at the du Pont $2.5\ m$ telescope within the LCO ($YJHK_S$ bands). The survey consists of $\sim$100 SNe Ia with hosts within a redshift range of $0.0037 < z < 0.0835$ (Krisciunas et al., 2017).

The CSP II project (Phillips et al., 2019; Hsiao et al., 2019; Burrow et al., 2020; Morrell et al., 2024) is a sequel to the CSP I survey. Most optical ($ugriBV$) and NIR (YJH) photometry for CSP II was obtained again from the LCO Swope and du Pont telescopes. The LCO *Magellan* Baade $6.5\ m$ telescope was also used to obtain NIR photometry of nine SNe Ia, as well as some NIR spectroscopy. Optical spectroscopy was acquired using the LCO du Pont and *Magellan* telescopes as well as the Observatorio del Roque de los Muchachos $2.5\ m$ Nordic Optical Telescope (NOT). The majority of NIR spectra, which consists of 661 spectra of 157 SNe Ia, were obtained using the Folded-port InfraRed Echellette (FIRE) in the high-throughput prism mode on the *Magellan* telescope. The host redshift of SNe Ia in the CSP II data lies in a range of $0.03 < z < 0.10$ (Phillips et al., 2019).

CHAPTER 3

GMM Formulation of Branch Groups

This chapter is taken from Burrow et al. (2020), published in *The Astrophysical Journal*, reformatted with slight modifications to improve the readability, consistency, and sequencing of this dissertation.

3.1. Introduction

Recently, Zheng et al. (2018) suggested an empirical fitting method for SN Ia light-curves using the rise-time and the Si II $\lambda 6355$ velocity at maximum light as a method for calibrating the B-band peak luminosity. Using the results of Zheng et al. (2018), Polin et al. (2019) plotted peak absolute magnitude in the B-band versus Si II $\lambda 6355$ velocity at maximum light. This plot, which is henceforth referred to as the M_B-vs-$v_{\text{Si II}}$ diagram, seemed to indicate a dichotomy in SN Ia explosion mechanisms which, when displayed so as to indicate $B_{\text{max}} - V_{\text{max}}$ color that has been corrected for Milky Way (MW) Galaxy extinction, led to an interpretation that the redder SNe Ia could be fit well by explosion models of helium detonations on sub-Chandrasekhar-mass progenitors (Polin et al., 2019). In this chapter, this M_B-vs-$v_{\text{Si II}}$ space is explored further to determine if a dichotomy in this space is statistically significant.

To study the M_B-vs-$v_{\text{Si II}}$ space in detail, a statistical treatment of the Branch classification system is first performed. We make use of these Branch groups to make connections between Si II pEWs and the M_B-vs-$v_{\text{Si II}}$ space. These connections are supportive in determining whether or not a dichotomy actually exists in the M_B-vs-$v_{\text{Si II}}$ diagram. We perform a detailed cluster analysis of the Branch groups in addition to a cluster analysis of the M_B-vs-$v_{\text{Si II}}$ space.

3.2. Data

We include SNe Ia whose time of maximum is known, whose M_B (the peak absolute magnitude in the B-band at maximum light) has been determined (Burns et al., 2014, 2018), and whose spectra fall into the range $|t_{epoch}| < 7$ days past the inferred time of B-band maximum. From the initial 364 SNe Ia in the CSP I+II set, our sample consists of a subset of 97 objects. We include additional SNe Ia available using data from Zheng et al. (2018). In doing so, there is a deviation from a completely homogeneous sample of data; however, we include this data set not only to increase the sample size, but also to compare our results with those of Polin et al. (2019) in Section 3.4.3. From the Zheng et al. (2018) set of 54 SNe Ia, 14 are shared with CSP I, and CSP I spectra were

preferred. We ultimately include 36 unique SNe Ia from the Zheng et al. (2018) sample. In total, our sample consists of 133 SNe Ia.

We examine spectroscopic classifications using CSP I+II samples (see Section 2.4 for a detailed description). The values of the photometric quantities $m_{B,\mathrm{max}}$, s_{BV}, host A_V, etc. were computed using `SNooPy` (Burns et al., 2014), which performs K-corrections. M_B is then determined from the value of $m_{B,\mathrm{max}}$, the MW extinction, the inferred host-galaxy extinction, and the distance modulus using $H_0 = 72$ km s^{-1} Mpc^{-1}. M_B and s_{BV} values for the CSP I+II sample used in this analysis are provided in Table 3.1.

Table 3.1. Spectroscopic and photometric information for all 133 SNe Ia from both the CSP I+II samples and the Zheng et al. (2018) sample. Values from the Zheng et al. (2018) sample are given in the bottom sector of this table.

SN	Epoch [days past B_{max}]	$v_{Si\ II}$ [1000 km s^{-1}]	pEW(Si II λ5972) [Å]	pEW(Si II λ6355) [Å]	M_B [mag]	s_{BV}
ASASSN-14hr	5.3	13.6 ± 0.3	30.5 ± 4.2	103.1 ± 5.3	-19.10 ± 0.12	0.79 ± 0.05
ASASSN-14hu	6.9	12.5 ± 0.2	7.4 ± 1.2	86.7 ± 2.8	-19.52 ± 0.12	1.08 ± 0.05
ASASSN-14kq	-2.3	10.5 ± 0.3	10.7 ± 1.5	82.5 ± 3.5	-19.44 ± 0.11	1.19 ± 0.05
ASASSN-14mf	5.2	9.9 ± 0.3	23.8 ± 2.0	103.8 ± 3.8	-19.45 ± 0.15	0.98 ± 0.05
ASASSN-14my	3.6	12.0 ± 0.3	26.1 ± 3.9	106.9 ± 4.3	-19.47 ± 0.12	0.91 ± 0.05
ASASSN-15al	5.9	11.7 ± 0.1	19.7 ± 4.2	81.7 ± 5.5	-19.21 ± 0.16	1.10 ± 0.06
ASASSN-15ba	4.0	11.4 ± 0.2	16.2 ± 1.7	117.3 ± 3.3	-19.31 ± 0.12	0.97 ± 0.05
ASASSN-15be	1.4	11.8 ± 0.4	3.8 ± 1.8	94.3 ± 3.8	-19.37 ± 0.12	1.20 ± 0.05
ASASSN-15bm	1.2	10.6 ± 0.2	9.7 ± 2.6	72.3 ± 4.9	-19.57 ± 0.14	1.00 ± 0.05
ASASSN-15dd	2.1	10.8 ± 0.3	21.0 ± 1.6	102.1 ± 2.7	-19.63 ± 0.15	0.83 ± 0.05
ASASSN-15fr	0.2	11.7 ± 0.2	20.7 ± 3.7	107.4 ± 6.1	-19.21 ± 0.09	0.88 ± 0.05
ASASSN-15ga	5.2	9.4 ± 0.2	52.4 ± 5.1	126.1 ± 5.0	-17.95 ± 0.40	0.50 ± 0.06
ASASSN-15hf	1.1	11.1 ± 0.3	21.3 ± 2.9	83.6 ± 2.9	-19.07 ± 0.38	0.91 ± 0.05
CSP14acl	5.0	10.1 ± 0.3	16.2 ± 3.2	116.0 ± 5.4	-18.91 ± 0.08	0.93 ± 0.05
CSP15B	2.2	13.9 ± 0.4	31.5 ± 4.5	175.4 ± 5.6	-19.15 ± 0.19	0.71 ± 0.05
CSP15aae	2.3	11.5 ± 0.4	56.8 ± 2.8	151.5 ± 3.7	-17.92 ± 0.15	0.47 ± 0.05
LSQ11bk	1.0	12.6 ± 0.3	5.6 ± 1.5	98.2 ± 3.5	-19.33 ± 0.06	1.08 ± 0.05
LSQ12fxd	3.8	10.9 ± 0.3	17.9 ± 1.9	81.0 ± 3.4	-19.78 ± 0.10	1.09 ± 0.05
LSQ12gdj	1.6	10.7 ± 0.5	8.7 ± 1.8	36.0 ± 2.2	-19.78 ± 0.10	1.14 ± 0.05
LSQ12hzj	3.2	9.8 ± 0.3	13.5 ± 2.2	59.1 ± 3.7	-19.11 ± 0.10	0.96 ± 0.05

Table 3.1 (cont'd)

SN	Epoch [days past $B_{\max}$]	$v_{\text{Si II}}$ [1000 km s^{-1}]	pEW(Si II λ5972) [Å]	pEW(Si II λ6355) [Å]	M_B [mag]	s_{BV}
LSQ13aiz	-6.8	14.2 ± 0.2	17.4 ± 2.7	133.6 ± 3.7	-19.67 ± 0.33	0.95 ± 0.05
LSQ13ry	4.7	12.6 ± 0.2	20.4 ± 1.2	82.3 ± 1.5	-19.15 ± 0.09	0.86 ± 0.05
LSQ15agh	3.8	9.9 ± 0.6	3.5 ± 2.6	55.3 ± 5.0	-19.54 ± 0.08	1.21 ± 0.05
LSQ15aja	3.4	10.2 ± 0.3	1.7 ± 3.2	78.1 ± 8.1	-19.57 ± 0.08	1.03 ± 0.05
PS1-14ra	3.6	11.2 ± 0.3	34.2 ± 3.9	124.0 ± 6.3	-19.11 ± 0.10	0.77 ± 0.05
PS1-14xw	1.2	11.7 ± 0.2	10.9 ± 2.1	77.8 ± 2.7	-19.72 ± 0.13	1.06 ± 0.05
PTF11pra	1.7	10.8 ± 0.4	55.5 ± 3.4	126.1 ± 4.0	-17.31 ± 0.24	0.40 ± 0.05
PTF13duj	1.6	14.2 ± 0.2	7.1 ± 1.7	100.4 ± 3.4	-19.35 ± 0.17	1.19 ± 0.05
PTF13ebh	0.6	10.8 ± 0.2	48.3 ± 2.3	125.0 ± 2.7	-18.76 ± 0.19	0.61 ± 0.05
PTF14w	4.6	12.2 ± 0.3	37.0 ± 2.3	124.0 ± 3.4	-18.80 ± 0.13	0.73 ± 0.05
2004ey	1.1	11.0 ± 0.3	13.8 ± 4.8	100.0 ± 8.7	-19.28 ± 0.17	1.01 ± 0.06
2004gs	1.8	11.2 ± 0.1	41.7 ± 6.3	136.7 ± 8.0	-18.86 ± 0.11	0.69 ± 0.06
2005M	0.5	9.2 ± 0.4	7.4 ± 3.7	47.5 ± 4.5	-19.61 ± 0.11	1.21 ± 0.06
2005bg	2.6	10.6 ± 0.3	7.3 ± 3.1	105.5 ± 9.9	-19.47 ± 0.16	1.00 ± 0.07
2005el	1.7	10.9 ± 0.3	19.5 ± 3.2	93.9 ± 4.2	-19.14 ± 0.15	0.84 ± 0.06
2005eq	6.7	9.8 ± 0.2	16.8 ± 4.8	77.3 ± 7.1	-19.60 ± 0.12	1.12 ± 0.06
2005hc	-6.0	10.3 ± 0.6	12.6 ± 13.5	101.5 ± 16.5	-19.46 ± 0.10	1.19 ± 0.06
2006D	1.8	11.2 ± 0.2	26.4 ± 0.9	100.4 ± 1.1	-19.26 ± 0.24	0.81 ± 0.06
2006ax	0.0	10.6 ± 0.3	9.6 ± 2.0	92.5 ± 2.9	-19.47 ± 0.13	0.99 ± 0.06
2006br	0.6	13.9 ± 0.4	8.5 ± 4.0	111.8 ± 8.5	-19.52 ± 0.27	0.91 ± 0.07

Table 3.1 (cont'd)

SN	Epoch [days past $B_{\max}$]	$v_{\mathrm{Si\,II}}$ [1000 km s^{-1}]	pEW(Si II λ5972) [Å]	pEW(Si II λ6355) [Å]	M_B [mag]	s_{BV}
2006hb	3.2	10.5 ± 0.3	41.4 ± 3.6	121.2 ± 5.2	-18.83 ± 0.18	0.67 ± 0.06
2006hx	3.5	10.5 ± 0.5	12.6 ± 5.9	42.6 ± 5.6	-19.64 ± 0.13	0.99 ± 0.06
2007S	0.1	10.1 ± 0.2	9.2 ± 3.2	58.7 ± 5.4	-19.85 ± 0.16	1.11 ± 0.06
2007af	1.1	11.2 ± 0.2	18.7 ± 5.2	104.2 ± 5.8	-19.12 ± 0.35	0.93 ± 0.06
2007as	3.5	12.6 ± 0.4	18.0 ± 2.7	138.3 ± 4.7	-19.24 ± 0.16	0.88 ± 0.06
2007ba	1.7	10.8 ± 0.2	49.1 ± 2.8	94.9 ± 3.0	-18.66 ± 0.10	0.54 ± 0.06
2007bc	1.1	10.6 ± 0.3	29.9 ± 1.8	100.0 ± 2.4	-19.32 ± 0.12	0.88 ± 0.06
2007bd	0.3	12.7 ± 0.3	12.3 ± 2.0	116.4 ± 4.4	-19.28 ± 0.10	0.88 ± 0.06
2007bm	3.6	10.6 ± 0.2	24.7 ± 1.5	101.8 ± 2.7	-19.59 ± 0.30	0.90 ± 0.06
2007ca	0.9	11.0 ± 0.3	10.5 ± 2.0	89.9 ± 2.4	-19.56 ± 0.18	1.06 ± 0.06
2007le	2.0	11.9 ± 0.3	12.7 ± 1.6	113.9 ± 3.5	-19.07 ± 0.40	1.02 ± 0.06
2007ol	2.4	11.1 ± 0.3	16.2 ± 4.1	77.2 ± 8.2	-18.87 ± 0.06	0.70 ± 0.07
2007on	1.4	11.2 ± 0.1	41.3 ± 3.9	117.4 ± 4.2	-19.05 ± 0.35	0.57 ± 0.06
2007ux	1.4	11.1 ± 0.2	54.3 ± 7.1	124.1 ± 9.6	-18.48 ± 0.10	0.59 ± 0.06
2008C	4.3	10.6 ± 0.4	15.5 ± 5.7	65.1 ± 9.2	-19.38 ± 0.21	0.95 ± 0.07
2008O	0.4	14.4 ± 0.7	47.6 ± 3.4	176.3 ± 4.7	-18.69 ± 0.12	0.65 ± 0.06
2008R	1.4	10.7 ± 0.2	49.4 ± 1.3	126.7 ± 1.9	-18.48 ± 0.18	0.59 ± 0.06
2008bc	3.6	11.7 ± 0.2	12.6 ± 3.3	105.0 ± 6.5	-19.45 ± 0.16	1.05 ± 0.06
2008bf	1.4	11.3 ± 0.2	10.7 ± 2.4	79.0 ± 4.8	-19.43 ± 0.10	1.02 ± 0.06
2008bq	0.5	10.6 ± 0.2	12.4 ± 3.4	92.0 ± 4.6	-19.74 ± 0.16	1.16 ± 0.06

Table 3.1 (cont'd)

SN	Epoch [days past $B_{\max}$]	$v_{\rm Si\,II}$ [1000 km s^{-1}]	pEW(Si II λ5972) [Å]	pEW(Si II λ6355) [Å]	M_B [mag]	$s_{\rm BV}$
2008cf	2.5	10.2 ± 0.3	7.9 ± 2.5	55.6 ± 5.0	-19.56 ± 0.10	1.12 ± 0.07
2008fl	3.2	10.7 ± 0.2	38.3 ± 4.9	112.3 ± 4.6	-19.38 ± 0.18	0.85 ± 0.06
2008fp	0.1	11.0 ± 0.1	12.4 ± 3.9	73.9 ± 4.7	-19.92 ± 0.39	1.08 ± 0.06
2008fr	4.6	10.1 ± 0.4	19.3 ± 3.6	93.3 ± 6.3	-19.35 ± 0.12	1.06 ± 0.06
2008fw	5.6	10.1 ± 0.2	13.0 ± 3.9	61.2 ± 5.7	-19.46 ± 0.27	1.11 ± 0.06
2008gg	4.9	12.7 ± 0.4	10.0 ± 3.3	149.8 ± 6.7	-19.47 ± 0.15	1.11 ± 0.07
2008gl	2.0	11.9 ± 0.4	20.6 ± 2.4	115.0 ± 3.3	-19.12 ± 0.09	0.85 ± 0.06
2008go	0.1	13.1 ± 0.4	8.1 ± 2.3	154.0 ± 4.9	-19.32 ± 0.10	0.91 ± 0.06
2008gp	6.8	11.0 ± 0.3	15.4 ± 3.0	43.7 ± 3.1	-19.42 ± 0.10	0.97 ± 0.06
2008hj	6.4	13.0 ± 0.5	16.0 ± 2.1	103.3 ± 3.5	-19.30 ± 0.09	1.01 ± 0.06
2008hu	3.7	12.4 ± 0.3	15.9 ± 4.7	127.0 ± 9.4	-19.04 ± 0.11	0.79 ± 0.06
2008hv	1.3	10.8 ± 0.4	21.3 ± 2.1	96.7 ± 3.3	-19.12 ± 0.17	0.85 ± 0.06
2008ia	2.3	11.4 ± 0.4	20.0 ± 2.4	105.9 ± 4.2	-19.31 ± 0.16	0.84 ± 0.06
2009D	0.7	9.7 ± 0.3	4.3 ± 5.3	77.5 ± 6.5	-19.65 ± 0.12	1.19 ± 0.06
2009Y	2.3	14.4 ± 0.2	8.6 ± 2.0	155.6 ± 4.7	-19.62 ± 0.23	1.19 ± 0.06
2009aa	0.1	10.9 ± 0.4	20.1 ± 2.9	63.2 ± 2.6	-19.40 ± 0.10	0.91 ± 0.06
2009ab	2.8	10.7 ± 0.2	28.3 ± 4.1	105.2 ± 5.4	-19.04 ± 0.23	0.87 ± 0.06
2009ad	1.2	10.2 ± 0.2	9.9 ± 3.8	67.4 ± 3.6	-19.52 ± 0.12	1.01 ± 0.06
2009ag	1.4	10.3 ± 0.2	21.4 ± 4.5	107.8 ± 5.4	-19.26 ± 0.31	0.96 ± 0.06
2009cz	0.1	10.1 ± 0.3	11.9 ± 3.2	75.6 ± 4.2	-19.59 ± 0.13	1.19 ± 0.06

Table 3.1 (cont'd)

SN	Epoch [days past $B_{\max}$]	$v_{\mathrm{Si\,II}}$ [1000 km s^{-1}]	pEW(Si II λ5972) [Å]	pEW(Si II λ6355) [Å]	M_B [mag]	s_{BV}
2009ds	4.0	12.6 ± 0.1	14.8 ± 2.3	71.8 ± 3.8	-19.65 ± 0.13	1.12 ± 0.06
2009le	-4.7	12.5 ± 0.2	8.2 ± 1.5	79.2 ± 2.6	-19.18 ± 0.16	1.16 ± 0.06
2011iv	5.0	10.9 ± 0.2	40.8 ± 2.2	88.6 ± 2.0	-19.67 ± 0.34	0.64 ± 0.05
2011jh	4.8	12.8 ± 0.2	35.0 ± 2.8	126.1 ± 3.9	-19.31 ± 0.29	0.80 ± 0.05
2012aq	6.5	11.2 ± 0.4	27.6 ± 2.2	105.5 ± 3.5	-19.53 ± 0.15	0.99 ± 0.05
2012bl	1.4	14.7 ± 0.3	2.3 ± 1.2	90.5 ± 2.4	-19.34 ± 0.14	1.11 ± 0.05
2012fr	0.9	12.1 ± 0.2	6.5 ± 1.7	70.5 ± 3.1	-19.46 ± 0.40	1.12 ± 0.05
2012gm	5.8	10.4 ± 0.3	24.1 ± 4.0	101.5 ± 7.5	-19.46 ± 0.17	0.98 ± 0.05
2012hl	3.5	12.9 ± 0.2	13.3 ± 3.4	142.0 ± 7.3	-18.86 ± 0.27	0.92 ± 0.06
2012hr	5.6	12.6 ± 0.3	15.3 ± 1.3	134.7 ± 2.1	-19.19 ± 0.29	0.96 ± 0.05
2012ht	1.6	10.9 ± 0.2	27.5 ± 1.9	115.2 ± 2.7	-19.06 ± 0.61	0.85 ± 0.05
2012ij	0.2	10.8 ± 0.3	53.3 ± 4.0	121.5 ± 4.7	-18.13 ± 0.21	0.53 ± 0.05
2013E	5.9	12.6 ± 0.3	3.8 ± 1.5	65.2 ± 3.0	-19.90 ± 0.25	1.12 ± 0.05
2013fy	5.1	10.8 ± 0.3	14.4 ± 1.4	93.4 ± 3.1	-19.67 ± 0.12	1.19 ± 0.05
2013gy	3.1	10.2 ± 0.3	28.9 ± 4.2	114.1 ± 4.3	-19.39 ± 0.18	0.89 ± 0.05
2014I	1.3	11.3 ± 0.2	23.5 ± 3.7	98.7 ± 4.7	-19.48 ± 0.10	0.90 ± 0.05
2014dn	3.0	10.4 ± 0.6	70.9 ± 6.2	140.1 ± 4.7	-17.68 ± 0.13	0.45 ± 0.05
1998dh	1.4	12.4 ± 0.2	26.1 ± 2.2	124.6 ± 2.0	-19.34 ± 0.22	$\cdots$
1998dm	2.0	11.0 ± 0.2	11.2 ± 1.9	73.2 ± 2.2	-18.68 ± 0.28	$\cdots$
1999cp	4.0	10.6 ± 0.2	22.7 ± 1.4	104.6 ± 2.4	-19.36 ± 0.18	$\cdots$

Table 3.1 (cont'd)

SN	Epoch [days past $B_{\max}$]	$v_{\rm Si\,II}$ [1000 km s^{-1}]	pEW(Si II λ5972) [Å]	pEW(Si II λ6355) [Å]	M_B [mag]	$s_{\rm BV}$
1999dq	0.5	11.1 ± 0.1	9.3 ± 1.6	44.8 ± 2.3	-19.82 ± 0.15	···
1999gp	0.1	11.1 ± 0.4	8.2 ± 2.2	53.8 ± 3.0	-19.61 ± 0.09	···
2000cx	0.0	11.8 ± 0.3	5.9 ± 2.2	40.3 ± 3.8	-19.31 ± 0.24	···
2000dn	1.8	9.7 ± 0.3	13.3 ± 3.2	102.4 ± 3.6	-19.15 ± 0.07	···
2000dr	0.0	10.3 ± 0.3	84.6 ± 3.2	131.4 ± 3.2	-18.52 ± 0.11	···
2000fa	0.4	11.9 ± 0.3	10.7 ± 2.3	83.4 ± 4.8	-19.54 ± 0.10	···
2001V	2.5	11.4 ± 0.1	13.2 ± 2.1	57.0 ± 1.9	-19.70 ± 0.13	···
2001en	0.4	12.6 ± 0.3	9.4 ± 1.5	135.0 ± 3.1	-18.86 ± 0.15	···
2001ep	0.4	10.7 ± 0.2	31.9 ± 2.1	111.7 ± 2.1	-19.24 ± 0.15	···
2002bo	2.0	13.4 ± 0.2	10.5 ± 5.1	150.4 ± 9.3	-19.31 ± 0.29	···
2002cr	0.3	10.1 ± 0.1	19.3 ± 1.8	104.6 ± 2.0	-19.34 ± 0.18	···
2002dj	5.0	14.0 ± 0.3	6.3 ± 1.3	151.0 ± 2.3	-19.26 ± 0.21	···
2002dl	6.6	12.5 ± 0.4	40.2 ± 4.5	90.1 ± 4.9	-18.28 ± 0.12	···
2002eb	3.0	10.1 ± 0.2	9.9 ± 2.1	65.0 ± 4.1	-19.60 ± 0.08	···
2002er	0.0	12.0 ± 0.0	24.2 ± 0.9	115.2 ± 1.2	-19.34 ± 0.21	···
2002fk	1.4	9.8 ± 0.3	18.6 ± 1.7	80.7 ± 2.2	-19.40 ± 0.24	···
2002ha	1.0	11.3 ± 0.3	30.6 ± 2.6	111.3 ± 2.8	-19.16 ± 0.14	···
2002he	0.6	12.6 ± 0.2	20.2 ± 1.1	118.3 ± 1.5	-19.01 ± 0.09	···
2003W	0.4	15.2 ± 0.5	6.4 ± 1.7	108.2 ± 3.5	-19.44 ± 0.10	···
2003Y	2.0	11.3 ± 0.4	71.8 ± 8.2	104.1 ± 5.9	-17.48 ± 0.14	···

Table 3.1 (cont'd)

SN	Epoch [days past $B_{\max}$]	$v_{\mathrm{Si\,II}}$ [1000 km s^{-1}]	pEW(Si II λ5972) [Å]	pEW(Si II λ6355) [Å]	M_B [mag]	s_{BV}
2003cg	0.2	11.1 ± 0.3	24.0 ± 2.5	96.1 ± 3.1	-19.49 ± 0.32	···
2003gn	3.3	13.2 ± 0.9	40.8 ± 9.3	146.8 ± 10.3	-18.81 ± 0.11	···
2003gt	4.8	11.3 ± 0.4	33.6 ± 3.6	76.8 ± 4.1	-19.47 ± 0.13	···
2004at	0.2	10.7 ± 0.2	10.5 ± 1.9	92.6 ± 2.3	-19.46 ± 0.09	···
2004dt	0.0	16.0 ± 0.0	21.5 ± 1.9	167.0 ± 3.0	-19.83 ± 0.10	···
2005cf	0.0	10.3 ± 0.3	17.0 ± 1.1	89.7 ± 1.2	-19.41 ± 0.25	···
2005de	1.7	10.7 ± 0.3	30.2 ± 3.2	102.9 ± 3.5	-19.06 ± 0.13	···
2006cp	3.7	14.9 ± 0.5	15.3 ± 2.4	159.0 ± 3.0	-19.27 ± 0.11	···
2006gr	0.3	11.3 ± 0.4	3.3 ± 2.5	65.0 ± 4.2	-19.41 ± 0.09	···
2006le	2.4	11.4 ± 0.2	13.0 ± 2.4	93.3 ± 5.2	-19.85 ± 0.11	···
2006lf	2.6	11.7 ± 0.1	25.7 ± 3.0	100.4 ± 4.4	-19.39 ± 0.16	···
2007ci	0.8	12.3 ± 0.4	50.0 ± 2.6	126.4 ± 2.8	-18.58 ± 0.11	···
2008ec	0.4	10.8 ± 0.7	36.7 ± 2.8	119.3 ± 3.1	-19.21 ± 0.13	···

Note. — s_{BV} values are not provided for the Zheng et al. (2018) subset and are not used in this work.

We use values of M_B that are corrected for MW extinction and host-galaxy extinction directly from Zheng et al. (2018) Table 1 with no remeasurement. These photometric values from Zheng et al. (2018) are supplemented with spectroscopy from the Open Supernova Catalog (Guillochon et al., 2017). The spectral epoch for the measurement is always chosen to be the epoch that is the closest available to the time of B-band maximum. The values of M_B determined by Zheng et al. (2018) were made with $H_0 = 70$ km s^{-1} Mpc^{-1} and host extinction was estimated using `MLCSk2` fitting (Jha et al., 2007) with R_V held fixed at a value of 1.8. Their sample has a maximum redshift of $z = 0.039$, but excluded all SNe Ia with host $E(B-V) > 0.3$ mag. K-corrections were not included. For CSP I SNe Ia in the Zheng et al. (2018) sample, we use the `SNooPy` determinations. The values used from the unique Zheng et al. (2018) objects included in this analysis are also provided in the bottom section of Table 3.1.

3.3. Methods

3.3.1. Velocities and Pseudo-Equivalent Widths

We use a modified version of the `Spextractor` code (Papadogiannakis, 2019) to measure velocities and pEWs. We modified this code to allow for the binned downsampling of spectra, with the constraint that the number of photons is conserved in each bin, in order to reduce computational cost. We have also made modifications that produce a more representative Gaussian process regression (GPR) model for a given spectrum. In the original program, the posterior was sampled at wavelengths given to the prior, whereas now we sample the posterior at uniformly spaced points at a higher resolution than the prior to account for point-to-point variance. Flux uncertainties are also added to the Matérn 3/2 GPR kernel in quadrature when available. See Appendix A for a more detailed technical description of `Spextractor`.

Figure 3.1 shows the basic function of the modified `Spextractor` code on ASASSN-14mf at epoch ~ 5 days past B-band maximum. The red curve is the mean function obtained using GPR, and the vertical dashed lines show the position of the flux minima of identified features. We identify the wavelength of the flux minima of the features as the shift due to the pseudo-photospheric velocities via the relativistic Doppler formula. The light-blue shading shows the area used to calculate pEWs. Only specified features are marked with `Spextractor`, and in this work, special care is taken only in retrieving accurate measurements for the Si II $\lambda5972$ and $\lambda6355$ lines. In order to obtain velocities, we assume that the identified flux minima correspond to the blueshifted absorption of Si II $\lambda5972$ and Si II $\lambda6355$ that is uncontaminated by line-blending. We assume a linear continuum approximation between maxima of selected wavelength ranges (as indicated by large black circles in Figure 3.1) to define features and their pEWs. The $v_{\text{Si II}}$, pEW(Si II $\lambda5972$), and pEW(Si II $\lambda6355$) values calculated by `Spextractor` for each SN Ia are given in Table 3.1.

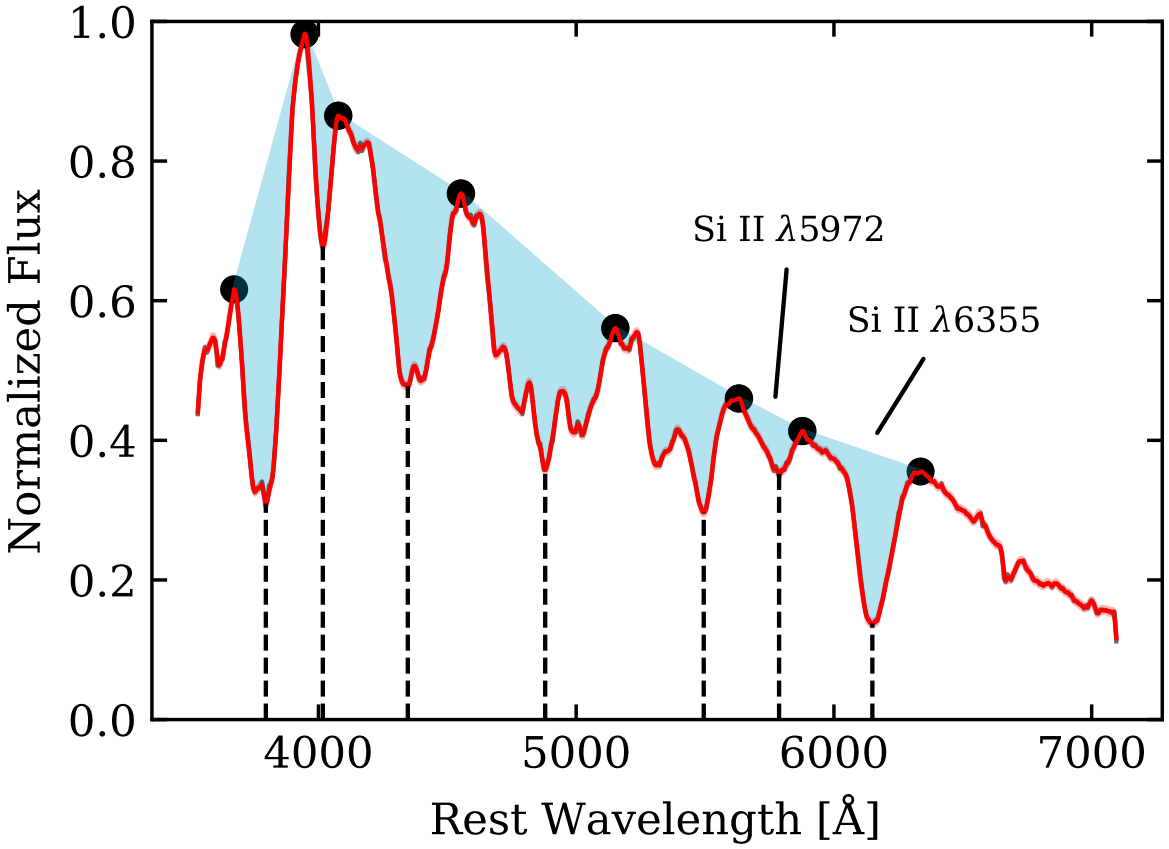


Figure 3.1: Representative `Spextractor` function acting on a spectrum of ASASSN-14mf at epoch ∼5 days past B-band maximum. The red line indicates the mean function of the corresponding GPR. With a large number of data points, the GPR mean function fits well to the original spectrum, which is displayed, albeit difficult to discern from the mean. The blue area represents the area integrated in pEW calculations for each line. Note that only the Si II λ5972 and λ6355 lines are studied in this work, so no special care was taken in extracting other features accurately. The large black dots indicate the points used to estimate the pseudo-continuum indicated by the solid blue line.

3.3.2. Cluster Analysis

We invoke a distribution-based cluster analysis using Gaussian mixture models (GMMs) to support grouping objects with related properties (Day, 1969; McLachlan & Peel, 2000). These GMMs are expectation-maximization algorithms that iteratively calculate which input properties maximize the likelihood of a set of m-dimensional (m-D) Gaussian distributions fitting a given training set of data. After this fit is calculated, a set of probability distributions is given, such that the probability of a point being associated with each of the n distributions can be determined. This method is chosen in order to describe group assignments probabilistically since there appears to be

a continuous distribution of multiple group clusters in our sample of data. There is no clear way of separating discrete groups with reasonable confidence — for example, ensuring that each group is disjoint with rigid boundaries. We show that by using Gaussian distributions and, therefore, a well-defined deviation σ for each group, the determined groups overlap within 3σ, and so we would consider the groups to be connected or non-discrete. The probability of a SN Ia belonging to any single group may be comparable to that for another group, even though it is within 3σ of either group's mean position.

The assumption is also made here that each of the measured quantities M_B, $v_{\mathrm{Si\ II}}$, pEW(Si II $\lambda6355$), and pEW(Si II $\lambda5972$) can be represented as jointly Gaussian-distributed to first approximation. This approximation is assumed only to have some measure of similarity between these four properties. These quantities are then used as input parameters to the different GMMs calculated and shown in Section 3.4, which yields n probabilities of each object being associated with each group. The groups from any GMM therefore describe similarity between the input quantities from which the GMM was calculated.

For each of our GMMs, the number of clusters (groups), n, assigned has been determined first with the assumption that each GMM that includes the two pEW quantities must have at least $n = 4$ clusters. Since the original work of Branch et al. (2006) identified four groups with several distinct characteristics, we do not consider clustering data involving pEW(Si II $\lambda6355$) and pEW(Si II $\lambda5972$) with fewer than four groups.

In general, the number of clusters used for each GMM may be decided based on standard testing that determines which value of n provides the model that best fits the data. For perfectly Gaussian clusters, n may be determined based on the GMM with the lowest Bayesian information criterion (BIC) value (Kass & Raftery, 1995). BIC values of each model presented in this work (see Section 3.4) were calculated for GMMs with $n = 1$ to $n = 6$. Each one of these calculations are the average of 20 single-trial calculations of the BIC for the model with a given n, which is necessary because the GMM algorithm operates with randomized initial conditions, potentially leading to different grouping systems for large values of n. Figure 3.2 (left panel) shows $\Delta(\mathrm{BIC})$ as a function of n for different sets of input parameters, where $\Delta(\mathrm{BIC})$ references the BIC of the $n = 1$ GMM such that $\Delta(\mathrm{BIC}) = \mathrm{BIC}(n) - \mathrm{BIC}(n = 1)$. Therefore, we choose n based on the model yielding the smallest $\Delta(\mathrm{BIC})$. Results from Figure 3.2 show that all models with pEW(Si II $\lambda6355$) and pEW(Si II $\lambda5972$) ("Branch" in the legend, including the 4-D model) prefer a value of $n = 2$. However, because it is assumed that $n \geq 4$ for these models, we use $n = 4$ for these GMMs.

Figure 3.2 (right panel) shows the mean Silhouette score s (Rousseeuw, 1987; de Souza et al., 2017) of the different GMMs for $n = 2$ to $n = 6$. The values of s were calculated using k-means clustering (MacQueen, 1967). While the Silhouette score is a good measure of how well the data is separable into clusters, we are guided by the $\Delta(\mathrm{BIC})$ in determining which value of n to use for

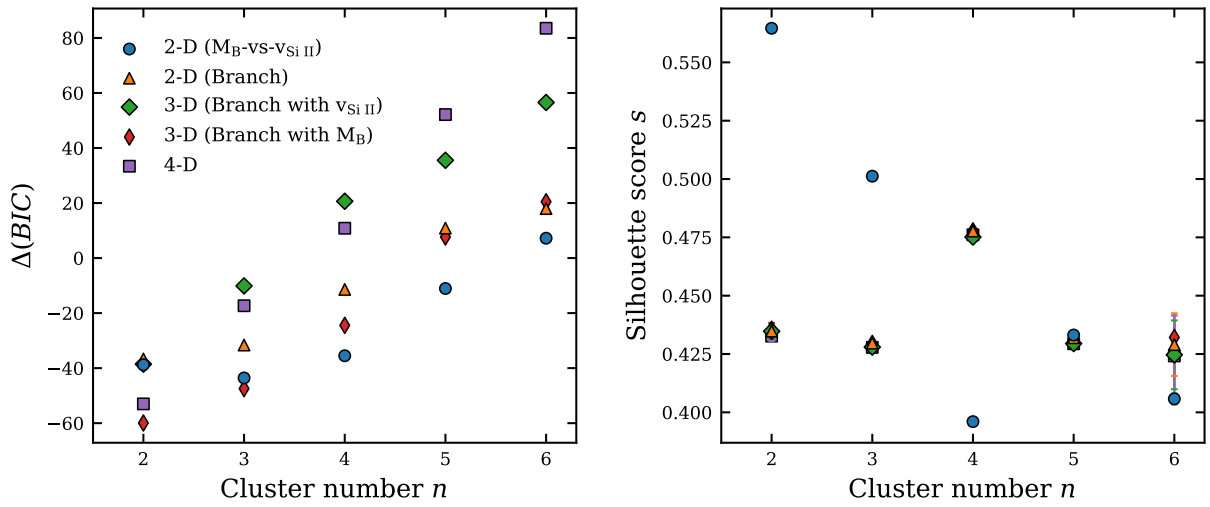


Figure 3.2: Left panel: $\Delta(\mathrm{BIC})$ versus the cluster number n for each GMM presented in this work. $\Delta(\mathrm{BIC})$ is referenced to the BIC of the $n = 1$ GMM. For all dimensionalities except the [M_B, $v_{\mathrm{Si\ II}}$] input set, models with $n < 4$ are not used, in order to follow historical precedent (Branch et al., 2006). Within these restrictions, we find the $n = 4$ case to have the smallest $\Delta(\mathrm{BIC})$; thus, these GMMs with $n = 4$ are used for this study (see Section 3.4.1 and Section 3.4.2). For the [M_B, $v_{\mathrm{Si\ II}}$] GMM, the $n = 3$ model has the smallest $\Delta(\mathrm{BIC})$ value and is therefore used to define M_B-vs-$v_{\mathrm{Si\ II}}$ groups discussed in Section 3.4.3. Right panel: the Silhouette score is displayed for each model having $n = 2$ to $n = 6$ clusters. See text for further description. The plot symbols and colors are shared between panels.

each model. This is because we establish clusters in this sample based on GMMs and not k-means clustering. It is interesting, however, that the Silhouette score consistently favors $n = 4$ (s closer to 1 is more preferable) for any model involving both pEW(Si II λ6355) and pEW(Si II λ5972).

