

# Machine learning-based inversion of nuclear responses

Krishnan Raghavan<sup>a</sup>, Prasanna Balaprakash<sup>a</sup>, Alessandro Lovato<sup>b,c,d</sup>, Noemi Rocco<sup>b,e</sup>, Stefan M. Wild<sup>a,f</sup>

<sup>a</sup>*Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, Illinois 60439, USA*

<sup>b</sup>*Physics Division, Argonne National Laboratory, Lemont, Illinois 60439, USA*

<sup>c</sup>*Computational Science Division, Argonne National Laboratory, Lemont, Illinois 60439, USA*

<sup>d</sup>*INFN-TIFPA Trento Institute of Fundamental Physics and Applications, Via Sommarive, 14, 38123 Trento, Italy*

<sup>e</sup>*Theoretical Physics Department, Fermi National Accelerator Laboratory, P.O. Box 500, Batavia, Illinois 60510, USA*

<sup>f</sup>*NAISE, Northwestern University, Evanston, Illinois 60208, USA*

(Dated: October 20, 2020)

A microscopic description of the interaction of atomic nuclei with external electroweak probes is required for elucidating aspects of short-range nuclear dynamics and for the correct interpretation of neutrino oscillation experiments. Nuclear quantum Monte Carlo methods infer the nuclear electroweak response functions from their Laplace transforms. Inverting the Laplace transform is a notoriously ill-posed problem; and Bayesian techniques, such as maximum entropy, are typically used to reconstruct the original response functions in the quasielastic region. In this work, we present a physics-informed artificial neural network architecture suitable for approximating the inverse of the Laplace transform. Utilizing simulated, albeit realistic, electromagnetic response functions, we show that this physics-informed artificial neural network outperforms maximum entropy in both the low-energy transfer and the quasielastic regions, thereby allowing for robust calculations of electron scattering and neutrino scattering on nuclei and inclusive muon capture rates.

## I. INTRODUCTION

Electron scattering experiments are powerful tools to simultaneously investigate the short- and long-range many-body dynamics of atomic nuclei. These experiments contributed to demonstrating the limitations of an independent particle picture of the nucleus that fails to provide a fully quantitative description of atomic nuclei [1]. At large momentum transfer, the large excess of neutron-proton correlated pairs with respect to the proton-proton and neutron-neutron pairs has highlighted the importance of the tensor component of the nuclear interaction and the interplay between nucleonic and partonic degrees of freedom [2–4]. The field has experienced a renewed interest also in view of its interplay with high-precision measurements of neutrinos and their oscillations [5]. This is the main focus of the accelerator-based neutrino oscillation program, which includes ongoing experiments such as NOvA [6] and T2K [7] and planned ones such as DUNE [8] and Hyper-K [9]. Nuclear targets are utilized in the detectors to increase the event rate. Hence, the determination of oscillation parameters requires accurate theoretical calculations of neutrino-nucleus interactions in a broad range of energy, in which a variety of reaction mechanisms are at play [10–12].

In the low-energy regime, the inclusive lepton-nucleus cross section is dominated by coherent scattering, excitations of low-lying nuclear states, and collective modes. At energies on the order of hundreds of MeV, the leading mechanism is quasielastic (QE) scattering, in which the probe interacts primarily with individual nucleons bound inside the nucleus. Corrections to this leading mechanism arise from processes in which the lepton couples to interacting nucleons, via nuclear correlations and two-body currents.

The inclusive lepton-nucleus scattering cross section

is completely determined by the electroweak response functions, which hold all information about the dynamics of the nuclear target. The Green’s function Monte Carlo (GFMC) method [13] has been successfully employed to compute the electromagnetic, neutral-current, and charged-current response functions of  $^4\text{He}$  and  $^{12}\text{C}$  in the QE region, up to moderate values of the momentum transfer [14–17] and the muon capture rates of  $^4\text{He}$  and  $^3\text{H}$  [18]. These calculations have unambiguously demonstrated the importance of properly treating nuclear correlations and meson exchange currents even for QE kinematics. Within this approach, the electroweak response functions are inferred from their Laplace transforms, denoted as Euclidean responses, that are estimated during the GFMC imaginary time propagation. Retrieving the energy dependence of the response functions from their Euclidean counterparts is nontrivial.

The maximum entropy method (MaxEnt) [19, 20] has been extensively employed to retrieve the energy dependence of the electroweak response functions. Despite its success in the QE region, MaxEnt appears to be inadequate to precisely reconstruct the low-energy structure of the nuclear response functions. In Ref. [15], experimental inputs on the low-lying nuclear transitions have been utilized to properly describe the longitudinal electromagnetic responses of  $^{12}\text{C}$  in the low-energy region. A comparison between GFMC and exact Faddeev results for the  $^3\text{H}$  muon capture rate has contributed to exposing the shortcomings of MaxEnt in reconstructing the charged-current response functions near the nuclear breakup threshold, corresponding to energies of few MeVs [18]. In addition, although heuristics have been used, to the best of our knowledge there is no rigorous way to propagate the statistical uncertainties of the Euclidean response into the response function and to quantify the systematic errors due to the approximate inver-

sion of the Laplace transform. These errors would propagate into the GFMC estimates of lepton-nucleus cross sections and are critical for informative comparisons with experiments.

In recent years, an increase in available computing resources has been accompanied by a prodigious rise of techniques based on machine learning (ML), which are now ubiquitous in physics [21]. Within low-energy nuclear physics, artificial neural networks (ANNs) have been used to estimate ground state energies and radii of nuclei by using results from no-core shell model calculations [22, 23]. Gaussian process emulators were used in Ref. [24] for Bayesian model mixing in order to predict bound nuclides between silicon and titanium. The authors of Ref. [25] represent the deuteron's wave function with ANNs. In Ref. [26] ANNs were used to model the Jastrow correlator of  $A \leq 4$  nuclei. Several works have demonstrated that ML approaches are suitable for solving inverse problems [27, 28]. In particular, Refs. [29, 30] utilized ANNs to recover the electron single-particle spectral density in the real frequency domain from the fermionic Green's function in the imaginary time domain. The same problem was tackled in Ref. [31] by utilizing an Adams-Bashforth residual ANN. In both cases, the ANN approaches have been found to outperform MaxEnt implementations.

In this work we develop a novel ANN architecture suitable for approximately inverting the Laplace transform of realistic nuclear electromagnetic response functions, similar to those computed with the GFMC method. The simulated responses utilized in the training dataset exhibit a sharp Gaussian peak corresponding to the low-energy elastic transition and an asymmetric broad peak in the QE region. The positions, heights, and widths of these two peaks are modeled consistently with their energy and momentum transfer behavior as measured by electron-scattering experiments. In contrast to previous approaches, we incorporate physics-grounded constraints into the neural-network architecture and use an entropic cost function. We demonstrate an improved accuracy of the inversion in the low  $\omega$  region with increased robustness to noise as compared with MaxEnt techniques.

This work is organized as follows. In Sec. II we state the problem to be solved and discuss the relevant features of the nuclear electromagnetic responses. In Sec. III we describe our ML algorithm. In Sec. IV we present our results, and in Sec. V we discuss our conclusions.

## II. NUCLEAR RESPONSES

In the one-photon exchange approximation, the inclusive electron-nucleus scattering cross section can be expressed in terms of the longitudinal and transverse response functions,  $R_L(\mathbf{q}, \omega)$  and  $R_T(\mathbf{q}, \omega)$ , respectively, where  $\mathbf{q}$  and  $\omega$  are the electron momentum and energy transfers. The response functions encode all information

on nuclear structure and dynamics and are defined as

$$R_\alpha(\mathbf{q}, \omega) = \sum_f \langle 0 | j_\alpha^\dagger(\mathbf{q}, \omega) | f \rangle \langle f | j_\alpha(\mathbf{q}, \omega) | 0 \rangle \times \delta(E_f - \omega - E_0), \quad (1)$$

for  $\alpha = L, T$ . In Eq. (1),  $|0\rangle$  and  $|f\rangle$  represent the initial and final nuclear states of energies  $E_0$  and  $E_f$ , respectively, and  $j_L(\mathbf{q}, \omega)$  and  $j_T(\mathbf{q}, \omega)$  are the electromagnetic charge and current operators, respectively.

