Draft version September 2, 2023
Typeset using LaTeX **twocolumn** style in AASTeX63

Nuclear Neural Networks - Implementing machine-learning methods for stellar nucleosynthesis

Aldana Grichener,[1,2] Mathieu Renzo,[3] Wolfgang Kerzendorf,[4] Earl Bellinger,[2,5] Stephen Justham,[2,6,7,8] Rob Farmer,[2] Nutan Chen,[9] and Selma de Mink[2,8]

[1] *Department of Physics, Technion, Haifa, 3200003, Israel; aldanag@campus.technion.ac.il*
[2] *Max-Planck-Institut für Astrophysik, Karl-Schwarzschild-Straße 1, 85741 Garching, Germany*
[3] *Center for Computational Astrophysics, Flatiron Institute, New York, NY 10010, USA*
[4] *Department of Physics and Astronomy, Michigan State University, East Lansing, MI 48824, USA*
[5] *Stellar Astrophysics Centre, Aarhus University, 8000 Aarhus, Denmark*
[6] *School of Astronomy and Space Science, University of the Chinese Academy of Sciences, Beijing 100012, China*
[7] *National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100012, China*
[8] *Anton Pannekoek Institute for Astronomy and GRAPPA, University of Amsterdam, NL-1090 GE Amsterdam, The Netherlands*
[9] *Machine Learning Research Lab, Volkswagen AG, Munich, Germany*

1. INTRODUCTION

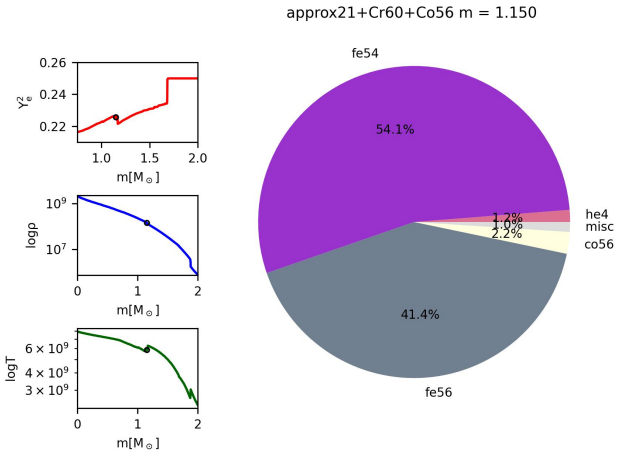
Massive stars usually end their life in core collapse supernovae (CCSNe) explosions where the core of the star collapses to form a neutron star (NS) or a black hole (BH) releasing gravitational energy during the collapse. The explosion mechanism of CCSNe has been broadly studied over the past decades (e.g., Bethe & Wilson 1985; Langer & Woosley 1996; Heger et al. 2000; Woosley & Janka 2005; Papish & Soker 2011; Janka et al. 2016; Kresse et al. 2021; Shishkin & Soker 2023). However it is still unclear how the gravitational energy from the collapse of the core is able to explode the star. One of the main challenges of modelling CCSNe explosions is the computational bottleneck of elements nucleosynthesis prior to and during the explosion. Due to the large number of isotopes that are formed modelling nucleosynthesis together with other physical processes crucial for stellar evolution with reasonable accuracy is hardly impossible even using powerful supercomputers. Therefore, in order to explore the explosion mechanism it is common to use small nets of isotopes, which leads to inaccurate results regarding the nucleosynthesis and energetic of the explosion.

The goal of our project is to improve the current understanding of supernova explosion physics by improving the accuracy of modelling the late stages of stellar nucleosynthesis. We are working on overcoming the nucleosynthesis computational barrier by replacing the solvers of the complicated differential equations of nucleosynthesis in stellar evolution codes with neural networks. Such an approach has the potential to accelerate detailed computation of nucleosynthesis by many orders of magnitude, hence enabling the inclusion of important physical processes that could not be included so far. This will enable the exploration of supernova explosions with stellar evolution codes and could shed light on the explosion mechanism and even on the formation of BHs.

