PROGENITOR IDENTIFICATION OF TYPE IA SUPERNOVAE THROUGH STATISTICAL ABUNDANCE TOMOGRAPHY FROM OPTICAL SPECTRA WITH MACHINE LEARNING AND RADIATIVE TRANSFER

By

John Thomas O'Brien

A DISSERTATION

Submitted to

Michigan State University

in partial fulfillment of the requirements

for the degree of

Astronomy & Astrophysics - Doctor of Philosophy

2024

ABSTRACT

Type Ia supernovae (SNe Ia) seed galaxies with intermediate mass and iron group elements while also serving as standardizable candles for testing cosmological models. Despite their importance to understanding the evolution of the universe, the progenitors of SNe Ia remain elusive. Cosmic distance estimation and the chemical evolution of the universe depended precisely on the exactly progenitor mechanism, therefore there is a need to identify their origins.

Several models have been developed over the past several decades in an attempt to explain their unique observational features, all involving the thermonuclear disruption of a carbon-oxygen (C/O) white dwarf (WD) in a binary system in one of three major regimes. 1) A WD accretes mass from a companion in order to become massive enough to initiate central carbon-burning 2) A helium layer formed through accretion generates accretion-induced thermonuclear instabilities on the surface that drive a converging shock into the core 3) A violent merger between a pair of WDs initiates a central detonation. The nature of the binary interaction has a large effect on which channel may lead to the explosion including the composition of the accreted material, the accretion rate, and the nature of the binary evolution of the system. The nucleosynthetic products as well as their distributions within the material ejected from the explosions of SNe Ia depend heavily on these proposed progenitor channels and therefore measurements of these features offer an opportunity to make testable predictions regarding their origins. The elemental composition within the ejecta is can be determined by modeling spectral observations with radiative transfer simulations. Traditionally, such methods have relied on qualitative metrics of model fits and manual adjustments of elemental compositions due to the heavy computational cost of performing spectral synthesis and the large number of parameters required to describe the composition of the ejecta leading to results devoid of information regarding uncertainties or parameter degeneracies resulting in uncertain progenitor identification.

This thesis presents a novel methodology for rapid probabilistic reconstructions of SNe Ia through the application of deep-learning accelerated radiative transfer simulations

under parametric ejecta models. This methodology is applied to explore the progenitors of SNe Ia in three different projects. First, analysis of the elemental composition of the outer ejecta of the archetypal SN Ia SN 2002bo shows that the parameter space is complex with multiple parameter degeneracies and multi-modalities but is overall inconsistent with traditional pure-deflagration models. Second, modeling the outer-ejecta of a population of the super-luminous silicon-deficient 1991T-like thermonuclear supernovae finds that they appear as an extension or extreme case of the normal SN Ia population with their unique observational signatures primarily dictated by small deviations in production of intermediate mass elements with higher ionization rates. Finally, progenitor channel probabilities are prescribed to the well observed SN Ia SN 2011fe by sampling a space of high-dimensional hydrodynamical models corresponding to a variety of SN Ia progenitor channels showing that it is best described by a core-detonation model of a sub-Chandrasekhar mass WD. These results both elucidate the progenitors of SNe Ia as well as provide insight regarding the limitations of current models to solve more detailed questions about their origins.

Contents

1 Introduction

SN Ia have garnered the attention and wonder of astronomers for hundreds of years. Their earliest observations changed the way astronomers think about the heavens and today they challenge the way we think about the cosmos. Their violent thermonuclear explosions create the elements that make up the world we see around us every day from the iron in our blood to the silicon in the computers on which you are likely reading this thesis. A paradoxical mystery, while they have critically advanced our understanding of the origins of the universe and of ourselves, the origin of Type Ia supernovae to this day remains elusive.

1.1 Early Observations

Supernovae historically were classified according to their spectra. Type I supernovae were originally the class of supernovae similar to SN1937C (Popper, 1937) characterized by the presence of a wide variety of optical emission features (Minkowski, 1941), while Type II supernovae showed a continuous spectrum. Comparison of high quality spectra over a large temporal range confirmed that Type I supernovae comprised a homogeneous group (Oke & Searle, 1974). More specifically, Branch & Patchett (1973) described the primary distinguishing feature of Type I supernovae as a lack of hydrogen emission in their spectra and applied the feature of homogeneity of the class in order to perform estimates for the rate of cosmic expansion. Later analysis by Elias et al. (1985) found that the group of Type I supernovae contained multiple spectroscopically distinct objects classifying those with strong variable absorption at around 1.2 um as Type Ia supernovae.

For decades a variety of models were developed to try to explain the origin and nature of these supernovae. Hoyle & Fowler (1960) proposed an explosion driven by carbon-burning in a degenerate stellar core as opposed to the catastrophic implosions we now know today to core-collapse supernovae, indicating an early theoretical progenitor separation between the Type I and Type II supernovae. Finzi & Wolf (1967) suggested that the electron-degenerate

stellar remnants, white-dwarfs, could serve as the origin of Type Ia supernovae through the process of inverse beta-decay in Mg or Ca. Arnett (1969) investigated a carbon-detonation mechanism for a near Chandrasekhar mass degenerate stellar core igniting a thermonuclear runaway burning alpha-elements such as oxygen and silicon all the way up to a significant amount of ^{56}Ni . Colgate & McKee (1969) found the decay chain of the isotope ^{56}Ni , the predominant end product of nucleosynthesis in thermonuclear explosions, drives the luminosity and structure of the light curves of Type Ia supernovae, and at the time predicted a required mass of 0.25 $M_\odot$of ^{56}Ni . Improved numerical simulations by Axelrod (1980) including steady-state non-LTE radiative transfer simulations including energy deposition from the radioactive decay of ^{56}Ni and its daughter nuclide, ^{56}Co, refined the estimate for the necessary mass of synthesized ^{56}Ni to be above 0.5 $M_\odot$.

Nomoto et al. (1976) proposed an alternative to the carbon-detonation scenario: a carbon deflagration which propagates through the core through convective heat transfer without growing into a detonation ending in the complete disruption of the star. This was an attempt to reconcile the deviations from the predictions of the carbon-detonation model such as over-production of iron-group elements. Nomoto & Sugimoto (1977) proposed a mechanism for rejuvenation of a helium dwarf through mass accretion of hydrogen from a companion star burning into helium. In this model a central helium flash grows into a detonation wave causing the supernova and disrupting the white dwarf completely.

By the 1980s, the field of Type Ia supernovae was both growing rapidly and relatively disorganized with many new proposed models, observations, and analysis methods aided by technological advances such as CCD cameras for telescopes and exponentially scaling computing power. The review by Wheeler (1981) attempted to condense the current state of the field as well as offer some commentary on the state of knowledge on the progenitors of Type Ia supernovae. By now it was clear that Type Ia supernovae must originate from at least C/O WDs as explosions of helium white dwarfs would generate ^{56}Ni at higher velocities than what had been observed. Since the spectra of Type Ia supernovae at maximum light

do not show lines of Nickel or Cobalt (Branch, 1980), it was presumed that these features are blanketed by an enriched helium shell that forms through accretion from a companion star in a binary system. The accretion rate must also be low enough that a hydrogen-rich envelope would not form otherwise hydrogen features would be seen in the spectrum. The double-detonation scenario was also disfavored as current models at the time showed they would completely burn the helium layer to ^{56}Ni with no helium blanket remaining (see e.g. Nomoto, 1980; Woosley et al., 1980). The favored model was the slow accretion model igniting carbon-burning in the center of the WD, but even these models had issues such as the helium shell being ejected with too low of a velocity to match spectral observations. At this point there was still no definitive model that could reproduce all of the features of Type Ia supernovae spectral and photometric observations, but it was clear that they must originate from accreting C/O WDs in binary systems.

1.2 Type Ia Progenitors

In the modern sense, there are three major mechanisms that can lead to the explosion of a carbon-oxygen white-dwarf each with their own unique predictions for the abundance stratification of the ejecta. The first is a near-Chandrasekhar mass white dwarf accretes material from a companion star to approach the density and temperature limit for nuclear burning in the core, resulting in either a pure-deflagration (Whelan & Iben, 1973; Nomoto et al., 1976) or delayed detonation (Khokhlov, 1991). In the case of a pure deflagration, less ^{56}Ni is produced leading to a less energetic explosion with a lower velocity ejecta and lower luminosity. 3D models show the turbulent deflagration causes the elemental abundances to be mixed at all depths into the explosion (Röpke et al., 2007). Conversely, in a delayed detonation, the ejecta is pre-expanded by an initial deflagration wave which transitions into a detonation which creates closer to the observed 0.6 $M_\odot$ of ^{56}Ni to be produced, with elements near the center being more mixed and elements further out being more stratified in their abundances. The second mechanism, called the double-detonation model (e.g. Taam, 1980),

involves the accretion of hydrogen or helium from a companion star or white-dwarf onto the surface of the primary, forming a helium shell on the surface either directly or through nuclear processes burning hydrogen. Eventually, a helium flash on the surface of the primary is ignited and wraps around the white-dwarf which in turn compresses to core leading to a central denotation. Double-detonation explosions tend to show more stratification near the core with excess heavy nuclear burning products near the surface resulting from the initial helium detonation, depending on the mass of the initial helium layer. The last prominent channel is the violent merger between a pair of white-dwarfs initiating a central detonation (e.g. Nomoto, 1982; Webbink, 1984; Iben & Tutukov, 1984; van Kerkwijk et al., 2010; Livio & Riess, 2003; Kashi & Soker, 2011). The super-sonic detonation wave rapidly moves through out the ejecta without much time for mixing of elements causing a much more stratified distribution of elements with heavy elements produced at deeper layers and lighter elements near the surface. In order to test each of these scenarios, a method for determination of the abundance stratification of the ejecta from SNe Ia explosions must be applied.

1.3   Abundance Tomography

Stehle et al. (2005) pioneered the method of abundance tomography for determination of the abundance stratification of the ejecta of SNe Ia. As the ejecta from the SN Ia expands and cools, the optical depth at a given ejecta velocity to decreases. This provides and effective photosphere below which light emitted from the supernova is thermalized and individual line interactions do not contribute substantially to the spectrum. Since only line information above this photosphere is imprinted onto spectral observations, early time spectra will provide information on the density and abundances present within the ejecta only in the outer most regions, and later time spectra will include information deeper into the explosion. By sequentially modeling spectral time-series with radiative transfer simulations, a full profile of the composition of the ejecta at all velocities can be reconstructed. Abundance tomography has since become a standard technique for determination of the abundance stratification of

various supernovae (e.g. Benetti et al., 2004; Mazzali et al., 2007) which has in turn allowed for the testing of predictions from various progenitor channels.

Abundance tomography has several limitations that prevent it from providing a full understanding of the composition of the ejecta and therefore the progenitors of SNe Ia. The first limitation is computational. Synthetic spectra are created from a model through radiative transfer simulations, which require between tens of minutes to hours to run (see Blondin et al., 2022) which leads to two major issues. The second limitation is dimensional. Modeling just 10 elements with 100 regions creates a 1000 dimensional parameter space that would be infeasible to sample even with computationally expedient radiative transfer. This means that a Bayesian analysis of spectra, which would require millions of sequential evaluations for each spectrum, are computationally infeasible so fits are performed using qualitative Chi-by-eye estimates through manual adjustments of elemental composition, density, and luminosity. Therefore, the resulting abundance stratification contains no quantitative information regarding the uncertainties in these quantities and may overlook potential parameter degeneracies in this extremely large parameter space. In this thesis, I will present a novel methodology that overcomes these limitations and apply it to a series of optical spectral observations to many SNe Ia.

1.4 Probabilistic Abundance Tomography

The following is a collection of 3 papers I have written throughout my graduate studies for my PhD in Astrophysics and Astronomy at Michigan State University and represent a culmination of both the growth of my own experience in attempting to understand SNe Ia as well as an advancement in the field's quest to pin down the underlying mechanisms that progenate and propagate their explosions. The first paper, O'Brien et al. (2021), is a published letter to the Astrophysical journal regarding determination of the progenitor of SN 2002bo through an advanced state-of-the-art method of abundance tomography as well as improved estimates for the chemical composition of the outer layer of the supernova's

ejecta. An optical spectrum taken approximately 10 days post-explosion is analysed by fitting deep-learning accelerated radiative transfer simulations through Bayesian inference. This paper concludes that the deflagration progenitor channel is not an adequate description of the early-time observations of SN 2002bo due to mismatches between the inferred abundance and density structure and those predicted by deflagration models. This paper is presented below in Chapter 2.

The second paper, O'Brien et al. (2023) is a publication in the Astrophysical Journal that analyses the relationship between the normal Type Ia population commonly used for cosmic distance estimation and the super-luminous subclass of Type Ia supernovae known as 1991T-like thermonuclear supernovae. An analytical ejecta model whose structure is constrained by theoretical progenitor models is applied to a population of normal and 1991T-like Type Ia supernovae to determine the abundance structure of their ejecta and determine if these two populations are directly separable. We find that the normal Type Ia population smoothly transitions into the 1991T-like population without a clear distinct clustering in the space of masses and compositions of their early-time ejecta. Furthermore, we investigate the underlying properties of their explosions leading to their observational separation, finding the objects classified by 1991T-likes appear as normal Type Ia supernovae with either lower production of intermediate mass elements or higher ionization states of those intermediate mass elements, indicating the possibility that the observational differences between the two classes come from a combination of factors. This paper is presented in Chapter 3.

The final paper is a direct attempt at progenitor identification of the archetypal Type Ia supernovae SN 2011fe and is currently awaiting publication. A large collection of hydro-dynamic simulations from different progenitor theories are pitted against one another to determine which channel best produces observations over the largest possible range of model configurations. We find that the outer ejecta of SN 2011fe in the early phases looks most like a pure core-detonation progenitor and therefore must be either a special case of a delayed detonation or double detonation, or possibly a novel unknown mechanism. Our comparison

between the double detonation and delayed detonation progenitor channel probabilities shows that one progenitor channel can not be statistically favored over another within the limitations of the tested spectral time-series and within the current limitations of modern radiative transfer and hydrodynamic modeling. Therefore, this paper also acts to encourage the direction of the field to focus on refining our simulations and understanding of these progenitor channels as well as their implications upon predictions of observables. This paper is presented in Chapter 4.

An appendix is provided at the end which includes the culmination of all appendices provided throughout these three papers for further reference on implementation details as well as data access in accordance with the formatting guide for PhD dissertations provided by Michigan State university.

2 Probabilistic Reconstruction of Type Ia Supernova SN 2002bo

ABSTRACT

Manual fits to spectral times series of Type Ia supernovae have provided a method of reconstructing the explosion from a parametric model but due to lack of information about model uncertainties or parameter degeneracies direct comparison between theory and observation is difficult. In order to mitigate this important problem we present a new way to probabilistically reconstruct the outer ejecta of the normal Type Ia supernova SN 2002bo. A single epoch spectrum, taken 10 days before maximum light, is fit by a 13-parameter model describing the elemental composition of the ejecta and the explosion physics (density, temperature, velocity, and explosion epoch). Model evaluation is performed through the application of a novel rapid spectral synthesis technique in which the radiative transfer code, TARDIS, is accelerated by a machine-learning framework. Analysis of the posterior distribution reveals a complex and degenerate parameter space and allows direct comparison to various hydrodynamic models. Our analysis favors detonation over deflagration scenarios and we find that our technique offers a novel way to compare simulation to observation.

2.1 Introduction

SNe Ia are a spectral class of supernovae defined by their lack of hydrogen lines and the presence of silicon lines. SNe Ia are caused by the thermonuclear explosion of carbon-oxygen white dwarfs in binary systems forming a large amount of ^{56}Ni, which drives the behavior of their light curves (Colgate & McKee, 1969). They contribute significantly to the chemical evolution of their host galaxies through the dispersion of iron-peak elements formed during the explosion (Kobayashi et al., 2020, see Figure 39).

Their ability to act as standardizable candles (Phillips, 1993) has served as a powerful tool in constraining cosmological parameters (Branch, 1992; Riess et al., 1998), though there

remains significant variation in their brightness that is unaccounted for (e.g. Blondin et al., 2012a). Furthermore, the identification of the ignition mechanism leading to SNe Ia remains an area of active research (see e.g. Polin et al., 2019).

The community has identified multiple promising pathways to explosions, many of which originate in a binary system. For example, nuclear burning may be ignited by either the merger of two CO white dwarfs (e.g. Nomoto, 1982; Webbink, 1984; Iben & Tutukov, 1984; van Kerkwijk et al., 2010; Livio & Riess, 2003; Kashi & Soker, 2011), or accretion from a companion star forming a near-Chandrasekhar mass CO white dwarf causing a central ignition (e.g. Whelan & Iben, 1973), or accretion of a helium layer onto a sub-Chandrasekhar mass white dwarf (e.g. Woosley & Weaver, 1994; Fink et al., 2010a; Shen et al., 2018; Polin et al., 2019) leading to a surface helium detonation that propagates inward triggering central ignition.

Various models have been proposed to describe the processes underlying SNe Ia. In particular, the speed at which the nuclear burning propagates through the star remains poorly understood. Reconstructing the explosion from spectral time series (also known as abundance tomography) is a crucial tool to understand the explosion scenario (see e.g. Mazzali et al., 2007). Previous work into abundance tomography (e.g. Stehle et al., 2005; Sauer & Mazzali, 2008) has begun to show us a picture of how SN Ia explosions compare to theoretical models, but they lack a probabilistic interpretation of their parameters.

SN 2002bo is a "Branch normal" (Branch et al., 1993; Benetti et al., 2004; Branch et al., 2006) SN Ia discovered in NGC 3190 that has been modeled extensively in the literature (e.g. Stehle et al., 2005; Sauer & Mazzali, 2008; Benetti et al., 2004; Kerzendorf, 2011). Specifically, Stehle et al. (2005) used a multi-line Monte-Carlo code to manually reconstruct the explosion mechanism using 13 epochs of spectra. Their inference suggests a Type Ia with moderate amounts of mixing of ^{56}Ni and intermediate-mass elements, as well as a lack of carbon in the ejecta, indicating a possible explosion asymmetry and orientation effects.

While these results offer a good foray into the investigation of the abundance tomography

of SNe Ia, the lack of uncertainty or error analysis limits our ability to constrain the range of possible explosion scenarios. Physical sources of uncertainty such a line-blending as well as potential parameter degeneracies warrant the need for probability distributions.

In this work, we present a method of Bayesian inference of supernova parameters by applying the radiative transfer code TARDIS (Kerzendorf & Sim, 2014), accelerated by a machine-learning framework (Kerzendorf et al., 2021), to a single spectrum of SN 2002bo taken 10 days before maximum light (Benetti et al., 2004). We begin with a description of our model and associated parameters in Section 2.2. The sampling of the parameter space, including a discussion on prior distributions and resulting posterior distributions, is given in Section 2.3. A summary of results can be found in Section 4.6. Appendices are included to provide general background on the techniques used for spectral synthesis acceleration as well as additional data used in our analysis. In Appendix A, we outline a machine-learning framework used to accelerate TARDIS evaluation. Finally, in Appendix B, links to data sources and data products are provided in order to assist researchers who wish to replicate our findings.

2.2 Explosion Model

The optical spectrum of SN 2002bo 10 days before maximum light is modeled with spectral synthesis produced by the radiative transfer code TARDIS. TARDIS is a modular framework that allows for the use of various physics modules and has been widely used for modeling a range of photospheric SNe (e.g. Magee et al., 2016; Boyle et al., 2017; Barna et al., 2017; Vogl et al., 2020a; Gillanders et al., 2020; Williamson et al., 2021). TARDIS approximates the radiation field in the ejecta with an optically thick inner boundary and an optically thin homologously expanding ejecta above. There is no energy generation in the simulation area and the energy injection is purely set by the temperature, T_{inner}, and radius, r_{inner}, of this inner boundary. The optically thin ejecta is divided into a series of concentric shells in velocity space. The velocity of each shell is determined by the inner boundary velocity, v_{inner},

and increases linearly up to an outer velocity boundary. The radius of the inner boundary, r_{inner}, and consequently the radius of the shells, are set by the product of v_{inner} with the time since the explosion, t_{exp}.

We employ a power law relationship of the density with the velocity parameterized by the power law index α_ρ such that $\rho_{\mathrm{shell}} \propto v_{\mathrm{shell}}^{\alpha_\rho}$[1]. In previous works (Stehle et al., 2005; Kerzendorf, 2011), the density profile of SN Ia ejecta has been described by a 1-dimensional parameterized explosion model known as W7 (see e.g. Nomoto et al., 1984) which can be approximated as a power law between velocity and density with an exponent of -7 (Branch et al., 1985). In order to account for deviations from the W7 power law profile we have left the power law index as a free parameter in our study, the prior for which can be found in Table 1.

We approximate the elemental composition of the ejecta by assuming a uniform distribution of abundances above the photosphere (the same abundance values are used in each shell). We explored a set of abundances commonly used in the literature (e.g. Stehle et al., 2005; Sauer & Mazzali, 2008; Kerzendorf, 2011), namely carbon, magnesium, silicon, sulfur, calcium, titanium, and chromium. Iron, cobalt, and nickel abundances were split up into the decay chain of the isotope ^{56}Ni and stable iron. These elements account for the majority of the mass in explosion models and are well constrained by the spectra of SNe Ia (Filippenko, 1997). The set of abundances (C, Mg, Si, S, Ca, Ti, Cr, Fe$_{\mathrm{stable}}$, and ^{56}Ni) and explosion parameters (T_{inner}, v_{inner}, t_{exp}, and α_ρ) all together compose a 13-dimensional parameter space to model our spectra.

For the plasma state, we have chosen the `nebular` ionization approximation implemented in TARDIS and the `dilute-lte` excitation approximation. The radiation-matter interaction is modeled using the `macroatom` prescription. We have also set the number of packets to be equal to $400\,000$. The final spectral calculation uses the formal integral method (Lucy, 1999a) rather than straight packet statistics. Configuration of TARDIS can be found in Appendix B.

---

[1]The reference density is pre-computed from the power law index to match that of the W7 model at $10\,000\,\mathrm{km/s}$

2.2.1 Model Evaluation

Spectral synthesis from our model with TARDIS, on average, takes approximately 10 minutes of CPU time on an Intel® Xeon® E5-2670 v2 CPU. Kerzendorf et al. (2021) estimates the time required to explore a 20-parameter toy-model at this rate to be ∼420 years. Such a time constraint on model evaluation imposes a restriction upon our ability to use radiative transfer codes as a method of exploring the posterior distribution of SN Ia models. In order to subvert this restriction, we have implemented a technique for speeding up our model evaluation by 8 orders of magnitude based upon the machine-learning framework developed by Kerzendorf et al. (2021). The estimation of our models through this technique is known as *emulation* and the machine-learning framework we used will from here on be referred to as the *emulator*. Details of the emulator including architecture, accuracy, and error analysis can be found in Appendix A. We find our emulator predicts the synthetic spectra produced by TARDIS given a set of model parameters within 1% and is therefore an effective and necessary substitute for model evaluation.

2.3 Parameter Inference

Vectors of candidate input abundances (carbon, magnesium, etc.) and explosion parameters, $\vec{\theta} = \{C, Mg, \ldots, t_{exp}, \alpha_\rho\}$, are drawn from a prior-distribution described in Section 2.3.1. Model spectra are then produced by the emulator, where the emulated synthetic spectrum is predicted using the input parameters $\vec{\theta}$. We determine the likelihood of a given model through the application of a likelihood function described in Section 2.3.2. We have developed a non-χ^2 likelihood function that takes into account systematic differences between our theoretical and observed spectra. Lastly, in Section 2.3.3, we outline the Monte Carlo sampling technique used to construct the posterior distribution.

2.3.1 Prior Distribution

We developed a distribution from which to draw our prior samples based on parameters of SN Ia abundances taken from the Heidelberg Supernova Model Archive (HESMA). We specifically used the set of abundance profiles provided from various SNe Ia hydrodynamic simulations (Fink et al., 2014; Noebauer et al., 2017; Kromer et al., 2013, 2015; Sim et al., 2010; Noebauer et al., 2017; Fink et al., 2018; Marquardt et al., 2015; Fink et al., 2010b; Kromer et al., 2010; Sim et al., 2012; Gronow et al., 2020) to determine the range of input parameters. We determined the bounds of our prior by taking the 60% quantile of the distribution of abundances from the HESMA models where the shell velocity was above $10\,000\,\mathrm{km/s}$ in order to be consistent with the expected structure of the outer shells.

Abundances were sampled uniformly in log-space with any remaining abundance fraction filled in with oxygen such that all abundance fractions summed to unity. Oxygen is often used as a "filler" element in supernova fitting (e.g. Hachinger et al., 2017) due to the insensitivity to changes in the spectrum with respect to the oxygen mass fraction (cf. Hachinger, 2011, Sec. 2.2.5.2). The oxygen abundance is therefore only determined implicitly and is not included as a model parameter.

For all other model parameters, we sampled along a uniform distribution. We used the values for explosion time, ejecta velocity, photospheric boundary temperature, and density profile power law exponent from the fit made by Kerzendorf (2011) as centroids. We then reviewed the works of Stehle et al. (2005) and Benetti et al. (2004) to determine reasonable ranges of uncertainties on these values which were used to set the edges of the distribution. The range of values sampled for each parameter can be found in Table 1.

2.3.2 Likelihood Estimation

While our emulator accurately recreates the behavior of TARDIS under our spectral synthesis model, observations of real spectra are subject to physical and systematic biases. In order to compare our model spectra, $\hat{f}(\vec{\theta})$, to observation, f_{obs}, we develop a likelihood function, $\mathcal{L}(\vec{\theta})$,

that corrects our model spectra and compares the results to our observed spectrum.

