

Physics-informed neural networks for predicting the asymptotic outcome of fast neutrino flavor conversions

Sajad Abbar¹, Meng-Ru Wu^{2,3,4} and Zewei Xiong⁵

¹Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), Boltzmannstr. 8, 85748 Garching, Germany

²Institute of Physics, Academia Sinica, Taipei 11529, Taiwan

³Institute of Astronomy and Astrophysics, Academia Sinica, Taipei 10617, Taiwan

⁴Physics Division, National Center for Theoretical Sciences, Taipei 10617, Taiwan

⁵GSI Helmholtzzentrum für Schwerionenforschung, Planckstraße 1, D-64291 Darmstadt, Germany



(Received 28 November 2023; accepted 30 January 2024; published 14 February 2024)

In the most extreme astrophysical environments, such as core-collapse supernovae (CCSNe) and neutron star mergers (NSMs), neutrinos can undergo fast flavor conversions (FFCs) on exceedingly short scales. Intensive simulations have demonstrated that FFCs can attain equilibrium states in certain models. In this study, we utilize physics-informed neural networks (PINNs) to predict the asymptotic outcomes of FFCs, by specifically targeting the first two moments of neutrino angular distributions. This makes our approach suitable for state-of-the-art CCSN and NSM simulations. Through effective feature engineering and the incorporation of customized loss functions that penalize discrepancies in the predicted total number of ν_e and $\bar{\nu}_e$, our PINNs demonstrate remarkable accuracies, with an error margin of $\lesssim 3\%$. Our study represents a substantial leap forward in the potential incorporation of FFCs into simulations of CCSNe and NSMs, thereby enhancing our understanding of these extraordinary astrophysical events.

DOI: [10.1103/PhysRevD.109.043024](https://doi.org/10.1103/PhysRevD.109.043024)

I. INTRODUCTION

Core-collapse supernovae (CCSNe) and neutron star mergers (NSMs) are among the most extreme astrophysical phenomena, representing the death of massive stars' life cycles and the collision of extremely dense remnants, respectively. These events not only signify the demise of massive stars and dense objects but also reveal some of the universe's most energetic and mysterious phenomena.

At the heart of these extraordinary settings lies a captivating process: the emission of neutrinos, released in copious amounts during both CCSNe and NSMs [1–6]. During their propagation through the extreme conditions within these events, neutrinos undergo a fascinating phenomenon known as *collective* neutrino oscillations [7–13] (for a recent review see Ref. [14]). This intriguing behavior emerges from the intricate interplay between the propagating neutrinos and the dense background neutrino gas, where coherent forward scatterings play a crucial role. This

phenomenon behaves in a nonlinear and collective manner, leading to a complex tapestry of flavor conversions.

Of particular interest are the so-called *fast* flavor conversions (FFCs), which occur on scales characterized by $\sim (G_F n_\nu)^{-1}$ (see, e.g., Refs. [15–64]). Here, G_F represents the Fermi coupling constant, and n_ν denotes the neutrino number density. These FFCs can take place on scales much shorter than what would be expected in the vacuum. FFCs occur *iff* the angular distribution of the neutrino lepton number, defined as,

$$G(\mathbf{v}) = \sqrt{2}G_F \int_0^\infty \frac{E_\nu^2 dE_\nu}{(2\pi)^3} [(f_{\nu_e}(\mathbf{p}) - f_{\nu_x}(\mathbf{p})) - (f_{\bar{\nu}_e}(\mathbf{p}) - f_{\bar{\nu}_x}(\mathbf{p}))], \quad (1)$$

crosses zero at some $\mathbf{v} = \mathbf{v}(\mu, \phi_\nu)$, with $\mu = \cos \theta_\nu$ [34]. Here, E_ν , θ_ν , and ϕ_ν are the neutrino energy, the zenith, and azimuthal angles of the neutrino velocity, respectively. The f_ν 's are the neutrino occupation numbers of different flavors, with ν_x and $\bar{\nu}_x$ denoting the heavy-lepton flavor of neutrinos and antineutrinos. When ν_x and $\bar{\nu}_x$ have similar angular distributions, a scenario commonly observed in state-of-the-art CCSN simulations, this expression transforms into the conventional definition of the neutrino electron lepton number, νELN .

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

FFCs tend to occur on spatial and temporal scales which are expected to be significantly shorter than those typically addressed in hydrodynamical simulations of CCSNe and NSMs. As a consequence, integrating FFC into these simulations presents a notable challenge. A potential strategy to address this challenge involves breaking down the problem into two scale hierarchies. This consideration motivates conducting local dynamical simulations at shorter scales and subsequently integrating the findings into some practical prescriptions. Such prescriptions can then be efficiently applied to the broader astrophysical modelings and hydrodynamic simulations [29,45,57,58,65–67].

The assessment of the outcome of FFCs has undergone thorough examinations of local dynamical simulations conducted within confined spaces employing periodic boundary conditions [44,48,51–54,68–73] (see also Ref. [74] for the possible impact of the choice of boundary conditions). Insights from these investigations indicate a tendency toward kinematic decoherence in flavor conversions, generally resulting in quasistationary states. These stationary states can be characterized by survival probabilities, which are governed by the conservation of neutrino lepton number and have been demonstrated to be possibly modeled by analytical formulation to a good accuracy [73].

Despite the existence of such analytical formulas for the angular distribution of survival probabilities, implementing FFCs in CCSN and NSM simulations remains a challenge. The obstacle lies in the requirement of having access to complete angular distributions of neutrinos to determine FFC outcomes through these analytical expressions. However, acquiring such detailed angular information proves challenging in most cutting-edge CCSN and NSM simulations due to their computationally intensive nature.

As a practical alternative to considering the full neutrino angular distributions, many state-of-the-art simulations opt for a more feasible approach by simplifying neutrino transport through a limited set of angular distribution moments [75–77]. In our specific investigation, we concentrate on radial moments, defined as,

$$I_n = \int_{-1}^1 d\mu \mu^n \int_0^\infty \int_0^{2\pi} \frac{E_\nu^2 dE_\nu d\phi_\nu}{(2\pi)^3} f_\nu(\mathbf{p}). \quad (2)$$

These moments effectively capture crucial aspects of the neutrino angular distribution while facilitating a computationally more manageable treatment.¹

In practical scenarios, one often encounters a situation where simulations directly provide only the first two moments, I_0 and I_1 , for (anti)neutrinos. The problem then

¹Our primary focus is presently on axisymmetric crossings, particularly emphasizing radial moments where the angular distribution integrates over ϕ_ν . It is crucial to highlight that our current study excludes nonaxisymmetric crossings. Exploring these aspects is a subject reserved for future investigations.

becomes determining the ultimate values of I_0 and I_1 following FFCs, based on their initial states. Note that despite the availability of analytical formulas for the angular distribution of the neutrino survival probabilities, determining these final values is inherently complex.

This paper represents a pioneering effort in predicting the asymptotic outcomes of FFCs in the moments scenario, using artificial neural network (NNs). NNs imitate closely the brain's network of connected neurons. Specifically, they have layers of artificial neurons that handle information. NNs have been proven to be useful in solving tricky problems due their strong learning capacities, resulting from adjusting the connections between neurons during the training phase. Their ability to learn from data without explicit programming sets NNs apart, making them attracting tools across various domains. In particular, NNs have been extensively used in the field of astrophysics and high energy physics [78,79].

Our approach involves the utilization of a NN, which takes the essential information extracted from the initial (anti)neutrino zeroth and first moments and then outputs the corresponding moments regarding the asymptotic outcome of FFCs. In particular, we employ physics-informed neural networks (PINNs), where the learning and performance of the NN can be enhanced with the utilization of the domain knowledge (specialized information specific to the problem that can be integrated into the NN) [80–82]. Our findings demonstrate the efficacy of a single hidden layer PINN, achieving a remarkable accuracy for the asymptotic values of I_0 and I_1 .

The paper is structured as follows. In Sec. II, we initiate by detailing our simulations of FFCs and elucidating the assumptions to deriving the outcomes of FFCs. Moving forward in Sec. III, we delve into the architecture of our NNs, shedding light on the requisite feature engineering and the deployment of customized loss functions. The ensuing discussion encompasses the results gleaned from both two- and three-flavor scenarios. Finally, we conclude in Sec. IV.

