Draft version December 30, 2023
Typeset using LaTeX **twocolumn** style in AASTeX631

Using Physics Informed Neural Networks for Supernova Radiative Transfer Simulation

Xingzhuo Chen,[1] David J. Jeffery,[2] Ming Zhong,[3] Levi McClenny,[4] Ulisses Braga-Neto,[4] and Lifan Wang[1]

[1] *George P. and Cynthia Woods Mitchell Institute for Fundamental Physics & Astronomy, Texas A. & M. University, Department of Physics and Astronomy, 4242 TAMU, College Station, TX 77843, USA*
[2] *Department of Physics & Astronomy and Nevada Center for Astrophysics (NCfA), University of Nevada, Las Vegas, Nevada, U.S.A.*
[3] *Texas A. & M. Institute of Data Science, Texas A. & M. University.*
[4] *Department of Electrical and Computer Engineering, Texas A. & M. University.*

(Revised January 1, 2024)

ABSTRACT

We use physics informed neural networks (PINNs) to solve the radiative transfer equation and calculate a synthetic spectrum for a Type Ia supernova (SN Ia) SN 2011fe. The calculation is based on local thermodynamic equilibrium (LTE) and 9 elements are included. Physical processes included are approximate radiative equilibrium, bound-bound transitions, and the Doppler effect. A PINN based gamma-ray scattering approximation is used for radioactive decay energy deposition. Note the physics ingredients are intended to implement a self-consistent SN Ia atmosphere to test PINN radiative transfer (including gamma-ray radiative transfer). The realism of the SN Ia atmosphere modeling is limited. The PINN synthetic spectrum is compared to an observed spectrum, a synthetic spectrum calculated by the Monte Carlo radiative transfer program TARDIS, and the formal solution of the radiative transfer equation. Qualitative agreement is achieved. The lack of quantitative agreement with the formal solution (which is the only test quantitative test of PINN radiative transfer) probably shows that we have not found an adequate way to apply PINN in supernova atmospheres. We discuss the challenges and potential of PINN radiative transfer. In fact, PINN offers the prospect of simultaneous solution of the atmosphere problem for both radiation field and thermal state throughout spacetime. We have made only modest steps to realizing that prospect with our calculations which required many approximations in order to be feasible at this point. Consequently, this paper is mostly of use just as supplementary material for future work.

Keywords: supernovae: general, galaxies: star formation

1. INTRODUCTION

Type Ia supernovae (SNe Ia) have been used as standard candles in cosmological studies (Riess et al. 2021) owing to empirical relations between the light curve properties and the maximum absolute magnitude (e.g., the Phillips relation (Phillips 1993) and Arnett rule (Arnett 1982)). However, the explosion mechanism of SNe Ia is still unclear, primarily due to the computational complexity of the physical processes, particularly nucleosynthesis and hydrodynamics needed in the supernova explosion simulation (Gronow et al. 2021).

Corresponding author: Lifan Wang
lifan@tamu.edu

In SNe Ia, hydrodynamic and nucleosynthesis processes are only significant in the first ∼100 seconds. Thereafter the supernova ejecta expands homologously and the observed optical spectra and light curves are generated by radiative transfer and the thermal state of the ejecta. Therefore, simulating the radiative transfer process is necessary to estimate the density profile and element abundances of the supernova ejecta so as to put constraints on the SNe Ia explosion mechanism. Several well known simulation programs have been developed for the calculation of synthetic spectra for explosion model supernova ejecta structure. SYNOW (Parrent et al. 2010; Thomas et al. 2011) uses the Sobolev method for radiative transfer calculation (e.g., Rybicki & Hummer 1978) and has been used for spectral line identification. PHOENIX (Hauschildt & Baron 2006) and CMFGEN (Hillier & Miller 1998) are more advanced simulation programs and are able to calculate non-local-thermodynamic-equilibrium (NLTE), time-dependent radiative transfer using the comoving frame equation of radiative transfer (e.g., Mihalas 1978, p. 490ff). Note the PHOENIX code has a 3-dimensional version PHOENIX/3D (Hauschildt & Baron 2006). Programs using the Monte Carlo method have also been developed for spectral simulation (e.g., SEDONA (Kasen et al. 2006), ARTIS (Kromer & Sim 2009)) and have been used for spectral polarization calculations (e.g., Kasen et al. 2006; Bulla et al. 2015; Livneh & Katz 2022). In particular, TARDIS (Kerzendorf & Sim 2014) is a one-dimensional radiative transfer program using the Monte Carlo method (of radiative transfer) in which several crude approximations for NLTE effects have been implemented. The research reported in this paper uses the spectra from TARDIS for comparison.

Although the aforementioned supernova spectrum simulation programs can provide results with different levels of approximation within reasonable amounts of computation time, the inverse problem, which estimates the supernova ejecta structure from an observed spectrum, still requires significant computational resources. In our previous study (Chen et al. 2020), a solution of the inverse problem is obtained by training a data-driven neural network on a simulated spectra data set, which contains 100,000 supernova spectra of different ejecta structure and costs ∼1,000,000 CPU-hours of computation time for spectral simulation and neural network training. Similarly, Kerzendorf et al. (2021) uses a data-driven neural network to accelerate the calculation of the forward modeling problem, and suggest the neural network could combine with the nested sampling algorithm (Buchner 2016) to solve the inverse problem.

Physics Informed Neural Networks (PINNs) have emerged recently as a powerful addition to traditional numerical partial differential equation (PDE) solvers (Raissi et al. 2019; Karniadakis et al. 2021). The PINN approach is based on constraining the output of a deep neural network to satisfy a physical model specified by a PDE. Using neural networks as universal function approximators to solve PDEs had been proposed already in the 1990's (Dissanayake & Phan-Thien 1994; Lagaris et al. 1998). PINN capabilities at solving PDEs have been enhanced in many different ways since then by utilizing the expressive powers of deep neural networks, which are made possible by the recent advances in GPU-computing and training algorithms (Abadi et al. 2016), as well as computational advances in automatic differentiation methods (Baydin et al. 2017). A significant advantage of PINNs over traditional time-stepping PDE solvers is that PINNs are mesh-less and can solve in space and time simultaneously. Combined with the regression capability of deep neural networks, PINNs are also suitable for PDE-related inverse problems.

Note that in Mishra & Molinaro (2021), PINNs were applied to solve several simple monochromatic and polychromatic radiative transfer problems.

In this paper, we employ PINNs to solve the radiative transfer equation (e.g., Hubeny & Mihalas 2014) in order to calculate an optical spectrum of SN Ia SN 2011fe at 12.35 days after explosion. The overall calculation is a self-consistent atmosphere solution of both radiative transfer and thermal state solution in order to test PINNs in the context of a qualitatively realistic self-consistent atmosphere solution. However, there is limited quantitative realism in many of the ingredients as we detail primarily in § 2. We are not doing a state-of-the-art SN Ia atmosphere modeling.

The paper is structured as follows. Section 2 introduces the theoretical background, including the optical radiative transfer equation, the atomic physics calculation method, and the approximate gamma-ray radiative transfer calculation method. Section 3 describes the PINN structure used in this research and the results from the PINN calculation. A summary and a discussion of the future challenges for PINN-based radiative transfer calculations are given in § 4. Appendix A presents the formal solution of the radiative transfer equation. The code used in this research is available on https://github.com/GeronimoChen/RTPI.

2. THE RADIATIVE TRANSFER EQUATION

In spherical symmetric coordinates, the time-independent radiative transfer equation in the rest frame (i.e., the frame defined by the center of mass of the spherically symmetric system) is

$$\cos(\varphi)\frac{\partial I}{\partial r} - \sin(\varphi)\frac{\partial I}{\partial \varphi}\frac{1}{r} - j_{\rm em}\left(\frac{\nu}{\bar{\nu}}\right)^{-2} + k_{\rm abs}\left(\frac{\nu}{\bar{\nu}}\right) I = 0 \ , \tag{1}$$

where I is specific intensity (here a function of spatial coordinate, viewing direction, and frequency), r is radius, φ is the angle between the viewing direction and the radius vector (i.e., the viewing angle), $k_{\rm abs}$ is the comoving frame opacity (not the rest frame opacity), $j_{\rm em}$ is the comoving frame emissivity (not the rest frame emissivity), and $\left(\frac{\nu}{\bar{\nu}}\right)$ is the ratio of comoving frame frequency to rest frame frequency, which is given by

$$\frac{\nu}{\bar{\nu}} = \gamma[1-\cos(\varphi)\beta] \ , \tag{2}$$

where $\gamma = (1-\beta^2)^{-0.5}$ is the Lorentz factor and $\beta = v/c$ is the velocity of the material divided by the speed of light (Castor e.g., 1972, eq. (1–3); see also Mihalas 1978, p. 31,33,495–496). The formal solution of the radiative transfer equation is presented in Appendix A.

Note we have dropped the time dependence term because we model only the atmosphere of a SN Ia above an inner core that provides inner boundary condition for the atmosphere. Time dependence in SN Ia atmospheres has generally been found to be relatively unimportant near and even prior to maximum light (e.g., Kasen et al. 2006, § 3.5) and can be neglected for our exploratory calculations.

Because SN Ia ejecta is expanding homologously by about the first 10 seconds after the explosion of the relatively small progenitor white dwarf (e.g., Röpke & Hillebrandt 2005), the radial velocity at all observable epochs is proportional to the radius at a given time and satisfies the relation $r = vt_{\rm exp} = c\beta t_{\rm exp}$, where $t_{\rm exp}$ is the time since the explosion. Therefore, we use the radial velocity to represent the radial coordinate in the figures and elsewhere as needed.

Note because of homologous expansion, surfaces of constant velocity $\beta_\varphi = \cos(\varphi)\beta$ in the direction to a distant observer obey

$$\begin{aligned}\beta_\varphi &= \cos(\varphi)\beta = \cos(\varphi)\frac{r}{ct_{\rm exp}} = \cos(\varphi)\frac{z}{\cos(\varphi)ct_{\rm exp}} \\ &= \frac{z}{ct_{\rm exp}} \ , \end{aligned} \tag{3}$$

where z is a constant length along beam paths to a distant observer. Because z is a constant, the surfaces of constant velocity β_φ in the direction to a distant observer are planes perpendicular to those beam paths. The plane for $z = 0$ passes through the center of mass of the SN Ia ejecta.

In the following subsections, we will introduce the different physical processes that contribute to the opacity $k_{\rm abs}$ and emissivity $j_{\rm em}$.

2.1. *Thermal State, LTE, and Temperature Profile*

For the calculation of the thermal state of our SN Ia atmosphere model, we assume local thermodynamic equilibrium (LTE) which means that all matter occupation numbers are determined by their thermodynamic equilibrium values calculated from a single temperature. Imposing conservation of energy including both radiation field and gamma-ray energy deposition (calculated as described in §§ 2.4 and 3.1), the LTE temperature T is calculated from the quasi steady state first law of thermodynamics equation with time derivatives of energy density and adiabatic cooling omitted as being negligible which is suitable for supernova atmospheres (e.g., Kasen et al. 2006, § 2.4). This equation written for our model is the radiative equilibrium equation (modified for gamma-ray energy deposition)

$$\int_0^\infty \kappa_\nu B_\nu(T)\, d\nu = \int_0^\infty \kappa_\nu J_\nu\, d\nu + \mathcal{E}_\gamma \ , \tag{4}$$

where ν is the comoving frequency, T is the LTE temperature to be solved for, κ_ν is the comoving absorption opacity (implicitly evaluated at T),

$$B_\nu(T) = \frac{2h\nu^3}{c^2}\frac{1}{e^{h\nu/(k_{\rm B}T)} - 1} \tag{5}$$

is the Planck law (with $k_{\rm B}$ being the Boltzmann constant),

$$J_\nu = \frac{1}{4\pi}\int_\Omega I\, d\Omega \tag{6}$$

is the comoving mean specific intensity (with the frequency dependence of I implicit which is the convention we adopt in this paper), and $\mathcal{E}_\gamma$ is the (rate of) gamma-ray energy deposition per unit solid angle (e.g., Mihalas 1978, p. 172: see also Kasen et al. 2006, § 2.4). Note

$$\mathcal{E}_\gamma = \frac{E_\gamma}{4\pi} \ , \tag{7}$$

where E_γ is the (rate of) gamma-ray energy deposition specified in § 2.4.