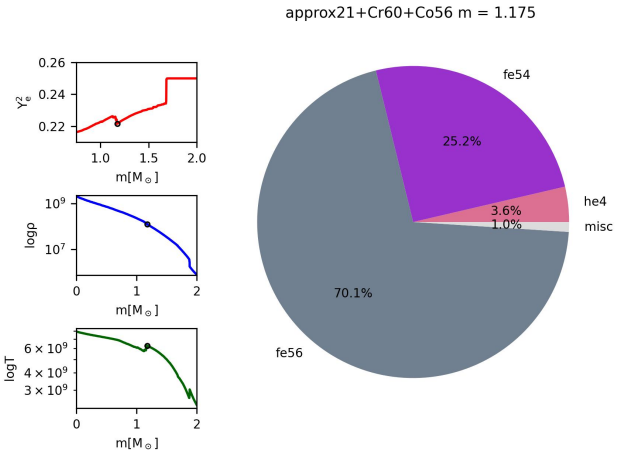
2. MODELLING THE LATE STAGES OF STELLAR NUCLEOSYNTHESIS

We use the stellar evolution code MESA (Modules for Experiments in Stellar Astrophysics; Paxton et al. 2011; Paxton et al. 2013; Paxton et al. 2015; Paxton et al. 2018; Paxton et al. 2019; Jermyn et al. 2023) to evolve a single star of mass $M_{\rm ZAMS} = 20 M_\odot$ and solar metallicity $Z = 0.02$ from zero age main sequence (ZAMS) to core collapse. We perform three separate runs of this stellar model using nets with different numbers of isotopes (22 isotopes, 80 isotopes and 151 isotopes) and explore the elemental composition and stellar profiles right before the collapse of the core. We examine the results of the electron fraction $Y_{\rm e}$ in late core burning phases as it is crucial to determine the 'explodability' of stars since electron captures lead to the collapse of the core prior to the CCSNe event.

The usual test case that have been used to evolve massive stars until core collapse with MESA contains 21 or 22 isotopes at the most. We show that this is problematic since that amount of isotopes does not provide enough accuracy to resolve nucleosynthesis in massive stars. One way this is manifested is by a non-physical behaviour of the electron fraction $Y_{\rm e}$ throughout the core. As we can see by looking at the upper left subpanel of Fig. 1a and Fig. 1b (red curves), in a certain part of the core $Y_{\rm e}$ drops outwards even though the material should be more neutronized as we go deeper in the core. This is since heavier stable isotopes, which are richer in neutrons, tend to be formed in hotter and denser areas where it is easier to overcome the Coulomb barrier. The reason for the nonphysical behavior of $Y_{\rm e}$ in the 22 isotopes net lies in the fact that MESA imple-



(a) Before the drop in Y_e, approx22



(b) After the drop in Y_e, approx22

Figure 1. (a) Left panels: Y_e^2 (upper panel), density (middle panel) and temperature (lower panel) profiles. Right panel: The composition in mass coordinate $M = 1.15 M_\odot$ before the decrease in the electron fraction Y_e in the MESA run with 22 isotopes. (b) Left panels: The same as in (a). Right panel: The composition in mass coordinate $M = 1.175 M_\odot$ after the decrease in Y_e for the same run.

ments electron capture in a single reaction, and when T, ρ conditions for this reaction are not met it stops abruptly, leading to the discontinuity between two mass shells in the star. This drop in the electron fraction become less and less prominent as the numbers of isotopes in the net grows.

In Figs. 1a and 1b we can also see the composition right before and right after the drop in Y_e for a net of 22 isotopes. We note the significant decrease in Fe54 and increase in Fe56 right after the nonphysical decrease in Y_e. This interchange depends on the (many) particular isotope that we are not including in our nets and on the rates of "fake reactions" through which the nucleosynthesis is performed in cases where isotopes that should be part of the reactants or products in the formation of other isotopes are not included in the net (see Fig. 14 in Marchant et al. 2019).