II. FFCs SIMULATIONS

To effectively train our NN, we require a substantial number of training samples containing initial values of the (anti)neutrino moments, I_0 and I_1 , within the neutrino gas. These samples should also encompass their corresponding final values, reflecting the asymptotic outcomes of FFCs.

Our chosen physical model involves the evolution of FFCs within a one-dimensional (1D) box, mirroring the setup outlined in Ref. [52]. This model assumes translation symmetry along the x and y axes, axial symmetry around the z axis, and periodic boundary conditions in the z direction. Notably, we omit considerations of vacuum mixing and neutrino-matter forward scattering in this model.

In this study, we prime the neutrino gas for tracking its flavor evolution by employing two widely utilized

parametric neutrino angular distributions documented in existing literature. The first one is the maximum entropy distribution defined as,

$$f_\nu^{\text{max-ent}}(\mu) = \exp[\eta + a\mu], \quad (3)$$

where we here consider the ϕ_ν -integrated distribution, i.e.,

$$f_\nu(\mu) = \int_0^\infty \int_0^{2\pi} \frac{E_\nu^2 dE_\nu d\phi_\nu}{(2\pi)^3} f_\nu(\mathbf{p}). \quad (4)$$

This is a very natural choice for the neutrino angular distribution since the maximum entropy closure [83] is currently very popular in the moment-based neutrino transport methods. This parametric distribution has been also used to detect ν ELN crossings using fitting and machine learning techniques [84–86]. Another angular distribution considered in the literature of FFCs (see, e.g., Refs. [52,87]) is the Gaussian distribution defined as,

$$f_\nu^{\text{Gauss}}(\mu) = A \exp\left[-\frac{(1-\mu)^2}{\xi}\right]. \quad (5)$$

Note that both of these distributions have a parameter which determines the overall neutrino number density, namely η and A , and the other parameters determining the shape of the distribution, i.e., a and ξ . Allowing for two distinct forms of angular distributions takes into consideration potential deviations in the shape of neutrino angular distributions in realistic simulations, which can occur, e.g., due to the use of different closure relations.

To ready our datasets, we begin with the initial angular distributions of neutrinos, which can either follow a maximum entropy distribution or a Gaussian distribution. Subsequently, we utilize analytical neutrino survival probabilities to determine the asymptotic outcome of FFCs. By performing integration over the neutrino angular distributions, we can obtain the initial and final values of I_0 and I_1 .

In our analytical treatment of the survival probability, we follow closely our recent work in Ref. [73]. We assume that $G(\mu)(= \int_0^{2\pi} d\phi_\nu G(\mathbf{v}))$ has only one zero crossing μ_c . This helps us to define,

$$\begin{aligned} \Gamma_+ &= \left| \int_{-1}^1 d\mu G(\mu) \Theta[G(\mu)] \right|, \\ \Gamma_- &= \left| \int_{-1}^1 d\mu G(\mu) \Theta[-G(\mu)] \right|, \end{aligned} \quad (6)$$

as the integration of positive and negative parts of $G(\mu)$. Here Θ is the Heaviside theta function. In the following, we specify the μ range over which the above integral is smaller (larger) by $\mu^<(\mu^>)$. For the survival probability in the two-flavor scenario, we use the analytical formula:

$$P_{\text{sur}}^{2f}(\mu) = \begin{cases} \frac{1}{2} & \text{for } \mu^<, \\ \mathcal{S}(\mu) & \text{for } \mu^>, \end{cases} \quad (7)$$

where the distribution over $\mu^>$ is formulated as,

$$\mathcal{S}(\mu) = 1 - \frac{1}{2} h(|\mu - \mu_c|/\xi). \quad (8)$$

Here, $h(x)$ is a μ -dependent function that monotonically decreases from 1 to 0 when x increases from 0 to infinity. To be specific, we here assume $h(x)$ to have a power-1/2 form, i.e., $h(x) = (x^2 + 1)^{-1/2}$. In addition, the parameter ξ can be found such that the survival probability function is continuous.

In the three-flavor case where ν_μ and ν_τ are indistinguishable, one can simply find the survival probabilities by the expression $P_{\text{sur}}^{3f}(\mu) = 1 - 4[1 - P_{\text{sur}}^{2f}(\mu)]/3$.

As illustrated in Table. 1 of Ref. [73], adopting a power-1/2 form for the survival probability proves to yield a comparatively low error in computing I_0 and I_1 analytically. This explains the rationale behind opting for this analytical survival probability in this work. In the concluding part of Sec. III, we also explore the scenario where the outcomes derived from actual simulations of FFCs are applied.

III. APPLICATIONS OF NEURAL NETWORKS

Before unveiling our findings, it is crucial to emphasize that to ensure high performance of our NN models in the test set, it is necessary to divide the dataset into three distinct sets. These sets are defined as follows: (i) training set: This set serves as the foundation for training the NN, allowing it to learn and adapt based on the provided data. (ii) development set: Also known as the validation set, this subset plays a pivotal role in determining the optimal hyper-parameters of the algorithm. It serves as a testing ground to fine-tune the model for optimal performance. (iii) test set: To assess the NN efficacy on novel, unseen data, the test set is utilized. This set provides a critical evaluation of the model's generalization capabilities beyond the training data.

A. The architecture of NNs

For a given arbitrary neutrino gas, one is provided with the initial values of I_0 's and I_1 's of ν_e , $\bar{\nu}_e$, and ν_x . In this context, we make the assumption that the initial distributions of $\bar{\nu}_x$ and ν_x are identical (though their final ones following FFCs could be different), a simplification that aligns with the majority of state-of-the-art CCSN and NSM simulations. In order to enhance the performance of our NNs, we introduce a layer of feature engineering, employing the following features as pertinent inputs in our NNs:

$$\alpha, \quad \alpha_{\nu_x}, \quad F_{\nu_e}, \quad F_{\bar{\nu}_e}, \quad \text{and} \quad F_{\nu_x},$$

with: $\alpha = \frac{n_{\bar{\nu}_e}}{n_{\nu_e}}, \quad \alpha_{\nu_x} = \frac{n_{\nu_x}}{n_{\nu_e}}, \quad \text{and} \quad F_{\nu} = \left(\frac{I_1}{I_0} \right)_{\nu}$ (9)

Note that the selection of these features offers explicit insights into the configuration of neutrino angular distributions, which plays a crucial role in understanding the asymptotic outcome of FFCs. Moreover, with the provided values of I_0 and I_1 for a specific neutrino species (derivable from $\alpha_{(\nu_x)}$ and F_{ν}), we use the root-finding function *fsolve* in PYTHON to determine the complete shape of the neutrino angular distribution, being either maximum entropy or the Gaussian one. Furthermore, it is worth highlighting that all quantities in this context are normalized by the initial ν_e number density, allowing the convenient choice of setting it to $n_{\nu_e}^{\text{initial}} = 1$. This simplification reduces the number of inputs to our NNs, and notably, there is no input parameter related to n_{ν_e} .

Though the aforementioned features serve as a necessary foundation for developing a NN, there remains room for further enhancement through more advanced feature engineering to optimize the performance of our NNs. This optimization can be achieved by gaining insights from the neutrino survival probability's shape, as expressed in Eq. (7). Substantial information pertaining to the distribution of the survival probability can be derived by learning the position of μ_c . Another valuable piece of information, given μ_c , is determining the specific side of μ_c on which equipartition occurs, while the behavior of the survival probability on the other side is governed by conservation laws. Information regarding the side on which equipartition occurs is provided in the quantity E_{RL} , a binary number which is 1 if the equipartition occurs for $\mu_c \leq \mu$, and 0 otherwise.

In our NN framework, we explore two distinct architectures, as illustrated in Fig. 1. In the foundational architecture, we integrate only $\alpha, \alpha_{\nu_x}, F_{\nu_e}, F_{\bar{\nu}_e},$ and F_{ν_x} into our NN. An alternative NN that we examine involves an additional layer of feature engineering, as discussed in the preceding paragraph, encompassing information about μ_c and E_{RL} .

Practically, this augmentation is accomplished by constructing a *separate* regression model trained on our dataset, from which information regarding μ_c and E_{RL} can be readily extracted. We have confirmed that the computation of E_{RL} and μ_c can be performed with small errors.