Using

$$B(T) = \frac{\sigma_{\rm SB}T^4}{\pi} \ , \tag{8}$$

the Planck law integrated over all frequency (with $\sigma_{\rm SB}$ being the Stefan-Boltzmann constant), we can rewrite Equation (4) in a form clearly implying there is a temperature to be solved for:

$$T = \left[\frac{\pi J}{\sigma_{\rm SB}}R_{\rm op} + \frac{\pi\mathcal{E}_\gamma}{\sigma_{\rm SB}\kappa_{\rm P}(T)}\right]^{1/4} \ , \tag{9}$$

where J is the mean intensity integrated over all frequency, $\kappa_{\rm P}(T)$ is the Planck mean opacity (see definition in the equation just below), and

$$R_{\rm op} = \frac{\kappa_J}{\kappa_{\rm P}(T)} = \frac{\int_0^\infty \kappa_\nu J_\nu\, d\nu/J}{\int_0^\infty \kappa_\nu B_\nu(T)\, d\nu/B(T)} \tag{10}$$

is the opacity ratio, where κ_J is the absorption mean opacity (e.g., Mihalas 1978, p. 60) and $\kappa_{\rm P}(T)$ is the aforementioned Planck mean opacity (e.g., Mihalas 1978, p. 59). Note that the opacity ratio is independent of the scales of J_ν, B_ν, and κ_ν. The independence of the scale of J_ν in the definition of the opacity ratio is just because J (loosely speaking the total driving radiation field) has been made to appear explicitly in Equation (9). The independence of the scale of B_ν just follows from the definition of the Planck mean opacity and that of κ_ν from the definition of the opacity ratio.

Note when $\mathcal{E}_\gamma = 0$, the LTE radiative equilibrium temperature itself is independent of the scale of κ_ν as seen from both Equations (4) and (9). This independence can be understood by seeing that when $\mathcal{E}_\gamma = 0$, the LTE radiative equilibrium temperature is independent of the scales of energy inflow and outflow to matter and these scales are the only things controlled by the scale of κ_ν.

The independence of the scales of J_ν, B_ν, and κ_ν in the opacity ratio $R_{\rm op}$ suggests that the opacity ratio can be of order 1 though it can also be very different from 1 if the shapes of J_ν and B_ν are very different which certainly happens in some atmosphere conditions. We show below that there are two limits where $R_{\rm op}$ does equal 1 exactly.

Equation (9) is, of course, still an implicit equation for LTE temperarure T. However, it will probably succeed as an iteration formula in most cases. However also, given that the opacity ratio $R_{\rm op}$ can be of order 1 and 1 is the natural choice in the absence of guiding information, we can set it to be 1 for two characteristic LTE temperatures derived from Equation (9). These are explicit solutions relative to the local thermal state, but depend on the radiation field, and so are implicit relative to the overall atmosphere. The first characteristic LTE temperature, which we call the Planck law temperature (PLT), is

$$T_{\rm PLT} = \left(\frac{\pi J}{\sigma_{\rm SB}}\right)^{1/4} , \tag{11}$$

where $\mathcal{E}_\gamma$ appearing in Equation (9) is set to zero. The second characteristic LTE temperature, which we call the Planck law temperature augmented (PLTA), is

$$\begin{aligned} T_{\rm PLTA} &= \left[\frac{\pi J}{\sigma_{\rm SB}} + \frac{\pi \mathcal{E}_\gamma}{\sigma_{\rm SB}\kappa_{\rm P}(T_{\rm PLT})}\right]^{1/4} \\ &= \left[T_{\rm PLT}^4 + \frac{\pi \mathcal{E}_\gamma}{\sigma_{\rm SB}\kappa_{\rm P}(T_{\rm PLT})}\right]^{1/4} , \end{aligned} \tag{12}$$

where $\mathcal{E}_\gamma$ appearing in Equation (9) is not set to zero.

There are two special cases where PLT is an exact explicit solution for the LTE temperature equation (i.e., Equation (9)) with $\mathcal{E}_\gamma = 0$. The two cases, of course, have $R_{\rm op} = 1$ exactly. The first case is in the optically thick limit where photons travel negligibly short distances compared to distances over which the thermal state changes. This case is usually at great optical depth in an atmosphere. Given the optically thick limit, $J_\nu = B_\nu(T)$, where T is the local temperature. It now follows from Equation (10) that $R_{\rm op} = 1$ exactly. The second case is where the absorption opacity is grey (i.e., frequency independent, but not necessarily independent of any other variable). It again follows from Equation (10) that $R_{\rm op} = 1$ exactly. The first case will actually be approached closely in realistic atmospheres in optically thick conditions and the second may be a good approximation in some cases.

As well as the two exact cases of PLT, there can be other cases not close to the two exact cases where PLT holds approximately provided that fortuitously the factors in Equation (10) multiply to 1. However, obviously there are realistic cases where PLT will be wrong by an order of magnitude or more. Extremely wrong, but unrealistic, cases are where the absorption opacity is a Dirac delta function. Given this opacity, it is clear that for extreme choices of J_ν and B_ν, the opacity ratio will have the range $R_{\rm op} \in [0, \infty]$.

An important fact about PLT is that it does conserve energy locally even if it gives a very wrong temperature. In Lambda iteration using the radiative transfer equation, local energy conservation is not sufficient to guarantee convergence and in fact in optically thick atmospheres generally fails in practice (e.g., Mihalas 1978, p. 147–150). However in Monte Carlo radiative transfer with indestructible photon packets and local energy conservation, the Lambda iteration converges robustly since the use of indestructible photon packets insures global conservation of energy (Lucy 1999; Kasen et al. 2006). Thus, PLT could be useful in determining a first or early iteration LTE temperature profile in a Lambda iteration since it avoids doing the integrations of Equation (4).

To conclude about PLT, despite the fact that PLT can be extremely wrong, it is still a reasonable choice for a characteristic LTE temperature for exploratory, example, or first or early Lambda iteration calculations given that it is exactly correct in two limits, may be fortuitously correct in other cases, conserves energy locally, and setting $R_{\rm op} = 1$ is, as aforesaid, the natural choice in the absence of guiding information. We can then conclude PLTA is therefore also a reasonable characteristic LTE temperature when $\mathcal{E}_\gamma$ is only 1st order correction to the PLT temperature (i.e., to the LTE thermal state). Note PLTA does not conserve energy locally exactly, but only to 1st order in small $\mathcal{E}_\gamma$.

An important point to make about PLT/PLTA is that if there are significant NLTE effects in an atmosphere, PLT/PLTA is, a priori, as good a characteristic temperature as the one obtained by solving Equation (4) (or, equivalently, Equation (9)) exactly since it is after all just an LTE equation and NLTE effects could make it as poor an approximate as PLT/PLTA.

For our pioneering PINN radiative transfer calculations, we do not need high realism in the thermal state solution (which for LTE is essentially solving for the temperature profile) and in particular do not need exact energy conservation. (Note we do not have radiative equilibrium in the exact sense of the expression because of the contribution of $\mathcal{E}_\gamma$ though that contribution turned out to be small in our calculations.) Therefore, we adopted the PLT formula to treat the radiation field contribution to temperature, but do a Lambda iteration for the LTE temperature (see procedure description below) and we made a simplifying approximation for the Planck mean opacity $\kappa_{\rm P}(T)$: we take it to be equal to the electron scattering opacity

$$k_e = \sigma_{\rm T} N_e \ , \tag{13}$$

where $\sigma_{\rm T}$ is the Thomson scattering cross section and N_e is free electron number density. Because N_e depends on the local LTE temperature, our equation for the local LTE temperature will be implicit relative to the local thermal state unlike Equations (11) and (12) for, respectively, PLT and PLTA.

Finally, the implicit LTE temperature formula adopted for our calculations (derived from Equation (9) and the assumptions given in the last paragraph) is

$$\begin{aligned} T_{\rm LTE} &= \left(\frac{\pi J}{\sigma_{\rm SB}} + \frac{\pi \mathcal{E}_\gamma}{\sigma_{SB} \sigma_{\rm T} N_e} \right)^{1/4} \\ &= \left(T_{\rm PLT}^4 + \frac{\pi \mathcal{E}_\gamma}{\sigma_{SB} \sigma_{\rm T} N_e} \right)^{1/4} \ , \end{aligned} \tag{14}$$

where the implicit $T_{\rm LTE}$ is hidden in the calculation of the electron density N_e.

It is, of course, formally wrong to use a scattering opacity as an absorption opacity. However, the usage is a calculational placeholder for a better treatment using a good Planck mean opacity from tables or a good approximate formula. Also, the electron scattering opacity is, in effect, just used as weighting for the gamma-ray energy deposition which our calculations show to be a relatively small contribution to temperature. Of course, the electron scattering opacity could be an overestimate/underestimate, and so could underestimate/overestimate in our calculations the effect of the gamma-ray energy deposition. Note that we also solve the gamma-ray radiative transfer with PINN to show that that can be done (see § 3.1).

The actual solution procedure for the temperature profile using the implicit LTE temperature equation (i.e., Equation (14)) is, as aforesaid, by the Lambda iteration (e.g., Mihalas 1978, p. 147–150). We assume an initial temperature profile from which we calculate the electron densities N_e and the bound-bound opacities (i.e., bb or line opacities) which are the only opacities we include in the radiative transfer (see 2.2): the electron scattering opacity is used only to solve for temperature from Equation (14). The bound-bound opacities are treated as pure absorption opacities: i.e., no line scattering is included and the lines emit thermally (see 2.2). Using the bound-bound opacities and emission from the inner boundary, we calculate J from the radiative transfer and then use Equation (14) to calculate a new temperature profile from which new electron densities N_e and new bound-bound opacities are calculated. We then calculate the radiative transfer again, and so on until the temperature profile converges. The Lambda iteration successfully converges in our calculations since the atmosphere is overall optically thin with just the bound-bound opacities included (e.g., Mihalas 1978, p. 147–150).

To speed up the Lambda iteration, we avoid the integrations for J by using a temperature (neural) network as described in § 3.3.