To find the smallest net of isotopes that will give reasonable results for Y_e we ran BBQ, which is a wrapper to the MESA one zone burner keeping $T = 7 \times 10^9$ K and $\rho = 3 \times 10^8$ K constant for the entire run. We ran ten nets with different numbers of isotopes from 22 to 3335. To find from which initial composition we should start the runs, we performed experiments trying different initial compositions, e.g., a uniform distribution, compositions that start with a single isotope, and the more physical case of a composition taken from the full MESA run with a net of 151 isotopes. For runs with larger nets we assumed zero contribution to begin with from isotopes that are not included in mesa151. Since the larger isotope nets do not necessarily fully contain smaller nets, we made sure to normalize the composition. We found that after a long enough time, the runs converged to the same final composition regardless of the initial one (for a given net), i.e. the system reached nuclear statistical equilibrium. Our purpose in these runs was to find the smallest net of isotopes that still gives the same result for Y_e and for the final composition as the large nets, in order to understand which net of isotopes we should choose to generate the input database for our neural networks.

As can seen in Fig. 2, we found that the equilibrium value of Y_e for the net of 22 isotopes is very different then for larger nets of isotopes implying that including only 22 isotopes in simulations of massive stars is not enough, in accordance with our conclusions based on Fig. 1a and Fig. 1b. The equilibrium value of the electron fraction is around the same value for the rest of the larger nets we examined. Therefore, we decided that we will begin our project by training our neural networks to predict the electron fraction and composition at late core burning phases for a net that contain 80 isotopes.

We generate training sets for the neural networks by running BBQ with a net of 80 isotopes for periods of time that are longer than the convergence time of the strong nuclear reactions but shorter than typical times of weak nuclear interactions. We sample the temperatures and densities quasi-randomly in the region 10^8 K $<$ T $<$ $10^{9.9}$ K, 10^7 g cm^{-3} $< \rho <$ 10^9 g cm^{-3}. We generate the compositions randomly since creating a quasi-random

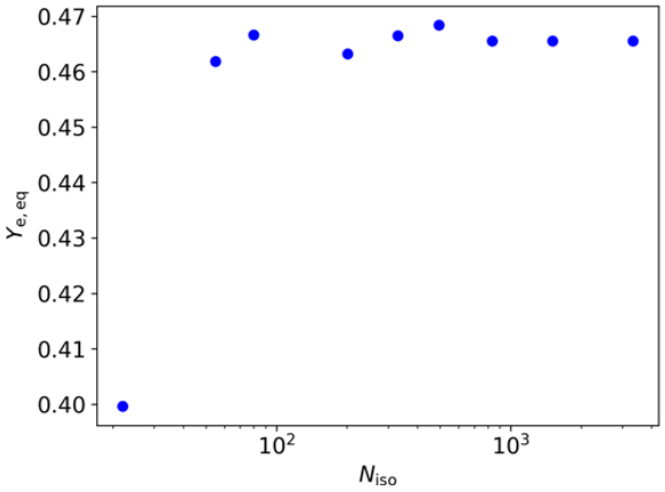


Figure 2. Equilibrium electron fraction at the end of the BBQ runs at fixed temperature and density vs the isotopes number in the net.

distribution of the composition is very challenging due to the fact that all the abundances in the composition are not independent (as they sum to 1). To further constrain our parameter space we intend to look at the typical composition from a MESA model, and set the main isotopes roughly at those values (within a few %) and pick the rest at random re-normalizing everything at the end.

3. A LITTLE DETOUR - A BUG IN MESA

Additionally to comparing the equilibrium value of the electron fraction for BBQ runs with different isotopes nets we examined the equilibrium compositions. We found that the equilibrium composition we get for running a net with 201 isotopes resembles the equilibrium abundances for runs with nets of 833-3335 isotopes. The equilibrium composition for the net of 80 isotopes is qualitatively consistent with these compositions as well. However, we got strange results for the final composition of two medium size nets (330 isotopes, 495 isotopes) that yield totally different results, as can be seen in Fig 3.