A correction function, $C(\hat{f}(\vec{\theta}))$, is applied to our model spectra. $C(\hat{f}(\vec{\theta}))$ first applies a redshift correction to set the frame of the model spectrum to the observed frame of SN 2002bo at z=0.0042 (Benetti et al., 2004). A host extinction correction is then performed using the model described by Cardelli et al. (1989) using $R_V = 3.1$ (Schlafly & Finkbeiner, 2011) and $E(B - V) = 0.3$ (Benetti et al., 2004). Finally, a continuum removal technique described by Tonry & Davis (1979) and Blondin & Tonry (2007a) is applied to the model spectrum. The continuum is estimated using a zero-mean 13-point cubic spline fit to the spectrum. We apply this continuum removal to our model spectra first, then we multiply by the continuum that would be removed by applying the same technique to the observed spectrum. Finally, the resulting continuum-removed model spectrum is linearly interpolated to the wavelength bins of the observed spectrum. Applying the corrections in this way allows us to compare our simulated spectra directly to the observed spectrum.

We compare our corrected model spectrum to the observed spectrum using a Gaussian likelihood function,

$$\log \mathcal{L}(\vec{\theta}) = -\frac{1}{2} \sum_{\lambda} \left[\frac{(C(\hat{f}(\vec{\theta})) - f_{\text{obs}})^2_{\lambda}}{s^2} + \log(2\pi s^2) \right],$$

where λ represents the wavelength bin of the observed spectrum of SN 2002bo in the observed frame. The parameter s^2 estimates the variance of our posterior distribution over model spectra which we infer as another parameter (Hogg et al., 2010) with a log-uniform prior.

2.3.3   Posterior Distribution

The topology of the posterior distribution is unknown a priori, and could contain complicated degeneracies or multimodalities. Nested sampling (Skilling, 2004; Buchner, 2021) is a robust Monte Carlo technique for this setting. We use the MLFriends algorithm (Buchner, 2014, 2017) implemented in the *UltraNest* package (Buchner, 2021). The posterior distribution was explored with 400 live points. It converged to the target distribution after 10 000 iterations

and required $1\,000\,000$ model evaluations.

2.4   Results

Figure 1 shows the converged parameter distributions from our statistical inference. Silicon and sulfur abundances contribute the largest fraction by mass of the ejecta which can be inferred from the spectral features present in SN 2002bo. Stehle et al. (2005) used a similar code to TARDIS to manually fit the spectral time series of SN 2002bo. However, due to differences in methodologies, direct comparison of elemental abundances is difficult and must be approximated. Since Stehle et al. (2005) does not provide uncertainties, we make the assumption that the uncertainty in their reported elemental abundances within various layers of the ejecta are comparable to those found in our study. Unfortunately, the full model inferred by Stehle et al. (2005) is not directly available for download so we estimate abundances in terms of mass fractions from the figures (Stehle et al., 2005, Figure 5).

We compare our findings to their range of abundances reported in the velocity interval from $10\,000\,\mathrm{km/s}$ to $15\,000\,\mathrm{km/s}$ and generally find good agreement within our uncertainty ranges. We find a significant lack of carbon in the ejecta consistent with their analysis. The range of abundances determined from their analysis of silicon (0.2 - 0.4), sulfur (0.06-0.1), and ^{56}Ni (0.09 - 0.11) all overlap with our 68% confidence interval in Table 1. Their abundances of iron ($< 10^{-4}$ - 0.04) and calcium (0.01 - 0.05) were slightly outside this region but are consistent if the level of uncertainty in their analysis is similar to ours. Individual values for both titanium and chromium are not available so performing a direct comparison is not particularly reasonable or reliable.

By far our largest deviation from Stehle et al. (2005) is our magnesium abundance. Magnesium has the largest range of uncertainty in our analysis, spanning nearly four orders of magnitude. Operating under the assumption that the uncertainties in Stehle et al. (2005) are comparable to ours, not much information can be gathered from a comparison of values between the two studies as the magnesium abundance is mostly uninformative.
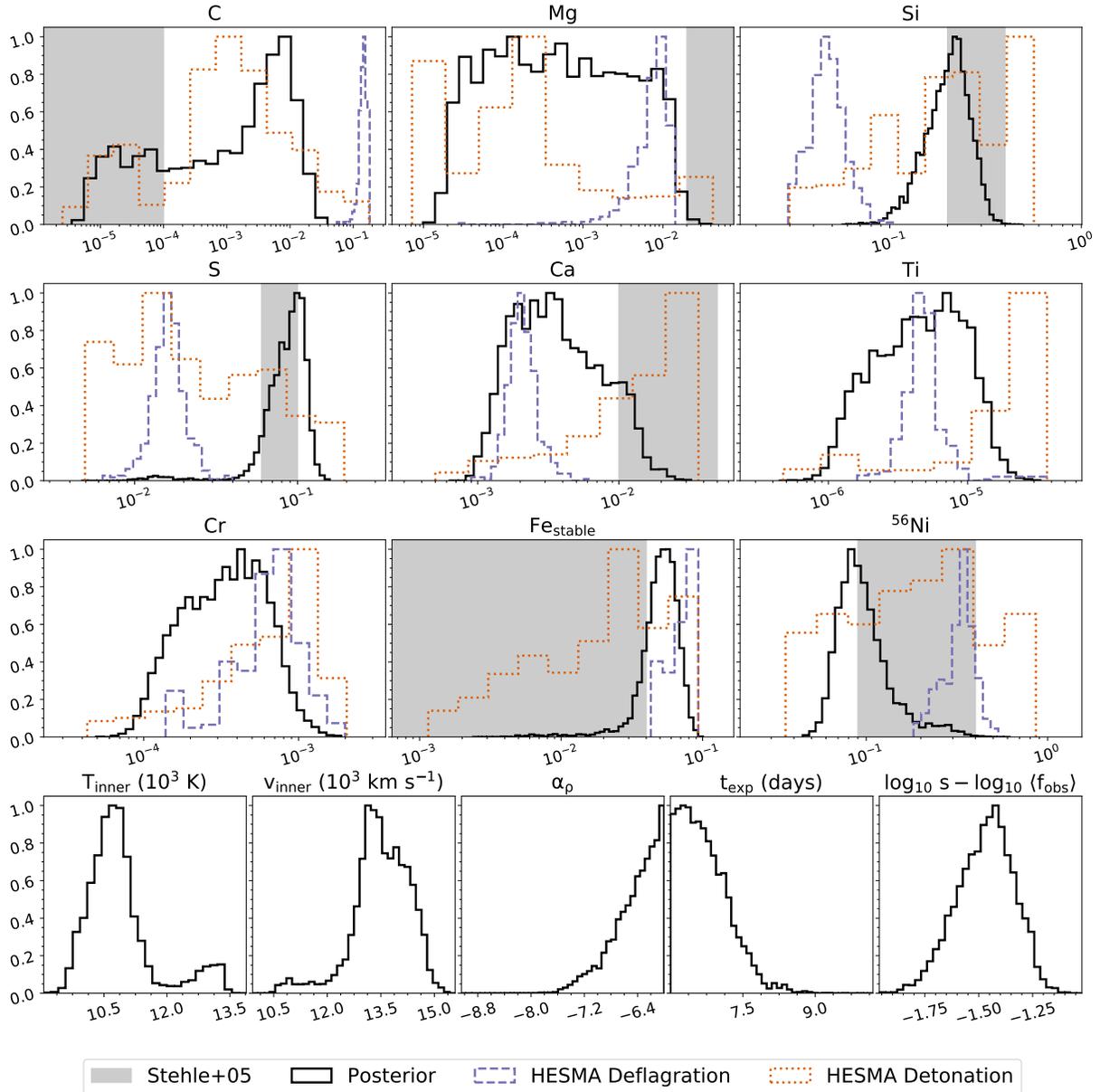
Figure 1: Posterior distribution of the parameter space sampled using nested sampling (Black). Overlaid are distributions of elemental abundances above 10 000 km/s taken from various HESMA models. Pure deflagration models are shown in green while pure detonation models are shown in orange. DDT models are not included as they would not be noticeably distinguishable from pure detonation models at this early epoch. Estimates of the range of abundances of elements in ejecta layers between 10 000 km/s and 15 000 km/s from Stehle et al. (2005) are represented by the grey shaded regions. Due to differences in methodology we do not have reliable estimates for the abundaces of titanium and chromium from Stehle et al. (2005).

Parameter	Prior Bounds		Posterior Percentiles		
	Minimum	Maximum	16%	50%	84%
C	2.3×10^{-6}	0.17	9.5×10^{-5}	0.0015	0.0085
Mg	8.3×10^{-6}	0.036	0.00011	0.00049	0.0047
Si	0.029	0.58	0.17	0.21	0.26
S	0.005	0.19	0.074	0.09	0.11
Ca	0.00043	0.039	0.0021	0.0034	0.0084
Ti	4.4×10^{-7}	3.7×10^{-5}	2.7×10^{-6}	4.7×10^{-6}	9.7×10^{-6}
Cr	3.8×10^{-5}	0.0022	0.00021	0.00034	0.00062
Fe_{stable}	0.0011	0.1	0.044	0.052	0.065
^{56}Ni	0.037	0.85	0.078	0.091	0.13
T_{inner} (K)	8000	18 000	10383	10720	11357
v_{inner} (km s^{-1})	7000	20 000	13100	13508	14291
α_ρ	-10	-6	-6.10	-6.36	-6.63
t_{exp} (days)	6	13	6.32	6.64	7.21
$\log_{10} s$	-18	-14	-15.91	-15.81	-15.69

Table 1: The range of parameters sampled from our prior distribution along with their estimates determined by the posterior distribution. The abundance distributions are based upon log-uniform sampling but modifications are made in order to assure that the sum of abundance parameters add to unity. All other values displayed are sampled uniformly. For a full description of the abundance sampling method see Section 2.3.1. Elemental abundances are shown in terms of mass fractions. Estimates from the posterior distribution are presented as the median with the edges of the 68% confidence interval.

We constrain $t_{\mathrm{exp}} = 6.64^{7.21}_{6.32}$ days[2] which is slightly below that of Benetti et al. (2004, $t_{\mathrm{exp}} = 7.9 \pm 0.5$ days) and Stehle et al. (2005, $t_{\mathrm{exp}} = 8.04$ days). Our estimates for both T_{inner} and v_{inner} are consistent with the range of values found by Stehle et al. (2005) for spectra between nearby epochs. The overall agreement of our results with similar previous attempts at manual fitting as well as theoretical models for SNe Ia explosion physics demonstrates that our model is consistent with the current literature.

There are a few notable mismatches between our posterior spectra and the observed spectrum (Figure 2). In the S Roman2 doublet our model over-fits the left peak and under-fits the right peak. This discrepancy is a common occurrence in radiative transfer model fits (see e.g. Stehle et al., 2005) to SN Ia spectra and is due to a poor understanding of the lines lists and occupation numbers in this region. Since our abundance distribution through the ejecta is approximated to be uniform, the iron abundance in the outer layers is generally overestimated. This causes line blanketing as the bluer packets are reflected back inwards resulting in a higher radiative temperature as well as less flux at the blue end of the spectrum. The higher temperatures affect the overall ionization state of the plasma causing the Si Roman2 to Si Roman3 ratio to decrease, weakening the Si Roman2 (5972 Å) feature. The poor fit to the Si Roman2 doublet is also seen in previous studies (see e.g. Benetti et al., 2004).

We are able to perform a direct comparison of inferred model parameters of a real SN Ia spectrum to statistical samples of theoretical explosion models. In addition to the posterior distributions of the model parameters inferred for SN 2002bo, Figure 1 shows the distribution of abundances from two classes of models taken from the HESMA data sets above $10\,000$ km/s corresponding to pure-deflagrations and pure-detonations. Deflagration to detonation transition (DDT) models are not included as they would be indistinguishable from pure-detonation models above the photosphere at these early times. The posterior distribution best matches with the distribution of abundances sampled from the HESMA detonation models, while mostly excluding the pure deflagration models. The unfavorability of

---

[2]See Table 1 for description of quantification
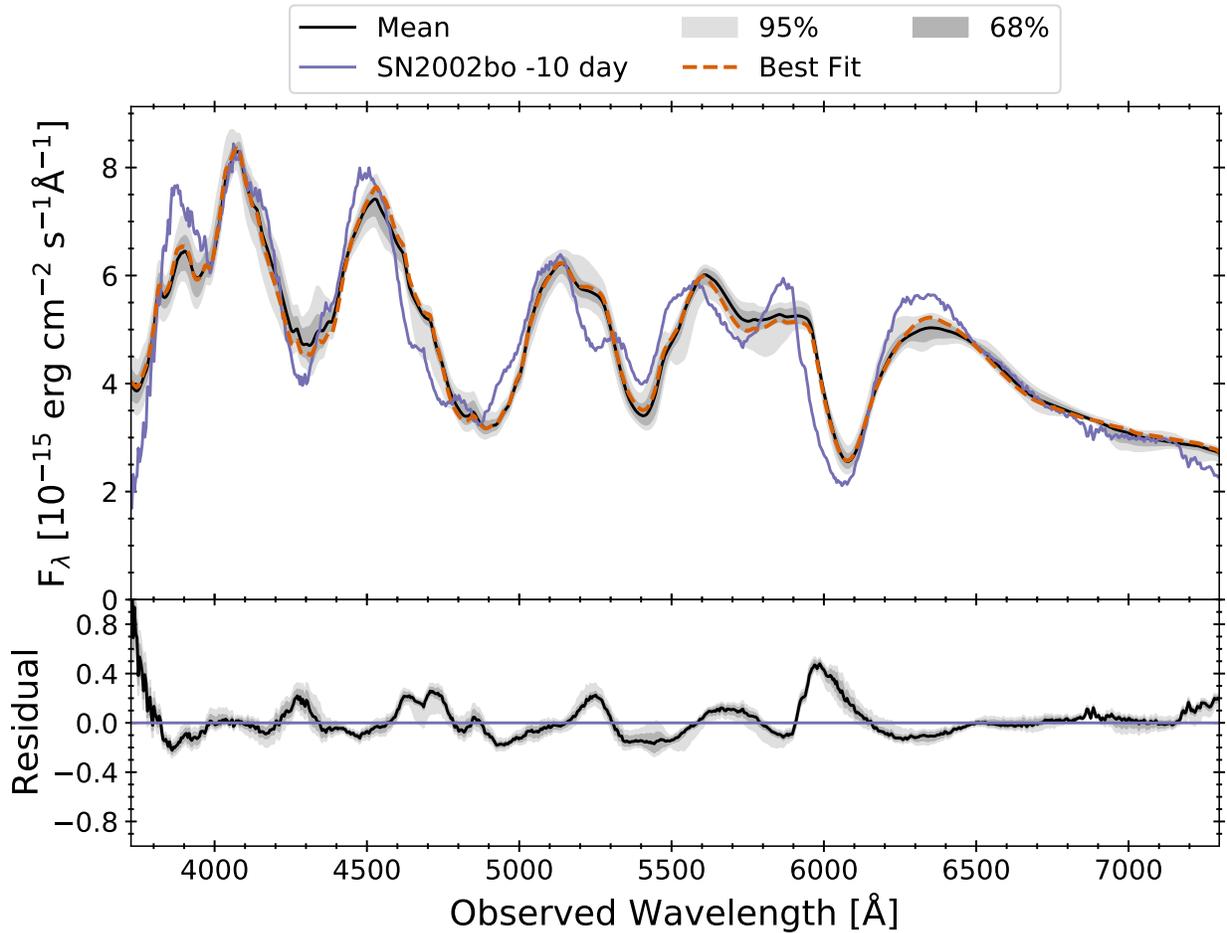
Figure 2: Fit to observed SN 2002bo −10 day spectrum (green) using nested sampling to sample the posterior distribution. The best fit spectrum (orange), represented by the maximum likelihood sample, shows a decent fit to the spectrum but misses features around 5972 Å and 3900 Å as well as much of the UV. The mean of the posterior distribution is shown in black with the 68% and 95% regions in grey and light grey respectively. Posterior spectra are presented after application of the correction function described in Section 2.3.2. The residual distribution is shown as the fractional error between our posterior and our observed spectrum.

pure-deflagration models is strongly apparent for the distribution of carbon, sulfur, and silicon abundances in Figure 1. Calcium and chromium abundances slightly favor pure-deflagration hydrodynamic models, though their distribution widths are large and stretch over a few orders of magnitude indicating that these abundances are not affecting the final shape of the spectrum significantly. We find that our initial modeling of the -10 day spectrum of SN 2002bo generally favors detonation or DDT models.

Figure 3 demonstrates the complexity of the posterior distribution of elemental abundances. A small multimodality in the sulfur abundance raises the possibility of manual fits becoming trapped in local minima. The joint probability distribution of stable iron with both silicon and ^{56}Ni is degenerate and multimodal. Such complexities indicate that any single set of model parameters may only describe one of a distribution of parameters that all appear to model the observed spectrum to similar accuracy. Despite some of the large variations and complexity in the posterior distribution of parameters (Figure 1), the distribution of model spectra produced by these parameters (Figure 2) is within 3% variation of the mean of the observed spectrum.

2.5   Conclusion

We present a probabilistic reconstruction of a SN Ia explosion. Our results generally agree with manual fits (see e.g. Stehle et al., 2005). We estimate the distribution of elemental abundances required to reproduce the observation of an early-time spectrum of SN 2002bo. Degeneracies and multimodalities in certain parameters showcase the need for a Bayesian treatment to draw secure physical conclusions since similar spectra may be synthesized over a wide and complex space of parameters. The posterior distribution is compared to the distribution of elemental abundances computed from various explosion models in HESMA. We find that our analysis favors detonation models over pure-deflagration models. Given the speed and effectiveness of our modeling technique, we have demonstrated a new avenue for investigating the inner mechanisms driving SN Ia explosions.
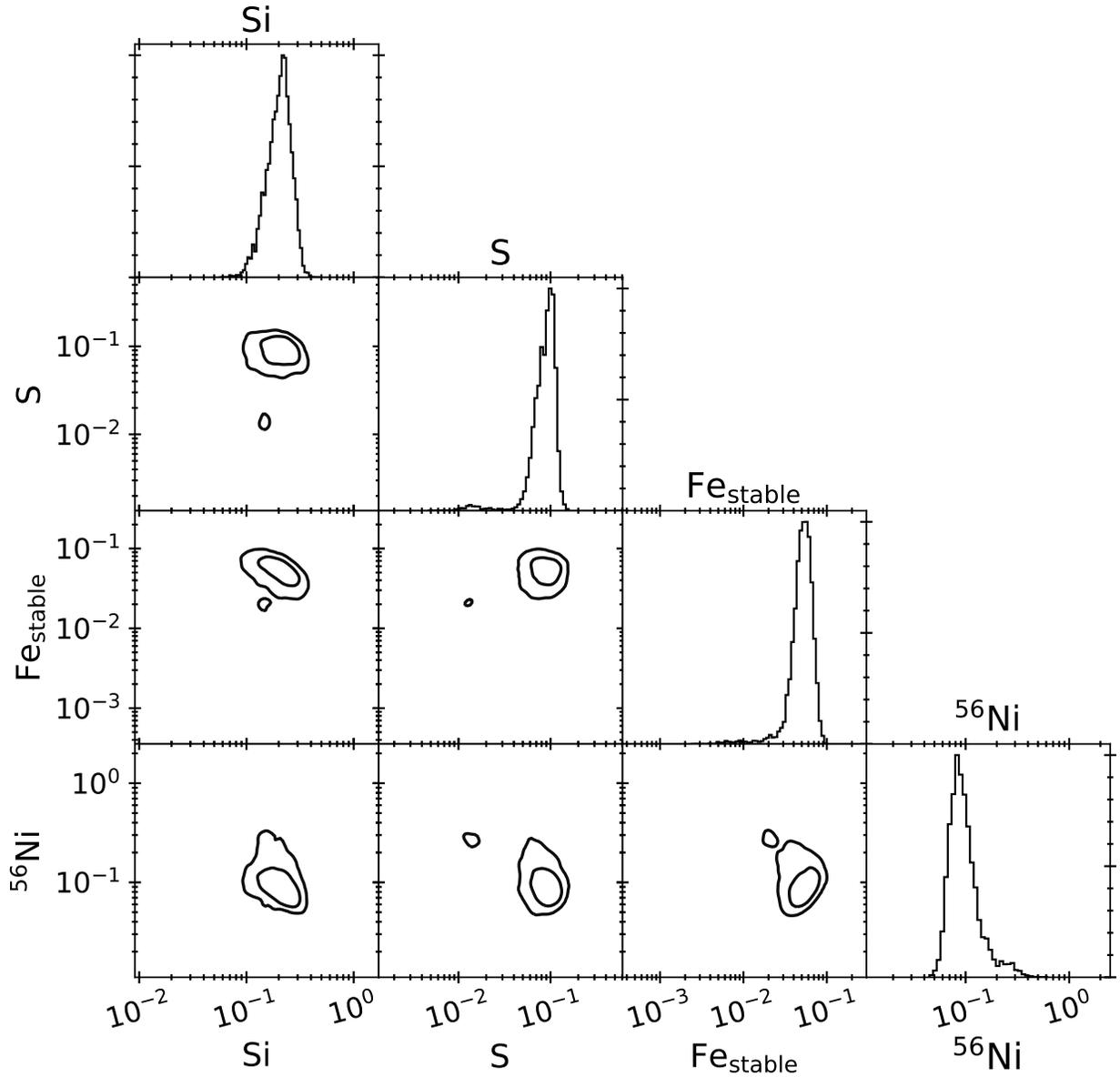
Figure 3: Posterior probability distribution of the elemental abundances of silicon, sulfur, stable iron, and ^{56}Ni. Contours show 68% and 95% confidence intervals of the Gaussian kernel density estimation (KDE) over the joint distribution of each parameter. Degeneracies and multimodalities in elemental abundances are apparent.

2.6 acknowledgements

---

[3] `https://www.astropy.org`
[4] `https://extinction.readthedocs.io/en/latest`
[5] `https://matplotlib.org`
[6] `https://numba.pydata.org`
[7] `https://numpy.org`
[8] `https://pandas.pydata.org`
[9] `https://scikit-learn.org`
[10] `https://www.scipy.org/`
[11] `https://www.tensorflow.org/`
[12] `https://johannesbuchner.github.io/UltraNest`

thank these communities for providing and maintaining the resources that allow science to be done in an open, replicable, and accessible way.

3 1991T-Like Type Ia Supernovae as an Extension of the Normal Population

ABSTRACT

Type Ia supernovae remain poorly understood despite decades of investigation. Massive computationally intensive hydrodynamic simulations have been developed and run to model an ever-growing number of proposed progenitor channels. Further complicating the matter, a large number of sub-types of Type Ia supernovae have been identified in recent decades. Due to the massive computational load required, inference of the internal structure of Type Ia supernovae ejecta directly from observations using simulations has previously been computationally intractable. However, deep-learning emulators for radiation transport simulations have alleviated such barriers. We perform abundance tomography on 40 Type Ia supernovae from optical spectra using the radiative transfer code TARDIS accelerated by the probabilistic DALEK deep-learning emulator. We apply a parametric model of potential outer ejecta structures to comparatively investigate abundance distributions and internal ionization fractions of intermediate-mass elements between normal and 1991T-like Type Ia supernovae in the early phases. Our inference shows that the outer ejecta of 1991T-like Type Ia supernovae are under-abundant in the typical intermediate mass elements that heavily contribute to the spectral line formation seen in normal Type Ia supernovae at early times. Additionally, we find that the intermediate-mass elements present in 1991T-like Type Ia supernovae are highly ionized compared to those in the normal Type Ia population. Finally, we conclude that the transition between normal and 1991T-like Type Ia supernovae appears to be continuous observationally and that the observed differences come out of a combination of both abundance and ionization fractions in these supernovae populations.

3.1 Introduction

Type Ia supernovae (SNe Ia), the thermonuclear explosions of Carbon/Oxygen (C/O) white dwarfs (WD), are critical tools for understanding the evolution of the cosmos. SNe Ia populate galaxies with iron-group and intermediate-mass elements (Kobayashi et al., 2020, see Figure 39) critical to the formation of planets and late-generation stars. As cosmic distance indicators (Phillips, 1993), SNe Ia have proved useful in both determining the size and age of the universe, as well as for probing the nature of dark energy (Branch, 1992; Riess et al., 1998; Perlmutter et al., 1999). However, despite their success as tools for probing galactic and cosmological evolution, the mechanism(s) underlying their ignition remain poorly understood.

An ever-increasing number of progenitor models have been proposed in the literature to explain SNe Ia, usually involving some sort of mass transfer from a binary companion. For example, ignition of a C/O WD has been suggested to be the result of mergers with a binary companion (e.g. Nomoto, 1982; Webbink, 1984; Iben & Tutukov, 1984; van Kerkwijk et al., 2010; Livio & Riess, 2003; Kashi & Soker, 2011), accretion from a companion star onto a near Chandrasekhar-mass (M_{Ch}) WD (e.g. Whelan & Iben, 1973) resulting in a turbulent deflagration, or accretion onto a sub-M_{Ch} WD resulting in a super-sonic detonation (e.g. Woosley & Weaver, 1994; Fink et al., 2010a; Shen et al., 2018; Polin et al., 2019; Pakmor et al., 2022). Despite intensive work and an ever-increasing number of proposed models, secure progenitor identification from spectral and photometric observations remains elusive.

Further complicating the matter of progenitor identification is the large spectroscopic diversity of thermonuclear SNe that have been identified over the past few decades. A large number of objects within the class of SNe Ia with unique spectral and photometric properties have resulted in a variety of classification schemes (e.g. Branch et al., 2006; Taubenberger, 2017). These objects range from the subluminous low-velocity Type Iax/02cx-like thermonuclear supernovae (Foley et al., 2013) to super-luminous shallow-silicon (Branch et al., 2006) 1991T-like SNe Ia (Filippenko et al., 1992; Phillips et al., 1992). The variation in

the properties of these objects leads us to consider the possibility of either distinct progenitor channels for these sub-types or a unified progenitor model that can describe massive variations in spectral properties.