2.2. *Bound-Bound Transitions*

The bound-bound opacity and emissivity are calculated using the local thermodynamic equilibrium (LTE) approximation. The spontaneous emissivity is

$$j_{\rm bb} = \frac{h}{4\pi} A_{ul} \nu_{ul} N_{i,j,u} \phi(\nu) \ , \tag{15}$$

where A_{ul} is the Einstein A coefficient for the bound-bound transition from the u-th level to the l-th level, $N_{i,j,u}$ is the number density, i, j, u are the indices of element, ionization, and energy level, respectively, ν_{ul} is the spectral line frequency, and $\phi(\nu)$ is the line shape profile which satisfies the normalization condition $\int_0^\infty \phi(\nu)\, d\nu = 1$ (e.g., Mihalas 1978, p. 78). The bound-bound opacity corrected for stimulated emission is

$$k_{\rm bb} = \frac{h}{4\pi} (N_{i,j,l} B_{lu} - N_{i,j,u} B_{ul}) \nu_{ul} \phi(\nu) \ , \tag{16}$$

where B_{ul} and B_{lu} are the Einstein B coefficients for the absorption and stimulated emission processes (e.g., Mihalas 1978, p. 78–79). All the Einstein coefficients are downloaded from NIST spectral database.

The realistic line profile $\phi(\nu)$ is usually the Voigt function which accounts for both natural line broadening and temperature Doppler broadening (e.g., Mihalas 1978, p. 279–281). These broadening effects are much smaller than the Doppler effect from the supernova ejecta velocity (typically $\sim 10000\,$km/s), and so we replace the usual realistic line profile with an artificial unrealistic one without significant error as long as it likewise has insignificant broadening. We choose a simple rectangular function with a 4 pixel width as the line shape profile in order to reduce the computation time. When the frequency grid is between $10^{14.4}\,$Hz (12000 Å) and $10^{15}\,$Hz (3000 Å) with 2048 sampling points uniformly sampled in the logarithmic space, the velocity resolution is 202 km/s and spectral line width is 808 km/s (which is much smaller ~ 10000 km/s). Note that the temperature Doppler broadening velocity in supernova ejecta (which typically have temperatures of order 10^4 K) is of order 10 km/s (e.g., Mihalas 1978, p. 279), and so we have introduced artificial line broadening large compared to temperature Doppler broadening, but still negligible for our calculations.

2.3. *Level Populations*

The level populations are calculated in LTE using Saha ionization equation and the Boltzmann equation. The relevant equations for solving for ionization state and electron density are as follows. First, the ratio between the two level populations (i.e., the Saha ionization equation in one version) is

$$\frac{N_{i,j,k}}{N_{i,j+1,0}} = N_e \frac{1}{2}\left(\frac{h^2}{2\pi m_e k_\mathrm{B}}\right)^{\frac{3}{2}} \frac{g_{i,j,k}}{g_{i,j+1,0}} T^{-\frac{3}{2}} e^{\frac{\chi_{i,j+1,0}-\chi_{i,j,k}}{k_\mathrm{B}T}} , \tag{17}$$

where N_e is the electron number density, $g_{i,j,k}$ is the degeneracy factor, $\chi_{i,j,k}$ is the level energy, i, j, k are as before the indices of element (which is the atomic number), ionization, and level, respectively, and 0 labels the ground state level (e.g., Mihalas 1978, p. 113). Second, the supernova ejecta plasma is assumed to be neutral, and so the electron number density satisfies

$$N_e = \sum_i \sum_j j \sum_k N_{i,j,k} . \tag{18}$$

Third, the element abundances N_i in supernova ejecta models satisfy

$$N_i = \sum_j \sum_k N_{i,j,k} . \tag{19}$$

We adopt a similar algorithm as is in the TARDIS code to solve the $N_{i,j,k}$ for the above equation group (i.e., Eqs. (17), (18), (19)). In the first step, we assume all the atoms are singly ionized and calculate an $N_{e,0}$. In the second step, the level population $N_{i,j,k,0}$ is calculated from Equation (17) and Equation (19) using the assumed $N_{e,0}$. In the third step, a newer electron density $N_{e,\mathrm{new}}$ is calculated from Equation (18) using the level population in the second step. In the fourth step, we calculate the mean of two electron densities: $N_{e,1} = (N_{e,0} + N_{e,\mathrm{new}})/2$, then go back to the second step and replace $N_{e,0}$ with $N_{e,1}$. We repeat the above 4 steps until there is a convergence of ionization state and electron density. In our calculations, the number of iterations is set to 30, which guarantees a convergent and sufficiently accurate solution in all of our tests.

2.4. *Gamma-Ray Energy Deposition*

Gamma-ray photons in SNe Ia are mainly generated from the decay chain $^{56}\mathrm{Ni} \rightarrow {}^{56}\mathrm{Co} \rightarrow {}^{56}\mathrm{Fe}$ and the photon energy is deposited as thermal energy via Compton scattering and photoelectric absorption processes. The pair-production process is negligible for our system (e.g., Kasen et al. 2006). The physical details of our gamma-ray treatment (discussed below) are taken from the appendix of Kasen et al. (2006). The Compton scattering opacity is

$$k_\mathrm{C} = \sigma_\mathrm{T} K(x) \sum_i i N_i , \tag{20}$$

where i is the atomic number as above, σ_T is the Thomson cross section for Thomson scattering (e.g., Mihalas 1978, p. 106), x is gamma-ray photon energy $x = h\nu/(m_e c^2)$, and $K(x)$ is the Klein-Nishina correction to the Thomson cross section for Compton scattering:

$$K(x) = \frac{3}{4}\left\{\frac{1+x}{x^3}\left[\frac{2x(1+x)}{1+2x} - \ln(1+2x)\right] + \frac{\ln(1+2x)}{2x} - \frac{1+3x}{(1+2x)^2}\right\} . \tag{21}$$

Note the summation is over all the bound and free electrons because the gamma-ray energies are much larger than the photoionization energies, and so all the electrons contribute to the Compton scattering effect.

The photoelectric opacity is

$$k_\mathrm{p} = \sigma_\mathrm{T} \alpha^4 \left(8\sqrt{2}\right) x^{-7/2} \sum_i i^5 N_i , \tag{22}$$

where α is the fine-structure constant.

The emissivity in the Compton scattering process is usually determined from the differential cross section $d\sigma/d\Omega$ in most of the Monte Carlo based radiative transfer programs (e.g., Kasen et al. 2006). The formula for $d\sigma/d\Omega$ is

$$\frac{d\sigma}{d\Omega} = \frac{3\sigma_\mathrm{T}}{16\pi} f(x,\Theta)^2 [f(x,\Theta) + f(x,\Theta)^{-1} - \sin^2\Theta] , \tag{23}$$

where Θ is the angle between the incoming and outgoing gamma-ray photon and $f(x,\Theta)$ is the energy ratio between incoming and outgoing gamma-ray photon which is given by

$$f(x,\Theta) = \frac{E_{\rm out}}{E_{\rm in}} = \frac{1}{1+x(1-\cos\Theta)} \ . \quad (24)$$

The average energy lost in an interaction is

$$F(x) = 1 - \frac{1}{4\pi}\int_\Omega f(x,\Theta)\, d\Omega \ . \quad (25)$$

We use Equation (25) to calculate a sequence of 7 discrete gamma-ray photon energy bins which correspond to 0, 1, 2, ... 6 times scattered photons. We use m as the general label for the 7 discrete energy bins which are specified in § 3.1. We then use PINN to model the number of photons over these discrete energy bins as a function of spatial coordinate r and viewing angle φ. The $(m+1)$-th emissivity is calculated from the integral of the absorbed photon energy in the m-th gamma-ray photon energy bin and averaged over solid angle and is written as

$$j_{{\rm C},m+1} = \frac{1}{4\pi}\int_\Omega k_{{\rm C},m} I_m \, d\Omega \ , \quad (26)$$

where m is the index of gamma energy bin.

The time independent gamma-ray radiative transfer equation in the static atmosphere approximation we adopt for gamma-ray transfer is

$$\cos(\varphi)\frac{\partial I_m}{\partial r} - \sin(\varphi)\frac{\partial I_m}{\partial \varphi}\frac{1}{r} + (k_{\rm C}+k_{\rm p})I_m - j_{\rm C,m} - j_{\rm r} = 0 \ , \quad (27)$$

where $j_{\rm r}$ is the gamma-ray source in the supernova atmosphere. The detailed calculation procedure for gamma-ray transfer is discussed in § 3.1. We assume the energy lost in gamma-ray Compton scattering and photoelectric absorption processes are deposited as thermal energy locally and, as aforesaid, neglect the gamma-ray photon pair-production process. Therefore the gamma-ray energy deposition is

$$E_\gamma = \sum_m h\nu_m \int_\Omega \left[(k_{\rm C,m} + k_{\rm p})I_m - j_{\rm C,m}\right] d\Omega \ , \quad (28)$$

where the summation is over all the allowed gamma-ray energy bins and we count $j_{\rm C,m}$ as a negative energy deposition. The energy deposition per unit solid angle $\mathcal{E}_\gamma = E_\gamma/(4\pi)$ is used in Equation (14) to calculate the plasma temperature.

3. THE PHYSICS INFORMED NEURAL NETWORK

The concept of solving partial differential equations (PDEs) using neural networks has a long history. The idea is commonly credited to Lagaris et al. (1998) though there is related work dating back to the late 1980s (see Viana & Subramaniyan (2021) for a historical review). Recently, Raissi et al. (2019) proposed Physics Informed Neural Network (PINN), a modern deep neural network approach to solve forward and inverse PDE-constrained problems. PINN uses a deep neural network to approximate a function over physical space and introduces constraints such as PDEs and boundary conditions directly in the loss function to train the parameters in the neural network. The neural network is called "physics informed" because the parameters and boundary conditions are mostly related to physical quantities and the PDE is usually a physical law.

In the present work, a PINN is used to approximate the specific intensity represented at given frequency points by vector $I_\nu = f(r,\varphi;w)$, where w represents the trainable parameters in the neural network. To train the neural network, three sets of data points are sampled in the physical space (r,φ): 1) the collocation points, where the PDE is enforced, are randomly sampled in the physical space from uniform distributions $r_{i,p} \in U(r_{\rm min}, r_{\rm max})$ and $\varphi_{i,p} \in U(0,\pi)$; 2) the inner boundary points where $r_{j,l} = r_{\rm min}$ and $\varphi_{j,l} \in U(0,\pi/2)$; 3) the outer boundary points where $r_{k,u} = r_{\rm max}$, $\varphi_{k,u} \in U(\pi/2,\pi)$. The PDE collocation points are used in the left-hand side of Equation (1) and Equation (27) to calculate the residuals $R_{i,p}$, which are used in the loss function. The inner and outer boundary points are directly used to calculate the predicted specific intensities: $I_{k,u} = f(r_{\rm max},\varphi_{k,u},w)$, $I_{j,l} = f(r_{\rm min},\varphi_{j,l},w)$ which are then used to calculate the residuals with respect to the pre-defined boundary conditions $R_{j,l}$ and $R_{k,u}$. The loss function is

$$L = w_p \sum_{i,\nu} R_{i,p}^2 + w_l \sum_{j,\nu} R_{j,l}^2 + w_u \sum_{k,\nu} R_{k,u}^2 \ , \quad (29)$$

where w_p, w_l, w_u are weight parameters, which should be specified before training and the summation is over the data points labeled i and the sampled frequency points labeled ν (i.e., spectral sampling pixels). The gradient of the loss function over the trainable parameters $\frac{\partial L}{\partial w}$ is calculated by reverse-mode automatic differentiation using the chain rule. Knowing the gradient, the trainable parameters can be adjusted with small steps in order to reduce the loss function. In practice, accelerated gradient-based algorithms (i.e., RMSprop (Hinton 2012), Adam (Kingma & Ba 2014)) are used to increase the training efficiency and avoid local minima of the loss function. When the neural network setup is appropriate for the problem, the loss function will be close to zero after several iterations and the neural network solution will be close to the true specific intensity solution.