Neither of these statistical merits indicate considerably good model fits, which likely indicates that the description of maximum-light properties of SNe Ia using Gaussian distributions is not ideal. However, the GMM's parametric nature, which allows for a maximized-likelihood approach, recommends itself over other, non-parametric methods. In particular, it provides stable probabilistic results, which can be interpreted using standard methods (de Souza et al., 2017) such as using the BIC or Silhouette score.

The dimensionality, m, of each GMM is determined by the number of SN Ia properties included in training the GMM, which is independent of the number of GMM components n. In training each GMM, we do not weight any points in the GMMs based on the uncertainty of any quantity.

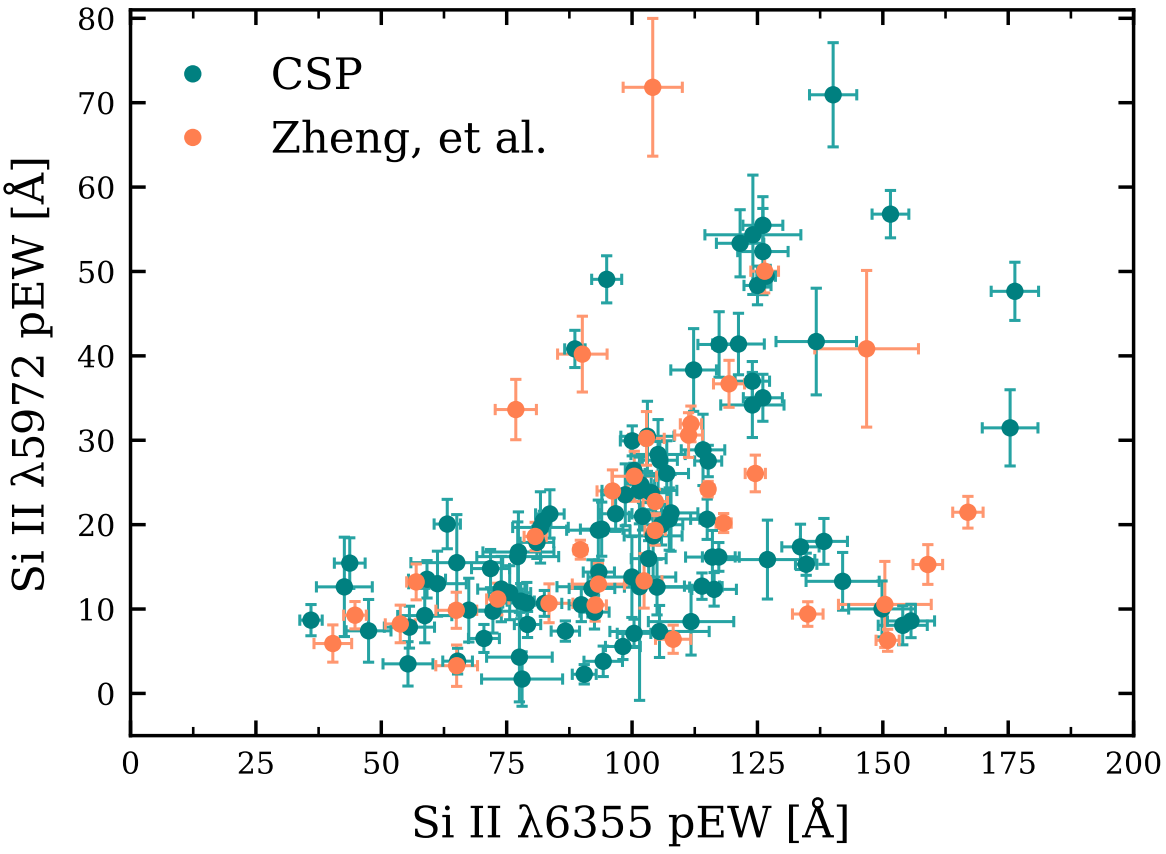


Figure 3.3: The Branch diagram that includes both the Zheng et al. (2018) and the CSP I+II samples. We see the expected Branch diagram trend with nine extended cool objects with pEW(Si II λ5972) $\geq$ 50 Å. Overall, pEW(Si II λ6355) in this sample exhibits a mean uncertainty of $\pm$ 4.35 Å and pEW(Si II λ5972) exhibits that of $\pm$ 3.02 Å.

3.4. Results

3.4.1. Branch Clustering

Figure 3.3 shows the Branch diagram obtained for the CSP I+II and Zheng et al. (2018) samples. With this large sample of data, it appears that there are no completely disconnected groupings: a similar and expected result compared to that obtained by Branch et al. (2006). We do see the expected pEW(Si II λ6355) $\leq$ 200 Å range. However, although the majority of the objects fall in the expected pEW(Si II λ5972) $\leq$ 50 Å range, we find a few extended CLs with pEW(Si II λ5972) $\geq$ 50 Å. Due to the lack of a discrete clustering found between these groups, we perform a cluster analysis statistically. Using this entire data set, we create a 2-D GMM in [pEW(Si II λ5972), pEW(Si II λ6355)] space with $n = 4$ components, as it is seen from Figure 3.2 that $n = 4$ minimizes Δ(BIC) for $n \geq 4$. For this and every other GMM that includes pEW(Si II λ5972) and

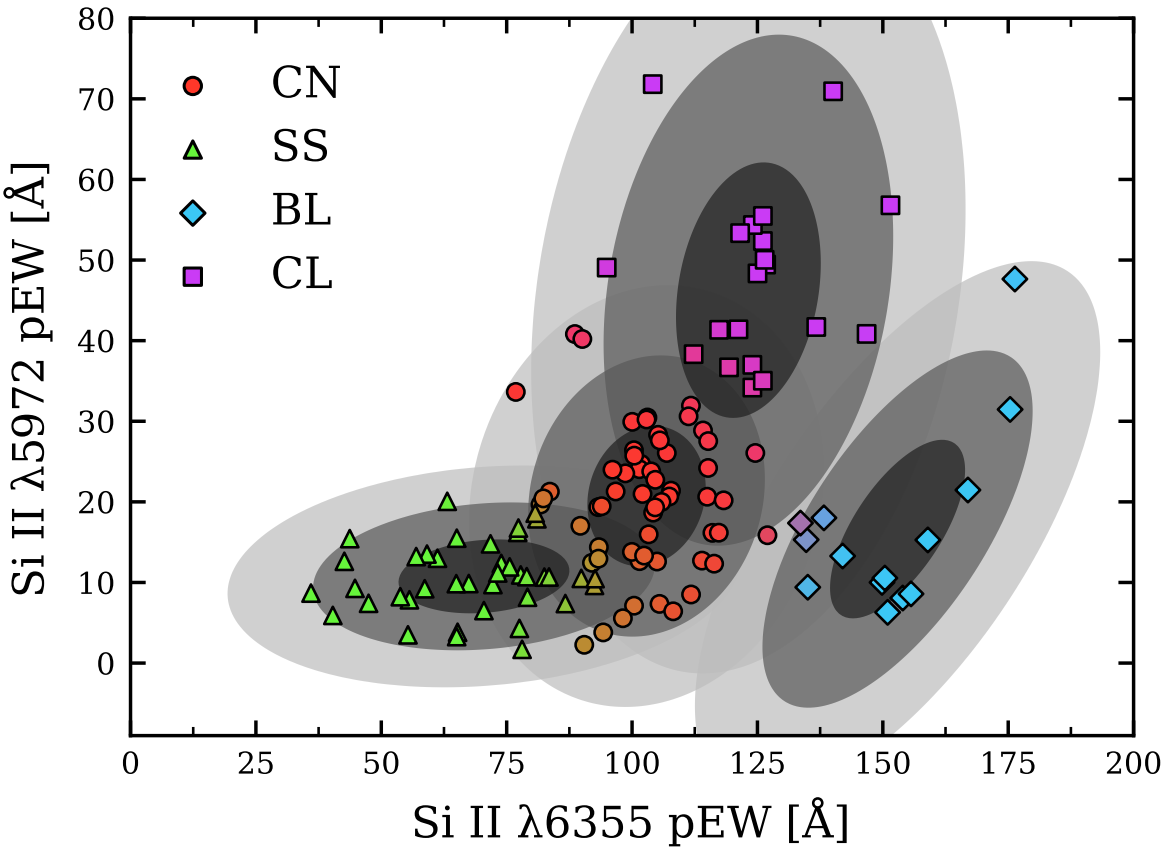


Figure 3.4: The Branch diagram colored with the [pEW(Si II λ5972), pEW(Si II λ6355)] GMM ($n = 4$ components). Contours indicate 1-, 2-, and 3σ from the mean of each group determined by the GMM. Different colors indicate group membership as a probability distribution of each point belonging to a given group, and different symbols indicate the group corresponding to the highest likelihood of membership.

pEW(Si II λ6355) in its input parameter space, this is also the case. The right panel of Figure 3.2 shows that the Silhouette score is maximized for these models at $n = 4$ and thus supports this number of clusters for this sample. This GMM is displayed in Figure 3.4. Different colors indicate group membership as a probability distribution of each point belonging to a given group. This is done using a linear mapping from probability space onto color space. Different symbols indicate the group corresponding to highest likelihood of membership. Contours correspond to 1-, 2-, and 3σ from the mean of each group determined by the GMM and is a representation of the covariance of the GMM groups. This figure clearly identifies four groups that indeed correspond to the originally identified Branch groups: core-normals, shallow-silicons, cools, and broad-lines. Note that, in this and future figures with contours representing group membership, we choose to exclude error bars for visual clarity, as the errors are always the same as those in Figures 3.3 and 3.9.

3.4.2. Higher-Dimensional Clustering

Although the robustness of the Branch groups is seen with a simple 2-D GMM, in intermediate areas between these groups many objects have comparable probabilities of membership to more than one group. For this reason, to achieve more certainty in membership to any single Branch group, we include additional input properties that are related to the two pEW parameters in the 2-D GMM. This is expected to provide further constraints for our sample and possibly provide insight into the Branch groups' relationship with the M_B-vs-$v_{\mathrm{Si\ II}}$ diagram, which is discussed in Section 3.4.3.

3.4.2.1. Inclusion of $v_{\mathrm{Si\ II}}$

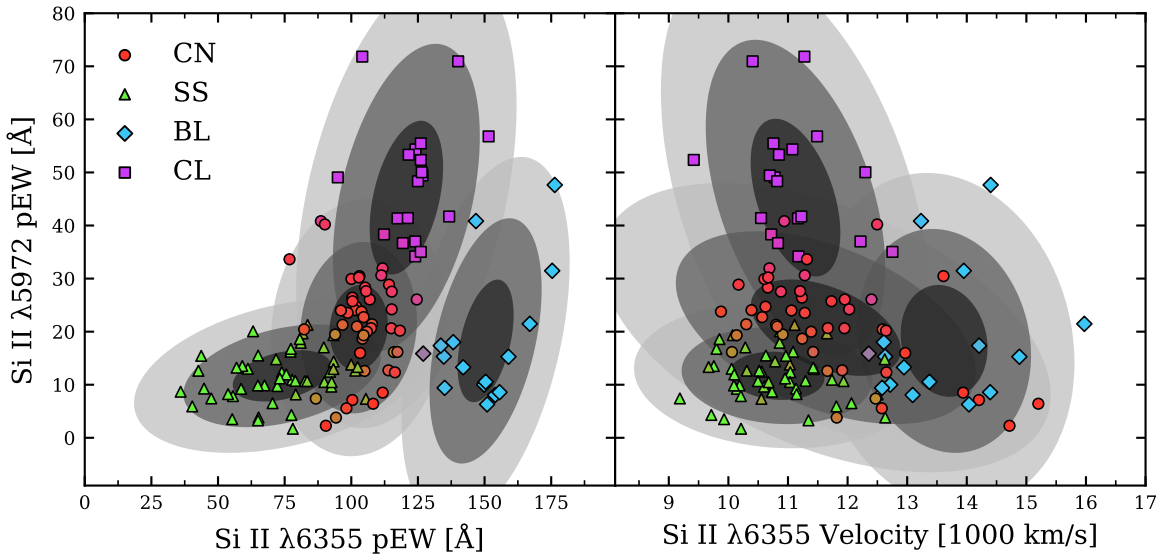


Figure 3.5: A 3-D GMM analysis of [$v_{\mathrm{Si\ II}}$, pEW(Si II λ5972), pEW(Si II λ6355)] ($n = 4$ components). The colors in each panel represent group membership and are shared between panels.

We first look at a 3-D GMM in [$v_{\mathrm{Si\ II}}$, pEW(Si II λ5972), pEW(Si II λ6355)] with $n = 4$ components (see Figure 3.2). Figure 3.5 shows the entire scope of this model. The 2-D contours in each panel are associated with the covariance of every determined group distribution for the respective 2-D slice. The left panel is a Branch diagram of the sample, and it again clearly separates the sample into Branch-like groups.

It appears that the inclusion of $v_{\mathrm{Si\ II}}$ information mostly alters the membership likelihoods in intermediate areas surrounding the CN group. This leads to changes in the covariances shown by the contours in Figure 3.5 as well as the general probability distribution (coloring). Comparing with those from the 2-D model in Figure 3.4, we find that the 3σ contour area of the CN group decreased by 13.9%, that of the BL group decreased by 1.8%, and of the CL group decreased by

18.1%, however the 3σ contour area of the SS group increased by 8.5%.

An important difference is that in areas between groups there is a steeper probability gradient in the 3-D model than the 2-D model in pEW space. More specifically, we refer to this probability gradient as the gradient at a given point of the group membership probability distribution projected into the 2-D subspace, which in this case is the pEW space. Qualitatively, this effect would narrow the shape or size of a group's contours, which corresponds to the aforementioned decreases in the 3σ areas of the CN, BL, and CL groups. In intermediate areas between groups, it more concretely defines the group membership of many objects that previously showed a nearly equal tendency toward two or more groups. This behavior is expected since naively one would expect a correlation between the pEW and velocity of the Si II $\lambda6355$ line, even though they are independently measured.

It is therefore seen that using a GMM to measure the similarity between the two quantities enforces a constraint that quantitatively defines groups. Most noticeably the contours for the CN and CL groups have reduced in size, meaning there is a narrower region in pEW space in which CNs are expected to lie. This illustrates that including the additional $v_{\mathrm{Si\ II}}$ parameter in the GMM more sharply defines Branch group membership and leads to more certainty in assignment compared to a GMM based on only pEW information.

3.4.2.2. Inclusion of M_B

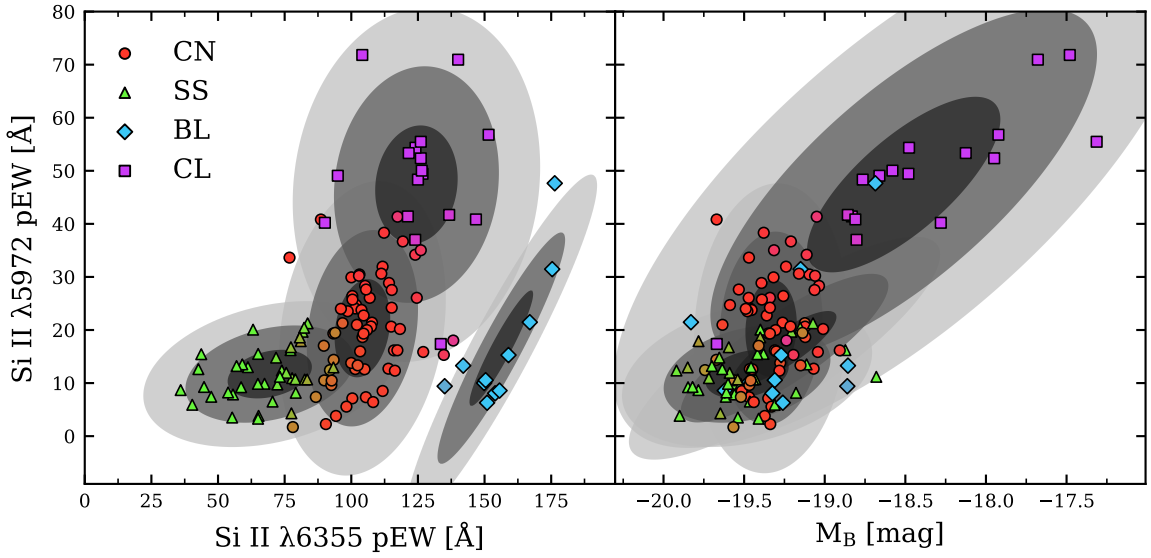


Figure 3.6: A 3-D GMM analysis of [M_B, pEW(Si II $\lambda5972$), pEW(Si II $\lambda6355$)] ($n = 4$ components).

We also investigate the 3-D GMM in [M_B, pEW(Si II $\lambda5972$), pEW(Si II $\lambda6355$)] with $n = 4$ components (see Figure 3.2). Figure 3.6 again shows the different slices of this model. In the left panel, we again find that the Branch diagram shown has more concretely defined Branch groups

than in the 2-D pEW space alone. Again comparing with the contours in the 2-D model displayed in Figure 3.4, we find that all 3σ contour areas decrease in size: CN (7.8%), SS (10.3%), BL (59.2%), and CL (23.3%). Clearly BL membership alters more drastically when M_B information is included in the GMM. The contours of the BLs are much narrower, so membership is generally contained in a narrower region in pEW space. In fact, with the inclusion of M_B the BL group becomes an almost completely distinguishable group, indicating that there may in fact be something distinct about the progenitor system or explosion mechanism that produces BLs.

It is clear from this GMM that using a Gaussian distribution in this cluster analysis is not a perfect method in predicting cluster membership. We begin to see some unexpected behavior in the GMM membership determination. For example, LSQ13aiz (at pEW(Si II $\lambda6355$) ≈ 133 and pEW(Si II $\lambda5972$) ≈ 17) indeed appears too bright to be a cool object as is suggested by the GMM. We see, then, that variations in one of the input properties can lead to outlier behavior. As will be seen in Section 3.4.2.3, this problem is partially solved with the inclusion of additional information.

The right panel also shows a strong functional relation between M_B and pEW(Si II $\lambda5972$). We fit a quadratic to this relation, and this is shown in Figure 3.7. It is interesting that this correlation, along with the Branch group classification, provides a rough M_B approximation that is purely based on spectroscopic information. Others have previously noted that pEW(Si II $\lambda5972$) and $\Delta m_{15}(B)$ are correlated (Hachinger et al., 2008; Folatelli et al., 2013), and we find that this effect is quite robust. Given the relationship between M_B and $\Delta m_{15}(B)$, this correlation is essentially a spectroscopic variant of the Phillips relation.

3.4.2.3. Inclusion of Both $v_{\text{Si II}}$ and M_B

Because we see a strong correlation in [M_B, pEW(Si II $\lambda5972$)], we attempt to constrain the Branch groups further by creating a 4-D GMM in [M_B, $v_{\text{Si II}}$, pEW(Si II $\lambda5972$), pEW(Si II $\lambda6355$)] with $n = 4$ components (see Figure 3.2). We show this Branch diagram in the left panel of Figure 3.8, which again is colored by this 4-D GMM. This GMM indeed still produces the four Branch groups. Even with all four dimensions there is still significant overlap between the CN and SS objects.

The inclusion of both M_B and $v_{\text{Si II}}$ constrains the groups further by similarly reducing the size of most contours associated with projected covariance, giving a more concrete assignment to more objects. This overall inclusion decreases the 3σ contour areas of the CN group (45.6%), the BL group (28.9%), and the CL group (28.7%). However, the 3σ contour area of the SS group is relatively unchanged, increasing by only 0.6%. In the projection shown, there is more overlap between groups than either 3-D model, which shows that, for our sample, M_B and $v_{\text{Si II}}$ contain independent information. That is to say, after including either quantity, including the fourth quantity will constrain the groups further.

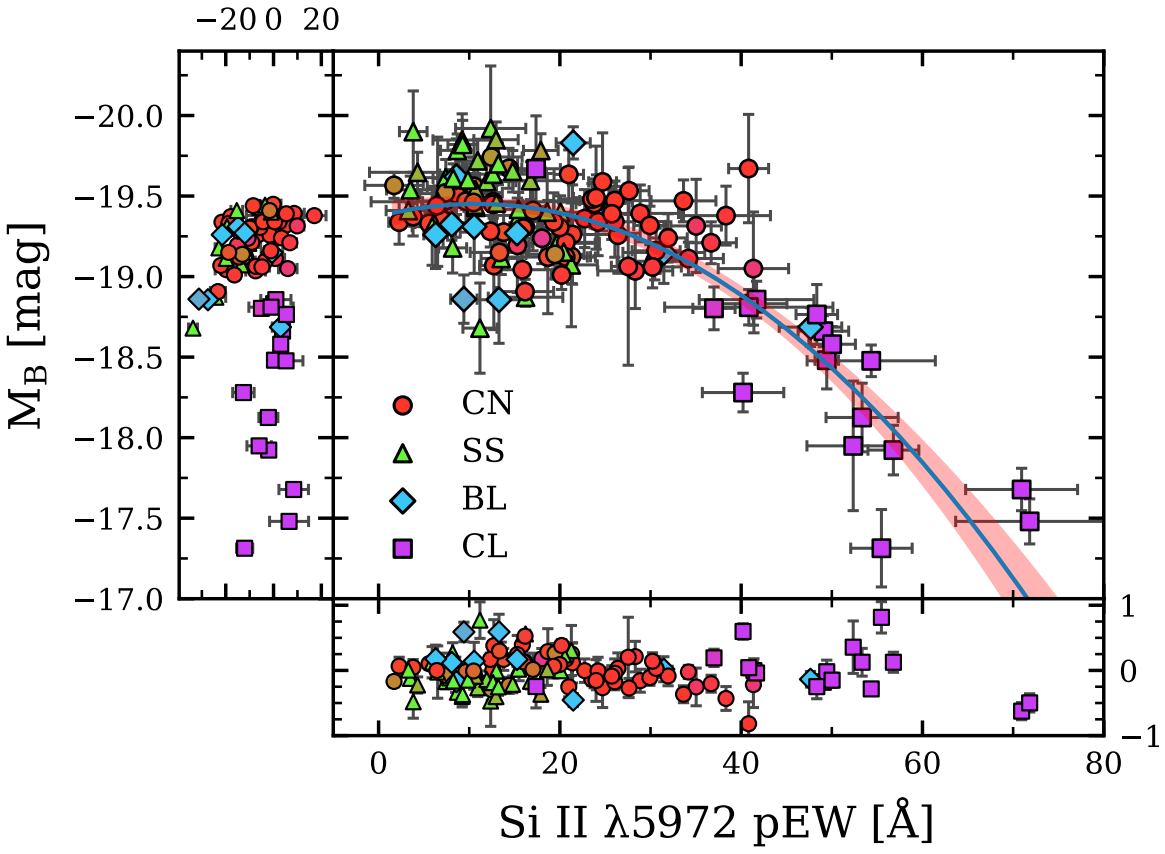


Figure 3.7: A quadratic fit of the M_B versus pEW(Si II λ5972) relation colored with the grouping from the 3-D GMM displayed by Figure 3.6. Residuals are given in panels to the left and below. The light red shaded region represents the uncertainty in the polynomial coefficients themselves. We find the quadratic fit coefficients (from highest to lowest order) to be $a = (6.623 \pm 1.168) \times 10^{-4}$, $b = (-1.425 \pm 0.657) \times 10^{-2}$, and $c = -19.37 \pm 0.07$.

Again we see that there are some objects that exhibit relatively substantial dispersion in one or more parameters that appear as outliers. For example, in the middle panel of Figure 3.8, SN 2008O (with $v_{\text{Si II}} \sim 14.4$ and pEW(Si II λ5972) ~ 47.6) and SN 2003gn (with $v_{\text{Si II}} \sim 13.2$ and pEW(Si II λ5972) ~ 40.8) seem most similar to the CL group, however they are deemed BL objects due to their pEW(Si II λ6355) and $v_{\text{Si II}}$ values.

Compared to the [pEW(Si II λ5972), pEW(Si II λ6355)] GMM defined in Figure 3.4, the inclusion of M_B and $v_{\text{Si II}}$ in the GMM allows us to constrain group membership in a way that is not apparent with only pEW(Si II λ5972) and pEW(Si II λ6355). That is to say, M_B and $v_{\text{Si II}}$ may be used to more concretely define Branch groups (remembering that we have a completely different definition for membership than was used by Branch et al., 2006), between which there would otherwise be more uncertainty and continuity in the probability distribution in pEW space.

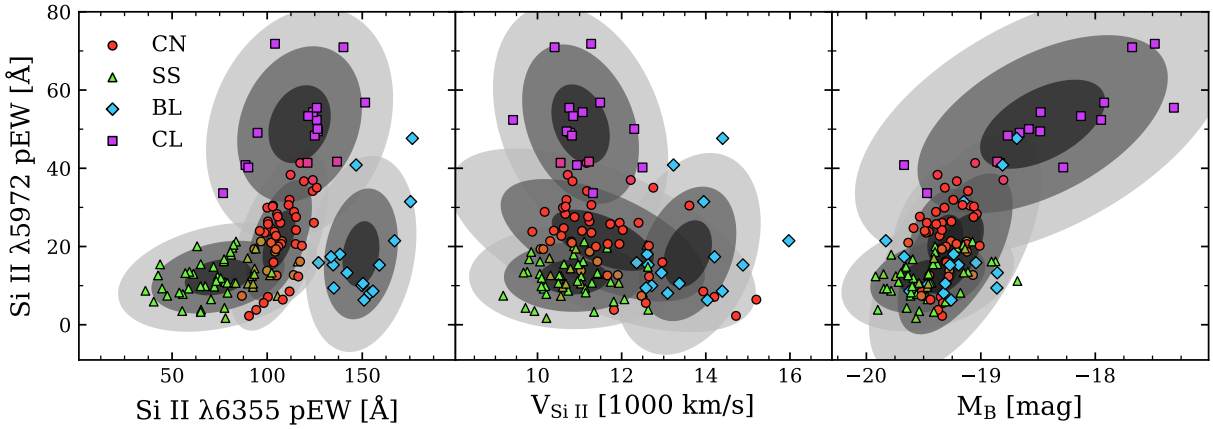


Figure 3.8: The Branch diagram (left) of the 4-D GMM analysis ($n = 4$ components) of all four parameters, showing the robustness of the Branch grouping system. The pEW(Si II λ5972)-versus-$v_{\rm Si\,II}$ (middle) and pEW(Si II λ5972)-versus-M_B (right) projections are shown and are color-coded by the same 4-D GMM group membership probabilities. The contours indicate 1-, 2-, and 3σ of group membership projected into each respective space. We see the groups have become much more constrained and display much overlap that could not be extracted solely with pEW information.

3.4.3. M_B-vs-$v_{\rm Si\,II}$ Clustering

Our version of the M_B-vs-$v_{\rm Si\,II}$ diagram with both Zheng et al. (2018) and CSP I+II samples is shown in Figure 3.9. The general structure of the original plot remains, although the CSP sample extends and fills in the low- to mid-velocity and dim portions of the plot. We see that the added continuity removes the original notion of a clear dichotomy between likely Chandrasekhar mass explosions and sub-Chandrasekhar mass helium detonations (Polin et al., 2019). We use a 2-D GMM in [M_B, $v_{\rm Si\,II}$] to account for this and to describe the M_B-vs-$v_{\rm Si\,II}$ group analysis statistically.

Figure 3.10 shows the M_B-vs-$v_{\rm Si\,II}$ diagram with this GMM with $n = 3$ components. This value of $n = 3$ minimizes Δ(BIC) (see Figure 3.2). While the Silhouette score favors two groups for the M_B-vs-$v_{\rm Si\,II}$ space, we choose this number based on Δ(BIC), which was calculated using a GMM instead of k-means clustering. Again, the lack of a strong preference is likely an indication of the inadequacy of the description of the groups with Gaussian distributions. We refer to these three determined groups as the M_B-vs-$v_{\rm Si\,II}$ groups — namely the Main group, the Dim group, and the Fast group.

The Main M_B-vs-$v_{\rm Si\,II}$ group seems to resemble very closely the more populous group of SNe in the original M_B-vs-$v_{\rm Si\,II}$ diagram that Polin et al. (2019) interpreted as near-Chandrasekhar-mass explosions. Conversely, the Dim and Fast groups together make-up the entirety of objects that Polin et al. (2019) identified with sub-Chandrasekhar models produced with a thin helium shell.

One may naively expect the Main group to have good overlap with the CN Branch group, the

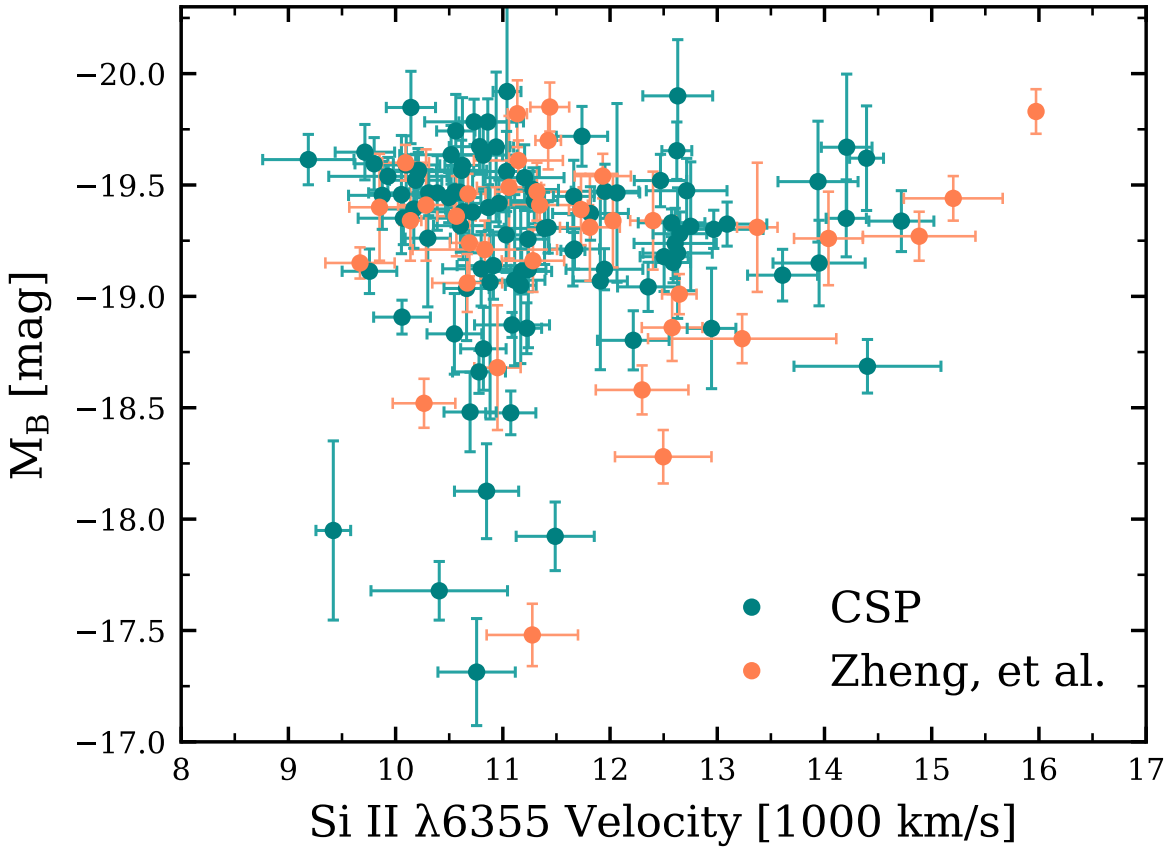


Figure 3.9: M_B-vs-$v_{\mathrm{Si\,II}}$ diagram of both Zheng et al. (2018) and CSP I+II samples. As the CSP sample is included, there are no longer two distinct groups, but rather a continuity of a similar form to the original M_B-vs-$v_{\mathrm{Si\,II}}$ diagram. For this sample we find M_B to have a mean uncertainty of $\pm$ 0.169 mag and $v_{\mathrm{Si\,II}}$ to have that of $\pm$ 290 km s^{-1}.

Fast group to overlap with the BLs [as with high $v_{\mathrm{Si\,II}}$ one would expect large pEW(Si II λ6355)], and the Dim group to overlap with the CLs. Figure 3.11 shows the Branch diagram colored with the M_B-vs-$v_{\mathrm{Si\,II}}$ groups defined by the GMM illustrated in Figure 3.10. Comparing with our 4-D GMM description of the Branch groups in Figure 3.8, it is clear that there is not a good match between the M_B-vs-$v_{\mathrm{Si\,II}}$ groups defined by [M_B, $v_{\mathrm{Si\,II}}$] and the Branch groups. The Main group is made up mostly of both CN and SS SNe. However, we also see many Fast group objects are either CNs or SSs. This is actually quite surprising, since *a priori* one would expect a strict relationship between the Fast M_B-vs-$v_{\mathrm{Si\,II}}$ group and the BLs. We see, then, that there is much dispersion in the relationship between $v_{\mathrm{Si\,II}}$ and pEW(Si II λ6355). The inconsistency here must be that the intermediate area between the Main and Fast groups cannot be established exclusively in [M_B, $v_{\mathrm{Si\,II}}$]. Finally, we remark that the Dim group tends to associate with the CLs nearly entirely, which

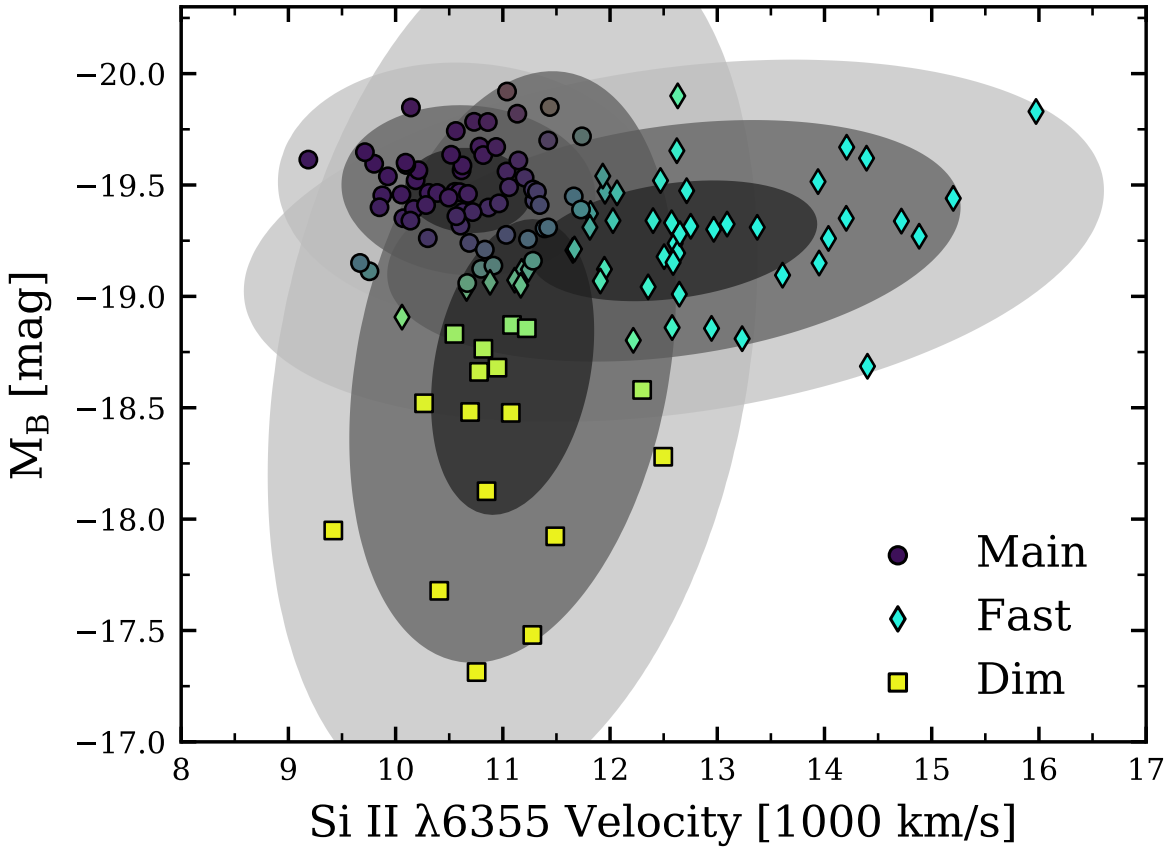


Figure 3.10: The M_B-vs-$v_{\mathrm{Si\,II}}$ diagram colored by a 2-D GMM in [M_B, $v_{\mathrm{Si\,II}}$] ($n = 3$ components) which separates the sample into what we label as the Main, Dim, and Fast groups. The contours correspond to covariance of group membership up to 3σ.

is expected.

In Figure 3.12 we show M_B-vs-$v_{\mathrm{Si\,II}}$ diagrams for CSP I+II data, excluding the Zheng et al. (2018) sample, as colors were not provided for this sample. The left panel is coded for $B_{\mathrm{max}} - V_{\mathrm{max}}$ color that is uncorrected for host-galaxy extinction, and the right panel is coded for $(B_{\mathrm{max}} - V_{\mathrm{max}})_0$ which is corrected for host extinction (Burns et al., 2018). Both colors are corrected for MW extinction. Contours in this figure are identical to those found in the [M_B, $v_{\mathrm{Si\,II}}$] GMM from Figure 3.10. From the left panel we see the similar result from Polin et al. (2019) that the bluest objects are consistently found within the Main group. However, comparing the two, we see that, after correcting for both MW and host extinction, $(B_{\mathrm{max}} - V_{\mathrm{max}})_0$ only seems to have a dependence on M_B. We find no evidence of red objects which are intrinsically bright. It is evident that both the Main and Fast groups primarily contain intrinsically bluer objects than the Dim group. After correcting for host extinction, the group that was claimed to follow sub-Chandrasekhar-mass models

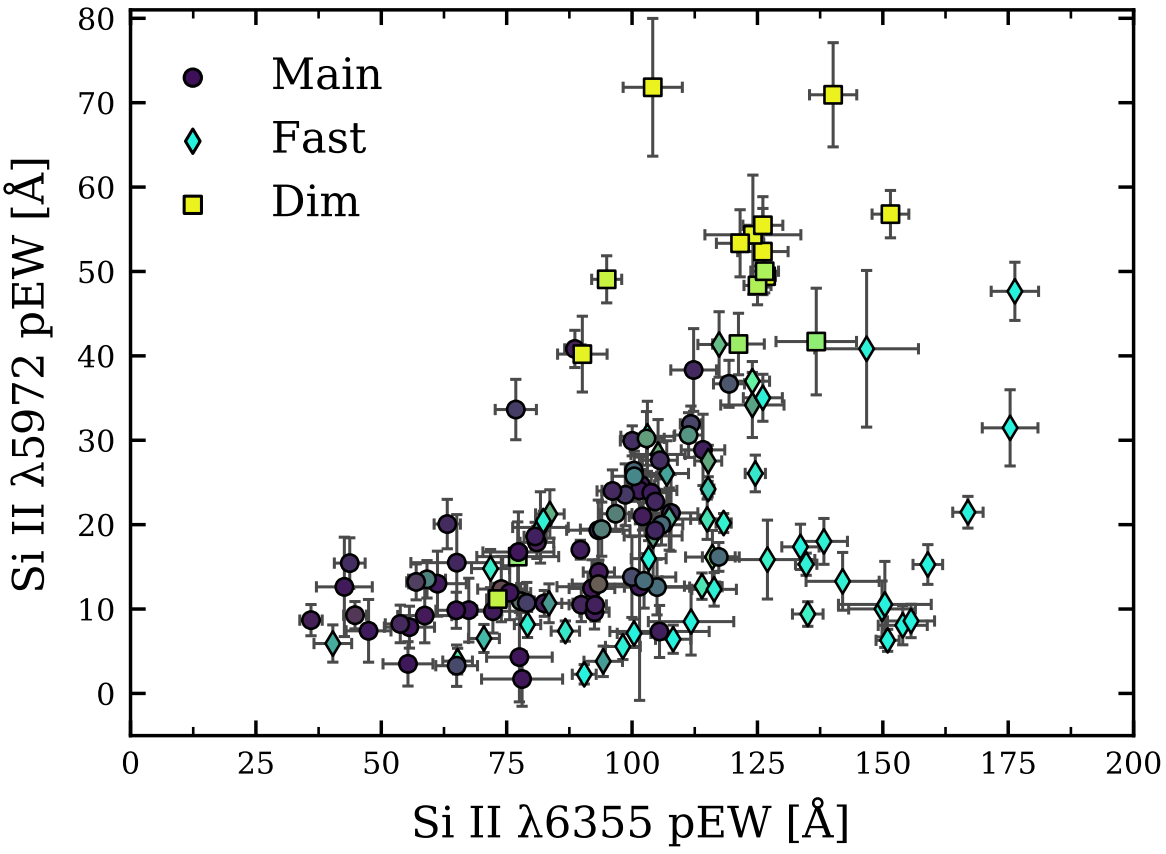


Figure 3.11: The Branch diagram with groups defined by the 2-D GMM displayed in Figure 3.10. We see that there is generally no absolute match between the original Polin et al. (2019) groups and the Branch groups as we see many Fast objects contained in BLs as well as CNs and SSs.

by Polin et al. (2019) does not display a more consistent redder color than the more populated clump of Main group objects, indicating that these high-velocity SNe Ia are not likely to come from sub-Chandrasekhar-mass explosions.