A direct calculation of  $R_\alpha(\mathbf{q}, \omega)$  requires evaluating all of the individual transition amplitudes induced by the charge and current operators and is therefore impractical except for very light nuclear systems [32, 33]. The use of integral transform techniques has proven helpful in circumventing these difficulties. One such approach is based on the calculation of the Euclidean response [34], which corresponds to the Laplace transform

$$E_\alpha(\mathbf{q}, \tau) = \int_0^\infty d\omega e^{-\omega\tau} R_\alpha(\mathbf{q}, \omega). \quad (2)$$

Fixing the intrinsic energy dependence of the charge and current operators to the QE peak,  $\omega_{QE} = \sqrt{\mathbf{q}^2 + m^2} - m$ , where  $m$  denotes the mass of the nucleon, one can express the Euclidean responses as ground-state expectation values

$$E_\alpha(\mathbf{q}, \tau) = \langle 0 | j_\alpha^\dagger(\mathbf{q}, \omega_{QE}) e^{-(H-E_0)\tau} j_\alpha(\mathbf{q}, \omega_{QE}) | 0 \rangle,$$

where  $H$  is the nuclear Hamiltonian. These expectation values can be evaluated by using the GFMC method on a uniform grid of  $n_\tau$  imaginary-time points [14, 34]. A set of noisy estimates for  $E_\alpha(\mathbf{q}, \tau_i)$  can be obtained by performing independent imaginary-time propagations, from which the average Euclidean response  $\bar{E}_\alpha(\mathbf{q}, \tau_i)$  and the covariance  $C_{ij}$  between the data at  $\tau = \tau_i$  and  $\tau = \tau_j$  can be readily estimated [15]. Note that, in general, the covariance matrix  $C$  is nondiagonal because of correlations among the imaginary-time points.

### Problem statement and the MaxEnt approach

In addition to the imaginary time  $\mathcal{T} = [\tau_1, \dots, \tau_{n_\tau}]$ , we discretize the continuous variables  $\omega$  on  $n_\omega$  grid points and thus define  $\Omega = [\omega_1, \dots, \omega_{n_\omega}]$  and the kernel  $K(\omega_i, \tau_j) = e^{-\omega_i \tau_j} \Delta\omega_i$ , where  $\Delta\omega_i$  is the discretization width at  $\omega_i$ . Dropping, for simplicity, the momentum transfer dependence and the subscript  $\alpha$  of the response functions, we can rewrite the Laplace transform of Eq. (2) in the compact matrix form

$$E(\mathcal{T}) = K(\Omega, \mathcal{T}) R(\Omega),$$

where  $E(\mathcal{T}) \in \mathbb{R}^{n_\tau}$ ,  $R(\Omega) \in \mathbb{R}^{n_\omega}$ , and  $K(\Omega, \mathcal{T}) \in \mathbb{R}^{n_\tau \times n_\omega}$ . The response function can thus be formally (for an appropriate definition of  $\cdot^{-1}$ ) recovered by

$$R(\Omega) = K(\Omega, \mathcal{T})^{-1} E(\mathcal{T}). \quad (3)$$

However, the inversion of  $K(\Omega, \mathcal{T})$  is numerically unstable because of the exponentially small tails in the kernel function for large  $\omega$ . Retrieving the response function from noisy GFMC estimates of  $E(\tau)$  involves significant difficulty; widely different response functions can correspond to very similar Euclidean responses.

Several algorithms have been developed for approximately inverting the Laplace transform by using prior knowledge about the solution. Arguably the most robust and popular of these is MaxEnt [19, 20], which has been used to reconstruct the (smooth) energy dependence of the nuclear response functions around the QE peak [15–17]. Within MaxEnt, the solution of the inverse problem is the response function that maximizes the posterior probability  $P(R|E)$  (i.e., the conditional probability of  $R(\Omega)$  given  $E(\mathcal{T})$ ). Bayes’ theorem states that the posterior probability is proportional to the product  $P(E|R) \times P(R)$ , where  $P(E|R)$  is the likelihood function and  $P(R)$  is the prior probability, containing information about the response function to be reconstructed. Arguments based on the central limit theorem show that the asymptotic limit of the likelihood function is given by  $P(E|R) \propto \exp(-\chi^2/2)$ , where

$$\chi^2 = \sum_{i,j=1}^{n_\tau} (E(\tau_i) - \bar{E}(\tau_i)) C_{ij}^{-1} (E(\tau_j) - \bar{E}(\tau_j)). \quad (4)$$

Since the response functions are positive and normalizable, they can be interpreted as probability distributions. The principle of maximum entropy states that the values of a probability distribution are to be assigned by maximizing the entropy, which is defined by

$$S = \sum_{i=1}^{n_\omega} \left( R(\omega_i) - M(\omega_i) - R(\omega_i) \ln \left( \frac{R(\omega_i)}{M(\omega_i)} \right) \right) \Delta\omega_i. \quad (5)$$

The positive-valued  $M(\omega)$  is the default model and encodes our prior knowledge about  $R(\omega)$  in the absence of data. The entropy measures how much the response function differs from the model. It vanishes when  $R(\omega) = M(\omega)$  and is negative when  $R(\omega) \neq M(\omega)$ .

MaxEnt improves upon the standard  $\chi^2$  minimization by using the prior information, whereby  $R(\omega)$  can be interpreted as a probability distribution. For given  $\bar{E}(\tau_i)$ ,  $C_{ij}$ , and default model  $M(\omega_i)$ , the response functions are found minimizing the quantity

$$Q = \frac{1}{2} \chi^2 - \alpha S, \quad (6)$$

where  $\alpha$  is a fixed parameter that controls the relative importance between the entropy and the error terms. Despite its tendency to underfit the data [35], in this work we adopt the *historic* MaxEnt approach [36], which consists in choosing  $\alpha$  so that  $\chi^2 = n_\tau$ . On the other hand, the more sophisticated *classic* MaxEnt [37] and *Bryan* MaxEnt [19]—both relying on the probability  $P(\alpha|E)$  to determine  $\alpha$ —tend to overfit the data since  $P(\alpha|E)$

is evaluated only approximately in practice [38, 39]. In general, the arbitrariness in choosing  $\alpha$  prevents a robust reconstruction of the rich structure that characterizes the low- $\omega$  region of  $R(\omega)$ , without running the risk of overfitting  $E(\tau)$  and hence causing spurious oscillation in the reconstructed response function.

### III. PHYSICS-INFORMED NEURAL NETWORK

As mentioned in the preceding section, the inversion of  $K(\Omega, \mathcal{T})$  is numerically unstable, and retrieving  $R(\Omega)$  from  $E(\mathcal{T})$  is an ill-posed inverse problem. To overcome this difficulty, we seek an approximate solution by designing a physics-informed neural network, which we dub “Phys-NN,” that is suitable for finding a controlled approximation  $\hat{R}(\Omega)$  for the right-hand side of Eq. (3).

#### A. The Phys-NN model

To model  $\hat{R}(\Omega)$ , we start by constructing a set  $\mathcal{R}$  of basis functions that takes into account the physics of the problem, while being as broadly applicable as possible. Note that each term in the matrix  $K(\Omega, \mathcal{T})$ , is proportional to  $e^{-\tau_j \omega_i}$  and therefore a reasonable choice to capture its structure is the Gaussian kernel basis, defined as

$$\phi(x, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}. \quad (7)$$

Finding the location and the scale of the Gaussian kernel, denoted by  $\mu \in \mathbb{R}$  and  $\sigma > 0$ , respectively, is part of the ML training problem. The first layer of the neural network, whose architecture is displayed in Fig. 1, takes as input the  $n_\tau$ -dimensional vector  $E(\mathcal{T})$ . To form a basis for each  $E(\tau)$ , we apply  $n_\eta$  Gaussian units of the form (7), where  $n_\eta$  is a hyperparameter. We then contract these Gaussian units with the resulting hidden layer outputs multiplied by weights  $w_{i,j}$  to obtain the output associated with  $\omega_i$ . Formally, the Phys-NN is given by

$$f(E(\mathcal{T}); \boldsymbol{\theta}) = \begin{bmatrix} \sum_{j=1}^{n_\tau} W_{1,j} \sum_{k=1}^{n_\eta} \phi(E(\tau_j), \mu_{j,k}, \sigma_{j,k}) \\ \vdots \\ \sum_{j=1}^{n_\tau} W_{n_\omega,j} \sum_{k=1}^{n_\eta} \phi(E(\tau_j), \mu_{j,k}, \sigma_{j,k}) \end{bmatrix}, \quad (8)$$

where we use  $\boldsymbol{\theta} = (\boldsymbol{\mu}, \boldsymbol{\sigma}, \mathbf{W})$  to denote the collection of training parameters  $\boldsymbol{\mu}, \boldsymbol{\sigma} \in \mathbb{R}^{n_\tau \times n_\eta}$  and  $\mathbf{W} \in \mathbb{R}^{n_\omega \times n_\tau}$ . We can express Eq. (8) componentwise by

$$f(\omega_i) = \sum_{j=1}^{n_\tau} W_{i,j} \sum_{k=1}^{n_\eta} \phi(E(\tau_j), \mu_{j,k}, \sigma_{j,k}), \quad i = 1, \dots, n_\omega.$$