PINN training is implemented in pytorch (Paszke et al. 2019). Section 3.1 specifies our solution procedure for the gamma-ray radiative transfer in SN Ia atmospheres. Section 3.2 gives the structure of the optical network, which is the major neural network for solving the optical radiative transfer problem and reports example synthetic spectra.

The spectra obtained from PINN are compared to those from the Monte Carlo radiative transfer code TARDIS using the same ejecta structure and the spectra from the formal solution calculated by the procedure specified in Appendix A. Section 3.3 introduces an approach to accelerate the calculation of plasma temperature using Equation (14). Section 3.4 discusses the computation time and the training schedule of the PINN.

Our synthetic spectra are compared to the observed spectrum of SN 2011fe at 12.35 days after explosion. This observed spectrum (Pereira et al. 2013) was obtained by Double Spectrograph (DBSP) mounted on Palomar 200-inch (P200) Telescope. The observed spectrum was also used to derive the supernova ejecta structure via the method discussed in Chen et al. (2020). The derived supernova ejecta density obeys

$$\rho = 3.87852 \times 10^{-14} \times 0.689^{\frac{v-12500\,\mathrm{km/s}}{1000\,\mathrm{km/s}}}\ \mathrm{g/cm^3}\ , \quad (30)$$

where v is the radial velocity in the ejecta. Figure 1 shows the derived density and element abundance of the SN Ia ejecta structure. We use only 9 elements, which are C, O, Mg, Si, S, Ca, Fe, Co, Ni, in this calculation for simplicity, and the highest ionization for these elements is limited to 3. Also in Figure 1, we compare our density profile with the density profile of model DDT-N100 (Röpke et al. 2012), which is also a SN Ia ejecta model used to fit SN 2011fe spectra. Both the calculations with TARDIS and PINN set the inner and outer boundary to be at, respectively, the velocity coordinates 10151.4 km/s and 35675.3 km/s.

3.1. *The Gamma-Ray Network*

The gamma-ray network calculates the gamma-ray specific intensity $I_\gamma(r, \varphi)$ in the supernova atmosphere using Equation (27) as the PDE. As a simplification, we assume the gamma-ray photon energies from the ^{56}Ni $\rightarrow$ ^{56}Co $\rightarrow$ ^{56}Fe decay chain are all 1 MeV to form the j_r term in Equation (27) and the input gamma-ray photon from the boundaries are also 1 MeV. We neglect the kinetic energy and gamma-ray energy of the positron released in the ^{56}Co decay. As discussed above in 2.4, we limit the allowed gamma-ray photon energies to 7 discrete energy bins in the gamma-ray network as a simplification. The 7 energy bins are [1,0.407,0.243,0.171,0.131,0.106,0.088] MeV. Similar to the optical network, the gamma-ray network is trained on PDE collocation points and outer boundary points and inner boundary points. The outer boundary condition is no photons entering the SN Ia atmosphere ($I_\gamma(r = r_\mathrm{max}, \varphi \in [\pi/2, \pi]) = 0$). The inner boundary condition is photons entering the supernova atmosphere from the inner boundary are in the highest energy bin ($I_{\gamma=1\mathrm{MeV}}(r = r_\mathrm{min}, \varphi \in [0, \pi/2]) = I_0$, $I_{\gamma<1\,\mathrm{MeV}}(r = r_\mathrm{min}, \varphi \in [0, \pi/2]) = 0$). Two values of I_0 are used as specified below.

The best treatment of gamma-ray radiative transfer would be to use the time-dependent radiative transfer

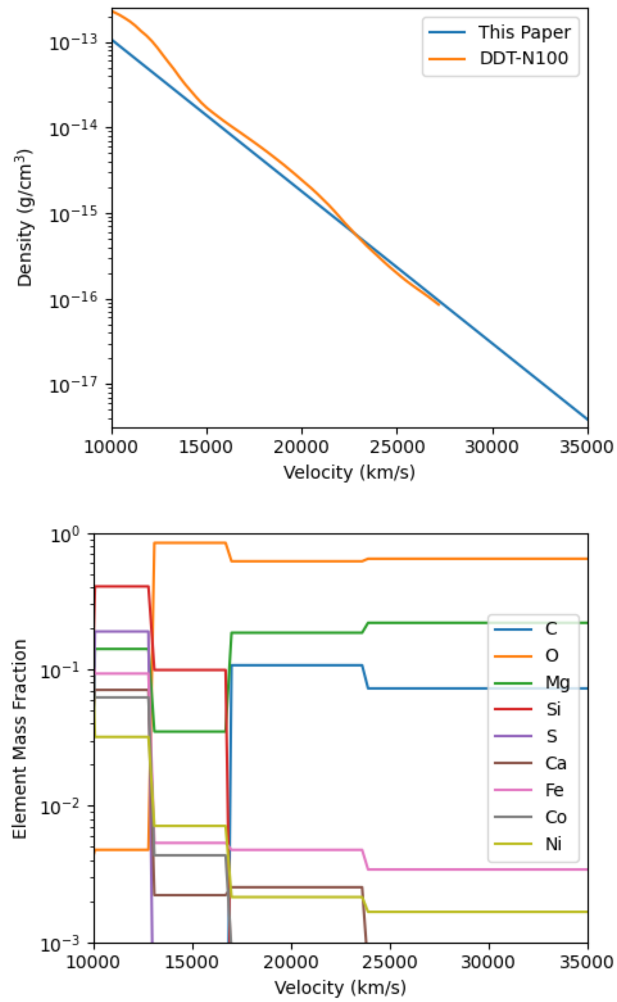

Figure 1. Upper Panel: The model SN 2011fe density profile at 12.35 days after explosion used in the TARDIS and PINN calculations (blue line) and for comparison the density profile of model DDT-N100 (Röpke et al. 2012) (orange line). Lower Panel: The model SN 2011fe element mass fractions at 12.35 days after explosion used in the TARDIS and PINN calculations.

equation and set the inner boundary to be the supernova center. However, this treatment requires an extra input dimension on the neural network, which could considerably increase the training time. Moreover, the source term j_r and the intensity I_γ would change several orders of magnitude throughout the whole supernova structure and the neural network does not give good performance over multiple orders of magnitude. Therefore, the gamma-ray radiative transfer calculation is limited to supernova upper atmosphere and the effect of the gamma-rays from the supernova center is approximated with the inner boundary condition.

We calculate two PINN models with different boundary conditions. In the first model, the inflow gamma-ray intensity is zero for both the inner and outer boundaries. In the second model, the inflow gamma-ray intensity from the inner boundary (i.e., I_0) is $10^{16}\,\mathrm{cm^{-2}s^{-1}sr^{-1}}$ with all gamma-ray energies set to 1 MeV and no inflow of gamma-rays from the outer boundary. The neural network is a 10-layered fully-connected neural network, the number of neurons is $[2, 128, 128, 128, 128, 128, 128, 128, 256, 256, 7]$, and the activation function is the hyperbolic tangent (i.e., tanh) for all the layers except the input and the output layers. The input and output layers use a linear activation function.

Note the original loss function in Equation (29) is not written with residuals in physically consistent units which makes assigning weights difficult. Therefore, we write the residuals in terms of natural units as follows:

$$R_{i,p,\mathrm{new}} = \frac{R_{i,p}}{\mathrm{Mean}(k_\mathrm{C} + k_\mathrm{p}) I_\mathrm{max}} \;, \tag{31}$$

$$R_{j,l,\mathrm{new}} = \frac{R_{j,l}}{I_\mathrm{max}} \;, \tag{32}$$

$$R_{k,u,\mathrm{new}} = \frac{R_{k,u}}{I_\mathrm{max}} \;, \tag{33}$$

where I_max is the maximum gamma-ray intensity and $\mathrm{Mean}(k_\mathrm{C} + k_\mathrm{p})$ is the mean opacity over all the PDE collocation points and all the gamma-ray energy bins. We set I_max using the equation

$$I_\mathrm{max} = I_l + \int_{r_\mathrm{min}}^{r_\mathrm{max}} j_\mathrm{r}\, dr \;, \tag{34}$$

where I_l is the inner boundary inflow intensity at 1 MeV, the I_l value is zero for the first model and $10^{16}\,\mathrm{cm^{-2}s^{-1}sr^{-1}}$ for the second model, and the integral is over the source term in the supernova atmosphere. In the loss function Equation (29), $R_{i,p}$, $R_{j,l}$, $R_{k,u}$ are replaced by $R_{i,p,\mathrm{new}}$, $R_{j,l,\mathrm{new}}$, $R_{k,u,\mathrm{new}}$, respectively. Using this modification of the loss function, the order of magnitude of the three residual terms will not change drastically with the change of supernova ejecta model or the boundary conditions. Thus, the modification helps to balance the importance of PDE and boundary conditions when training the PINN. We found the PINN results are stable when the weight parameters in Equation (29) are $w_p = 1$, $w_l = 3000$, $w_u = 3000$, respectively.

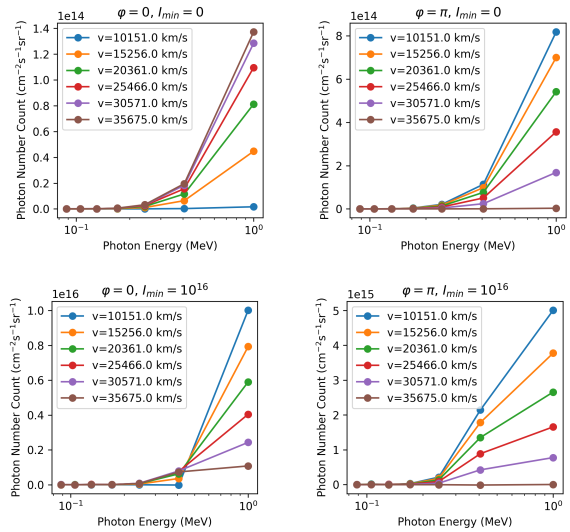

Figure 2. The gamma-ray specific intensity as a function of gamma-ray photon energy at representative velocity points and viewing angles in the two supernova ejecta models used to investigate gamma-ray radiative transfer in the ejecta. The vertical axis unit is shown at the left side of each panel. The upper panels show the results from the first model and the lower panels show the results from the second model. The left panels show the results at viewing angle $\varphi = 0$ and the right panels show the results at viewing angle $\varphi = \pi$.

Figure 2 shows the results of the two models for gamma-ray radiative transfer. Note that both models obey the inner and outer boundary conditions accurately. In the first model, we note that the intensity at the viewing angle $\varphi = 0$ and 1 MeV energy bin increases with the increase of radial velocity due to the ^{56}Ni and ^{56}Co in the supernova atmosphere shown in the lower panel of Figure 1. In the second model, we note the intensity is much larger than that of the first model, which means the input energy from the inner boundary dominates over the radioactive energy in the supernova atmosphere. We note that there is relatively little scattering of gamma-ray photons to energy bins below 1 Mev in all cases. The most such scattering occurs for the second model at the viewing angle $\varphi = \pi$. This is to be expected since inward moving gamma-ray photons will be mostly scattered gamma-ray photons when the gamma-ray intensity is dominated by a central source.