We begin our investigation into the relationship between SNe Ia sub-types from the bright end of thermonuclear transients by focusing on the super-luminous 1991T-like SNe Ia. On the observational side, 1991T-like SNe Ia appear spectroscopically similar to the normal (Branch et al., 1993; Benetti et al., 2004; Branch et al., 2006) SNe Ia population after their light curves achieve maximum brightness (Phillips et al., 1992), however, in their early phases they are quite distinct. Their early-time spectra contain strong absorption lines of high-velocity Fe II/Fe III and lack the characteristic strong Si II absorption features of normal SNe Ia (Filippenko et al., 1992; Filippenko, 1997). Additionally, 1991T-like SNe Ia lie close to the normal SNe Ia in the space of the luminosity-decline rate relation, potentiality contaminating SNe Ia samples used for cosmic distance measurements due to Malmquist bias at high redshift (Sasdelli et al., 2014). On the theoretical side, Filippenko et al. (1992) originally proposed that 1991T-like supernovae may either be the results of either a double-detonation initiated at an intermediate layer in the progenitor WD, or a delayed-detonation model, in order to explain the large amount of the progenitor WD that is burned into ^{56}Ni and the apparent narrow region of IMEs present with the ejecta. Since then, many hypotheses have been proposed to explain the deviations in photometric and spectroscopic properties of 1991T-like SNe Ia from the normal SNe Ia population with mixed success (e.g. Ruiz-Lapuente et al., 1992; Mazzali et al., 1995; Liu et al., 1997; Marquardt et al., 2015; Seitenzahl et al., 2016). A definitive connection between the theoretical progenitor channels for 1991T-like SNe Ia and their observed spectral properties requires constraining the possible theoretical models to the observations directly.

In this paper, we present ejecta reconstructions from inference and a direct statistical comparison of the internal ejecta state between populations of 35 normal and five 1991T-like SNe Ia. The ejecta models are presented as probability distributions determined through

Bayesian inference performed on single-epoch early-time optical spectra. Our parameterized ejecta model is based on hydrodynamical simulations of a variety of proposed progenitor systems from the Heidelberg Supernova Model Archive (HESMA Kromer et al., 2017). We use a radiative transport scheme based on the open-source radiative transfer code TARDIS (Kerzendorf & Sim, 2014) accelerated by the probabilistic DALEK deep-learning emulator (Kerzendorf et al., 2022) to generate predictions of synthetic spectra over our space of model parameters. We compare distributions of ejecta compositions and ionization states between the normal and 1991T-like SNe Ia populations and identify a relationship between their internal structure and observed spectral features. These results allow us to better understand the relationship between normal SNe Ia and 1991T-like SNe Ia.

In Section 3.2, we describe the selection criteria for the observed spectra samples of normal and 1991T-like SNe Ia that we chose to model. Section 3.3 describes the parametric ejecta model implemented to model these spectra as well as details of the radiative transfer simulation and its acceleration through emulation. Section 3.4 describes the inference framework for estimating the posterior distributions of our model parameters, including the form of the likelihood function and the priors placed on our parameters. Results of our modeling are presented in Section 3.5 along with a discussion of their physical implications. Finally, our conclusions and final discussion are summarized in Section 3.6.

3.2   Data

We select a sample of normal and 1991T-like SNe Ia with spectra between 7 and 14 days before the B-band maximum in the light curve as these observations are well into the photospheric phase (see Section 3.3.4) when the ejecta are still optically thick. This selection was designed to model spectral observations taken 8 to 12 days post-explosion given a rise-time of 19.5 days with a 2.5 day rise-time uncertainty. Some studies (e.g. Phillips et al., 2022) will discern between the transitional shallow-silicon 1999aa-like SNe Ia and the 1991T-like SNe Ia due to the presence of early-time Calcium features and larger Si II absorption features. For the

purposes of this study, we group together 1999aa-like SNe Ia with 1991T-like SNe Ia and refer to the joint group as 1991T-like SNe Ia.

Our sample of selected SNe Ia is based on the sample investigated by Polin et al. (2021) as these objects are well studied. We queried WISeREP (Yaron & Gal-Yam, 2012) for each selected SN, filtering to only objects labeled as either Ia or Ia-pec with spectra within our time interval, and found a total of 158 spectra covering 44 objects. For each object found, we select a single spectrum to model according to two criteria relating to the quality and coverage of the data. We first attempt to limit our sets of spectra to those with coverage of more than 90% of the wavelength range from $3400\,\text{Å}$ to $7600\,\text{Å}$ which corresponds to the wavelength range of our model. If no spectra for a single object fully encompass this range, we keep them for the next step of selection to maximize the number of objects we model. We then select the spectrum from each object with the highest average signal-to-noise ratio. If a spectrum does not include the flux error, we assume the signal-to-noise ratio for that spectrum is below that of all spectra containing a flux error column when making this cut.

We classify the spectra into two categories: 1991T-like SNe Ia and normal SNe Ia based on spectral template fitting. We use the Supernova Identification tool (SNID Blondin & Tonry, 2007b) to determine the sub-type, and all objects that are found to be 1991T-like objects are further investigated through a literature search (See footnotes of Table 2) in order to properly classify objects whose photospheric phase spectra can commonly be mistaken with 1991T-likes such as 02cx-likes/Type Iax (see e.g. Phillips et al., 2022). The final selection includes five 1991T-like SNe Ia and 35 normal SNe Ia spectra. The list of objects, with their phase from maximum light, classification, and references can be found in Table 2.

3.3   Supernova Model

We present a condensed parametric ejecta model designed to fit a wide variety of predicted SNe Ia spectra corresponding to different progenitor systems. In Section 3.3.1 we introduce the hydrodynamic models upon which these parameters and their ranges are based. Section 3.3.2

SN	Phase (d)	$\lambda_{\min}$ (Å)	$\lambda_{\max}$ (Å)	Date (MJD)	Telescope	Instrument	Reference
				1991T-likes			
1991T	-9.00	3100.00	9840.00	48365.00	Lick-3m	UV-Schmidt	Filippenko et a
2001V[13]	-9.67	3720.00	7540.50	51963.33	FLWO-1.5m	FAST	Matheson et al.
2003fa[14]	-9.66	3720.00	7540.50	52797.34	FLWO-1.5m	FAST	Blondin et al. (
1999dq[??]	-9.55	3380.00	9040.00	51426.45	FLWO-1.5m	FAST	Matheson et al.
1999aa	-11.67	3440.00	7220.00	51223.33	FLWO-1.5m	FAST	Matheson et al.
				Normal			
1998dm	-11.49	3300.00	10100.00	51049.51	Lick-3m	KAST	Silverman et al.
2005ki	-8.50	3708.77	7151.80	53697.00	LCO-duPont	Mod-spec	Guillochon et a
2005mz	-7.67	3490.00	7409.02	53738.13	FLWO-1.5m	FAST	Blondin et al. (
2006X	-10.00	4134.97	6794.63	53775.00	Nayuta	MALLS	Yamanaka et a
2006ax	-8.70	3486.00	7407.96	53818.30	FLWO-1.5m	FAST	Blondin et al. (
2006cp	-9.74	3482.00	7403.96	53887.26	FLWO-1.5m	FAST	Blondin et al. (
2006gr	-7.70	3479.00	7415.66	54005.30	FLWO-1.5m	FAST	Blondin et al. (
2000dn	-7.91	3720.00	7540.50	51816.29	FLWO-1.5m	FAST	Blondin et al. (
2006lf	-7.60	3477.00	7413.66	54037.40	FLWO-1.5m	FAST	Blondin et al. (
2007af	-10.00	3182.61	5271.20	54163.00	ESO-NTT	EMMI	Guillochon et a
2007bd	-9.32	3476.00	7412.66	54197.18	FLWO-1.5m	FAST	Blondin et al. (
2007ci	-8.20	3480.00	7416.66	54238.20	FLWO-1.5m	FAST	Blondin et al. (
1998dh	-8.50	3720.00	7540.50	51021.40	FLWO-1.5m	FAST	Matheson et al.
2007qe	-8.89	3476.00	7417.07	54420.11	FLWO-1.5m	FAST	Blondin et al. (
1998aq	-7.74	3499.50	7140.00	50922.26	FLWO-1.5m	FAST	Branch et al. (2
2005cf	-8.71	3485.00	7411.37	53524.29	FLWO-1.5m	FAST	Wang et al. (20
2006le	-7.57	3476.00	7412.66	54040.43	FLWO-1.5m	FAST	Blondin et al. (
2004eo	-10.00	3741.26	9092.24	53268.00	LCO-duPont	WFCCD	Guillochon et a
2004at	-7.58	3720.00	7540.50	53084.42	FLWO-1.5m	FAST	Blondin et al. (
2000fa	-11.52	3680.00	7541.00	51881.48	FLWO-1.5m	FAST	Matheson et al.
2001ep	-7.51	3720.00	7540.50	52192.49	FLWO-1.5m	FAST	Blondin et al. (
2001gc	-8.64	3720.00	7540.50	52235.26	FLWO-1.5m	FAST	Blondin et al. (
2002bo	-7.66	3720.00	7540.50	52349.34	FLWO-1.5m	FAST	Blondin et al. (
2002cr	-11.31	3720.00	7540.50	52397.29	FLWO-1.5m	FAST	Blondin et al. (
2002cs	-8.61	3720.00	7540.50	52401.39	FLWO-1.5m	FAST	Blondin et al. (
2004ef	-8.70	3479.00	7414.19	53255.30	FLWO-1.5m	FAST	Blondin et al. (
2002dj	-7.83	3720.00	7560.00	52443.17	FLWO-1.5m	FAST	Blondin et al. (
2002er	-8.00	3500.47	9294.97	52516.00	Ekar	AFOSC	Kotak et al. (20
2002he	-8.52	3720.00	7500.00	52577.48	FLWO-1.5m	FAST	Blondin et al. (
2003W	-11.65	3200.00	8800.00	52668.35	MMT	MMT-Blue	Blondin et al. (
2003cg	-8.00	3700.00	9347.83	52721.00	CA-2.2m	CAFOS	Elias-Rosa et a
2003du	-7.76	3720.00	7540.50	52757.24	FLWO-1.5m	FAST	Blondin et al. (
2008ar	-8.71	3476.00	7418.54	54525.39	FLWO-1.5m	FAST	Blondin et al. (
2002dl	-7.55	3720.00	7540.50	52444.45	FLWO-1.5m	FAST	Blondin et al. (
2011fe	-11.00	3500.91	9498.69	55803.00	WHT-4.2m	ISIS	Parrent et al. (

Table 2: Table of selected SNe with photospheric phase spectra. The phase of the spectrum represents the time before maximum B-band magnitude that the spectrum was taken. Classification of the SNe Ia sub-types was performed with SNID for all models and further classification of those initially labeled as 91T-likes is determined through a literature search to avoid possible contamination.

introduces the way that the density structure of the ejecta is parameterized in the regime of the photospheric outer ejecta. Section 3.3.3 describes the method by which we parameterize the relative abundances according to the masses of individual elements present throughout the ejecta and how these masses are folded into a general multi-zone model for SNe Ia ejecta. Sections 3.3.4 and 3.3.5 describe the physical assumptions made when performing spectral synthesis for comparison between model parameters and observed spectra. Finally, Section 3.3.6 describes the deep-learning framework implemented to perform the acceleration of our spectral synthesis over our space of model parameters.

3.3.1 Parameterized Ejecta Model

We develop a parametric model of the ejecta of SNe Ia based on the structure of spherically averaged ejecta profiles taken from HESMA. HESMA contains a database of a wide range of simulations of a variety of proposed SNe Ia progenitor scenarios (Fink et al., 2014; Noebauer et al., 2017; Kromer et al., 2013, 2015; Sim et al., 2010; Noebauer et al., 2017; Fink et al., 2018; Marquardt et al., 2015; Fink et al., 2010b; Kromer et al., 2010; Sim et al., 2012; Gronow et al., 2020) which provide an approximation to the space of potential ejecta structures that describe SNe Ia observations at various times. A visualization of a randomly generated ejecta profile from a set of model parameters drawn from our space is presented in Figure 4. The ejecta model is parameterized by density and abundance profiles, described in the next two sections.

3.3.2 Density Profile

We adopt a velocity-dependent power-law density profile in homologous expansion to model the outer ejecta of the supernova (Equation 1). The outer ejecta of HESMA models can be well fit by power-law at early times. A power-law index, α_ρ is left as a free parameter which allows the model to cover the full range of outer-ejecta density profiles present in the HESMA models (see Section 3.4.2 for a description). A fixed reference velocity for our density

profile, $v_0 = 8000 \ \mathrm{km \, s^{-1}}$, is statically set for all models as a reference density, ρ_0, is solved to constrain the density of the model. The constructed density profile extends from v_0 to an outer boundary velocity, v_{outer}, set such that the density at the outer-boundary velocity is a fixed $\rho(v_{\mathrm{outer}}, t = t_0) = 10^{-14} \ \mathrm{g \, cm^3}$ which is the cutoff value of the density profiles present in the HESMA models at a fixed $t_0 = 2 \, \mathrm{days}$. The value of v_0 is an arbitrary choice as a reference coordinate from where we define our model, so the value was selected as the lower bound of the inner boundary velocity prior (Section 3.4.2) for simplicity.

$$\rho(v) = \rho_0 \left(\frac{t_0}{t} \right)^3 \left(\frac{v}{v_0} \right)^{\alpha_\rho} \tag{1}$$

We constrain the values for v_{outer} and ρ_0 from a given total ejecta mass above v_0, M_{tot}, and a given α_ρ by integrating Equation 1 at a time $t = t_0$ by applying the substitution $v \, t_0 = r$ from homologous expansion.

$$M_{\mathrm{tot}} = \frac{\rho_0 t_0^3 4\pi}{v_0^{\alpha_\rho}} \int_{v_0}^{v_{\mathrm{outer}}} v^{\alpha_\rho} v^2 dv \tag{2}$$

The value for M_{tot} is determined from the total of the masses of the individual elements contributing to the ejecta above v_0.

3.3.3 Abundance Profile

We model the abundances of the same elements explored by O'Brien et al. (2021) in our ejecta model as these elements account for the majority of line formation in the resulting spectrum as well as trace the general nucleosynthetic products of the supernova (see e.g. Filippenko, 1997). We parameterize these elements in terms of total masses above v_0 in order to better constrain the total ejecta mass as well as simplify the sampling procedure. Masses for Carbon (M_{C}), Oxygen (M_{O}), Magnesium (M_{Mg}), Silicon (M_{Si}), Sulfur (M_{S}), Calcium (M_{Ca}), Chromium (M_{Cr}), Titanium (M_{Ti}), stable Iron (M_{Fe}), and initial ^{56}Ni at t_0, $M_{^{56}\mathrm{Ni}}$, are aggregated into three quantities corresponding to the mass of Iron Group

Elements (IGEs, $M_{\text{IGE}} = M_{^{56}\text{Ni}} + M_{\text{Cr}} + M_{\text{Ti}} + M_{\text{Fe}}$), Intermediate Mass Elements (IMEs, $M_{\text{IME}} = M_{\text{Si}} + M_{\text{S}} + M_{\text{Mg}} + M_{\text{Ca}}$), and Unburned Elements (UBEs, $M_{\text{UBE}} = M_{\text{C}} + M_{\text{O}}$), as well as a total ejecta mass ($M_{\text{tot}} = M_{\text{IGE}} + M_{\text{IME}} + M_{\text{UBE}}$). We place these three categories of elements into three distinct regions of the ejecta corresponding to a general structure seen in the HESMA abundance profiles as well as tomography results presented by Aouad et al. (2022, Figure 18) in which IGEs resulting from complete nuclear burning are placed below a layer of IMEs resulting from incomplete burning, with UBEs placed in the outer-most regions (see Figure 4). The fractional abundance of each region is parameterized by a set of functions, $A_{UBE}(v; v_c, w)$, $A_{IME}(v; v_c, w)$, $A_{IGE}(v; v_c, w)$, where the sum of the profiles at each velocity adds up to unity. A modified Gaussian is used to represent the distribution of IMEs which is parameterized by a width, w, and a centroid, v_c, in velocity space. The form of this profile was selected to allow for the model to parameterize various amounts of mixing between regions of the ejecta as well as explore the depth at which the properties of the ejecta are changing. The model results in a mass-fraction profile that follows a Gaussian bubble of IMEs over the ejecta velocity and serves as an approximation to the profiles present in the HESMA dataset.

$$A_{\text{IME}}(v; v_c, w) = A_0 v^{-(\alpha_\rho + 2)} \exp\left[-\frac{1}{2} \frac{(v - v_c)^2}{w^2} \right] \tag{3}$$

Where A_0 is a normalization constant set to the inverse of the maximum value of $A_{IME}(v = v_{\max}; v_c, w)$. The velocity corresponding to the distribution's maximum value is determined from v_c and w through the relation

$$v_c = \frac{w^2(\alpha_\rho + 2) + v_{\max}}{v_{\max}}. \tag{4}$$

The values for $v_{\max}$ and w are then determined from the relative masses of each region of

elements by numerically solving the following system of equations

$$M_{\text{IME}} = \frac{\rho_0 t_0^3 4\pi}{v_0^{\alpha_\rho}} \int_{v_0}^{v_{\text{outer}}} v^{\alpha_\rho} A_{\text{IME}}(v) v^2 dv, \tag{5}$$

$$M_{\text{IGE}} = \frac{\rho_0 t_0^3 4\pi}{v_0^{\alpha_\rho}} \int_{v_0}^{v_{\text{max}}} v^{\alpha_\rho} \left[1 - A_{\text{IME}}(v)\right] v^2 dv, \tag{6}$$

$$M_{\text{UBE}} = \frac{\rho_0 t_0^3 4\pi}{v_0^{\alpha_\rho}} \int_{v_{\text{max}}}^{v_{\text{outer}}} v^{\alpha_\rho} \left[1 - A_{\text{IME}}(v)\right] v^2 dv \tag{7}$$

which results in a complete ejecta profile.

3.3.4 Explosion Model

Our analysis of early-phase spectra relies on the photospheric approximation in which the rapidly increasing optical depth of the ejecta towards the center is approximated as a hard inner boundary in velocity space, v_{inner}. Thermalized radiation is injected into the ejecta above from a black-body distribution at a given temperature, T_{inner}. A parameter representing the time since the explosion, t_{exp}, scales the density profile (Equation 1) as well as sets the abundances of decay products of ^{56}Ni in the final ejecta profile.

3.3.5 Radiative Transfer

We calculate synthetic spectra from our ejecta model using the open-source Monte Carlo radiative transfer code TARDIS (Kerzendorf & Sim, 2014; Kerzendorf et al., 2021). TARDIS is a 1D steady-state code that iteratively solves for the excitation and ionization state of the plasma. TARDIS uses an inner boundary photosphere approximation that injects radiative packets into a homologously expanding ejecta.

In this work, we use TARDIS version `022.5.9.dev5+gf27fa30` together with atomic data being produced by the TARDIS sub-package CARSUS (Pássaro et al., 2019) version `0.1.dev677+gd623c94`. The generated atomic data takes ionization energies from CMFGEN (Hillier & Lanz, 2001) for O I, O II, S I, S II, Si I, and Si II. Ionization energies for other
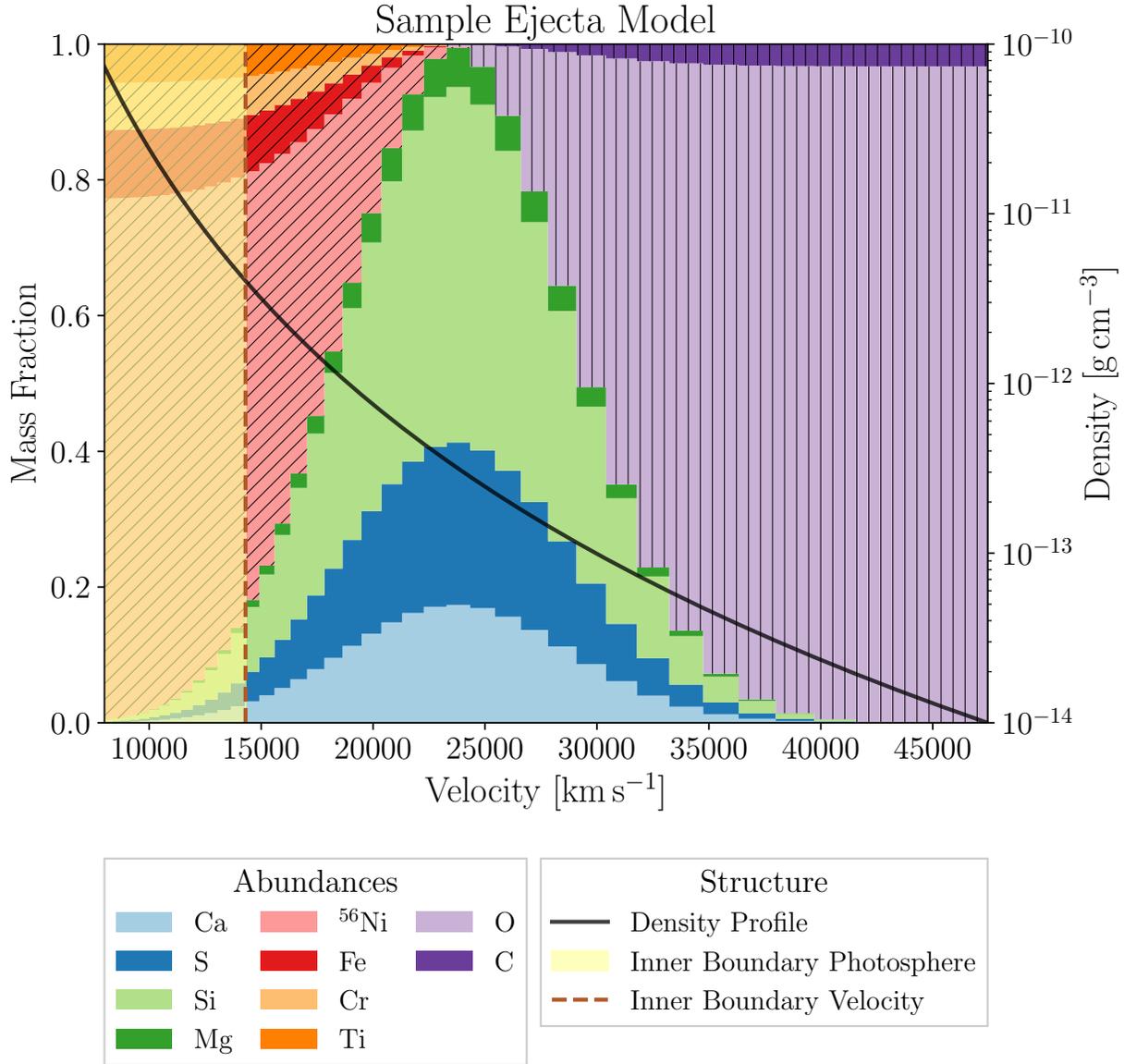
Figure 4: Visualization of a random realization of an abundance profile produced from our model in our prior space. Elemental abundances are presented as stacked histograms. The diagonally hatched regions correspond to the inner iron-group elements, the central unhatched region corresponds to the abundance of IMEs, and the vertically hatched region corresponds to the unburned elements in the outer layers of the ejecta. A red dashed vertical line represents the inner boundary velocity from which thermalized radiative packets are injected into the ejecta above. The solid black line represents the density of the ejecta through velocity space and the value of the density if provided by the right-hand axis.

species used in this work were taken from NIST (Ralchenko, 2005) with lines and level data taken from Kurucz GFALL (Kurucz & Bell, 1995).

Ionization populations are solved using the "nebular" approximation (Equation 3 in Kerzendorf & Sim, 2014) and excitation populations are solved using the "dilute-lte" prescription (Equation 5 in Kerzendorf & Sim 2014; Equation 4 in Lucy 1999b). The "nebular" approximation assumes that the expanding envelope is optically thin in all ionization continua (Mazzali & Lucy, 1993) which acts as a good approximation for radiative-transfer in the photospheric phase. The "dilute-lte" prescription acts as an approximation for NLTE excitation levels again in the optically thin limit. Line interactions are handled using a macro-atom model (Lucy, 2002). Models were generated using 40 shells of ejecta and run until plasma state convergence with 10^5 packets per Monte Carlo iteration. Further configuration information for TARDIS including links to a reproducible setup and the atomic data file created with CARSUS can be found in Appendix C.

3.3.6 Emulator

Spectral synthesis with TARDIS is too computationally expensive to be used directly for fitting. For example, a single TARDIS simulation takes approximately 30 CPU minutes, which would require hundreds of years to effectively sample a posterior distribution which requires over a million sequential simulation runs.

In recent years emulation of radiative transfer models has served as a powerful tool for directly probing the properties of a variety of supernovae and other astrophysical objects (see e.g. Vogl et al., 2020b; O'Brien et al., 2021; Fullard et al., 2022). To expedite model evaluation we employ an emulator for TARDIS which performs spectral synthesis from model input parameters through an analytic approximation. O'Brien et al. (2021) applied a deep-learning emulator for TARDIS based on Kerzendorf et al. (2021) to simulate single-zone ejecta models for normal SNe Ia which, for the first time, allowed for fully-probabilistic reconstructions of the outer ejecta of a SN Ia. Kerzendorf et al. (2022) expanded upon

the utility of such emulators by incorporating a probabilistic deep-learning architecture for emulated SNe Ia spectral synthesis which includes the added functionality of providing uncertainties in the emulated spectra.

We combine our ejecta model with a probabilistic emulator architecture based upon that of Kerzendorf et al. (2022) to rapidly generate synthetic spectra from our model's parameters with improved uncertainty estimates. Our emulator serves as a surrogate for evaluation of our parametric ejecta model with the TARDIS radiative transfer code and is only subject to the constraints of the model from the training samples. We train a deep ensemble (Lakshminarayanan et al., 2017) of 12 probabilistic emulators to emulate our spectral synthesis. Model evaluation is performed by aggregating the resulting spectra from each emulator with their associated uncertainty. Scripts and data files containing the emulator and its training data can be found in Appendix C.

3.4   Model Inference

We perform Bayesian inference in order to find the posterior distribution of model parameters given our observed spectra. In order to model the posterior distribution we require a method of likelihood estimation, presented in Section 3.4.1, to effectively compare simulated spectra to observed spectra in the context of physical and systematic uncertainties and biases. The constraints we place on the parameters of our model are discussed in Section 3.4.2 and the method of sampling the posterior distribution is discussed in Section 3.4.3. A short discussion of our method of lowering the emulation uncertainty for regions of parameter space that are both high in likelihood and under-sampled in our emulator's training data is presented in Section 3.4.5.