As illustrated in Fig. 1, our feedforward NN has a single hidden layer containing 50 neurons, unless stated otherwise. The rationale for this choice is illustrated in Fig. 3 and the text around it. Also regarding the output layer, our NNs provide the values of I_0 and I_1 for both ν_e and $\bar{\nu}_e$, effectively utilizing a total of 4 neurons. The determination of I_0 and I_1 for ν_x and $\bar{\nu}_x$ can be deduced by applying the

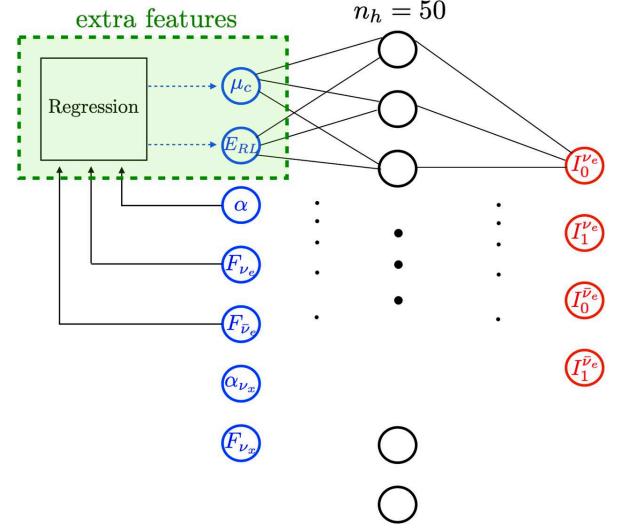


FIG. 1. Schematic architecture of our NNs. The green zone shows the implementation of the extra features, μ_c , and E_{RL} , which are obtained through an extra layer of regression, using linear and logistic regressions, respectively. Here, μ_c is the crossing direction and E_{RL} is a binary, which is 1 if the equilibrium occurs for $\mu_c \leq \mu$, and 0 otherwise. In our basic NN, referred to as the NN with no extra features, the NN only takes the inputs highlighted in Eq. (9). However and in our PINN, we provide our NN with the extra features μ_c and E_{RL} .

principles of neutrino and antineutrino number density and momentum conservation. In simpler terms, our NN's ensures that the fundamental laws governing neutrino conservation are respected without any exceptions.

B. Loss functions

The loss function, \mathcal{L} , is a crucial component in training NNs, serving as a measure of the model's predictive performance. It quantifies the disparity between predicted values and actual target values, providing a guide for the model to adjust its parameters during the optimization process.

When it comes to neutrino flavor conversions in CCSNe and NSMs, a critical parameter of utmost significance is the number of neutrinos in the electron channel, i.e., $N_{\nu_e + \bar{\nu}_e} = n_{\nu_e} + n_{\bar{\nu}_e}$, as opposed to the number of neutrinos in the heavy-lepton channel, $N_{\nu_x + \bar{\nu}_x} = n_{\nu_x} + n_{\bar{\nu}_x}$. Leveraging this crucial physical insight to enhance the performance of our NN, we incorporate an additional loss term in the optimization of the NN model with the extra features. This loss term is designed to penalize discrepancies in $N_{\nu_e + \bar{\nu}_e}$, and is defined as,

$$\mathcal{L}_{\text{extra}} = \frac{1}{N_{\text{sample}}} \sum_i (\Delta N_{\nu_e + \bar{\nu}_e, i})^2, \quad (10)$$

where Δ , N_{sample} , and \sum_i denote the difference between the true and predicted values, the number of samples in the

training set, and the summation over the training samples, respectively. This loss term provides an additional constraint for the model. The integration of the domain knowledge characterizes this particular NN architecture as a PINN, given that its distinctive nature is shaped by our insights into the underlying physics of the problem [80–82]. The PINN should be compared with our basic NN, referred to as NN with *no extra features*, for which the loss term only includes the ordinary mean squared errors of the output parameters. Note that the incorporation of the domain knowledge in our PINN includes both architectural aspects, utilizing additional features, and learning-based enhancements via the loss function. It's also important to highlight that unlike what typically observed in PINNs, our approach does not involve a loss term associated with some partial differential equations.

C. Three-flavor scenario

In this section, we discuss the evaluation of our NNs concerning their predictions for the asymptotic outcome of FFCs in the three-flavor scenario. To train and assess our model, we utilize a dataset comprising a well-balanced

combination of maximum entropy and Gaussian initial neutrino angular distributions. The total size of our dataset is 2×10^5 points. The ultimate outcome of FFCs is determined through a three-flavor survival probability, as detailed in Eq. (7) and the surrounding text. To better encapsulate realistic conditions regarding the values of n_ν 's and the hierarchy among F 's, we prepare each sample by randomly selecting the inputs for our NNs. Specifically, we set $\alpha \in (0, 2.5)$, $\alpha_{\nu_x} \in (0, 3)$, $F_{\nu_x} \in (0, 1)$, $F_{\bar{\nu}_e} \in (0.4F_{\nu_x}, F_{\nu_x})$, and $F_{\nu_e} \in (0.4F_{\bar{\nu}_e}, F_{\bar{\nu}_e})$. This selection process ensures consistency with the expected hierarchy $F_{\nu_e} \lesssim F_{\bar{\nu}_e} \lesssim F_{\nu_x}$, characteristic of CCSN environment. Given these quantities, one can then determine the initial angular distributions of neutrinos. Using the analytical survival probability, the final I 's can be derived.

In the left panels of Fig. 2, we present the performance results of our PINN model. Here, an epoch refers to a single pass through the entire training dataset during the training phase. Notably, the relative error in the electron neutrino number density, defined as $|\Delta(n_{\nu_e} + n_{\bar{\nu}_e})|/(n_{\nu_e} + n_{\bar{\nu}_e})$, can attain values as low as 2.5%. Furthermore, we observe that the mean absolute error in the output quantities, defined as