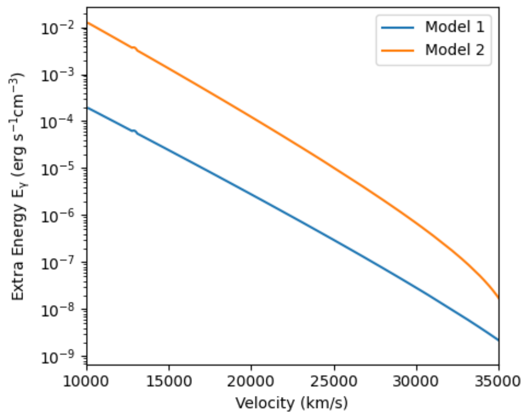

Figure 3. The gamma-ray energy E_γ calculated from the two supernova ejecta models used to investigate gamma-ray radiative transfer in the ejecta. The model names are labeled in legend.

Figure 3 shows the gamma-ray energy deposition E_γ as a function of radial velocity for the two models. Note the gamma-ray radiative transfer equation (Eq. (27)) is linear since the gamma-ray sources and opacities are fixed inside the atmosphere. Therefore, the general gamma-ray intensity (or energy deposition) for our supernova ejecta structure with a general inner boundary condition can be written as a linear combination of our two PINN-calculated gamma-ray intensity solutions (or energy deposition solutions). To be explicit for the general gamma-ray energy deposition, one has

$$\begin{aligned} E_\gamma(I_l, r) &= E_\gamma(I_{l,1}, r) + \frac{I_l - I_{l,1}}{I_{l,2} - I_{l,1}} \left[E_\gamma(I_{l,2}, r) - E_\gamma(I_{l,1}, r)\right] \\ &= E_\gamma(I_{l,1}, r) + \frac{I_l}{I_{l,2}} \left[E_\gamma(I_{l,2}, r) - E_\gamma(I_{l,1}, r)\right] \ , \end{aligned} \quad (35)$$

where $I_{l,1} = 0$ is the inner boundary condition of the first model, $I_{l,2} = 10^{16}\,\mathrm{cm^{-2}s^{-1}sr^{-1}}$ is the inner boundary condition of the second model, and I_l is the target inner boundary condition. Using this relation, the gamma-ray energy deposition for our supernova ejecta structure with different inner boundary conditions can be generated without extra PINN calculations.

3.2. *The Optical Network*

The optical neural network calculates the optical specific intensity $I_{r,\varphi}$ using Equation (1) as PDE. The output of the PINN is the specific intensity sampled in a frequency grid between $10^{14.4}$ Hz (11,935 Å) and 10^{15} Hz (3000 Å) with 2048 pixels uniformly sampled in the logarithmic space. The frequency upper limit is set as 10^{15} Hz for two simplification reasons. First, several strong Fe-group element spectral lines, which could lead to order-of-magnitude problems in the training of PINN, lie above this frequency. Second, bound-free opacity is more significant above this frequency, and we have not included bound-free opacity as a simplification. In fact, the specific intensity is suppressed by the high-opacity spectral lines and the bound-free opacity above 10^{15} Hz. Therefore, removing the specific intensity calculation above this frequency will probably not lead to significant error in a direct sense. However, the thermal state of the ejecta can only be crudely approximated without the high frequency region, bound-free opacity, and NLTE effects.

Considering the output of the neural network is a vector with the length of 2048, while other PINN applications have typically one or a few dimensions as output (e.g., Mishra & Molinaro 2021), we prepared a large 14-layered neural network structure (hereafter N14). The number of neurons in each layer is, respectively, [2, 256, 256, 256, 256, 256, 512, 512, 512, 512, 512, 2048, 2048, 2048, 2048] and the activation function is the hyperbolic tangent (i.e., tanh) for all the layers except the input and the output layers: the input and output layers use a linear activation function. We also prepared a smaller 6-layered neural network (hereafter N6) for comparison: the number of neurons in each layer is [2, 512, 2048, 2048, 2048, 2048].

The outer boundary condition is that no radiation flow into the material ($I(r_{\max}, \varphi \in [\pi/2, \pi]) = 0$) and the inner boundary condition is that the input radiation flow is an isotropic blackbody spectrum with a pre-defined boundary temperature T_{Bo}:

$$I_{r_{\min}, \varphi \in [0, \pi/2]} = \frac{2h\nu^3}{c^2\left(e^{\frac{h\nu}{k_{\mathrm{B}} T_{\mathrm{Bo}}}} - 1\right)} \ . \quad (36)$$

We found the synthetic spectra are close to the observed spectra when $T_{\mathrm{Bo}} = 11500$K, and thus we adopted this value for all our calculations. As a simplification, we used $I_{r_{\min}}$ as a rest frame specific intensity though. Formally it should be a comoving frame specific intensity. The distinction between the two quantities is small.

Similar to the gamma-ray network, residuals in the loss function Equation (29) are written in terms of natural units as follows:

$$R_{i,p,\mathrm{new}} = \frac{R_{i,p}}{I_{\max}\mathrm{Mean}(k_{\mathrm{e}} + k_{\mathrm{bb}})} \ , \quad (37)$$

$$R_{j,l,\mathrm{new}} = \frac{R_{j,l}}{I_{\max}} \ , \quad (38)$$

$$R_{k,u,\mathrm{new}} = \frac{R_{k,u}}{I_{\max}} \ , \quad (39)$$

where $I_{\max}$ is the maximum pixel value of the inner boundary condition Equation (36) and the mean is over all the frequency pixels and PDE collocation points. The weight parameters used in Equation (29) are $w_p = 1$, $w_l = 3000$, and $w_u = 3000$. These are the same as for the gamma-ray radiative transfer calculation and were found comparably good for the optical spectrum calculations.

For our example synthetic spectrum calculation, we use the gamma-ray energy E_γ calculated in § 3.1 and assume the gamma-ray intensity inner boundary condition is $I_0 = 10^{16}\,\mathrm{cm^{-2}s^{-1}sr^{-1}}$ and the plasma temperature is calculated using the temperature network, which is described in § 3.3.

Figure 4 shows the specific intensities as functions of rest frame wavelength sampled at representative radial velocities and viewing angles calculated by the N14 neural network (i.e., the N14 PINN) and the formal solution (see Appendix A). Note that at the inner boundary ($v = 10151.4\,\mathrm{km/s}$, $\varphi < \pi/2$) and the outer boundary ($v = 35675.3\,\mathrm{km/s}$, $\varphi > \pi/2$), the PINN solutions satisfy the boundary condition virtually exactly as should be the case. Also note that there were no specific intensity values less than zero as should be the case. Blueshifted absorption lines can be observed in the $\varphi = 0$ spectra and emission lines of varying shift can be seen in the spectra with $\varphi > 0$.

To explicate the absorption and emission specific intensity features formed inside the SN Ia model ejecta and illustrated in Figure 4, note the following equation for comoving frame wavelength λ derived from the Doppler shift formula Equation (2):

$$\lambda = \frac{\bar{\lambda}}{\gamma[1 - \cos(\varphi)\beta]} \;, \tag{40}$$

where $\bar{\lambda}$ is the rest frame wavelength. Now consider a line wavelength (i.e., bound-bound transition wavelength) in the comoving frame λ_{line} which is, of course, the laboratory line wavelength. Next consider a beam of rest frame wavelength $\bar{\lambda}$ (which is invariant, of course, as the beam propagates) that starts from the inner boundary at a point A and reaches a point B in the atmosphere where we evaluate the rest frame specific intensity at that wavelength $\bar{\lambda}$. Note that a beam starting from the inner boundary always has $\cos(\varphi)\beta > 0$, and thus always redshifts in the comoving frame provided γ stays sufficiently close to 1 which it always does for supernovae, except in extreme cases which we will not consider. If $\bar{\lambda}$ is blueward of λ_{line}, but not too blueward, the beam will redshift such that its comoving frame wavelength $\lambda = \lambda_{\mathrm{line}}$ at point C that is in between point A and point B. There will be absorption from the beam by the line at point C. There will also be emission into the beam by the line at point C, but in supernovae in the photospheric phase (when the overall ejecta is optically thick enough to give rise to a photosphere which in our modeling is the inner boundary), the absorption usually dominates and the beam specific intensity at rest frame wavelength $\bar{\lambda}$ is diminished passing through point C. Since there is a continuous rest frame wavelength range of emission from the inner boundary, there will be a continuous rest frame wavelength range of line absorption which happens a continuous range of spatial points along a single beam path. The foregoing explicates the absorption features seen in Figure 4 for the beams with $\varphi = 0$ and velocity greater than the inner boundary velocity: all these beams are radial, in fact. For beams at the inner boundary velocity with $\varphi \le \pi/2$ (shown in the top panels of Figure 4), the specific intensity is just the inner boundary condition specific intensity and all the curves for these beams overlap. For beam paths not starting on the inner boundary, there is net emission into the beam and this explains the emission features seen in Figure 4 for beam paths with $\varphi \neq 0$ and not starting on the inner boundary.

Continuing the explication of beams in the SNe Ia model ejecta, the points C for a single rest frame wavelength $\bar{\lambda}$ for beams heading toward a distant observer are on planes perpendicular to the direction to the distant observer and this true for all of supernovae, in fact. These planes have constant velocity in the direction to the distant and were discussed in § 2. On the planes there is usually net absorption from all beams heading toward the distant observer. The upshot is there is a broad blueshifted absorption in the spectrum of the supernova as seen by the distant observer. The absorption cuts into the continuum level and the emission feature (which we describe in the next paragraph).

What of beams heading toward the distant observer that do not start on the inner boundary, but rather from line emission at comoving frame wavelength λ_{line}? Given spherically symmetric geometry, the emission tends to be strongest from the plane in the atmosphere perpendicular to beams to the distant observer that passes through the center of mass of the supernova: we will call this plane the central plane. The beams that start on the central plane have $\bar{\lambda} = \lambda_{\mathrm{line}}$ since the central plane has zero velocity in the direction of the distant observer in the rest frame. Beams from (parallel) planes closer/farther relative to the central plane are blueshifted/redshifted in rest frame wavelength from λ_{line} (since they have higher/lower velocity in the direction toward the distant observer) and are usually weaker in intensity the more they are blueshifted/redshifted since they come from lower-density-on-average planes. The upshot is the emission from the planes toward the