3.4.3.1. Color of the Fast Group SNe Ia

It is seen from Figure 3.12 that nine Fast SNe Ia (chosen by $v_{\mathrm{Si\,II}} > 13{,}000$ km s^{-1}) from the CSP I+II sample generally exhibit a redder $B_{\mathrm{max}} - V_{\mathrm{max}}$ color (when not corrected for host-galaxy extinction). These SNe Ia are listed in Table 3.2, along with their $(B_{\mathrm{max}} - V_{\mathrm{max}})_0$ color and extinction values for Milky Way and host galaxies. This is in line with the results from Polin et al. (2019) who showed a similar plot to the left panel of Figure 3.12. Note that Polin et al. (2019) do not correct for host-galaxy extinction. Comparing the two panels of Figure 3.12 suggests, then, that these Fast SNe Ia seem to be quite reddened in their host galaxy. Figure 3.13 shows that in [$V_{\mathrm{max}} - r_{\mathrm{max}}$, $B_{\mathrm{max}} - V_{\mathrm{max}}$] space, non-CL SNe Ia generally follow a line that is relatively insensitive to the value

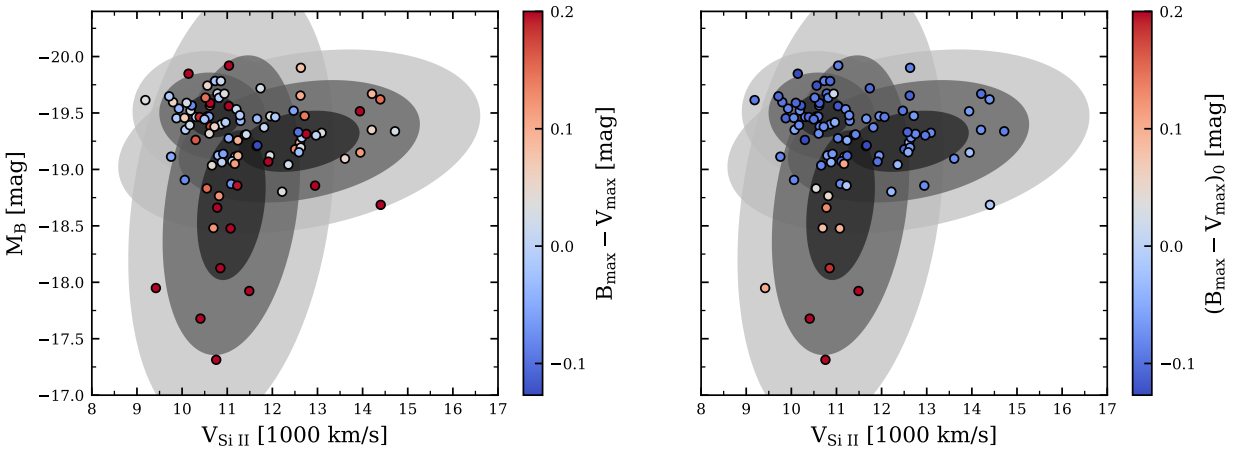


Figure 3.12: M_B-vs-$v_{\mathrm{Si\,II}}$ diagram of the CSP I+II samples (without the Zheng et al. (2018) sample, as they did not provide colors for their sample). The left panel is coded for $B_{\mathrm{max}} - V_{\mathrm{max}}$ color that is uncorrected for host galaxy extinction, whereas the right panel is coded for $(B_{\mathrm{max}} - V_{\mathrm{max}})_0$ color that is corrected for host extinction. The color scale is truncated for $(B_{\mathrm{max}} - V_{\mathrm{max}})_0 > 0.2$ mag due to relatively large separation of 19 outliers with 0.2 mag $< B_{\mathrm{max}} - V_{\mathrm{max}} \leq 0.939$ mag (left panel) and four outliers with 0.2 mag $< (B_{\mathrm{max}} - V_{\mathrm{max}})_0 \leq 0.536$ mag (right panel). The contours are exactly those from Figure 3.10, which indicate the [M_B, $v_{\mathrm{Si\,II}}$] GMM covariance. M_B plotted in both panels is the same, always including host extinction.

of R_V, giving confidence that these Fast SNe Ia are for some reason preferentially reddened in their host galaxies and that the inferred host extinction is not due to some underlying assumption in the `SNooPy` templates that is misinterpreted. Since Zheng et al. (2018) used `MLCSk2` and a fixed R_V to infer host extinction and M_B, although similar but not identical to the methods of `SNooPy`, it lends some credence to the argument that `SNooPy` is not somehow miscorrecting the Fast subsample. However, we have introduced a correlation between pEW(Si II λ5972) and M_B as well as pEW(Si II λ5972) and $(B_{\mathrm{max}} - V_{\mathrm{max}})_0$ just by using `SNooPy` to fit for these quantities. The same would be true if we were to use other light-curve fitters such as `MLCS2k2` (Jha et al., 2007) or `SALT2`/`SALT3` (Guy et al., 2007; Kenworthy et al., 2021), because their light-curve-shape parameters would also be highly correlated with pEW(Si II λ5972). Nonetheless, while these objects have distinct spectroscopic properties, we are assuming they all follow the same intrinsic color–s_{BV} relation from Burns et al. (2014) in order to perform the extinction corrections.

Wang et al. (2009) noted this importance of $v_{\mathrm{Si\,II}}$, splitting SNe Ia into two groups and noting that the group with $v_{\mathrm{Si\,II}} \gtrsim 11{,}800$ km s^{-1} preferred a lower value of $R_V \sim 1.6$. Figure 3.13 shows that these Fast SNe Ia follow a dust trajectory which is insensitive to R_V, which is an indication that the trend seen in Figure 3.12 is likely due to extinction from dust in the host. This preference for these Fast objects to appear in dusty regions (perhaps near the core, see Uddin et al., 2020) should be the subject of future work.

Table 3.2. Subset of Fast CSP I+II SNe Ia with $v_{\text{Si II}} > 13{,}000$ km s^{-1}.

SN	$(B_{\text{max}} - V_{\text{max}})_0$ (mag)	A_V^{MW} (mag)	A_V^{Host} (mag)
ASASSN-14hr	-0.074 ± 0.018	0.04	0.31
CSP15B	-0.024 ± 0.022	0.20	0.36
LSQ13aiz	-0.100 ± 0.024	0.24	0.39
PTF13duj	-0.073 ± 0.024	0.21	0.15
2006br	-0.064 ± 0.101	0.06	2.28
2008go	-0.078 ± 0.016	0.10	0.09
2008O	-0.014 ± 0.029	0.24	0.37
2009Y	-0.070 ± 0.025	0.27	0.28
2012bl	-0.085 ± 0.017	0.09	0.01

We also note that, in contrast to the $B_{\text{max}} - V_{\text{max}}$ values presented in Polin et al. (2019), when corrected for host extinction, the $(B_{\text{max}} - V_{\text{max}})_0$ range is much narrower, mostly falling within $-0.2 < (B_{\text{max}} - V_{\text{max}})_0 < 0.2$, having only four outliers with $(B_{\text{max}} - V_{\text{max}})_0 > 0.2$ (shown more explicitly in Figure 3.14). Figure 3.14 shows $(B_{\text{max}} - V_{\text{max}})_0$ versus pEW(Si II λ5972) for the CSP I+II sample. We see that, at low pEW(Si II λ5972), there is no correlation between the two quantities. This is expected from Figure 3.12, because only Dim objects have a large increase in $(B_{\text{max}} - V_{\text{max}})_0$ compared to brighter objects from the Main and Fast groups that exhibit bluer color with no other noticeable trend. Generally, there is only a spread of about 0.15 mag in $(B_{\text{max}} - V_{\text{max}})_0$ for most SNe Ia.

We also include a Branch diagram colored in $(B_{\text{max}} - V_{\text{max}})_0$ for reference to the Branch groups in Figure 3.15. Contours in this figure are exactly those determined by the 4-D GMM that are displayed in Figure 3.8. We see that, as expected, the CLs are generally the reddest objects, where there is no other trend between the other three groups.

3.5. Discussion

3.5.1. Constraining the Branch GMM

Supplementing the two quantities pEW(Si II λ5972) and pEW(Si II λ6355) with more information provides more certainty in Branch group assignment. Quantitatively, the 4-D GMM has been shown to decrease the size of GMM covariance contours by making a constraint that values must also be similar in additional dimensions. When we compare the three different models in Section 3.4.2, however, we see that uncertainty in the [pEW(Si II λ5972), pEW(Si II λ6355)] GMM mostly decreases with the inclusion of only a single quantity. Therefore, although we take the 4-D

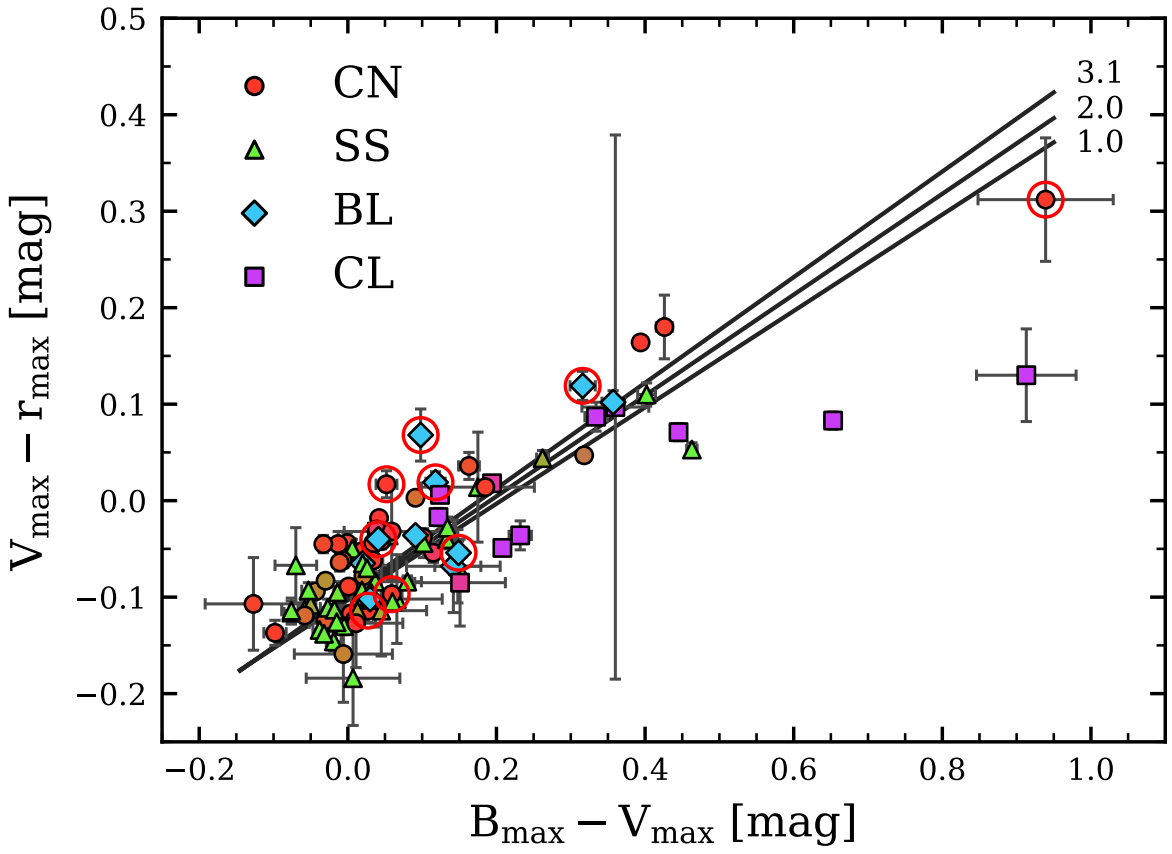


Figure 3.13: Color-color figure of MW-extinction-corrected $V_{\rm max} - r_{\rm max}$ versus $B_{\rm max} - V_{\rm max}$ for the CSP I+II subsample. The figure is colored by 4-D Branch group membership, and red circles indicate objects with $v_{\rm Si\,II} > 13{,}000$ km s^{-1}. R_V values of 1.0, 2.0, and 3.1 are indicated. The R_V lines show the general direction of displacement due to a correction for reddening. In general, non-CL SNe Ia are generally insensitive to the value of R_V and follow a trajectory expected from a dust reddening law. These Fast group SNe Ia also follow the dust trajectory, suggesting that, for some reason, they tend to be relatively highly extinguished in their host galaxies. While this does not prove that they have the correct M_B determined by `SNooPy`, it makes it plausible.

GMM as our defining model for Branch groups, as it includes additional M_B information that we show is non-linearly correlated with pEW(Si II λ5972), it can still be approximated with only the inclusion of $v_{\rm Si\,II}$. In this way, the Branch groups can be approximated solely with spectroscopic data. The effects obtained above are interesting given works showing that the inclusion of $v_{\rm Si\,II}$ reduces the scatter of Hubble residuals (Foley & Kasen, 2011) and that there appears to be a correlation in the sign of the Hubble residuals and the value of $v_{\rm Si\,II}$ (Siebert et al., 2020). Recently, it was suggested that SNe Ia with high $v_{\rm Si\,II}$ are closely associated with massive host environments and that high $v_{\rm Si\,II}$ is due to a variation in explosion mechanism (Pan, 2020). We find that the BL group

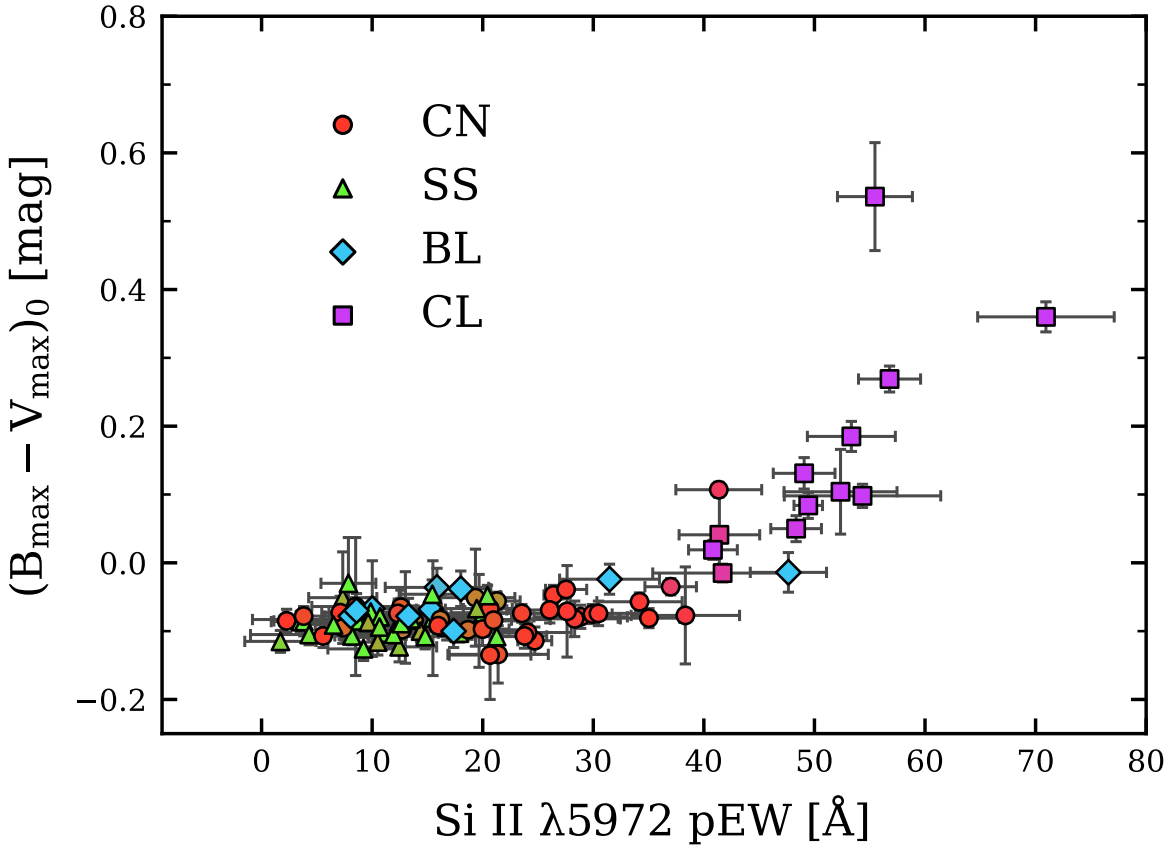


Figure 3.14: $(B_{\rm max} - V_{\rm max})_0$ versus pEW(Si II λ5972) of the CSP I+II samples with 4-D Branch groups colored. We find that at pEW(Si II λ5972) $\lesssim$ 40 Å, there is no correlation between the two quantities, which is expected by the M_B versus pEW(Si II λ5972) relation we find in Section 3.4.2.2.

is more likely to be distinct, but it would be good to identify the environments of objects in different Branch groups. We do note that the tendency of the Fast objects to be strongly extinguished in their hosts is likely an indication about the nature of the environment of these SNe Ia.

It is also seen in Figure 3.8 that 26 SNe Ia change their most probable group assignment between the 2-D GMM, the 3-D GMM with $v_{\rm Si\ II}$ inclusion, and the 4-D GMM. Table 3.3 details the changes involved for this sample. Out of the total 26 SNe Ia, 24 are reassigned with the inclusion of both M_B and $v_{\rm Si\ II}$ to the GMM, which is substantial compared to the total sample size of 133 SNe Ia. We find that including these two parameters has the greatest effect on CN and CL objects, with little to no effect on SS and BL objects that are defined in the 2-D model. This may be due to the relationship we find between M_B and pEW(Si II λ5972) coupled with a dispersion in the $v_{\rm Si\ II}$–pEW(Si II λ6355) relation; however, more work must be done to determine the dominant quantities that govern the changes between each group. Ultimately, group assignment and any physical attributes that objects

Table 3.3. SNe Ia with Branch group membership changes between the 2-D GMM, the spectroscopic 3-D GMM with $v_{\mathrm{Si\,II}}$ dependence, and the 4-D GMM. The group for each model is based on the most probable assignment determined by each GMM.

SN	2-D Group	3-D ($v_{\mathrm{Si\,II}}$) Group	4-D Group
CSP14acl	CN	CN	SS
LSQ13ry	CN	CN	SS
2005el	CN	CN	SS
2004ey	CN	SS	CN
2008fr	CN	SS	CN
ASASSN-15al	CN	SS	SS
ASASSN-15hf	CN	SS	SS
2000dn	CN	SS	SS
2005bg	CN	SS	SS
2005cf	CN	SS	SS
2005hc	CN	SS	SS
2006le	CN	SS	SS
2008bq	CN	SS	SS
2013fy	CN	SS	SS
2008hu	CN	BL	BL
2002dl	CN	CN	CL
2003gt	CN	CN	CL
2011iv	CN	CN	CL
ASASSN-14hu	SS	CN	CN
2003gn	CL	BL	BL
PS1-14ra	CL	CL	CN
PTF14w	CL	CL	CN
2007on	CL	CL	CN
2008ec	CL	CL	CN
2008fl	CL	CL	CN
2011jh	CL	CL	CN

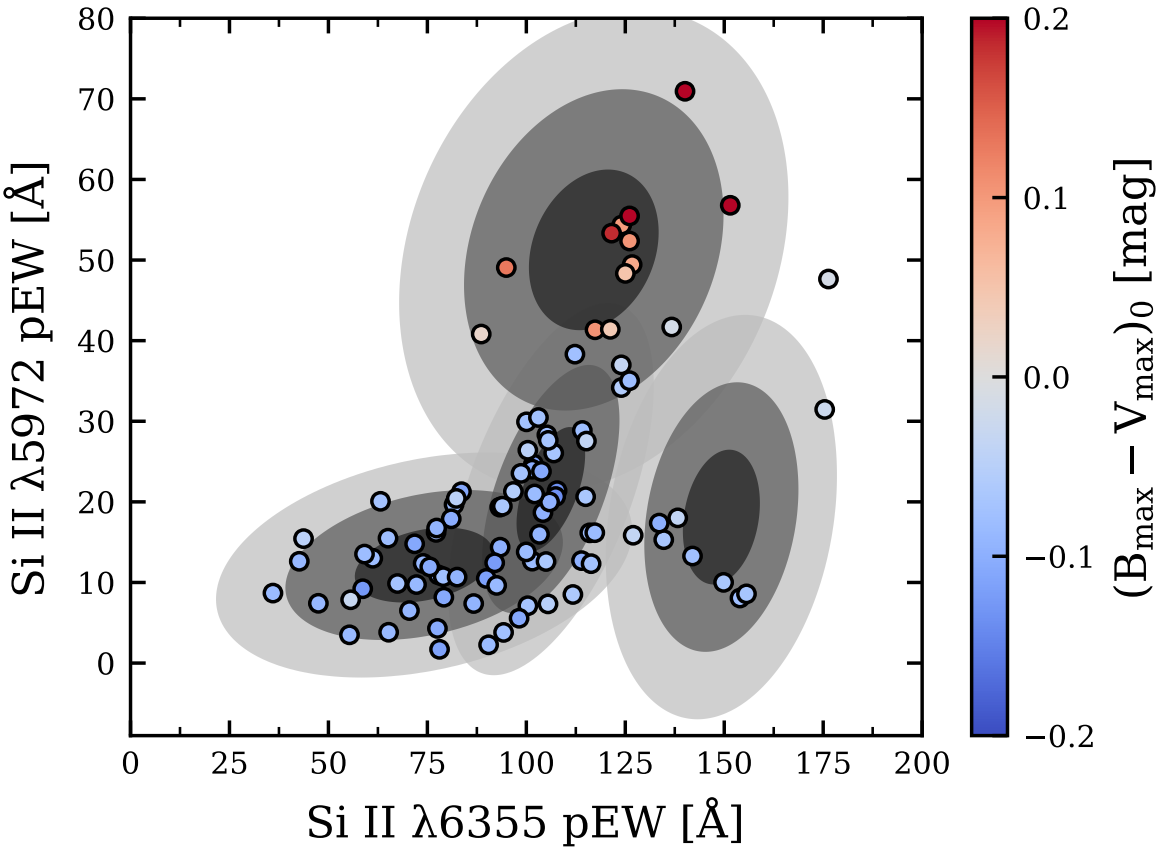


Figure 3.15: Branch diagram of the CSP I+II samples coded for $(B_{\rm max} - V_{\rm max})_0$ color. The color scale is truncated for $(B_{\rm max} - V_{\rm max})_0 > 0.2$ in the same way as Figure 3.12. The contours are exactly those from Figure 3.8, indicating our defined 4-D Branch groups.

of each group may possess are not solely a function of pEWs, but also of their other properties such as M_B and $v_{\rm Si\,II}$, as is shown in this work. Interestingly, the classification changes from CN to SS are fairly stable going from the 3-D GMM to the 4-D GMM, and also the identification of BL is stable from 3-D to 4-D. The most noticeable difference from 3-D to 4-D is the change from CL to CN for a half-dozen SNe Ia.

3.5.2. M_B-vs-$v_{\rm Si\,II}$ Groups

Our analysis does not show the clear dichotomy in [M_B, $v_{\rm Si\,II}$] space found in Polin et al. (2019) when we include the CSP I+II samples in addition to the Zheng et al. (2018) sample. In particular, their result seems to be due to not correcting for reddening in the host, when plotting $(B_{\rm max} - V_{\rm max})_0$. The only intrinsically red objects are the Dims/CLs. When all four parameters are accounted for in the GMM, as shown in Figure 3.16, we see that the Main group consists primarily of CN and SS objects, the Fast group is dominated by BL objects with only a few CNs, and the Dim group is again

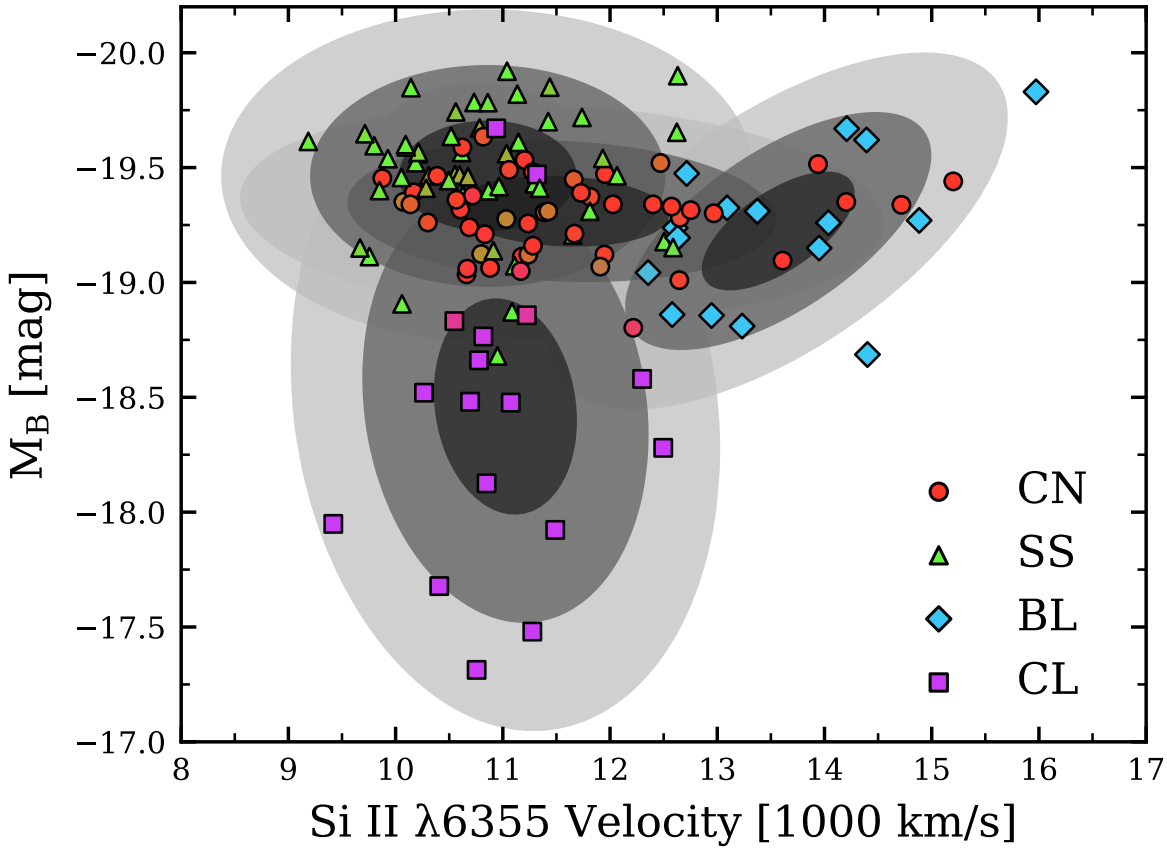


Figure 3.16: The M_B-vs-$v_{\mathrm{Si\,II}}$ diagram of the 4-D GMM analysis from Section 3.4.2.3. The contours indicate covariance (up to 3σ) of group membership projected into [M_B, $v_{\mathrm{Si\,II}}$] space. We see that the Main group consists of both CN and SS objects and that the Fast group is separated into both BL and CN objects. This separation implies there is a dispersion in the relationship between $v_{\mathrm{Si\,II}}$ and pEW(Si II λ6355). As expected, the Dim group almost exclusively contains CL objects.

made up primarily of CL objects. The Fast objects can indeed be divided using their pEW properties. Therefore, the Fast group SNe Ia do not need to stem from sub-Chandrasekhar explosions, and additional information is required to more concretely determine the underlying physics involved.

If we assume the Main group to be the union of both SS and CN objects, we do find that, in comparing Figure 3.10 with the modified M_B-vs-$v_{\mathrm{Si\,II}}$ diagram presented in Figure 3.16, the projected covariances into [M_B, $v_{\mathrm{Si\,II}}$] space show significant qualitative change. The BL contours that are associated with the Fast group are noticeably smaller, meaning it is a more well-constrained grouping with less uncertainty between the BLs and other groups. The Main group contains both CN and SS objects, so it is important to determine if there are other distinctions between CN and SS objects that can be related to variations in the progenitors or explosion mechanisms.

From Figure 3.12 we see that the 4-D grouping system projected onto [M_B, $v_{\rm Si\ II}$] space does not sort SNe Ia by $(B_{\rm max} - V_{\rm max})_0$ color. We find that $(B_{\rm max} - V_{\rm max})_0$ is mostly distinct only for Dims, and the Fast group color is bluer than Dims, regardless of whether the object is a CN or BL SN Ia. More work must be done to further sub-classify these SNe Ia.

3.5.3. Branch Group Relation with $s_{\rm BV}$

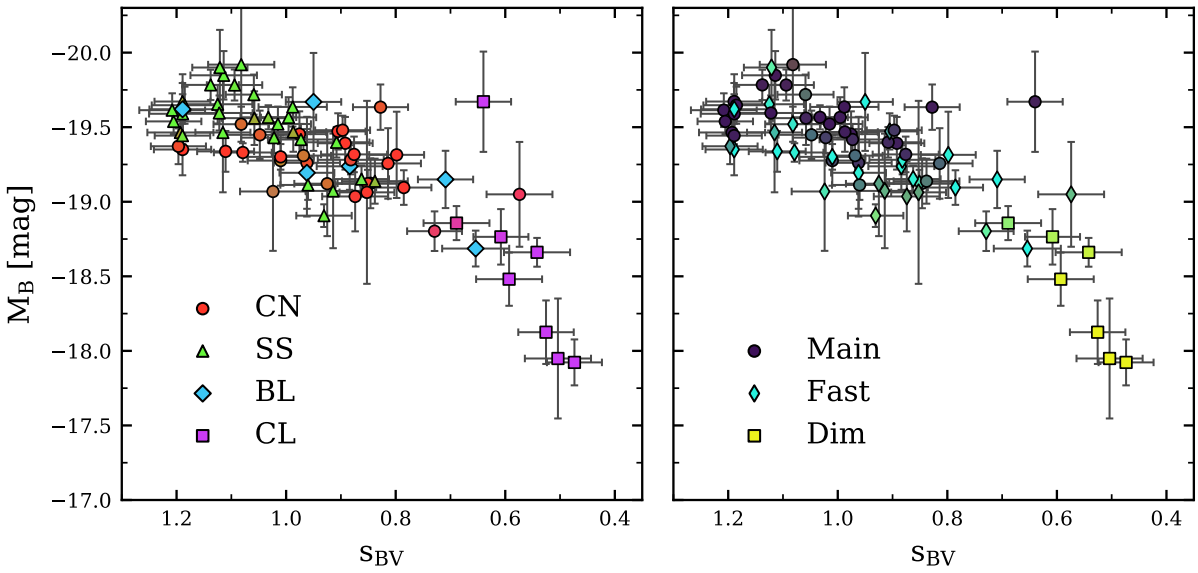


Figure 3.17: Illustration of the Phillips relation as M_B versus $s_{\rm BV}$. Points are color-coded by the 4-D Branch groups (left panel) and the M_B-vs-$v_{\rm Si\ II}$ groups (right panel).

Figure 3.17 shows the Phillips/Burns relation color-coded by Branch group (left panel) and by M_B-vs-$v_{\rm Si\ II}$ group (right panel). Both the CL and Dims are intrinsically dim, and there is much overlap between them, as seen in Section 3.5.2. The Branch diagram coded for $s_{\rm BV}$ is displayed in Figure 3.18, and it shows that $s_{\rm BV}$ does not distinguish between CN and BL SNe Ia with high values of $v_{\rm Si\ II}$ as is seen in Figure 3.16. It will be interesting to understand what causes this variation to better improve the use of SNe Ia as cosmological probes.

In Figure 3.17 there appears to be a transition region of SS into CN that corresponds to a transition of the Main group into the Fast group. It could be that there is some underlying parameterization involved that better explains the dispersion in this Phillips/Burns relation, such as a dependency upon M_B-vs-$v_{\rm Si\ II}$ group membership or $v_{\rm Si\ II}$. This interesting behavior is left to be studied in future work.

Figure 3.19 shows that a Phillips-like relation exists for pEW(Si II λ5972) and the color stretch parameter, $s_{\rm BV}$. In other words, pEW(Si II λ5972) acts as a stand-in for M_B (see Figures 3.6 and 3.7). Figure 3.19 also shows the 4-D Branch group placement for this relation. At high $s_{\rm BV}$ ($s_{\rm BV}$

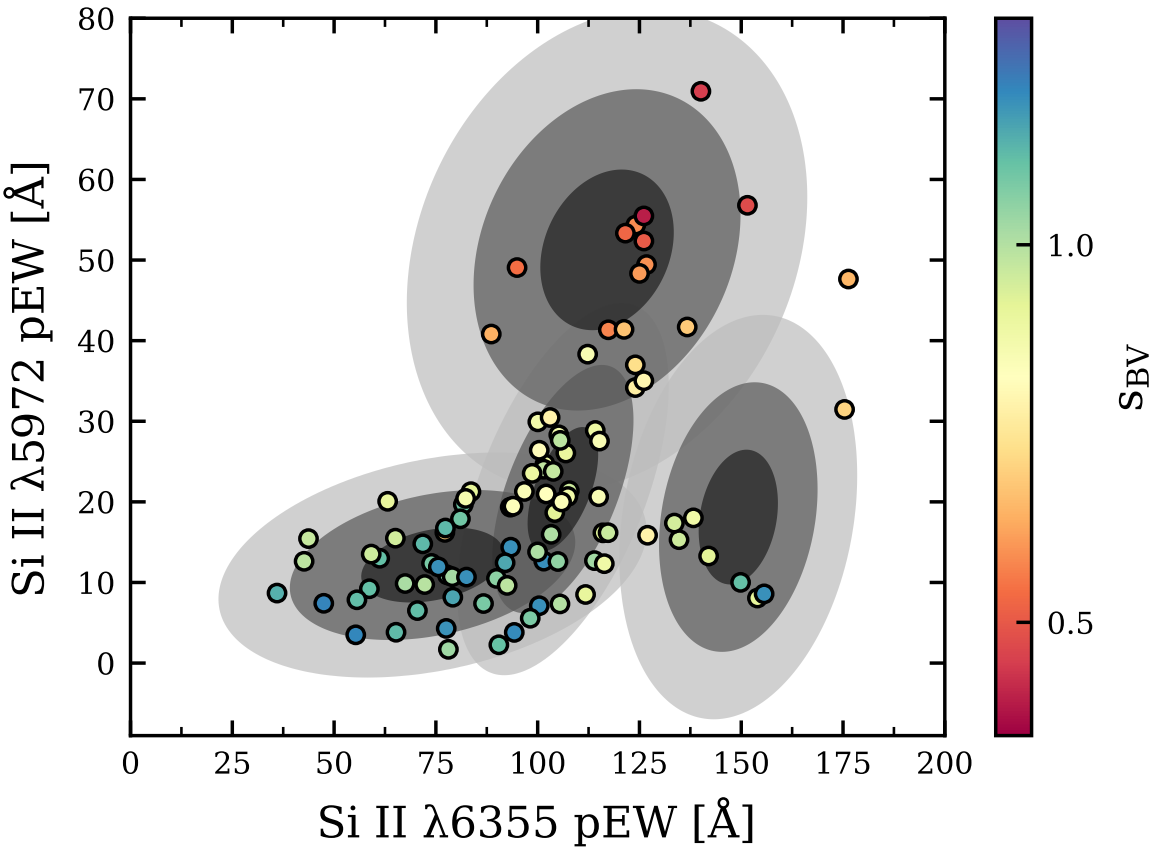


Figure 3.18: Branch diagram of the CSP I+II samples coded for color stretch parameter $s_{\rm BV}$. The contours are exactly those from Figure 3.8, indicating our defined 4-D Branch groups.

$\gtrsim 1.0$), little correlation is found between pEW(Si II λ5972) and $s_{\rm BV}$, which is expected because the dependency of $s_{\rm BV}$ on M_B decreases for brighter SNe Ia (seen in Figure 3.17). However, a linear trend is found for $s_{\rm BV} \lesssim 1.0$.