From comparing the lists of isotopes between the different nets we concluded that surprisingly, the presence of tritium in the medium nets (and its absence in the smaller and larger nets) led to this prominent difference in final compositions. Narrowing this down, we found that there is a specific reaction that causes this effect:

$$\mathrm{n} + \mathrm{n} + \mathrm{He}^4 + \mathrm{He}^4 \rightarrow \mathrm{h}^3 + \mathrm{Li}^7, \qquad (1)$$

which was even more surprising since the probability of a four-particle reaction is expected to be negligible. We found that the rate of this reaction in MESA is 20-24 orders of magnitude larger than the rate in the literature (Malaney & Fowler 1989).

We discovered that MESA miscalculates the rates of reactions that consume energy for all reactions involving more than two reactants and/or products. This is since MESA calculates these reaction rates from detailed balance considerations and seems to omit a phase space factor related to the number of particles involved in the reactions. Even though there is a relatively small number of reactions that this bug affects, it can have a huge impact on the nucleosynthesis results in high temperatures and densities in nets with a large number of isotopes. However, we do not expect this bug to affect most stellar evolution calculations, especially for models that include the standard (commonly used) isotopes nets. For more information on this bug you can see the GitHub issue we opened. We are currently working on fixing this bug.

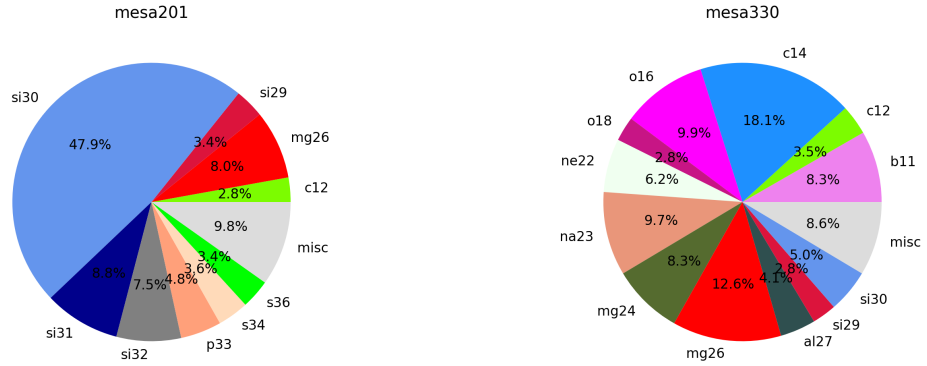
4. NUCLEAR NEURAL NETWORKS

We used PyTorch Lightning to build our nuclear neutral networks (NNNs). Our current neural networks architecture is composed from an input layer that contains the temperature, density and initial abundance of 80 isotopes, two hidden layers and an output layer of the composition after a fixed timestep (see Fig. 4). In the future, we will generalize our neural networks to include the timestep as an input, and to predict the energy lost by neutrinos during reactions that involve the weak nuclear force.

In Figs. 5a - 5c we present the results of the NNN trained on our largest training set that contains 6×10^4 data points for the current architecture. In the left panels we show the final composition obtained from our BBQ runs, and the right panels show the predictions of the NNN for the final composition. The best results so far (Fig. 5a) has 4% error in the composition, the average results (Fig. 5b) has 28% error and the worse results (Fig 5c) has 131% error. In Fig 6 we can see the relative error of the abundance losses for every isotope that is included in our initial net.

To test whether enlarging our data sets has the potential to substantially increase the accuracy of our results we plot the abundance loss function of the NNN as a function of the training set size in Fig. 7. We can see from the figure that the loss keeps decreasing as the training sets get larger, implying that we can significantly improve our results by generating more training sets, as we expected due to the high dimensionality of our problem.

5. SUMMARY AND DISCUSSION



(a) Equilibrium composition of mesa201

(b) Equilibrium composition of mesa330

Figure 3. (a) Composition in nuclear statistical equilibrium of a BBQ run where $T = 7 \times 10^9$ K, $\rho = 3 \times 10^8$ K for a net that contains 201 isotopes. (b) Composition in nuclear statistical equilibrium for a BBQ run with the same initial parameters except the initial net that contains 330 isotopes.