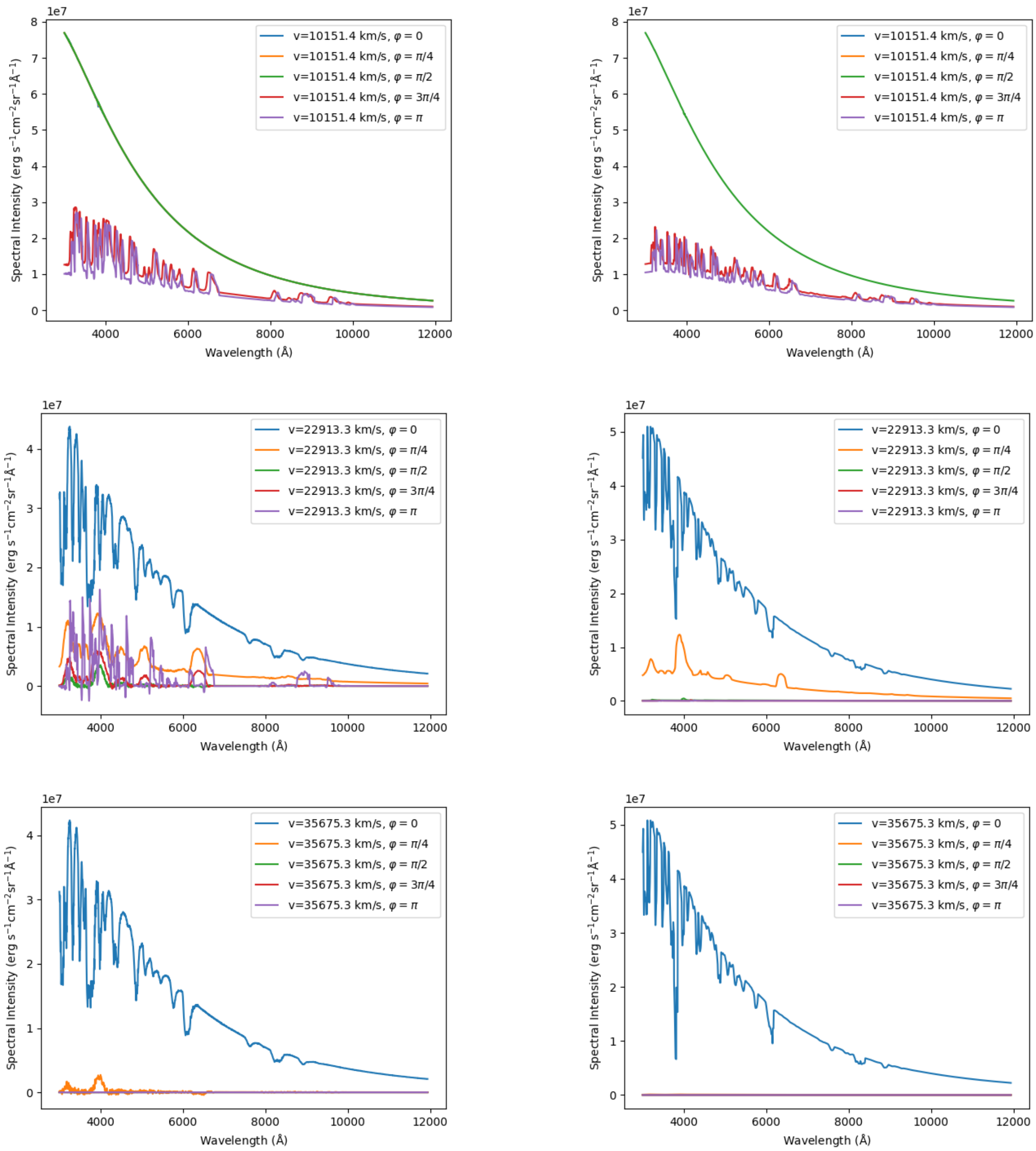

Figure 4. The rest frame specific intensity as a function of rest frame wavelength $\bar{\lambda}$ at representative radial velocities and viewing angles. Note specific intensity is in the wavelength representation rather then in the frequency representation which we use in the text. The rest frame is the frame defined by the center of mass of the spherically symmetric SN Ia ejecta. The rest frame is the same for all following figures. The left panels show specific intensities from PINN calculations using the N14 neural network structure and the right panels show the corresponding specific intensities from the formal solution of the PDE. The coordinates and viewing angles are shown in legends. In the upper panels, the specific intensity curves for $\varphi \leq \pi/2$ are just the inner boundary specific intensity, and so all overlap and give a net green color. In the middle right panels, the specific intensity curves for $\varphi \geq \pi/2$ are all nearly zero, and so overlap and give a net purple color. In the lower left panel (lower right panel), the specific intensity curves for $\varphi \geq \pi/2$ ($\varphi \geq \pi/4$) are all nearly zero, and so overlap and give a net purple color.

observer tends to give a broad emission feature in rest frame wavelength centered on $\lambda_{\rm line}$ and superimposed on the spectrum continuum level. However, the blueshifted absorption in beams that start on the inner boundary cuts into the continuum level and the absorption feature (as we described in the last paragraph) and the result is a P-Cygni line: a broad observed spectrum line in rest frame wavelength consisting of a blueshifted aborption and an asymmetric emission roughly centered on a line wavelength.

A detailed explication of P-Cygni line formation in supernovae is given by Jeffery & Branch (1990, p. 173–194).

Having finished our explication of the formation of the features in Figure 4, we now discuss the difference between the PINN spectra and formal solution spectra. When comparing them, we notice significant differences at coordinates ($v = 22913.3\,\mathrm{km/s},\ \varphi \geq \pi/2$) and ($v = 35675.3\,\mathrm{km/s},\ \varphi = \pi/4$). The PINN spectrum emission lines are much larger than the formal solution emission lines (which in fact are close to zero). The same discrepancy also occurs when the PINN spectra are calculated by the N6 neural network structure.

To investigate the cause of the discrepancies between the PINN and formal solution specific intensity spectra, we have plotted in Figure 5 several example rest frame opacity and emissivity terms at different coordinates as functions of rest frame wavelength. We note that the opacity and emissivity due to the bound-bound transitions (which are seen as sharp spikes) can be up to $\sim$ 1000 times the corresponding electron scattering terms for v = 12207.4 km/s and up to $\sim$ 10 times the corresponding electron scattering terms for v = 23211.5 km/s. We surmise the discrepancy between the PINN solution and the formal solution for the specific intensity spectra is due to the large order-of-magnitude variation in the opacity and emissivity terms in the PDE loss function that goes beyond the dynamic range of neural networks.

Using the specific intensity calculated by PINN or the formal solution, a synthetic spectrum that can be compared to observations is calculated by the integral

$$\mathrm{Spec}(\nu) = \int_0^{\pi/2} d\varphi\ 2\pi r_{\max}^2 I(r_{\max}, \varphi, \nu) \sin(\varphi) \cos(\varphi) \tag{41}$$

(e.g., Mihalas 1978, p. 11–12). In the wavelength representation (not the frequency representation), Figure 6 shows two PINN synthetic spectra from the N14 and N6 neural network structures, the corresponding formal solution synthetic spectrum, the TARDIS synthetic spectrum (fitted to the observations using the method of Chen et al. (2020)), and the observed spectrum for SN 2011fe at 12.35 days after explosion. Because the method in Chen et al. (2020) is specifically designed for TARDIS, as well as the supernova ejecta structure used in this paper, the TARDIS synthetic spectrum fits observed spectrum with reasonable accuracy. However, the PINN and formal solution spectra (which are obtained without the optimized fitting of Chen et al. (2020)) do not fit to the same level of accuracy and cannot be expected to do so.

The test of the PINN synthetic spectrum calculation is the comparison to the formula solution synthetic spectrum calculation which is based on exactly the same atmosphere structure and thermal state and is calculated with guaranteed numerical accuracy. First, we note the PINN spectra and the formal solution spectrum are qualitatively alike. In particular, they both exhibit typical P-Cygni lines as are also seen in the observed and TARDIS spectra. P-Cygni lines are characteristic of supernova atmospheres and expanding atmospheres in general and, as explicated above, have a broad emission feature centered around the laboratory line wavelength and a blueshifted absorption feature. Note line blending can distort P-Cygni line behavior to unrecognizablity.

Second, in the spectra in Figure 6, P-Cygni lines are recognizable for several conspicuous spectral lines: Ca K&H 3934 Å, 3968, Å, Si II 6355 Å, S II 5468 Å (multiplet average), S II 5640 Å (rough average of several lines); Ca II 8498 Å, 8542 Å, 8662 Å. Qualitatively, the agreement between the PINN spectra and the formal solution spectrum for these lines is moderate. Overall, the PINN spectra show stronger emission features. This is to be expected given that the PINN line emission specific intensities in Figure 4 were generally too strong in the PINN case. The PINN spectra also show noise which is to be expected for PINN calculations. We note that neither the PINN sprectra nor the formal solution spectrum produce the "W" shaped S II feature around 5500 Å seen in the observed spectrum. The TARDIS spectrum does produce this shape qualitatively and this is probably attributable to TARDIS's better thermal state calculation compared to ours. TARDIS uses a dilute-blackbody approximation in the temperature calculation and the macroatom approximation in the source function calculation. Our thermal state calculation is simpler and is described in §§ 2 and 3.3.

3.3. *The Temperature Network*

The plasma temperature calculated directly from Equation (14) requires numerical integration. However, it is known from previous simulations (e.g., Chen et al. 2020), that the temperature profile in supernovae above the photosphere is a smooth function of radial velocity. Therefore, we use a simple neural network to interpolate the temperature profile during the PINN calculation in order to curtail the computational time in numerical integration. The neural network is a simple fully-connected neural network, the number of neurons per layer is [1,64,64,64,1], and the activation function

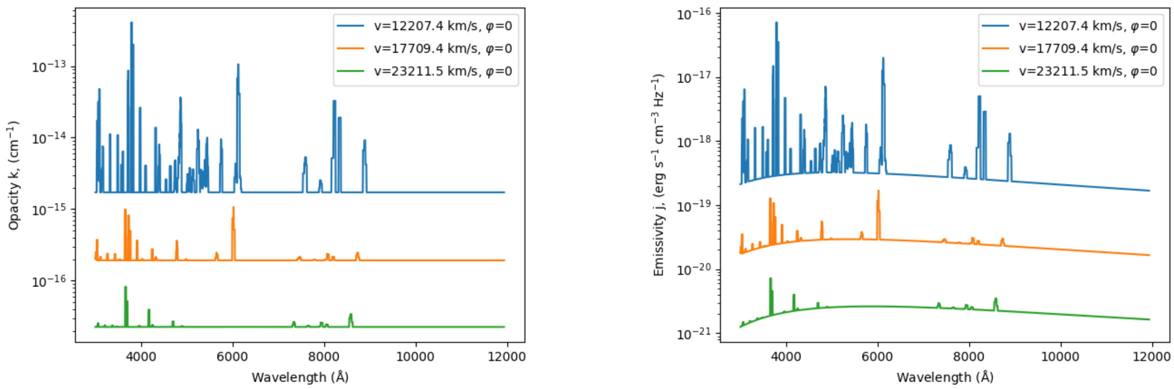

Figure 5. Left panel is the rest frame opacity k at various velocity coordinates. Right panel is the rest frame emissivity j at various velocity coordinates. The wavelength is rest frame wavelength $\bar{\lambda}$.

is the SELU function (Klambauer et al. 2017). During the training of the optical network, the radius values of the PDE collocation points are input into the temperature neural network, then the predicted temperature values are used to calculate the electron density N_e, the level populations and the opacity and emissivity used in Equation (1). During the training of the temperature network, we randomly sample 200 radius values, then use the following loss function to train the neural network:

$$L_T = \sum_n \Big[4\sigma_{\rm SB}\sigma_{\rm T}N_e T(r_n)^4 - E_\gamma(r_n) - \sigma_{\rm T}N_e\int_{\nu_{\rm min}}^{\nu_{\rm max}} d\nu \int_\Omega d\Omega\, I(r_n,\varphi)\Big]^2 , \tag{42}$$

where n labels the random sample, r_n is the radius from the random sample, $T(r_n)$ is the neural network predicted temperature, and the integral is calculated using trapezoid integration over the 2048 frequency pixels and Monte Carlo integration with 200 sampling points over the viewing angle φ. The loss function forces an approximate solution of Equation (14) throughout the atmosphere. Note in the above equation we used $I(r_n,\varphi)$ which is a rest frame specific intensity. Properly, we should use a comoving frame specific intensity. However, the distinction between the two quantities is small.