3.6. Conclusion

We have shown that a 4-D GMM analysis of [M_B, $v_{\rm Si\,II}$, pEW(Si II λ5972), pEW(Si II λ6355)] with $n = 4$ components yields robust groupings that strongly identify the four Branch groups: core-normals, shallow-silicons, broad-lines, and cools. We have shown that there seems to be a strong correlation between M_B and pEW(Si II λ5972), and because of this we suggest that the quantities pEW(Si II λ6355), pEW(Si II λ5972), and $v_{\rm Si\,II}$ can be used to approximate this 4-D model with only spectroscopic quantities. Furthermore, a 3-D GMM makes the BL group nearly distinct; it seems reasonable that this subclass of SNe Ia would be distinct from other SNe Ia as suggested previously by Wang et al. (2013). One possibility is that the high velocity of the photosphere is

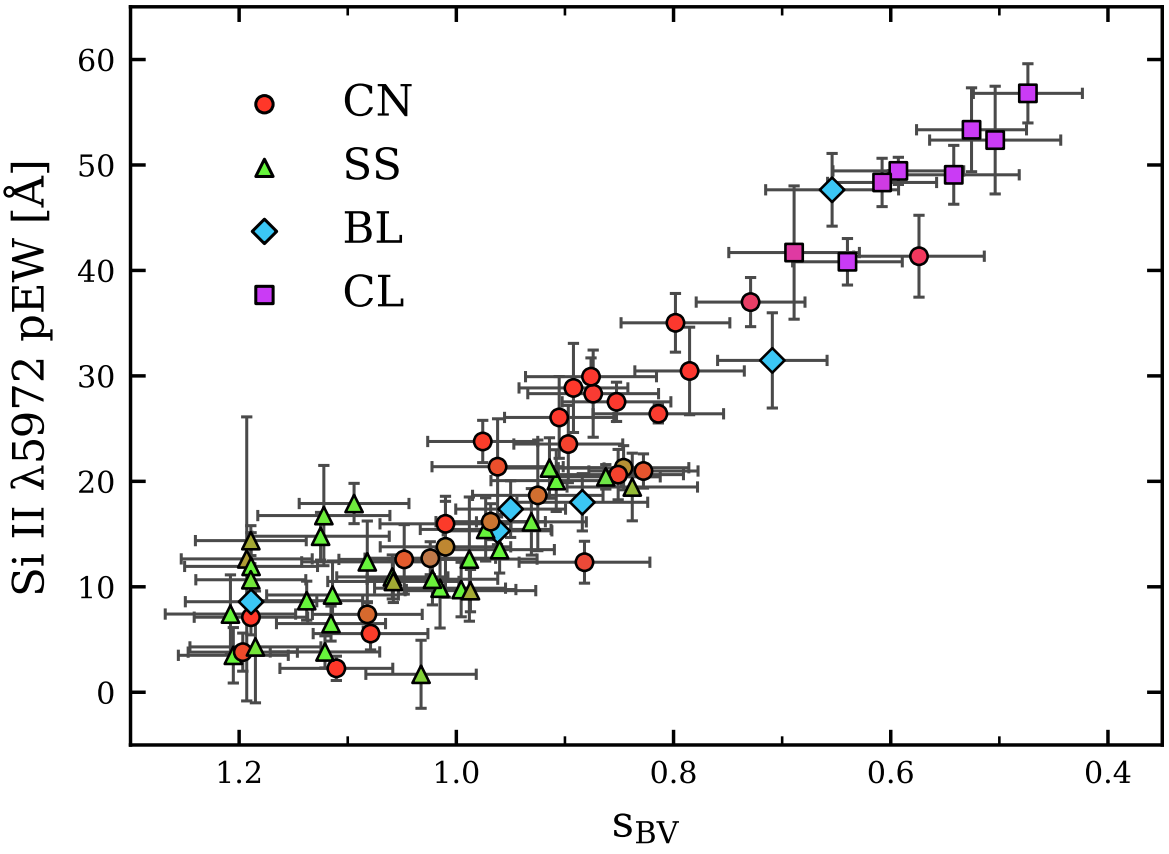


Figure 3.19: pEW(Si II λ5972) versus $s_{\rm BV}$ of the CSP I+II samples colored by the 4-D Branch groups. We see a linear trend for SNe Ia with $s_{\rm BV} \lesssim 1.0$.

produced by a shell, such as that produced by a pulsating delayed detonation (Hoeflich et al., 1996; Dessart et al., 2014).

The original M_B-vs-$v_{\rm Si\,II}$ diagram was interpreted to separate SNe Ia into two groups based on M_B and $v_{\rm Si\,II}$ that corresponded to Chandrasekhar-mass explosions and sub-Chandrasekhar explosions (Polin et al., 2019). We find that this is an incomplete description; rather than a clear dichotomy in [M_B, $v_{\rm Si\,II}$] space, there are three connected groups. Ultimately we find that SN Ia subtypes are better-delineated by the Branch groups using GMMs. We define these three M_B-vs-$v_{\rm Si\,II}$ groups as: the Main group consisting of CNs and SSs, the Fast group consisting of BLs and a subset of CNs, and the Dim group that consists of CLs. It is shown that Fast objects can be divided using their Branch group membership (Si II pEWs), and this division may aid in predicting their underlying explosion mechanisms. We find that this separation is generally unexplained by the color-stretch parameter $s_{\rm BV}$, and so future work must be done to explain this dispersion.

Open access to each of these GMMs is provided in a Python wrapper named `SNIaDCA`. See

Section A.2 for a brief discussion.

CHAPTER 4

Variation of the NIR Secondary Peak

4.1. Introduction

The optical light-curves of spectroscopically normal SNe Ia are typically characterized by a global maximum that occurs ∼18 days (in the B-band) after explosion. This is followed by a decline in the light-curve until the nebular phase, typically beginning ∼200 days after maximum light (see Friesen et al., 2017). However, for NIR light-curves (e.g., $izYJHK$ bands), a secondary maximum is observed ∼20–30 days after the global maximum for that band (see, e.g., Elias et al., 1981; Hamuy et al., 1996; Meikle, 2000). Kasen (2006) finds that the origin of this NIR secondary maximum can be traced to the transition between singly and doubly ionized states of iron-group elements. As the SN Ia cools over time, a temperature $T_{21} \approx 7000$ K is achieved that allows the recombination of iron-rich ejecta, which leads to an increase in NIR radiation. This "wave" of recombination recedes into the ejecta, which is indicated by the beginning and end of the secondary maximum of the NIR light-curves.

The variation of the NIR secondary peak observed has not yet been fully captured with sufficient accuracy with theoretical or empirical models. This chapter provides an illustration of this problem and is intended as a preface to the following Chapter 5, which describes the work done by Burrow et al. (2024). The motivation behind the techniques employed in Chapter 5 stems from the inability for the color-stretch light-curve parameter $s_{\rm BV}$ to describe the variation seen in the NIR secondary peak. This may lead to poor estimations of the NIR SED made for some SNe Ia by spectral templates that are parameterized by $s_{\rm BV}$, such as those by Lu et al. (2023).

`SNooPy` is a light-curve fitter that models light-curves with single-valued parameters such as $\Delta m_{15}(B)$ and $s_{\rm BV}$. It has shown to be useful in various SN Ia studies, including those given here in Chapter 3 and Chapter 5, to accurately estimate photometric properties such as the time of maximum-light (in various bands), $s_{\rm BV}$, K-corrections, host-galaxy reddening, etc. However, because it is solely parameterized by optical quantities $\Delta m_{15}(B)$ and $s_{\rm BV}$, the NIR light-curve models are much more limited. In Section 4.2, we provide a brief review of how these `SNooPy` models are created and some of the functionality of `SNooPy`. Then, in Section 4.3, we show an example of the ineffectiveness of capturing the NIR secondary peak with $s_{\rm BV}$.

4.2. Review of SNooPy

In Chapter 3, we use values of M_B and other photometric quantities such as $B_{\max} - V_{\max}$ color and host extinction — all of which were determined by the light-curve-fitting tool SNooPy. In this section, we briefly review how SNooPy (Burns et al., 2014, 2018) models light-curves from observations of multi-band photometry. This will provide further context to Chapter 5 and Chapter 6, as the results of those chapters are directly applicable to tools like SNooPy which use observed photometry of SNe Ia for precision cosmology.

The goal of SNooPy is to solve for a set of parameters such that a chosen model, which is dependent on those parameters, fits the observed photometry with maximized likelihood. There are three types of built-in models for SNooPy that depend on $s_{\rm BV}$: EBV_model2 with parameters $t_{B_{\max}}$, $s_{\rm BV}$, $E(B-V)_{\rm host}$, and distance modulus μ; max_model with parameters $t_{B_{\max}}$, $s_{\rm BV}$, and $X_{\max}$ for each band X being fit; and color_model with parameters $t_{B_{\max}}$, $s_{\rm BV}$, $B_{\max}$, $E(B-V)_{\rm host}$, and $R_{V,\rm host}$. Each of these models also depend on templates for each band to be fit, which are trained using CSP I+II light-curves that are fit using splines (see Burns et al., 2011). The parameters of these models can be estimated using either a Levenberg-Marquardt least-squares fitter or a Markov chain Monte Carlo fitter; however, the latter is particularly useful because it is a Bayesian framework, allowing one to specify priors on the parameters being fit.

For any model used with SNooPy, there are more obstacles to overcome, such as determining K-corrections and reddening coefficients R_X. K-corrections (Oke & Sandage, 1968) are corrections to photometry made to correct for host-galaxy redshift. This process is complex due to the wavelength-dependency of this redshift and requires an estimate for the SED of the target SN Ia that is being fit to obtain an accurate K-correction for a given photometric band. Accurately estimating an SED for K-corrections of NIR bands is more difficult due to the relatively low number of spectroscopic observations in the NIR. The recent work of Lu et al. (2023) has produced NIR spectral templates for this very reason. These templates are parameterized by $s_{\rm BV}$ and the phase with respect to time of $B_{\max}$. Another method of obtaining the NIR part of SED is introduced in Chapter 5; this method does not rely on the value of $s_{\rm BV}$, but rather the observed optical spectrum. An accurate estimate of a K-correction is needed for every band that SNooPy is fitting, as these K-corrections are not negligible when inferring absolute photometric quantities of SNe Ia and distances to their host galaxies.

4.3. SNooPy Residuals in the NIR

In Figure 4.1, we show residuals of three SNooPy models from observations taken between -10 and $+100$ days in four filters, $BViY$, indicated in the top right of each panel. The three models are based on observed photometry from SN 2007af, SN 2009Y, and SN 2012hr, which all have similar $s_{\rm BV}$ ($0.95 < s_{\rm BV} < 0.99$). The models for each filter were made by interpolating the

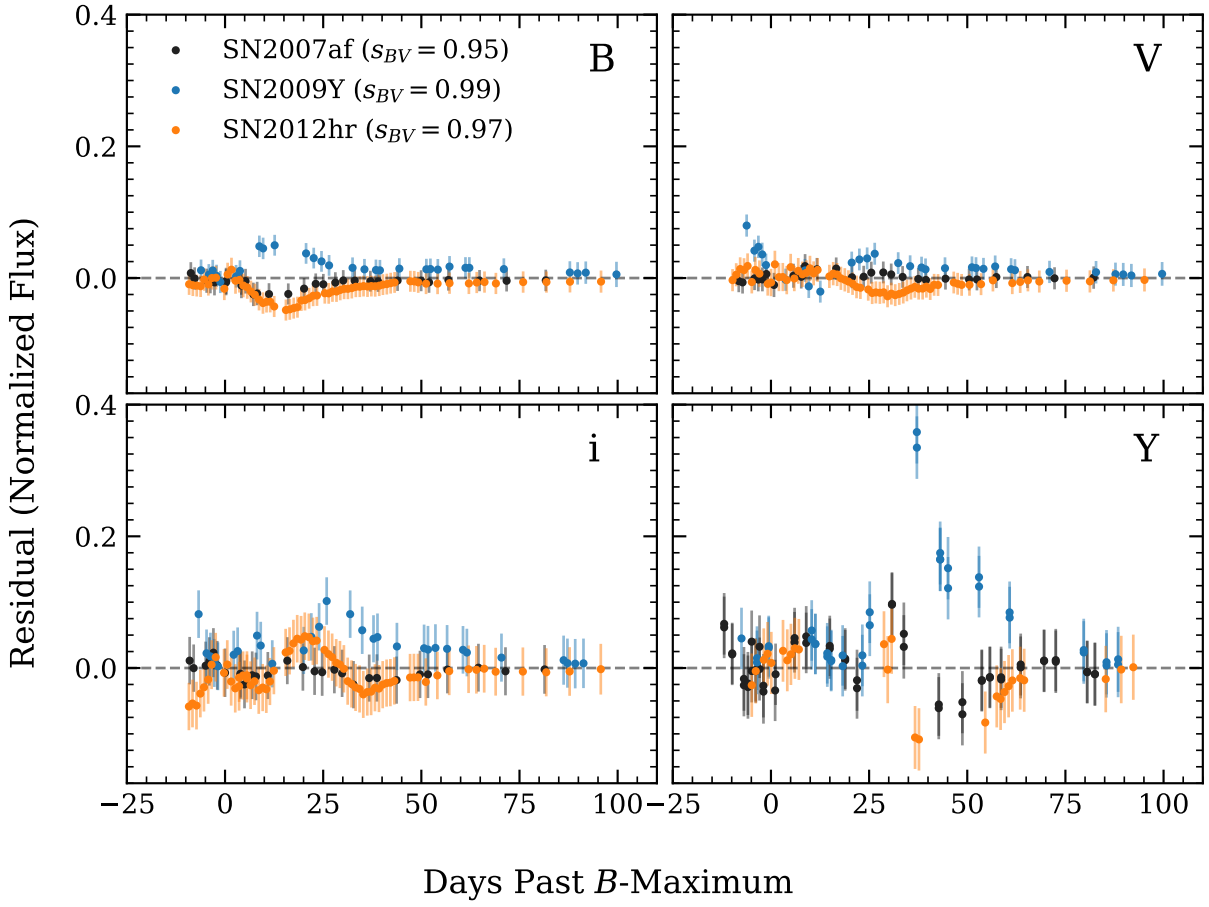


Figure 4.1: SNooPy residuals for SN 2007af, SN 2009Y, and SN 2012hr, all of which have similar $s_{\rm BV} \sim 0.97$.

observed photometry of each individual filter of the SN Ia, instead of using SNooPy multi-band light-curve templates. Regardless, the conclusion would be the same because they share similar values of $s_{\rm BV}$. The residuals are calculated from observations and models on the scale of normalized flux (as opposed to magnitudes). Each epoch shown for residuals from each filter is on the same phase with respect to B-band maximum.

For the BV light-curves in the top two panels, we find that there is generally good fits between the SNooPy light-curves and the observed photometry. The highest degree of variation for these bands is seen around the time of the NIR secondary maximum, around +15–25 days past B-maximum for these three SNe Ia. This variation of the residuals between different SNe Ia increases for the i-band light-curve at the same phases in time. At earlier times, near B-maximum, there is also more deviation from the observed photometry seen for each model. The variation between different SNe Ia is even greater in the Y-band light-curve, especially at and after the secondary maximum. Because the SNooPy models should stay nearly equivalent between these three SNe Ia

sharing such similar s_{BV} values, the residuals indicate that the observed photometry of these SNe Ia are more intrinsically variable in the NIR.

4.4. Conclusion

Figure 4.1 demonstrates that empirical models of the NIR behavior of SNe Ia cannot be parameterized completely by the photometric parameter s_{BV}. For SNe Ia of similar s_{BV}, `SNooPy` models show considerable variation in their NIR light-curves. Because NIR spectral templates for SEDs such as those of Lu et al. (2023) are parameterized by phase and s_{BV}, it could be that they do not account for this NIR variation seen for SNe Ia of similar s_{BV}. If the NIR part of an SED is not precisely estimated for a SN Ia, the K-corrections and corrected light-curves will also suffer as a result. As seen in Figure 4.1, this would occur especially near the secondary maximum shown in NIR bands.

An alternative to NIR spectral templates is discussed in the next chapter. This method prepares the spectra for the empirical models in a similar manner to spectral templates by performing PCA on a set of NIR spectra. However, this method avoids parameterization by s_{BV}; to summarize, this is done by using optical spectra in place of s_{BV}. By doing so, more complex wavelength-dependent correlations could be found between optical and NIR spectra, theoretically allowing for the NIR SED to be estimated with less uncertainty. In this way, Chapter 5 is complimentary to the work of Lu et al. (2023).

CHAPTER 5

Extrapolation of Spectra into the NIR

This chapter is taken from Burrow et al. (2024), published in *The Astrophysical Journal*, reformatted with slight modifications to improve the readability, consistency, and sequencing of this dissertation. The Discussion sections that are excluded here (Sections 5.1–5.2 of Burrow et al. (2024)) are instead discussed in Chapter 7.

5.1. Introduction

Recently, the behavior of SN Ia spectra have been studied by the use of advanced statistical methods and machine-learning techniques. As sample sizes of observed SNe Ia are fairly low and time-dependent, it becomes difficult and even unwise to make use of complex methods such as neural networks in most cases. However, one approach that handles this limitation relatively well due to its simplicity is principal component analysis (PCA; Pearson, 1901). PCA is a process of dimensionality reduction that is used to describe the highest degrees of variation in a given sample. This is done by algorithmically calculating a set of basis eigenvectors — or principal components (PCs) — that iteratively describe the most variation in the sample of data that is orthogonal to each subsequent PC.

PCA has been employed to study the diversity of SNe Ia many times in the recent past. `SALT2` and `SALT3` (Guy et al., 2007; Kenworthy et al., 2021) use a technique very similar to principal component decomposition which includes a reddening factor, constraining SN Ia diversity to two parameters. This allows for estimates of distance and photometric redshift as well as aiding in SN Ia spectroscopic line identification. `SNEMO`, an empirical principal-component model has been generated to describe three different levels of SN Ia diversity that is comparable to `SALT2` (Saunders et al., 2018). PCA has also been useful in constructing SED templates to model the baseline spectroscopic behavior of various SNe Ia in both the optical and NIR wavelength regimes (see, e.g., Hsiao et al., 2007; Hsiao, 2009; Lu et al., 2023). These templates have been parameterized by stretch-like quantities such as $\Delta m_{15}(B)$ and s_{BV}, as well as relative phase in time to describe time-dependence. By doing so, these templates can conveniently infer an SED that may be used to determine time-dependent K-corrections to observed photometry, which are required to correct observed photometry for host-galaxy redshift (Oke & Sandage, 1968).

Our primary goal in this project is to employ the use of PCA to create time-dependent models that describe the variations observed in both optical and NIR spectra. In this way, it should be

possible to provide optical spectroscopy of any SN Ia to a model and perform an extrapolation process that infers the spectroscopic behavior of that SN Ia at that phase in time in the NIR with some known uncertainty. This estimation could be used in a similar manner to templates — to estimate an SED in order to perform K-corrections and bolometric luminosity calculations at any point in time. This is especially useful as optical and NIR photometry of the same SN Ia are not often obtained simultaneously, especially at early times.

This method is similar to the template method of inferring SNe Ia SEDs established by Hsiao et al. (2007) and extended by Lu et al. (2023), as both techniques make use of PCA to formulate a model. However, there are some key differences that should lead to different results. Most importantly, this work does not strictly parameterize models by $s_{\rm BV}$, but rather by using inherent correlations between the spectroscopic features themselves. In general, this may allow more flexibility in how NIR behavior is predicted using optical information. It may be shown that some quantities that describe the shape of light-curves cannot fully parameterize spectroscopic or photometric behavior. For example, Burrow et al. (2020) show spectroscopically dissimilar SNe Ia that exhibit similar values of $s_{\rm BV}$ (see Figure 3.18 in Chapter 3). SNe Ia with similar $s_{\rm BV}$ may also show variability of the secondary photometric light-curve peak seen in the NIR, which is seen in Chapter 4. Papadogiannakis et al. (2019) show varying levels of spread in the time of secondary maximum in the r and i bands for SNe Ia of similar $s_{\rm BV}$. Therefore, it could be that models that are not restricted to parameterization by $s_{\rm BV}$ may produce less uncertain results. In this way, this work is intended to be complementary to Lu et al. (2023) and provide an alternative outlook on the problem.

In addition, we also provide an analysis of how our results in the NIR may be influenced by the classification of a SN Ia established by Branch et al. (2006), which we refer to here as the Branch group classification. These Branch groups consist of the core-normal (CN), shallow-silicon (SS), broad-line (BL) and cool (CL) groups, and they have been shown to be robust in a detailed cluster analysis (Burrow et al., 2020), as shown in Chapter 3. This classification system typically uses the measured pEWs of Si II $\lambda6355$ and Si II $\lambda5972$ to classify SNe Ia into these four groups with what have been shown in previous studies to be unique spectroscopic and photometric properties. Photometrically, CNs and BLs have been shown to exhibit moderate values of $s_{\rm BV}$, typically between $0.7 \lesssim s_{\rm BV} \lesssim 1.1$, with clearly distinct Si II $\lambda6355$ features between them (see Figure 3.18, taken from Burrow et al., 2020). SSs are usually associated with luminous, slow-declining SNe Ia, whereas CLs are dimmer with faster-declining light-curves. In this chapter, we examine how this optical classification scheme may relate to varying spectroscopic behavior seen in both the optical and the NIR using PCA.

In Section 5.2, samples and selection criteria used to inform our PC models are discussed. In Section 5.3, the methods used to create such models are discussed, as well as how these models are

used along with a given optical SN Ia spectrum to predict its NIR spectrum. Section 5.4 illustrates examples of the prediction process for cases at different phases and for different subtypes of SNe Ia. Finally in Section 5.5, we discuss how the resulting PCs may relate to the Branch groups as well as $s_{\rm BV}$, and this extrapolation method is compared with spectral templates, both in spectroscopy and bolometric luminosity.

5.2. Data

5.2.1. Data Source & Preprocessing

To represent the population of observable SN Ia spectra in both the optical and NIR, we make use of two different series of spectroscopic SNe Ia observations. The compiled spectra used in this project are described here.

The optical range of our data is provided by both CSP I+II sets of spectra (see Section 2.4 for a detailed description). This combined data set consists of spectra of 364 unique SNe Ia. The NIR component of the data used here is the FIRE NIR spectra (also detailed in Section 2.4), which is the same used by Lu et al. (2023). FIRE captures adequate signal at NIR wavelengths where considerable telluric absorption typically occurs, allowing for reasonable corrections to these regions, albeit still with a relatively low signal-to-noise ratio and high uncertainties. Acquiring as accurate spectroscopic measurements as possible of SNe Ia in these wavelength regimes is crucial to achieving an SED suitable enough to calculate accurate bolometric light-curves and K-corrections. We choose to use this NIR set of spectra because this allows for consistency when comparing with NIR spectral templates of Lu et al. (2023), which is one goal of this project. In addition, the SNe Ia observed with FIRE were purposefully selected to have good overlap with many CSP II targets captured in the optical. These FIRE spectra include observations of 142 unique SNe Ia in total; however, using the same selection criteria as Lu et al. (2023), a total number of 94 SNe Ia are included in this NIR sample. These spectra typically extend redward to around 2.4 μm.

Every SN Ia used to inform models here also has optical photometric observations from CSP I+II (Krisciunas et al., 2017) such that light-curves have been fit to them using the light-curve-fitting software, `SNooPy` (Burns et al., 2014, 2018, N. Suntzeff et al, in prep). This is so that photometric quantities required for this analysis, such as time of B-maximum and host reddening, are inferred for each SN Ia. There are 115 SNe Ia from CSP I with both spectra and these photometric quantities available, and similarly 161 SNe Ia from CSP II.

Ideally, each spectrum in the data set would be corrected such that its synthetic photometry would match the observed photometry at the same time. This may be done using `SNooPy`, for example, where a best-fit smooth basis spline with knots at the effective wavelengths of each photometric filter is multiplied by the spectrum such that the result yields synthetic photometry closest to the observed values. However, as this project deals with both optical and NIR spectra,

this means simultaneous values of optical and NIR photometry must be available for each spectrum used. Even with light-curve interpolation for each individual filter, only a fraction of SNe Ia in our data set have optical photometry available to allow for this correction, and a much smaller fraction have NIR (Y, J, and H) photometry. Because of the non-parametric nature of the spline-based color-matching technique, the correction is useful for a study such as this only when both optical and NIR photometry are available. Therefore, the spectra used here are not corrected for observed photometry, and our models do not account for this source of variation. However, the CSP I spectra have been shown to give accurate optical colors (Stritzinger et al., 2023) including in $r - i$, which includes the wavelength region where the optical and NIR spectra are merged here. Hsiao et al. (2019) also showed that the CSP II NIR colors obtained from synthetic photometry were accurate to $0.03 - 0.08$ mag.

To eliminate as much other known, systematic variation in the training data sample as possible before a model is calculated, some additional preliminary processing has been performed. Three optical telluric features at 5876–5910 Å, 6849–6929 Å, and 7576–7688 Å were removed from each spectrum. Then, each spectrum was put into the rest frame and corrected for Milky Way (MW) extinction using reddening estimates of Schlafly & Finkbeiner (2011) with $R_V = 3.1$ and a host-galaxy color-excess value $E(B - V)_{\text{host}}$ inferred by `SNooPy`. These corrections were performed using the CCM law given by Cardelli et al. (1989). Note that host-galaxy extinction is less certain than MW extinction; therefore, some variability due to host extinction could be present in PCs described later in Section 5.3.

5.2.2. Sample Selection

As discussed in Section 5.1, the goal of this extrapolation method is to use the available data to train a time-dependent model that describes both optical and NIR spectroscopy. Rather than compiling all available observed spectra into a single model by the use of some interpolation technique such as Gaussian processes, we have chosen to generate multiple models, with each model representing SN Ia spectra at some point in time. Specifically, each model will represent an integer day past B-band maximum (i.e. day -1, day $+0$, day $+1$, etc.). We chose to incorporate time-dependence in this way to eliminate possible bias; for example, two SNe Ia of different subtypes may appear spectroscopically similar in some aspects at different phases in time (e.g., Yarbrough et al., 2023). In doing this, the effects of time evolution should be minimally captured by our individual models. In addition, hereafter in this paper, maximum light will always refer to the time of B-band maximum.

The sample of spectra selected from our total data set for each individual model representing some time t is chosen in the following way. First, for every SN Ia that has spectra observed before and after t, we interpolate the spectra using a GPR if there are enough spectra and there is an

observed spectrum close enough to t. We find that 3–4 observed spectra within approximately -10 to $+50$ days of maximum light provide a good interpolation with low propagated uncertainties from the GPR if at least two of those observation times surround t — otherwise it would require extrapolation rather than interpolation. If the SN Ia is sparsely observed or is otherwise unable to be interpolated, we choose the spectrum observed closest to time t within -5 to $+10$ days relative to t. This window is asymmetric to allow for more spectra to lie within the window, while assuming observed spectra change more dynamically for the same SN Ia at earlier times. Although using such a large window may introduce some time evolution into the model, we find that this is favorable *in lieu* of the smaller sample sizes that a smaller window would generate. Larger sample sizes would produce smaller uncertainties, which are discussed in Section 5.3.4.

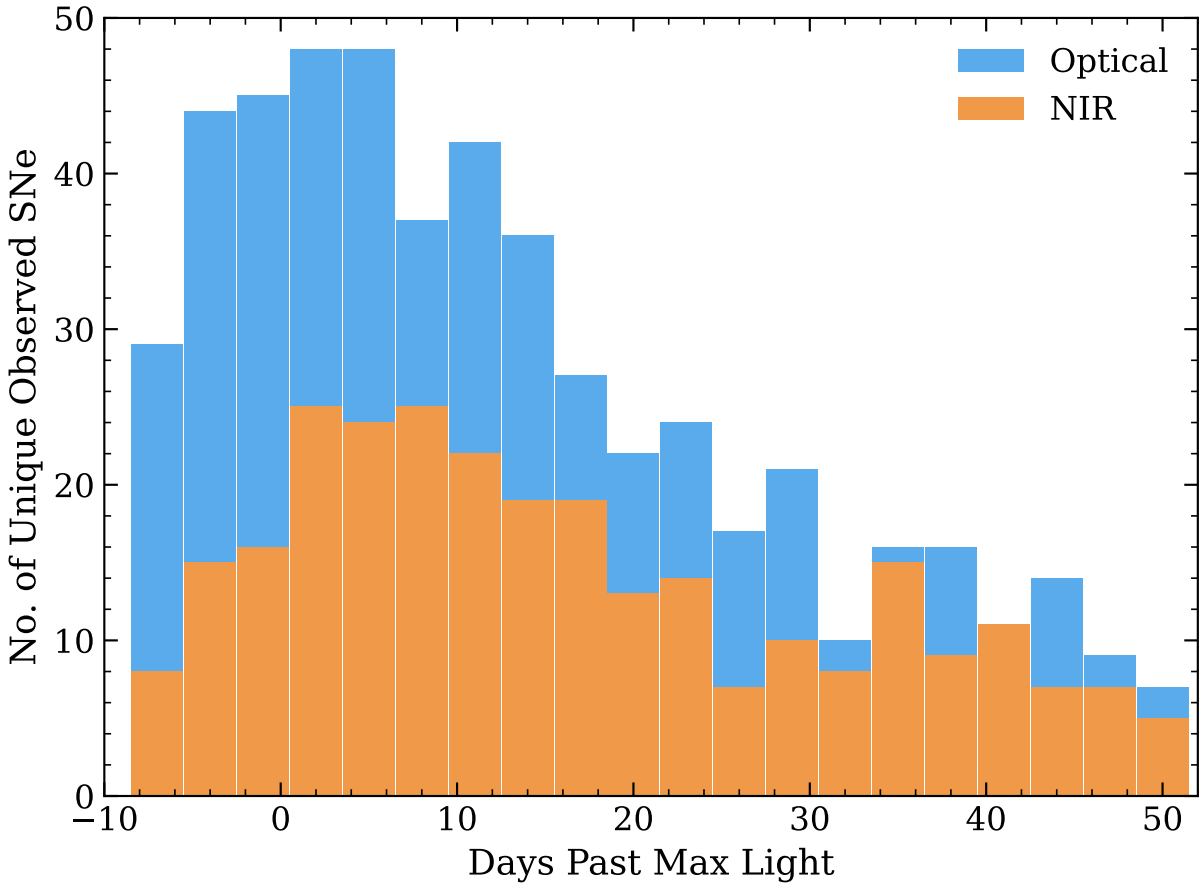


Figure 5.1: Bars here show the number of unique observed SNe Ia in the optical and NIR data sets as a function of time past B-band maximum in three-day bins.

To reiterate, small sample size is the primary difficulty in this analysis. To illustrate the optical and NIR data sets used, in Figure 5.1 we show a histogram of the number of unique SNe Ia from each data set with at least one observation within the given bins, which are three days wide, ranging

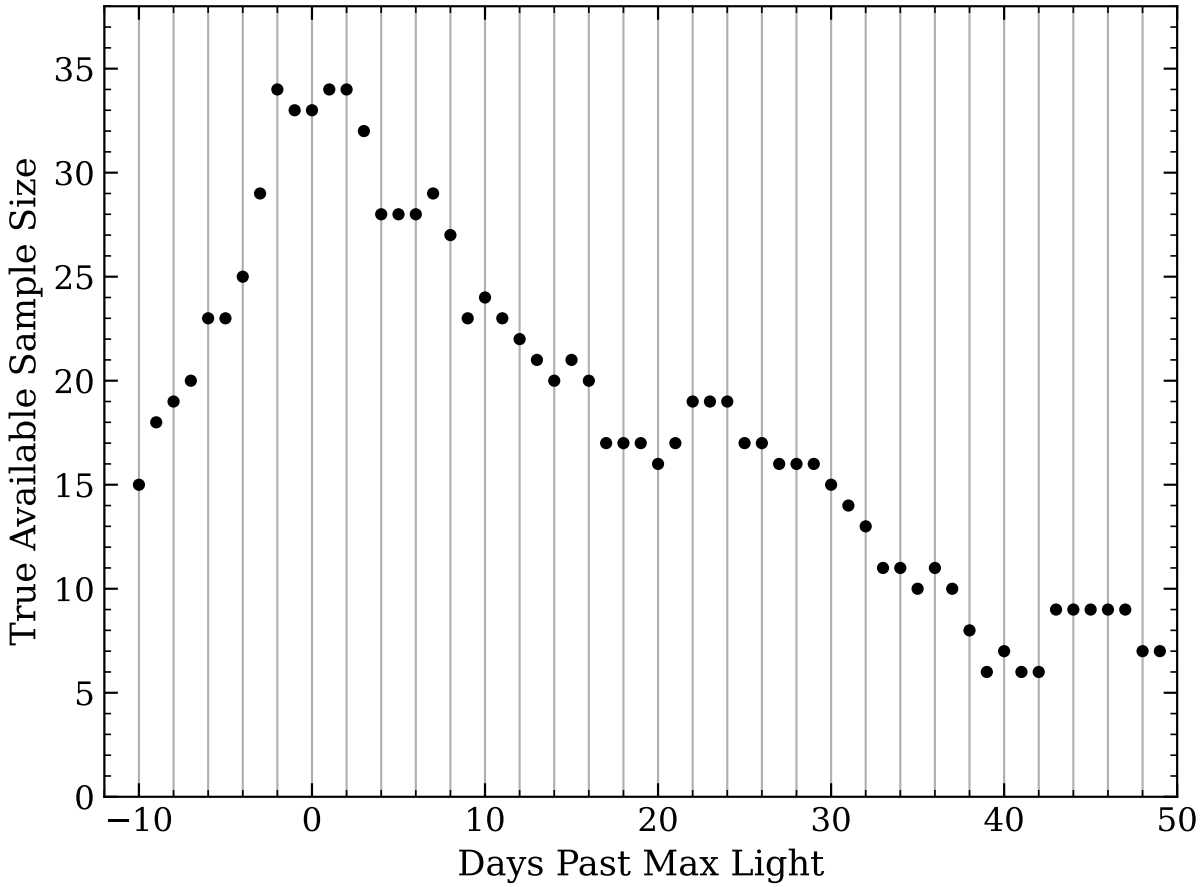


Figure 5.2: The true available sample size for our models as a function of time, accounting for both optical and NIR spectra eligible to have a representative spectrum for that time (described in Section 5.2.2).

from -10 to $+50$ days relative to maximum light. This only accounts for those we use for the analysis; that is, those that have had light-curves fit, etc. Both samples contain more observations at times near maximum light, implying that our models will be more accurate when used to extrapolate spectra observed near the epoch of maximum light. This may be more useful as it is often more difficult to arrange observations of the same SNe Ia concurrently in the optical and NIR at earlier times, and the early phase characterizes the light-curves of SNe Ia.

In addition, Figure 5.2 shows the true sample size available to each model at each phase t relative to maximum light. This accounts for the specifications given in the previous paragraph, including SNe Ia whose spectra may be interpolated or otherwise are within -5 to $+10$ days of t. Accordingly, the sample sizes available for our models at early times are typically around 20–30 SNe Ia, peaking at 34 SNe Ia available when attempting to do an extrapolation of a spectrum taken around $t = +0$ days. This is still a low sample size, however PCA is chosen for this analysis because it tends to

work well with low sample sizes due to its simplicity.

Fortunately, the optical and NIR data sets have good combined wavelength coverage. This allows for two spectra chosen from the two data sets to be merged into a continuous spectrum of a SN Ia representing time t that spans the optical and NIR. Nearly all optical and NIR spectra have shared wavelength coverage between 0.81–0.84 μm. Choosing the higher value of this range to remove optical fringing in the NIR data, we choose to scale the NIR spectrum by a constant factor that allows the two spectra to be equal at 0.84 μm. As the optical and NIR spectra are separately smoothed prior to merging, described in more detail in Section 5.3.1, noise around 0.84 μm should not affect this scaling factor much. The two spectra are then merged into a full optical+NIR spectrum for that SN Ia, ignoring everything redward of 0.84 μm for the optical spectrum and blueward of 0.84 μm for the NIR spectrum. This entire process is performed for each supernova until a complete sample of spectra representative of each time t that span both the optical and NIR is achieved.

As stated in Section 5.2.1, the individual optical and NIR spectra are unable to be consistently color-corrected. Consequently, each final merged spectrum is also not corrected to match the photometric colors. We discuss potential impacts of this on the results in more detail in Section 5.5.1.

5.3. Methods

5.3.1. PCA Method

PCA is a dimensionality-reduction technique that allows one to describe the variation in a sample of training data made up of N vectors of the same dimension. In this case, the sample here is a set of compiled spectra described in Section 5.2.2. To transform each spectrum into a vector in the same space, each spectrum in the training sample is interpolated to the same points in wavelength space. This interpolation is performed via a GPR using `Spextractor` (see Section A.1). As a result, the spectra are also smoothed, eliminating much of the noise each spectrum may contain.

To account for any spectra in the sample that have relatively high flux uncertainty, for example in telluric regions, a variant of standard PCA called expectation-maximization PCA (EMPCA) (Roweis, 1997; Bailey, 2012) is used. EMPCA uses estimates of uncertainty of input data to weight each sample such that the variation described by eigenvectors is not dominated by high-noise samples or extreme outliers. It is also particularly useful when working with data sets with missing points.

Typically one weights each point by a quantity similar to the inverse variance $1/\sigma^2$; however, in our testing we found this particular weighting scheme leads to noisy eigenvectors, possibly due to the fluctuation in the errors from point to point on the spectra. This does not significantly affect the end results when reconstructing spectra with more eigenvectors (discussed in Section 5.3.3), however when attempting to reconstruct spectra with fewer eigenvectors, the predictions become much noisier. We instead make use of a weighting scheme that is linear in variance, such that the

weight of the ith wavelength point of each sample is given by $1 - 0.75\ \sigma_i^2/\sigma_{i,max}^2$, where $\sigma_{i,max}^2$ is the largest variance exhibited at that wavelength across all samples. The factor of 0.75 is inserted to prevent high-variance data from being given zero weight, and it is an arbitrary parameter that describes the relative weighting between high- and low-variance data points. This min-max scaling of flux variance chosen for the weighting scheme is not standard, but we find it works well in limiting the weight of spectra containing relatively high uncertainties.

5.3.2. PC Training

Once the sample spectra are interpolated in wavelength space (as discussed in Section 5.3.1), applying PCA will effectively be looking at how each individual wavelength point varies for each spectrum in the sample, and connecting them to form spectrum-like eigenvectors. Because of this, we may choose any set of wavelength points for our training data. However, to be consistent and eliminate as many variables from this analysis as possible, we choose to use spectra only in the range of 0.50–2.30 μm. We typically do not use data blueward of this region because we find in testing that including Fe-group features blueward of 0.50 μm adds too much variation to the training sample given the available sample sizes, leading to lower quality fits and less consistent predicted results. This variation stems from sources such as line blanketing and Ni mixing (see, e.g., Baron et al., 2006; Hoeflich, 2017; DerKacy et al., 2020). The exact reason that adding more information to the model blueward of 0.50 μm leads to inferior results is unknown. However, this seems to suggest that the features in this bluer optical region are not correlated with much of the NIR in a way that PCA can describe easily with the limited data available. Perhaps studying the variation in the UV+optical region is the subject of future work when more data is available, but with our current data limitations it is outside the scope of this project.

Once a set of spectra that represent the same time and the same points in wavelength space is compiled, they need to be normalized and standardized before performing PCA. First, each spectrum is individually normalized by its mean value of flux between 0.50–0.84 μm. This is chosen because normalizing by the mean (rather than, for example, the maximum flux value) eliminates some of the variation between spectra due to intrinsic differences in color or other related quantities. It is important to note that PCA is quite dependent on the choice of normalization used, so another normalization may produce different results.

After this normalization, the training data are all uniformly standardized. This includes subtracting the total mean spectrum μ_F of all spectra in the sample, and subsequently dividing them by the standard deviation of the sample, σ_F, such that the variation of flux in the sample at each wavelength point has a standard deviation of 1. This ensures that the eigenvectors are describing true variation from the mean, and that each wavelength point is on the same scale. This is important because there would be a larger magnitude of variation in the optical part of a spectrum than the

NIR region, as spectra typically exhibit higher values of optical flux.

5.3.3. Prediction Using PCA Eigenvectors

EMPCA produces a basis of eigenvectors that describe most of the variation in the training data sample. It is desirable for these eigenvectors to capture either physical or non-physical correlations between different regions in wavelength space. As an example, the third PC may describe that the sample typically exhibits increased O I and NIR Ca II simultaneously, which would hint that the two features may be correlated to some extent for SNe Ia. Indeed this is the case for spectra at maximum light, as shown in Section 5.4.1. A more in-depth discussion of the behavior of eigenvectors for maximum-light spectra is also given in Section 5.4.1.

By making use of these correlations or anti-correlations, if the spectroscopic behavior of a SNe Ia can be predicted in the NIR using an observed optical spectrum, the optical spectrum can be projected onto the optical part of each eigenvector in the basis to get a set of eigenvalues. This projection is done via a simple dot product, treating the spectrum and eigenvectors as vectors in wavelength space. It should be noted that the spectrum first needs to be normalized and standardized in the exact same manner as the training data (discussed in Section 5.3.1). The resulting projection values just correspond to the eigenvalues needed to reconstruct the optical part of the spectrum — which is already known — using the eigenvector basis. In other words, these projection values describe the extent that each eigenvector describes the optical spectrum. To extrapolate the spectrum in the NIR region, the optical eigenvalues may then be used with the full eigenvectors that span both the optical and NIR regions to form a linear combination that reconstructs the spectrum in both the optical and NIR. This is possible because each eigenvector individually describes both optical and NIR variation simultaneously. Because the data have been standardized, the variation seen in the optical and NIR will be on the same unit-scale. This process may also be performed with any subset of the eigenvector basis, but we assume that the reconstruction is more precise with more PCs used, as each PC accounts for further variation in the sample. However, higher-degree PCs could be capturing pure noise and are therefore often irrelevant to much of the analysis and the prediction.