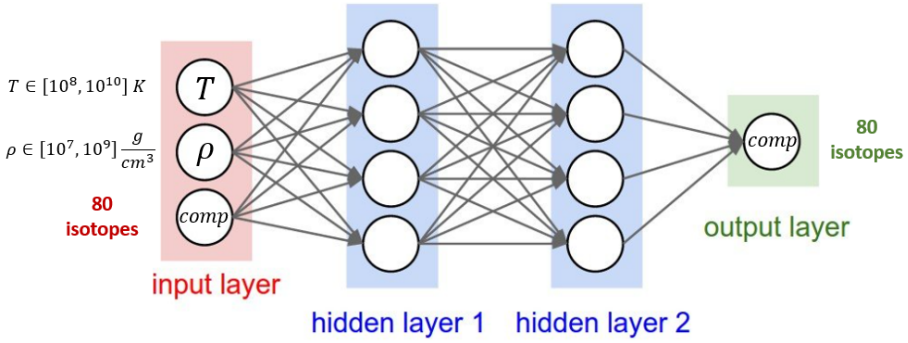
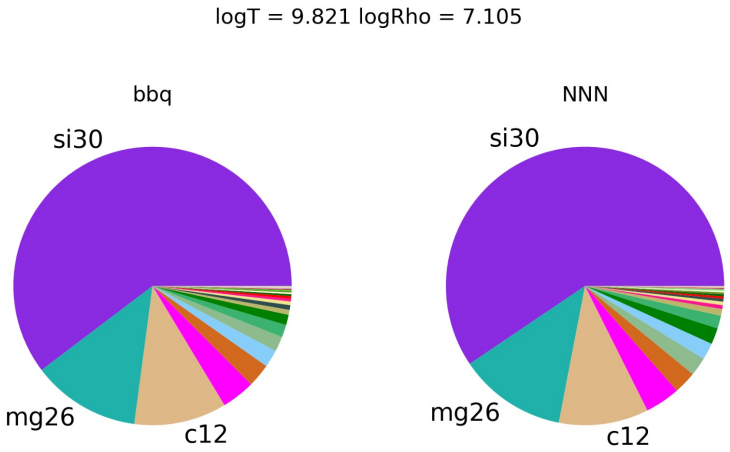


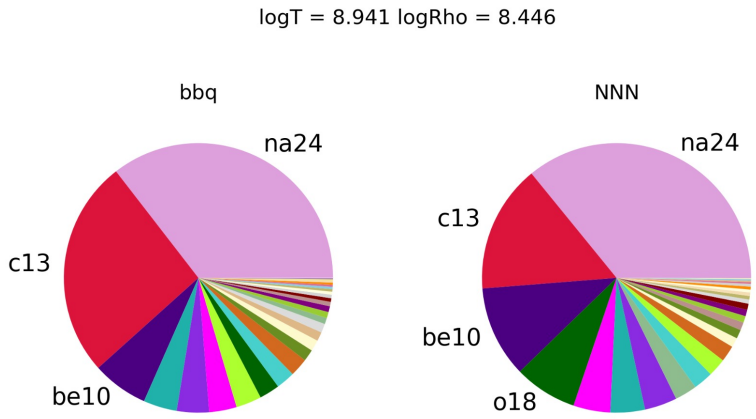
Figure 4. Our nuclear neural network (NNN) architecture. The NNN is composed two hidden layers. At the moment the input layer maps the temperature, density and initial abundance of 80 isotopes to 128 neurons, the first hidden layer maps 128 neurons to 256 neurons, the second hidden layer maps these 256 neurons to other 256 neurons and the output layer maps the 256 neurons to the final abundance of 80 isotopes. Adapted from Gaudio et al. (2021).

In this work we designed and trained neural networks to predict the abundances of isotopes at late core burning phases. In section 2 we used the stellar evolution code MESA to evolve a single star with ZAMS mass $M_{\rm ZAMS} = 20 M_\odot$ and showed the inadequacy of exploring CCSNe with small isotopes nets. We run BBQ and found that a net of 80 isotopes gives the same result for the electron fraction right before the collapse as nets with thousands of isotopes, making it sufficient to train our NNNs to predict the composition of BBQ runs with 80 isotopes given fixed temperatures and densities. We used PyTorch Lightning to build neural networks and trained them on sets that contain 6×10^4 data points (section 4).