Figure 7 shows the neural network predicted temperature and the temperature calculated from the integration of the specific intensity I. There is reasonably good agreement between the two temperature curves as seen in the lower panel. At most radial velocities, the temperature difference is smaller than 50 K. However, the temperature difference near the inner boundary is ~ 500 K and near the outer boundary is ~ 250 K. Although other methods (i.e., linear interpolation, cubic spline) may be as good as, or even better than, the neural network in approximating the temperature profile as a function of radius $T(r)$ from 200 sampling points, the neural network will be a better interpolation function in higher dimensional problems (e.g., the 3D radiative transfer problem). So we will continue to use the neural network as the temperature interpolation function for the upgrades in the future.

We need reiterate that our atmosphere calculation relies on many approximations, and many results, including the temperature profile in Figure 7, have low quantitative reliability. However, the temperature profile is roughly consistent with expectations from detailed NLTE calculations insofar as we can tell. One of the few papers to publish temperature profiles from detailed NLTE calculations for supernovae is DerKacy et al. (2020). Their Figure 11 shows temperature profiles for SN 2011fe for a similar ejecta model to the one we use, but with an outer boundary at about 25,000 km/s. Recall in our ejecta model the inner and outer boundaries are at, respectively, the velocity coordinates 10151.4 km/s and 35675.3 km/s. The profiles of DerKacy et al. (2020) are overall about 2000 K lower than ours, but they are modeling SN 2011fe for an epoch 10 days later than our spectrum when an overall decline in temperatures in the outer layers of order 2000 K is to be expected. Also the overall density of their ejecta model is a factor of ~ 6 lower than that of our ejecta model because of ejecta expansion in the interval between the two epochs. The higher density for the epoch for our model usually means NLTE effects will be lower for our epoch. The temperature profiles of DerKacy et al. (2020) are for a range of model luminosities. The lower luminosities give a monotonic decline with velocity and the higher ones give a rise in temperature above 20,000 km/s. The reasons for the rises are not explicitly discussed by DerKacy et al. (2020). However, we conclude from their results that NLTE effects in and of

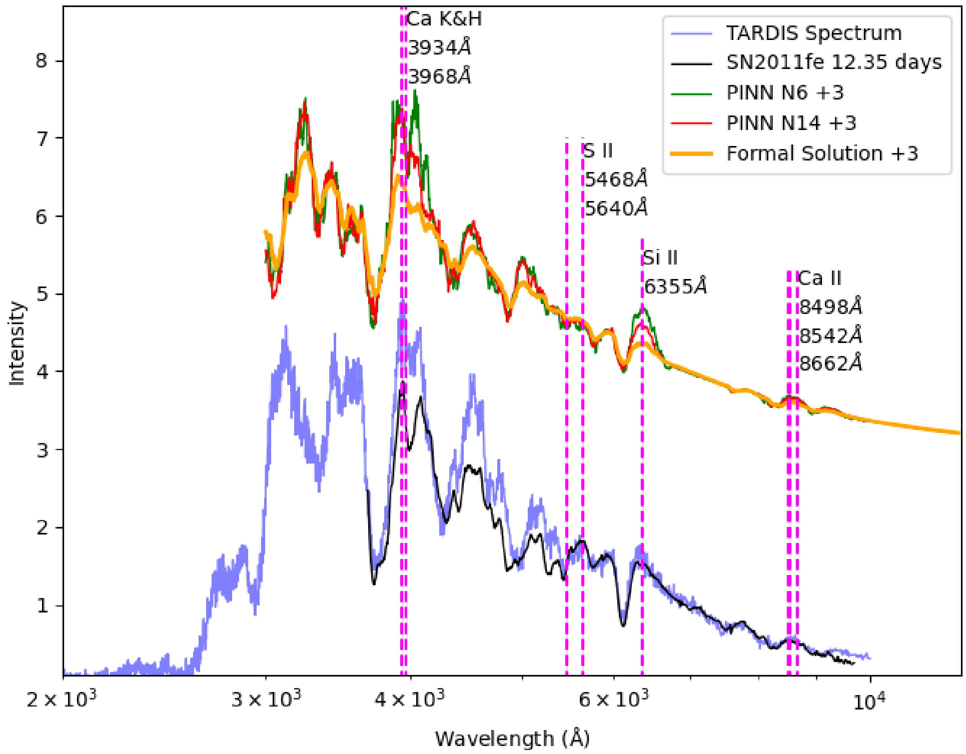

Figure 6. In the wavelength representation, the spectrum of SN 2011fe at 12.35 days after explosion (black line), the TARDIS synthetic spectrum (blue line), the PINN spectra from N14 and N6 neural network structures (red line and green line), and the formal solution spectrum (orange line). The intensity is in arbitrary units and rest frame wavelength $\bar{\lambda}$ is on a logarithmic scale. The PINN spectra and formal solution spectrum are moved upward by 3 units (as indicated in the legend) for clarity. Several spectral lines are marked with magenta dashed lines.

themselves do not lead to rising temperature in the outer layers of SNe Ia. Combinations of effects may lead to rises or not as the case may be.

3.4. *The Training Procedure*

When training the gamma-ray neural network, we use the Adam algorithm (Kingma & Ba 2014) to update the trainable parameters. In each epoch of training, the neural network reads a small batch of data with 2000 PDE collocation points, 2000 outer boundary points, and 2000 inner boundary points to update the trainable parameters. An epoch has 200 batches and there are 140 epochs. During the training, the learning rate changes from 1×10^{-4} to 1×10^{-6}. The first few epochs have larger learning rates to efficiently train the neural network to an approximate solution, and then the smaller learning rates in the following epochs increase the training precision. The total training time is about 36 minutes using one Nvidia-A100 GPU card.

The optical neural network with the N14 structure is much more sophisticated and requires more computation time: 62 hours in total (see below). The computation time of different sub-steps of the training on a Nvidia-A100 GPU card are as follows:

- Calculating the $k_{\rm abs}$ and $j_{\rm em}$ values of 1500 PDE collocation points on 2048 frequency sampling pixels takes 965 ms.
- Updating the optical neural network with the Adam algorithm using 1500 PDE collocation points, 1500 outer boundary points, and 1500 inner boundary points takes 71 ms.
- Calculating the integral in Equation (14) over 200 sampling points takes 863 ms.
- Updating the temperature with the Adam algorithm using 200 sampling points (see § 3.3) takes 4.5 ms.

Note the calculation of the $k_{\rm abs}$ and $j_{\rm em}$ values and the integral in Equation (14) take much longer times than

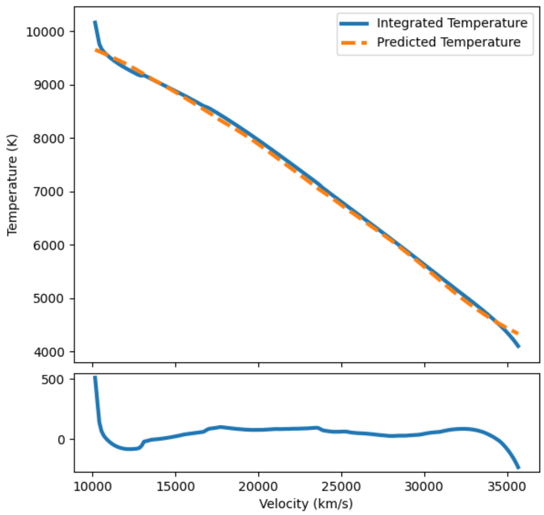

Figure 7. Upper panel: The predicted temperature from the temperature neural network (orange dashed line) and the temperature integrated from the specific intensity using Equation (14) as functions of radial velocity. Lower panel: The difference between two temperatures.

the other two computation processes. In order to accelerate the training, we make two modifications. First, we repeat the updating of optical neural network 20 times on the same batch of data points. Second, because the temperature neural network is much simpler than the optical neural network, we set the learning rate to be 10 times of that of the optical neural network.

As aforesaid for the optical network, the training batch size is 1500 for the PDE collocation points and for each of the inner and outer boundary points. An epoch has 400 batches and the total training procedure has 136 epochs. Similar to the gamma-ray neural network, the learning rate of the optical neural network changes from 1×10^{-4} to 1×10^{-6} during the training epochs. The total training time of the N14 neural network is 62 hours using one Nvidia-A100 GPU. The total training time of the N6 neural network is 60 hours using one Nvidia-A100 GPU, which means a simplified neural network structure did not significantly reduce the large amount of training time in the present work.

4. CONCLUSION

We used PINNs to calculate an optical spectrum of SN Ia SNe SN 2011fe at 12.35 days after explosion. The specific intensity throughout the supernova atmosphere is roughly solved and the synthetic spectrum is in qualitative agreement with the observed spectrum and the formal solution spectrum, noting especially that the spectrum line profiles caused by several important atomic transitions (e.g., Si II 6355 Å; Ca II 8498 Å, 8542 Å, 8662 Å) are qualitatively reproduced.

However, there are several challenges to the further exploration of the supernova explosion mechanism via the PINN-based method. First, the PINN-based method is inefficient at integration. The only integral calculation in the current PINN setup is Equation (14) for temperature which requires a significant amount of the computation time.

Second, apart from the integration calculations, PINN calculation is slow. Despite the temperature neural network, the refined training strategy, and other tricks we have introduced, which have already accelerated the training procedure significantly, the computation cost of a full simulation is about several GPU-days In contrast, TARDIS typically uses several CPU hours to run a simulation.

Third, the PINN spectrum is not quantitatively accurate as shown by comparison to the formal solution spectrum. We surmise that this is due to large order-of-magnitude variations in emissivity $j_{\rm em}$ and opacity $k_{\rm abs}$ (see § 3.2). Using XPINN (Hu et al. 2021), which can separate the parameter space into different subdomains and connect the neural networks in different subdomains with extra boundary conditions, may alleviate this order-of-magnitude variation problem. However, we did not attempt this method in this paper because it can drastically increase the computational resources required.

To summarize, using PINN in the forward modeling problem of supernova radiative transfer calculation faces multiple challenges in computational efficiency, and therefore in applying it to a large grid of supernova ejecta models. The challenges to PINN radiative transfer equally apply to the construction of a PINN inverse problem solver which encodes the supernova ejecta structure parameters into the input of the PINN and fits observed spectra. If the challenges are not overcome, an inverse problem solver will be too computationally demanding for use. The high dimensionality of the parameter space for an inverse problem solver adds to the challenges.

Clearly, innovative upgrades are necessary to significantly accelerate the PINN training process and radiative transfer calculation for either forward or inverse modeling. Those upgrades may include combining PINN with other methods: e.g., Monte Carlo or traditional numerical PDE methods.

ACKNOWLEDGEMENTS

Portions of this research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing. Ulisses Braga-Neto and Ming Zhong are supported by the National Science Foundation through NSF award CCF-2225507. Lifan Wang acknowlege the NSF grant AST-1817099 for supports on this work.