5.3.4. Uncertainty Measurement

A measurement of uncertainty for these predictions is given by accounting for two different sources of variation. The first of which comes from uncertainty in how well the eigenvectors are able to be used to make the exact same predictions for the training sample itself, including the same region projected onto the eigenvectors to make a prediction. More specifically, the same prediction is made for each SN Ia in the training sample, and then the sample standard deviation of these predictions is used as part of the total uncertainty. This treats each wavelength point as

purely uncorrelated, that is, the noise is treated as white noise. However, the wavelength points in the spectrum are correlated by the underlying physics. Chapter 6 provides a solution to this by modeling the residuals used to calculate uncertainty. However, the method requires testing and analysis (which is the subject of future work), and therefore we adopt the white-noise approximation for this analysis.

Secondly, if there is uncertainty in the flux in the known part of the spectrum that is being extrapolated, this could affect the prediction and therefore must be represented in the total uncertainty. To account for this, we assume that the flux uncertainty is distributed normally and draw 500 samples of the spectrum. Using these samples we then make the same predictions using the same projection and prediction regions, and the standard deviation in these predictions is calculated. This is then added in quadrature to the previous source of uncertainty to make up the total uncertainty in the reconstructed spectrum.

5.4. Results

5.4.1. Maximum-Light Eigenvectors

When a standardized training data set has been established to represent some point in time, a model may be created that includes a set of N eigenvectors which describe the N highest degrees of variation around the mean of that sample of spectra. Figure 5.3 shows the first four eigenvector spectra that describe the highest amount of variance in the training sample of 33 unique SNe Ia spectra observed at or around maximum light ($t = 0$) given by the EMPCA algorithm. Note again that these eigenvectors are standardized, showing unit-scaled variation around the mean spectrum μ_F.

In Figure 5.3, it can be seen that the first PC seems to represent a broad correction to spectra. This could be related to some stretch- or color-like quantity; however, this is unlikely since in Section 7.3 we find that the first PC is not significantly correlated with the color-stretch quantity $s_{\rm BV}$, but rather a linear combination of PCs may describe $s_{\rm BV}$. The other PCs seem to describe more individual feature behavior. The typical location of some of the major SN Ia spectroscopic features are indicated and colored in Figure 5.3. Because of the abstract nature of these eigenvectors, it is unclear visually if an individual PC captures all information about individual features. For example, both the third and fourth PCs seem to explain variation around Si II $\lambda6355$, although one may be explaining its depth or pEW while the other makes corrections in accordance with the Si II $\lambda6355$ velocity extent. This is discussed in a more quantitative manner in Section 7.2.

The cumulative fraction of explained variance for each subsequent eigenvector is shown in Figure 5.4. As may be seen, 10 eigenvectors explain up to 94.8% total variation of the training sample. Because the spectra have been standardized at all wavelengths, this demonstrates that 10 eigenvectors is sufficient in reconstructing typical spectra belonging to the same population of

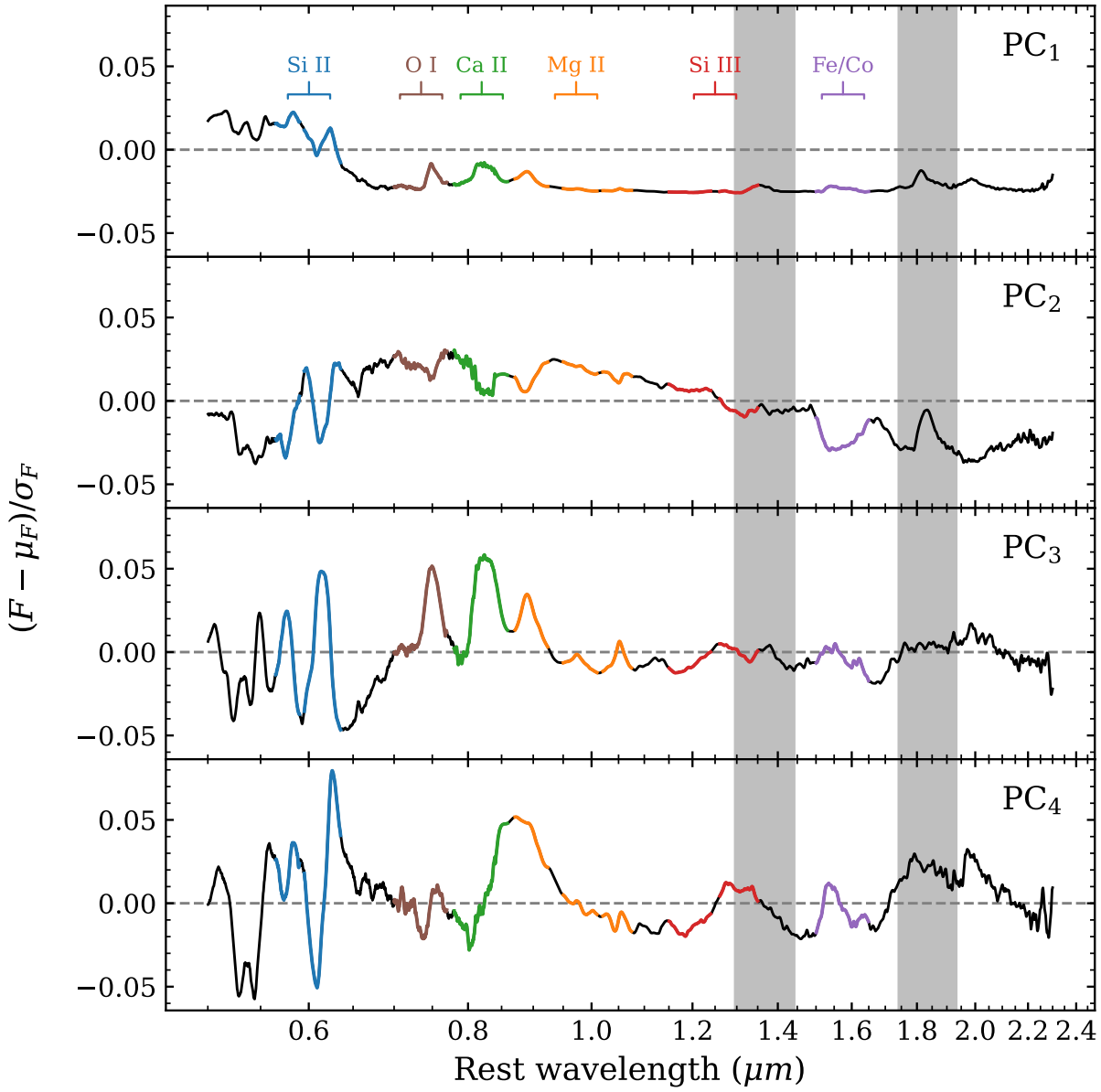


Figure 5.3: The first four PC eigenvectors representing $t = 0$ that describe the highest degrees of variation in the near-maximum-light standardized training data set. These vectors lie within wavelength space, and note that these eigenvectors represent unit-scaled variation from the mean spectrum. Grey shaded regions indicate the larger telluric regions in the NIR.

SNe Ia.

We speculate that much of the variance that is unexplained by the PCs may stem from the relatively high uncertainty in the corrections to NIR telluric regions made for the FIRE data set. In the near future, the negative impact of telluric features on data driven models may be mitigated with JWST observations of SNe Ia (see, e.g., Kwok et al., 2023).

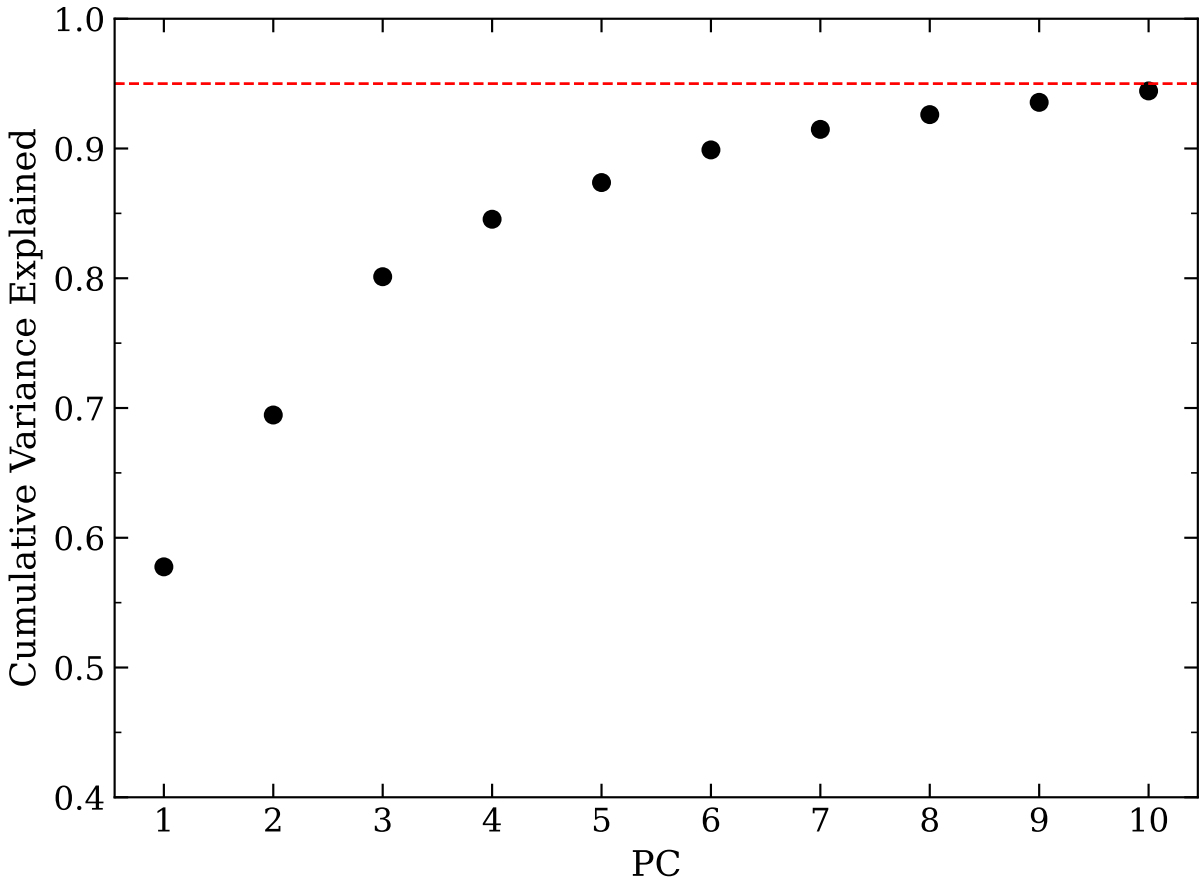


Figure 5.4: Cumulative fractional variance of the training spectra explained by the eigenvectors in Figure 5.3. The red dashed line represents when a total of 95% explained variance is achieved.

5.4.2. Prediction Near Maximum Light

Unfortunately, there is much complexity in the time- and wavelength-dependence of these predictions. There are also not many homogeneously observed spectra available in both the optical and NIR at nearby phases in time outside the sample used in the PCA here. Because of this, a formal test and quantitative statistical analysis of how well this method actually performs is difficult to achieve. Instead, to illustrate the potential of this method, in this section we show a few near-maximum-light examples of this extrapolation procedure. In particular, we show this in Figure 5.5 for SNe 2011fe, 2021fxy, 2021aefx, and 2022hrs, which are a few diverse, well-observed SNe Ia that are all outside the training sample used to create the maximum-light eigenvectors shown in Figure 5.3 and discussed in Section 5.4.1. Because the full optical–NIR range of these SNe Ia is required here for comparisons, the closest optical and NIR observations to maximum light available to us were merged into the black line shown in Figure 5.5. It should be noted that some of these NIR spectra were observed days apart from their optical counterparts, which may affect the results,

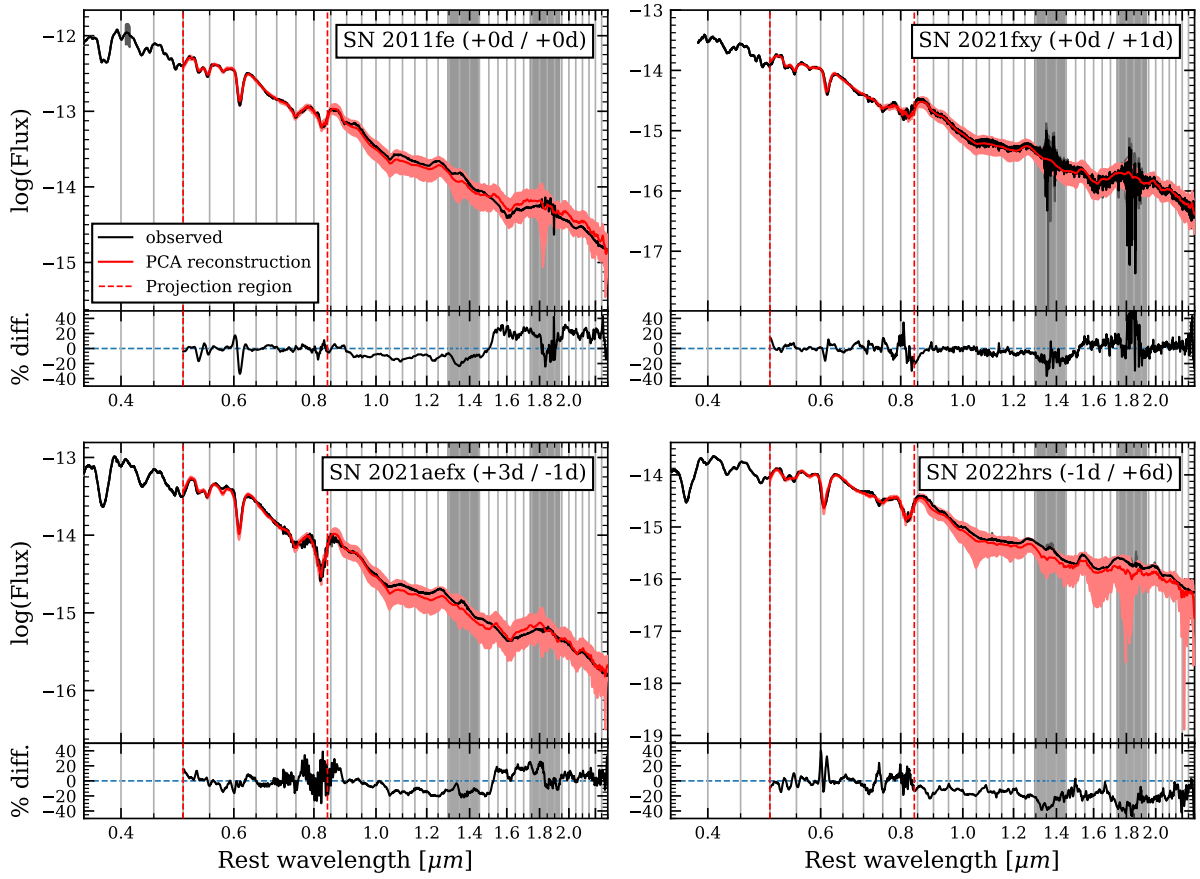


Figure 5.5: The extrapolation of different SNe Ia near maximum light shown in red, compared with the observed spectrum in black. Uncertainties in the predictions are given by the red shaded areas. Grey shaded areas specify telluric regions in the NIR. The region between red dashed lines is projected onto the maximum-light eigenvectors to make this prediction. Percent-difference between the prediction and flux is given on the bottom of each panel; note this difference is not calculated in log-space. Phases shown are the phases of the optical and NIR spectra, respectively, that were merged to create the observed spectrum.

although this is unavoidable due to data limitations.

Each of the panels of Figure 5.5 show a prediction for 0.50–2.30 μm indicated in red. They are compared with their respective rest-frame observed spectrum in black, part of which was used to generate a prediction. These observations are referenced and further elaborated upon below. Vertical dashed lines indicate the 0.50–0.84 μm region that is projected onto the PC eigenvectors to make this prediction. A measurement of uncertainty for the prediction (see Section 5.3.4) is shown in the light red shaded area. The two NIR telluric regions between 1.2963–1.4419 μm and 1.7421–1.9322 μm are shaded in gray. Percent-differences between the PCA reconstructions and

observed spectra are shown below each spectrum; note that these differences are in flux space — not in log-flux space.

The top-left panel of Figure 5.5 shows the extrapolation of a near-maximum spectrum of SN 2011fe (Mazzali et al., 2014). Calculating Si II line velocity $v_{\text{Si II}}$, pEW(Si II $\lambda6355$), and pEW(Si II $\lambda5972$) using this spectrum, and using values with GMMs from Chapter 3, SN 2011fe is classified to be a CN SN Ia with a probability of 74.2%. Because SN 2011fe has historically been so well-observed and studied, this SN Ia is included here to illustrate our extrapolation performance on a quintessential test case using this spectroscopically normal SN Ia.

Looking at the prediction for SN 2011fe in Figure 5.5, the reconstruction of the spectrum performs well within the region projected onto the eigenvectors (indicated by dashed lines), and the uncertainty is relatively small. There is at most a nearly 35% difference in the flux of observation and prediction in this region where the depth of Si II $\lambda6355$ is not fully reproduced. Outside this region — past 0.84 μm where the extrapolation actually occurs — the general trend of flux is captured by the prediction, as well as most individual feature behaviors outside the two large telluric regions in the NIR. However, in this region the uncertainties are much larger. In addition, at around 1.5 μm the prediction begins to deviate from the observations. Between 0.9–1.5 μm, the median percent-difference in the flux is -10.5% — not including flux in telluric regions, which will continue to be ignored for other median values provided. After 1.5 μm the same median jumps to $+18.8\%$ difference in the flux. Although these figures seem high, when integrating the quantity λF_λ across the extrapolated spectrum between 0.84–2.30 μm, ignoring telluric regions in this integral, the difference is only -3.1%. This difference in the integral gives perhaps a better figure of merit when making corrections to photometry than looking at only the median percent-difference in the flux.

In the top-right panel of Figure 5.5, a near-max spectrum of SN 2021fxy (DerKacy et al., 2023) is shown alongside its NIR prediction. The observation is, similar to the SN 2011fe spectrum from Mazzali et al. (2014), a merge of an optical spectrum around $+0$ days and a NIR one around $+1$ days. This process of merging normalizes the NIR spectrum to the optical spectrum by the integrated flux within 0.84–0.88 μm, and performs a weighted average to get the new merged flux for that region. SN 2021fxy was determined by GMMs to be a SS SN Ia with a probability of 95.5%. The prediction here is quite similar to that of SN 2011fe, however all merits of fit discussed for SN 2011fe were more true to the observation for SN 2021fxy, even with a noisier spectrum, aside from a poorer fit to NIR Ca II around 0.8 μm. In the 0.9–1.5 μm region, the median difference in flux was -5.6%, and past 1.5 μm the median is $+4.0\%$. Integrating the observed and the predicted spectra in the same way as before shows a difference of -2.9%. It seems that although the eigenvectors shown in Figure 5.3 appear to show substantial variation of Si II features, shallower Si II lines that are projected onto these eigenvectors do not lead to a less accurate prediction.

The bottom-left panel of Figure 5.5 shows a near-maximum spectrum of SN 2021aefx (Burns et al., 2021; Hosseinzadeh et al., 2022) as well as its extrapolation. The observed spectrum is merged in the same way as SN 2021fxy with an optical spectrum at $+3$ days and a NIR spectrum at -1 days. SN 2021aefx is classified with GMMs as a CN with a probability of only 57.6% and therefore may exhibit some atypical spectroscopic behavior relative to most CN SNe Ia. Within 0.50–0.84 μm, the fit to the observation including Si II is good except for the noisier region around 0.80 μm where there is more fringing of the original optical spectrum. As such, the fit to NIR Ca II is poor. Still, the NIR prediction of this spectrum performs reasonably, with a median percent difference in flux of $+7.8\%$ until 1.5 μm where this difference increases, reaching a median of -13.6%. SN 2021aefx continues to compare well to the observation with a difference in integrated flux of -2.7%.

Finally, the bottom-right panel of Figure 5.5 shows a 100%-likely BL SN Ia (rounded up from a GMM calculation of 99.98% probability), SN 2022hrs (Burns et al., 2021, P. Brown et al., in prep.). The observation shown here is made from merging an optical spectrum at -1 days and a NIR spectrum at $+6$ days. This SN Ia provides an interesting test case as no BL SNe Ia exist in the maximum-light training data set used to make this prediction. This possibly explains the relatively poor fit to the observation in the 0.50–0.84 μm region, and especially in the Si II $\lambda6355$ feature. It seems that a linear combination of PCs that is able to describe the depth of Si II $\lambda6355$ alongside other features not obtainable without more BL data included in the training sample. There are also larger uncertainties involved with this spectrum than those of our other test cases. The most glaring difference between this SNe Ia and the other three examples is that there seems to be a $\sim$20% near-constant negative offset between the prediction and observation after 0.84 μm. This again may be due to the eigenvector training sample not including BL SNe Ia, and therefore a BL such as SN 2022hrs is not represented well enough to accurately reconstruct its spectrum in either the optical or the NIR. In fact, we see this later in Section 5.5.1 with SN 2014J, a CN SN Ia with high $v_{\mathrm{Si\,II}}$, similar to BLs. This offset may also be due to the seven-day difference in observation time between the optical and NIR spectrum used to create the observation shown. With the offset given by the prediction for SN 2022hrs, the difference in integrated flux between the prediction and observation is -17.7%, substantially greater than the other three example SNe Ia.

Although Figure 5.5 demonstrates a good comparison between predictions and their respective observed spectra, it is difficult to determine any significant difference between the predictions, and whether or not the characteristic differences between SN Ia subtypes are being well-represented by the predictions. For this reason, we also provide a direct comparison of the predictions themselves in Figure 5.6. The left panel shows the observed spectra from each SN Ia in Figure 5.5 past 0.84 μm. The spectrum in the telluric regions of SN 2021fxy was omitted for visual clarity due to the high amount of noise. In the right panel, the corresponding predictions (red lines from Figure 5.5) are

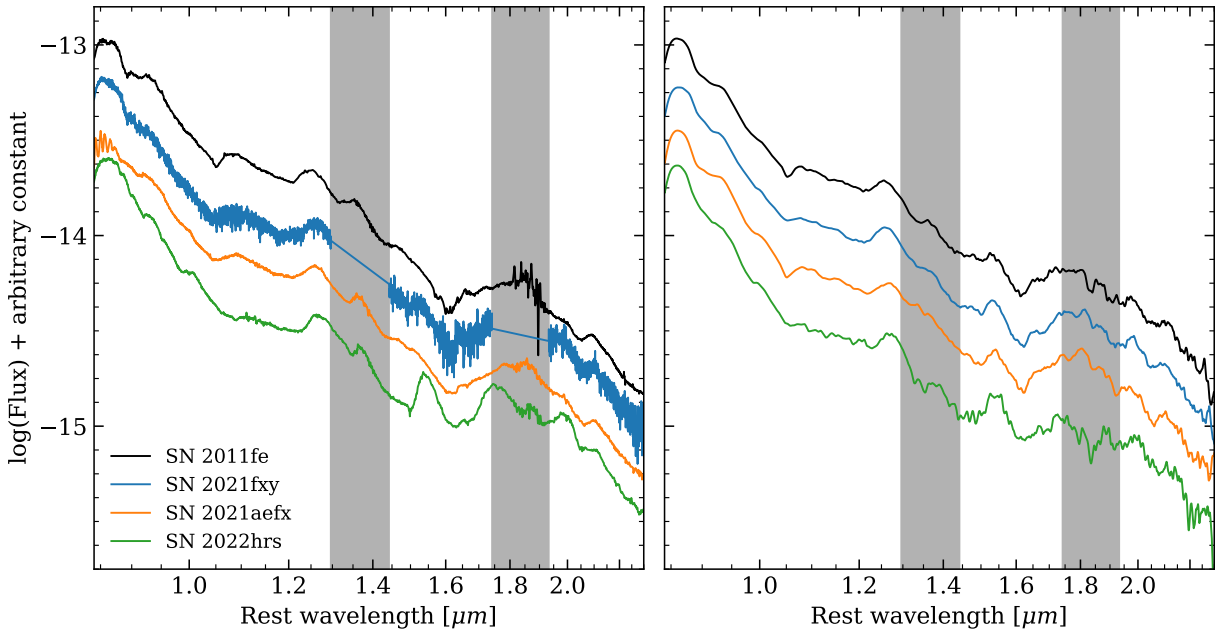


Figure 5.6: A direct comparison of the four near-maximum-light observations from Figure 5.5 (left panel) and their corresponding extrapolations (right panel) in the NIR. The extrapolations appear to capture many of the distinctions shown between the observed spectra, except for the Fe II feature near 1.50 μm, which is too strong in the predictions.

plotted. Each spectrum was normalized by the integrated flux of the 0.84–2.30 μm region, ignoring the telluric regions in the grey shaded region, with SN 2011fe as the reference. The spectra are then equally spaced from each other in the figure.

Here it is beneficial to summarize the Branch groups to which each SN Ia here belongs: SN 2011fe and SN 2021aefx are classified CNs, SN 2021fxy is a SS, and SN 2022hrs is a BL SN Ia. Keeping this in mind, the predictions seem to do very well at the feature level in capturing the differences between these SNe Ia. Using line identifications by Hsiao et al. (2019), Mg II around 0.92 μm, 1.09 μm is conserved in the prediction, being stronger for the CNs, slightly weaker for the SS, and weakest for the BL SN Ia. However, the predicted Mg II at 1.09 μm appears much more narrow than the observation for SN 2021aefx. Mg II near 2.10 μm appears strong for the BL SN Ia as well, and the prediction captures the observations well for this feature, including the feature shape. Si III near 1.25 μm is similar for each SN Ia in both the observation and the prediction. Using line identifications displayed by Friesen et al. (2014), the Co II feature around 1.65 μm is also differentiated well in the prediction for each subtype, including the feature shape. The biggest departure from the observation that the prediction contains is the Fe II feature around 1.50 μm. Fe II here is much more shallow for CNs, which is not reproduced by the predictions. In general, though, the feature behavior between SNe Ia of different Branch groups appears to be captured well

by these predictions.

5.4.3. Time-Dependent Predictions

We showed in Section 5.4.2 that the extrapolation process performs fairly well for spectroscopically normal SNe Ia near maximum light. However, when extrapolating at other phases, the results and performance of this technique may lead to less certain predictions, especially due to decreasing training sample sizes. As expected from the sample sizes given in Figure 5.2, at later times the available number of training spectra decreases significantly as there are not as many concurrent optical and NIR observations in the CSP I+II data set. Interpolation at these later times is often not possible, as either the total number of observations of each individual SN Ia are too few, or the observations were too far apart in time to give reasonable interpolations. However, up to around 20 days past maximum light, the available sample size is still around 20 SNe Ia— twice the number of eigenvectors we use to model spectra, which may generate reasonable extrapolations of spectra observed at these phases.

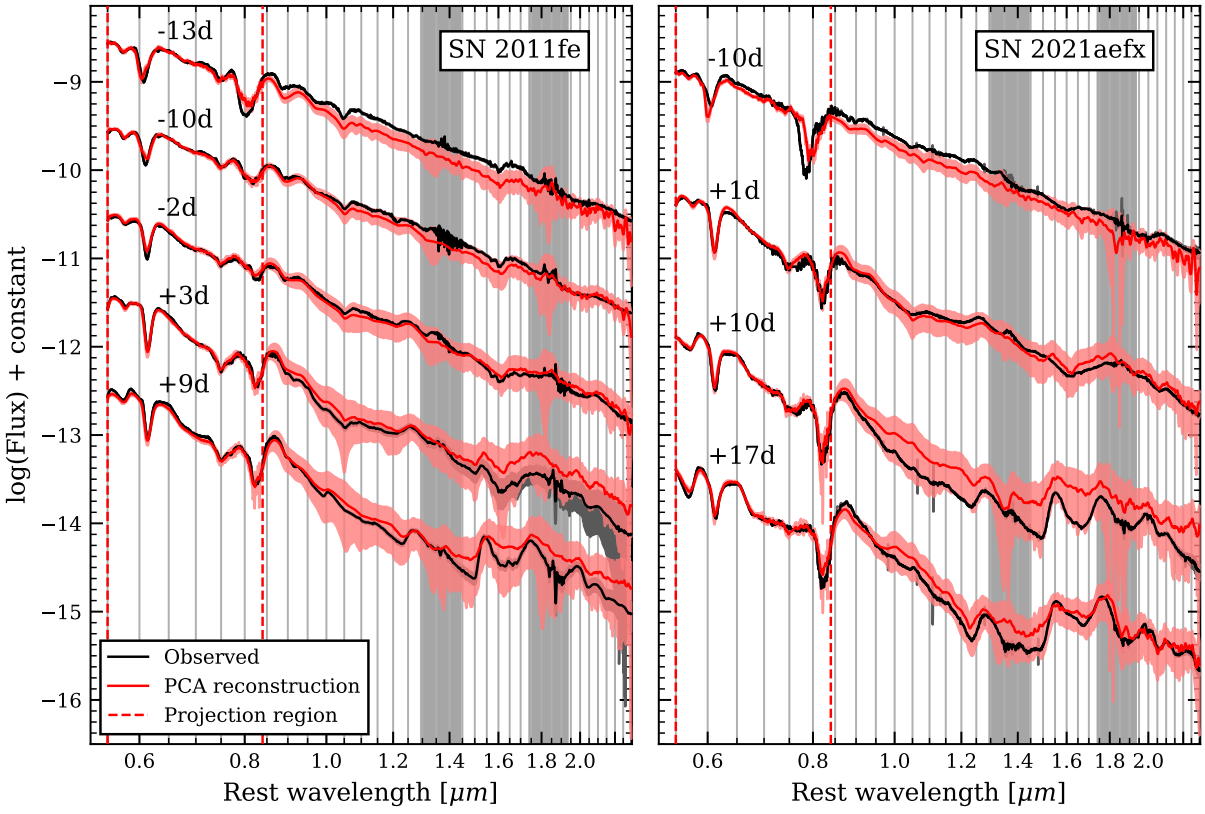


Figure 5.7: Extrapolation for a time series of SN 2011fe and SN 2021aefx. All formatting and notation is the same as in Figure 5.5.

In Figure 5.7 we show the same prediction process as in Section 5.4.2 on a time series of early

spectra of SN 2011fe (Mazzali et al., 2014) and SN 2021aefx (Burns et al., 2021; Hosseinzadeh et al., 2022), as they have good concurrent observations in both the optical and NIR surrounding maximum light. Both SNe Ia yield similar results. At earlier times, most feature behavior is captured, aside from Ca II around 0.80 μm. The prediction for SN 2011fe also introduces a small constant offset from the prediction across the NIR that is not exhibited for SN 2021aefx. This may be due to differences in the flux calibrations between the optical and NIR spectra, which cannot be corrected for in the merging process due to a lack of photometric data for the entire data set (see Section 5.5.1).

At later times, the general trends of the flux of both SNe Ia are captured; however, much of the feature information beyond 1.5 μm appears to be lost, and the uncertainties are much larger. This could mean that there is more variation in this region than can be explained with the limited sample size available.

5.4.4. Prediction Using Si II

For consistency we used the 0.50-0.84 μm region to project onto the eigenvector basis in order to make predictions. However, this region itself is a variable and may be changed. Increasing this range is impractical as, again, using information blueward of 0.50 μm introduces much of the UV-optical variation that is difficult to explain using a limited training sample. Instead, this section focuses on the outcome of narrowing this region to a more meaningful one.

It is clear that the Si II $\lambda6355$ and Si II $\lambda5972$ features in the optical are important observables in the diversity of SNe Ia, as they make up the basis of the Branch classification system, which has been shown to be yield a statistically significant set of groups with distinct characteristics (Chapter 3). Figure 5.8 shows the result of projecting only the 0.55–0.64 μm region consisting of these two Si II lines onto the eigenvector basis to extrapolate into the NIR up to 2.3 μm. This was done for the same SNe Ia as in Section 5.4.2, and the figures show the same information as shown in Figure 5.5; the only change is the region used to make the prediction, shown as red dashed lines. It is important to reiterate that the only information provided to the maximum-light model is this narrow Si II window — aside from the mean flux value of the 0.50-0.84 μm region required to normalize the spectrum in the same way as the training data.

The predictions shown in Figure 5.8 are quite surprising, given the small amount of information provided. Comparing SN 2011fe and SN 2021fxy in Figures 5.5 and 5.8, the predictions are quite similar. For each SN Ia, both cases seem to predict the general flux behavior of the observation fairly well. The uncertainties shown are nearly the same across each spectrum between the two cases, except they are much improved in the case of SN 2022hrs where spikes of larger uncertainties are shown at some wavelengths past 1.5 μm. The most notable difference between these two cases is the inability to reproduce the NIR Ca II triplet well. This seems to suggest that Ca II behavior is

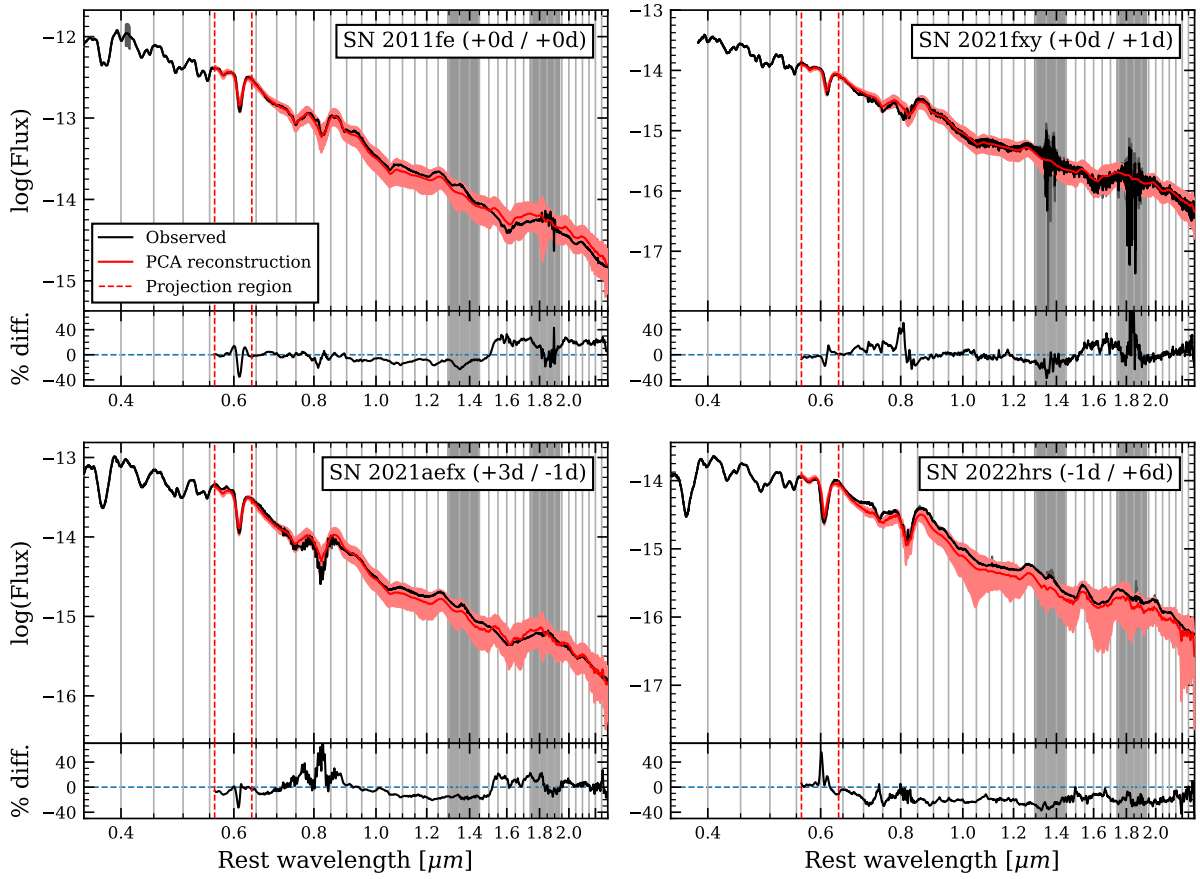


Figure 5.8: This shows the same spectra as in Figure 5.5, however the predictions are now made with a narrower optical region dominated mostly by Si II $\lambda 6355$ and Si II $\lambda 5972$. All formatting and notation is the same as in Figure 5.5.

not strongly dependent on Si II $\lambda 6355$ and Si II $\lambda 5972$ near maximum light.

The integrated flux differences from the observations in this case are -2.4% for SN 2011fe, $+2.0\%$ for SN 2021fxy, -0.4% for SN 2021aefx, and -18.5% for SN 2022hrs. This means that the prediction was marginally improved for SN 2011fe, SN 2021fxy, and SN 2021aefx, and only weakened for SN 2022hrs, which is likely because it is a BL, unlike the SNe Ia in the training set. Given just these few test spectra shown here, no conclusion is drawn regarding which case performs better. However, the similarity in results between the two cases indicates that Si II may be the only requirement for a reasonable prediction in the NIR in general.

5.5. Discussion

5.5.1. Color-Matching with Broad-Band Photometry

As discussed in Section 5.2, neither training spectra nor any spectra shown in Section 5.4 have been color-matched to the broad-band photometry. As such, the differences in the flux calibrations between optical and NIR spectra may result in some training spectra having improperly scaled NIR flux. In other words, the photometric color of the predicted spectrum will be incorrect. This may lead to substantial variation in the training data in the NIR. However, if this variation is in any way systematic, the eigenvectors may describe this variation and be able to correct for it, albeit not perfectly. It is not possible to know if this hypothesis is true without observed photometry for all of our training SNe Ia; however, we can still make this assumption. This assumption may be supported by the first PC in Figure 5.3, which seems to describe a broad correction that accounts for around 58% of the variation (see Figure 5.4) in the training sample. We show later in Section 7.3 that PC_1 does not correlate by itself with s_{BV}, and therefore the first maximum-light eigenvector likely does not describe variation of a stretch-like property of SN Ia light-curves. Instead, it may be correcting for differences in the flux calibration between optical and NIR spectra — if so, the final NIR prediction should be on the appropriate scale of the optical spectrum.

An illustration of this problem is shown in Figure 5.9, which shows a maximum-light prediction of SN 2014J using a merged optical-NIR spectrum (Galbany et al., 2016; Srivastav et al., 2016) in the same way as those in Section 5.4.2. SN 2014J is known to be highly reddened (Amanullah et al., 2014; Foley et al., 2014). We assume host-galaxy color-excess values of $E(B-V)_{\rm host} = 1.24$ mag and $R_{V,\rm host} = 1.44$ (Ashall et al., 2014) to correct the spectrum for extinction using the CCM law (Cardelli et al., 1989), similar to spectra in the PC training data. Even though the merging and prediction process is the same as the other four SNe Ia that yields good fits, the case of SN 2014J is unexpectedly poor. Inside the 0.50–0.84 μm region which was projected onto the eigenvectors, the PCA reconstruction fails to fit the observation for most features in this region, and this is especially true for Si II. It seems as though SN 2014J does not exhibit the same behavior as the training sample and therefore could be an interesting case study on its own. It is worth noting that, although we classify SN 2014J as a CN with a probability of 67.6%, it is also on the faster end of CNs with $v_{\rm Si\,II}$ of 12.0×10^3 km s^{-1} and could nearly be considered a BL SN Ia by traditional classification with only pEWs. However, the GMM probability of belonging to the BL group is 13.5%. It was shown in Section 5.4.2 that this prediction process is not suitable for BLs such as SN 2022hrs, as a constant offset from the prediction was produced. Perhaps the same underlying effect occurs for the case of SN 2014J. This shows that more observations of BL and faster SNe Ia are needed for a complete understanding of the diversity of SNe Ia.