3.4.1   Likelihood Estimation

We apply an extended form of the likelihood function used by O'Brien et al. (2021) to incorporate emulator uncertainties determined by the probabilistic DALEK emulator by

adding them in quadrature to the other sources of uncertainty. We aim to best reconstruct the composition of the ejecta, so we remove the continuum when determining the quality of a fit in order to maximize contributions from line formation. We incorporate a spectral continuum removal process, $C(\hat{F}_\lambda(\vec{\theta}))$ which normalizes the synthetic spectrum estimate, $\hat{F}_\lambda(\vec{\theta})$ to the continuum of the observed spectrum, F_λ. This continuum removal process fits a 3rd order polynomial to the ratio between the observed spectrum and the simulated spectrum then multiplies the simulated spectrum by the polynomial. Such removal is necessary to remove the effects of the continuum, distance, and reddening from the observation to ensure our fits are driven by the line features. The total form of the log-likelihood is

$$\log \mathcal{L}(\vec{\theta}) = -\frac{1}{2} \sum_\lambda \left[\left(\frac{C(\hat{F}_\lambda(\vec{\theta})) - F_\lambda}{\sigma_\lambda(\vec{\theta})} \right)^2 + \log \left(2\pi\sigma_\lambda^2(\vec{\theta}) \right) \right],$$

where

$$\sigma_\lambda^2(\vec{\theta}) = \sigma_{\mathrm{obs},\lambda}^2 + f_\sigma^2 C^2(\hat{F}_\lambda(\vec{\theta})) + \sigma_{\mathrm{emu},\lambda}^2(\vec{\theta})$$

where f_σ represents an inferred fractional uncertainty (Hogg et al., 2010) over our spectrum and $\sigma_{\mathrm{obs},\lambda}$ is the observational uncertainty of the spectrum we are fitting. Observational uncertainties are taken from the spectra data source if available, otherwise, a constant uncertainty of 1% of the mean of the spectrum is assumed. $\sigma_{\mathrm{emu},\lambda}$ is the estimate of the emulator's uncertainty (Equation 4 in Kerzendorf et al., 2022) in the region corresponding to the fit.

3.4.2   Prior Bounds

Table 3 lists our prior distributions of model parameters. Multiple constraints are placed on the prior distribution of model parameters in order to accurately reflect the limits of currently explored hydrodynamic simulations of progenitor scenarios for SNe Ia. A large variety of hydrodynamical simulations of various SNe Ia progenitor systems are found in the HESMA models and offer information about the expected general properties of the ejecta structure

Distribution	Model Parameter	Distribution Parameters	
Uniform		Low	High
	T_{inner} (K)[15]	8000	15000
	v_{inner} (km s^{-1})[??]	8000	16000
	α_ρ	-10	-5
Normal		μ[16]	σ[17]
	t_{exp} (days)	19.5 + Phase[18]	2.5
Multivariate Log-normal		μ[19]	σ[20]
	M_{Si} (M$_\odot$)	7.84×10^{-2}	0.93
	M_{Ca} (M$_\odot$)	1.10×10^{-2}	1.08
	M_{S} (M$_\odot$)	3.94×10^{-2}	0.89
	M_{Mg} (M$_\odot$)	1.19×10^{-2}	1.02
	M_{Ni56} (M$_\odot$)	1.11×10^{-1}	1.49
	M_{Cr} (M$_\odot$)	3.17×10^{-3}	1.47
	M_{Ti} (M$_\odot$)	1.48×10^{-3}	1.87
	M_{Fe} (M$_\odot$)	2.04×10^{-2}	1.39
	M_{O} (M$_\odot$)	7.12×10^{-2}	1.34
	M_{C} (M$_\odot$)	2.59×10^{-2}	0.87

Table 3: The prior distributions from which our model parameters are sampled during posterior inference. Parameters are sampled over different distributions according to their range of physical applicability determined from hydrodynamical models in the HESMA data set.

such as the relative typical ratios of nucleosynthetic products present within the ejecta as well as full density profiles. We generate a prior space for total elemental masses by integrating models taken from HESMA above v_0 so that the final masses of each element follow the same general correlation structure as the sum of all hydrodynamic models, ensuring a reasonable estimate of the distribution of likely supernovae ejecta profiles. The prior distribution of elemental masses is drawn from a multivariate Gaussian distribution whose covariance is set as the covariance of the log of elemental masses taken from the HESMA models with a centroid taken as the log of the mean of HESMA masses in linear space as to not bias the distribution towards models with little or no mass of certain elements. Drawing from this distribution offers a good balance between tracing the general covariance structure of the models found in the HESMA while also permitting nearly any parameter combination to be tested, albeit with a smaller probability. Since these distributions are based upon the distributions of elemental masses present in the HESMA models, the total mass of the

outer-ejecta is also constrained by this distribution.

We set a uniform prior on the distribution of values of α_ρ by fitting linear models to the HESMA density profiles above v_0 and taking the minimum and maximum value to the nearest integer. Velocity and temperature distributions are initially sampled uniformly over the ranges specified in Table 3, with cuts placed on the luminosity of the supernovae under homologous expansion with an assumed rise time of 19.5 days (Riess et al., 1999) according to the Stephan-Boltzmann law as an estimate for the range of realistic maximum light luminosities. The prior distribution for the time since the explosion, t_{exp}, is determined on a spectrum-by-spectrum basis. The distribution is always represented by a Gaussian distribution centered at a time of 19.5 days plus the phase of the spectrum from maximum light (see Table 2) with a standard deviation of 2.5 days to account for rise-time uncertainty based on the spread of rise-times between normal and 1991T-like SNe Ia (see Figure 6 in Ganeshalingam et al., 2011).

3.4.3 Sampling the posterior

3.4.4 UltraNest

The posterior inference was performed with nested sampling (Skilling, 2004; Buchner, 2021) with the MLFriends Monte Carlo algorithm (Buchner, 2014, 2017). Nested sampling is ideal for generating posterior samples from complex high-dimensional distributions. We used the nested sampling package ULTRANEST[21] (Buchner, 2021) to sample the posterior distribution for each observed spectrum. Each spectrum returned between 10 000 and 30 000 effective posterior samples which are presented in Figures 6 and 7.

3.4.5 Active Learning

The high dimensionality of the parameter space and unknown apriori parameter constraints required to effectively model individual spectra observations create difficulty in selecting

---

[21]https://johannesbuchner.github.io/UltraNest/

an optimal training set for our emulator. We resolve this issue by iteratively selecting new training points that are predicted to best improve emulator accuracy in the regions of the parameter space that are most likely to model the spectra we are attempting to model.

We apply *Active Learning* (AL Cohn et al., 1996; Beluch et al., 2018) iterations to the emulator training to improve accuracy in regions of high importance. After an initial draw of 250 000 random samples, the emulator is trained to reproduce the results of TARDIS (see Section 3.3.6). We sample the posterior distribution, using this emulator, of parameters best matching our observed spectra using a modified AL likelihood function, $\mathcal{L}_{\mathrm{AL}}(\vec{\theta})$. This likelihood function weighs the likelihood of a proposed $\vec{\theta}$ by the relative fraction of emulator uncertainty to total uncertainty, encouraging exploration into regions of the parameter space where the emulator has less information. The AL likelihood function is computed as

$$\log \mathcal{L}_{\mathrm{AL}}(\vec{\theta}) = \log \mathcal{L}(\vec{\theta}) + \frac{1}{2} \sum_{\lambda} \log \frac{\sigma^2_{\mathrm{emu},\lambda}(\vec{\theta})}{\sigma^2_{\lambda}(\vec{\theta})}$$

An equal number of posterior samples are selected for each observed spectrum and are evaluated by TARDIS. Synthetic TARDIS spectra are then appended to the original training data to provide the emulator with more information around areas that are simultaneously high in likelihood while also high in emulation uncertainty. Each acquisition process yields approximately 200 000 additional samples per iteration. Two iterations of active learning were performed on the data.

In regards to the final emulator's performance in predicting TARDIS spectra under this model, we have performed an analysis across a hold-out set, or "test set", of model-spectra pairs that were not included in the training or validation set of the emulator. The probabilistic DALEK emulator applied in this paper has some key differences from the original DALEK emulator, namely that the probabilistic DALEK emulator produces a distribution of spectra for a given input parameter vector (represented by a mean and standard deviation) rather than a single prediction point. In evaluating the performance of the mean of the prediction,

which is the closest comparison we can directly make to the original DALEK emulator, we find that the Mean Fractional Error averages approximately 1% which is below the total aggregate uncertainty used for inference in our likelihood function. We also measured a z-score over the test set at each wavelength as the difference between the mean prediction and the true spectrum divided by the predicted standard deviation and evaluated the fraction of data-points that fall within 1, 2, and 3 standard deviations. We find that at each interval, the fraction of data corresponds to the 1, 2, and 3-sigma tails of a normal distribution or better, with 88.5% of predicted means falling within 1 predicted standard deviation of the true spectrum and 99.7% of predictions falling within 3-standard deviations of the true spectrum.

3.5   Results

The posterior probability distribution of spectra for the five 1991T-like SNe Ia in our samples along with their maximum likelihood estimate and total uncertainty is presented in Figure 5. For comparison, a selection of five of the normal SNe Ia from our sampled are shown in Figure 5 as well. Our fits accurately reproduce major line features that distinguish 1991T-like SNe Ia from the normal SNe Ia population. Specifically, our models generate the high-velocity Fe III features around $4250\,\text{Å}$ and $4950\,\text{Å}$ as well as the Si II feature near $6150\,\text{Å}$.

While our emulator may appear more effective at modeling the general SNe Ia population than the 1991T-like population, most of the deviations from the means of the posteriors are within 68% uncertainty. This results in apparent inconsistencies with the 1991T-like SNe Ia in 2 ways. First, in the UV where there is heavy blanketing from IGEs, there is a large Monte-Carlo uncertainty produced by TARDIS due to the radiation being sampled from a black-body which sharply drops off in the blue. As can be seen in Figure 5, while the means of the posterior spectra do not always line up with the observations, there are very large uncertainties and these uncertainties still encompass the data. Second, on the red end, since the likelihood is evaluated over the entire spectrum shallow silicon features will contribute less to the weight of the overall likelihood and, due to Monte-Carlo uncertainty, shallower
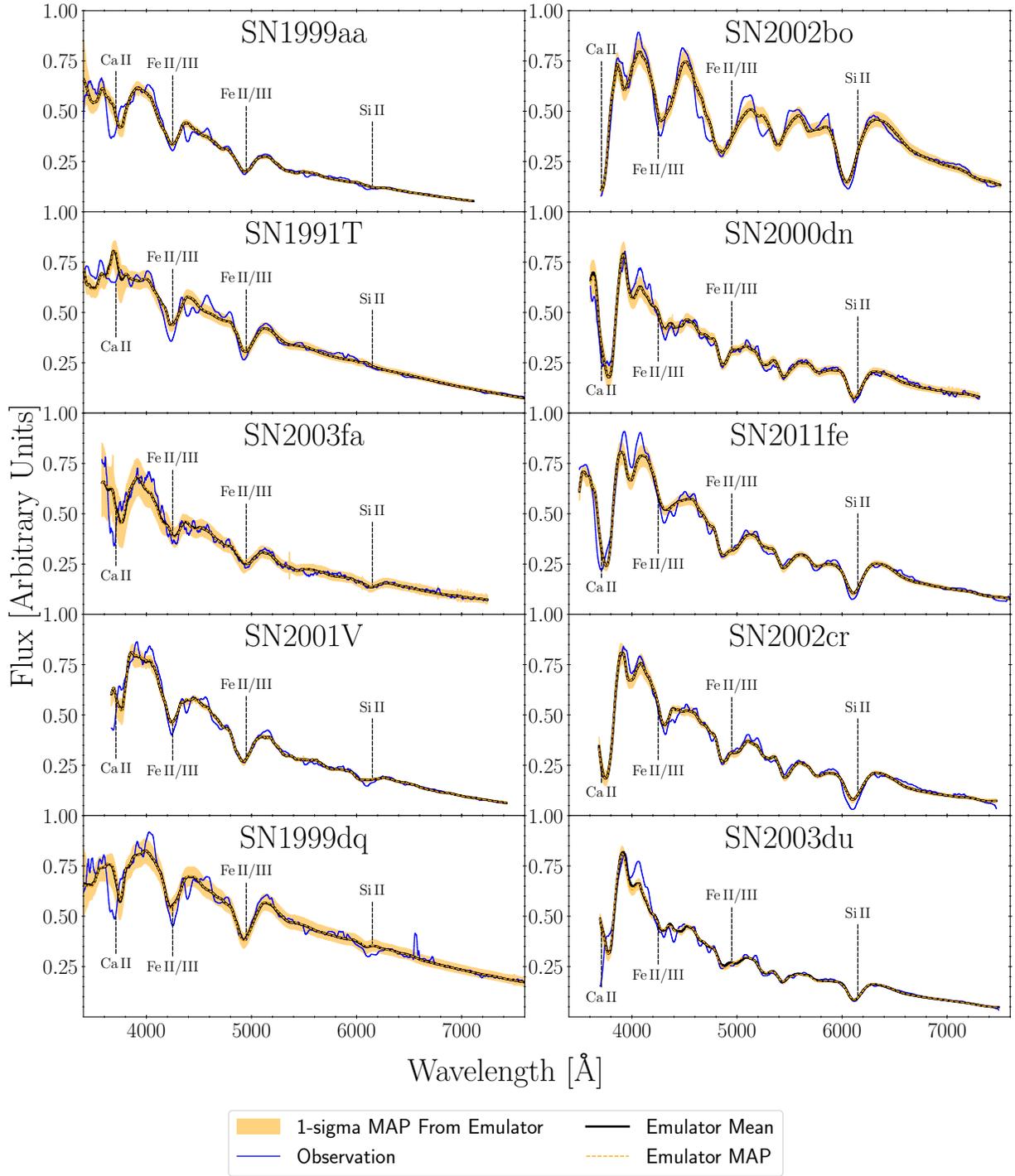
Figure 5: Posterior spectra of 1991T-like (left) and normal (right) SNe Ia scaled and offset for visualization. The mean of the posterior is represented in black with the best fit (maximum a-posteriori sample) in orange dashed and the shaded orange region representing the total uncertainty of the best-fit sample at 1-σ.

silicon features will be closer to the amplitude of the noise. As can be seen in Figure 5, our 68% posterior still encompasses the shallow silicon features, but the posterior mean alone may be misleading.

3.5.1 Ejecta Properties

The peculiar nature of early-time 1991T-like spectra has been well identified, but their origin remains unclear. 1991T-like spectra show the presence of high-velocity Fe III emission and lack the strong characteristic Si II and Ca H&K absorption commonly seen in Branch-normal SNe Ia (see e.g. Filippenko, 1997). After maximum light, 1991T-like spectra begin to behave similarly to normal Type Ia spectra, with Si II features reappearing in the spectra (see e.g. Taubenberger, 2017). There have been two suggested causes behind the lack of singly-ionized IME absorption at early times. Namely, a lack of total IME production and higher ionization states of IMEs produced in the ejecta (e.g. Jeffery et al., 1992; Ruiz-Lapuente et al., 1992; Sasdelli et al., 2014).

We find a variety of parameters that indicate the differences between 1991T-like and Normal SNe Ia. The distribution of inner boundary temperatures for 1991T-like SNe Ia are higher than the average inner boundary temperatures for the normal SNe Ias, though still exists within the range of nearly half of the normal SNe Ias in our sample (Figure 6) indicating that high-ionization states of IMEs, in particular Silicon, are not due to a difference in temperature of the ejecta alone, but a combination of factors. This leads us then to investigate two other possible causes for the lack of Si II formation in the photospheric phase: a decrease in the electron density at the primary location of IME composition or a decrease in the total mass of IMEs contributing to the line features seen in the ejecta.

The material below the photosphere, parameterized through the inner boundary velocity, does not contribute to features in the resultant spectra. Therefore, constraints of physical properties of the ejecta must rely strictly upon material above the inner boundary photosphere. Since the selected range of phases for these spectral encompass a wide range in photospheric
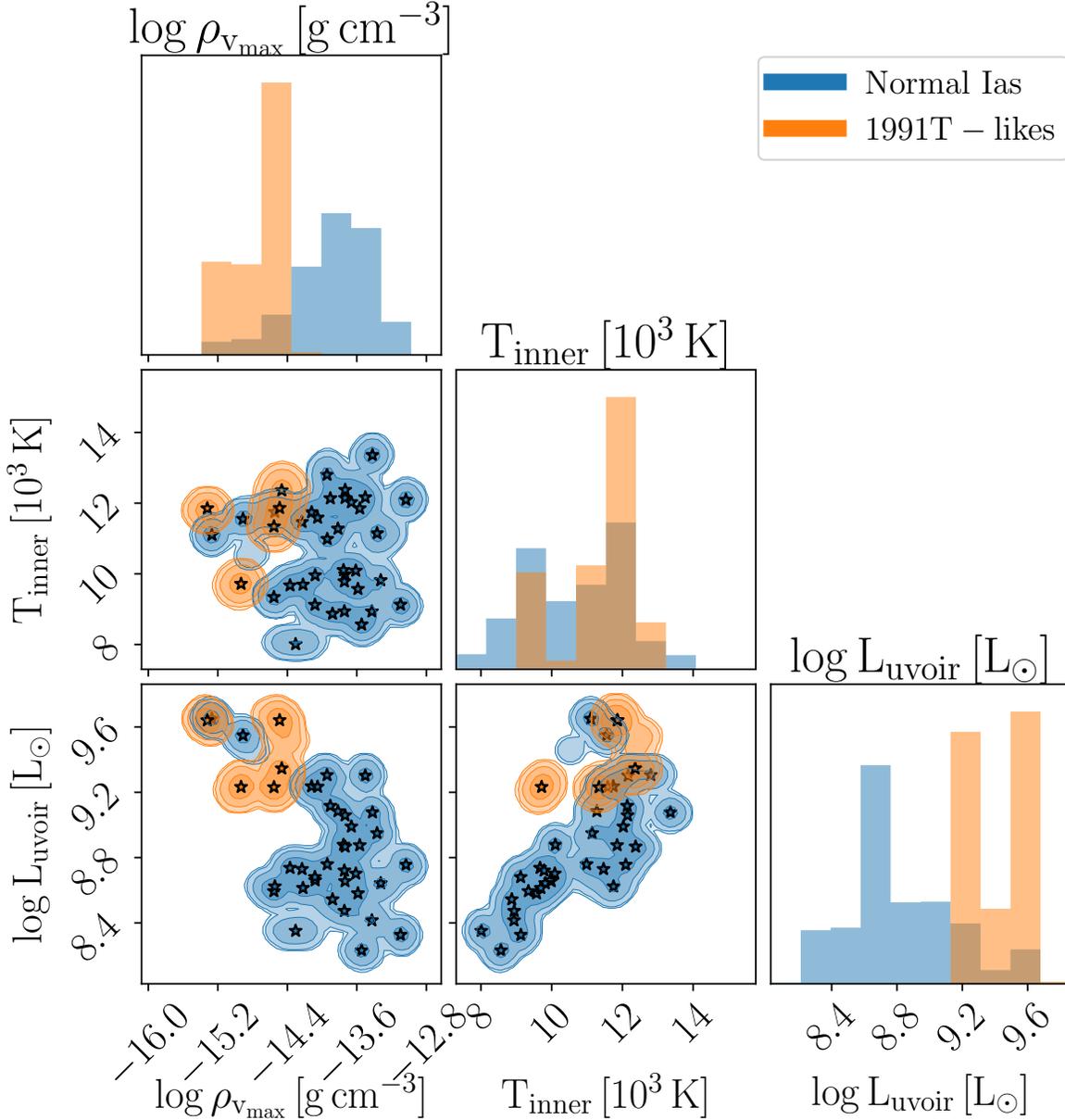
Figure 6: Posterior contours of SNe Ia probed in this study. Normal SNe Ia are shown with blue contours and 1991T-like SNe Ia are shown in orange. The contours cover 68%, 95%, and 99.5% quantiles. The posterior means for each object are shown as stars. The plots show the joint distributions between ejecta density at $v = v_{\mathrm{max}}$, inner boundary temperature T_{inner}, and integrated UVOIR luminosity from the model spectrum. While 1991T-like SNe Ia are generally brighter than the Normal SNe Ia population, the increase in brightness does not seem to be driven by substantially higher photospheric temperatures. The lower ejecta density in the region of highest intermediate mass element abundance shows that higher ionization fractions in 1991T-like SNe Ia are influenced by the lower electron density.

velocities, such analysis must be viewed in the context of comparison between the bulk-populations of each supernova type as the two populations have similar mean photospheric velocities over all of the posterior samples. We determine the total mass of each contributing element above the photosphere by integrating Equations 5, 6, and 7 with their lower bounds set to the inner boundary velocity, v_{inner}. We compute the mass fraction of each element as the integrated mass of each element above the photosphere divided by the total mass above the photosphere. The mass fraction offers a direct probe of the nucleosynthetic products that are visible in the photospheric phase and which can be directly compared to hydrodynamic models without a need to convert abundance fractions into total masses.

Figure 7 shows the posterior probability distributions of the IME fractions from 1991T-like SNe Ia demonstrating a clear deficit compared to that of normal SNe Ia coupled with a small increase of IGEs as a fraction of the total ejecta. The marginal distribution of the fraction of unburned elements does not demonstrate a discernible difference between 1991T-likes and Normal SNe Ia, though the joint distribution between IGEs and unburned elements shows an interesting correlation in 1991T-likes in which the fraction of unburned elements in the ejecta is slightly higher for 1991T-like SNe Ia compared to Normal SNe Ia given the same iron-group element fraction. The consistent lack of IME mass fractions changing with respect to UBE fractions along with the correlation between UBE and IME fractions in 1991T-like SNe Ia implies a rapid and consistent drop-off in the rate of production of nucleosynthetic products with respect to depth into the explosion.

While many 1991T-like SNe Ia show generally lower mass fractions of IMEs compared to the normal Ia population, there are cases of overlap (see Figures 6, 7) where low mass fractions alone are not enough to explain the observed lack of IME features, such as the Si II 6150 Å doublet, in the resulting spectra. Additionally, we note that the 1991T-like SNe Ia population has generally lower ejecta densities at the location of the peak of the fractional abundance of IMEs in our model implying a lower electron density and therefore a higher ionization state. The combination of low IME mass fraction and higher ionization states
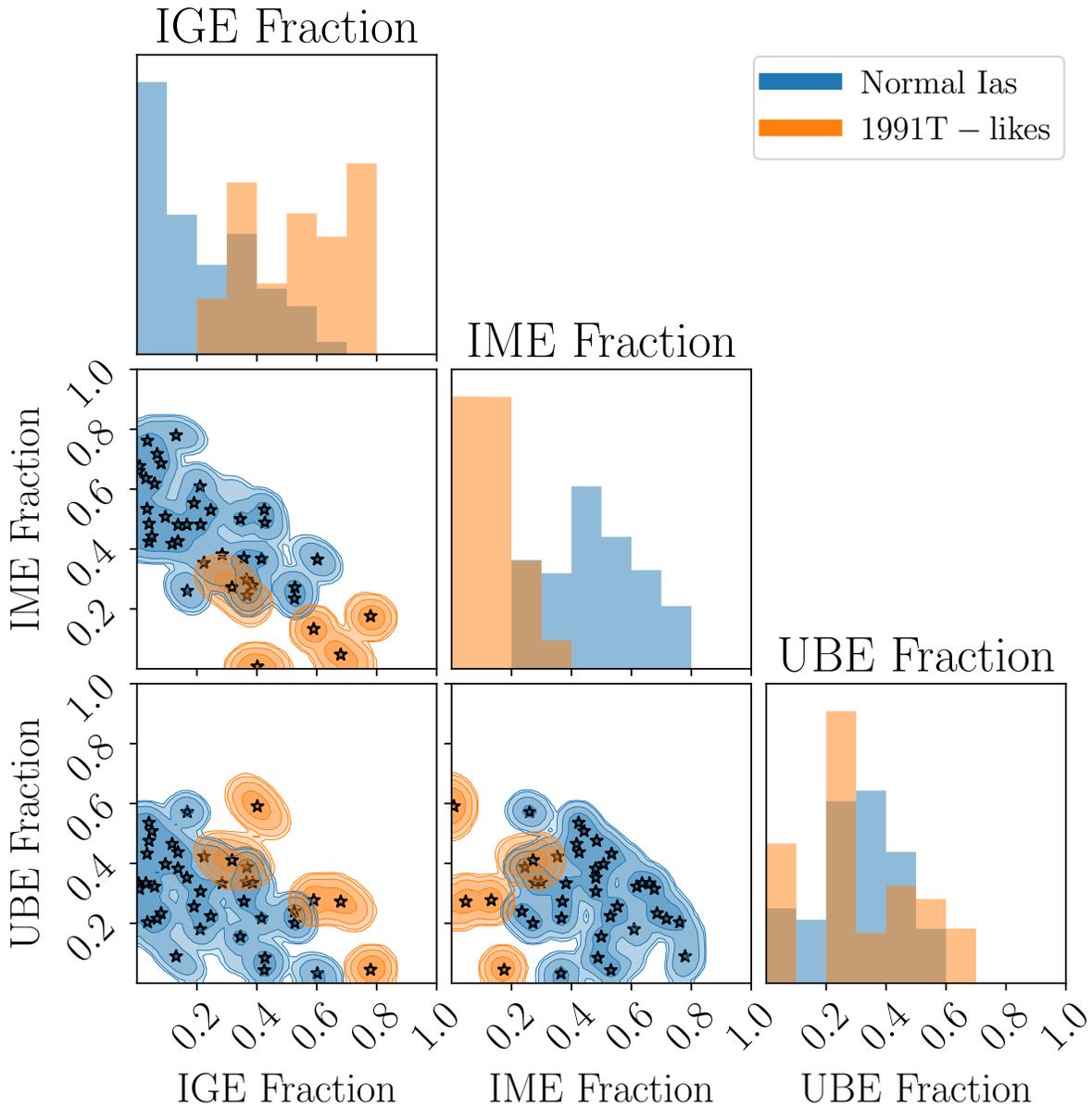
Figure 7: Posterior contours of SNe Ia probed in this study. Normal SNe Ia are shown with blue contours and 1991T-like SNe Ia are shown in orange. The contours cover 68%, 95%, and 99.5% quantiles. The posterior means for each object are shown as stars. Each plot shows the marginal distribution of mass fractions of the various ejecta compositions above the inner boundary velocity by integrating equations 6, 5, and 7. It can clearly be seen that 1991T-like SNe Ia lie on the edge of IME mass fraction distribution describing normal SNe Ia.

leads to a dual effect where the observed properties of 1991T-like SNe Ia in comparison to the normal Ia population is not due to a single underlying mechanism, but a combination of different physical processes which result in similar looking spectra observationally.