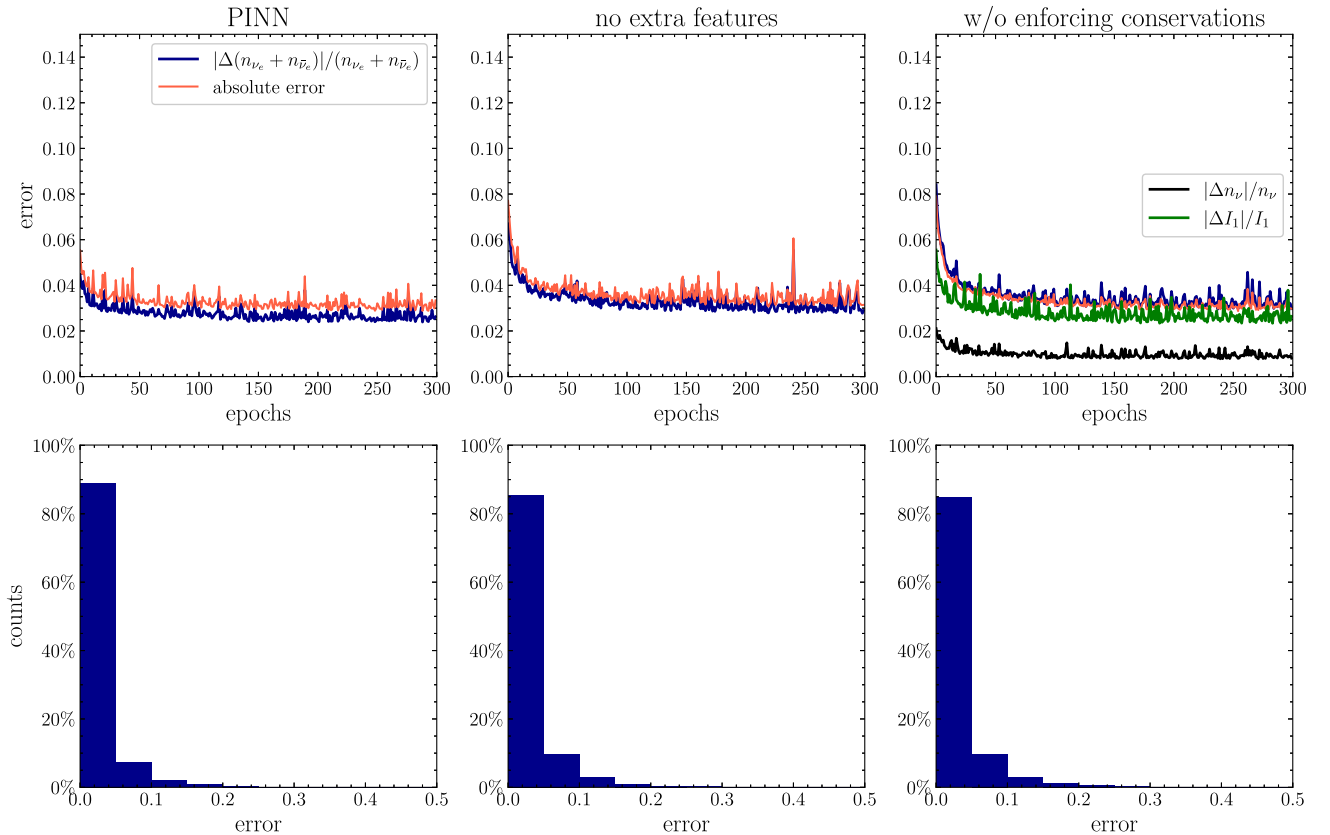


FIG. 2. Performance evaluation of various NN models. In the upper panels, we present the absolute error in the output parameters of the NN models, along with the relative error in the total number of neutrinos within the electron channel, $N_{\nu_e + \bar{\nu}_e}$. Moving to the lower panels, one can find the distribution in the error of $N_{\nu_e + \bar{\nu}_e}$. The results are showcased for our PINN, our basic NN with no extra features, as well as a NN model that does not enforce the conservation laws for neutrinos, respectively. In the upper right panel, the black and green curves represent the relative error in the total number and its first moments for (anti)neutrinos.

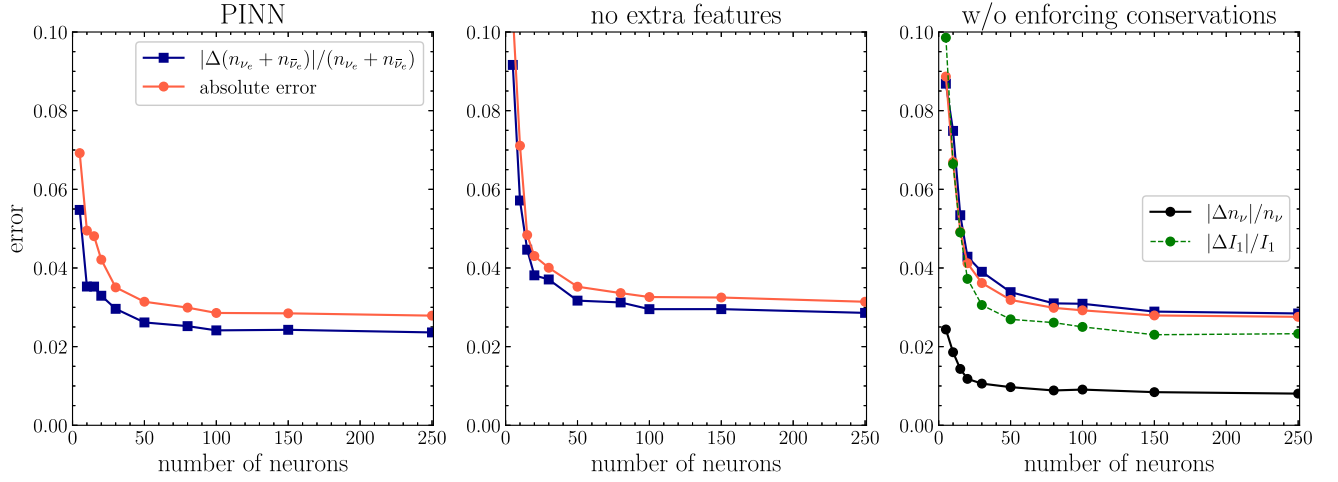


FIG. 3. Performance evaluation of various NN models (on the validation set) as a function of the number of neurons in the hidden layer. It is evident that the NNs perform best on the validation set once $n_h \gtrsim 50$. The labels and NN models are the same as those in Fig. 2.

$(|\Delta I_0^{\nu_e}| + |\Delta I_1^{\nu_e}| + |\Delta I_0^{\bar{\nu}_e}| + |\Delta I_1^{\bar{\nu}_e}|)/4$, can reach values $\sim 3\%$. Additionally, from the lower panel, it is evident that almost 90% of the predictions exhibit errors of $\lesssim 5\%$ in $N_{\nu_e + \bar{\nu}_e}$.

In the middle panels of Fig. 2, we present the performance of our basic NN with no extra feature and no loss term for enhancing the accuracy of $N_{\nu_e + \bar{\nu}_e}$. It is evident that the errors in this case are a bit greater than those of the PINN model. It is also worth noting that while the error in $N_{\nu_e + \bar{\nu}_e}$ is smaller than the absolute error in the output quantities, the gap between them has been reduced. This could be attributed to the lack of a specific loss term targeting the reduction of errors in $N_{\nu_e + \bar{\nu}_e}$. In addition and by examining the lower panel, we can observe that almost 85% of the predictions still exhibit errors of $\lesssim 5\%$ in $N_{\nu_e + \bar{\nu}_e}$.

Instead of rigidly enforcing the conservation of neutrino quantities, an alternative approach involves a more flexible NN model for which strict conservation laws are not forcefully respected. Instead, one can aim to derive all the neutrino moments (I_0 's and I_1 's for all flavors) as outputs, while introducing an additional loss term, which *effectively* enforces the conservation of (anti)neutrino number densities and momenta. This increased flexibility, combined with the loss term addressing conservation, has the potential to enhance the training of the NN, resulting in reduced errors in its predictions. Consider that this NN architecture could also be referred to as a PINN due to its loss term encompassing domain-specific knowledge, accounting for the conservation laws governing neutrino number density and momentum.

However, our findings emphasize the potential significant risks associated with applying such an informed NN to our specific problem. This concern is evident in the right panel of Fig. 2. While the absolute error and the error observed in $N_{\nu_e + \bar{\nu}_e}$ are comparable to what one observes in the left and middle panels, there is a new type of error

which appears in the total number of neutrinos and its first moment, soaring to values as high as 1%–3%. Such errors, whether in total neutrino number density or momentum, have the potential to distort the physics of CCSNe and NSMs. Note that such errors can generate greater risks in scenarios where a disparity might exist between the training and test datasets.

All calculations presented so far employ a feedforward NN with a single hidden layer containing $n_h = 50$ neurons. The rationale for this choice of the number of neurons is illustrated in Fig. 3, where different errors are shown for different NN architectures. It is evident that the NNs perform best on the validation set once $n_h \gtrsim 50$. It is also illuminating to note that if, for any reason such as computational constraints, a simpler NN with a smaller n_h is used, the performance of the PINN surpasses notably that of the model without additional features. However, this performance gap diminishes as larger n_h values are utilized. Furthermore, it is evident that even the model without explicitly enforced conservation laws achieves its optimal performance when $n_h \gtrsim 50$.