Similar to SN 2022hrs, SN 2014J shows a large and nearly constant offset seen between the

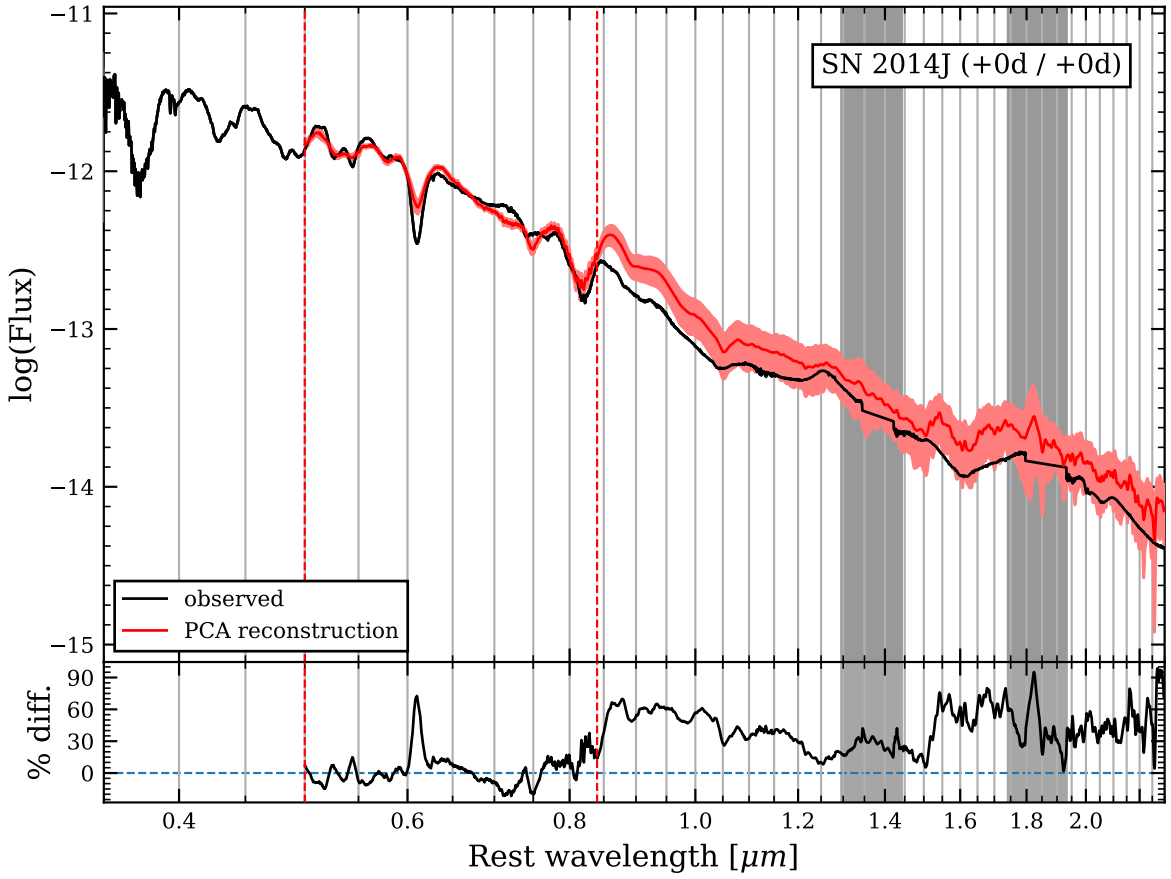


Figure 5.9: Maximum-light extrapolation for SN 2014J. All formatting is the same as that in Figure 5.5. Although SN 2014J is classified as a CN SN Ia, a constant offset is prominently seen between the observation and prediction. This may suggest that photometric corrections to spectra are necessary to strengthen the performance of these extrapolations, which requires more photometric observations in both the optical and NIR.

prediction and the observation after 0.84 μm. Because the observations shown in black are simply merged by scaling the NIR to the red end of the optical flux, there could still be large discrepancies between it and the true SED of the SN Ia, and one would need observed photometry to correct for them. If the first PC were to describe some sort of correction that adjusts the NIR prediction to match the optical, then the offset shown in Figure 5.9 is not a problem with the model, but rather the reference observation. However, it is again difficult to confirm this without observed photometry of the training set of SNe Ia.

Because the differences in the flux calibration between the optical and NIR parts of the training spectra are not completely systematic due to the inhomogeneity of the spectroscopic observations, corrections should still be made to the final predictions. We chose not to do so in Section 5.4 for consistency with the lack of NIR photometric observations. However, it will yield a more accurate

SED approximation with correct synthetic photometric colors if observed photometry is available. It is also worth noting that spectroscopic templates that are parameterized by light-curve quantities are also photometry-corrected separately after a first approximation is given to light-curve fitting procedures such as `SNooPy`.

5.5.2. Comparison with Templates – Spectroscopy

In this section we provide a brief comparison between the extrapolation procedure described in this work and the parameterized spectral templates of Lu et al. (2023). In short, Lu et al. (2023) make use of 331 NIR spectra of 94 SNe Ia and performs PCA on multiple regions in wavelength space to create a model that is parameterized by $s_{\rm BV}$ using a GPR. A substantial statistical comparison of the two methods is not provided here due to the complexity of the predictions and lack of data with which to perform tests.

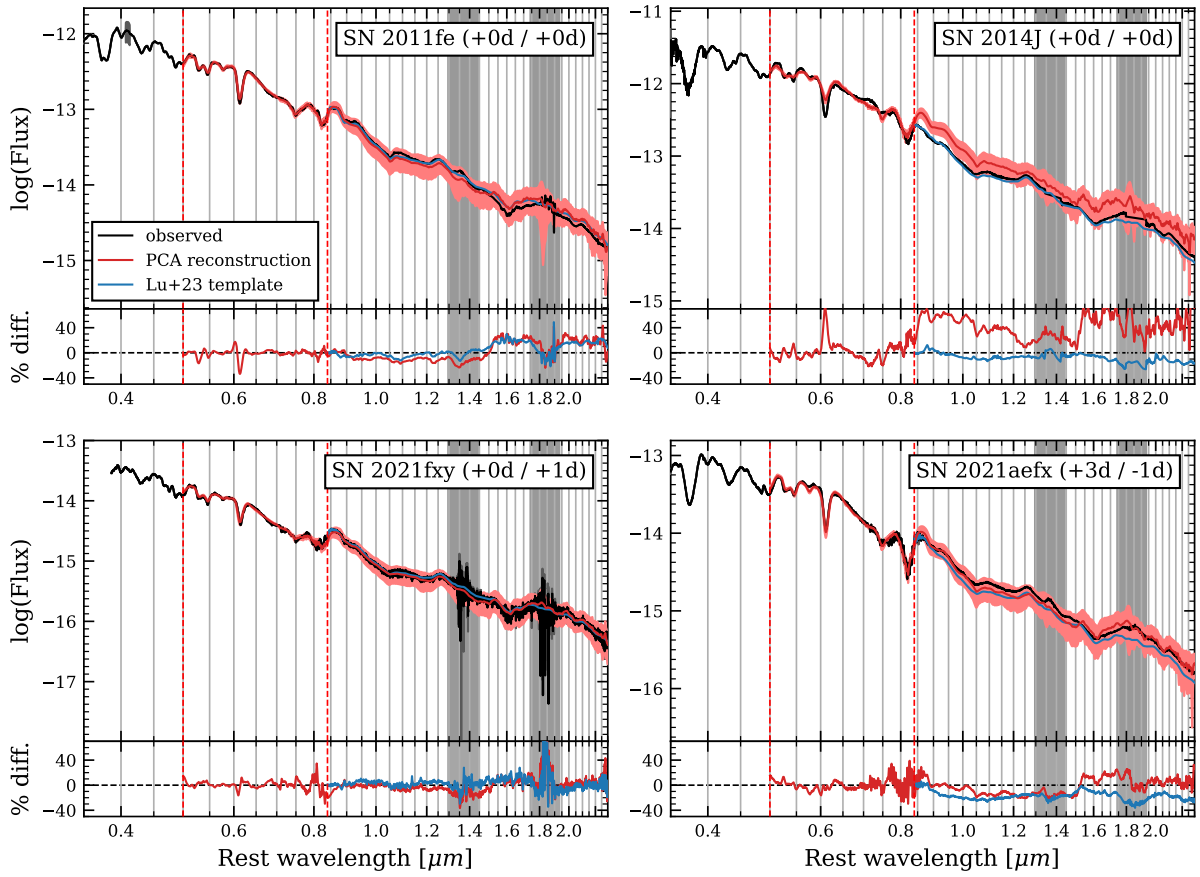


Figure 5.10: Comparison of near-maximum-light extrapolations of SNe 2011fe, 2014J, 2021fxy, and 2021aefx with Lu et al. (2023) spectral templates given their respective values of $s_{\rm BV}$ and phase.

In Figure 5.10, the template approximation of SNe 2011fe, 2014J, 2021fxy, and 2021aefx are

overlaid on top of the extrapolations shown here in earlier sections. The templates were given the same phase as well as values of s_{BV} for each SN Ia: $s_{\mathrm{BV}} = 0.95$ for SN 2011fe (Ashall et al., 2019), $s_{\mathrm{BV}} = 0.99$ for SN 2021fxy (DerKacy et al., 2023), and $s_{\mathrm{BV}} = 1.05$ for SN 2021aefx (Ni et al., 2023). We also use $s_{\mathrm{BV}} = 1.00$ for SN 2014J by converting its measured value of $\Delta m_{15}(B) = 0.98$ (Li et al., 2019) to s_{BV} using the relation determined by Burns et al. (2018). The template predictions were then scaled to the optical part of the observations in the same way as the NIR observations.

From Figure 5.10, for SNe 2011fe, 2021fxy, and 2021aefx, the extrapolation and template methods produce similar results overall. This is expected as models for both the extrapolation method and the spectral templates were built using subsets of the same FIRE NIR spectra. In the case of SN 2011fe, the template seems to be more true to the observation, however after 1.5 μm both methods deviate significantly from the observation. As this effect is prominently seen in both the extrapolation and the template, this may be the result of the inhomogeneity between the CSP NIR spectra used here and the other observations, rather than the fault of the prediction method itself. Both methods perform well for SN 2021fxy, consistently fitting the observed spectrum in each wavelength region. The fits for SN 2021aefx are inadequate for both methods, but interestingly redward of 1.5 μm the extrapolation begins to overestimate the flux, leading to a near-zero difference in integrated flux, whereas the template continues to underestimate the flux. The spectroscopic behavior exhibited by SN 2021aefx and the inability of both the extrapolation and template to reconstruct the observed spectrum may be due to the low GMM probability for SN 2021aefx belonging to the CN Branch group. The probability that it lies within the CN group is only 57.6%, suggesting that it exhibits some Si II and other spectroscopic behavior that is atypical of CNs.

Finally, it is clear that the template prediction is a drastically better fit to the observation of SN 2014J than that given by the extrapolation, even if the extrapolation were to be corrected by a constant factor. This is likely due to the training sample used by Lu et al. (2023) being more representative of SNe Ia similar to SN 2014J and even BLs, as there were far fewer constraints on data selection. It is surprising that the template produces such a good fit to this spectrum, given that SN 2014J seems also atypical for CN SNe Ia as it has such a high value of $v_{\mathrm{Si\,II}}$. Although these four SNe Ia all exhibit similar values of s_{BV} with $0.95 < s_{\mathrm{BV}} < 1.05$, they each display fairly unique spectroscopic behavior. Because of this, it is also unexpected that the template method seems to perform better overall compared to the extrapolation method, which makes predictions using spectroscopic information instead of the single s_{BV} parameter. The likely explanation for this is the much more limited samples sizes used for the extrapolation method, and therefore to improve this method to be consistently on par with spectral templates, more concurrent spectroscopic observations in both the optical and NIR are required.

Because the template method and this extrapolation method both make use of PCA to generate

models, a similar sample size of around 100 SNe Ia (rather than the maximum of 34 used here) would likely be sufficient in capturing the spectroscopic information needed in both the optical and NIR to make models that are consistently comparable to the spectral template method. It may be the case that training these models on a comparable sample size of spectra allows the advantages of the extrapolation method to produce better results in some cases. Despite such a drastically smaller sample size, the extrapolation method does well in comparison to the spectral template results.

5.5.3. Comparison with Templates – Photometry

Similar to Section 5.5.2, we also provide in this section a comparison between the extrapolation and the spectral templates of Lu et al. (2023) in the context of photometry. In particular, we make this comparison by comparing bolometric light-curves obtained from `SNooPy`, which has been modified slightly to reflect the procedure discussed below. This comparison is a more realistic and applicable approach to comparing integrated flux as a function of time.

For the same four SNe Ia as in Section 5.5.2 — SNe 2011fe, 2014J, 2021fxy, and 2021aefx — we calculate the bolometric light-curves of each SN Ia in the following way. First, we gather BVR photometry of SN 2011fe (Richmond & Smith, 2012), BVR photometry of SN 2014J (Foley et al., 2014), $uBVgri$ photometry of SN 2021fxy (DerKacy et al., 2023), and $BVgri$ photometry of SN 2021aefx (Hosseinzadeh et al., 2022). This photometry allows for a measurement of $s_{\rm BV}$ using the `EBV_model2` model in `SNooPy` which in turn is able to produce an approximate SED from the spectral templates of Hsiao et al. (2007), which we will henceforth refer to as the Hsiao templates, that are built into `SNooPy`. From here, three different methods are used to determine bolometric light-curves for each SN Ia: the default `SNooPy` method, which integrates the Hsiao templates to calculate the bolometric luminosity (L_{Bol}); the extrapolation method, which does the same with the extrapolation procedure discussed in this work; and the NIR template method, which again does the same, but using the spectral template of Lu et al. (2023). For clarification, we emphasize that L_{Bol} is calculated by integrating an estimated SED as opposed to using the aforementioned photometric observations of the four SNe Ia with differing bands available. The listed photometry is primarily used to determine $s_{\rm BV}$, which parameterizes the spectral templates, which are then integrated to achieve L_{Bol}.

For the extrapolation and NIR template methods, the Hsiao template is used in the optical regime until the red-most wavelength in the effective range of the red-most band of the photometry given. For example, this is about 9000 Å when BVR photometry is given. Redward of this wavelength, either the extrapolation or NIR template method replaces the Hsiao template by merging the SEDs between 8750–9250 Å. For the extrapolation method, the optical part of the Hsiao template is treated as an observation and projected onto the relevant eigenvectors so that a prediction can be made. This is necessary for testing purposes as an observed spectrum of each supernova is not

available for each epoch that L_{Bol} is calculated. After the full optical+NIR SED is approximated, it follows the same process in `SNooPy` by correcting for the observed photometry provided and then computing the bolometric luminosity.

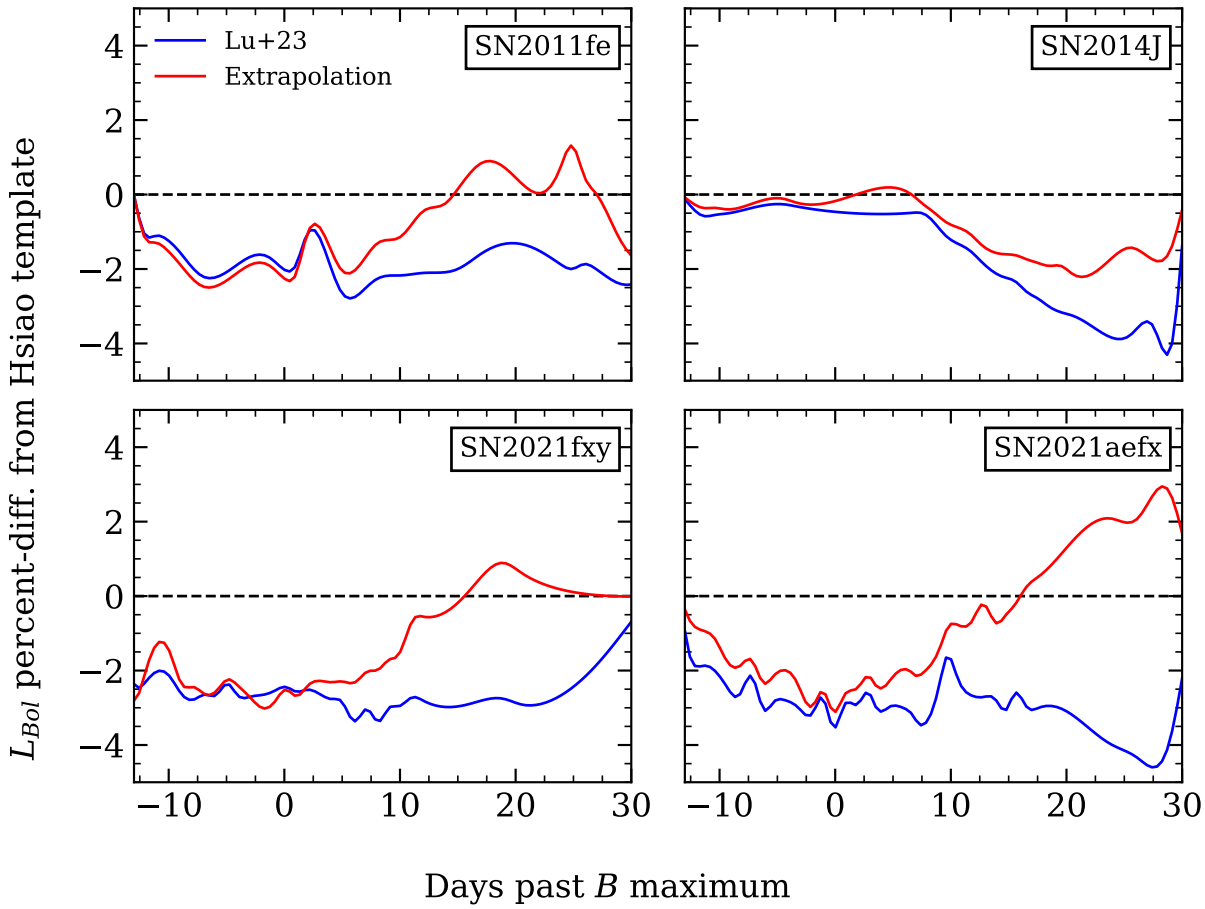


Figure 5.11: Percent-difference of calculated L_{Bol} light-curves from using the Hsiao et al. (2007) template for the method presented in this work (Extrapolation) and for the Lu et al. (2023) (NIR) template. Note that the Hsiao et al. (2007) template does not yield light-curves taken as truth, but rather it is used as a reference to compare the two other methods.

The three bolometric light-curves for each SN Ia were calculated between -16 and $+30$ days relative to B-maximum, and, for the purpose of comparison, the percent-difference of L_{Bol} given the extrapolation and NIR template methods from that of the Hsiao templates is shown in Figure 5.11. Note that the Hsiao templates and their resulting light-curves are not taken as truth, and this difference is displayed to give them a shared point of reference. From Figure 5.11 we see that in L_{Bol} space, the two SED estimation methods yield very similar light-curves between around -16 to $+10$ days until they begin to deviate consistently for each SN Ia. After $+10$ days the extrapolation method yields around $+2\%$ or higher bolometric luminosity than the NIR templates when compared

to the Hsiao templates. This could be due to the lower sample sizes available at later times for the extrapolation method (see Figure 5.2).

Finally, it is worth noting that, although the extrapolated spectrum of SN 2014J exhibits sizable offset from the NIR template (seen in Figure 5.10), the bolometric light-curves show very small differences between the SED estimates around maximum light. This is likely caused by differences between the spectrum used to generate the extrapolation shown in Figure 5.10 and the optical part of the Hsiao template used to generate the extrapolations that yield light-curves from Figure 5.11. In addition, the merging process occurs in different wavelength regions. Therefore, there are no real inconsistencies in what is seen between Figures 5.10 and 5.11. Ultimately, we find that the two methods yield very similar results in both maximum-light spectroscopy and bolometric luminosity surrounding maximum light. Most differences that occur between the two methods are at times later than $+10$ days where sample sizes are smaller, leading to a less representative sample of the population of SNe Ia with which to train models.

5.6. Conclusion

In this study, we make use of optical spectra alongside NIR FIRE spectra from the CSP I+II data set to study correlations between optical and NIR spectroscopy in an attempt to model optical-NIR behavior using PCA. This differs from spectral template methods as this method is not intrinsically parameterized by s_{BV} or other stretch-like quantities, as s_{BV} does not capture the full diversity of SNe Ia in the NIR regime. It is shown here that using patterns exhibited by SN Ia spectra to extrapolate these spectra into the NIR up to 2.3 μm results in predictions that are in accordance with observations. Our results show absolute differences typically of 3% or less in integrated flux from observations, although this may vary greatly depending on the supernova, as is evident from the SN 2014J trial. In cases such as this one, we find that there are large offsets between extrapolated flux and observed flux. These offsets may be due to differences in the flux calibrations between optical and NIR spectra that are unaccounted for in this study. This problem may be addressed by correcting both training and testing spectra by matching the colors of the synthetic photometry and observed photometry at concurrent phases. This effect, as well as large uncertainties shown for these extrapolations, clearly illustrate the need for more observations to train these data-informed models. Specifically, the importance of *concurrent* optical and NIR spectroscopy cannot be overstated for the study of SNe Ia at early phases due to the variation in their spectra at these times.

We also compare this extrapolation method to the NIR templates of Lu et al. (2023) and find that, although the two yield similar results for practical purposes, the NIR templates still have an edge on extrapolation. NIR templates give SED approximations that were slightly closer to observation in most cases. In one case shown here, the study of SN 2014J, the NIR template was much closer to the observation. It is also seen that, when applied to bolometric light-curves, these two methods

produce consistently similar light-curves between around -16 to $+10$ days relative to maximum light. We speculate that, with additional data added to the training sample, this method would indeed be capable of fitting NIR spectroscopic observations of SNe Ia that contain fundamental differences that are independent of light-curve shape. This is because we have shown that projections onto the PCs determined here appear to have some relationship with $s_{\rm BV}$ and spectroscopic quantities such as Si II pEWs and Si II $\lambda6355$ velocity. We surmise that, if future observations allow for a comparable sample size as that used by Lu et al. (2023) ($N \sim 100$), the flexibility of this non-parametric extrapolation method may allow for improvements compared to the spectral template in some cases. It offers a solution to the fact that NIR light-curves, especially around the secondary maximum, cannot be parameterized by $s_{\rm BV}$; with more data, this can be tested more completely. This method is limited by the data available, but it is shown here to have potential in predicting the NIR behavior of SNe Ia.

This extrapolation method along with precalculated time-dependent eigenvectors is openly maintained in the software package named `SNEx` (Burrow, 2024). This program is described further in Section A.3. There are many variables in the work described here, many of which were assumed in this paper for simplicity and not studied in detail. Therefore, some parameters may yield better results when changed for certain SNe Ia, which should be the subject of future work when a more extensive training data sample is established.

CHAPTER 6

Modeling Uncertainties of NIR Extrapolation

6.1. Introduction

Section 5.3.4 reviews the method used by Burrow et al. (2024) to calculate wavelength-dependent uncertainties in extrapolations shown in figures from Sections 5.4 and 5.5. These uncertainties were given by determining residuals between the NIR part of the training spectra and their respective NIR prediction by the model, and then calculating the standard deviation of these residuals at every wavelength point, thus treating these residuals as white noise. However, in doing so, this method does not capture any wavelength-correlated noise or uncertainty in these residuals. It is clear that spectra will always contain at least some correlated uncertainty, because features at different wavelength points are correlated by the underlying atomic physics. If correlated noise is not captured, then the given uncertainties could be underestimated, especially when the uncertainty in any integrated quantity (e.g. bolometric flux, K-corrections, etc.) is desired.

A more robust solution to estimating the uncertainty of these extrapolations is to instead model the residuals with a component describing wavelength-dependent correlation. In this chapter, a method using Gaussian processes is suggested. This method involves creating distributions of logistic function parameters and Gaussian process hyperparameters as sub-models, which is reminiscent of hierarchical Bayesian modeling. Samples are drawn from these parameter distributions that will lead to a measurement of uncertainty that accounts for correlated noise.

6.2. The Logistic Model

The top panel of Figure 6.1 displays the residuals of the training sample of 33 SNe Ia that produced the maximum-light eigenvectors shown in Figure 5.3. These residuals come from the difference in the observed flux of each training sample and the prediction from 0.50–2.30 μm of that sample. The residuals shown are in the standardized frame; recall that standardized flux in this context means that the mean of the training spectra has been subtracted from the flux, and subsequently the resulting flux is scaled by the standard deviation of the training spectra. A prediction is given for the entire 0.50–2.30 μm wavelength range; however, an extrapolation only occurs redward of 0.84 μm, which is indicated by the vertical dashed lines in Figure 6.1 where the model begins to deviate much more drastically from the observed flux. For reference, the standard deviation of these residuals is what is added in quadrature to the second type of uncertainty discussed in Section 5.3.4 to achieve what is displayed as uncertainty in the observed flux everywhere in

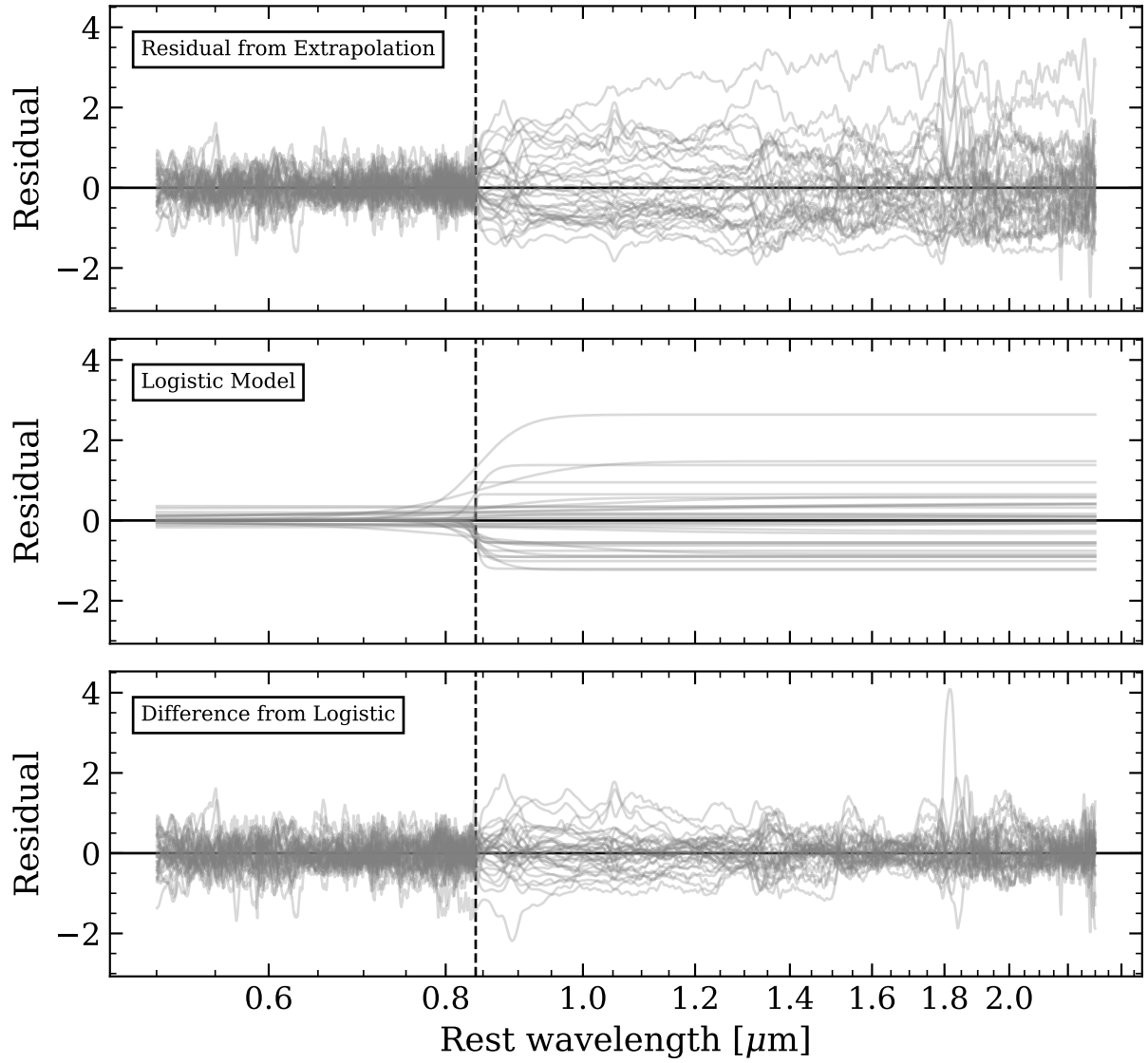


Figure 6.1: Top: Residuals between 33 training spectra of maximum-light PCs from Figure 5.3 and their respective predicted spectra; Middle: Set of logistic functions fit to each residual shown in the top panel; Bottom: The difference between the residuals in the top panel and their logistic model in the middle panel. Vertical dashed lines indicate 0.84 μm where true extrapolation begins.

Sections 5.4 and 5.5.

Looking at these residuals in the top panel of Figure 6.1, there is an apparent degree of asymmetry to these residuals. In other words, there are a small number of outliers which seem to only exist for positive residuals. This appears to be possibly related to the near-constant offsets seen from SN 2022hrs in Figure 5.5 or even from SN 2014J in Figure 5.9. In addition, each individual residual seems to stay either strictly positive or negative redward of 0.84 μm, aside from those that are consistently near zero. Here we choose to model both of these effects using the sigmoidal logistic function centered around $\lambda_0 = 0.84$ μm described by

$$R(\lambda|L,k) = \frac{L}{1+e^{-k(\lambda-\lambda_0)}}, \tag{6.1}$$

where R is the plotted residual, L is a horizontal asymptote as λ approaches infinity, and k is the logistic growth rate which describes the steepness of the curve around λ_0. Note that we choose to include the logistic function into our overall model because Gaussian processes will not account for the asymmetries seen in the residual with the kernel that is used (see Section 6.3). Each residual was fit with this logistic function using a least-squares fit to determine the pair of parameters L and k, and these models are shown in the middle panel of Figure 6.1. These models seem to capture the asymmetry and outliers of the sample residuals well.

The differences between the residuals in the top panel of Figure 6.1 and the models in the middle panel are shown in the bottom panel to illustrate the amount of information captured by these logistic models. Interestingly, aside from the large regions of telluric effects and the area just redward of 0.84 μm, the scale of these uncertainties have almost matched that of the optical residuals blueward of 0.84 μm. There is also much wavelength dependence in these logistic-corrected residuals shown in the bottom panel. This wavelength dependence is the leftover residual that is desirable to be modeled with a Gaussian process, which is the subject of the next section.

6.3. Fitting with Gaussian Processes

After the set of residuals is modeled by logistic functions, the leftover residual — shown in the bottom panel of Figure 6.1 for the maximum-light example we discuss in this chapter — is to then be modeled using Gaussian processes (see Appendix B for a discussion on Gaussian processes). For our residuals, we make use of the Matérn 3/2 kernel to populate the covariance matrix which describes the correlation between different wavelength points. This kernel contains two hyperparameters: the lengthscale (ℓ), which describes the typical scale of wavelength dependence that leads to some characteristic change in the dependent variable, and the variance (σ^2), which is an amplitude-like quantity for the average distance of the dependent variable from its mean.

Here we assume a mean function of zero for the Gaussian process, and therefore the model

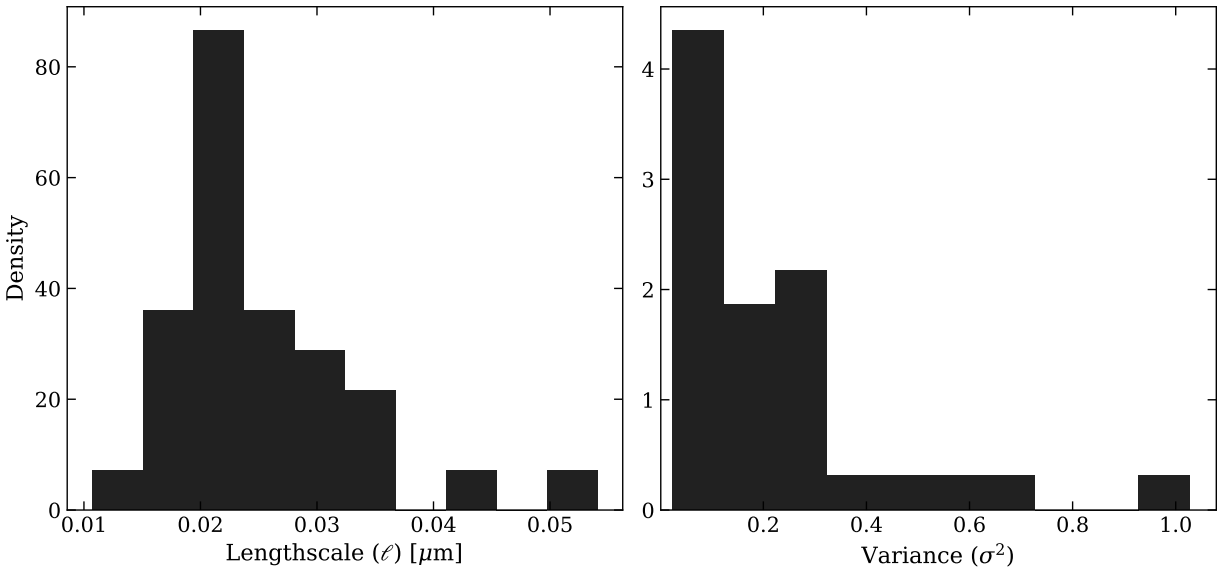


Figure 6.2: Distribution of lengthscale (left panel) and variance (right panel) hyperparameters determined by optimization of Gaussian processes fit to the post-logistic residuals shown in the bottom panel of Figure 6.1.

will not be conditioned on data. This conditioning is what allows for the solution of a posterior distribution whose mean can act as a model for a single piece of data. For example, `Spextractor` conditions a Gaussian process on a spectrum, and the mean of that posterior distribution is the smoothed output spectrum which may act as an interpolator. Here, we instead just use the prior distribution of the Gaussian process to describe the normally distributed spread of many data around a mean of zero.

Each post-logistic residual is fit with an individual Gaussian process whose hyperparameters are optimized such that the likelihood of obtaining data from such a model is maximized in a process known as Bayesian optimization. Unfortunately, performing an optimization process for each residual can start to become somewhat computationally expensive; however, this only needs to be performed once per set of residuals — or in other words, once per epoch that a PCA model for extrapolation is created. After this, a distribution of hyperparameters can be determined, which will be used thereafter.

The lengthscale and variance hyperparameters that have been determined from this process for each maximum-light residual are compiled into a histogram shown in Figure 6.2. Looking at the distribution of lengthscales in the left panel of Figure 6.2, the median correlation length is around 0.023 μm, or 230 Å. This aligns well with the typical width of some of the strongest features seen in SN Ia spectra. The variance distribution in the right panel of Figure 6.2 is not as conceptually significant, but it is important to understand the distributions of these hyperparameters, or if a

distribution exists at all. If there is a distribution to both of these best-fit hyperparameters, then a model of these residuals cannot be appropriately represented with only a single value of either the lengthscale or variance when using Gaussian processes with the Matérn 3/2 kernel we have chosen.

6.4. Sampling

6.4.1. Parameter Distributions

Because there is a distribution to the best-fit Gaussian process hyperparameters for the entire sample of residuals, we choose to model the data probabilistically. Note that there is also a distribution in the logistic parameters that have not yet been explicitly shown; this is known because the logistic models in the middle panel of Figure 6.1 have varying horizontal asymptotes as well as growth rates around 0.84 μm. The probabilistic model used should therefore be a distribution in the space of these four parameters: L, k, ℓ, and σ^2. However, we assume that L and k are independent of ℓ and σ^2, as the logistic function is accounted for before the Gaussian process takes place. In addition, it is shown in Figure 6.2 that ℓ and σ^2 have individual distributions (shown as histograms); however, there could be a functional relationship between the lengthscale and variance determined for each sample, and the same could be inferred for L and k. To account for this, these distributions will be two-dimensional. Therefore, there will be two desired 2-D distributions — one for L and k, and another for ℓ and σ^2.

One simple method of inferring a distribution from a collection of points in any number of dimensions is by using kernel density estimation (KDE) (Rosenblatt, 1956; Parzen, 1962). A KDE places a kernel distribution at every point in a data set and averages them with optional weights to estimate the overall distribution of the data. The individual distributions are parameterized by a bandwidth, which acts analogously to the bin size of discretized histograms. In this case, for both distributions, we use a Gaussian kernel with a bandwidth given by the so-called Scott's Rule (Scott, 1992), which gives the bandwidth as $N^{-1/(m+4)}$, where N is the number of points in the sample, and m is the dimensionality ($m = 2$ in this case).

Figure 6.3 plots the parameters L and k that are found for the models shown in Figure 6.1. These values are displayed as black points, and the surrounding contours are those of the distribution given by the KDE fit using those points. Aside from a few outlying points, there appears to be somewhat of a function relationship between the two parameters; the growth rate k is found to be larger for residuals with mostly negative offsets (L), meaning the model is closer to a step function at $\lambda_0 =$ 0.84 μm. Otherwise, at near-zero and positive offsets, the growth rate becomes smaller or even slightly negative. The models with these low k values are effectively constant functions at their value of L with little to no growth.

This logistic-parameter space is approximated with a distribution given by the KDE shown as contours in Figure 6.3. It is clear from the contours that this distribution captures outliers such as

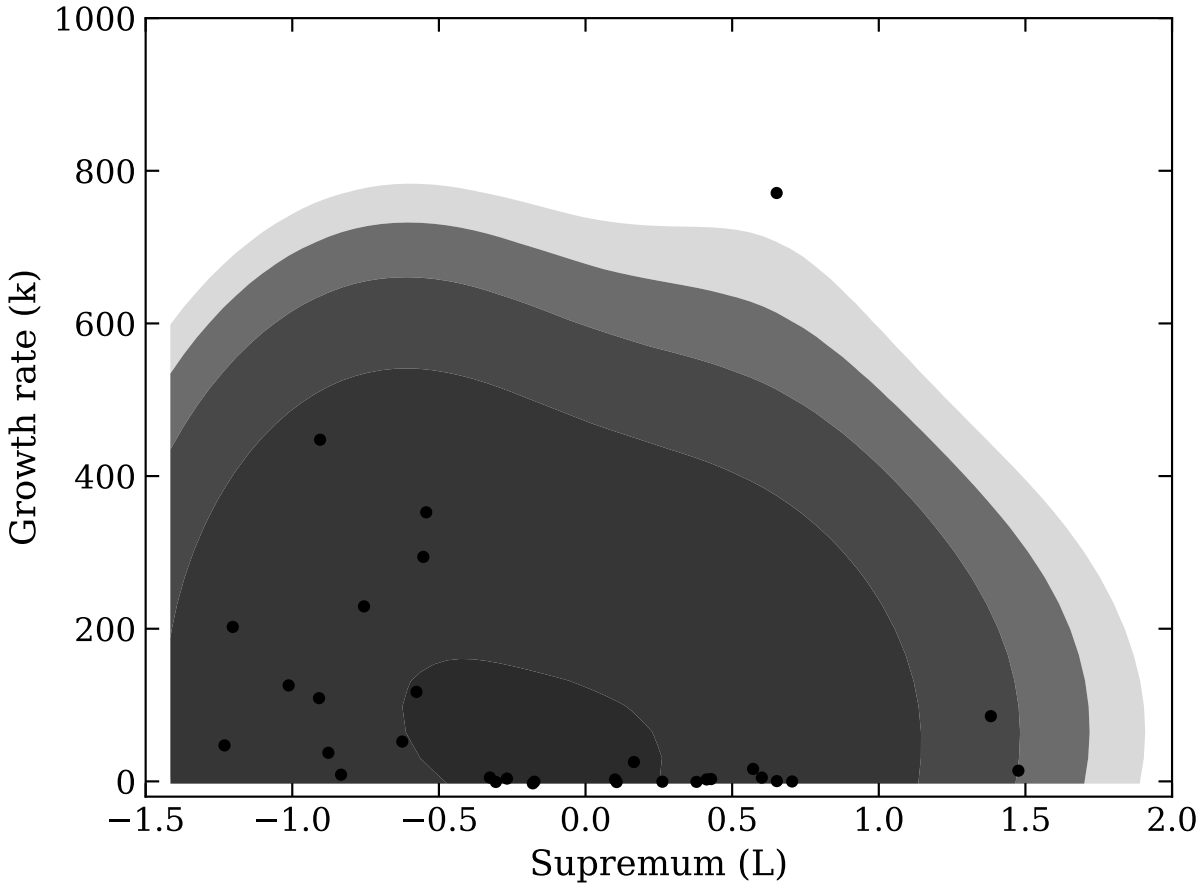


Figure 6.3: Points indicate the set of L and k parameters of the logistic function (Equation 6.1) found for models of the residuals shown in the middle panel of Figure 6.1. Contours map out the distribution of a Gaussian KDE applied to these parameters in two dimensions.