We selected the maximum likelihood sample for each spectrum and simulated the ejecta radiation field using TARDIS. The ionization fraction of Si III to Si II was determined at the shell containing the velocity v_{max} providing a look into the ionization state of the plasma at the location with the maximum abundance of intermediate-mass elements. 1991T-like supernovae were found to have overall higher ionization fractions than the vast majority of normal SNe Ia, though some overlap was found within the normal SNe Ia population (Figure 8). The normal SNe Ia with comparable ionization fractions to 1991T-like SNe Ia all had a higher mass fraction of IMEs than 1991T-like SNe Ia at the same ionization fraction. Additionally, 1991T-like SNe Ia with lower ionization fractions among the 1991T-like population also had a lower mass fraction of intermediate-mass elements within their ejecta. The suppressed Si II absorption features observed in 1991T-like SNe Ia, therefore, is a result of a combination of low IME fraction and high ionization state, creating a region of space where there is a turnover in the classification between the spectral types.

3.6   Conclusion

We have performed single-epoch outer ejecta reconstructions for 35 normal SNe Ia and five 1991T-like SNe Ia. Elemental abundance distributions and their ionization fractions in the early phases of the explosion have provided a picture linking the internal properties of the outer ejecta to the observational properties of their spectra. Comparison between the ejecta properties between the two populations provides insight into the relationship between normal SNe Ia and 1991T-like SNe Ia.

We find 1991T-like SNe Ia both under-produce IMEs relative to the normal SNe Ia population and these IMEs exist in higher ionization states than the IMEs in the normal SNe Ia population. The cause of the higher ionization fractions is primarily driven by a lower
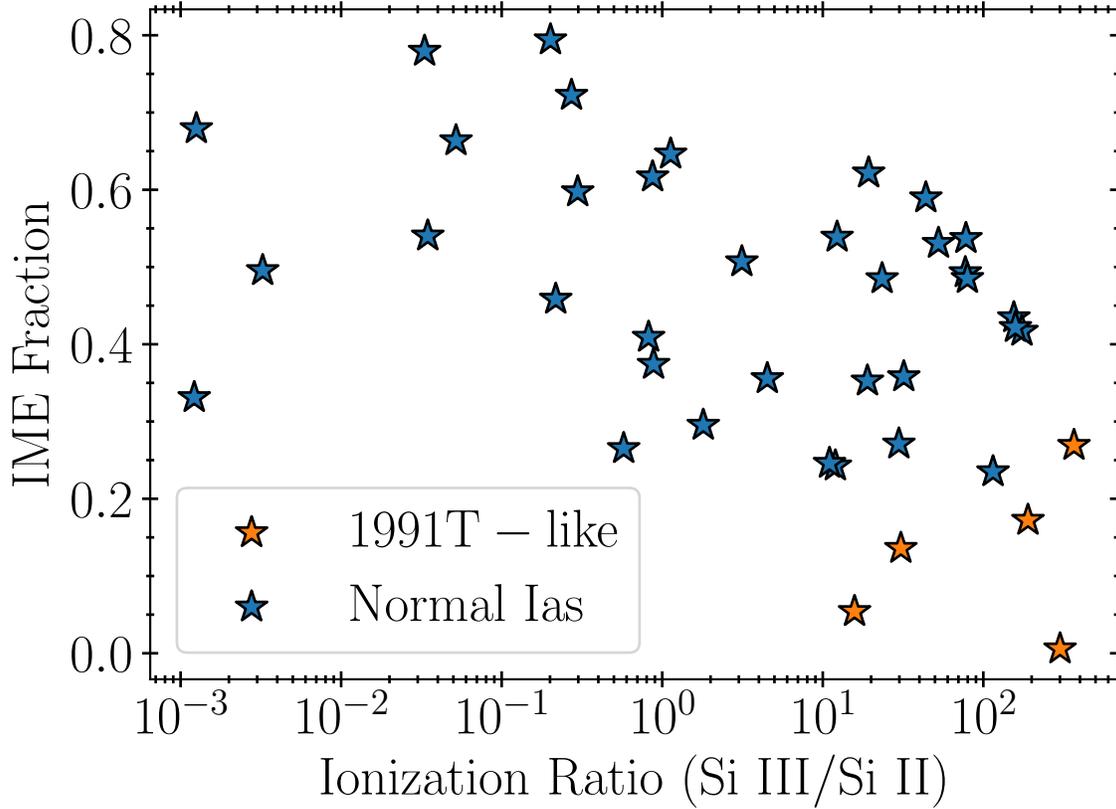
Figure 8: Ratio of Si III to Si II ion density at the shell encompassing v_{max} versus the fraction of intermediate-mass elements in the ejecta above the inner boundary. Samples are taken from the maximum likelihood parameters of each SNe Ia in our sample run through TARDIS to calculate the properties of the radiation field. Normal SNe Ia are shown in blue and 1991T-like SNe Ia are shown in orange. A clear turnover between normal SNe Ia and 1991T-like SNe Ia is evident in the regions of low intermediate mass element fraction and high ionization state of silicon. The combination of the lack of material coupled with high ionization states creates a boundary between the spectral types.

overall electron density in the ejecta. The lower overall electron density may be a result of a relative overabundance of IGEs relative to the abundance of IMEs in the ejecta of 1991T-like SNe Ia resulting in an ejecta composition dominated by high-neutron number elements, while normal SNe Ia with depleted IMEs may have the remainder of the ejecta filled with unburned Carbon and Oxygen.

Neither the low abundance fraction of IMEs nor the high ionization states of IMEs alone are enough to explain the peculiar properties of 1991T-like SNe Ia; instead, a combination of the two effects drives their unique spectral signatures at early times. We have found 1991T-like SNe Ia that contain a similar IME fraction to some of the normal SNe Ia in our sample, but these 1991T-like SNe Ia have a higher overall IME ionization than a normal SNe Ia at a similar IME mass fraction. Conversely, we have found 1991T-like SNe Ia with similar IME ionization fractions to the normal SNe Ia but these objects have a lower mass fraction of IMEs than the normal SNe Ia given their ionization state.

Our findings are consistent with normal SNe Ia and 1991T-like SNe Ia arising from a similar population or progenitor system. The observational spectral properties that traditionally separate the two groups result from a sharp change in the amplitude of spectral features corresponding to IMEs over small changes in both composition and ionization state. This results in small deviations in ejecta composition leading to a sharp contrast in observed spectral features. More detailed ejecta models will be required to definitively distinguish or unify the progenitor channels responsible normal and 1991T-like SNe Ia.

3.7   Acknowledgments

synthesis in supernovae (Kerzendorf & Sim, 2014; Kerzendorf et al., 2021). The development of TARDIS received support from the Google Summer of Code initiative, from ESA's Summer of Code in Space program, and from NumFOCUS's Small Development Grant. TARDIS makes extensive use of Astropy[22] (Astropy Collaboration et al., 2013, 2018)

Matplotlib[23](Hunter, 2007), Numba[24] (Lam et al., 2015), NumPy[25] (Harris et al., 2020), pandas[26] (Wes McKinney, 2010), scikit-learn[27] (Pedregosa et al., 2011), SciPy[28] (Virtanen et al., 2020), TARDIS[29] (Kerzendorf & Sim, 2014),Pytorch[30] (Paszke et al., 2019), and UltraNest[31] (Buchner, 2014, 2019)

---

[22]https://www.astropy.org
[23]https://matplotlib.org
[24]https://numba.pydata.org
[25]https://numpy.org
[26]https://pandas.pydata.org
[27]https://scikit-learn.org
[28]https://www.scipy.org/
[29]https://tardis-sn.github.io/tardis
[30]https://pytorch.org/
[31]https://johannesbuchner.github.io/UltraNest

4 Constraining Type Ia Supernova Progenitors with Generative Models: Insights from SN 2011fe Spectral Time Series

ABSTRACT

The progenitors and explosion mechanisms driving Type Ia supernovae (SNe Ia) continue to remain elusive. A variety of progenitor channels have been proposed and simulated leading to their own unique observational predictions but none have yet been successful in describing all properties of observed SNe Ia. We aim to provide a probabilistic ranking of proposed progenitor channels to describe the SN Ia 2011fe from a set of photospheric phase spectral observations based on simulated progenitor channels already present in the literature. We train a conditional variational auto encoder (cVAE) on a set of 1D ejecta profiles produced by hydrodynamic simulations of SNe Ia explosions taken from the Heidelberg Supernova Model Archive conditioned on their progenitor channels to create a low-dimensional embedding of SNe Ia ejecta profiles. This embedding is mapped to synthetic spectral observations through a deep-learning emulator trained on the radiative transfer code TARDIS. The spectral time series of SN 2011fe is fit with the combined cVAE-Emulator model conditioned on each progenitor channel to determine the posterior distribution of model parameters which are integrated to produce a Bayesian evidence for each progenitor channel. We find that the early-time spectral time series of SN 2011fe is best described by a pure-detonation explosion of a sub-Chandrasekhar mass carbon-oxygen white-dwarf. We compare these results to previous abundance tomography studies of SN 2011fe discuss the potential mechanisms that can lead to such an explosion.

4.1 Introduction

Despite Type Ia supernovae (SNe Ia) being critical to our understanding of cosmic expansion (Branch, 1992; Phillips, 1993; Riess et al., 1998; Perlmutter et al., 1999) and chemical

evolution of the cosmos (Kobayashi et al., 2020, see Figure 39), the mechanism(s) leading to their cataclysmic explosions remain elusive. SNe Ia are generally agreed to result from the thermonuclear runaway of a carbon-oxygen (C/O) white dwarf (WD) in a binary system as a result of mass transfer from a binary companion. An ever-growing variety of mechanisms have been proposed to explain their observational properties such as delayed-detonations of near Chandrasekhar-mass (M_{Ch}) WD (Khokhlov, 1991), a surface helium detonation compressing the core of a sub-M_{Ch} WD (e.g. Woosley & Weaver, 1994; Fink et al., 2010a; Shen et al., 2018; Polin et al., 2019; Pakmor et al., 2022), or mergers between a pair of WDs (e.g. Nomoto, 1982; Webbink, 1984; Iben & Tutukov, 1984; van Kerkwijk et al., 2010; Livio & Riess, 2003; Kashi & Soker, 2011).

Direct determination of the progenitor system for SNe Ia is a computationally difficult problem. Hydrodynamic simulations produce predictions for nucleosynthetic byproducts, energy deposition, and ejecta densities which are used to generate synthetic observables such as spectral time-series and light-curves which can be compared to data. Matching the resulting ejecta profiles of the simulations to synthetic observables directly requires running massive numbers of hydrodynamic simulations with varying parameters. While full hydrodynamic simulations provide physically realistic parameterizations, running single models is extremely costly and time-intensive making them impractical for direct inference.

An alternative approach is to perform inference on the ejecta profile directly as radiative transfer simulations are less computationally expensive than full hydrodynamics. Studies have shown that the full diversity of Type Ia spectra can be described by only a few parameters (see e.g. Saunders et al., 2018; Boone et al., 2021; Murakami et al., 2023) and typically fitting SNe Ia ejecta profiles involves application of reasonable approximations in order to reduce the size of the space of potential ejecta profiles (see e.g. O'Brien et al., 2021; Magee et al., 2021; O'Brien et al., 2023).

We extend this approach by using a Conditional Variational Auto Encoder (cVAE) to restrict the space of abundance and density profiles to a low-dimensional latent distribution

from which Monte-Carlo sampling can be performed efficiently. The conditions of the cVAE correspond to a set of progenitor channels associated with each model. We sample over the latent distribution conditioned on each progenitor channel along with parameters for a time and luminosity. This is then used to evaluate the model, by matching the synthetic spectral time series corresponding to each channel to an observed spectral time series. We do this for the well-studied SNe Ia SN 2011fe using emulated radiative transfer (Kerzendorf et al., 2021, 2022) for model evaluation. By integrating the Bayesian evidence of the posterior distribution under each channel, we determine a relative ranking of each progenitor system to accurately model the observed spectral time series and provide the relative probabilities that SN 2011fe originated from each channel. We decode the posterior distribution of latent parameters for each model to provide probabilistic posterior distributions of relative elemental mass fractions under each progenitor channel.

We outline the observational spectral time-series data in Section 4.2. In Section 4.3 we outline the space of the ejecta model under the cVAE as well as the progenitor channels being explored. We discuss the radiative transfer scheme used to evaluate the ejecta models in Section 4.4 and the details of the conditional radiative transfer emulator. In Section 4.5 we discuss the sampling and model-selection procedure, present our likelihood function used for spectral time-series inference, and explain the components of our prior-distribution. The results of our inference are presented in Section 4.6 and are discussed in Section 4.7. Details of the cVAE including architecture, training procedure, and data formats are presented in Appendix E.

4.2 Spectral Observations

We model a spectral time-series of SN 2011fe produced by the SNFactory (Aldering et al., 2002) taken between 8 to 12 days post explosion from Pereira et al. (2013). The spectra have been corrected for galactic extinction and shifted to the rest-frame. Our spectral time-series contains 4 spectral epochs interpolated to a logarithmic wavelength grid from $3300\,\text{Å}$ to

9700 Å with 500 points in order to re-weight line-contributions over their velocities.

4.3   Supernova Ejecta Model

We develop a model of the supernova ejecta which includes both the density profile and abundance profile for each element in a manner consistent with expectations derived from theoretical predictions from studies made using hydrodynamic simulations of the various progenitor channels. The ejecta profile consists of 100 velocity bins with 10 elements per bin represented as a 1100 dimensional vector of 10×100 elemental masses and 100 velocity edges. We use a variational auto-encoder (Rezende et al., 2014; Kingma & Welling, 2014) to compress the space of potential ejecta models to a dimension that is feasible to sample. This variational auto-encoder has the advantage that we may select the distribution of the compressed space, or latent space, to follow a normal distribution with mean of zero and standard deviation of one. We condition this auto encoder using a label corresponding to each progenitor system from which the training data originates (Sohn et al., 2015).

This model is trained on a public sample of SNe Ia ejecta profiles taken from the Heidelberg Supernovae Model Archive (HESMA Kromer et al., 2017). HESMA contains a set of ejecta models resulting from various studies using hydrodynamic simulations to model different progenitor systems under different scenarios. These models are spherically averaged isotopic abundances and densities over 3D models and we restrict the elements to only those that contribute to major line features of SNe Ia. We specifically use the abundances of carbon, oxygen, iron, titanium, chromium, silicon, sulfur, magnesium, calcium, and radioactive ^{56}Ni at $t_0 = 100$ seconds post explosion. The abundance profiles provided by HESMA are in the form of a table of velocity values, which define the edges of each cell, density values for each shell, and mass fractions of each element contained within each cell. The cVAE allows us to randomly draw complete ejecta profiles that are used for evaluation with the radiative transfer code TARDIS to generate synthetic spectra for inference. The decoder will produce full 1200 dimensional ejecta models that are physically realistic by randomly sampling in the

latent space which allows us to sample from a significantly smaller distribution and reduces the requirement for the number of samples by a factor of 10^{200}.

Our training data consists of 104 isotopic abundance and density profiles from 4 progenitor channels. These models include 12 pure detonation models (DET) from Sim et al. (2010); Marquardt et al. (2015), 45 pure deflagration models (DEF) from Fink et al. (2014); Kromer et al. (2015); Lach et al. (2022), 19 delayed detonation models (DDT) from Seitenzahl et al. (2013); Ohlmann et al. (2014), and 26 double detonation models (DOUBLEDET) from Sim et al. (2012); Gronow et al. (2020, 2021). Details of the architecture and training procedure for the cVAE can be found in the appendix.

4.3.1 Double Detonations

Double detonation models are systems in which an accretion-induced surface detonation propagates towards the center of a sub-M_{Ch} secondary WD imploding the core leading to a supersonic nuclear burning wave front. Nucleosynthetic burning products from the original surface detonation create the presence of heavier elements in the outer regions of the ejecta leading to reddening (Polin et al., 2021) and may explain the initial bump and red colors seen in some of the early lightcurves of SNe Ia (Polin et al., 2019; Bulla et al., 2020; Jiang et al., 2017; De et al., 2019). For the purpose of this study, we restrict our definition of double detonation models to those where there is a thick helium shell ($M_{He} \sim 0.1 M_\odot$) making significant contribution to the helium burning ash in the outer layers of the explosion. Thin helium shell models ($M_{He} \sim 0.01 M_\odot$) are represented by the pure detonation models discussed next.

4.3.2 Pure Detonations

Many studies have been performed to analyze the effects of thin-helium shell double detonations by approximating the explosion as a pure detonation of a sub-M_{Ch} WD with an artificial explosion placed at the center. These models allow us to understand the evolution

of the supernova when contributions from surface helium burning ash as insignificant to the observations. Violent mergers between WDs with a primary mass of around 0.9 $M_\odot$ also produce similar explosions to the pure detonation models (Pakmor et al., 2013). For the purpose of this study, pure detonation models encompass models in which surface helium burning has an insignificant contribution to observables and stands to represent either the class of thin helium shell double detonations or violent mergers.

4.3.3 Pure Deflagrations

Pure deflagration models are an older model for the explosions of near M_{Ch} WDs in binary systems. This explosion produce less ^{56}Ni in their ejecta and consequently lower overall luminosities in their light curves. These models are often used as a more accurate description of lower luminosity thermonuclear transients such as Type Iax Supernovae. We include them in our analysis to investigate if there is any potential for these models to be able to accurately reconstruct an archetypal Type Ia spectral time-series and to help serve as a baseline for the quality of our method of model comparison.

4.3.4 Deflagration-Detonation Transitions

Deflagration-Detonation Transitions, also known as delayed detonations, offer an avenue for a near M_{Ch} SNe Ia to produce the correct amount of ^{56}Ni to explain the light curve evolution of normal SNe Ia like SN 2011fe. These models represent an intermediate mechanism between pure detonations and pure deflagrations.

4.4 Radiative Transfer

Evaluation of a set of model parameters to produce a synthetic spectrum is performed with the Monte-Carlo radiative-transfer code TARDIS (Kerzendorf & Sim, 2014; Kerzendorf et al., 2021). We used the `nebular` approximation (Equation 3 in Kerzendorf & Sim, 2014) for solving ionization populations which is appropriate in the regime of an optically thin

expanding envelope in all ionization continua (Mazzali & Lucy, 1993), which is the case in the early phases being explored in this work. Excitation populations are solved with the `dilute-lte` prescription (Equation 5 in Kerzendorf & Sim, 2014) which approximates NLTE excitation levels in the optically thin limit. Line interactions are handled with the `macroatom` model (Lucy, 2002). Atomic data used with TARDIS in this study is taken from Kurucz GFALL (Kurucz & Bell, 1995). Full ejecta profiles are provided to TARDIS along with a time since explosion parameter and a requested luminosity which represents the total emitted luminosity over the wavelength range provided to TARDIS of $1000 \overset{\circ}{A}$ to $10\,000\,\text{Å}$.

4.4.1 Photospheric Inner Boundary Velocity Estimation

A photospheric inner boundary approximation is used as the opacity of the ejecta rises quickly with ejecta depth under the regime in which we are exploring. The location of this inner-boundary velocity is solved automatically to avoid including an extra sampling dimension and to limit the space of synthetic spectra to only those that are physically consistent with the dilution factor (see Kogure & Leung, 2007, for discussion). As this is similar to a approximately half of the energy of a black-body passing through the photosphere, we use an iterative approach to adjust the inner-boundary velocity of the simulation to a targeted mean optical depth of $\frac{2}{3}$ (see e.g. Dessart et al., 2014) computed using the Rosseland mean optical depth (Rosseland, 1924). The Rosseland mean optical depth is an appropriate approximation for the optical depth in the diffusion limit where the distances in question are smaller than the mean free path and an optical depth of $\frac{2}{3}$ corresponds to approximately $\frac{1}{2} \approx \exp\left[-\frac{2}{3}\right]$ of the radiation passing through the ejecta above the inner boundary.

Since TARDIS in its default mode already iteratively solves for the temperature of the inner-boundary, we only need to provide a requested luminosity to the code in order to solve for the location of the inner-boundary velocity. This inner-boundary is achieved through a damped convergence using the same damping constants used to solve to the dilution-factor and radiative temperatures by recomputing the location of the inner boundary velocity in the

ejecta at each step at which the optical depth would be $\frac{2}{3}$ in the current iteration through exponential extrapolation from the previous iterations. Our resulting parameterization then only requires the ejecta profile from the cVAE decoder (determined by the latent variables), a luminosity to set the inner boundary velocity and temperature, and a time since explosion to evolve the density and radioactively decay the ejecta.

4.4.2 Radiative Transfer Emulator

In order to accelerate the evaluation of TARDIS, we emulate the synthesis of synthetic spectra produced by a set of model parameters using a labeled version of the PDALEK emulator (Kerzendorf et al., 2022) using the same architecture as (O'Brien et al., 2023). An ensemble of 6 emulators is trained and their results are aggregated to produce probabilistic predictions in the form of a distribution consisting of a mean and standard deviation for the expected synthetic spectrum to be produced given a set of inputs.

To generate the training data for the spectral emulator, random points from the latent distribution are drawn fed through the decoder layer of the cVAE to generate random abundance profiles for each progenitor system. A corresponding set of time and luminosity parameters is drawn from a distribution based on the limits of the possible values for the SN 2011fe spectral time-series in our sample. The generated abundance profiles along with the time and luminosity parameters are input into TARDIS running with our routine for inner-boundary optimization to produce synthetic spectra. The training set consists of the 6 latent dimensions of the cVAE along with a time and luminosity parameter as the inputs. The outputs are the spectra produced by TARDIS through the formal integral method (Lucy, 1999c). In total, 65 536 training samples are generated for each progenitor channel with a total of 262 144 training samples across all progenitors labels.

4.5 Abundance Tomograpahy and Model Selection

The ejecta composition and structure is inferred for each potential progenitor system by conditioning the spectral emulator with the label corresponding to each progenitor channel. The entire spectral time series is evaluated under a single likelihood function using predictions from the spectral emulator. We integrate the posterior distribution corresponding to each progenitor channel to determine the evidence which is used to estimate the relative probabilities that each progenitor system best models the observed spectral time-series.

4.5.1 Prior Distribution

We construct the prior distribution from the distribution of the latent parameters of the cVAE along with independent distributions for the time and luminosity parameters. The time since explosion parameter is sampled uniformly between 8 and 12 days as this is the range of the data we selected as well as the range of the training data provided to the spectral emulator. The luminosity of each spectral epoch corresponds to the emitted luminosity over the wavelength range of the TARDIS models. The luminosity prior distribution is based on the range of bolometric luminosities estimated by Zhang et al. (2016) over the time interval for the observed spectra in the time series. Bolometric corrections are applied to the luminosity range using templates[32] from Nugent et al. (2002) to convert the bolometric luminosities to the luminosities emitted within the range of requested luminosities by TARDIS (see Section 4.4). An additional factor of 5% is added above and below this range to account for observational uncertainty and luminosity excess produced by TARDIS in the red end of the spectrum due to lack of complete continuum modeling (e.g. Mazzali et al., 2008). The final luminosity distribution is sampled logarithmically from a uniform distribution. The full parameterization of the prior distribution is shown in Table 4.

---

[32]https://c3.lbl.gov/nugent/nugent_templates.html

Model Parameter	Distribution
$z_{1:6}$[33]	Normal(0, 1)
t_{exp} (days)	Uniform(8, 12)
$L_{1:4}$[34] (erg/s/Å)	Log-Uniform($10^{42.22}$, $10^{43.27}$)
σ_{frac}	Log-Uniform(10^{-6}, 10^{-2})

Table 4: The prior distribution of model parameters for inferring our spectral time-series observations.

4.5.2 Likelihood Function

We apply a Gaussian likelihood of a model given the data. Continuum removal is approximated by multiplying the synthetic spectrum, $\hat{f}_{c;\lambda}$ with label c at wavelength λ, by a 5$^{\mathrm{th}}$-order polynomial fit to the ratio between the observed spectrum, f_λ, and $\hat{f}_{c;\lambda}$ determined over λ. Let $\hat{\bar{f}}_{c;\lambda}(\vec{\theta}_t)$ be the continuum-corrected synthetic spectrum in units of luminosity density in wavelength produced by a subspace, $\vec{\theta}_t$, of the model parameters, $\vec{\Theta}$, corresponding to an observation at a particular observation epoch, t, and $\hat{\bar{\sigma}}_{c,\lambda}(\vec{\theta}_t)$ be the continuum corrected emulator uncertainty.