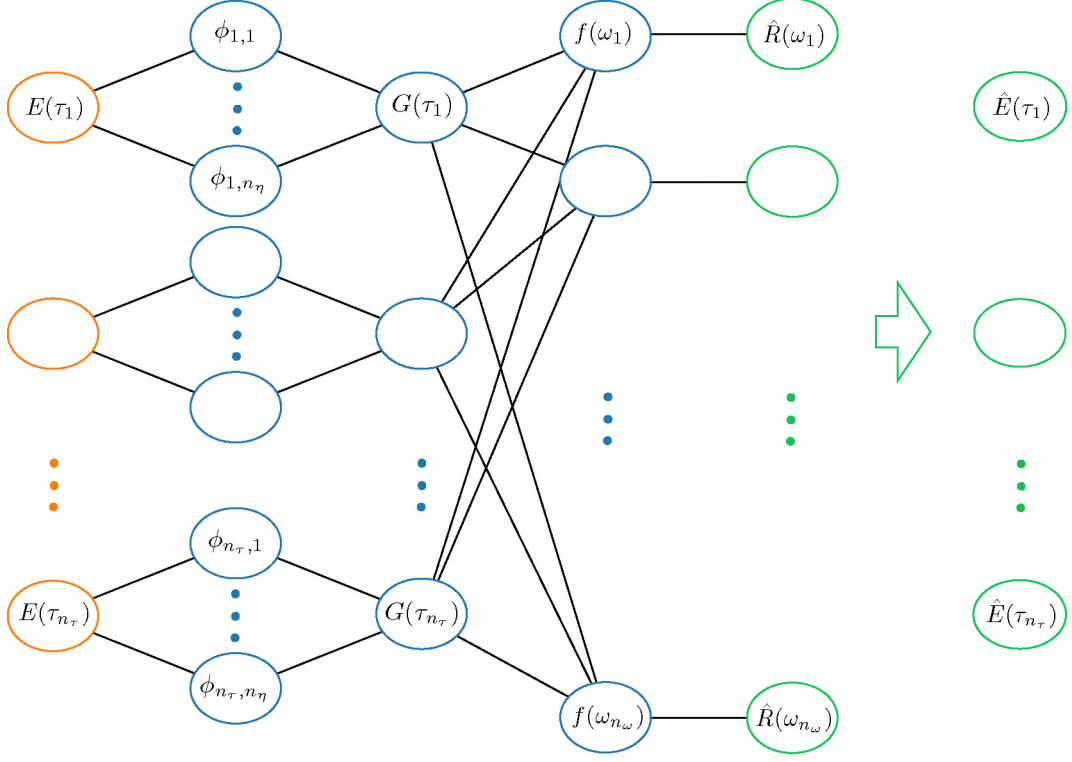


FIG. 1: Schematic overview of the Phys-NN approach.

In order to ensure that the response function is positive for all  $\omega \in \Omega$ , the output is passed through an exponential function, and the final approximation of the response functions is given by

$$\hat{R}(\Omega; \theta) = \frac{1}{\mathcal{N}_0} e^{f(E(\tau); \theta)}.$$

The normalization factor  $\mathcal{N}_0$  ensures that the integral of  $\hat{R}(\Omega; \theta)$  coincides with  $E(\tau_0)$ , so that the output of the Phys-NN automatically satisfies the sum rule of the response function.

### B. Simulated data

To train the Phys-NN, we use two distinct datasets of physically meaningful  $R(\omega)$ ,  $E(\tau)$  pairs that are simulated as follows. The responses belonging to the first dataset—a few of which are displayed in Fig. 2—are characterized by a single asymmetric peak, corresponding to the QE reaction mechanism, modeled by a skew-normal distribution

$$R_{\text{QE}}(\omega) = N_{\text{QE}} \phi(\omega, \omega_{\text{QE}}, \sigma_{\text{QE}}) \Phi\left(\frac{\alpha(\omega - \omega_{\text{QE}})}{\sigma_{\text{QE}}}\right),$$

where  $\phi(\omega, \omega_{\text{QE}}, \sigma_{\text{QE}})$  is the Gaussian density defined in Eq. (7) and

$$\Phi(x) = \frac{1}{2} \left( 1 + \text{erf} \left( \frac{x}{\sqrt{2}} \right) \right)$$

is the Gaussian's cumulative distribution function. The values of  $N_{\text{QE}}$ ,  $\sigma_{\text{QE}}$ , and  $\alpha$  are obtained according to arguments based on the scaling of the response functions [40].

First, we sample the variable  $q$ , corresponding to the momentum transfer, from a uniform distribution between 100 and 700 MeV. Consistent with non-relativistic calculations of the electromagnetic response functions, we assume that  $\omega_{\text{QE}} = q^2/(2m_N) + \epsilon$ , where  $m_N$  is the nucleon mass and  $\epsilon = 25$  MeV is the nuclear binding. A suitable definition for the QE region corresponds to the interval  $\psi = [-1, 1]$  for the scaling variable. Hence, in the non-relativistic case, the width of the QE peak is approximately  $2q_F/m_N$ , and we take the Fermi momentum

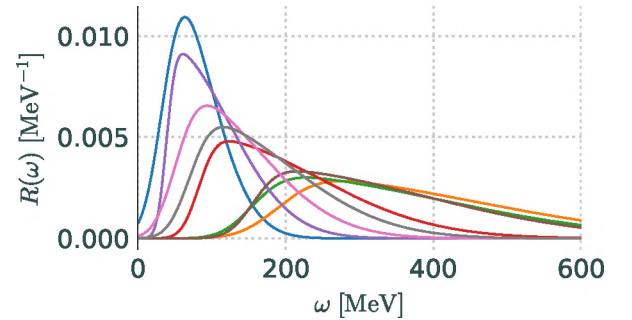


FIG. 2: Training data examples of response functions exhibiting a single asymmetric QE peak.

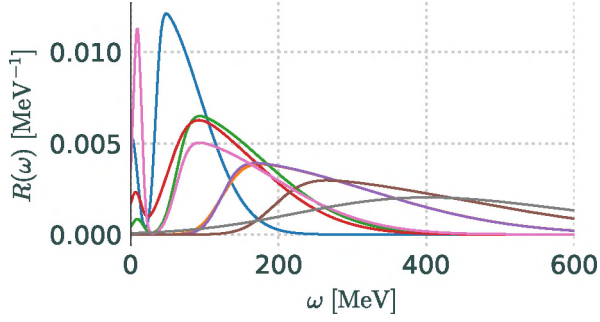


FIG. 3: Training data examples of response functions characterized by an EL narrow peak in addition to the QE peak.

to be  $k_F = 225$  MeV [41]. In the simulated responses, we encode this behavior by allowing 20% fluctuations of  $\sigma_{QE}$  around its central value  $2qk_F/m_N$ . The height of the quasielastic peak is  $N_{QE}/\sigma_{QE}$  and  $N_{QE}$  guarantees that  $R_{QE}(\omega)$  is normalized to unity. The skewness parameter  $\alpha$  is randomly sampled between 2 and 10—the normal distribution is recovered for  $\alpha = 0$ . This interval has been chosen to reproduce the typical asymmetry displayed by the electromagnetic responses of light nuclei.