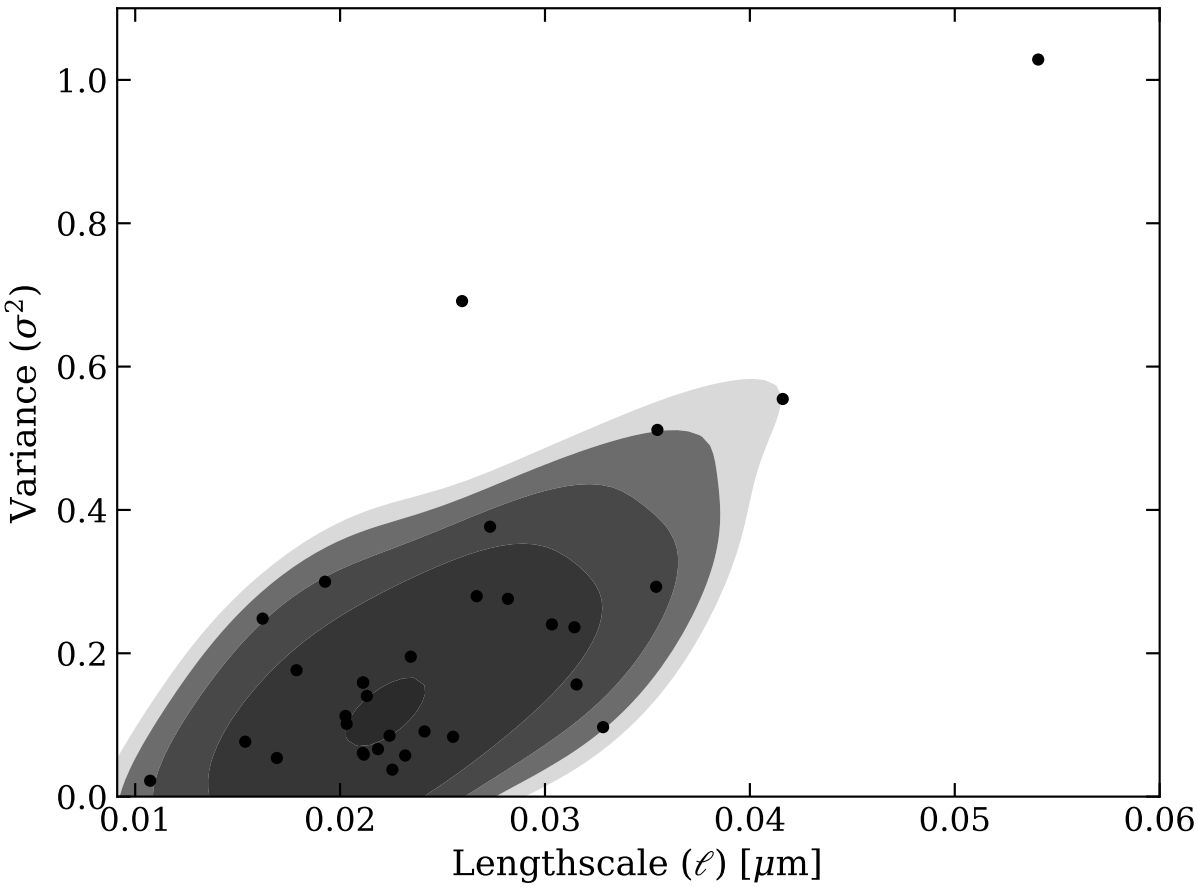


Figure 6.4: The Gaussian process hyperparameters ℓ and σ^2 along with their 2-D distribution given by a KDE, similar to Figure 6.3 and with the same formatting.

that near $k = 800$ and those around $1.4 < L < 1.5$. However, by normalizing the set of Gaussian distributions made in the KDE, the relatively low frequency of outliers appearing is accounted for when sampling parameters from this distribution. By using a probabilistic model such as this, when sampling from this distribution, the majority of samples will lie closer to the more populous clusters of points in the figure. Therefore, the logistic part of the realizations that come from sampling this parameter distribution will likely look similar to the logistic models shown in Figure 6.1, which is desired.

Similarly, the best-fit hyperparameters ℓ and σ^2 of the post-logistic Gaussian processes are shown in Figure 6.4. The functional relationship between the two hyperparameters seems much more one-to-one in this case, with longer lengthscales being found for larger average deviations from the mean. The contours of the distribution given by the KDE on these points are also shown in Figure 6.4. This distribution seems much more compact, and therefore it is not as influenced by the two outliers with $\sigma^2 > 0.6$. Thus, when sampling this hyperparameter distribution for different

realizations of a Gaussian process, the hyperparameters of the Gaussian processes from which realizations are taken will be fairly close to one another.

Unfortunately, a Gaussian KDE does not account for physical or mathematical limitations to the data. For example, the distribution shown in Figure 6.4 extends to the region of the parameter space where $\sigma^2 < 0$. However, σ^2 can never be negative, else the covariance matrix of the Gaussian process will not be positive semi-definite. We also choose to limit the sampled parameters from straying too far outside the range of parameters that are initially found to fit the residuals. To address both of these concerns, any samples from these two distributions are resampled until the following conditions are met: (a) $-1.4 < L < 3.0$, (b) $k > -3$, (c) $0.009 < \ell < 0.06$, and (d) $\sigma^2 > 0$. This resampling effectively renormalizes the distribution, where the distributions are those shown explicitly by the cut-off contours drawn in Figures 6.3 and 6.4.

6.4.2. Resulting Realizations

Using the distributions shown in Figures 6.3 and 6.4, a sample of each can be taken independently to get a set of parameters L, k, ℓ, and σ^2. Using those values of L and k with the logistic function and adding it together with a sample of a Gaussian process prior distribution given the value of ℓ and σ^2, a full realization of a modeled residual can be created. If this process is performed many times, the full parameter space will be fully mapped out, and the realizations will show a spread of possible residuals from the model.

For example, in the bottom panel of Figure 6.5, 100 realizations of the residuals were created from our model. These realizations are compared with the originally determined residuals, shown in the top panel of Figure 6.5. The red line in each panel is the mean of all residuals shown at each wavelength point. Note that this mean will always be nearly zero for the real residuals in the top panel because these are residuals given by projecting the training spectra onto eigenvectors that were trained using those same spectra. The eigenvectors were mathematically determined such that these residuals are effectively minimized at each wavelength point. The mean of the realizations of the model (in the bottom panel) is slightly offset from zero, showing a slight tendency of the residuals to have higher positive offsets, which is due to the effect of the highest outlier in the top panel of Figure 6.5 on the distribution of the logistic parameters. This verifies that this model does capture the asymmetry of the residuals we originally find, which was one of the desired qualities of this model.

Comparing the realizations from the model with the real residuals in the top panel of Figure 6.5, the realizations seem to capture a similar spread in the residuals. The largest deviation from the real residuals is that there is no information on any individual spectroscopic features. It is clear from the top panel that the residual can be much higher around certain features such as Mg II around 1.05 μm. Feature-dependent residuals are not captured by our model given only by a logistic function and

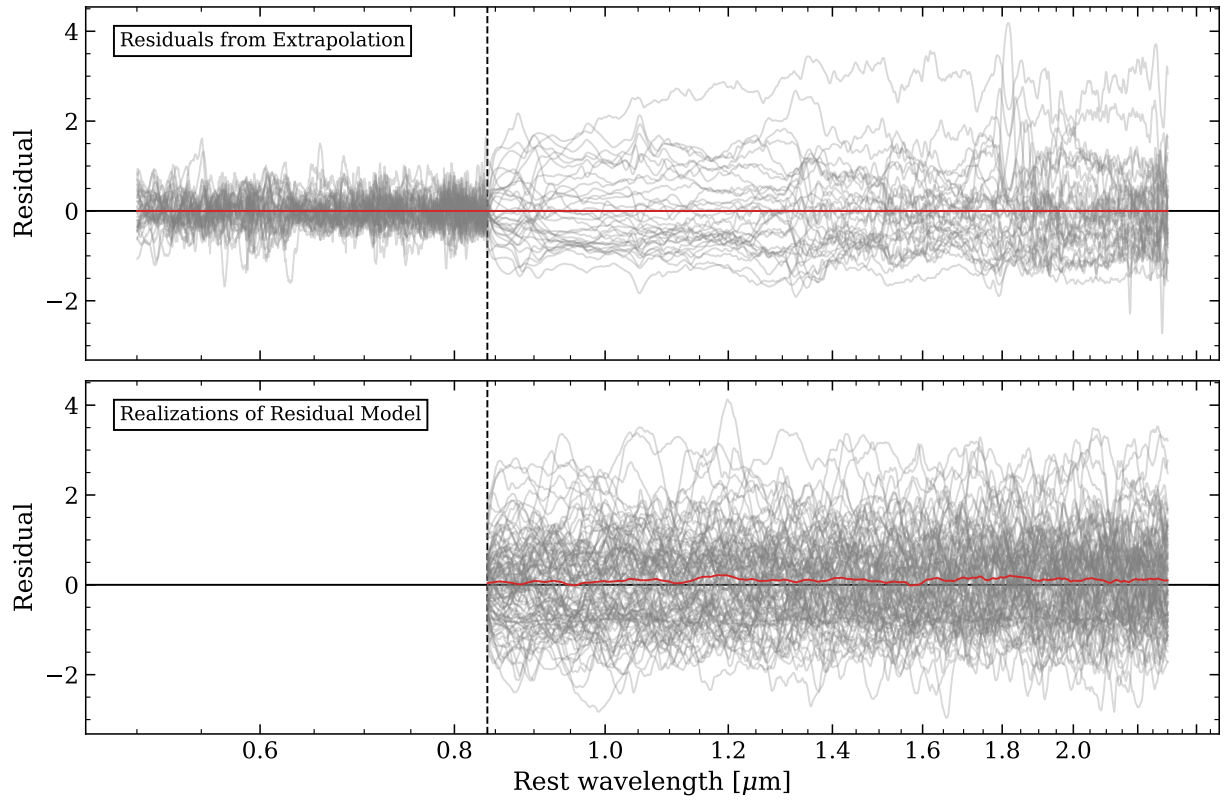


Figure 6.5: Top: The same original residuals shown in the top panel of Figure 6.1; Bottom: Realizations of the complete model established with a logistic function and Gaussian processes obtained by sampling parameters from distributions shown in Figures 6.3 and 6.4. Red lines show the mean of all samples in the panel at each wavelength point.

a Gaussian process with a single kernel. It could be possible to capture some feature-dependent residuals by adding, for example, a narrow Gaussian function to our overall model at each feature. However, the benefit of this would likely be small, because the residual around these features are not actually described by a Gaussian or any other simple function.

6.5. Conclusion

Here we have modeled an example set of 33 NIR residuals that come from comparing the spectra used to train the maximum-light eigenvectors shown in Figure 5.3 with their extrapolated spectra, given their optical flux. This model was created by assuming an offset starting at $\lambda_0 = 0.84$ μm that is described with a logistic function. After correcting for this logistic offset, the remainder of the residuals may then be described using Gaussian processes with a single Matérn 3/2 kernel. We find that when these residuals are fit to this model, the logistic parameters L and k and the Gaussian process hyperparameters ℓ and σ^2 form distributions, which we assume may be represented by a Gaussian KDE.

After determining distributions of L, k and ℓ, σ^2, one can sample from these distributions to acquire parameters that can form any number of different realizations of the 33 training data residuals. These realizations account for any wavelength-dependent correlation of the residuals on the lengthscale ℓ used; however, they do not account for differences in individual feature residuals because of the simple model used. The resulting realizations of residuals can then be added back to the extrapolated spectrum (while the extrapolated spectrum is still standardized). This, in turn, may be used for finding wavelength-dependent uncertainty or any desired photometric uncertainty for spectroscopy and photometry around maximum light. This process would then be repeated if other epochs are desired, such as for use in calculating uncertainty of light-curves. An analysis of this entire process should be the subject of future work, along with comparing the resulting photometric uncertainties obtained through this method and those of the white noise case (described in Section 5.3.4) and those of using templates by Lu et al. (2023), as those spectral templates provide uncertainties as well. This uncertainty model has not yet been implemented into `SNEx`; however, it will begin production with high priority to initiate a formal analysis.

CHAPTER 7

Correlations of PC Projections

This chapter is taken from Burrow et al. (2024), published in *The Astrophysical Journal*, reformatted with slight modifications to improve the readability, consistency, and sequencing of this dissertation. The content in this chapter covers Sections 5.1–5.2 from Burrow et al. (2024).

7.1. Introduction

In Chapter 3 we find that Si II plays a significant role in partially explaining the diversity of SN Ia spectroscopy, which is to be expected given the well-studied classification system of Branch et al. (2006). It is also shown in Section 5.4.4 that performing an extrapolation process given only Si II information results in similar and sometimes better fits to observations than using a wider range of optical spectra.

Because the NIR spectrum seems well-inferred by projecting only the two Si II features onto a set of maximum-light PCs, one may think that the PCs themselves have some relation to the Branch group classification scheme. In this section, we look at how these PCs relate to the Branch groups to explore the possibility that each individual Branch group has unique spectroscopic features that exhibit a relationship with features in the NIR. In addition, we explore further relationships these PCs have with s_{BV}.

7.2. PC Relation to Branch Groups

We project 101 near-maximum-light spectra (within ± 5 days of maximum light) available from the optical CSP I+II data set onto the maximum-light eigenvectors shown in Figure 5.3. However, we only project the region between 0.55–0.66 μm in order to account for the extent of Si II $\lambda 6355$ and Si II $\lambda 5972$ on the eigenvectors. These values — denoted as PC_1, PC_2, etc., corresponding to the projection onto the respective eigenvectors — are then plotted against each other in the correlation matrix shown in Figure 7.1. The pEWs of Si II $\lambda 6355$ and Si II $\lambda 5972$ as well as Si II $\lambda 6355$ expansion velocity $v_{\mathrm{Si\,II}}$ of each spectrum were calculated using `Spextractor`. By using a 3-D GMM from Section 3.4.2.1, the Branch group that each SN Ia is assigned to lie within has been colored in a similar fashion to Figure 3.5. In addition, SNe Ia exhibiting higher velocities with $v_{\mathrm{Si\,II}} > 12{,}500$ km s^{-1} are symbolized with stars in Figure 7.1. Many of these high-velocity (HV) SNe Ia are classified as CNs or BLs, which is clear from Figure 3.16.

It is immediately apparent that a few subspaces, for example PC_1 vs. PC_3 or PC_2 vs. PC_4, seem to show CN, SS, and CL SNe Ia lying in distinct regions, similar to a classic Branch diagram. There

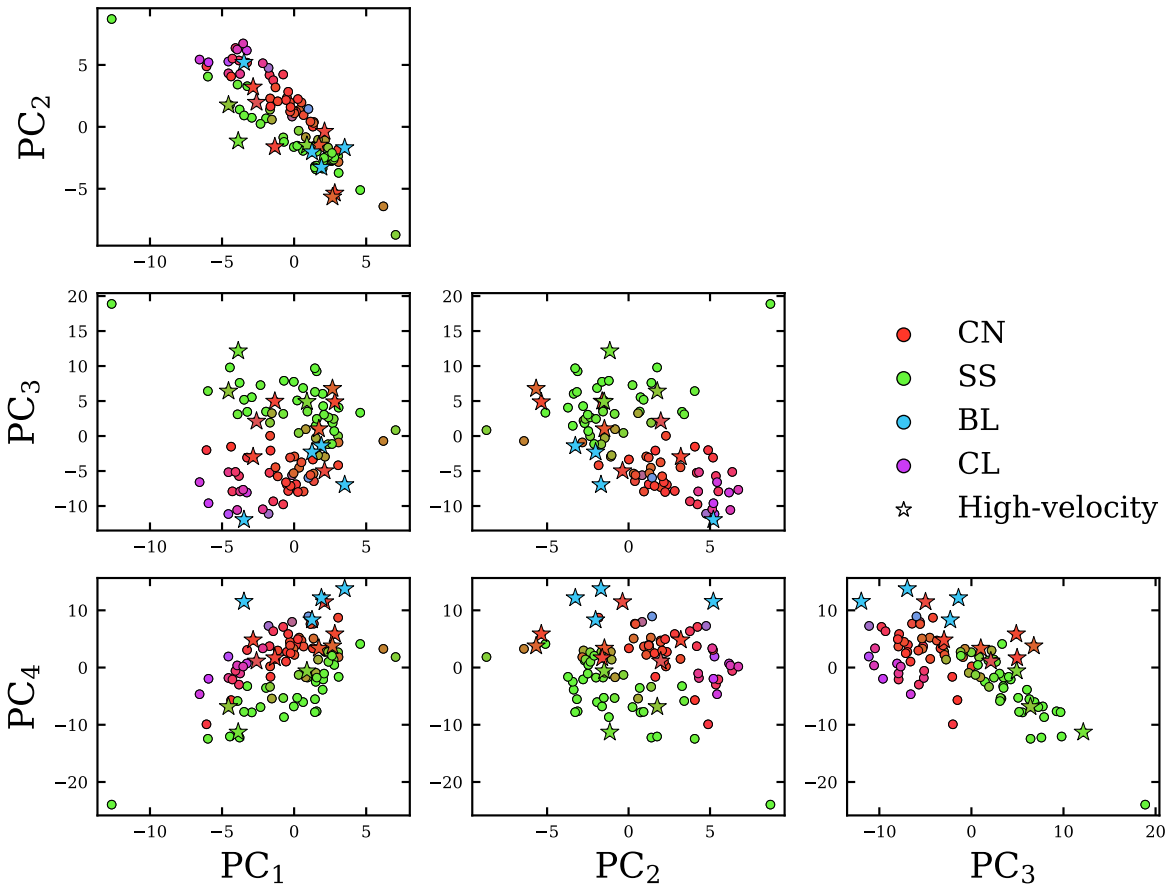


Figure 7.1: The projections of spectra in the optical data set onto the first four maximum-light PCs in Figure 5.3 only in the 0.55 μm $< \lambda <$ 0.66 μm region. Points are colored for the probability of Branch group classification in the same way as in Section 3.4 using a 3-D GMM. Points shown as stars are high-velocity SNe Ia exhibiting $v_{\mathrm{Si\ II}} > 12{,}500$ km s^{-1} at maximum light.

are some exceptions to this, for example in the PC_1 vs. PC_3 panel, several CNs lie in the region dominated by SS SNe Ia. This illustrates that there is not a one-to-one relationship between these projections and the measured pEWs of Si II λ6355 and Si II λ5972, and therefore this is not a recreation or rotation of the classic Branch diagram. This is further explored in the discussion of Table 7.1 later in this subsection. This suggests that there may be some intrinsic differences between these CNs lying within the SS-dominated region in this subspace and the other more concentrated CNs. Many of these CNs within the SS region exhibit higher velocities; however, this likely cannot be the only characteristic that sets these SNe Ia apart from other CNs, because other HV CNs are seen in the concentrated region as well. Finally, note that BL SNe Ia are not as constrained to their own regions in Figure 7.1. This could be due to the unfortunate lack of BL representation in the training sample used to calculate these eigenvectors. Because of this, it is difficult to arrive at

concrete conclusions for BL SNe Ia from this figure.

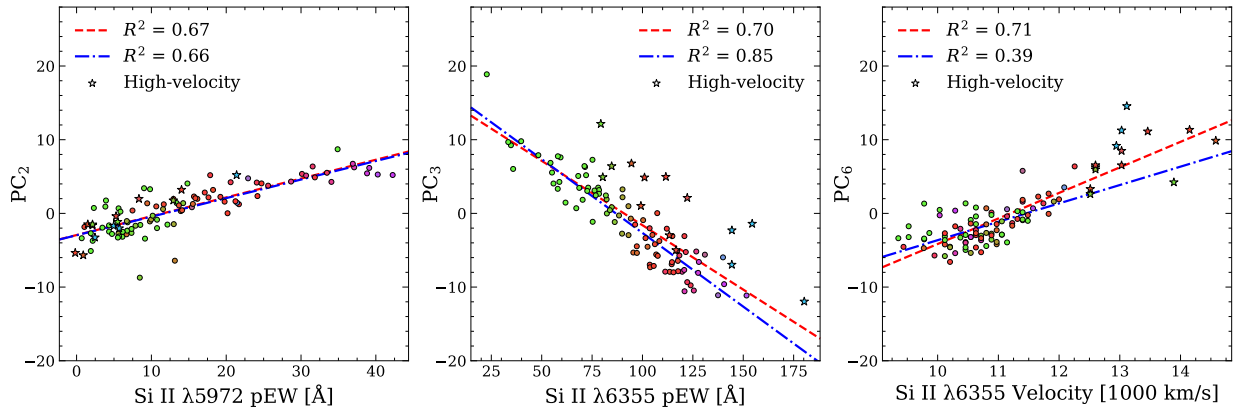


Figure 7.2: PC$_2$, PC$_3$, and PC$_6$ plotted against pEW(Si II λ5972), pEW(Si II λ6355), and $v_{\text{Si II}}$, respectively. Colors here are the same 3-D GMM colors shown in Figure 7.1, and points shown as stars indicate the same HV SNe Ia. Lines here indicate linear regression lines for all the points (in red, dashed) and for all points except HV SNe Ia (in blue, dash-dotted).

Investigating the PC$_i$ subspaces further, Figure 7.2 plots the projections of Si II onto PCs 2, 3, and 6 against pEW(Si II λ5972), pEW(Si II λ6355), and $v_{\text{Si II}}$, respectively. Each point is colored by Branch group classification in the same way as Figure 7.1. Linear regression lines are shown for each panel; the red line shows the regression for all points, and the blue line shows the same for all points except high-velocity SNe Ia. The R^2 coefficient of determination for each regression line is shown in the legend of each panel. A summary of the R^2 values for each projection up to PC$_8$ against all three spectroscopic properties is given in Table 7.1, where R^2 values from fits excluding HV SNe Ia are given in parentheses. PCs 2, 3, and 6 were chosen to be displayed because they exhibit the highest R^2 for correlations with the three properties for both the general and non-HV cases. Although a low value of R^2 does not rule out a non-linear correlation, other values of PC$_i$ are not displayed here because there is no clear visual trend in them, or the spread of the points is too large. This leads to the low values of R^2 for the majority of projections seen in Table 7.1.

The highlight of this analysis is obtained by observing how PC$_3$ relates to the pEW(Si II λ6355) shown in the middle panel of Figure 7.2. There is a significant spread of SNe Ia around the regression line, however when ignoring the clearly distinct HV SNe Ia in this panel, the spread is greatly reduced, leading to a significantly improved R^2, which is indicative of a stronger linear correlation between PC$_3$ and pEW(Si II λ6355). The effect of HV SNe Ia being outliers from the trend line is also not dependent on Branch group classification — i.e., these outlying points do not only correspond to BL SNe Ia. It appears that the measurement of PC$_3$ seems to be accounting for HV behavior in a way that the classic measurement of pEW(Si II λ6355) cannot.

In addition, we find a mild linear correlation between PC$_2$ and pEW(Si II λ5972), which itself

Table 7.1. Summary of R^2 values of all PC_i linear correlations with various spectroscopic properties. Values in parentheses indicate the R^2 value of the linear regression excluding HV SNe Ia.

Projection	R^2 [pEW(Si II λ5972)]	[pEW(Si II λ6355)]	[$v_{\mathrm{Si\,II}}$]
PC_1	0.42 (0.44)	0.02 (0.03)	0.03 (0.06)
PC_2	0.67 (0.66)	0.19 (0.29)	0.03 (0.00)
PC_3	0.34 (0.35)	0.70 (0.85)	0.00 (0.08)
PC_4	0.00 (0.00)	0.49 (0.43)	0.23 (0.35)
PC_5	0.51 (0.51)	0.01 (0.03)	0.42 (0.20)
PC_6	0.04 (0.00)	0.17 (0.06)	0.71 (0.39)
PC_7	0.09 (0.11)	0.61 (0.62)	0.01 (0.05)
PC_8	0.03 (0.03)	0.46 (0.52)	0.02 (0.16)

has been shown to exhibit a connection with s_{BV} (see Figure 3.19). However, we show and discuss in Section 7.3 that more information than PC_2 is needed to extract a estimate of s_{BV} from a spectrum using these projections. The slope estimator of each linear regression shown in Figure 7.2 has a two-sided P-value of much less than 0.001, heavily suggesting that there is at least some dependency of these three PC_i and the respective spectroscopic quantities shown. Visually, though, the wide spread of measured PC_2 seems comparable to the change in the trend of PC_2 as pEW(Si II λ5972) changes and therefore, this linear relationship seems untrustworthy as an estimator. This also seems true for PC_6 versus $v_{\mathrm{Si\,II}}$, as removing HV SNe Ia in this panel heavily diminishes the quality of this linear correlation, which is suggested by the much lower R^2 statistic.

These projections may offer an alternative method of classifying SNe Ia that does not rely on the assumption of a pseudo-continuum that can lead to inconsistent measurements. It would be interesting to perform further testing with these PC_i measurements to decide their suitability as classification criteria that may offer solutions to some known discrepancies in SNe Ia classification. For example, in Figure 3.16 it is shown that several CNs exhibit high velocities that would be expected of BLs, given the relationship between the velocity of the blueshifted absorption part of a P Cygni profile and its equivalent width. Measuring PC_i values instead of pEWs could reclassify these HV CNs to be similar to SNe Ia with broader Si II λ6355 features.

7.3. PC Relation to s_{BV}

Similar to Lu et al. (2023), it is interesting to see how these PCs may be related to light-curve parameterization such as the color-stretch parameter s_{BV}. This relation may play an important role

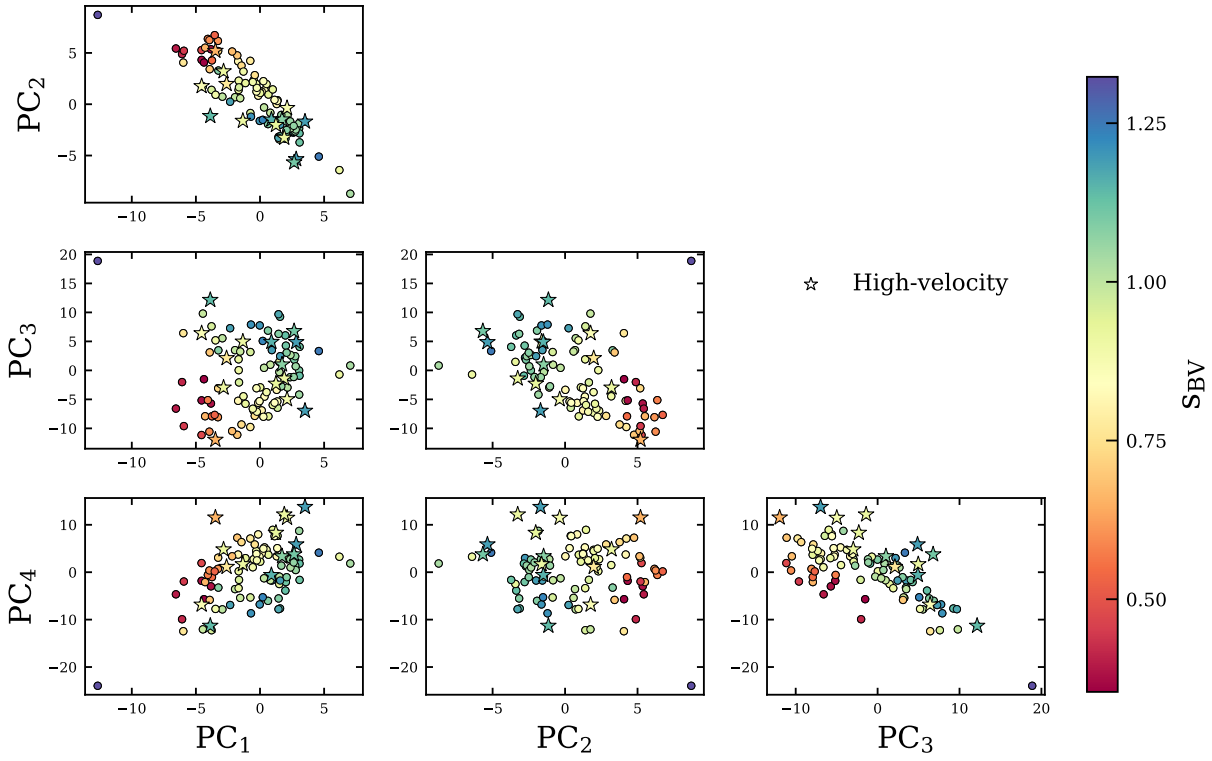


Figure 7.3: The same projections as in Figure 7.1, however now colored for s_{BV} determined by `SNooPy`. It seems visually plausible that s_{BV} may be estimated here as a function of two or more PCs.

in the connection between observed photometry and the underlying SED of SNe Ia. In Figure 7.3, the exact same PC projections as those shown in Figure 7.1 are displayed, now colored for the s_{BV} value of each SN Ia calculated using `SNooPy`. There appears to be a multi-dimensional trend between s_{BV} and many of the PC values. In other words, projecting the Si II region of a spectrum onto the maximum-light eigenvectors produces values that should allow one to predict s_{BV} given only spectroscopic information. For example, just using PC_1 and PC_2, it may be seen that SNe Ia with high values of PC_1 and low values of PC_2 tend to have higher s_{BV}, and vice versa.

Looking at the PC_1 column, it appears that PC_1 does not directly correlate with s_{BV} by itself. This is verified by Figure 7.4, which plots PC_1 against s_{BV}, showing the large amount of spread in the relationship between the two quantities. If one were to use this relationship with a measured value of PC_1, it would not be possible to estimate an accurate s_{BV} without large uncertainty. This heavily suggests that the first PC shown in Figure 5.3 does not describe the same variation of SNe Ia within the training sample that s_{BV} explains.

Examining each subplot in Figure 7.3, the regression of s_{BV} visually appears to define a direction in each subspace, suggesting that s_{BV} may be determined by the projections. This idea of predicting

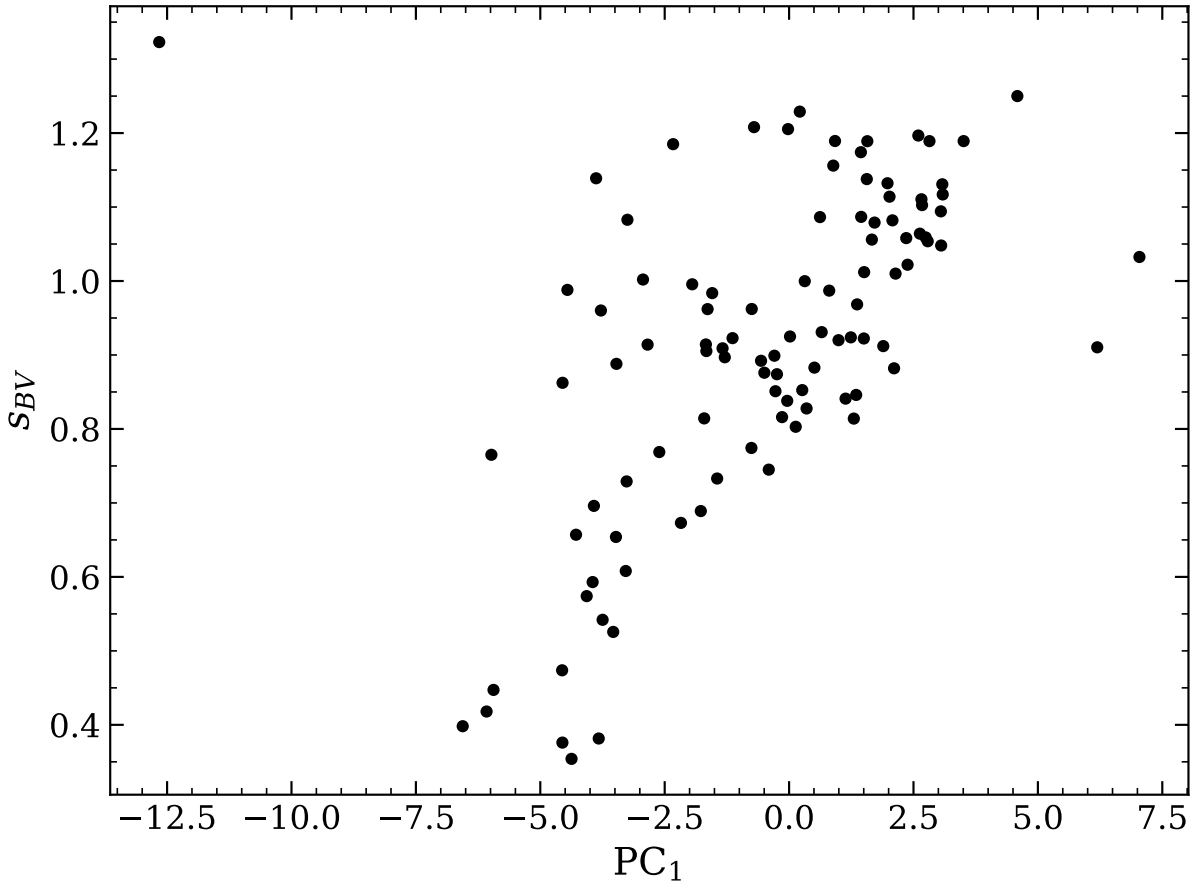


Figure 7.4: A direct view of PC$_1$, from Figure 7.3, versus $s_{\rm BV}$. This shows the lack of correlation between the two, showing that more information than PC$_1$ is needed to estimate $s_{\rm BV}$.

$s_{\rm BV}$ may be put to the test with a simple linear model. This model may be a function of any number of PC values, i.e.,

$$s_{\rm BV} = a_0 + \sum_n^N a_n {\rm PC}_n, \tag{7.1}$$

where a_n are coefficients for each projection PC$_n$ up to the projection PC$_N$. After some testing, we find that the sum of least-squares are reduced as more projections are added to this model (and thus N increases). For this example we extend this model up to $N = 8$, and therefore the linear model will contain nine coefficients to be fit. These coefficients are determined by fitting to the data in Figure 7.3 using least-squares fitting.

To simulate a test of this procedure, we take a random 80% (82 points) of the SNe Ia in Figure 7.3 to fit the coefficients of the model, then use this model to estimate $s_{\rm BV}$ of the same SNe Ia. These values are displayed in Figure 7.5 as black points. The remaining points colored red in Figure 7.5 are the $s_{\rm BV}$ model estimates of the remaining 19 SNe Ia, from which no information was provided

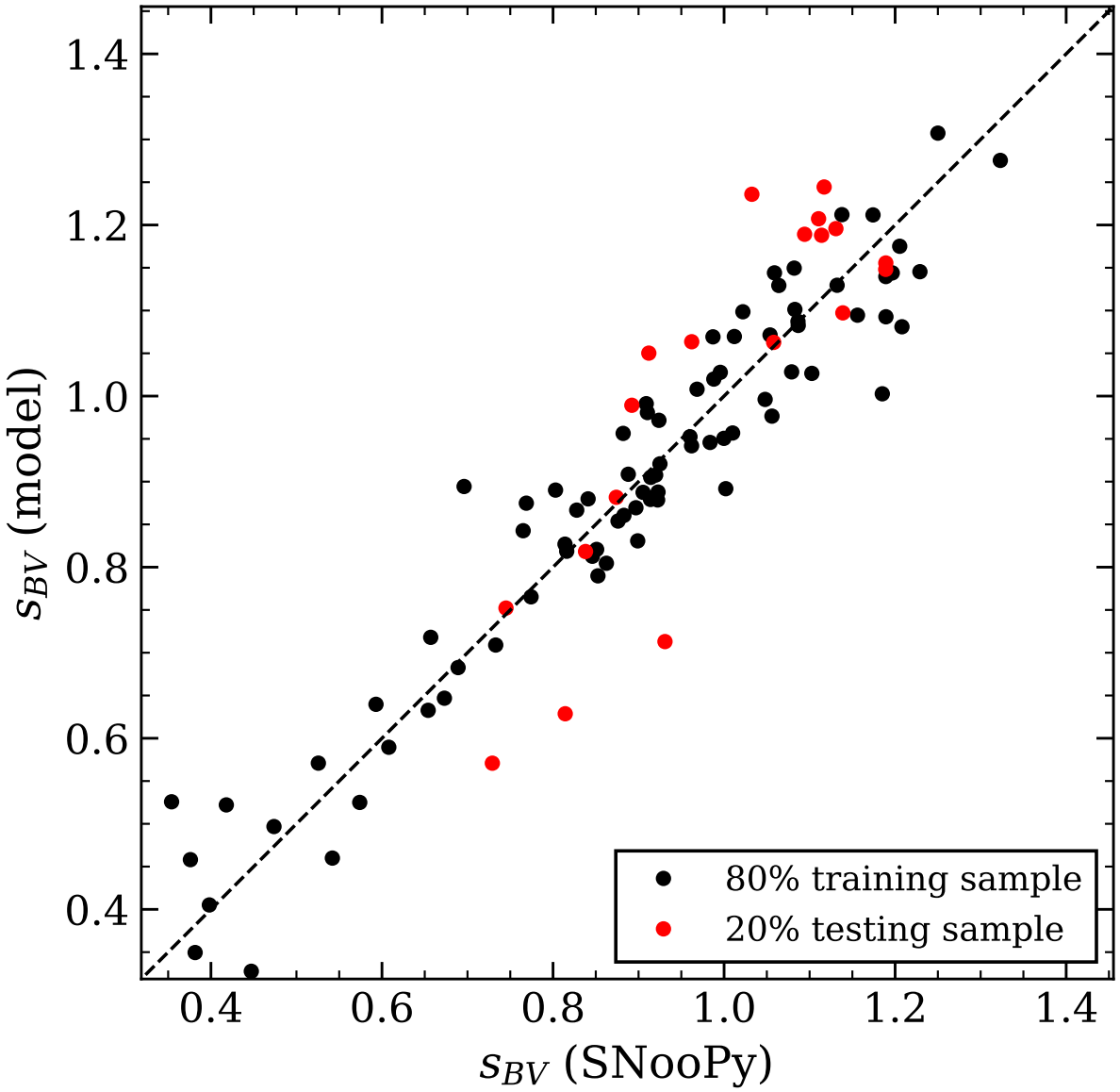


Figure 7.5: On the x-axis, s_{BV} calculated by `SNooPy` is shown for the same SNe Ia that produce the points in Figure 7.3. On the y-axis, s_{BV} is shown, calculated from a linear model described by Equation 7.1, where the a_n coefficients are determined via a least-squares fit to a random 80% of the available spectra representing black points. The red points are the remaining 20%, illustrating this s_{BV} model fits well with the majority of values given by `SNooPy`.

to the model. These s_{BV} values are then compared with those obtained by the light-curve-fitting tool `SNooPy`.