In Fig. 4, we present an analysis of the performance of our PINN as a function of the size of the training set. The red curve represents the absolute error in the PINN's output, while the blue curve illustrates the relative error in $N_{\nu_e + \bar{\nu}_e}$. It is noteworthy that as the training dataset expands to incorporate several thousand data points, the error rapidly diminishes to values below 5% and the difference between the error in the validation and training set disappears. This establishes the absolute minimum number of data points essential for conducting dependable calculations using NN's. However, it is important to bear in mind that this requisite number is expected to inherently grow as one explores increasingly intricate models involving more inputs and outputs. It is also interesting to observe that the performance of our NN remains satisfactory even

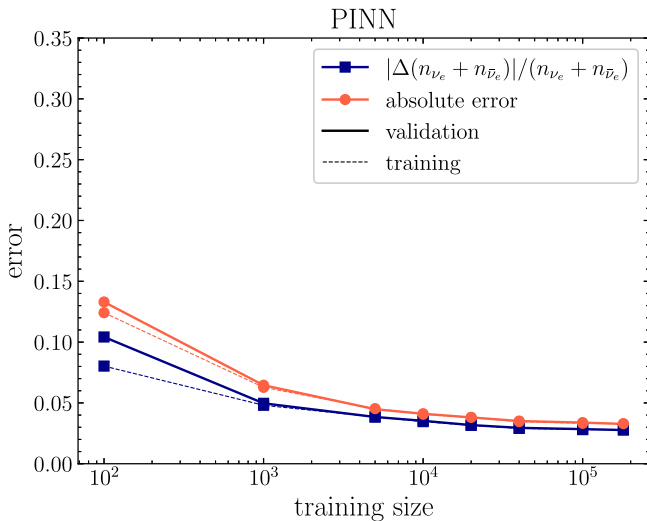


FIG. 4. Absolute error in the output of our PINN (red curve) vs the relative error in the number of neutrinos in electron channel, i.e., $N_{\nu_e + \bar{\nu}_e}$ (blue curve). Note that having a few thousand data points in the training set already leads to an error $\lesssim 5\%$, and also the disappearance of error variations between the validation and training sets.

when trained on relatively small datasets, comprising just a few hundred data points.

D. Two-flavor scenario

In this section, we conduct an evaluation of the performance of our NNs within a two-flavor scenario. The architectural configuration of the NNs and the training process closely mirror those detailed in the preceding section. A notable departure from the prior section lies in the fact that, in this case, we employ the two-flavor version of the survival probability [Eq. (7)]. Additionally, we have considered the results derived from our 1D box simulations of FFCs. These simulations encompass 10,000 data points characterized by the Gaussian angular distributions. It is crucial to emphasize that when working with the simulation results, no analytical prescription is employed for the survival probability. Instead, we derive the outcomes of FFCs directly from the simulations. This distinctive approach offers the advantage of enabling us to assess the reliability of the NN models trained on artificial data, when tested on outcomes derived from actual simulations of FFCs.

The left panels of Fig. 5 present the performance of our NNs trained using the artificial data, when tested on the simulation data. Notably, the PINN consistently outperforms the basic NN with no extra features, particularly in the error associated with $N_{\nu_e + \bar{\nu}_e}$. It is crucial to observe that the errors exhibit more pronounced variations compared to previous cases. This increased variability might be attributed to inherent systematic errors in the analytical prescription (for an estimation of the error associated with the analytical prescription, see Table. 1 of Ref. [73]).

During the evaluation of our NNs' performance on simulation data, we noted the critical importance of ensuring the equivalence between the input parameter space covered in the training set and the test set. This holds particularly true for the ranges of α and α_{ν_x} , as well as for the hierarchical relationship among F_ν 's. To elaborate further, if the test set encounters regions within the input parameter space that were completely unseen during the training phase, it could significantly degrade the performance of the NN, potentially resulting in very poor performance.

The middle and right panels of Fig. 5 display errors encountered during the computations in which both training and testing were performed on identical datasets—either artificial data generated using analytical formulas or data obtained from simulations. As discussed before, we have considered here a single form of neutrino angular distributions, namely the Gaussian one. This analysis provides insights into the inherent errors presented in each of the training sets. Then by comparing them with the left panels, one gets an idea of the error existing in the analytical formula.

The results depicted in Fig. 5 demonstrate the notable enhancement achieved by utilizing our PINN method in the case of two-flavor scenario. Specifically, a noticeable disparity is evident between the performance of PINN and that of the NN with no extra features, surpassing the distinctions observed in Fig. 2 for the three-flavor scenario. This fundamental discrepancy between two- and three-flavor scenarios highlights a greater degeneracy in the former, ultimately resulting in a more overall performance improvement when additional information, such as the implementation of PINN, is incorporated.

It is also worth noting a shift in the hierarchy between the absolute error and the error in $N_{\nu_e + \bar{\nu}_e}$, in the two- and three-flavor scenarios (compare Fig. 5 with the left and middle panels of Fig. 2). Although this observation is intriguing, it is crucial to recognize that comparing an absolute error with a relative error may not be entirely equitable. Such a hierarchy is anticipated to be sensitive to changes in the data, and thus, its intrinsic merit is limited. To address this concern, we investigated the hierarchy between the *relative* absolute error and the relative error in $N_{\nu_e + \bar{\nu}_e}$. This led us to find that the hierarchy remains consistent when considering these two types of errors, providing a fair basis for comparison.

Despite what discussed above, we opted to utilize absolute error instead of relative absolute error throughout this study for two primary reasons. First, in our calculations, we have already normalized all quantities by n_{ν_e} , resulting in having relative values for each quantity. This implies that any absolute error could be interpreted already to be relative in spirit. Second, to avoid excessive sensitivity to the small values associated with some of the neutrino quantities, we found it more appropriate to employ

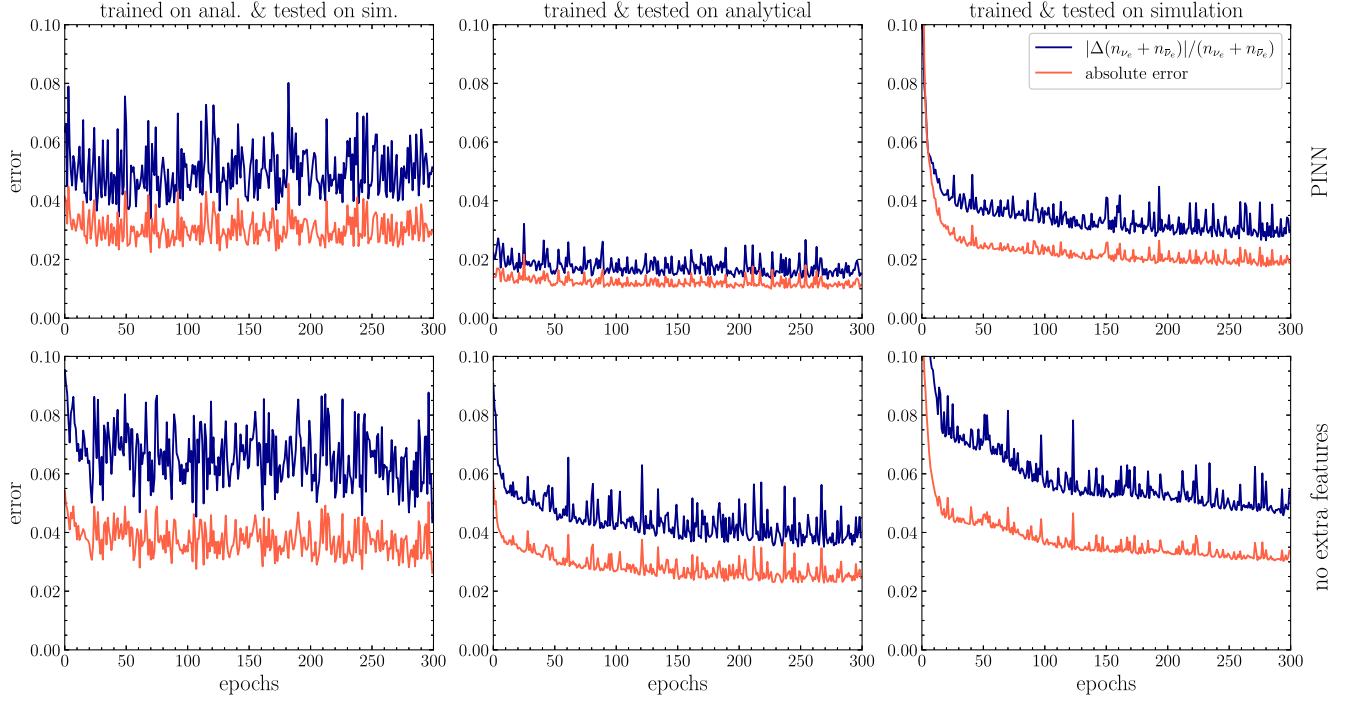


FIG. 5. The performance of our NNs in the two-flavor scenario for the PINN and the NN with no extra features. The left, middle, and right panels present the performance of the NN model trained using artificial data (with analytical prescription) and tested on simulation data, the one trained and tested on the artificial data, and the model trained and tested on the simulation data, respectively. Note that the PINN continues to outperform especially when it comes to the error in the total number of neutrinos in the electron channel.

absolute error as a metric in our study. This decision ensures a balanced and meaningful evaluation of our results.