As shown in Fig. 3, the responses belonging to the second dataset exhibit two distinct peaks, corresponding to the elastic (EL) and QE transitions, namely,  $R_{EL}(\omega) + R_{QE}(\omega)$ . The elastic transition contributes in the low  $\omega$  region, and it is characterized by a  $\delta$ -like peak centered at  $\omega_{EL} = q^2/(2M_A)$ , with  $M_A \approx 4m_N$  being the mass of the  ${}^4\text{He}$  nucleus. We model the EL response with a Gaussian distribution

$$R_{EL}(\omega) = \phi(\omega, \omega_{EL}, \sigma_{EL}),$$

where  $\sigma_{EL}$  is uniformly sampled between 5 and 10 MeV to get a much narrower peak than the QE one. The integrated strength of the EL transition is proportional to the square of the elastic transition form factor  $F_{EL}(q)$ . Inspired by the sum-of-Gaussians parameterizations of  $F_{EL}(q)$  in Ref. [42], we sample  $N_{EL}$  proportional to  $e^{-\frac{\gamma}{2}q^2}$ , where we take  $\gamma = 400$  MeV to reproduce the low-momentum behavior of  $F_{EL}(q)$  for the  ${}^4\text{He}$  nucleus. A direct consequence of this choice is that the strength of the EL peak decreases with the momentum transfer. Consistent with the one-peak case, we enforce the normalization

$$E(\tau_0) = \int d\omega (R_{EL}(\omega) + R_{QE}(\omega)) = 1.$$

The response functions are conveniently tabulated on a uniform  $\omega$  grid between 0 and 2 GeV with  $n_\omega = 2000$ . The corresponding Euclidean responses are obtained by numerically integrating  $R(\omega)$ . Since the simulated response are smooth functions of  $\omega$ , the numerical integration error on the Euclidean responses is smaller than  $10^{-5}$ . To mimic the statistical error of GFMC calcu-

tion, we “corrupt” the simulated  $E(\tau)$  by adding stochastic noise [30]:

$$E(\tau_i) + \epsilon_i, \quad (9)$$

where  $\epsilon_i$  are independent samples from a Gaussian distribution with mean zero and standard deviation  $\sigma$ . Consistent with typical GFMC calculations of the Euclidean electromagnetic responses of  ${}^4\text{He}$ , we take  $\sigma = 10^{-4}$  in most of our tests.

For each of the one-peak and two-peaks cases, we generate a total of 500,000 pairs  $(R_k(\Omega), E_k(\mathcal{T})) \in \mathbb{R}^{n_\omega + n_\tau}$  of responses and corresponding Euclidean responses, which we then partition into training ( $\mathbb{T}$ ), validation ( $\mathbb{V}$ ), and test/out-of-sample ( $\mathbb{O}$ ) datasets. The one-peak and two-peaks test datasets comprise 1,000 pairs each; the combined test dataset is just the union of these two sets. We use 80% and 20% of the remaining data for training the network and validation, respectively. Since MaxEnt is relatively slow—taking about 5 seconds to perform one inversion of the Laplace transform—our comparison is limited to the test dataset.

### C. Training

Values for the parameters  $\theta$  are found by the standard supervised learning approach of approximately solving

$$\min_{\theta} \frac{1}{|\mathbb{T}|} \sum_{k \in \mathbb{T}} \ell(E_k(\mathcal{T}), R_k(\Omega), \hat{R}_k(\Omega; \theta)) \quad (10)$$

by using a minibatch-based stochastic gradient descent procedure to minimize an empirical loss function. Our overall objective in Eq. (10) is the average loss over the  $|\mathbb{T}|$  points in the training set. For each data and model output, we employ a loss function that is the sum of a *response cost* and a *Euclidean cost*,

$$\ell(E_k, R_k, \hat{R}_k) = \gamma_R S_R(R_k, \hat{R}_k) + \gamma_E \chi_E^2(E_k, \hat{R}_k),$$

where  $\gamma_R, \gamma_E > 0$  are user-defined parameters. The response cost is defined according to the entropy measure of Eq. (5), namely

$$S_R(R, \hat{R}) = \sum_{i=1}^{n_\omega} \left( R(\omega_i) - \hat{R}(\omega_i) - R(\omega_i) \ln \left( \frac{R(\omega_i)}{\hat{R}(\omega_i)} \right) \right) \Delta\omega_i, \quad (11)$$

and ensures that the reconstructed response functions are close to the original ones. The Euclidean cost, which is aimed at aligning the Laplace transform of  $\hat{R}(\Omega; \theta)$  with the original Euclidean response, is the reduced  $\chi^2$  per degrees of freedom

$$\chi_E^2(E, \hat{R}) = \frac{1}{n_\tau} \sum_{j=1}^{n_\tau} \frac{1}{\sigma_j^2} \left( E(\tau_j) - \hat{E}(\tau_j) \right)^2. \quad (12)$$

Compared with Eq. (4), in Eq. (12) we have assumed a diagonal covariance matrix, with the diagonal elements corresponding to variance of the independent Gaussian distributions of Eq. (9):  $\sigma_j^2 = \sigma^2 = 10^{-8}$  for all  $j$ . This assumption can be easily relaxed when dealing with correlated data. We evaluate  $\hat{E}(\mathcal{T}; \boldsymbol{\theta}) = K(\mathcal{T}, \Omega) \hat{R}(\Omega; \boldsymbol{\theta})$  by using a simple trapezoidal rule

$$\hat{E}(\tau_j) = \sum_{i=1}^{n_\omega} e^{-\omega_i \tau_j} \hat{R}(\omega_i) \Delta \omega_i. \quad (13)$$

As discussed in the following section, the positive values of  $\gamma_R$  and  $\gamma_E$  are chosen to compensate for the fact that  $\chi_E^2(E, \hat{R})$  is typically much larger than the entropy  $S_R(R, \hat{R})$ .

Since the inversion of the Laplace transform is an ill-posed problem, there are many response functions whose Laplace transform are compatible with the original Euclidean responses. Consequently, there are instances in which  $\chi_E^2$  is small even when the reconstructed response is not similar to the original one, leading to potential instabilities in the minimization procedure. To tame this behavior, we split the training into two phases.

In the first phase, we take  $\gamma_R = 10^7$  and  $\gamma_E = 10^{-7}$  and optimize the network using the Adam [43] optimizer with a learning rate of  $10^{-3}$ . Since  $\gamma_R \gg \gamma_E$ , the entropy response cost dominates the loss function and drives the reconstructed response functions close to the original ones. Once the  $S_R$  has reduced significantly, we enter the second phase of the optimization, where we keep  $\gamma_R = 10^7$  but increase the relative importance of the Euclidean cost by taking  $\gamma_E = 1$  so that Phys-NN learns to keep the Laplace transform of the response function close to the original Euclidean response. Reducing the learning rate in the second phase to  $10^{-5}$  is necessary in order to keep the reconstructed response functions close to the optimal ones found in the previous phase.

The neural-network variants are implemented in Python 3.6 by using TensorFlow 2.0 libraries [44]. Training, validation, and testing are performed by using NVIDIA Tesla V100 SXM2 with 32GB HBM2 and NVIDIA Tesla K80 with dual GPUs hosted at Argonne's Joint Laboratory for System Evaluation and Leadership Computing Facility.

#### IV. RESULTS

We consider three realizations of Phys-NN that differ in the datasets used for training, validation, and testing purposes: one-peak data only, two-peak data only, and combined one-peak and two-peak data. We quantify the accuracy of our approach using three metrics averaged over the associated test/out-of-sample dataset  $\mathbb{O}$ . We use the average entropy

$$\overline{S_R} = \frac{1}{|\mathbb{O}|} \sum_{k \in \mathbb{O}} S_R(R_k, \hat{R}_k),$$

with the entropy  $S_R$  defined in Eq. (11), as well as the average reduced  $\chi_E^2$ ,

$$\overline{\chi_E^2} = \frac{1}{|\mathbb{O}|} \sum_{k \in \mathbb{O}} \chi_E^2(E_k, \hat{R}_k),$$

with  $\chi_E^2$  defined in Eq. (12). We also employ a metric  $\overline{R_R^2}$  for the response functions, which is defined as an average over  $|\mathbb{O}|$  terms of the form

$$R_R^2(R_k, \hat{R}_k) = \frac{\sum_{i=1}^{n_\omega} (\hat{R}_k(\omega_i) - R_k(\omega_i))^2}{\sum_{i=1}^{n_\omega} (\hat{R}_k(\omega_i) - \bar{R}_k(\omega))^2}. \quad (14)$$

##### A. Out-of-sample tests

The values for the three testing metrics for the single-peak, two-peak, and combined datasets are listed in Table I. For both Phys-NN and MaxEnt, the one-peak reconstructions are the closest to their original inputs, the two-peak reconstructions are the worst, and the combined dataset reconstructions rest between those of the other two datasets. This behavior is not unexpected, since the response functions characterized by two peaks, with the EL one in the low- $\omega$  region, are notoriously more difficult to reconstruct than those having a single broad QE peak.

For Phys-NN, the one-peak response function metrics  $1 - \overline{R_R^2}$  and  $\overline{S_R}$  are on the order of  $10^{-5}$ . The reduced  $\chi^2$  is also close to one; smaller values indicate potential overfitting [45]. When reconstructing responses belonging to the two-peak dataset, we observe slightly worse, although still satisfactory, performance compared with the one-peak case, as quantified by the larger values of all three metrics; for the combined dataset,  $\overline{\chi_E^2}$  is only slightly larger than 2.