The total form of the likelihood function is

$$\ln \mathcal{L}_c(\vec{\Theta}) = -\frac{1}{2} \sum_{t,\lambda} \left[\frac{\left(\hat{\bar{f}}_{c;\lambda}(\vec{\theta}_t) - f_{\lambda,t} \right)^2}{\sigma_{c;\lambda}^2(\vec{\theta}_t)} + \ln(2\pi\sigma_{c;\lambda}^2(\vec{\theta}_t)) \right]$$

where

$$\sigma_{c;\lambda}(\vec{\theta})^2 = \hat{\bar{\sigma}}_{c;\lambda}(\vec{\theta})^2 + \hat{\bar{f}}_{c;\lambda}(\vec{\theta})^2 \sigma_{\mathrm{frac}}^2 + \sigma_{\mathrm{obs};\lambda}^2$$

is the the aggregate of the corrected emulator uncertainty, $\hat{\bar{\sigma}}_{c;\lambda}(\vec{\theta}_t)$, the intrinsic fractional uncertainty $\hat{\bar{f}}_{c;\lambda}(\vec{\theta}_t)\sigma_{\mathrm{frac}}$, and the observational error $\sigma_{\mathrm{obs};\lambda}$. $\vec{\theta}_{t;0} = L_t$, $\vec{\theta}_{t;1} = t_{\mathrm{exp}} + \Delta t_t$, and $\vec{\theta}_{t;3:8} = z_{1:6}$ where Δt_t is the total time difference of a spectral observation from the first observation indexed as $t = 1$ where $\Delta t_1 = 0$ corresponding to the first spectrum in the time series.

4.5.3 Inference

Nested sampling (Skilling, 2004; Buchner, 2021) is performed using the UltraNest[35](Buchner, 2021) package to both sample the posterior distribution as well as compute the evidence integral which will be used for model selection. Sampling was performed with the "PopulationSliceSampler" to accelerate live-point selection up likelihood contours leveraging the computational efficiency of vectorized inputs to our full likelihood evaluation running on the GPU. The prior-distribution, $p(\Theta)$ is the same for each progenitor channel condition so the Bayesian evidence computed for each progenitor system is dominated by the likelihood function. The Bayesian evidence, Z, is computed as the integration over the posterior distribution from Bayes' Theorem (see e.g. Goodman, 2005)

$$Z_c = \int \mathcal{L}_c(\vec{\Theta})p(\vec{\Theta})d\vec{\Theta},$$

where $\mathcal{L}_c$ is the likelihood function conditioned on progenitor channel c.

4.5.4 Model Selection

The best progenitor system is determined by comparing the values of the evidence integrals for each inference run conditioned on each progenitor channel. Bayes factors are commonly used to relate two models to one another in determining the relative probability of those models being the best description of the system. As all 4 models are sampled from the sample prior-distribution, we are able to compare the relative contributions of their likelihoods to the posterior distribution. For the purposes of this study, we present the probabilities in terms of the total relative probabilities of the 4 progenitor conditions and so the sum of probabilities over all conditions sums to unity. The probability of a progenitor channel best describing the model is computed as

$$P_c = \frac{Z_c}{\sum_{c'=1}^{4} Z_{c'}}$$

---

[35]https://johannesbuchner.github.io/UltraNest/

Progenitor (c)	$\ln Z_c$	Probability (P_c)
Delayed Detonation	60321.90	0.0
Pure Deflagration	59699.81	0.0
Pure Detonation	**60602.97**	1.0
Double Detonation	60322.85	0.0

Table 5: The relative probabilities of each progenitor system based on the value of the evidence integral of the posterior distribution conditioned on each progenitor channel. The probabilities are normalized such that they sum to unity. We found that the spectral time-series of SN 2011fe is best described by a pure-detonation model (bold).

where c is the label corresponding to the progenitor system and Z_c is the the value of the evidence integral of the posterior distribution corresponding to that progenitor channel. The denominator is the sum over all 4 tested progenitor channels.

4.6   Results

Our Bayesian evidence estimation for each progenitor condition is presented in Table 5. The relative probability of one model, $a \in \{c\}$ compared to another $b \in \{c\}$ can be computed as $\exp\left[\ln Z_a - \ln Z_b\right]$. Across the 4 progenitor channels we have sampled, our Bayesian evidence estimates most favor the pure detonation progenitor channel. The double detonation and delayed detonation progenitor channels are nearly equally favored and the pure deflagration channel is significantly disfavored. It is important to note that we may only rank these progenitor channels out of the progenitor channels that we have investigated and can not not directly conclude that SN 2011fe was produced by a specific progenitor channel as there exist models this analysis has not considered and we do not discount the possibility that a future proposed models may better explain observations of SNe Ia.

Posterior distributions for the spectral time-series for the pure detonation progenitor channel is shown in Figure 9(a). Our emulator produces distributions of spectra given by a mean and standard deviation (see Section 4.4.2) so the posterior distribution of spectra is a distribution of distributions. We therefore represent the posterior distribution of spectra as the median of posterior means of each spectrum with the shaded region representing the

median of the posterior distribution of spectra at 1 standard deviation.

The resulting posterior distributions of the latent variables from our fits to the spectral time series of SN 2011fe are transformed back through the decoder layer of the cVAE to provide physical descriptions the posterior abundance profiles. Abundances and densities presented at low velocities less than approximately $10\,000\,\mathrm{km/s}$ are not directly informed by information in the spectral observations as they lie below the inner boundary velocities used in our photospheric approximations and no information from this region is imprinted on the simulated spectral time series. Therefore their distributions are informed by the expected correlations determined by the cVAE for the inner region abundances based on models whose outer layer abundances are consistent with the results from our spectral inference and should not be interpreted as data-informed estimate, but instead resulting from strong priors on our models.

The reconstructed abundance tomography from the decoded posterior distribution of latent variables is shown in Figure 9(b). We use point estimates from the decoder layer to produce the abundance profiles to be emulated for spectral synthesis so, unlike for the spectral time series posterior, each posterior sample represents a point estimate of the distribution. Our posterior abundance distributions produce individual samples in both fractional abundance and velocity for each point in the ejecta, so the posterior distribution of abundance profiles shown has each sample interpolated to a common velocity grid. The shaded regions show the posterior estimates at the 99.7% quantile and the solid lines represent the median of each distribution. The equivalent figures for the pure-deflagration, delayed-detonation, and double detonation models are shown in Appendix F for reference.

4.7 Discussion

The appearance of of pure-detonation may be the result of either a double detonation with a thin helium shell (Pakmor et al., 2022; Shen et al., 2021), a fast delayed detonation with a transition occurring at a higher velocity (Blondin et al., 2012), or a violent merger between two

(a)

(b)
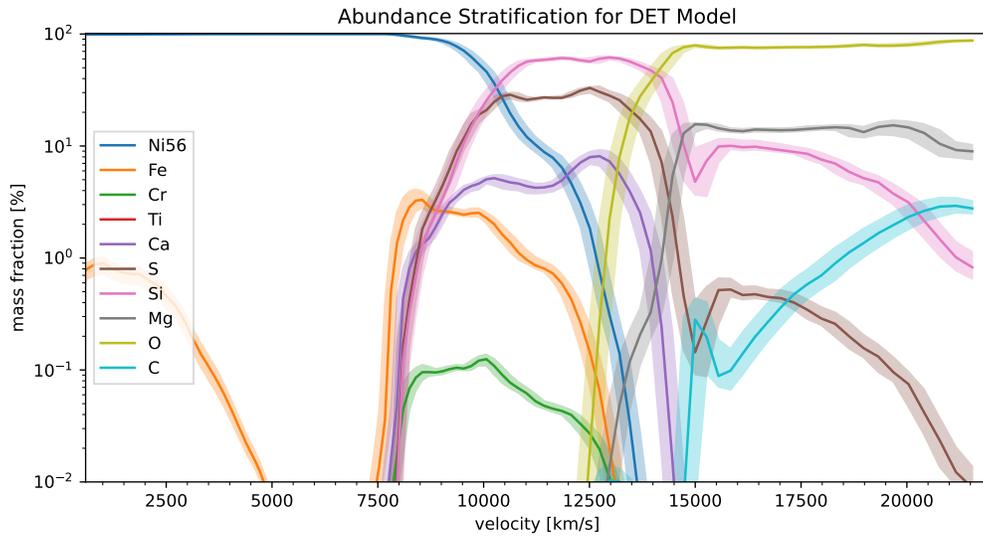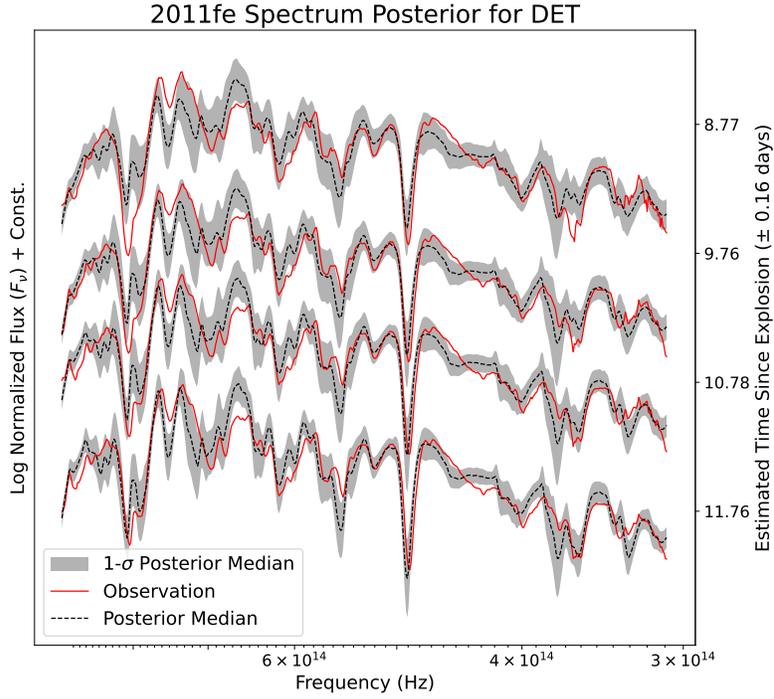
Figure 9: 9(a) Spectral time series from the best fit Pure-Detonation progenitor model for SN 2011fe. 9(b) Abundance distribution of reconstructed ejecta profiles for SN 2011fe under the DET model.

degenerate WDs (Pakmor et al., 2012, 2013). The Bayesian evidence between the DDT and DOUBLEDET progenitor channels from our inference are too close to firmly draw conclusions about which channel from which SN 2011fe may have originated given that the estimated uncertainties on $\ln Z$ are approximately 0.5. Violent merger models were not included in the cVAE so we cannot provide a quantitative ranking for these models to describe SN 2011fe but such models may still be considered a reasonable progenitor channel given their similarity to pure-detonation models and previous work (e.g. Bloom et al., 2012; Röpke et al., 2012; Nugent et al., 2011) suggesting SN 2011fe may have originated in a double-degenerate binary system.

We compare the distribution of latent variables from our DET posterior to the positions in latent space of the models in the training and validations sets for the cVAE. We find the closest model in latent space is the 0.88 $M_\odot$ model from Sim et al. (2010) indicating that our predicted distribution is more consistent with a pure-detonation of a C/O WD than the ONe WD models from Marquardt et al. (2015) which serves as a sanity test of our methodology but since we cannot make strong statements regarding the density and abundances within the inner layers of the ejecta we cannot confidently put constraints on the progenitor mass itself.

Our posterior-abundance tomography for the DET model shows many similarities to that of Mazzali et al. (2015) especially in the outer-layers of the ejecta. We find a similar range of peak silicon and sulfur abundances between 10 000 and 12 000 km/s but with a significantly elevated calcium abundance in our model. We find that most iron-group elements are much more concentrated below 10 000 km/s in our model with a sharper drop-off in the outer layers more consistent of the high mass thin-shell double-detonation models of Polin et al. (2019). We find an earlier drop-off in Oxygen into the inner layers of the explosion compared to Mazzali et al. (2015) but later than Polin et al. (2019) but this may be a result of the insensitivity of photospheric phase spectra to changes in oxygen abundance (e.g. Hachinger et al., 2017).

We find more disagreement with the models of Mazzali et al. (2015) in the deeper regions

of the ejecta instead favoring the thin-shell double-detonation model compared in Polin et al. (2021), though this may be a result of the inability of our model to sample deeper into the ejecta at early times due to the photospheric inner boundary approximating cutting off information about the ejecta profile deeper than the placement of the inner boundary velocity which will require nebular-phase models in order to fully resolve.

Overall, due to the constraints on the detonation models requiring extreme stratification of elements in the ejecta we find a more stratified model though we are still able to accurately reproduce the spectral time-series of SN 2011fe over the 4-day interval of our sample. Interestingly, our abundance stratification for the DDT model in Figure F.3 is in most agreement with the results of Mazzali et al. (2015) which used a DDT density profile to perform their abundance tomography, demonstrating the influence of an assumed progenitor model on the resulting predicted abundance stratification.

4.8   Conclusions

We have performed a generalized probabilistic abundance tomography of the outer ejecta of SN 2011fe from 4 photospheric phase optical spectra conditioned on different progenitor channels. Integrating the posterior distributions of our generative model for each progenitor condition reveals that the spectral time-series of SN 2011fe is most consistent with a pure-detonation sub-M_{Ch} C/O WD. This progenitor channel may either represent a thin helium shell double-detonation, a delayed detonation with an early deflagration-detonation transition, or a violent merger between two C/O WDs. Our Bayesian evidence is not strong enough to directly favor the double-detonation model over the delayed detonation model or vice versa, but our abundance stratification for our posterior distribution of pure-detonation models is most similar to that of the double-detonation model for SN 2011fe explored in Polin et al. (2021). Violent merger progenitor channels were not included in our cVAE model but their ejecta stratification is similar to that of the pure-detonation models and may represent a promising progenitor scenario.

Extending our study to a larger temporal range of spectra and including more progenitor systems will be required in order to further differentiate between these models given that our inference is insensitive to the compositions and densities of the inner regions of the ejecta due to our analysis being limited to the photospheric phase. As hydrodynamic simulations of the evolution of SNe Ia improve in fidelity and become more widely available, we will be able to improve the priors set on our cVAE model and explore a larger range of physically realistic ejecta configurations for each progenitor channel as well as a larger variety of progenitor channels overall.

4.9   Acknowledgements

---

[36]`https://www.astropy.org`

5 Conclusions and Future Work

For the first time, rapid probabilistic abundance tomography and, with it, a path towards progenitor identification is now possible. Radiative transfer simulations of the outer-ejecta of SNe Ia have been successfully accelerated through the use of deep-learning models to reduce evaluation time from nearly an hour down to a few milliseconds, allowing full Bayesian posteriors of ejecta models to be determined from direct spectral observations.

Modeling the outer-ejecta of the prototypical SNe Ia SN 2002bo has not only revealed the composition of the outer layers of its explosion, but shown that the probability space is degenerate and complex, with multi-modalities and large uninformative ranges in elemental abundances yielding qualitatively similar spectra. Comparison to expected elemental compositions in the outer-ejecta from theoretical models have shown that SN 2002bo is best described by a detonation-driven model, ruling out the pure-deflagration progenitor channel for this supernova.

Expanding this analysis to a multi-zone model applied to two populations of SNe Ia has provided valuable insight into the differences between normal SNe Ia and the super-luminous 1991T-like SNe Ia. The investigation into the nucleosyntheitic products in their outer ejecta has shown that the 1991T-like SNe Ia appear as an extension of the normal SNe Ia population, without any clear clustering that would imply a distinct progenitor system. Analysis of the ionization states of the intermediate mass elements of the ejecta of supernovae from both populations has revealed that 1991T-like SNe Ia appear observationally distinct for a combination of reasons, where a slight decrease in the intermediate mass elements in the outer-ejecta combined with slightly higher ionization states lead to the suppression of IME line formation in their early-time spectra, giving them an observationally distinct appearance.

Improving the ejecta model to be more in-line with theoretical predictions has provided a novel framework for direct probabilistic ranking of various progenitor channels to describe a given SN Ia. A set of labeled ejecta profiles produced by hydrodynamic simulations of various progenitor scenarios was compressed to a low-dimensional i.d.d. latent distribution to create

a more realistic parameter space to describe SNe Ia explosions. Inference over a spectral time series of the well studied SN Ia SN 2011fe conditioned on each progenitor channel provided estimates of the Bayesian evidence that a given progenitor channel describes its explosion. Our evidence favors a model that appears as a pure-detonation of a sub-M_{ch} WD, which may have been the result of either a double-detonation of WD with a thin helium shell or a violent merger between a pair of low-mass WDs.

Further modeling will be required to narrow down the exact explosion mechanism and progenitor system. Additional physics such as full continuum modeling, gamma-ray deposition, and the inclusion of more progenitor models applied to a temporally longer optical spectral time series will further constrain the progenitors of SNe Ia and drive the direction of research towards the most promising models. Higher-fidelity hydrodynamical simulations will improve our sampling space by restricting it to higher quality models that better reflect realistic physical scenarios and improved machine-learning models involving techniques such as active learning updates will rapidly allow our parameterizations to converge towards the true distribution of SNe Ia ejecta compositions.

A Emulator

Emulation is the practice of developing some analytic function that approximates the behavior of another function. TARDIS can be thought of as a function mapping a vector of supernova parameters to a vector representing a spectrum. We extend the techniques described in the Kerzendorf et al. (2021) paper to make an emulator for the $-10\,\text{day}$ spectrum of SN 2002bo. The method proposed by Kerzendorf et al. (2021) uses an ensemble of feed-forward neural networks to emulate the spectrum computation. Our neural network is trained from a set of pre-computed data points, composed of training spectra over a grid spanning a physically plausible parameter space for a SN Ia. The goal for the emulator is to be used in our parameter inference so we ensure that the training set parameter space contains the final prior fitting space (see Section 2.3.1).

We changed several parts of the procedure when compared to the emulator described by Kerzendorf et al. (2021). One key difference is the addition of two parameters: the power law index α_ρ and the time since explosion t_{exp}. The bounds on parameters corresponding to computed spectra were also modified to encompass elemental abundances corresponding to shells above $8000\,\text{km/s}$ in HESMA models. Kerzendorf et al. (2021) presented an ensemble of different neural network architectures that could reproduce simulated TARDIS spectra to a high degree of precision. For this paper, for computational efficiency, we chose only a single network from the neural networks described by Kerzendorf et al. (2021). Specifically, we used a model which propagates the 14 inputs through three subsequent hidden layers of 400 neurons each, reaching 500 outputs. The hidden units used the "softplus" activation function. We trained our emulator with the "nadam" optimizer on a $91\,000$ sample training set and $39\,000$ sample validation set in a 70%/30% training/validation split. Training time was 20 minutes on an NVIDIA® GeForce® RTX 2080Ti GPU.

The measured accuracy of our emulator using the mean and maximum fractional error (Figure A.1) is similar to that of the initial DALEK emulator. Figure A.1 shows that our mean fractional error is almost always below 1% over our validation set. The final fit presented
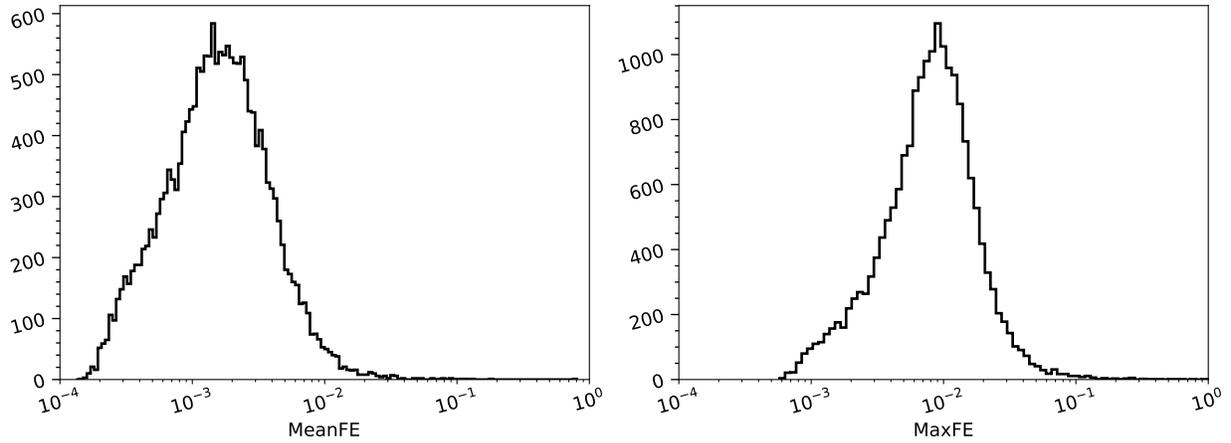
Figure A.1: Mean and Maximum fractional error for our TARDIS emulator. Test spectra are compared to emulated spectra generated using the same parameter set. The low level of error demonstrates that our emulator is effective at modeling the physics of TARDIS. Descriptions of the mean and maximum fractional error can be found in Kerzendorf et al. (2021)

in Section 4.6 has a mean fractional error of 10% between the observed spectrum and the maximum posterior model indicating that any uncertainty from our emulation is less than systematics for the presented work.

B   External Links to Data

The TARDIS configuration file, posterior samples with their associated weights, and the parameter grid and corresponding spectra used in training the emulator are provided through Zenodo: [10.5281/zenodo.5007378]10.5281/zenodo.5007378.

The observed spectrum of SN 2002bo used in this paper is hosted by the Open Supernova Catalog (Guillochon et al., 2017).

C   Data Products

TARDIS configuration information, emulator weights and training data, and example scripts can be found at the following location can be found at the following link: 10.5281/zenodo.7818303

D   Outer Ejecta Inference from Single Spectrum

Photospheric phase abundance tomography generally relies upon analysis of a spectral time-series in order to infer the distribution of abundances at each velocity interval. We show that information of the composition of the ejecta at all velocities above our photospheric inner boundary is resolvable from a single-spectrum fit in the early phases. As packets travel through the ejecta, they interact with material seated at different velocities and radiative temperatures leading to interactions at different ionization states Doppler shifted by different velocities. All of these interactions contribute to the observed spectral features at this epoch and therefore contribute to the likelihood function when inference is performed on an observed spectrum.

E   Conditional Variational Auto-Encoder
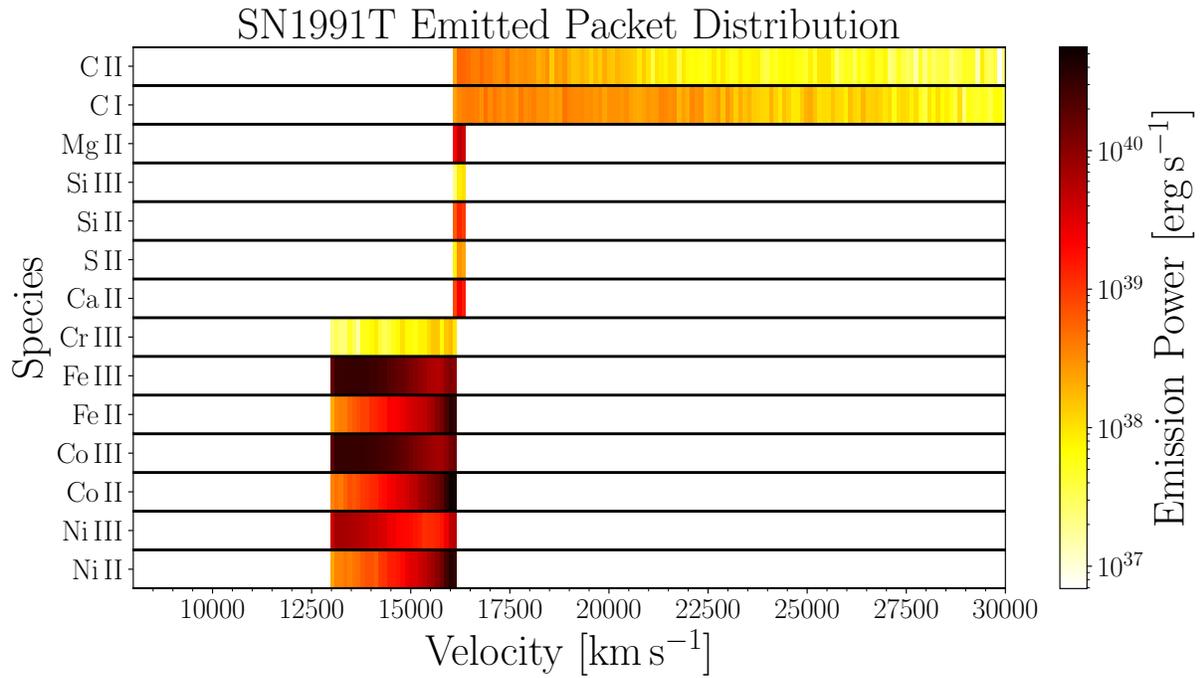
Figure D.2: Power emitted from final line interactions of emitted packets in TARDIS simulation of ejecta profile constructed from the maximum-likelihood sample for SN 1991T. The packet interactions that contribute to emission features in the spectrum span a wide range in velocity-space in the outer-ejecta.

Algorithm 1 Training algorithm.

Hyper-parameters: $n_{\text{batch}}, \xi_{\text{rec}}, \eta_{\text{rec}}$
Constants: $\lambda_{\text{rec}}^0, \lambda_{\text{rec}}^{\max}, \alpha$
Initialise $t = 0$
Initialise $\lambda_{\text{rec}} = \lambda_{\text{rec}}^0$
Initialise InitialPhaseRec = True
while training **do**
   Read current data batch m_b, v_b, c_b of size n_{batch}
   Sample from variational posterior $z_b \sim q_\phi(\cdot | m_b, v_b, c_b)$
   Compute $h_{\text{rec}} = L_{\text{rec}} - \xi_{\text{rec}}$ (batch average)
   $\hat{h}_{\text{rec}} \leftarrow (1 - \alpha)\, \hat{h}_{\text{rec}} + \alpha\, h_{\text{rec}},\ (h_{\text{rec}}^{(0)} = h_{\text{rec}})$
   if $h_{\text{rec}} < 0$ **and** InitialPhaseRec **then**
     InitialPhaseRec = False
   end if
   if $\neg$InitialPhaseRec **then**
     $\lambda_{\text{rec}} \leftarrow \lambda_{\text{rec}} \cdot \exp\{\eta_{\text{rec}} \cdot \hat{h}_{\text{rec}}\}$
     $\lambda_{\text{rec}} \leftarrow \text{clip}(\lambda_{\text{rec}}, \lambda_{\text{rec}}^{\max})$
   end if
   Compute loss $L(\theta, \phi) \leftarrow \lambda_{\text{rec}}(L_{\text{rec}} - \xi_{\text{rec}})$
   Compute $L(\theta, \phi) \leftarrow L(\theta, \phi) + \text{KL}[q_\phi(z_b; m_b, v_b, c_b) \,\|\, p(z)]$ (batch average)
   update (θ, ϕ) using $(\partial_\theta L(\theta, \phi), \partial_\phi L(\theta, \phi))$
   $t \leftarrow t + 1$
end while

Table 6: Model architecture and hyper-parameters. FC refers to fully connected layer.