We see that the vast majority of these SNe Ia adhere to this model within a spread of about ± 0.1 units of s_{BV} for all values of measured s_{BV}. The testing sample seems to be commensurate

with the training sample for all $s_{\rm BV}$, aside from three outlying test points around $0.7 < s_{\rm BV} < 1.0$, although they do not lie too far outside this trend. In general, for such a simple linear model and method (using of a small region of the spectrum), this fit to $s_{\rm BV}$ values measured with SNooPy is surprisingly good. One could possibly use a more complex regression technique or test other regions in wavelength space to produce better fits (at the risk of over-fitting); however, the purpose of this is to show the simplicity of the correlation of Si II properties with $s_{\rm BV}$. In addition, for future studies, it could be interesting to see if any outliers to this model exhibit larger residuals in its NIR light-curve as determined by SNooPy. If an observed spectrum yields a substantially different value of $s_{\rm BV}$ for a supernova using these PCs compared to using the observed photometry, this could indicate intrinsic differences in its spectrum which could be seen in its NIR light-curves.

7.4. Conclusion

Here we have shown that projecting Si II onto the set of maximum-light eigenvectors gives a set of projection values that are shown to contain information that is similar to a supernova's Branch classification; however, the projections also contain Si II velocity information. This velocity information may be useful in explaining the dispersion seen in the Si II $\lambda6355$ pEW-velocity relation that leads to such high-velocity CN SNe Ia (see Figure 3.16. These projections or a similar quantity may be good candidates to replace pEW measurements, as obtaining pEWs can be inconsistent, since SNe Ia do not show a formal continuum in their spectrum, and pEWs are easily contaminated due to line-blending.

It is also seen that these projections extract the SNooPy measurement of $s_{\rm BV}$ well from a simple linear model, within a spread of about ±0.1 units. Interestingly, this provides a method of predicting the photometric $s_{\rm BV}$ for a SN Ia given only a maximum-light spectrum, therefore the maximum-light PCs contain enough information to sub-categorize SNe Ia. The photometric quantity $s_{\rm BV}$ is obtained roughly over the first 30 days from the B- and V-band light-curves. Nevertheless, a single maximum-light spectrum contains enough information to estimate $s_{\rm BV}$ to an accuracy of $\sim$10%, indicating that the maximum light spectrum contains the explosion history of the ejecta. This shows the complementarity of spectroscopy and photometry for understanding complex astrophysical phenomena.

CHAPTER 8

Conclusion

In this work, we have probed the observational diversity of SNe Ia using spectroscopy and photometry. We have created empirical models that connect SN Ia spectroscopy with their photometric properties.

Using pEWs of Si II and the line velocity $v_{\rm Si\,II}$, we apply GMMs to define a robust classification scheme based on that of Branch et al. (2006). When M_B information is included in these models, these classifications become more constrained, leading to less uncertainty in the predictions of group membership for individual SNe Ia. Similar GMMs have been applied to the M_B-vs-$v_{\rm Si\,II}$ space, wherein Polin et al. (2019) suggested that SN Ia observations can be distinguished as originating from either Chandrasekhar- or sub-Chandrasekhar-mass explosions. We find through a cluster analysis that there is no statistical significance that the M_B-vs-$v_{\rm Si\,II}$ space contains a dichotomous distribution of SNe Ia. Instead, it is shown that three groups best describe this space: a Main group that coincides SS SNe Ia and most CN SNe Ia, a Fast group mostly containing BL SNe Ia, and a Dim group containing CL SNe Ia. Using the GMM classification scheme for the Branch groups, we find that the Fast group also contains several CN SNe Ia, which also suggests that there is more complexity in the M_B-vs-$v_{\rm Si\,II}$ space than a simple dichotomous structure. In fact, this implies that there is an unexpected dispersion in the relationship between pEW(Si II $\lambda6355$) and $v_{\rm Si\,II}$ observed in SNe Ia.

Additionally, a method of approximating NIR SEDs of SNe Ia has been presented. This method is similar to the NIR spectral templates of Lu et al. (2023) for which PCA is performed on a set of FIRE spectra to understand the variable behavior of NIR spectra. However, instead of parameterizing the model by phase and $s_{\rm BV}$, we perform PCA simultaneously on merged optical+NIR spectra of SNe Ia. By doing so, the PCs describe how the optical and NIR spectra are correlated, and thus a NIR spectrum can be predicted given optical spectroscopy. It is shown in Chapter 4 that $s_{\rm BV}$ by itself does not fully capture the observed diversity of SNe Ia in the NIR, especially near and just after the secondary peak. Therefore, this new extrapolation technique is an attempt at accounting for more variation seen in the NIR that a single optical photometric parameter cannot capture. We find that this method yields comparable results to the NIR spectral templates; however, the spectral templates still typically outperform this extrapolation method in reproducing the observed NIR flux. This is likely due to the more strict data constraints needed for the extrapolation, because near-simultaneous optical and NIR observations are required to compare two wavelength regimes

at the same phases. Finally, we have also made good progress in modeling correlated noise and uncertainties from our extrapolation method using Gaussian processes, which may provide very useful information for light-curve fitters such as `SNooPy` in the future. A much more detailed analysis on the implementation of these modeled uncertainties is the subject of future work.

To connect the NIR variation seen via PCA with the optical in more detail, we also study the projections of Si II onto the maximum-light PCs used for NIR extrapolation. In doing so, we find that these projections seem to reproduce the Branch groups similar to pEW(Si II $\lambda5972$) and pEW(Si II $\lambda6355$); however, it also appears that the same projections contain $v_{\mathrm{Si\,II}}$ information, and they are not correlated with the two pEW quantities directly. The additional information that these projections provide may be indicative that they may be good candidates for replacing pEW measurements in a true Branch classification scheme. This is because measuring pEWs can be inconsistent, due to SNe Ia lacking a true continuum. These projections could also aid in explaining the dispersion seen between the Si II $\lambda6355$ pEW and velocity. We also find that the photometric color-stretch parameter s_{BV} can be estimated within a spread of about ±0.1 units using the projections onto PC_1 and PC_2. This indicates an interesting connection between spectroscopy and photometry of SNe Ia, where the maximum-light spectrum appears to contain information about the explosion history of ejecta over roughly 30 days past B-band maximum.

Bibliography

Amanullah, R., Goobar, A., Johansson, J., et al. 2014, ApJL, 788, L21, doi: 10.1088/2041-8205/788/2/L21

Anderson, J. P., González-Gaitán, S., Hamuy, M., et al. 2014, ApJ, 786, 67, doi: 10.1088/0004-637X/786/1/67

Arnett, W. D., Bahcall, J. N., Kirshner, R. P., & Woosley, S. E. 1989, ARA&A, 27, 629, doi: 10.1146/annurev.aa.27.090189.003213

Ashall, C., Mazzali, P., Bersier, D., et al. 2014, MNRAS, 445, 4427, doi: 10.1093/mnras/stu1995

Ashall, C., Hsiao, E. Y., Hoeflich, P., et al. 2019, ApJL, 875, L14, doi: 10.3847/2041-8213/ab1654

Ashall, C., Lu, J., Burns, C., et al. 2020, ApJL, 895, L3, doi: 10.3847/2041-8213/ab8e37

Bailey, S. 2012, PASP, 124, 1015, doi: 10.1086/668105

Barbon, R., Ciatti, F., & Rosino, L. 1979, A&A, 72, 287

Baron, E., Bongard, S., Branch, D., & Hauschildt, P. H. 2006, ApJ, 645, 480, doi: 10.1086/504101

Bersten, M. C., & Mazzali, P. A. 2017, in Handbook of Supernovae, ed. A. W. Alsabti & P. Murdin, 723, doi: 10.1007/978-3-319-21846-5_25

Blondin, S., Matheson, T., Kirshner, R. P., et al. 2012, AJ, 143, 126, doi: 10.1088/0004-6256/143/5/126

Branch, D., & Wheeler, J. C. 2017, Supernova Explosions, doi: 10.1007/978-3-662-55054-0

Branch, D., Dang, L. C., Hall, N., et al. 2006, PASP, 118, 560, doi: 10.1086/502778

Burns, C., Hsiao, E., Suntzeff, N., et al. 2021, The Astronomer's Telegram, 14441, 1

Burns, C. R., Stritzinger, M., Phillips, M. M., et al. 2011, AJ, 141, 19, doi: 10.1088/0004-6256/141/1/19

—. 2014, ApJ, 789, 32, doi: 10.1088/0004-637X/789/1/32

Burns, C. R., Parent, E., Phillips, M. M., et al. 2018, ApJ, 869, 56, doi: 10.3847/1538-4357/aae51c

Burrow, A. 2024, anthonyburrow/SNEx: v1.0.0: First release, 1.0.0, Zenodo, doi: 10.5281/zenodo.10849861

Burrow, A., Baron, E., Ashall, C., et al. 2020, ApJ, 901, 154, doi: 10.3847/1538-4357/abafa2

Burrow, A., Baron, E., Burns, C. R., et al. 2024, ApJ, 967, 55, doi: 10.3847/1538-4357/ad3c45

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245, doi: 10.1086/167900

Clifford, F. E., & Tayler, R. J. 1965, MmRAS, 69, 21

Dan, M., Rosswog, S., Brüggen, M., & Podsiadlowski, P. 2014, MNRAS, 438, 14, doi: 10.1093/mnras/stt1766

Day, N. E. 1969, Biometrika, 56, 463–474, doi: https://doi.org/10.1093/biomet/56.3.463

de Souza, R. S., Dantas, M. L. L., Costa-Duarte, M. V., et al. 2017, MNRAS, 472, 2808, doi: 10.1093/mnras/stx2156

Debes, J. H., Hoard, D. W., Wachter, S., Leisawitz, D. T., & Cohen, M. 2011, ApJS, 197, 38, doi: 10.1088/0067-0049/197/2/38

DerKacy, J. M., Baron, E., Branch, D., et al. 2020, ApJ, 901, 86, doi: 10.3847/1538-4357/abae67

DerKacy, J. M., Paugh, S., Baron, E., et al. 2023, MNRAS, 522, 3481, doi: 10.1093/mnras/stad1171

Dessart, L., Blondin, S., Hillier, D. J., & Khokhlov, A. 2014, MNRAS, 441, 532, doi: 10.1093/mnras/stu598

Domínguez, I., Piersanti, L., Bravo, E., et al. 2006, ApJ, 644, 21, doi: 10.1086/503534

Doull, B. A., & Baron, E. 2011, PASP, 123, 765, doi: 10.1086/661023

Elias, J. H., Frogel, J. A., Hackwell, J. A., & Persson, S. E. 1981, ApJL, 251, L13, doi: 10.1086/183683

Filippenko, A. V. 1988, AJ, 96, 1941, doi: 10.1086/114940

Filippenko, A. V., Porter, A. C., & Sargent, W. L. W. 1990, AJ, 100, 1575, doi: 10.1086/115618

Filippenko, A. V., Richmond, M. W., Matheson, T., et al. 1992a, ApJL, 384, L15, doi: 10.1086/186252

Filippenko, A. V., Richmond, M. W., Branch, D., et al. 1992b, AJ, 104, 1543, doi: 10.1086/116339

Finzi, A., & Wolf, R. A. 1967, ApJ, 150, 115, doi: 10.1086/149317

Fisher, A., Branch, D., Hatano, K., & Baron, E. 1999, MNRAS, 304, 67, doi: 10.1046/j.1365-8711.1999.02299.x

Folatelli, G., Phillips, M. M., Burns, C. R., et al. 2010, AJ, 139, 120, doi: 10.1088/0004-6256/139/1/120

Folatelli, G., Morrell, N., Phillips, M. M., et al. 2013, ApJ, 773, 53, doi: 10.1088/0004-637X/773/1/53

Foley, R. J., & Kasen, D. 2011, ApJ, 729, 55, doi: 10.1088/0004-637X/729/1/55

Foley, R. J., Narayan, G., Challis, P. J., et al. 2010, ApJ, 708, 1748, doi: 10.1088/0004-637X/708/2/1748

Foley, R. J., Challis, P. J., Chornock, R., et al. 2013, ApJ, 767, 57, doi: 10.1088/0004-637X/767/1/57

Foley, R. J., Fox, O. D., McCully, C., et al. 2014, MNRAS, 443, 2887, doi: 10.1093/mnras/stu1378

Friesen, B., Baron, E., Wisniewski, J. P., et al. 2014, ApJ, 792, 120, doi: 10.1088/0004-637X/792/2/120

Friesen, B., Baron, E., Parrent, J. T., et al. 2017, MNRAS, 467, 2392, doi: 10.1093/mnras/stx241

Galbany, L., Moreno-Raya, M. E., Ruiz-Lapuente, P., et al. 2016, MNRAS, 457, 525, doi: 10.1093/mnras/stw026

Garnavich, P. M., Bonanos, A. Z., Krisciunas, K., et al. 2004, ApJ, 613, 1120, doi: 10.1086/422986

Gaskell, C. M., Cappellaro, E., Dinerstein, H. L., et al. 1986, ApJL, 306, L77, doi: 10.1086/184709

GPy. 2012, GPy: A Gaussian process framework in python, http://github.com/SheffieldML/GPy

Guillochon, J., Dan, M., Ramirez-Ruiz, E., & Rosswog, S. 2010, ApJL, 709, L64, doi: 10.1088/2041-8205/709/1/L64

Guillochon, J., Parrent, J., Kelley, L. Z., & Margutti, R. 2017, ApJ, 835, 64, doi: 10.3847/1538-4357/835/1/64

Guy, J., Astier, P., Nobili, S., Regnault, N., & Pain, R. 2005, A&A, 443, 781, doi: 10.1051/0004-6361:20053025

Guy, J., Astier, P., Baumont, S., et al. 2007, A&A, 466, 11, doi: 10.1051/0004-6361:20066930

Hachinger, S., Mazzali, P. A., Tanaka, M., Hillebrandt, W., & Benetti, S. 2008, MNRAS, 389, 1087, doi: 10.1111/j.1365-2966.2008.13645.x

Hamuy, M., Phillips, M. M., Suntzeff, N. B., et al. 1996, AJ, 112, 2438, doi: 10.1086/118193

Hamuy, M., Maza, J., Pinto, P. A., et al. 2002, AJ, 124, 417, doi: 10.1086/340968

Hamuy, M., Phillips, M. M., Suntzeff, N. B., et al. 2003, Nature, 424, 651, doi: 10.1038/nature01854

Hamuy, M., Folatelli, G., Morrell, N. I., et al. 2006, PASP, 118, 2, doi: 10.1086/500228

Hicken, M., Garnavich, P. M., Prieto, J. L., et al. 2007, ApJL, 669, L17, doi: 10.1086/523301

Hillebrandt, W., & Hoflich, P. 1989, Reports on Progress in Physics, 52, 1421, doi: 10.1088/0034-4885/52/11/002

Hoeflich, P. 2017, in Handbook of Supernovae, ed. A. W. Alsabti & P. Murdin, 1151, doi: 10.1007/978-3-319-21846-5_56

Hoeflich, P., & Khokhlov, A. 1996, ApJ, 457, 500, doi: 10.1086/176748

Hoeflich, P., Khokhlov, A., Wheeler, J. C., et al. 1996, ApJL, 472, L81, doi: 10.1086/310363

Hoeflich, P., Khokhlov, A. M., & Wheeler, J. C. 1995, ApJ, 444, 831, doi: 10.1086/175656

Hosseinzadeh, G., Sand, D. J., Lundqvist, P., et al. 2022, ApJL, 933, L45, doi: 10.3847/2041-8213/ac7cef

Howell, D. A., Sullivan, M., Nugent, P. E., et al. 2006, Nature, 443, 308, doi: 10.1038/nature05103

Hoyle, F. 1946, MNRAS, 106, 343, doi: 10.1093/mnras/106.5.343

Hoyle, F., & Fowler, W. A. 1960, ApJ, 132, 565, doi: 10.1086/146963

Hsiao, E. Y., Conley, A., Howell, D. A., et al. 2007, ApJ, 663, 1187, doi: 10.1086/518232

Hsiao, E. Y., Phillips, M. M., Marion, G. H., et al. 2019, PASP, 131, 014002, doi: 10.1088/1538-3873/aae961

Hsiao, Y. C. E. 2009, PhD thesis, University of Victoria, Canada

Iben, I., J., & Tutukov, A. V. 1984, ApJS, 54, 335, doi: 10.1086/190932

Ilkov, M., & Soker, N. 2012, MNRAS, 419, 1695, doi: 10.1111/j.1365-2966.2011.19833.x

—. 2013, MNRAS, 428, 579, doi: 10.1093/mnras/sts053

Imshennik, V. S., & Nadezhin, D. K. 1989, Astrophys. Space Phys. Res., 8, 1

Jeffery, D. J., Leibundgut, B., Kirshner, R. P., et al. 1992, ApJ, 397, 304, doi: 10.1086/171787

Jha, S., Riess, A. G., & Kirshner, R. P. 2007, ApJ, 659, 122, doi: 10.1086/512054

Jha, S. W. 2017, in Handbook of Supernovae, ed. A. W. Alsabti & P. Murdin, 375, doi: 10.1007/978-3-319-21846-5_42

Kasen, D. 2006, ApJ, 649, 939, doi: 10.1086/506588

—. 2010, ApJ, 708, 1025, doi: 10.1088/0004-637X/708/2/1025

Kashi, A., & Soker, N. 2011, MNRAS, 417, 1466, doi: 10.1111/j.1365-2966.2011.19361.x

Kass, R., & Raftery, A. 1995, Journal of the American Statistical Association, 90, 773, doi: 10.2307/2291091

Kawabata, M., Maeda, K., Yamanaka, M., et al. 2020, ApJ, 893, 143, doi: 10.3847/1538-4357/ab8236

Kenworthy, W. D., Jones, D. O., Dai, M., et al. 2021, ApJ, 923, 265, doi: 10.3847/1538-4357/ac30d8

Khokhlov, A. M. 1991, A&A, 245, L25

Krisciunas, K., Contreras, C., Burns, C. R., et al. 2017, AJ, 154, 211, doi: 10.3847/1538-3881/aa8df0

Kromer, M., Pakmor, R., Taubenberger, S., et al. 2013, ApJL, 778, L18, doi: 10.1088/2041-8205/778/1/L18

Kutsuna, M., & Shigeyama, T. 2015, PASJ, 67, 54, doi: 10.1093/pasj/psv028

Kwok, L. A., Jha, S. W., Temim, T., et al. 2023, ApJL, 944, L3, doi: 10.3847/2041-8213/acb4ec

Leibundgut, B., Kirshner, R. P., Phillips, M. M., et al. 1993, AJ, 105, 301, doi: 10.1086/116427

Li, W., Filippenko, A. V., Chornock, R., et al. 2003, PASP, 115, 453, doi: 10.1086/374200

Li, W., Wang, X., Hu, M., et al. 2019, ApJ, 882, 30, doi: 10.3847/1538-4357/ab2b49

Lu, J., Hsiao, E. Y., Phillips, M. M., et al. 2023, ApJ, 948, 27, doi: 10.3847/1538-4357/acc100

MacQueen, J. B. 1967, in Proc. of the fifth Berkeley Symposium on Mathematical Statistics and Probability, ed. L. M. L. Cam & J. Neyman, Vol. 1 (University of California Press), 281–297

Maeda, K., Kutsuna, M., & Shigeyama, T. 2014, ApJ, 794, 37, doi: 10.1088/0004-637X/794/1/37

Marietta, E., Burrows, A., & Fryxell, B. 2000, ApJS, 128, 615, doi: 10.1086/313392

Mazzali, P. A., Danziger, I. J., & Turatto, M. 1995, A&A, 297, 509

Mazzali, P. A., Sullivan, M., Hachinger, S., et al. 2014, MNRAS, 439, 1959, doi: 10.1093/mnras/stu077

McCray, R. 1993, ARA&A, 31, 175, doi: 10.1146/annurev.aa.31.090193.001135

McLachlan, G. J., & Peel, D. 2000, Finite mixture models, Wiley series in probability and statistics (New York: J. Wiley & Sons)

Meikle, W. P. S. 2000, MNRAS, 314, 782, doi: 10.1046/j.1365-8711.2000.03411.x

Miyaji, S., Nomoto, K., Yokoi, K., & Sugimoto, D. 1980, PASJ, 32, 303

Morrell, N., Phillips, M. M., Folatelli, G., et al. 2024, ApJ, 967, 20, doi: 10.3847/1538-4357/ad38af

Murphy, K. 2012, Machine Learning: A Probabilistic Perspective, 1st edn., Adaptive Computation and Machine Learning series (Cambridge, MA: MIT Press)

Ni, Y. Q., Moon, D.-S., Drout, M. R., et al. 2023, ApJ, 959, 132, doi: 10.3847/1538-4357/ad0640

Nomoto, K. 1980, in Texas Workshop on Type I Supernovae, ed. J. C. Wheeler, 164–181

Nomoto, K. 1982, ApJ, 253, 798, doi: 10.1086/159682

—. 1984, ApJ, 277, 791, doi: 10.1086/161749

—. 1987, ApJ, 322, 206, doi: 10.1086/165716

Nomoto, K., Thielemann, F. K., & Yokoi, K. 1984, ApJ, 286, 644, doi: 10.1086/162639

Nugent, P., Baron, E., Branch, D., Fisher, A., & Hauschildt, P. H. 1997, ApJ, 485, 812, doi: 10.1086/304459

Nugent, P., Phillips, M., Baron, E., Branch, D., & Hauschildt, P. 1995, ApJL, 455, L147, doi: 10.1086/309846

Oke, J. B., & Sandage, A. 1968, ApJ, 154, 21, doi: 10.1086/149737

Paczynski, B. 1985, in Cataclysmic Variables and Low-Mass X-ray Binaries, ed. D. Q. Lamb & J. Patterson, 1, doi: 10.1007/978-94-009-5319-2_1

Paczynski, B., & Zytkow, A. N. 1978, ApJ, 222, 604, doi: 10.1086/156176

Pakmor, R., Hachinger, S., Röpke, F. K., & Hillebrandt, W. 2011, A&A, 528, A117, doi: 10.1051/0004-6361/201015653

Pakmor, R., Kromer, M., Röpke, F. K., et al. 2010, Nature, 463, 61, doi: 10.1038/nature08642

Pakmor, R., Kromer, M., Taubenberger, S., et al. 2012, ApJL, 747, L10, doi: 10.1088/2041-8205/747/1/L10

Pakmor, R., Kromer, M., Taubenberger, S., & Springel, V. 2013, ApJL, 770, L8, doi: 10.1088/2041-8205/770/1/L8

Pan, Y.-C. 2020, ApJL, 895, L5, doi: 10.3847/2041-8213/ab8e47

Papadogiannakis, S. 2019, PhD thesis, Stockholm University, Stockholm, doi: https://tinyurl.com/yc4gyah4

Papadogiannakis, S., Dhawan, S., Morosin, R., & Goobar, A. 2019, MNRAS, 485, 2343, doi: 10.1093/mnras/stz493

Parzen, E. 1962, The Annals of Mathematical Statistics, 33, 1065 , doi: 10.1214/aoms/1177704472

Patra, K. C., Yang, Y., Brink, T. G., et al. 2022, MNRAS, 509, 4058, doi: 10.1093/mnras/stab3136

Paz Arribas, M. 2017, PhD thesis, doi: 10.18452/18056

Pearson, K. 1901, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2, 559, doi: 10.1080/14786440109462720

Pedregosa, F., Varoquaux, G., Gramfort, A., et al. 2011, Journal of Machine Learning Research, 12, 2825, doi: http://jmlr.org/papers/v12/pedregosa11a.html

Pellegrino, C., Howell, D. A., Sarbadhicary, S. K., et al. 2020, ApJ, 897, 159, doi: 10.3847/1538-4357/ab8e3f

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, ApJ, 517, 565, doi: 10.1086/307221

Phillips, M. M. 1993, ApJL, 413, L105, doi: 10.1086/186970

Phillips, M. M., Lira, P., Suntzeff, N. B., et al. 1999, AJ, 118, 1766, doi: 10.1086/301032

Phillips, M. M., Wells, L. A., Suntzeff, N. B., et al. 1992, AJ, 103, 1632, doi: 10.1086/116177

Phillips, M. M., Contreras, C., Hsiao, E. Y., et al. 2019, PASP, 131, 014001, doi: 10.1088/1538-3873/aae8bd

Phillips, M. M., Ashall, C., Brown, P. J., et al. 2024, arXiv e-prints, arXiv:2405.15027. https://arxiv.org/abs/2405.15027

Polin, A., Nugent, P., & Kasen, D. 2019, ApJ, 873, 84, doi: 10.3847/1538-4357/aafb6a

Rakavy, G., Shaviv, G., & Zinamon, Z. 1967, ApJ, 150, 131, doi: 10.1086/149318

Rasmussen, C. E., & Williams, C. K. I. 2006, Gaussian Processes for Machine Learning (Cambridge, MA: MIT Press), doi: http://www.gaussianprocess.org/gpml/

Rebassa-Mansergas, A., Toonen, S., Korol, V., & Torres, S. 2019, MNRAS, 482, 3656, doi: 10.1093/mnras/sty2965

Richmond, M. W., & Smith, H. A. 2012, JAAVSO, 40, 872, doi: 10.48550/arXiv.1203.4013

Rosenblatt, M. 1956, The Annals of Mathematical Statistics, 27, 832 , doi: 10.1214/aoms/1177728190

Rousseeuw, P. J. 1987, Journal of Computational and Applied Mathematics, 20, 53, doi: http://dx.doi.org/10.1016/0377-0427(87)90125-7

Roweis, S. 1997, in Advances in Neural Information Processing Systems 10 (NIPS 1997), ed. M. Jordan, M. Kearns, & S. Solla. https://papers.nips.cc/paper_files/paper/1997/hash/d9731321ef4e063ebbee79298fa36f56-Abstract.html

Ruiz-Lapuente, P., Cappellaro, E., Turatto, M., et al. 1992, ApJL, 387, L33, doi: 10.1086/186299

Saunders, C., Aldering, G., Antilogus, P., et al. 2018, ApJ, 869, 167, doi: 10.3847/1538-4357/aaec7e

Scalzo, R. A., Aldering, G., Antilogus, P., et al. 2010, ApJ, 713, 1073, doi: 10.1088/0004-637X/713/2/1073

Schlafly, E. F., & Finkbeiner, D. P. 2011, ApJ, 737, 103, doi: 10.1088/0004-637X/737/2/103

Schlegel, E. M. 1990, MNRAS, 244, 269

Scott, D. W. 1992, Multivariate Density Estimation: Theory, Practice and Visualization, Wiley Series in Probability and Statistics (New York: J. Wiley & Sons), doi: 10.1002/9780470316849

Shen, K. J., & Bildsten, L. 2014, ApJ, 785, 61, doi: 10.1088/0004-637X/785/1/61

Shen, K. J., Kasen, D., Miles, B. J., & Townsley, D. M. 2018, ApJ, 854, 52, doi: 10.3847/1538-4357/aaa8de

Siebert, M. R., Foley, R. J., Jones, D. O., & Davis, K. W. 2020, MNRAS, 493, 5713, doi: 10.1093/mnras/staa577

Sienkiewicz, R. 1980, A&A, 85, 295

Sion, E. M., Acierno, M. J., & Tomczyk, S. 1979, ApJ, 230, 832, doi: 10.1086/157143

Smartt, S. J. 2009, ARA&A, 47, 63, doi: 10.1146/annurev-astro-082708-101737

—. 2015, PASA, 32, e016, doi: 10.1017/pasa.2015.17

Smith, N. 2014, ARA&A, 52, 487, doi: 10.1146/annurev-astro-081913-040025

Smith, N., Li, W., Filippenko, A. V., & Chornock, R. 2011, MNRAS, 412, 1522, doi: 10.1111/j.1365-2966.2011.17229.x

Sparks, W. M., & Stecher, T. P. 1974, ApJ, 188, 149, doi: 10.1086/152697

Srivastav, S., Ninan, J. P., Kumar, B., et al. 2016, MNRAS, 457, 1000, doi: 10.1093/mnras/stw039

Stritzinger, M. D., Holmbo, S., Morrell, N., et al. 2023, A&A, 675, A82, doi: 10.1051/0004-6361/202243376

Uddin, S. A., Burns, C. R., Phillips, M. M., et al. 2020, ApJ, 901, 143, doi: 10.3847/1538-4357/abafb7

Valenti, S., Howell, D. A., Stritzinger, M. D., et al. 2016, MNRAS, 459, 3939, doi: 10.1093/mnras/stw870

Wang, X., Wang, L., Filippenko, A. V., Zhang, T., & Zhao, X. 2013, Science, 340, 170, doi: 10.1126/science.1231502

Wang, X., Filippenko, A. V., Ganeshalingam, M., et al. 2009, ApJL, 699, L139, doi: 10.1088/0004-637X/699/2/L139

Webbink, R. F. 1984, ApJ, 277, 355, doi: 10.1086/161701

Werner, K., & Rauch, T. 2015, A&A, 584, A19, doi: 10.1051/0004-6361/201527261

Wheeler, J. C., Harkness, R. P., Clocchiatti, A., et al. 1994, ApJL, 436, L135, doi: 10.1086/187651

Whelan, J., & Iben, Icko, J. 1973, ApJ, 186, 1007, doi: 10.1086/152565

Woosley, S. E., & Weaver, T. A. 1994, ApJ, 423, 371, doi: 10.1086/173813

Xi, G., Wang, X., Li, W., et al. 2022, MNRAS, 517, 4098, doi: 10.1093/mnras/stac2848

Yarbrough, Z., Baron, E., DerKacy, J. M., et al. 2023, MNRAS, 521, 3873, doi: 10.1093/mnras/stad758

Zheng, W., Kelly, P. L., & Filippenko, A. V. 2018, ApJ, 858, 104, doi: 10.3847/1538-4357/aabaeb

Zhou, W.-H., Wang, B., Meng, X.-C., Liu, D.-D., & Zhao, G. 2015, Research in Astronomy and Astrophysics, 15, 1701, doi: 10.1088/1674-4527/15/10/007

CHAPTER A

Software Produced

A.1. `Spextractor`

`Spextractor` is a spectrum-smoothing tool based on a wavelength-dependent Gaussian process regression (GPR), which is discussed in more detail in Appendix B. This tool is a modified version of the original `Spextractor` created by Papadogiannakis (2019)[1]. Both versions perform a GPR using the Python module `GPy` (GPy, 2012) on a user-provided spectrum using a Matérn 3/2 kernel. In doing so, the mean of the model posterior distribution acts as a smoothed version of the input spectrum. The noise from the input spectrum is effectively removed from this mean, as the GPR kernel contains a noise hyperparameter that is optimized to fit the input spectrum, leading to a smooth mean function as output. In addition, the GPR also acts as an interpolator, as you can sample from the posterior distribution at any point in wavelength space.

An advantage to using GPR as an interpolator is that variance (uncertainty) in the output is also estimated. For interpolated points, there will be some uncertainty in the prediction based on the covariance matrix built from the data points of the observed spectrum. This modified version of `Spextractor` also adds the uncertainty in the observed spectrum to the GPR kernel in quadrature (see Appendix B). Propagating the errors in this way provides a much more realistic expectation of the uncertainty in predicted flux, especially when attempting to interpolate spectra with large uncertainties. Otherwise, the uncertainty given by the GPR is unrealistically small because the resolution of the observed spectrum is much higher than the typical wavelength scale that large changes of flux are seen, indicated by the lengthscale hyperparameter of the kernel.

Another major modification to `Spextractor` is the ability to downsample the input spectrum prior to performing the final GPR done on all the points in the spectrum. A GPR is typically computed with $\mathcal{O}(n^3)$ time complexity (Murphy, 2012), and with spectra containing hundreds to thousands of points, the use of GPR can become quite costly in time and resources. The $n \times n$ covariance matrix must also be stored in some way as well, so memory and data storage management plays an important role in performing multiple or larger instances of GPR. To mitigate these problems in `Spextractor`, the input spectrum is binned by a user-provided downsampling factor, and in each bin an integral of λF_λ is performed that allows one to solve for a weighted-average flux value that ensures that photon count is conserved in the bin. In this way, the downsampled spectrum, on average, contains much of the same information as the original spectrum. A GPR can

[1] https://github.com/astrobarn/spextractor

then be performed on the downsampled spectrum with minimal loss from the original spectrum.

In the original program, the posterior was sampled at wavelength points given to the prior. This was also changed to sample the posterior at uniformly spaced points at a higher resolution than the input spectrum such that an interpolation takes place. Any point-to-point variance in the interpolation is given in this way, which better represents the true uncertainty of the spectrum.

There are many more features that have been added to `Spextractor`, including spectrum preprocessing, which allows one to correct for observational effects such as host-galaxy redshift and extinction. Given observed photometry, it may also color-correct the spectrum using `SNooPy`. `Spextractor` also searches user-defined wavelength regions to determine spectral features in order to make consistent calculations of the pEWs and velocities of spectroscopic features of SNe Ia. These are the quantities used in Chapter 3 to formulate GMMs to describe the Branch groups and in Section 7.2 to search for correlations between maximum light PCs and common spectroscopic quantities.

`Spextractor` is openly maintained at https://github.com/anthonyburrow/spextractor with examples of its use and descriptions of the major functions and parameters available inside docstrings.

A.2. `SNIaDCA`

In Chapter 3, multi-dimensional GMMs were determined that describe CSP I+II data. These GMMs were created using the Python package `scikit-learn` (Pedregosa et al., 2011). Using the GMMs, one may input combinations of pEW(Si II λ5972), pEW(Si II λ6355), $v_{\text{Si II}}$, and M_B (depending on the desired model), and the probability that a SN Ia with those values belonging to each group of the GMM will be given. This is particularly useful in determining a consistent, statistical classification scheme that may be beneficial outside these projects. For convenience, a wrapper module called `SNIaDCA` has been created that uses these pre-determined GMMs to provide numerical probabilities of Branch-group membership of any given SN Ia.

`SNIaDCA` is available at https://github.com/anthonyburrow/SNIaDCA for public use, with examples of use provided. Updating these GMMs to be trained on more SN Ia data may be the subject of future work. It may be the case that providing more data to these models will further constrain the groups, reducing the uncertainty in Branch-group membership, which could be beneficial in the future study of the diversity of SNe Ia.

A.3. `SNEx`

`SNEx` (Burrow, 2024) is intended as a generalized spectrum-extrapolation tool written in Python. However, the primary focus of this software is to implement the NIR extrapolation of SN Ia spectra developed by Burrow et al. (2024). The conceptual technique behind this extrapolation procedure is also detailed here in Section 5.3. To summarize, a NIR spectrum is extrapolated using PCA by

projecting the optical spectrum onto a set of eigenvectors that statistically describe the intrinsic variation of SNe Ia spectra. CSP I+II optical+NIR spectra are used to model this variation.

The eigenvectors are precalculated and are saved in Python objects that contain other information such as training data, sample mean and standard deviation, etc. This software takes a spectrum as input along with user-defined parameters such as phase and the projection region itself. The spectrum is standardized and then projected onto the set of eigenvectors based on the spectrum's phase of observation (in days past B-band maximum) to produce the resulting linear combination spectrum. This linear combination acts as an extrapolation outside of the region that was projected onto the eigenvectors. A measurement of uncertainty is also provided — see Section 5.3.4 for a description of white-noise uncertainty currently accounted for by this software. In the future, after some testing, the technique for measuring uncertainties introduced in Chapter 6 will be implemented into `SNEx`.

`SNEx` is available at https://github.com/anthonyburrow/SNEx along with precalculated models that include the eigenvectors representing phases between -10 and $+40$ days relative to B-band maximum.

CHAPTER B

Gaussian Process Regression

This section briefly introduces Gaussian processes from a mathematical perspective and gives context to most major changes made to `Spextractor` with respect to fitting spectra via Gaussian process regression (GPR). The background information given paraphrases a more detailed explanation provided by Murphy (2012).

A Gaussian process is used to infer a distribution over functions $p(f|\mathbf{X}, \mathbf{y})$ given some observed input set $\mathbf{X}$ and output set $\mathbf{y}$ such that $y_i = f(x_i)$ for $y_i \in \mathbf{y}$ and $x_i \in \mathbf{X}$. In the context of fitting spectra with a GPR, $\mathbf{y}$ is the set of flux measurements, and $\mathbf{X}$ is the set of corresponding wavelengths at which these measurements were taken. The Gaussian process is defined to assume that $p(f(x_1), ..., f(x_N))$ is jointly Gaussian for an arbitrary set of inputs $x_1, ..., x_N$ for N observation points within $\mathbf{X}$ and therefore in $\mathbf{y}$. This joint Gaussian distribution has a mean $\mu(\mathbf{X})$ and covariance $\mathbf{K}(\mathbf{X})$ given that by $K_{ij} = \kappa(x_i, x_j)$, where κ is the positive semi-definite kernel function for which we choose the Matérn 3/2 covariance function (Rasmussen & Williams, 2006) as does the original `Spextractor` code. The functional distribution is then normally distributed as

$$\begin{pmatrix} \mathbf{f} \\ \mathbf{f}* \end{pmatrix} = \mathcal{N}\left(\begin{pmatrix} \mu \\ \mu_* \end{pmatrix}, \begin{pmatrix} \mathbf{K} & \mathbf{K}_* \\ \mathbf{K}_*^\mathsf{T} & \mathbf{K}_{**} \end{pmatrix} \right), \tag{B.1}$$

where $\mathbf{f}_*$ are functional outputs of $\mathbf{X}$, and $\mathbf{K} = \kappa(\mathbf{X}, \mathbf{X})$, $\mathbf{K}_* = \kappa(\mathbf{X}, \mathbf{X}_*)$, and $\mathbf{K}_{**} = \kappa(\mathbf{X}_*, \mathbf{X}_*)$, where $\mathbf{X}_*$ is some test (or prediction) input set. The output values $\mathbf{f}_*$ are then predicted with a GPR as mean values and associated variances for each test point. In this work we modify `Spextractor` such that $\mathbf{X}_* \neq \mathbf{X}$ to allow a representation of the variance between observed data points. More specifically, $\mathbf{X}_*$ was selected as a uniform distribution of 2,000 values that spanned the given spectrum.

If there are independently and identically distributed uncertainties ϵ_i (see, for example, Murphy, 2012) in the observed output such that

$$y_i = f(x_i) + \epsilon_i, \tag{B.2}$$

where in general each ϵ_i may not be equal to one another, we may write the distribution as

$$\begin{pmatrix} \mathbf{y} \\ \mathbf{f}* \end{pmatrix} = \mathcal{N}\left(\begin{pmatrix} \mu \\ \mu_* \end{pmatrix}, \begin{pmatrix} \mathbf{K}_y & \mathbf{K}_* \\ \mathbf{K}_*^{\mathsf{T}} & \mathbf{K}_{**} \end{pmatrix} \right), \tag{B.3}$$

where $\mathbf{K}_y$ is the covariance matrix defined by $\mathbf{K}_y \equiv \mathbf{K} + (\sigma^2)^{\mathsf{T}} I_N$ with σ^2 being a vector of associated output uncertainties and I_N the $N \times N$ identity. The quantity $(\sigma^2)^{\mathsf{T}} I_N$ is a white-noise kernel that is added to the Matérn 3/2 kernel in this case. We therefore add flux uncertainties σ^2_{flux} in quadrature to the kernel when flux uncertainties were provided.