In Table I one can see in what ways Phys-NN outperforms MaxEnt: both the  $1 - \overline{R_R^2}$  and  $\overline{S_R}$  values obtained with MaxEnt are significantly worse, up to two orders of

TABLE I: Phys-NN and MaxEnt testing metrics  $\overline{S_R}$ ,  $1 - \overline{R_R^2}$ , and  $\overline{\chi_E^2}$  for the one-peak, two-peak, and combined datasets. The standard errors on the last digit of  $\overline{\chi_E^2}$  are given in parentheses.

	$1 - \overline{R_R^2}$ $\times 10^{-4}$	$\overline{\chi_E^2}$	$\overline{S_R}$ $\times 10^{-4}$
<b>Phys-NN</b>			
One-peak	0.42	1.171(13)	0.72
Two-peak	9.04	3.220(87)	9.16
Combined	0.61	2.335(14)	3.66
<b>MaxEnt</b>			
One-peak	29.7	1.015 (1)	60.4
Two-peak	84.8	1.016 (1)	107
Combined	57.2	1.015 (1)	83.7



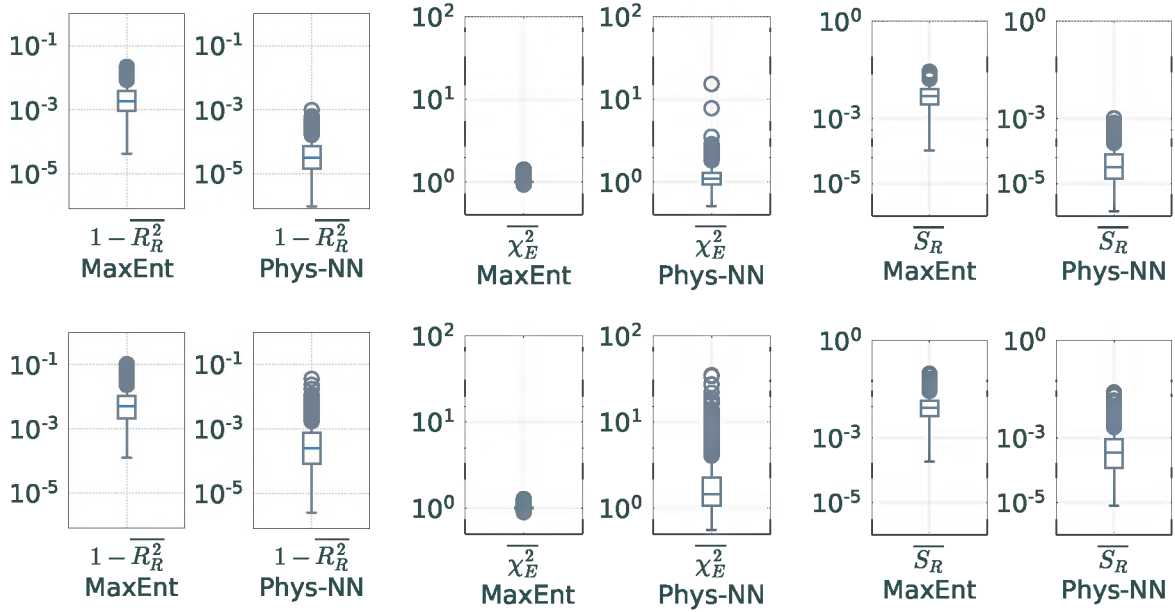


FIG. 4: Box plots of (left)  $R^2$ , (middle)  $\chi_E^2$ , and (right)  $S_R$  for the Phys-NN and MaxEnt methods. The top and bottom rows refer to the one-peak and two-peaks datasets, respectively. The line in the middle of the box denotes the median, and the box represents the range between the 25% and 75% quantiles. Whiskers cover the area between the 1% and 99% quantiles; data beyond these whiskers are outliers and are indicated by circles.

magnitude, than those of Phys-NN. This is a clear indication that Phys-NN captures the energy dependence of the response functions better than does MaxEnt. Since historic MaxEnt finds the optimal response function by fixing  $\alpha$  of Eq. (6) so that  $\chi_E^2 = 1$ , it is not surprising that MaxEnt's reduced  $\chi^2$  values are closer to one than those found by Phys-NN. As evidenced by the other two metrics, because of the ill-posed nature of the problem, achieving  $\chi_E^2 \approx 1$  does not guarantee an accurate reconstruction of the original response functions.

To further examine the performance of Phys-NN and MaxEnt, in Fig. 4 we display box plots of the distributions of the  $1 - \overline{R_R^2}$ ,  $\overline{S_R}$ , and  $\overline{\chi_E^2}$  metrics for the one-peak (top row) and two-peak (bottom row) datasets. Consistent with the results listed in Table I, for both Phys-NN and MaxEnt, the one-peak  $1 - \overline{R_R^2}$  and  $\overline{S_R}$  distributions are narrower and centered on smaller values than are the two-peak ones, while the combined dataset results are intermediate between the two. Since Phys-NN is trained to keep the reconstructed response function as close as possible to the original ones, we observe a much smaller spread of  $1 - \overline{R_R^2}$  and  $\overline{S_R}$  values compared with MaxEnt. This behavior, which is exhibited across the one-peak, two-peak, and combined datasets, provides additional support for Phys-NN's reconstruction performance.

Because the historic MaxEnt algorithm is based on  $\overline{\chi_E^2}$  minimization, the resulting distributions of  $\overline{\chi_E^2}$  for both the one-peak dataset and the two-peak dataset are narrow and centered on one. The spread associated with the Phys-NN results is larger. To investigate correlations

between  $\overline{\chi_E^2}$  and  $\overline{S_R}$ , in Fig. 5 we show scatter plots for the one-peak and two-peak datasets. Some correlation is visible in the Phys-NN results, displayed in the top two panels, especially for the two-peak dataset. Conversely, the MaxEnt scatter plots show no correlation between  $\overline{\chi_E^2}$  and  $\overline{S_R}$ , since the  $\overline{\chi_E^2}$  values are relatively constant around one, even for widely different  $\overline{S_R}$ . The correlations between  $\overline{\chi_E^2}$  and  $1 - \overline{R_R^2}$  exhibit an almost identical pattern and are thus not included here.

Direct comparison of Phys-NN and MaxEnt outputs is presented in Fig. 6, where we display the Phys-NN *best* (left panels), *average* (central panels), and *worst* (right panels) reconstructed response functions, according to the  $\overline{S_R}$  values of the Phys-NN results, and the corresponding Euclidean responses from the one-peak dataset. Here, the training is performed on the combined dataset, to better test whether Phys-NN is able to learn how to simultaneously reconstruct one-peak and two-peak response functions. The uncertainty associated with the random initialization of the Phys-NN parameters is estimated by performing ten independent training procedures, each corresponding to a distinct random seed used by the training procedure. We gather the predictions obtained from each of these ten runs to estimate the error band displayed by the shaded area in Fig. 6. Not only the best and the average but also the worst response functions reconstructed with the Phys-NN are in better agreement with the original ones than are those obtained with the MaxEnt algorithm. The Laplace transform of the Phys-NN response functions are also in ex-

cellent agreement with the original Euclidean responses: the  $\chi_E^2$  values are 1.071, 0.902, and 1.834 for the best, average, and worst reconstructions, respectively. As discussed previously, by design the MaxEnt  $\chi_E^2$  values are all very close to one.

An analogous pattern emerges in the two-peak dataset. In this case, the best and the average Phys-NN responses, represented in the left and central of Fig. 7, respectively, are in excellent agreement with the original ones. Only minor discrepancies are visible in the worst reconstruction, displayed in the right panels. Although larger than in the one-peak case, the Phys-NN reduced  $\chi_E^2$  values are more than satisfactory: the values for the best, average, and worst reconstructions are 1.102, 1.024, and 6.996, respectively. This behavior is reflected in the excellent agreement between the original and reconstructed Euclidean responses. On the other hand, despite the MaxEnt values for  $\chi_E^2$  again being very close to one, MaxEnt consistently fails to resolve the EL peak in the low-energy region. In addition, it often yields QE peaks that are shifted to higher energy transfer than in the original response functions.