NAME	HYPER-PARAMETERS
INPUT m, v DIMENSION	1100
LABEL c DIMENSION	4
LATENT DIMENSION	6
$p(z)$	NORMAL DISTRIBUTION
ACTIVATION OF μ OF $p_\theta(z|m, u, c)$	LINEAR
ACTIVATION OF σ OF $p_\theta(z|m, u, c)$	SOFTPLUS
ACTIVATION OF μ OF $q_\phi(m, u|z, c)$	LINEAR
ACTIVATION OF σ OF $q_\phi(m, u|z, c)$	SOFTPLUS
$q_\phi(m, u|z, c)$	(FC 256 LEAKYRELU) $\times$ 2 LAYERS
$p_\theta(z|m, u, c)$	(FC 256 LEAKYRELU) $\times$ 2 LAYERS
BATCH SIZE	FULL TRAINING DATA
LEARNING RATE	0.0008
OPTIMIZER	ADAM
GRADIENT CLIP VAL	0.5
WEIGHT DECAY	0.0005
DROPOUT	0.1

E.1 Latent variable model for abundance profile

To learn a conditional joint distribution over u, m, and c, we propose the hierarchical probabilistic model

$$p_\theta(u, m \mid c) = \int p_{\theta_u}(u \mid c, z) \, p_{\theta_m}(m \mid c, z) \, p_0(z) \, dz \tag{8}$$

where we use $\theta = (\theta_v, \theta_m)$ to denote all parameters of the model and we choose the prior $p_0(z) = \mathcal{N}(z; 0, I)$.

To learn this model we consider our data $\{(u_k, m_k, c_k)\}_k$ to be independent and identically distributed. We would like to use the maximum-likelihood method $\theta^\star = \mathrm{argmax}_\theta \sum_k \log p_\theta(u_k, m_k \mid c_k)$, however, it is not directly applicable since $p_\theta(u, m \mid c)$ can be analytically intractable. For this reason we use a variational lower bound on $\log p_\theta(u, m \mid c)$ called the *evidence lower bound* (ELBO) defined as

$$\log p_\theta(u_k, m_k \mid c_k, z_k) \geq \max_{q_k} \Big\{ \mathrm{E}_{q_k(z_k)}\big[\log p_\theta(u_k, m_k \mid c_k, z_k)\big] - \mathrm{KL}\big[q_k(z_k) \,||\, p_0(z_k)\big] \Big\} \tag{9}$$

via a variational distribution $q_k(z_k)$. The ELBO is tight for $q_k^*(z_k) = p_\theta(z_k \mid u_k, m_k, c_k)$, however, we are often restricted to choose $q_k(z_k)$ from an analytically tractable class of distribution such as the class of multivariate Normal distributions. The resulting algorithm is known in the literature as a variational expectation maximization (Neal & Hinton, 1998). Optimizing the ELBO—solving for $q_k(z_k)$—for each data item (u_k, m_k, c_k) is computationally demanding and hence additional approximation techniques have been introduced such as *amortization* where we jointly optimize the bounds for all data k by parameterizing $q_k(z_k)$ as $q_k(z_k) = q_\phi(z_k \mid u_k, m_k, c_k)$ and optimizing w.r.t. ϕ (Kingma & Welling, 2014). Let us use $\hat{p}(u, m, c)$ to denote the empirical distribution of the data. The learning problem for our

model and dataset is thus formulated as

$$\min_{\theta,\phi} E_{\hat{p}(u,m,c)} \left[E_{q_\phi(z;u,m,c)}[-\log p_\theta(u,m \mid c,z)] + \text{KL}[q_\phi(z;u,m,c) \mid\mid p_0(z)]\right]. \qquad (10)$$

To model the conditional distributions we use multivariate Normal distributions with diagonal covariance matrices and neural network models to define the conditional means and variances, that is, $p_{\theta_u}(u \mid c,z) = \mathcal{N}(u; \mu_{\theta_u}^{\text{NN}}(c,z), \sigma_{\theta_u}^{\text{NN}}(c,z)^2)$, $p_{\theta_m}(m \mid c,z) = \mathcal{N}(m; \mu_{\theta_m}^{\text{NN}}(c,z), \sigma_{\theta_m}^{\text{NN}}(c,z)^2)$, and $q_\phi(z \mid u,m,c) = \mathcal{N}(z; \mu_\phi^{\text{NN}}(u,m,c), \sigma_\phi^{\text{NN}}(u,m,c)^2)$.

Training probabilistic generative models such as VAEs (Kingma & Welling, 2014) can be a complex task due to issues like suboptimal local minima (Sønderby et al., 2016) or posterior collapse (Bowman et al., 2015; Kingma et al., 2016). Additionally, it's not a given that higher Evidence Lower Bound (ELBO) values will lead to better prediction performance or more informative latent spaces (Alemi et al., 2018; Higgins et al., 2017). To mitigate the risk of over-regularization of q_ϕ towards p_0, a variety of annealing strategies have been introduced that gradually "switch on" the KL-divergence term in the ELBO. Particularly, scheduling schemes that are derived from constrained optimization methods (Rezende & Viola, 2018) can notably enhance the training process in hierarchical generative models (Klushyn et al., 2019). For this reason, instead of optimizing (10) we solve the optimization problem

$$\min_{\theta,\phi} E_{\hat{p}(u,m,c)} \left[\text{KL}\big[q_\phi(z;u,m,c) \mid\mid p_0(z)\big]\right] \qquad (11)$$

$$\text{s.t. } E_{\hat{p}(u,m,c)} E_{q_\phi(z;u,m,c)}[-\log p_\theta(u,m \mid c,z)] \leq \xi. \qquad (12)$$

The resulting Lagrangian has a similar form as (10) with an additional Lagrange multiplier λ for the expected log-likelihood. The resulting saddle-point optimization adaptively balances the relative weight of the two terms via λ to alleviate (some) of the above mentioned problems. We follow the method proposed in (Chen et al., 2022) to solve the saddle-point optimization.

E.2 Training and evaluation

The dataset sizes for the four abundance profiles is imbalanced between samples from the DDT, DET, DEF, and DOUBLEDET progenitor channels. For this reason, during the training process we adjust the sampling frequency from the data corresponding to these profiles to have an even coverage of each dataset. To solve the saddle-point optimization resulting from (11) we use a stochastic batch gradient descent-ascent. For the descent we use ADAM (Kingma & Ba, 2014) gradient-descent steps while for the Lagrange multiplier λ we use the EMM (Bertsekas, 2003) quasi-gradient-ascent steps $\lambda^{(t+1)} = \lambda^{(t)} \exp\{\eta(\mathrm{E}_{\hat{p}(u,m,c)}\mathrm{E}_{q_\phi(z;u,m,c)}[-\log p_\theta(u,m \mid z,c)] - \xi)\}$. Intuitively, when the constraints are not satisfied, λ, and thus the weight of the relative weight of the reconstruction term, increases. When the constraints are satisfied the opposite relative reweighting takes place. We run the optimization until it converges. We select the optimal hyperparameters based on which model yields the lowest KL loss after satisfying the constraint. The complete optimization algorithm and the neural network architectures used in the model are presented in detail in Algorithm 2 and Table 7.

Table 7: Model architecture and hyper-parameters. FC refers to fully connected layer.

NAME	HYPER-PARAMETERS
INPUT m, v DIMENSION	1100
LABEL c DIMENSION	4
LATENT DIMENSION	6
$p(z)$	NORMAL DISTRIBUTION
ACTIVATION OF μ OF $p_\theta(z \mid m, u, c)$	LINEAR
ACTIVATION OF σ OF $p_\theta(z \mid m, u, c)$	SOFTPLUS
ACTIVATION OF μ OF $q_\phi(m, u \mid z, c)$	LINEAR
ACTIVATION OF σ OF $q_\phi(m, u \mid z, c)$	SOFTPLUS
$q_\phi(m, u \mid z, c)$	(FC 256 LEAKYRELU) $\times$ 2 LAYERS
$p_\theta(z \mid m, u, c)$	(FC 256 LEAKYRELU) $\times$ 2 LAYERS
BATCH SIZE	FULL TRAINING DATA
LEARNING RATE	0.0008
OPTIMIZER	ADAM
GRADIENT CLIP VAL	0.5
WEIGHT DECAY	0.0005
DROPOUT	0.1

F Extra Figures

Algorithm 2 Training algorithm.

  Hyper-parameters: $n_{\mathrm{batch}}, \xi_{\mathrm{rec}}, \eta_{\mathrm{rec}}$
  Constants: $\lambda_{\mathrm{rec}}^0, \lambda_{\mathrm{rec}}^{\max}, \alpha$
  Initialize $t = 0$
  Initialize $\lambda_{\mathrm{rec}} = \lambda_{\mathrm{rec}}^0$
  Initialize InitialPhaseRec = True
  while training **do**
    Read current data batch m_b, u_b, c_b of size n_{batch}
    Sample from variational posterior $z_b \sim q_\phi(\cdot \mid m_b, u_b, c_b)$
    Compute $h_{\mathrm{rec}} = L_{\mathrm{rec}} - \xi_{\mathrm{rec}}$ (batch average)
    $\hat{h}_{\mathrm{rec}} \leftarrow (1 - \alpha)\,\hat{h}_{\mathrm{rec}} + \alpha\,h_{\mathrm{rec}},\ (h_{\mathrm{rec}}^{(0)} = h_{\mathrm{rec}})$
    if $h_{\mathrm{rec}} < 0$ **and** InitialPhaseRec **then**
      InitialPhaseRec = False
    end if
    if ¬InitialPhaseRec **then**
      $\lambda_{\mathrm{rec}} \leftarrow \lambda_{\mathrm{rec}} \cdot \exp\{\eta_{\mathrm{rec}} \cdot \hat{h}_{\mathrm{rec}}\}$
      $\lambda_{\mathrm{rec}} \leftarrow \mathrm{clip}(\lambda_{\mathrm{rec}}, \lambda_{\mathrm{rec}}^{\max})$
    end if
    Compute loss $L(\theta, \phi) \leftarrow \lambda_{\mathrm{rec}}(L_{\mathrm{rec}} - \xi_{\mathrm{rec}})$
    Compute $L(\theta, \phi) \leftarrow L(\theta, \phi) + \mathrm{KL}[q_\phi(z_b; m_b, u_b, c_b) \,\|\, p(z)]$ (batch average)
    update (θ, ϕ) using $(\partial_\theta L(\theta, \phi), \partial_\phi L(\theta, \phi))$
    $t \leftarrow t + 1$
  end while
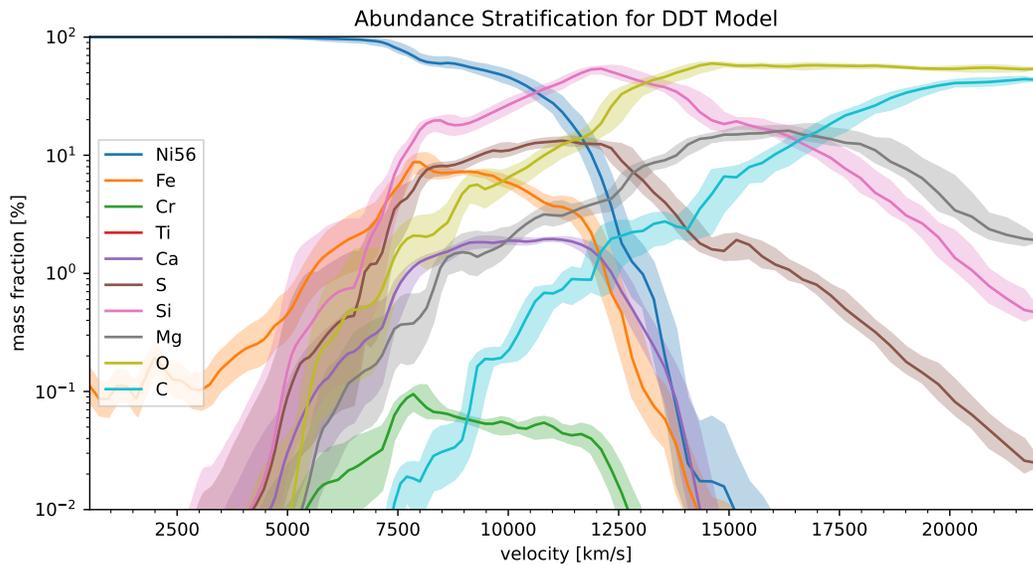
Figure F.3: Posterior distribution of reconstructed ejecta profiles for SN 2011fe under the DDT model.

References

Abadi, M., Agarwal, A., Barham, P., et al. 2015, TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. `http://tensorflow.org/`

Aldering, G., Adam, G., Antilogus, P., et al. 2002, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 4836, Survey and Other Telescope Technologies and Discoveries, ed. J. A. Tyson & S. Wolff, 61–72, doi: `10.1117/12.458107`

Alemi, A. A., Poole, B., Fischer, I., et al. 2018, ICML

Aouad, C. J., Mazzali, P. A., Hachinger, S., et al. 2022, Monthly Notices of the Royal Astronomical Society, 515, 4445, doi: `10.1093/mnras/stac2024`

Arnett, W. D. 1969, APSS, 5, 180, doi: `10.1007/BF00650291`

Astropy Collaboration, Robitaille, T. P., Tollerud, E. J., et al. 2013, A&A, 558, A33, doi: `10.1051/0004-6361/201322068`
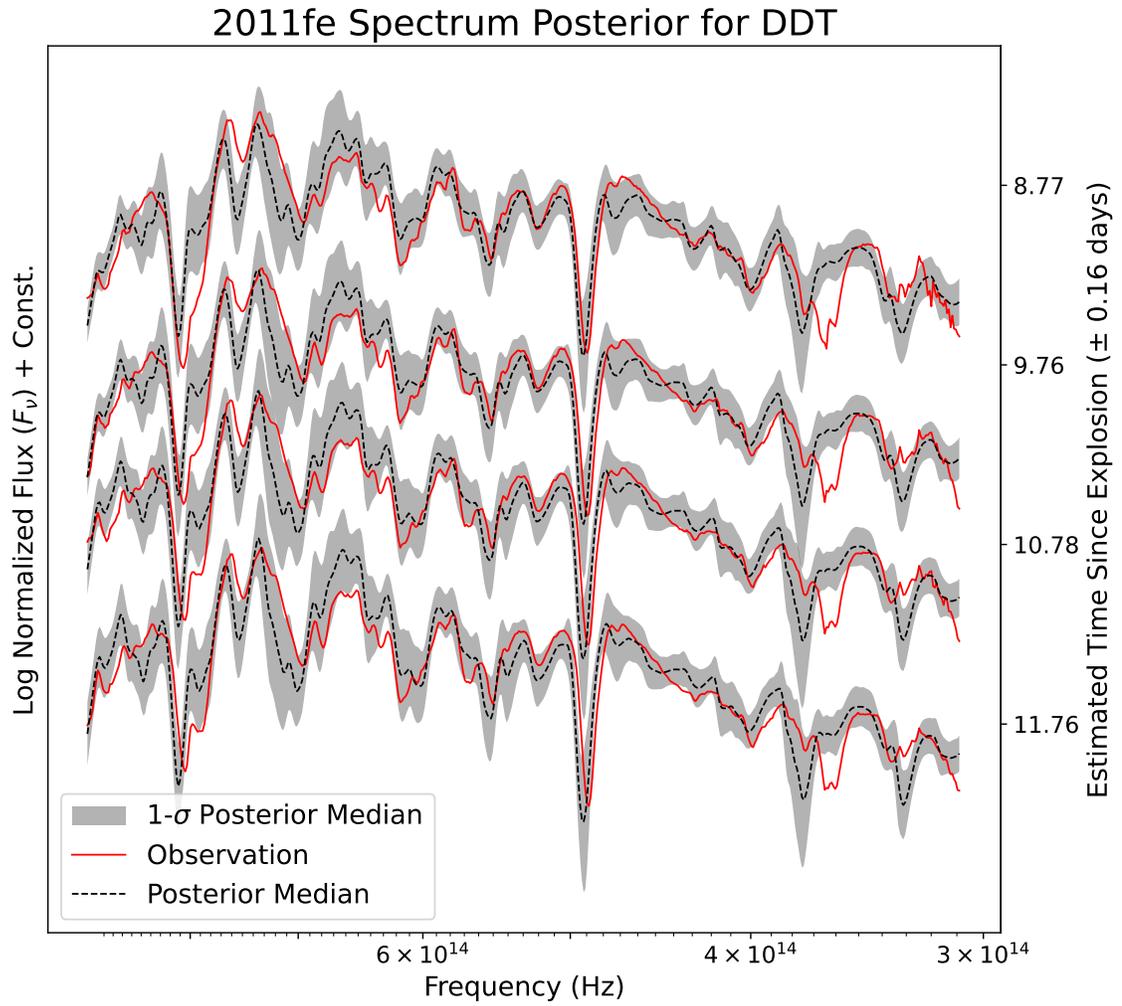
Figure F.4: Spectra from the best fit Delayed Detonation progenitor model for SN 2011fe.
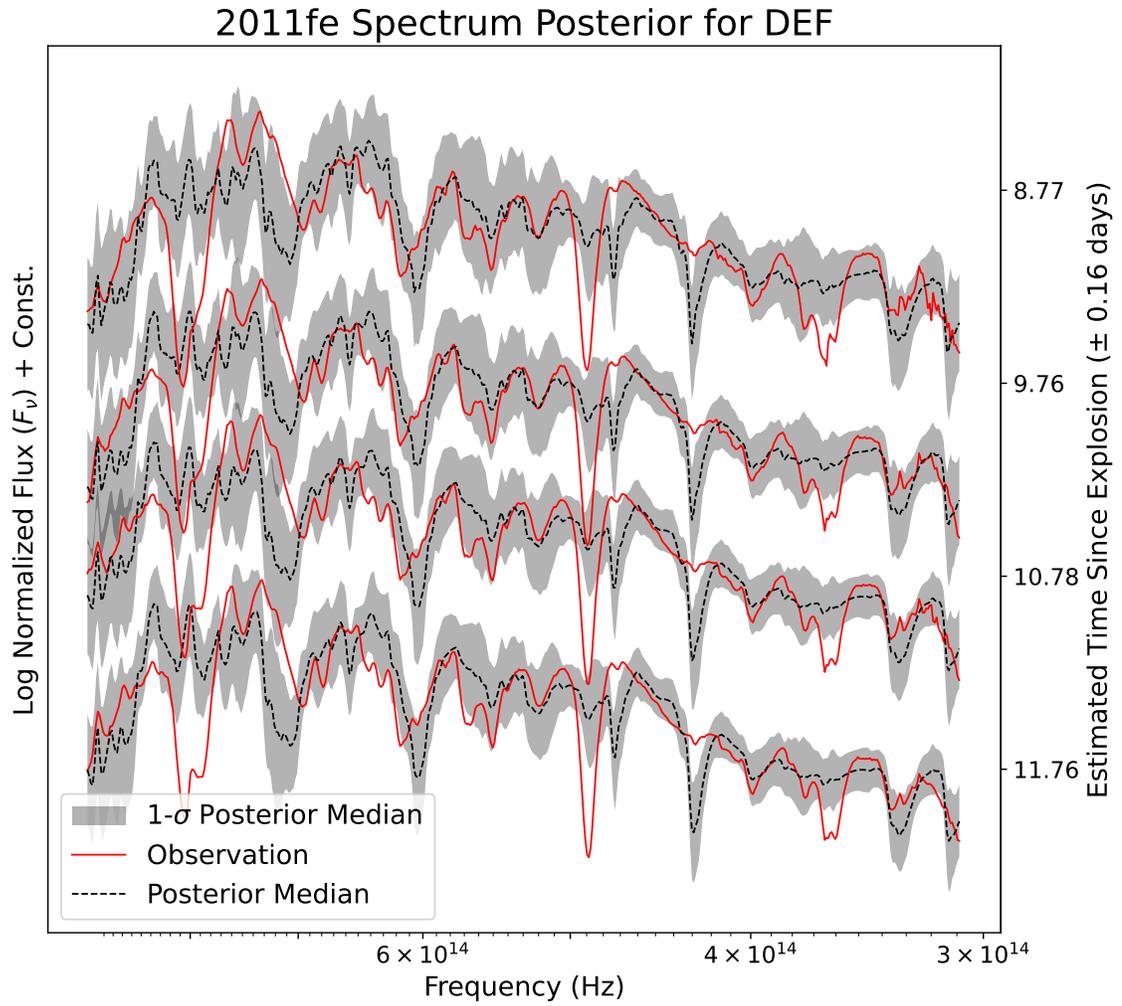
Figure F.5: Spectra from the best fit Pure-Deflagration progenitor model for SN 2011fe.
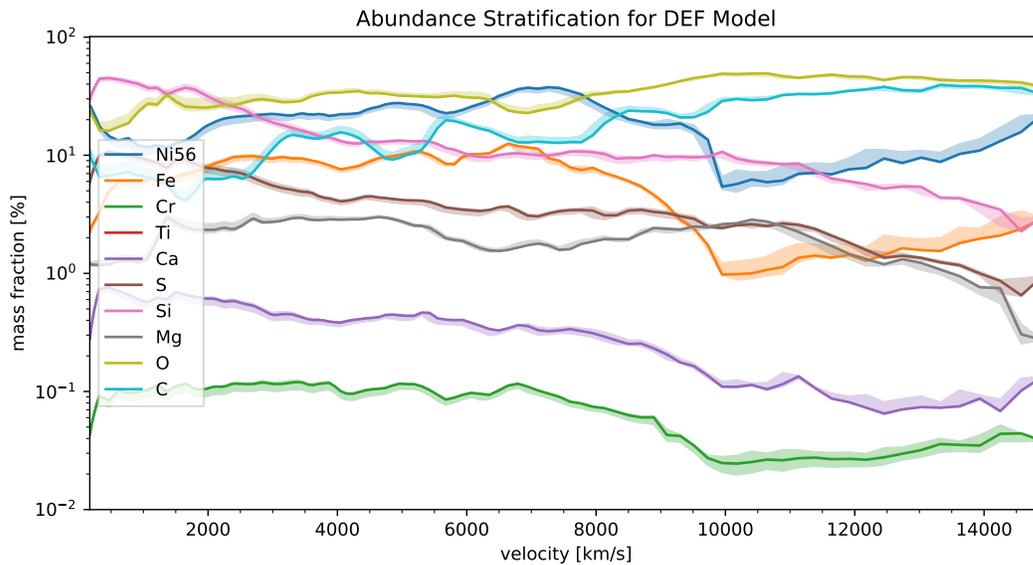
Figure F.6: Posterior distribution of reconstructed ejecta profiles for SN 2011fe under the DEF model.