It is important to note that our NNs are not specifically designed to capture ν ELN crossings. Hence, a practical implementation involves first employing one of the classical ML methods outlined in Refs. [85,86]. Subsequently, if fast modes are identified, our NNs can be effectively utilized. Despite this, given the broad range of our training set, which includes cases with both narrow and shallow crossings, our method is expected to provide accurate results, even in the absence of a crossing (where our approach should return outputs that are very close to the initial values).

IV. DISCUSSION AND OUTLOOK

Intensive simulations have demonstrated that FFCs can achieve equilibrium states in some models. In this study, we have employed neural networks (NNs) to predict the asymptotic outcome of FFCs in a three-flavor neutrino gas within a 1D box with periodic boundary conditions. Specifically, our focus was on the first two moments of neutrino angular distributions as inputs/outputs, making our NN models applicable to cutting-edge CCSNe and NSM simulations. We have shown that our NNs can predict the asymptotic outcomes of the (anti)neutrino I_0 's and I_1 's with a notable accuracy, corresponding to an error of $\lesssim 3\%$.

In order to enhance the performance of our NNs, we implement some novel features aiming at capturing the characteristics of the expected neutrino survival probability distributions. First, we incorporate a new feature related to the position of the zero crossing in the distribution of ν ELN, μ_c . Additionally, we introduce another feature indicating on which side of μ_c the expected equipartition between different neutrino flavors occurs. Both of these features are derived through a layer of regression applied to the initial inputs of the NN (see Fig. 1).

In the context of neutrino flavor conversions in CCSNe and NSMs, a critical parameter is the quantity of neutrinos and antineutrinos in the electron channel. To further optimize our NNs, we incorporate a supplementary loss term penalizing any discrepancies in predicting $N_{\nu_e + \bar{\nu}_e}$. The results demonstrate a relative improvement in our customized physics-informed neural network (PINN) due to the incorporation of extra features and a tailored loss function, outperforming a basic neural network that uses a standard mean squared error loss function and lacks these extra features.

We have also conducted a comprehensive evaluation of the performance of our NNs focusing on the variance between the training and validation sets (Fig. 4). Our findings reveal that the observed variance almost disappears when considering a minimum of a few thousand data points. This establishes an absolute minimum number of

data points essential for developing a dependable NN for predicting the outcome of FFCs in our model.

An intriguing observation arising from our study is that, even with the utilization of relatively small datasets, the variance remains modest, with an associated error limited to $\lesssim 15\%$. This insight further emphasizes the potential of NNs in scenarios where obtaining extensive datasets may be challenging or resource-intensive.

Instead of rigidly adhering to the strict conservation of neutrino quantities, we have also assessed the performance of a NN with a more flexible approach, where all the neutrino moments (I_0 's and I_1 's for all flavors) are treated as outputs. Here, we introduced an additional loss term effectively ensuring the conservation of (anti)neutrino number densities and momenta. However, our research has underscored noteworthy concerns associated with the application of such an informed NN to our specific problem. Indeed, we have shown that there could exist unignorable errors in the total number of neutrinos and their first moments, indicating a capacity to distort the physics of CCSNe and NSMs.

In our research, our NN models were predominantly trained on artificial data derived from two initial parametric angular distributions: the maximum entropy and Gaussian distributions. Additionally, we assessed the performance of our NN models using simulation data in a two-flavor scenario. Our findings indicate that the observed errors can be rationalized by accounting for the anticipated discrepancies between analytical and numerical results, as well as the inherent errors present in the training set.

In summary, our findings underscore the viability of NNs in forecasting the asymptotic outcomes of FFCs, once only the initial two moments of neutrino angular distributions are taken into account. This marks a significant advancement in the potential integration of FFCs into simulations of CCSNe and NSMs. However, there are still vital avenues for further exploration. First, our NN models were notably constrained to scenarios where neutrino distributions were assumed to be axisymmetric. Additionally, we operated

under the assumption that ν_x and $\bar{\nu}_x$ exhibit similar distributions. Relaxing these assumptions necessitates access to training datasets derived from actual simulations of FFC evolution in models without imposed axisymmetry, and where ν_x and $\bar{\nu}_x$ distributions may differ. Furthermore, our results are founded on a single-energy neutrino gas, prompting a crucial question regarding expectations in a multienergy neutrino environment. This consideration is especially relevant, given that almost all practical applications involve predicting FFC outcome regarding the neutrino energy spectrum. Given the efficacy of NNs in this domain, taking these crucial steps enhances remarkably the feasibility of incorporating FFCs into CCSN and NSM simulations.

ACKNOWLEDGMENTS

We are deeply grateful to Georg Raffelt for insightful conversations and reading our manuscript. We also thank Gabriel Martínez-Pinedo and Oliver Just for fruitful discussions. S. A. was supported by the German Research Foundation (DFG) through the Collaborative Research Centre “Neutrinos and Dark Matter in Astro- and Particle Physics (NDM),” Grant NO. SFB-1258, and under Germany’s Excellence Strategy through the Cluster of Excellence ORIGINS EXC-2094-390783311. M.-R. W. acknowledges supports from the National Science and Technology Council under Grant No. 111-2628-M-001-003-MY4, the Academia Sinica under Project No. AS-CDA-109-M11, and the Physics Division, National Center for Theoretical Sciences, as well as the resource of the Academia Sinica Grid-computing Center (ASGC). Z. X. was supported by the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (ERC Advanced Grant KILONOVA No. 885281). We would also like to acknowledge the use of the following softwares: SCIKIT-LEARN [88], KERAS [89], MATPLOTLIB [90], NUMPY [91], SCIPY [92], and IPYTHON [93].

-
- [1] A. Burrows and D. Vartanyan, Core-collapse supernova explosion theory, *Nature (London)* **589**, 29 (2021).
 - [2] H.-T. Janka, Explosion mechanisms of core-collapse supernovae, *Annu. Rev. Nucl. Part. Sci.* **62**, 407 (2012).
 - [3] F. Foucart, Neutrino transport in general relativistic neutron star merger simulations, *Living Rev. Comput. Astrophys.* **9**, 1 (2023).
 - [4] K. Kyutoku, M. Shibata, and K. Taniguchi, Coalescence of black hole–neutron star binaries, *Living Rev. Relativity* **24**, 5 (2021).
 - [5] S. A. Colgate and R. H. White, The hydrodynamic behavior of supernovae explosions, *Astrophys. J.* **143**, 626 (1966).
 - [6] J. M. Lattimer and D. N. Schramm, Black-hole-neutron-star collisions, *Astrophys. J. Lett.* **192**, L145 (1974).
 - [7] J. T. Pantaleone, Neutrino oscillations at high densities, *Phys. Lett. B* **287**, 128 (1992).
 - [8] G. Sigl and G. Raffelt, General kinetic description of relativistic mixed neutrinos, *Nucl. Phys.* **B406**, 423 (1993).
 - [9] S. Pastor and G. Raffelt, Flavor oscillations in the supernova hot bubble region: Nonlinear effects of neutrino background, *Phys. Rev. Lett.* **89**, 191101 (2002).

