2024, **25** (4), 464–475. doi 10.26089/NumMet.v25r435

doi 10.26089/NumMet.v25r435

On the possibility of using the NNQS for the Klein–Gordon–Fock equation

Aleksandr M. Kalitenko
Lomonosov Moscow State University, Faculty of Physics, Moscow, Russia
ORCID: 0000-0002-0788-2291, e-mail: am.kalitenko@physics.msu.ru

Petr I. Pronin
Lomonosov Moscow State University, Faculty of Physics, Moscow, Russia
ORCID: 0009-0008-6388-8744, e-mail: petr_pro_iv@mail.ru

Abstract: In this article, we present a method for finding quantum stationary states of the Klein–Gordon–Fock (KGF) equation using neural networks (NNs). The method has been tested on two well-known systems: a relativistic spinless particle in a Coulomb potential, and a one-dimensional relativistic harmonic oscillator. The results of training the neural network for these two systems are presented, as well as the analysis of the training process. The neural network method shows a good agreement with analytical calculations (if they can be found explicitly), providing a promising approach for solving more complex problems in quantum physics and quantum chemistry.

Keywords: quantum mechanics, neural network, Klein–Gordon–Fock equation.

For citation: A. M. Kalitenko, P. I. Pronin, "On the possibility of using the NNQS for the Klein–Gordon–Fock equation," Numerical Methods and Programming. **25** (4), 464–475 (2024). doi 10.26089/NumMet.v25r435.

О возможности применения метода NNQS для уравнения Клейна–Гордона–Фока

А. М. Калитенко
Московский государственный университет имени М. В. Ломоносова, физический факультет, Москва, Российская Федерация
ORCID: 0000-0002-0788-2291, e-mail: am.kalitenko@physics.msu.ru

П. И. Пронин
Московский государственный университет имени М. В. Ломоносова, физический факультет, Москва, Российская Федерация
ORCID: 0009-0008-6388-8744, e-mail: petr_pro_iv@mail.ru

Аннотация: В этой статье мы представляем метод вычисления стационарных состояний уравнения Клейна—Гордона–Фока с помощью нейронных сетей. Метод был апробирован на двух хорошо известных системах: релятивистской бесспиновой частице в кулоновском потенциале и одномерном релятивистском гармоническом осцилляторе. Представлены результаты обучения нейронной сети для этих двух систем, а также анализ процесса обучения. Метод нейронных сетей показывает хорошее соответствие с результатами аналитических вычислений (если они могут быть найдены в явном виде), что открывает перспективы для решения более сложных задач в области квантовой физики и квантовой химии.

Ключевые слова: квантовая механика, нейронные сети, уравнение Клейна–Гордона–Фока.

Для цитирования: Калитенко А.М., Пронин П.И., О возможности применения метода NNQS для уравнения Клейна–Гордона–Фока // Вычислительные методы и программирование. 2024. **25**, № 4. 464–475. doi 10.26089/NumMet.v25r435.

2024, **25** (4), 464–475. doi 10.26089/NumMet.v25r435

1. Introduction. Machine learning models and algorithms have been applied in various fields of physics [1–3]. Machine learning has its advantages and shortcomings over numerical methods [4, 5], complements and develops methods of computational physics. In quantum theory, there are some equations that are difficult to solve analytically. In 2017, two physicists, Giuseppe Carlo and Matthias Troyer, proposed the Neural Network Quantum States method (NNQS) [6, 7]. This method is based on the use of neural networks to approximate the ground state of a quantum system. The NNQS method has been shown to be very effective in solving difficult problems in quantum physics [8–11]. Nowadays, the NNQS is used in many problems: PauliNet [12], a deuteron [13], quantum wells [14], etc. The variational principle for quantum field theories was proposed in [15].