Among the shortcomings of the MaxEnt technique, the most problematic is probably its poor performance in the low-energy transfer region. The results shown in Fig. 7 clearly indicate that Phys-NN performs much better there. To quantify this behavior, we define an  $\omega$ -dependent entropy,  $S_R(\omega)$ , by restricting the integral of Eq. (11) to an interval of 5 MeV around each value of the energy transfer grid  $\omega_i$  in the region  $0 < \omega < 200$

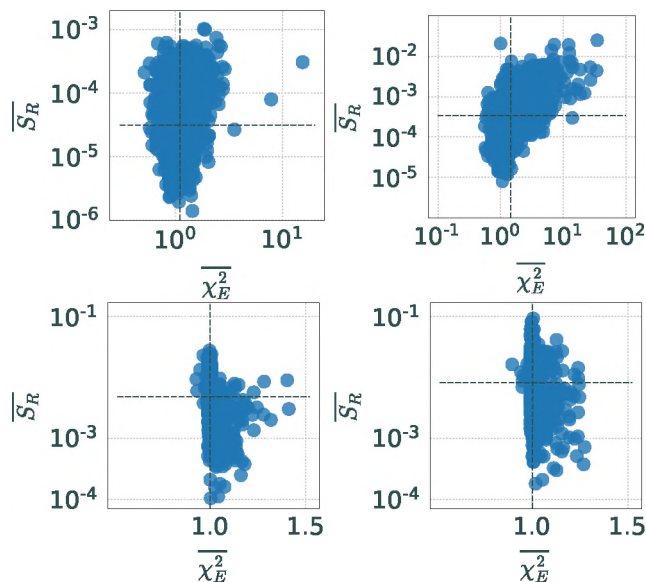


FIG. 5: Correlation plots of  $\overline{\chi_E^2}$  versus  $\overline{S_R}$  as obtained with the Phys-NN (top row) and MaxEnt (bottom row) algorithm. The left and right columns refer to the one-peak dataset and the two-peak dataset, respectively. The reference lines indicate the median  $\overline{\chi_E^2}$  and  $\overline{S_R}$  values.

MeV. First, we compute  $S_R(\omega)$  for all the responses in the test datasets; then we calculate the average and the standard error of this quantity, displayed by the shaded areas in Fig. 8 for the one-peak (left panel) and two-peak (right panel) case. The Phys-NN results are consistently below the MaxEnt ones, indicating better reconstruction performance for both one-peak and two-peak data. This fact will likely have important implications for GFMC calculations of the inclusive lepton-nucleus cross section in the low-energy regime.

## B. Predictions on noisier inputs

An important feature of any reconstruction technique is its robustness to the noise level of the input Euclidean response functions. We analyze how the performance of the Phys-NN and MaxEnt methods deteriorate when the standard deviation of the Gaussian noise of Eq. (9) is increased from  $\sigma = 10^{-4}$  to  $\sigma = 10^{-3}$ . For the results in this section, we indicate the dataset used for training by including the training data standard deviation in parentheses. Training is always done on the combined dataset, and the training strategy and hyperparameters are unchanged from those used for the noise level  $\sigma = 10^{-4}$ .

In Fig. 9 we compare sample reconstructed response functions when the noise on the input Euclidean is increased from  $\sigma = 10^{-4}$  to  $\sigma = 10^{-3}$ . In both the one-peak (top panel) and two-peak (bottom panel) response, MaxEnt clearly is more susceptible to the increased noise level than is Phys-NN. In the one-peak case, MaxEnt( $10^{-3}$ ) significantly overestimates the height of the QE peak and shifts its maximum to higher energies compared with the original response function; this behavior is not present in the Phys-NN reconstructions. In the two-peak case, Phys-NN captures the EL peak in the low-energy region for both values of  $\sigma$ . On the other hand, the MaxEnt reconstruction, already not fully satisfactory for  $\sigma = 10^{-4}$ , fails to reproduce the EL peak for  $\sigma = 10^{-3}$ . As with the one-peak case, for this higher noise level MaxEnt( $10^{-3}$ ) overestimates the height QE peak, and its position is shifted toward higher energies than in the original response function.

To further quantify these results, we calculate the change in the entropy due to the increase in the noise level in the input in the test dataset. The average values of  $S_R$  obtained from Phys-NN and MaxEnt calculations are plotted in Fig. 10. We observe that the change in the entropy due to the increase in the noise level is one order of magnitude larger for MaxEnt than that for Phys-NN( $10^{-3}$ ). In Fig. 10, we also report results for Phys-NN( $10^{-4}$ ), obtained by training Phys-NN on the low-noise data. In this case, the entropies increase by  $4.00 \times 10^{-4}$  and  $51.3 \times 10^{-4}$  for the one-peak and two-peak test datasets, respectively. While still a significant improvement compared with MaxEnt, the results for Phys-NN( $10^{-4}$ ) are not as good as those obtained by Phys-NN( $10^{-3}$ ). We conclude that Phys-NN is able to capture



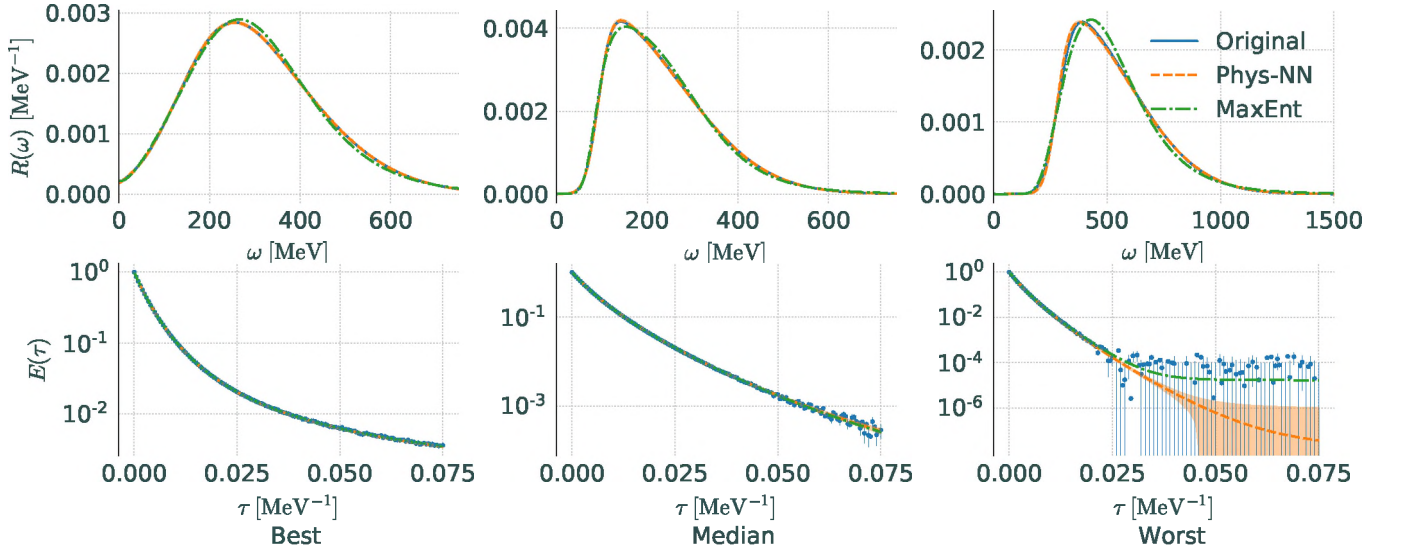


FIG. 6: Comparison between the Phys-NN and MaxEnt reconstructions for the one-peak dataset. The top row displays the response functions and the bottom row the corresponding Euclidean responses.

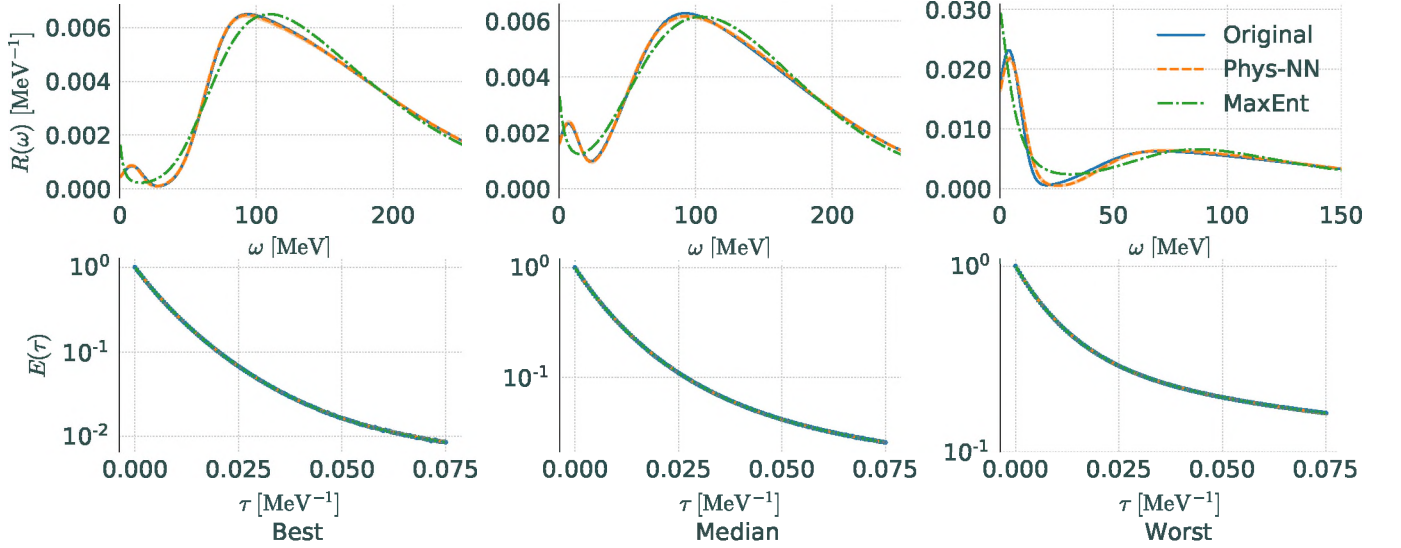


FIG. 7: Same as Fig. 6 for the two-peaks dataset.