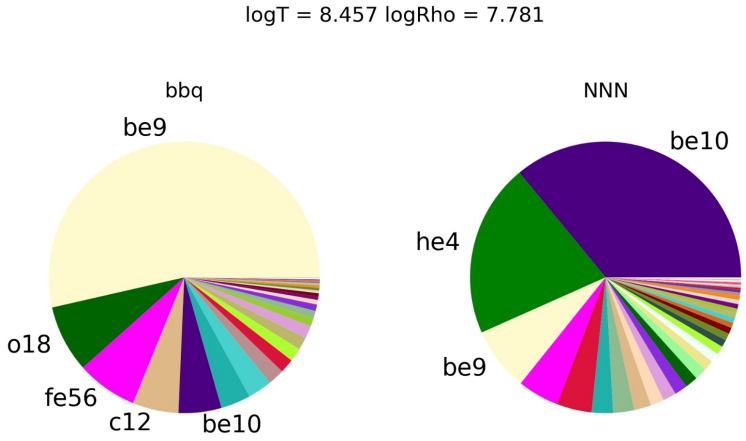
Even though we only reproduced the results qualitatively, based on Fig. 7 we are optimistic that by en-



(a) The best results from our nuclear neural networks (NNNs) so far. Left panel: The equilibrium composition obtained from running BBQ with logT = 9.821, logρ = 7.105 and uniform composition. Right panel: The prediction of our NNN for the same input parameters.



(b) The average result from our NNNs so far. Left and right panels are the same as in Fig 5a for logT = 8.941, logρ = 8.446



(c) The worst result from our NNNs so far. Left and right panels are the same as in Fig 5a for logT = 8.457, logρ = 7.781

Figure 5. Best (a), average (b) and worst (c) predictions of the equilibrium compositing by our nuclear neural networks.

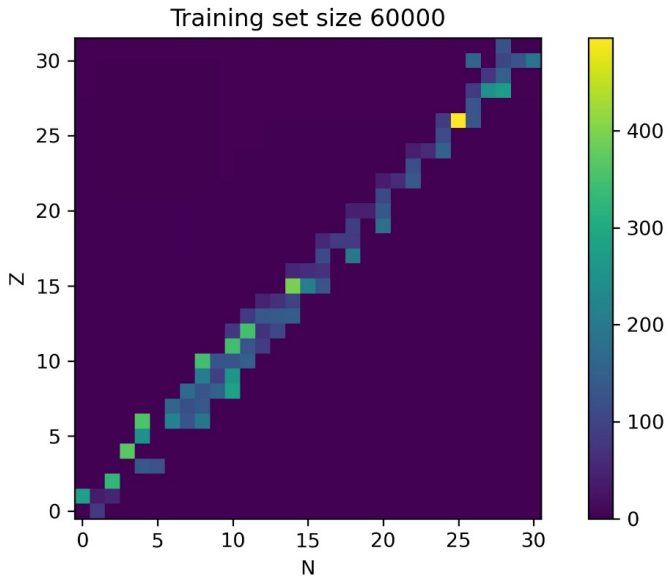


Figure 6. Relative error of the abundance losses of isotopes with N neutrons and Z protons for our largest data sample.

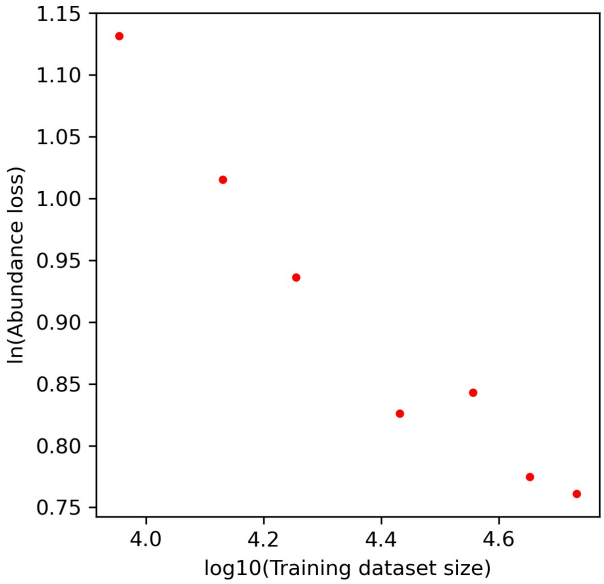


Figure 7. Abundance loss as a function of the training set size on a logarithmic scale.