- [10] H. Duan, G. M. Fuller, J. Carlson, and Y.-Z. Qian, Simulation of coherent non-linear neutrino flavor transformation in the supernova environment. 1. Correlated neutrino trajectories, *Phys. Rev. D* **74**, 105014 (2006).
- [11] H. Duan, G. M. Fuller, J. Carlson, and Y.-Z. Qian, Coherent development of neutrino flavor in the supernova environment, *Phys. Rev. Lett.* **97**, 241101 (2006).
- [12] H. Duan, G. M. Fuller, and Y.-Z. Qian, Collective neutrino oscillations, *Annu. Rev. Nucl. Part. Sci.* **60**, 569 (2010).
- [13] A. Mirizzi, I. Tamborra, H.-T. Janka, N. Saviano, K. Scholberg, R. Bollig, L. Hüdepohl, and S. Chakraborty, Supernova neutrinos: Production, oscillations and detection, *Riv. Nuovo Cimento* **39**, 1 (2016).
- [14] M. C. Volpe, Neutrinos from dense: Flavor mechanisms, theoretical approaches, observations, new directions, [arXiv: 2301.11814](https://arxiv.org/abs/2301.11814).
- [15] R. F. Sawyer, Speed-up of neutrino transformations in a supernova environment, *Phys. Rev. D* **72**, 045003 (2005).
- [16] R. F. Sawyer, Neutrino cloud instabilities just above the neutrino sphere of a supernova, *Phys. Rev. Lett.* **116**, 081101 (2016).
- [17] S. Chakraborty, R. S. Hansen, I. Izaguirre, and G. Raffelt, Self-induced neutrino flavor conversion without flavor mixing, *J. Cosmol. Astropart. Phys.* **03** (2016) 042.
- [18] I. Izaguirre, G. Raffelt, and I. Tamborra, Fast pairwise conversion of supernova neutrinos: A dispersion-relation approach, *Phys. Rev. Lett.* **118**, 021101 (2017).
- [19] F. Capozzi, B. Dasgupta, E. Lisi, A. Marrone, and A. Mirizzi, Fast flavor conversions of supernova neutrinos: Classifying instabilities via dispersion relations, *Phys. Rev. D* **96**, 043016 (2017).
- [20] M.-R. Wu and I. Tamborra, Fast neutrino conversions: Ubiquitous in compact binary merger remnants, *Phys. Rev. D* **95**, 103007 (2017).
- [21] M.-R. Wu, I. Tamborra, O. Just, and H.-T. Janka, Imprints of neutrino-pair flavor conversions on nucleosynthesis in ejecta from neutron-star merger remnants, *Phys. Rev. D* **96**, 123015 (2017).
- [22] S. Abbar and H. Duan, Fast neutrino flavor conversion: Roles of dense matter and spectrum crossing, *Phys. Rev. D* **98**, 043014 (2018).
- [23] S. Abbar and M. C. Volpe, On fast neutrino flavor conversion modes in the nonlinear regime, *Phys. Lett. B* **790**, 545 (2019).
- [24] F. Capozzi, B. Dasgupta, A. Mirizzi, M. Sen, and G. Sigl, Collisional triggering of fast flavor conversions of supernova neutrinos, *Phys. Rev. Lett.* **122**, 091101 (2019).
- [25] J. D. Martin, C. Yi, and H. Duan, Dynamic fast flavor oscillation waves in dense neutrino gases, *Phys. Lett. B* **800**, 135088 (2020).
- [26] S. Abbar, H. Duan, K. Sumiyoshi, T. Takiwaki, and M. C. Volpe, On the occurrence of fast neutrino flavor conversions in multidimensional supernova models, *Phys. Rev. D* **100**, 043004 (2019).
- [27] S. Abbar, H. Duan, K. Sumiyoshi, T. Takiwaki, and M. C. Volpe, Fast neutrino flavor conversion modes in multidimensional core-collapse supernova models: The role of the asymmetric neutrino distributions, *Phys. Rev. D* **101**, 043016 (2020).
- [28] F. Capozzi, G. Raffelt, and T. Stirner, Fast neutrino flavor conversion: Collective motion vs. decoherence, *J. Cosmol. Astropart. Phys.* **09** (2019) 002.
- [29] M. George, M.-R. Wu, I. Tamborra, R. Ardevol-Pulpillo, and H.-T. Janka, Fast neutrino flavor conversion, ejecta properties, and nucleosynthesis in newly-formed hypermassive remnants of neutron-star mergers, *Phys. Rev. D* **102**, 103015 (2020).
- [30] L. Johns, H. Nagakura, G. M. Fuller, and A. Burrows, Neutrino oscillations in supernovae: Angular moments and fast instabilities, *Phys. Rev. D* **101**, 043009 (2020).
- [31] J. D. Martin, J. Carlson, V. Cirigliano, and H. Duan, Fast flavor oscillations in dense neutrino media with collisions, *Phys. Rev. D* **103**, 063001 (2021).
- [32] I. Tamborra and S. Shalgar, New developments in flavor evolution of a dense neutrino gas, *Annu. Rev. Nucl. Part. Sci.* **71**, 165 (2021).
- [33] G. Sigl, Simulations of fast neutrino flavor conversions with interactions in inhomogeneous media, *Phys. Rev. D* **105**, 043005 (2022).
- [34] T. Morinaga, Fast neutrino flavor instability and neutrino flavor lepton number crossings, *Phys. Rev. D* **105**, L101301 (2022).
- [35] H. Nagakura, L. Johns, A. Burrows, and G. M. Fuller, Where, when, and why: Occurrence of fast-pairwise collective neutrino oscillation in three-dimensional core-collapse supernova models, *Phys. Rev. D* **104**, 083025 (2021).
- [36] H. Sasaki and T. Takiwaki, A detailed analysis of the dynamics of fast neutrino flavor conversions with scattering effects, *Prog. Theor. Exp. Phys.* **2022**, 073E01 (2022).
- [37] I. Padilla-Gay, I. Tamborra, and G. G. Raffelt, Neutrino flavor pendulum reloaded: The case of fast pairwise conversion, *Phys. Rev. Lett.* **128**, 121102 (2022).
- [38] Z. Xiong and Y.-Z. Qian, Stationary solutions for fast flavor oscillations of a homogeneous dense neutrino gas, *Phys. Lett. B* **820**, 136550 (2021).
- [39] F. Capozzi, M. Chakraborty, S. Chakraborty, and M. Sen, Fast flavor conversions in supernovae: The rise of mu-tau neutrinos, *Phys. Rev. Lett.* **125**, 251801 (2020).
- [40] S. Abbar, F. Capozzi, R. Glas, H.-T. Janka, and I. Tamborra, On the characteristics of fast neutrino flavor instabilities in three-dimensional core-collapse supernova models, *Phys. Rev. D* **103**, 063033 (2021).
- [41] F. Capozzi, S. Abbar, R. Bollig, and H. T. Janka, Fast neutrino flavor conversions in one-dimensional core-collapse supernova models with and without muon creation, *Phys. Rev. D* **103**, 063013 (2021).
- [42] M. Delfan Azari, S. Yamada, T. Morinaga, W. Iwakami, H. Okawa, H. Nagakura, and K. Sumiyoshi, Linear analysis of fast-pairwise collective neutrino oscillations in core-collapse supernovae based on the results of Boltzmann simulations, *Phys. Rev. D* **99**, 103011 (2019).
- [43] A. Harada and H. Nagakura, Prospects of fast flavor neutrino conversion in rotating core-collapse supernovae, *Astrophys. J.* **924**, 109 (2022).
- [44] S. Abbar and F. Capozzi, Suppression of fast neutrino flavor conversions occurring at large distances in core-collapse supernovae, *J. Cosmol. Astropart. Phys.* **03** (2022) 051.