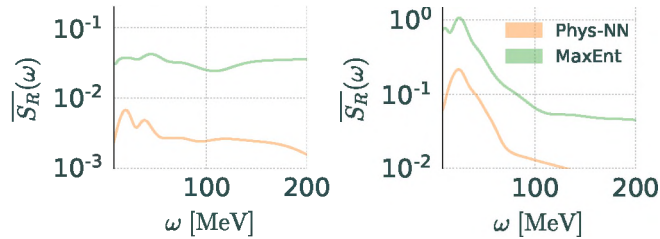


FIG. 8: Energy-dependent entropy for the Phys-NN and MaxEnt results for the one-peak (left panel) and two-peak (right panel) datasets.

the main characteristics of the response functions even

from noisier Euclidean responses. We note that it is beneficial to be able to train on a set of responses having noise levels comparable to those of the target Euclidean responses.

## V. CONCLUSIONS

This work introduces Phys-NN, a physics-informed ANN approach to approximately invert the Laplace transform and reliably reconstruct the electromagnetic response functions of atomic nuclei from their corresponding Euclidean responses.

We train, validate, and test Phys-NN, using 1 million response functions that exhibit the same features as

those measured in electron scattering experiments. Half of the simulated responses are characterized by a single asymmetric broad peak in the quasielastic region; the other half possess an additional sharp Gaussian peak to model the low-energy transfer elastic transition. Unbiased Gaussian noise ( $\sigma = 10^{-4}$ ) is added to the Euclidean responses to simulate the statistical error of typical GFMC calculations for the  ${}^4\text{He}$  nucleus. For training, we use a loss function with two terms. The first, inspired by the MaxEnt method, is an entropic loss to keep the reconstructed response functions close to the original ones. To avoid flat directions and improve the convergence of the optimization, we include a second term that seeks to keep the Laplace transform of the reconstructed responses close to the input Euclidean responses.

On a test dataset independent of that used in the training, we demonstrate that Phys-NN significantly outperforms MaxEnt in terms of both the  $S_R$  and  $1 - R_R^2$  metrics, especially on response functions characterized by two peaks. Direct examination of the reconstructed responses shows that Phys-NN is capable of capturing the low-energy structures of the responses that are often completely missed by MaxEnt. We also find that Phys-NN better reproduces the position and height of the QE peak. Phys-NN produces about an order of magnitude improvement over MaxEnt in an energy-dependent entropy measure, especially for energy transfer up to 200 MeV. This feature of Phys-NN is promising for the reliable reconstruction of the low-energy structure of nuclear response functions and muon capture rates from GFMC calculations of the Euclidean responses.

Our results show that Phys-NN is robust on a number of levels. First, Phys-NN has only two hyperparameters (the number of ANN Gaussians and the learning

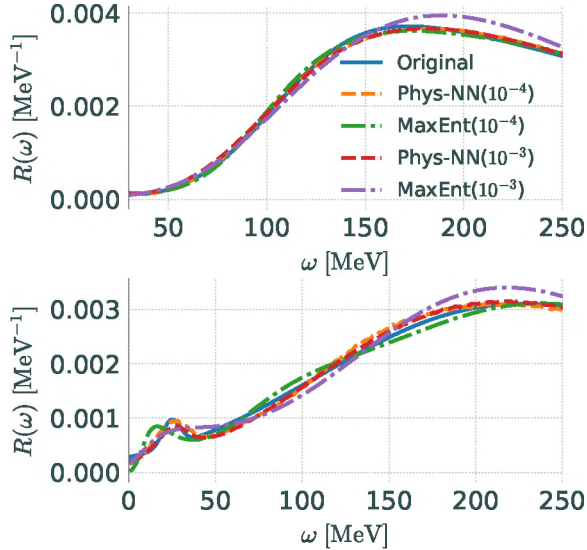


FIG. 9: Phys-NN and MaxEnt reconstruction performance with increasing level of noise in the input Euclidean responses.

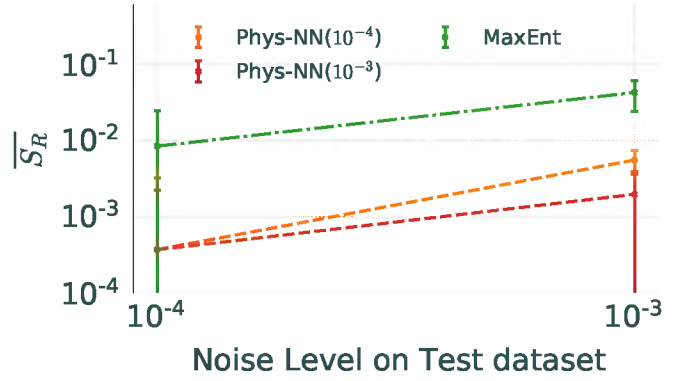


FIG. 10: Change in the entropy from increasing the standard deviation of the Gaussian noise in the input Euclidean responses from  $\sigma = 10^{-4}$  to  $\sigma = 10^{-3}$ .

rate), and the relatively small amount of validation data used for determining values for these proved to be sufficient. Second, the Phys-NN outputs from ten independent training trials show remarkably little spread among the predicted responses, indicating a desirable insensitivity within the training process employed. We stress that the associated uncertainty bands do not represent the full theoretical error of our predictions, which in principle requires propagating the statistical errors of the Euclidean response through the response functions. In future work, we intend to include full uncertainty quantification and propagation by leveraging the linearity of the Laplace transform. Third, when deployed on noisier testing data, Phys-NN maintains its advantage over MaxEnt.

In addition to the Laplace transform, primarily utilized within the GFMC method, the Lorentz kernel is commonly used in the nuclear physics community [46]. While initially restricted to light nuclear systems [47–49], its domain of applicability has recently been extended to study electron-nucleus interactions of medium-mass nuclei [50–52]. Similarly the Gaussian kernel has been found to be applicable in quantum algorithms with near-optimal computational cost to study the problem of spectral density estimation [53]. We plan on generalizing the Phys-NN method to accommodate the inversion of both the Lorentz and Gaussian kernels, with the goal of improving existing techniques.

## ACKNOWLEDGMENTS

This work was supported in part by the U.S. Department of Energy (DOE), Office of Science, Offices of Advanced Scientific Computing Research and Nuclear Physics, by the Argonne LDRD program, and by the NUCLEI, FASTMath, and RAPIDS SciDAC projects under contract number DE-AC02-06CH11357. This work was also supported in part by the National Science Foundation CSSI program under award number OAC-2004601 (BAND Collaboration). P.B., A.L., and S.M.W. were

also supported by DOE Early Career Research Program awards. We are grateful for the computing resources from the Joint Laboratory for System Evaluation at Argonne

National Laboratory.