The Schrödinger equation is a well-known equation that plays a crucial role in quantum mechanics [16, 17], but Schrödinger equation is not relativistic, which is noticeable in some systems, for example, an atom. The relativistic Klein–Gordon–Fock equation [18–21] describes spinless particles but can be generalized for particles with integer and half-integer spins and has many applications for describing macroscopically isotropic crystals, pi-mesons, and Cooper pairs. A physics-informed neural network (PINN) approach has been considered in [22]. In the case of the Klein–Gordon–Fock equation, it is difficult to write out the Hamiltonian explicitly; only the square of the Hamiltonian is obvious. Additionally, the Klein–Gordon–Fock equation contains two extra degrees of freedom [23, 24], which makes it challenging to apply variational techniques based on searching for the minimum energy. In this article, we will discuss and analyze methods for overcoming the above-mentioned issues and investigate two well-known physical systems: the relativistic spinless particle in a Coulomb potential and the one-dimensional relativistic harmonic oscillator.

2. Method.

2.1. Theory. The KGF equation is a four-dimensional equation (x_1, x_2, x_3, t coordinates). However, in this paper, we will focus on using neural networks to study quantum systems with stationary solutions. These systems do not include time derivatives, so we will not consider them in this discussion. Many three-dimensional problems can be simplified to one dimension, for example, in the case of a particle in a spherically symmetric potential. We will demonstrate this below.

In relativistic quantum mechanics, for a free particle with rest mass m, the relationship between energy E and momentum p is

$$E^2 = p^2c^2 + m^2c^4, \tag{1}$$

where c is the speed of light in a vacuum. If the particle is in a potential $V(x)$, the expression (1) is modified in the following way:

$$[E - V]^2 = p^2c^2 + m^2c^4. \tag{2}$$

The corresponding wave equation in one dimension is

$$[E - V]^2\psi(x) = \left[-\hbar^2c^2\frac{d^2}{dx^2} + m^2c^4\right]\psi(x), \tag{3}$$

where $\hbar$ is the reduced Planck constant, $\psi(x)$ is the wavefunction.

It is a so-called relativistic Klein–Gordon–Fock equation. We can rewrite it in the Schrödinger-like form [25, 26]

$$E_{\text{eff}}\,\psi(x) = \left[-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V_{\text{eff}}(x)\right]\psi(x), \tag{4}$$

where we denote

$$E_{\text{eff}} = \frac{E^2 - m^2c^4}{2mc^2}, \tag{5}$$

$$V_{\text{eff}} = \frac{2EV(x) - V^2(x)}{2mc^2}, \tag{6}$$

$$\hat{H}_{\text{eff}} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V_{\text{eff}}(x), \tag{7}$$

$$E_{\text{eff}} = \frac{\langle\psi \mid \hat{H}_{\text{eff}} \mid \psi\rangle}{\langle\psi \mid \psi\rangle} \geqslant E_0, \tag{8}$$

E_0 is the minimum energy for our system. Since the energy E is extracted from E_{eff}, and V_{eff} also contains E, strictly speaking, this is not the eigenvalue problem. Nevertheless, in practice, this procedure works for a number of problems, as we will see below.

2024, **25** (4), 464–475. doi 10.26089/NumMet.v25r435

2.2. Learning process. To solve our problems we apply the *PyTorch* library of the Python programming language. A neural network is a tool which can be used to find the wavefunctions of stationary states. We want to select a neural network that will lead us to the bound states. Therefore, we solve the optimization problem and reduce a loss function, i.e. minimize the energy. However, as mentioned above, when solving the KGF equation, we face the problem of quadratic energy, which leads to modes with negative and positive total energies. To eliminate negative modes, we introduce an ansatz for wavefunctions $\psi(x) = network(x) \cdot f(x)$, where $f(x)$ is the envelope, or analytical approximation, or network from the Shrödinger solution, and $network(x)$ is a neural network. The coordinate of point x is the input of the neural network, and in the output we get the value of the wavefunction at this point, the phase of the wavefunction is not required, so only the real-valued wavefunction is considered. The loss function solely depends on the solutions obtained by the neural network while the training process is fully data-free. This formulation results in an unsupervised learning method. We also tested the energy values in a set of points different from those used during training, and obtained similar results. In all cases, we apply a multilayer perceptron (MLP) with one hidden layer (50–100 nodes for different problems), see Fig. 1. In our opinion, a single hidden layer is sufficient for our tasks, although two hidden layers could also be appropriate. We use the hyperbolic tangent ($\tanh(x)$) as the activation function. There are other activation functions available, such as ReLU, SiLU, and others. The choice of them depends on the specific task. In our research, we choose the hyperbolic tangent activation function because of its symmetrical domain of definition. Additionally, we use a scheduler to decrease the learning rate and L_2 regularization [27, 28] to avoid instabilities associated with the growth of weights