- [45] O. Just, S. Abbar, M.-R. Wu, I. Tamborra, H.-T. Janka, and F. Capozzi, Fast neutrino conversion in hydrodynamic simulations of neutrino-cooled accretion disks, *Phys. Rev. D* **105**, 083024 (2022).
- [46] I. Padilla-Gay, I. Tamborra, and G. G. Raffelt, Neutrino fast flavor pendulum. II. Collisional damping, *Phys. Rev. D* **106**, 103031 (2022).
- [47] F. Capozzi, M. Chakraborty, S. Chakraborty, and M. Sen, Supernova fast flavor conversions in 1 + 1D: Influence of mu-tau neutrinos, *Phys. Rev. D* **106**, 083011 (2022).
- [48] M. Zaizen and H. Nagakura, Simple method for determining asymptotic states of fast neutrino-flavor conversion, *Phys. Rev. D* **107**, 103022 (2023).
- [49] S. Shalgar and I. Tamborra, Neutrino decoupling is altered by flavor conversion, *Phys. Rev. D* **108**, 043006 (2023).
- [50] C. Kato and H. Nagakura, Effects of energy-dependent scatterings on fast neutrino flavor conversions, *Phys. Rev. D* **106**, 123013 (2022).
- [51] S. Bhattacharyya and B. Dasgupta, Fast flavor depolarization of supernova neutrinos, *Phys. Rev. Lett.* **126**, 061302 (2021).
- [52] M.-R. Wu, M. George, C.-Y. Lin, and Z. Xiong, Collective fast neutrino flavor conversions in a 1D box: Initial conditions and long-term evolution, *Phys. Rev. D* **104**, 103003 (2021).
- [53] S. Richers, D. E. Willcox, N. M. Ford, and A. Myers, Particle-in-cell simulation of the neutrino fast flavor instability, *Phys. Rev. D* **103**, 083013 (2021).
- [54] S. Richers, D. Willcox, and N. Ford, Neutrino fast flavor instability in three dimensions, *Phys. Rev. D* **104**, 103023 (2021).
- [55] B. Dasgupta, Collective neutrino flavor instability requires a crossing, *Phys. Rev. Lett.* **128**, 081102 (2022).
- [56] H. Nagakura and M. Zaizen, Time-dependent and quasisteady features of fast neutrino-flavor conversion, *Phys. Rev. Lett.* **129**, 261101 (2022).
- [57] J. Ehring, S. Abbar, H.-T. Janka, G. Raffelt, and I. Tamborra, Fast neutrino flavor conversion in core-collapse supernovae: A parametric study in 1D models, *Phys. Rev. D* **107**, 103034 (2023).
- [58] J. Ehring, S. Abbar, H.-T. Janka, G. Raffelt, and I. Tamborra, Fast neutrino flavor conversions can help and hinder neutrino-driven explosions, *Phys. Rev. Lett.* **131**, 061401 (2023).
- [59] Z. Xiong, M.-R. Wu, and Y.-Z. Qian, Symmetry and bipolar motion in collective neutrino flavor oscillations, *Phys. Rev. D* **108**, 043007 (2023).
- [60] D. F. G. Fiorillo and G. G. Raffelt, Flavor solitons in dense neutrino gases, *Phys. Rev. D* **107**, 123024 (2023).
- [61] H. Nagakura, Global features of fast neutrino-flavor conversion in binary neutron star mergers, *Phys. Rev. D* **108**, 103014 (2023).
- [62] J. D. Martin, D. Neill, A. Roggero, H. Duan, and J. Carlson, Equilibration of quantum many-body fast neutrino flavor oscillations, *Phys. Rev. D* **108**, 123010 (2023).
- [63] D. F. G. Fiorillo and G. G. Raffelt, Slow and fast collective neutrino oscillations: Invariants and reciprocity, *Phys. Rev. D* **107**, 043024 (2023).
- [64] E. Grohs, S. Richers, S. M. Couch, F. Foucart, J. Froustey, J. Kneller, and G. McLaughlin, Two-moment neutrino flavor transformation with applications to the fast flavor instability in neutron star mergers, [arXiv:2309.00972](https://arxiv.org/abs/2309.00972).
- [65] Z. Xiong, A. Sieverding, M. Sen, and Y.-Z. Qian, Potential impact of fast flavor oscillations on neutrino-driven winds and their nucleosynthesis, *Astrophys. J.* **900**, 144 (2020).
- [66] X. Li and D. M. Siegel, Neutrino fast flavor conversions in neutron-star postmerger accretion disks, *Phys. Rev. Lett.* **126**, 251101 (2021).
- [67] R. Fernández, S. Richers, N. Mulyk, and S. Fahlman, Fast flavor instability in hypermassive neutron star disk outflows, *Phys. Rev. D* **106**, 103003 (2022).
- [68] S. Bhattacharyya and B. Dasgupta, Late-time behavior of fast neutrino oscillations, *Phys. Rev. D* **102**, 063018 (2020).
- [69] M. Zaizen and T. Morinaga, Nonlinear evolution of fast neutrino flavor conversion in the preshock region of core-collapse supernovae, *Phys. Rev. D* **104**, 083035 (2021).
- [70] S. Bhattacharyya and B. Dasgupta, Elaborating the ultimate fate of fast collective neutrino flavor oscillations, *Phys. Rev. D* **106**, 103039 (2022).
- [71] E. Grohs, S. Richers, S. M. Couch, F. Foucart, J. P. Kneller, and G. C. McLaughlin, Neutrino fast flavor instability in three dimensions for a neutron star merger, *Phys. Lett. B* **846**, 138210 (2023).
- [72] S. Richers, H. Duan, M.-R. Wu, S. Bhattacharyya, M. Zaizen, M. George, C.-Y. Lin, and Z. Xiong, Code comparison for fast flavor instability simulations, *Phys. Rev. D* **106**, 043011 (2022).
- [73] Z. Xiong, M.-R. Wu, S. Abbar, S. Bhattacharyya, M. George, and C.-Y. Lin, Evaluating approximate asymptotic distributions for fast neutrino flavor conversions in a periodic 1D box, *Phys. Rev. D* **108**, 063003 (2023).
- [74] M. Zaizen and H. Nagakura, Characterizing quasisteady states of fast neutrino-flavor conversion by stability and conservation laws, *Phys. Rev. D* **107**, 123021 (2023).
- [75] M. Shibata, K. Kiuchi, Y.-i. Sekiguchi, and Y. Suwa, Truncated moment formalism for radiation hydrodynamics in numerical relativity, *Prog. Theor. Phys.* **125**, 1255 (2011).
- [76] C. Y. Cardall, E. Endeve, and A. Mezzacappa, Conservative 3 + 1 general relativistic variable Eddington tensor radiation transport equations, *Phys. Rev. D* **87**, 103004 (2013).
- [77] K. S. Thorne, Relativistic radiative transfer: Moment formalisms, *Mon. Not. R. Astron. Soc.* **194**, 439 (1981).
- [78] D. Guest, K. Cranmer, and D. Whiteson, Deep learning and its application to LHC physics, *Annu. Rev. Nucl. Part. Sci.* **68**, 161 (2018).
- [79] M. J. Smith and J. E. Geach, *Astronomia ex machina: A history, primer and outlook on neural networks in astronomy*, *R. Soc. Open Sci.* **10**, 221454 (2023).
- [80] G. E. Karniadakis, I. G. Kevrekidis, L. Lu, P. Perdikaris, S. Wang, and L. Yang, Physics-informed machine learning, *Nat. Rev. Phys.* **3**, 422 (2021).
- [81] S. Cuomo, V. S. Di Cola, F. Giampaolo, G. Rozza, M. Raissi, and F. Piccialli, Scientific machine learning through physics-informed neural networks: Where we are and what's next, *J. Sci. Comput.* **92**, 88 (2022).
- [82] M. Raissi, P. Perdikaris, and G. E. Karniadakis, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.* **378**, 686 (2019).

- [83] J. Cernohorsky and S. A. Bludman, Maximum entropy distribution and closure for Bose-Einstein and Fermi-Dirac radiation transport (1994).
- [84] S. Richers, Evaluating approximate flavor instability metrics in neutron star mergers, *Phys. Rev. D* **106**, 083005 (2022).
- [85] S. Abbar, Applications of machine learning to detecting fast neutrino flavor instabilities in core-collapse supernova and neutron star merger models, *Phys. Rev. D* **107**, 103006 (2023).
- [86] S. Abbar and H. Nagakura, Detecting fast neutrino flavor conversions with machine learning, [arXiv:2310.03807](https://arxiv.org/abs/2310.03807).
- [87] C. Yi, L. Ma, J. D. Martin, and H. Duan, Dispersion relation of the fast neutrino oscillation wave, *Phys. Rev. D* **99**, 063005 (2019).
- [88] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, Scikit-learn: Machine learning in PYTHON, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- [89] F. Chollet *et al.*, Keras, <https://keras.io> (2015).
- [90] J. D. Hunter, Matplotlib: A 2d graphics environment, *Comput. Sci. Eng.* **9**, 90 (2007).
- [91] S. van der Walt, S. C. Colbert, and G. Varoquaux, The NUMPY array: A structure for efficient numerical computation, *Comput. Sci. Eng.* **13**, 22 (2011).
- [92] P. Virtanen *et al.*, SciPy 1.0: Fundamental algorithms for scientific computing in PYTHON, *Nat. Methods* **17**, 261 (2020).
- [93] F. Pérez and B. E. Granger, IPYTHON: A system for interactive scientific computing, *Comput. Sci. Eng.* **9**, 21 (2007).