REFERENCES

Abadi, M., Barham, P., Chen, J., et al. 2016, in 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16) (Savannah, GA: USENIX Association), 265–283. https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi

Arnett, W. D. 1982, ApJ, 253, 785, doi: 10.1086/159681

Baydin, A. G., Pearlmutter, B. A., Radul, A. A., & Siskind, J. M. 2017, J. Mach. Learn. Res., 18, 5595–5637

Buchner, J. 2016, Statistics and Computing, 26, 383, doi: 10.1007/s11222-014-9512-y

Bulla, M., Sim, S. A., & Kromer, M. 2015, MNRAS, 450, 967, doi: 10.1093/mnras/stv657

Castor, J. I. 1972, ApJ, 178, 779, doi: 10.1086/151834

Chen, X., Hu, L., & Wang, L. 2020, The Astrophysical Journal Supplement Series, 250, 12, doi: 10.3847/1538-4365/ab9a3b

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

Dissanayake, M. W. M. G., & Phan-Thien, N. 1994, Communications in Numerical Methods in Engineering, 10, 195, doi: https://doi.org/10.1002/cnm.1640100303

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

Hauschildt, P. H., & Baron, E. 2006, A&A, 451, 273, doi: 10.1051/0004-6361:20053846

Hillier, D. J., & Miller, D. L. 1998, ApJ, 496, 407, doi: 10.1086/305350

Hinton, G. 2012. http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture_slides_lec6.pdf

Hu, Z., Jagtap, A. D., Karniadakis, G. E., & Kawaguchi, K. 2021, arXiv e-prints, arXiv:2109.09444. https://arxiv.org/abs/2109.09444

Hubeny, I., & Mihalas, D. 2014, Theory of stellar atmospheres: An introduction to astrophysical non-equilibrium quantitative spectroscopic analysis, Vol. 26 (Princeton, New Jersey: Princeton University Press)

Jeffery, D., & Branch, D. 1990, in Jerusalem Winter School for Theoretical Physics, Volume 6: Supernovae, ed. J. C. Wheeler, T. Piran, & S. Weinberg (Singapore: World Scientific), 149–247

Karniadakis, G. E., Kevrekidis, I. G., Lu, L., et al. 2021, Nature Reviews Physics, 3, 422

Kasen, D., Thomas, R. C., & Nugent, P. 2006, ApJ, 651, 366, doi: 10.1086/506190

Kerzendorf, W. E., & Sim, S. A. 2014, MNRAS, 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

Kingma, D. P., & Ba, J. 2014, arXiv e-prints, arXiv:1412.6980. https://arxiv.org/abs/1412.6980

Klambauer, G., Unterthiner, T., Mayr, A., & Hochreiter, S. 2017, arXiv e-prints, arXiv:1706.02515. https://arxiv.org/abs/1706.02515

Kromer, M., & Sim, S. A. 2009, MNRAS, 398, 1809, doi: 10.1111/j.1365-2966.2009.15256.x

Lagaris, I. E., Likas, A., & Fotiadis, D. I. 1998, IEEE transactions on neural networks, 9, 987

Livneh, R., & Katz, B. 2022, MNRAS, doi: 10.1093/mnras/stab3787

Lucy, L. B. 1999, A&A, 344, 282

Mihalas, D. 1978, Stellar atmospheres (San Francisco: W.H. Freeman and Company)

Mishra, S., & Molinaro, R. 2021, Journal of Quantitative Spectroscopy and Radiative Transfer, 270, 107705, doi: 10.1016/j.jqsrt.2021.107705

Parrent, J., Branch, D., & Jeffery, D. 2010, SYNOW: A Highly Parameterized Spectrum Synthesis Code for Direct Analysis of SN Spectra. http://ascl.net/1010.055

Paszke, A., Gross, S., Massa, F., et al. 2019, in Advances in Neural Information Processing Systems 32, ed. H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, & R. Garnett (Curran Associates, Inc.), 8024–8035. http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-libr pdf

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

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

Raissi, M., Perdikaris, P., & Karniadakis, G. 2019, Journal of Computational Physics, 378, 686, doi: https://doi.org/10.1016/j.jcp.2018.10.045

Riess, A. G., Yuan, W., Macri, L. M., et al. 2021, arXiv e-prints, arXiv:2112.04510. https://arxiv.org/abs/2112.04510

Röpke, F. K., & Hillebrandt, W. 2005, A&A, 431, 635, doi: 10.1051/0004-6361:20041859

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

Rybicki, G. B., & Hummer, D. G. 1978, ApJ, 219, 654, doi: 10.1086/155826

Thomas, R. C., Nugent, P. E., & Meza, J. C. 2011, PASP, 123, 237, doi: 10.1086/658673

Viana, F. A., & Subramaniyan, A. K. 2021, Archives of Computational Methods in Engineering, 28, 3801

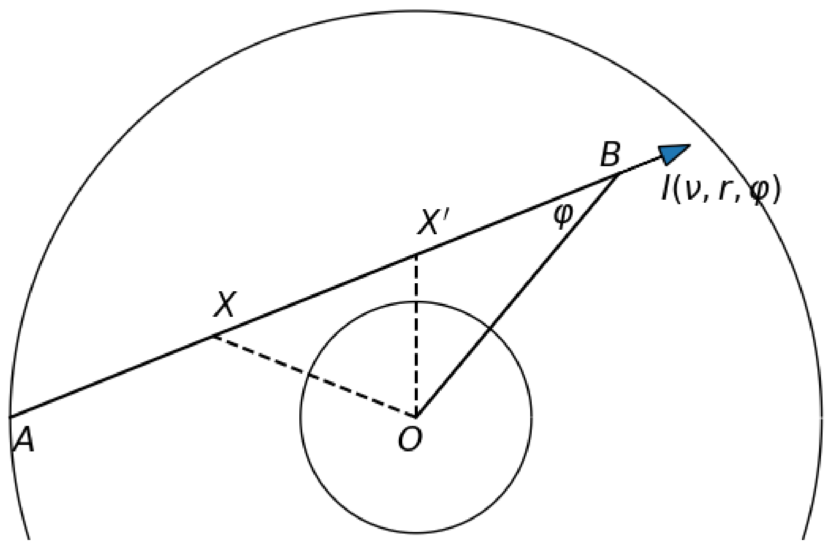

Figure 8. An illustration of the geometry of the atmosphere and the beam path used in our presentation of the formal solution.

APPENDIX

A. THE FORMAL SOLUTION OF THE RADIATIVE TRANSFER EQUATION

In this appendix, we present the analytical formal solution of the specific intensity I at a given coordinate r and a viewing angle φ for the radiative transfer equation (Eq. (1)) written in terms of beam path coordinate x, where x increases in the direction of radiation flow (i.e., the beam path direction). For our presentation, Figure 8 illustrates the geometry of the atmosphere and the beam path for the case that the beam path intersects the outer boundary of the atmosphere. Note the viewing angle φ is the angle between outward radial direction and the beam path: we leave it unsubscripted for point B and subscripted by x and x' for the corresponding points shown in Figure 8. In terms, of x, the radiative transfer equation (neglecting time dependence) is

$$\frac{\partial I}{\partial x} - j_{\rm em}\left(\frac{\nu}{\bar{\nu}}\right)^{-2} + k_{\rm abs}\left(\frac{\nu}{\bar{\nu}}\right) I = 0 \ , \tag{A1}$$

where the frequency dependence is implicit for I, $j_{\rm em}$, and $k_{\rm abs}$ (Castor e.g., 1972, eq. (1–3); see also Mihalas 1978, p. 31,33,495–496) and, as in § 2, $\bar{\nu}$ is rest frame frequency and ν is comoving frame frequency. As a simplification, we define the opacity and emissivity in rest frame as $K = k_{\rm abs}\left(\frac{\nu}{\bar{\nu}}\right)$ and $J = j_{\rm em}\left(\frac{\nu}{\bar{\nu}}\right)^{-2}$, and note that they both depend on the viewing angle φ via the $\nu/\bar{\nu}$ factor as seen from Equation (2) in § 2. The formal solution follows straightforwardly using the integrating factor $e^{\int_A^x K(r_{x'},\varphi_{x'})\,dx'}$:

$$I(\nu, r, \varphi) = I_{\rm BC}\ e^{-\int_A^B K(r_x,\varphi_x)\,dx} + \int_A^B J(r_x,\varphi_x) e^{-\int_x^B K(r_{x'},\varphi_{x'})\,dx'}\,dx \ , \tag{A2}$$

where $I_{\rm BC}$ is the boundary condition value. If point A is on the inner boundary, then $I_{\rm BC}$ is the inner boundary condition, which is Equation (36) as in the main text. If point A is on the outer boundary, then $I_{\rm BC}$ is the outer boundary condition, which is zero. The formal solution is calculated by numerical integration.

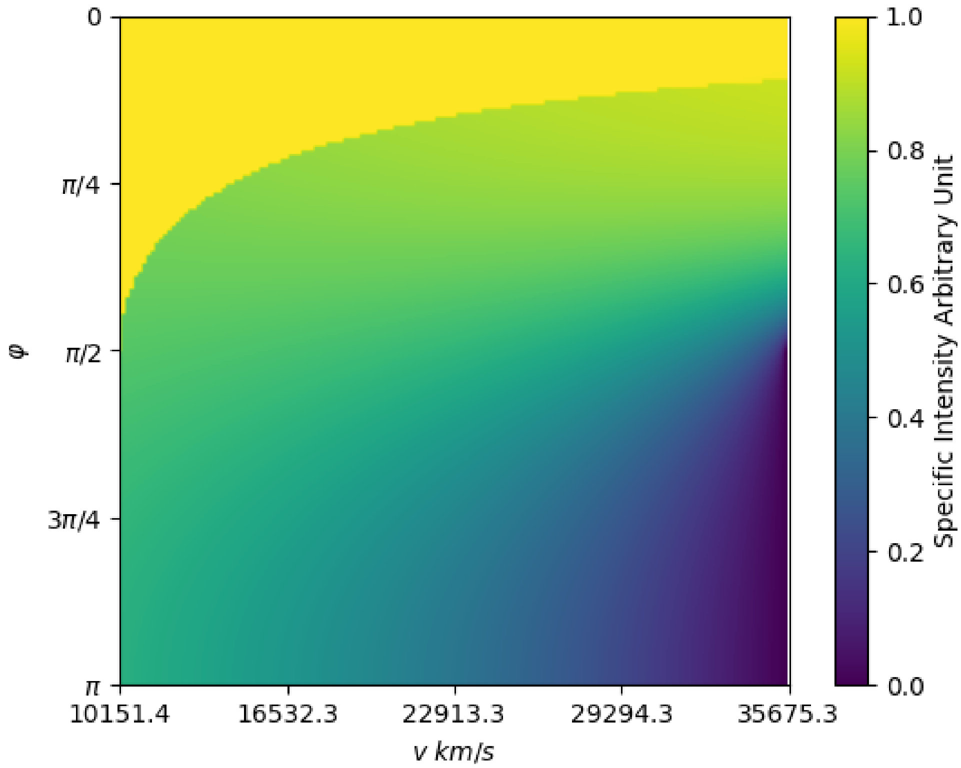

Figure 9. The specific intensity formal solution given in color format for our SN Ia atmosphere with specific intensity as a function of velocity coordinate and viewing angle φ for a single representative frequency with $k_{\rm abs}$ and $j_{\rm em}$ set to constant values. The specific intensity increases as color varies from purple to yellow.

Note for a given r, the φ parameter space is divided into two regions: one where point A is on the inner boundary and one where it is on the outer boundary. The dividing line $\varphi_{\rm div}$ is given by

$$\varphi_{\rm div} = \arcsin\left(\frac{r_{\rm min}}{r}\right) \tag{A3}$$

which always satisfies $0 \leq \varphi_{\rm div} \leq \pi/2$. The specific intensity is discontinuous across the dividing line since the boundary condition $I_{\rm BC}$ changes discontinuously across the dividing line. Therefore, we use two neural networks, one for each region.

Figure 9 shows an example of the formal solution. The discontinuity at the dividing line can clearly be seen as the curve separating the yellow color (which characterizes beams starting on the inner boundary) and the green and bluer colors (which characterize beams starting on the outer boundary).