larging our training sets and considering different neural networks architectures we will be able to sufficiently increase the accuracy of our results to the point of integrating the NNNs into MESA, hence making stellar nucleosynthesis computations faster and more robust. To enable this we intend to train our NNNs to predict the composition for arbitrary timesteps, and to predict the energy lost by neutrinos during the nucleosynthesis of elements. When we sufficiently increase the accuracy of our results, the NNNs could be used to replace the nucleosynthesis computations in hydrodynamical simulations as well. Our main conclusion from the work performed during the Kavli summer program is that neural networks are a promising tool for nucleosynthesis computations and could be used to successfully model late core burning phases in stars, shedding light on the physics of CCSNe explosions.

ACKNOWLEDGEMENTS

We thank Alexander Heger for helpful discussion. We thank the Max-Planck-Institut für Astrophysik for its hospitality while part of this work was completed, and the Kavli Foundation for supporting the 2023 Kavli Summer Program.

DATA AVAILABILITY

The data underlying this article will be shared on reasonable request to the corresponding author.

REFERENCES

Bethe, H. A. & Wilson, J. R. 1985, ApJ, 295, 14. doi:10.1086/163343

Gaudio, M. T., Coppola, G., Zangari, L., et al. 2021, Earth Systems and Environment, 5, 385. doi:10.1007/s41748-021-00220-x

Heger, A., Langer, N., & Woosley, S. E. 2000, ApJ, 528, 368. doi:10.1086/308158

Janka, H.-T., Melson, T., & Summa, A. 2016, Annual Review of Nuclear and Particle Science, 66, 341. doi:10.1146/annurev-nucl-102115-044747

Jermyn, A. S., Bauer, E. B., Schwab, J., et al. 2023, ApJS, 265, 15. doi:10.3847/1538-4365/acae8d

Kresse, D., Ertl, T., & Janka, H.-T. 2021, ApJ, 909, 169. doi:10.3847/1538-4357/abd54e

Langer, N. & Woosley, S. E. 1996, From Stars to Galaxies: the Impact of Stellar Physics on Galaxy Evolution, 98, 220

Malaney, R. A. & Fowler, W. A. 1989, ApJL, 345, L5. doi:10.1086/185538

Marchant, P., Renzo, M., Farmer, R., et al. 2019, ApJ, 882, 36. doi:10.3847/1538-4357/ab3426

Papish, O. & Soker, N. 2011, MNRAS, 416, 1697. doi:10.1111/j.1365-2966.2011.18671.x

Paxton, B., Smolec, R., Schwab, J., et al. 2019, ApJS, 243, 10. doi:10.3847/1538-4365/ab2241

Paxton, B., Schwab, J., Bauer, E. B., et al. 2018, ApJS, 234, 34. doi:10.3847/1538-4365/aaa5a8

Paxton, B., Marchant, P., Schwab, J., et al. 2015, ApJS, 220, 15. doi:10.1088/0067-0049/220/1/15

Paxton, B., Cantiello, M., Arras, P., et al. 2013, ApJS, 208, 4. doi:10.1088/0067-0049/208/1/4

Paxton, B., Bildsten, L., Dotter, A., et al. 2011, ApJS, 192, 3. doi:10.1088/0067-0049/192/1/3

Shishkin, D. & Soker, N. 2023, MNRAS, 522, 438. doi:10.1093/mnras/stad889

Woosley, S. & Janka, T. 2005, Nature Physics, 1, 147. doi:10.1038/nphys172