Astropy Collaboration, Price-Whelan, A. M., Sipőcz, B. M., et al. 2018, AJ, 156, 123, doi: `10.3847/1538-3881/aabc4f`

Axelrod, T. S. 1980, PhD thesis, University of California, Santa Cruz

Barbary, K. 2016, extinction v0.3.0, Zenodo, doi: `10.5281/zenodo.804967`

Barna, B., Szalai, T., Kromer, M., et al. 2017, Monthly Notices of the Royal Astronomical Society, 471, 4865, doi: `10.1093/mnras/stx1894`

Beluch, W. H., Genewein, T., Nürnberger, A., & Köhler, J. M. 2018, in Proceedings of the IEEE conference on computer vision and pattern recognition, 9368–9377

Benetti, S., Meikle, P., Stehle, M., et al. 2004, Monthly Notices of the Royal Astronomical Society, 348, 261, doi: `10.1111/j.1365-2966.2004.07357.x`

Bertsekas, D. P. 2003, Nonlinear Programming: Second Edition (Athena Scientific)

Blondin, S., Dessart, L., Hillier, D. J., & Khokhlov, A. M. 2012, Monthly Notices of the Royal Astronomical Society, 429, 2127, doi: `10.1093/mnras/sts484`
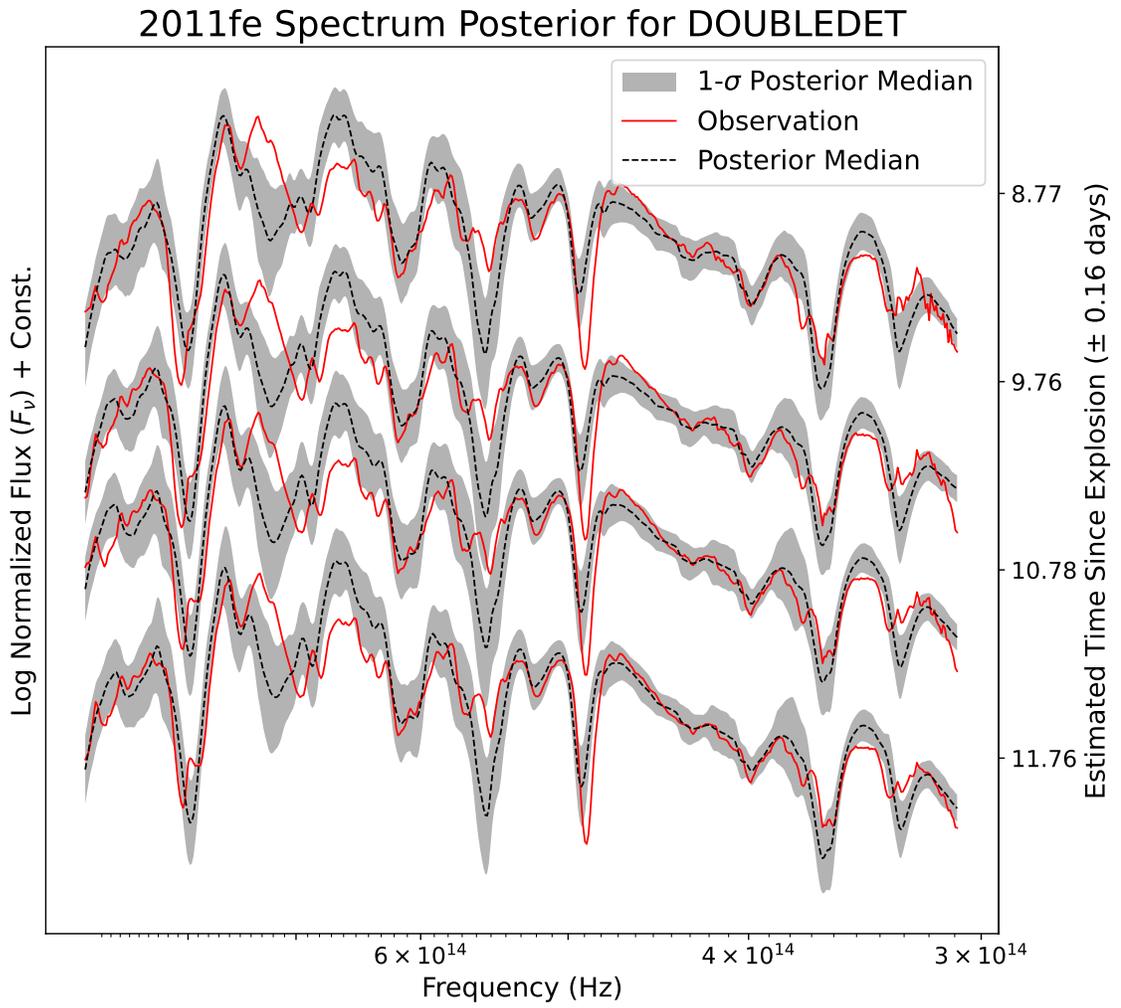
Figure F.7: Spectra from the best fit Double Detonation (Thick Helium Shell) progenitor model for SN 2011fe.
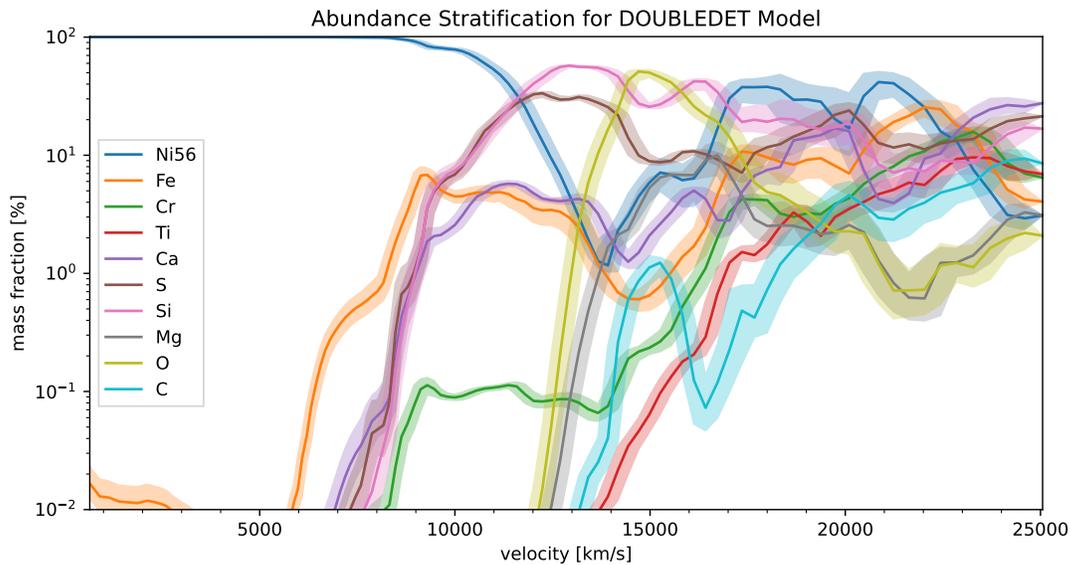
Figure F.8: Posterior distribution of reconstructed ejecta profiles for SN 2011fe under the DOUBLEDET model.

Blondin, S., & Tonry, J. L. 2007a, The Astrophysical Journal, 666, 1024, doi: `10.1086/520494`

—. 2007b, The Astrophysical Journal, 666, 1024, doi: `10.1086/520494`

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

—. 2012b, AJ, 143, 126, doi: `10.1088/0004-6256/143/5/126`

Blondin, S., Blinnikov, S., Callan, F. P., et al. 2022, A&A, 668, A163, doi: `10.1051/0004-6361/202244134`

Bloom, J. S., Kasen, D., Shen, K. J., et al. 2012, ApJL, 744, L17, doi: `10.1088/2041-8205/744/2/L17`

Boone, K., Aldering, G., Antilogus, P., et al. 2021, The Astrophysical Journal, 912, 70, doi: `10.3847/1538-4357/abec3c`

Bowman, S. R., Vilnis, L., Vinyals, O., et al. 2015, arXiv preprint arXiv:1511.06349

Boyle, A., Sim, S. A., Hachinger, S., & Kerzendorf, W. 2017, A&A, 599, A46, doi: `10.1051/0004-6361/201629712`

Branch, D. 1980, in Texas Workshop on Type I Supernovae, ed. J. C. Wheeler, 66–71

Branch, D. 1992, The Astrophysical Journal, 392, 35, doi: `10.1086/171401`

Branch, D., Doggett, J. B., Nomoto, K., & Thielemann, F. K. 1985, The Astrophysical Journal, 294, 619, doi: `10.1086/163329`

Branch, D., Fisher, A., & Nugent, P. 1993, AJ, 106, 2383, doi: `10.1086/116810`

Branch, D., & Patchett, B. 1973, Monthly Notices of the Royal Astronomical Society, 161, 71, doi: `10.1093/mnras/161.1.71`

Branch, D., Garnavich, P., Matheson, T., et al. 2003, AJ, 126, 1489, doi: `10.1086/377016`

Branch, D., Dang, L. C., Hall, N., et al. 2006, Publications of the Astronomical Society of the Pacific, 118, 560, doi: `10.1086/502778`

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

Buchner, J. 2014, arXiv e-prints, arXiv:1407.5459, doi: `10.1007/s11222-014-9512-y`

—. 2017, arXiv e-prints, arXiv:1707.04476. `https://arxiv.org/abs/1707.04476`

—. 2019, PASP, 131, 108005, doi: `10.1088/1538-3873/aae7fc`

—. 2021, arXiv e-prints, arXiv:2101.09675. `https://arxiv.org/abs/2101.09675`

Buchner, J. 2021, Journal of Open Source Software, 6, 3001, doi: `10.21105/joss.03001`

Bulla, M., Miller, A. A., Yao, Y., et al. 2020, The Astrophysical Journal, 902, 48, doi: `10.3847/1538-4357/abb13c`

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, The Astrophysical Journal, 345, 245, doi: `10.1086/167900`

Chen, N., van der Smagt, P., & Cseke, B. 2022, in Topological, Algebraic and Geometric Learning Workshops 2022, 55–66

Cohn, D. A., Ghahramani, Z., & Jordan, M. I. 1996, Journal of artificial intelligence research, 4, 129

Colgate, S. A., & McKee, C. 1969, The Astrophysical Journal, 157, 623, doi: `10.1086/150102`

De, K., Kasliwal, M. M., Polin, A., et al. 2019, ApJL, 873, L18, doi: `10.3847/2041-8213/ab0aec`

Dessart, L., Blondin, S., Hillier, D. J., & Khokhlov, A. 2014, Monthly Notices of the Royal Astronomical Society, 441, 532, doi: `10.1093/mnras/stu598`

Elias, J. H., Matthews, K., Neugebauer, G., & Persson, S. E. 1985, The Astrophysical Journal, 296, 379, doi: `10.1086/163456`

Elias-Rosa, N., Benetti, S., Cappellaro, E., et al. 2006, Monthly Notices of the Royal Astronomical Society, 369, 1880, doi: `10.1111/j.1365-2966.2006.10430.x`

Filippenko, A. V. 1997, A&A, 35, 309, doi: `10.1146/annurev.astro.35.1.309`

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

Fink, M., Kromer, M., Hillebrandt, W., et al. 2018, A&A, 618, A124, doi: `10.1051/0004-6361/201833475`

Fink, M., Röpke, F. K., Hillebrandt, W., et al. 2010a, A&A, 514, A53, doi: `10.1051/0004-6361/200913892`

—. 2010b, A&A, 514, A53, doi: `10.1051/0004-6361/200913892`

Fink, M., Kromer, M., Seitenzahl, I. R., et al. 2014, Monthly Notices of the Royal Astronomical Society, 438, 1762, doi: `10.1093/mnras/stt2315`

Finzi, A., & Wolf, R. A. 1967, The Astrophysical Journal, 150, 115, doi: `10.1086/149317`

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

Fullard, A. G., O'Brien, J. T., Kerzendorf, W. E., et al. 2022, The Astrophysical Journal, 930, 89, doi: `10.3847/1538-4357/ac589e`

Ganeshalingam, M., Li, W., & Filippenko, A. V. 2011, Monthly Notices of the Royal Astronomical Society, 416, 2607, doi: `10.1111/j.1365-2966.2011.19213.x`

Gillanders, J. H., Sim, S. A., & Smartt, S. J. 2020, Monthly Notices of the Royal Astronomical Society, 497, 246, doi: `10.1093/mnras/staa1822`

Goodman, S. N. 2005, Clinical Trials, 2, 282, doi: `10.1191/1740774505cn098oa`

Gronow, S., Collins, C., Ohlmann, S. T., et al. 2020, A&A, 635, A169, doi: `10.1051/0004-6361/201936494`

Gronow, S., Collins, C. E., Sim, S. A., & Röpke, F. K. 2021, A&A, 649, A155, doi: `10.1051/0004-6361/202039954`

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

Hachinger, S. 2011, PhD thesis, TU München

Hachinger, S., Röpke, F. K., Mazzali, P. A., et al. 2017, Monthly Notices of the Royal Astronomical Society, 471, 491, doi: `10.1093/mnras/stx1578`

Hachinger, S., Röpke, F. K., Mazzali, P. A., et al. 2017, Monthly Notices of the Royal Astronomical Society, 471, 491, doi: `10.1093/mnras/stx1578`

Harris, C. R., Millman, K. J., van der Walt, S. J., et al. 2020, Nature, 585, 357–362, doi: `10.1038/s41586-020-2649-2`

Higgins, I., Matthey, L., Pal, A., et al. 2017, ICLR

Hillier, D. J., & Lanz, T. 2001, in Astronomical Society of the Pacific Conference Series, Vol. 247, Spectroscopic Challenges of Photoionized Plasmas, ed. G. Ferland & D. W. Savin, 343

Hogg, D. W., Bovy, J., & Lang, D. 2010, arXiv e-prints, arXiv:1008.4686. `https://arxiv.org/abs/1008.4686`

Hoyle, F., & Fowler, W. A. 1960, The Astrophysical Journal, 132, 565, doi: `10.1086/146963`

Hunter, J. D. 2007, Computing in Science Engineering, 9, 90, doi: `10.1109/MCSE.2007.55`

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

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

Jiang, J.-A., Doi, M., Maeda, K., et al. 2017, Nature, 550, 80, doi: `10.1038/nature23908`

Kashi, A., & Soker, N. 2011, Monthly Notices of the Royal Astronomical Society, 417, 1466, doi: `10.1111/j.1365-2966.2011.19361.x`

Kerzendorf, W., Chen, N., O'Brien, J., Buchner, J., & van der Smagt, P. 2022, arXiv e-prints, arXiv:2209.09453, doi: `10.48550/arXiv.2209.09453`

Kerzendorf, W., Chen, N., O'Brien, J., Buchner, J., & van der Smagt, P. 2022, arXiv preprint arXiv:2209.09453

Kerzendorf, W., Sim, S., Vogl, C., et al. 2021, tardis-sn/tardis: TARDIS v3.0.dev4010, v3.0.dev4010, Zenodo, doi: `10.5281/zenodo.4995779`

Kerzendorf, W. E. 2011, PhD thesis, Australian National University, Research School of Astronomy & Astrophysics, doi: `10.25911/5d78d6c48f314`

Kerzendorf, W. E., & Sim, S. A. 2014, Monthly Notices of the Royal Astronomical Society, 440, 387, doi: `10.1093/mnras/stu055`

Kerzendorf, W. E., Vogl, C., Buchner, J., et al. 2021, ApJL, 910, L23, doi: `10.3847/2041-8213/abeb1b`

Kerzendorf, W. E., Vogl, C., Buchner, J., et al. 2021, The Astrophysical Journal Letters, 910, L23

Khokhlov, A. M. 1991, A&A, 245, 114

Kingma, D. P., & Ba, J. 2014, arXiv preprint arXiv:1412.6980

Kingma, D. P., Salimans, T., Jozefowicz, R., et al. 2016, Advances in neural information processing systems, 29

Kingma, D. P., & Welling, M. 2014, ICML

Klushyn, A., Chen, N., Kurle, R., Cseke, B., & van der Smagt, P. 2019, Advances in Neural Information processing Systems, 32

Kobayashi, C., Karakas, A. I., & Lugaro, M. 2020, The Astrophysical Journal, 900, 179, doi: `10.3847/1538-4357/abae65`

Kogure, T., & Leung, K.-C. 2007, Formation of Emission Lines (New York, NY: Springer New York), 135–186, doi: `10.1007/978-0-387-68995-1_4`

Kotak, R., Meikle, W. P. S., Pignata, G., et al. 2005, A&A, 436, 1021, doi: `10.1051/0004-6361:20052756`

Kromer, M., Ohlmann, S., & Röpke, F. K. 2017, Mem. Soc. Astron. Italiana, 88, 312. `https://arxiv.org/abs/1706.09879`

Kromer, M., Sim, S. A., Fink, M., et al. 2010, The Astrophysical Journal, 719, 1067, doi: `10.1088/0004-637X/719/2/1067`

Kromer, M., Fink, M., Stanishev, V., et al. 2013, Monthly Notices of the Royal Astronomical Society, 429, 2287, doi: `10.1093/mnras/sts498`

Kromer, M., Ohlmann, S. T., Pakmor, R., et al. 2015, Monthly Notices of the Royal Astronomical Society, 450, 3045, doi: `10.1093/mnras/stv886`

Kurucz, R., & Bell, B. 1995, Atomic Line Data (R.L. Kurucz and B. Bell) Kurucz CD-ROM No. 23. Cambridge, 23

Lach, F., Callan, F. P., Bubeck, D., et al. 2022, A&A, 658, A179, doi: `10.1051/0004-6361/202141453`

Lakshminarayanan, B., Pritzel, A., & Blundell, C. 2017, Advances in neural information processing systems, 30

Lam, S. K., Pitrou, A., & Seibert, S. 2015, in Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC, LLVM '15 (New York, NY, USA: Association for Computing Machinery), doi: `10.1145/2833157.2833162`

Liu, W., Jeffery, D. J., & Schultz, D. R. 1997, ApJL, 486, L35, doi: `10.1086/310832`

Livio, M., & Riess, A. G. 2003, The Astrophysical Journal, 594, L93, doi: `10.1086/378765`

Lucy, L. B. 1999a, A&A, 345, 211

—. 1999b, A&A, 345, 211

—. 1999c, A&A, 345, 211

—. 2002, A&A, 384, 725, doi: `10.1051/0004-6361:20011756`

Magee, M. R., Gillanders, J. H., Maguire, K., Sim, S. A., & Callan, F. P. 2021, Monthly Notices of the Royal Astronomical Society, 509, 3580, doi: `10.1093/mnras/stab3123`

Magee, M. R., Kotak, R., Sim, S. A., et al. 2016, A&A, 589, A89, doi: `10.1051/0004-6361/201528036`

Marquardt, K. S., Sim, S. A., Ruiter, A. J., et al. 2015, A&A, 580, A118, doi: `10.1051/0004-6361/201525761`

Matheson, T., Kirshner, R. P., Challis, P., et al. 2008, AJ, 135, 1598, doi: `10.1088/0004-6256/135/4/1598`

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

Mazzali, P. A., & Lucy, L. B. 1993, A&A, 279, 447

Mazzali, P. A., Röpke, F. K., Benetti, S., & Hillebrandt, W. 2007, Science, 315, 825, doi: `10.1126/science.1136259`

Mazzali, P. A., Sauer, D. N., Pastorello, A., Benetti, S., & Hillebrandt, W. 2008, Monthly Notices of the Royal Astronomical Society, 386, 1897, doi: `10.1111/j.1365-2966.2008.13199.x`

Mazzali, P. A., Sullivan, M., Filippenko, A. V., et al. 2015, Monthly Notices of the Royal Astronomical Society, 450, 2631, doi: `10.1093/mnras/stv761`

Minkowski, R. 1941, PASP, 53, 224, doi: `10.1086/125315`

Murakami, Y. S., Riess, A. G., Stahl, B. E., et al. 2023, Journal of Cosmology and Astroparticle Physics, 2023, 046, doi: `10.1088/1475-7516/2023/11/046`

Neal, R., & Hinton, G. 1998, Learning in graphical models, 89, 355. `http://www.cs.toronto.edu/~radford/ftp/emk.pdf`

Noebauer, U. M., Kromer, M., Taubenberger, S., et al. 2017, Monthly Notices of the Royal Astronomical Society, 472, 2787, doi: `10.1093/mnras/stx2093`

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

Nomoto, K. 1982, The Astrophysical Journal, 257, 780, doi: 10.1086/160031

Nomoto, K., & Sugimoto, D. 1977, PASJ, 29, 765

Nomoto, K., Sugimoto, D., & Neo, S. 1976, APSS, 39, L37, doi: 10.1007/BF00648354

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

Nugent, P., Kim, A., & Perlmutter, S. 2002, Publications of the Astronomical Society of the Pacific, 114, 803, doi: 10.1086/341707

Nugent, P. E., Sullivan, M., Cenko, S. B., et al. 2011, Nature, 480, 344, doi: 10.1038/nature10644

O'Brien, J. T., Kerzendorf, W. E., Fullard, A., et al. 2021, ApJL, 916, L14, doi: 10.3847/2041-8213/ac1173

—. 2023, arXiv e-prints, arXiv:2306.08137, doi: 10.48550/arXiv.2306.08137

Ohlmann, S. T., Kromer, M., Fink, M., et al. 2014, A&A, 572, A57, doi: 10.1051/0004-6361/201423924

Oke, J. B., & Searle, L. 1974, A&A, 12, 315, doi: 10.1146/annurev.aa.12.090174.001531

O'Brien, J. T., Kerzendorf, W. E., Fullard, A., et al. 2021, The Astrophysical Journal Letters, 916, L14, doi: 10.3847/2041-8213/ac1173

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

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

Pakmor, R., Callan, F. P., Collins, C. E., et al. 2022, Monthly Notices of the Royal Astronomical Society, 517, 5260, doi: `10.1093/mnras/stac3107`

Pakmor, R., Callan, F. P., Collins, C. E., et al. 2022, Monthly Notices of the Royal Astronomical Society, 517, 5260, doi: `10.1093/mnras/stac3107`

Parrent, J. T., Howell, D. A., Friesen, B., et al. 2012, ApJL, 752, L26, doi: `10.1088/2041-8205/752/2/L26`

Paszke, A., Gross, S., Massa, F., et al. 2019, in Advances in Neural Information Processing Systems 32 (Curran Associates, Inc.), 8024–8035. `http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf`

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

Pereira, R., Thomas, R. C., Aldering, G., et al. 2013, A&A, 554, A27, doi: `10.1051/0004-6361/201221008`

Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999, The Astrophysical Journal, 517, 565, doi: `10.1086/307221`

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

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

Phillips, M. M., Ashall, C., Burns, C. R., et al. 2022, The Astrophysical Journal, 938, 47, doi: `10.3847/1538-4357/ac9305`

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

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

Polin, A., Nugent, P., & Kasen, D. 2021, The Astrophysical Journal, 906, 65, doi: `10.3847/1538-4357/abcccc`

Popper, D. M. 1937, PASP, 49, 283, doi: `10.1086/124853`

Pássaro, E. A., E., K. W., A., F., & C., V. 2019, An Open Source Framework for Atomic Data, Zenodo, doi: `10.5281/zenodo.4062427`

Ralchenko, Y. 2005, Memorie della Societa Astronomica Italiana Supplementi, 8, 96

Rezende, D. J., Mohamed, S., & Wierstra, D. 2014, in ICML, Vol. 32, 1278–1286

Rezende, D. J., & Viola, F. 2018, CoRR

Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, AJ, 116, 1009, doi: `10.1086/300499`

Riess, A. G., Filippenko, A. V., Li, W., et al. 1999, The Astronomical Journal, 118, 2675, doi: `10.1086/301143`

Röpke, F. K., Kromer, M., Seitenzahl, I. R., et al. 2012, ApJL, 750, L19, doi: `10.1088/2041-8205/750/1/L19`

Rosseland, S. 1924, Monthly Notices of the Royal Astronomical Society, 84, 525, doi: `10.1093/mnras/84.7.525`

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

Röpke, F. K., Hillebrandt, W., Schmidt, W., et al. 2007, The Astrophysical Journal, 668, 1132, doi: `10.1086/521347`

Sasdelli, M., Mazzali, P. A., Pian, E., et al. 2014, Monthly Notices of the Royal Astronomical Society, 445, 711, doi: `10.1093/mnras/stu1777`

Sauer, D., & Mazzali, P. 2008, New Astronomy Reviews, 52, 370, doi: `https://doi.org/10.1016/j.newar.2008.06.023`

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

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

Seitenzahl, I. R., Ciaraldi-Schoolmann, F., Röpke, F. K., et al. 2013, Monthly Notices of the Royal Astronomical Society, 429, 1156, doi: `10.1093/mnras/sts402`

Seitenzahl, I. R., Kromer, M., Ohlmann, S. T., et al. 2016, A&A, 592, A57, doi: `10.1051/0004-6361/201527251`

Shen, K. J., Boos, S. J., Townsley, D. M., & Kasen, D. 2021, The Astrophysical Journal, 922, 68, doi: `10.3847/1538-4357/ac2304`

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

Silverman, J. M., Foley, R. J., Filippenko, A. V., et al. 2012, Monthly Notices of the Royal Astronomical Society, 425, 1789, doi: `10.1111/j.1365-2966.2012.21270.x`

Sim, S. A., Fink, M., Kromer, M., et al. 2012, Monthly Notices of the Royal Astronomical Society, 420, 3003, doi: `10.1111/j.1365-2966.2011.20162.x`

Sim, S. A., Röpke, F. K., Hillebrandt, W., et al. 2010, ApJL, 714, L52, doi: `10.1088/2041-8205/714/1/L52`

Skilling, J. 2004, in American Institute of Physics Conference Series, Vol. 735, Bayesian Inference and Maximum Entropy Methods in Science and Engineering: 24th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, ed. R. Fischer, R. Preuss, & U. V. Toussaint, 395–405, doi: `10.1063/1.1835238`

Sohn, K., Lee, H., & Yan, X. 2015, Advances in neural information processing systems, 28

Sønderby, C. K.and Raiko, T., Maaløe, L., Sønderby, S. K., & Winther, O. 2016, NeurIPS

Stehle, M., Mazzali, P. A., Benetti, S., & Hillebrandt, W. 2005, Monthly Notices of the Royal Astronomical Society, 360, 1231, doi: `10.1111/j.1365-2966.2005.09116.x`

Taam, R. E. 1980, The Astrophysical Journal, 242, 749, doi: `10.1086/158509`

Taubenberger, S. 2017, in Handbook of Supernovae, ed. A. W. Alsabti & P. Murdin, 317, doi: `10.1007/978-3-319-21846-5_37`

Tonry, J., & Davis, M. 1979, AJ, 84, 1511, doi: `10.1086/112569`

van Kerkwijk, M. H., Chang, P., & Justham, S. 2010, The Astrophysical Journal, 722, L157, doi: `10.1088/2041-8205/722/2/l157`

Virtanen, P., Gommers, R., Oliphant, T. E., et al. 2020, Nature Methods, 17, 261, doi: `10.1038/s41592-019-0686-2`

Vogl, C., Kerzendorf, W. E., Sim, S. A., et al. 2020a, A&A, 633, A88, doi: `10.1051/0004-6361/201936137`

—. 2020b, A&A, 633, A88, doi: `10.1051/0004-6361/201936137`

Wang, X., Li, W., Filippenko, A. V., et al. 2009, The Astrophysical Journal, 697, 380, doi: `10.1088/0004-637X/697/1/380`

Webbink, R. F. 1984, The Astrophysical Journal, 277, 355, doi: `10.1086/161701`

Wes McKinney. 2010, in Proceedings of the 9th Python in Science Conference, ed. Stéfan van der Walt & Jarrod Millman, 56 – 61, doi: `10.25080/Majora-92bf1922-00a`

Wheeler, J. C. 1981, Reports on Progress in Physics, 44, 85, doi: `10.1088/0034-4885/44/2/001`

Whelan, J., & Iben, Icko, J. 1973, The Astrophysical Journal, 186, 1007, doi: `10.1086/152565`

Williamson, M., Kerzendorf, W., & Modjaz, M. 2021, The Astrophysical Journal, 908, 150, doi: `10.3847/1538-4357/abd244`

Woosley, S. E., & Weaver, T. A. 1994, The Astrophysical Journal, 423, 371, doi: `10.1086/173813`

Woosley, S. E., Weaver, T. A., & Taam, R. E. 1980, in Texas Workshop on Type I Supernovae, ed. J. C. Wheeler, 96–112

Yamanaka, M., Naito, H., Kinugasa, K., et al. 2009, PASJ, 61, 713, doi: `10.1093/pasj/61.4.713`

Yaron, O., & Gal-Yam, A. 2012, PASP, 124, 668, doi: `10.1086/666656`

Zhang, K., Wang, X., Zhang, J., et al. 2016, The Astrophysical Journal, 820, 67, doi: `10.3847/0004-637X/820/1/67`

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