## BIBLIOGRAPHY

- 
- [1] O. Benhar, D. Day, and I. Sick, *Rev. Mod. Phys.* **80**, 189 (2008), arXiv:nucl-ex/0603029.
  - [2] R. Wiringa, R. Schiavilla, S. C. Pieper, and J. Carlson, *Phys. Rev. C* **89**, 024305 (2014), arXiv:1309.3794 [nucl-th].
  - [3] C. Ciofi degli Atti, *Phys. Rept.* **590**, 1 (2015).
  - [4] O. Hen, G. A. Miller, E. Piasetzky, and L. B. Weinstein, *Rev. Mod. Phys.* **89**, 045002 (2017), arXiv:1611.09748 [nucl-ex].
  - [5] J. E. Amaro, M. B. Barbaro, J. A. Caballero, R. González-Jiménez, G. D. Megias, and I. Ruiz Simo, (2019), arXiv:1912.10612 [nucl-th].
  - [6] The NOvA Experiment, <http://www-nova.fnal.gov>.
  - [7] The T2K Experiment, <http://t2k-experiment.org>.
  - [8] The Deep Underground Neutrino Experiment, <http://www.dunescience.org>.
  - [9] Hyper-Kamiokande, <http://www.hyperk.org>.
  - [10] O. Benhar, P. Huber, C. Mariani, and D. Meloni, *Phys. Rept.* **700**, 1 (2017), arXiv:1501.06448 [nucl-th].
  - [11] T. Katori and M. Martini, *J. Phys. G* **45**, 013001 (2018), arXiv:1611.07770 [hep-ph].
  - [12] L. Alvarez-Ruso et al. (NuSTEC), *Prog. Part. Nucl. Phys.* **100**, 1 (2018), arXiv:1706.03621 [hep-ph].
  - [13] J. Carlson, S. Gandolfi, F. Pederiva, S. C. Pieper, R. Schiavilla, K. Schmidt, and R. Wiringa, *Rev. Mod. Phys.* **87**, 1067 (2015), arXiv:1412.3081 [nucl-th].
  - [14] J. Carlson, J. Jourdan, R. Schiavilla, and I. Sick, *Phys. Rev. C* **65**, 024002 (2002), arXiv:nucl-th/0106047.
  - [15] A. Lovato, S. Gandolfi, J. Carlson, S. C. Pieper, and R. Schiavilla, *Phys. Rev. Lett.* **117**, 082501 (2016), arXiv:1605.00248 [nucl-th].
  - [16] A. Lovato, S. Gandolfi, J. Carlson, E. Lusk, S. C. Pieper, and R. Schiavilla, *Phys. Rev. C* **97**, 022502 (2018), arXiv:1711.02047 [nucl-th].
  - [17] A. Lovato, J. Carlson, S. Gandolfi, N. Rocco, and R. Schiavilla, (2020), arXiv:2003.07710 [nucl-th].
  - [18] A. Lovato, N. Rocco, and R. Schiavilla, *Phys. Rev. C* **100**, 035502 (2019), arXiv:1903.08078 [nucl-th].
  - [19] R. Bryan, *Eur. Biophys. J.* **18**, 165 (1990).
  - [20] M. Jarrell and J. Gubernatis, *Phys. Rept.* **269**, 133 (1996).
  - [21] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, *Rev. Mod. Phys.* **91**, 045002 (2019).
  - [22] G. A. Negroita, J. P. Vary, G. R. Luecke, P. Maris, A. M. Shirokov, I. J. Shin, Y. Kim, E. G. Ng, C. Yang, M. Lockner, and G. M. Prabhu, *Phys. Rev. C* **99** (2019), 10.1103/physrevc.99.054308.
  - [23] W. G. Jiang, G. Hagen, and T. Papenbrock, *Phys. Rev. C* **100** (2019), 10.1103/physrevc.100.054326.
  - [24] L. Neufcourt, Y. Cao, W. Nazarewicz, E. Olsen, and F. Viens, *Phys. Rev. Lett.* **122**, 062502 (2019).
  - [25] J. Keeble and A. Rios, *Phys. Lett. B* **809**, 135743 (2020), arXiv:1911.13092 [nucl-th].
  - [26] C. Adams, G. Carleo, A. Lovato, and N. Rocco, (2020), arXiv:2007.14282 [nucl-th].
  - [27] M. T. McCann, K. H. Jin, and M. Unser, *IEEE Signal Process. Mag.* **34**, 85 (2017), arXiv:1710.04011 [eess.IV].
  - [28] L.-F. Arsenault, R. Neuberg, L. A. Hannah, and A. J. Millis, *Inverse Probl.* **33**, 115007 (2017).
  - [29] H. Yoon, J.-H. Sim, and M. J. Han, *Phys. Rev. B* **98**, 245101 (2018), arXiv:1806.03841 [cond-mat.str-el].
  - [30] R. Fournier, L. Wang, O. V. Yazyev, and Q. Wu, *Phys. Rev. Lett.* **124**, 056401 (2020).
  - [31] X. Xie, F. Bao, T. Maier, and C. Webster, arXiv (2019), arXiv:1905.10430 [physics.comp-ph].
  - [32] G. Shen, L. E. Marcucci, J. Carlson, S. Gandolfi, and R. Schiavilla, *Phys. Rev. C* **86**, 035503 (2012), arXiv:1205.4337 [nucl-th].
  - [33] J. Golak, R. Skibiński, K. Topolnicki, H. Wiśniewski, A. Grassi, H. Kamada, and L. E. Marcucci, *Phys. Rev. C* **88**, 015501 (2018), arXiv:1805.00103 [nucl-th].
  - [34] J. Carlson and R. Schiavilla, *Phys. Rev. Lett.* **68**, 3682 (1992).
  - [35] D. M. Titterton, *Astron. Astrophys.* **144**, 381 (1985).
  - [36] S. F. Gull and G. J. Daniell, *Nature (London)* **272**, 686 (1978).
  - [37] J. Skilling, “Classic maximum entropy,” in *Maximum Entropy and Bayesian Method*, edited by J. Skilling (Springer Netherlands, Dordrecht, 1989) pp. 45–52.
  - [38] W. Von Der Linden, R. Preuss, and V. Dose, in *Maximum Entropy and Bayesian Methods*, edited by W. von der Linden, V. Dose, R. Fischer, and R. Preuss (Springer Netherlands, Dordrecht, 1999) pp. 319–326.
  - [39] M. Hohenadler, D. Neuber, W. von der Linden, G. Wellein, J. Loos, and H. Fehske, *Phys. Rev. B* **71**, 245111 (2005), arXiv:cond-mat/0412010 [cond-mat.str-el].
  - [40] T. W. Donnelly and I. Sick, *Phys. Rev. C* **60**, 065502 (1999), arXiv:nucl-th/9905060 [nucl-th].
  - [41] N. Rocco, L. Alvarez-Ruso, A. Lovato, and J. Nieves, *Phys. Rev. C* **96**, 015504 (2017), arXiv:1701.05151 [nucl-th].
  - [42] I. Sick, *Prog. Part. Nucl. Phys.* **47**, 245 (2001), arXiv:nucl-ex/0208009 [nucl-ex].
  - [43] D. P. Kingma and J. Ba, arXiv preprint arXiv:1412.6980 (2014).
  - [44] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg, D. Mané, R. Monga, S. Moore, D. Murray, C. Olah, M. Schuster, J. Shlens, B. Steiner, I. Sutskever, K. Talwar, P. Tucker, V. Vanhoucke, V. Vasudevan, F. Viégas, O. Vinyals, P. Warden, M. Wattenberg, M. Wicke, Y. Yu, and X. Zheng, “TensorFlow: Large-scale machine learning on heterogeneous systems,”

- (2015), software available from tensorflow.org.
- [45] R. T. Birge, Phys. Rev. **40**, 207 (1932).
  - [46] V. D. Efros, W. Leidemann, and G. Orlandini, Phys. Lett. B **338**, 130 (1994), arXiv:nucl-th/9409004.
  - [47] V. D. Efros, W. Leidemann, and G. Orlandini, Phys. Rev. Lett. **78**, 432 (1997), arXiv:nucl-th/9701004.
  - [48] S. Bacca, M. A. Marchisio, N. Barnea, W. Leidemann, and G. Orlandini, Phys. Rev. Lett. **89**, 052502 (2002), arXiv:nucl-th/0112067.
  - [49] S. Bacca, N. Barnea, W. Leidemann, and G. Orlandini, Phys. Rev. Lett. **102**, 162501 (2009), arXiv:0811.4624 [nucl-th].
  - [50] S. Bacca, N. Barnea, G. Hagen, G. Orlandini, and T. Papenbrock, Phys. Rev. Lett. **111**, 122502 (2013), arXiv:1303.7446 [nucl-th].
  - [51] J. Birkhan *et al.*, Phys. Rev. Lett. **118**, 252501 (2017), arXiv:1611.07072 [nucl-ex].
  - [52] J. Simonis, S. Bacca, and G. Hagen, Eur. Phys. J. A **55**, 241 (2019), arXiv:1905.02055 [nucl-th].
  - [53] A. Roggero, Phys. Rev. A **102**, 022409 (2020), arXiv:2004.04889 [quant-ph].