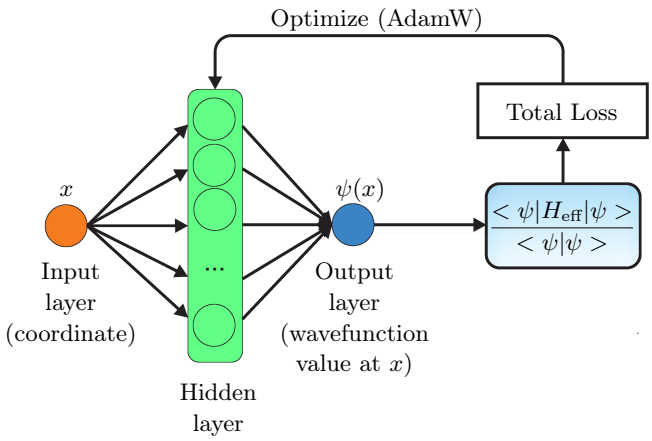

Figure 1. The neural network architecture used in our study and the optimization process are illustrated in the feedback loop.

$$L_{2,\,\text{regul}} = \tilde{\lambda} \sum_i \tilde{\alpha}_i^2, \tag{9}$$

where $\tilde{\alpha}_i$ are the weights of the NN, and $\tilde{\lambda}$ is the Lagrange multiplier which is introduced for optimization, namely for minimization of the predefined loss function. We found it necessary to implement a regularization in a custom loss function. With the AdamW optimizer [29], one can use weight decay instead of L_2 regularization. We apply the AdamW optimizer in our experiments. In the code, we left the regularization in case of other optimizer. Since the effective potential V_{eff} includes the total energy of the particle, we iteratively replace it with the previous value during training. The initial energy E is set to be closed to mc^2, since the effective potential V_{eff} contains E as a variable. In our work, we use the term "epoch" to refer to one cycle of weight optimization, based on minimizing a sum of loss functions across the coordinate mesh x_i.

3. Application.

3.1. Relativistic spinless particle in a Coulomb potential. Historically, Schrödinger solved the problem of the relativistic hydrogen atom in the formulation of wave quantum mechanics. For the Coulomb potential, the transformed equation (4) is similar to the Schrödinger equation, and an exact expression for the energy can be obtained (the spinless relativistic electron model) [26]

$$E_{nl} = \frac{mc^2}{\sqrt{1 + \dfrac{Z^2\alpha^2}{(n - \delta_l)^2}}} - mc^2, \tag{10}$$

with

$$\delta_l = l + \frac{1}{2} - \sqrt{\left(l + \frac{1}{2}\right)^2 - Z^2\alpha^2}, \tag{11}$$

where n is the principal quantum number, l is the azimuthal quantum number, $\alpha \approx 1/137$ is the fine-structure constant, Z is the charge number.

2024, **25** (4), 464–475. doi 10.26089/NumMet.v25r435

The neural network becomes unstable around $Z = 70$ with $l = 0$. This can be explained by the impact of the large attractive potential contributions. The manifestation of this feature is also contained in the analytical solution (10), (11) for $l = 0$ and for the large values of Z δ_l becomes complex. Therefore, there are different constraints on Z for each value of l. The physical reason of this feature is the incorrectness of the application of a single-particle equation in strong fields [26], which leads to large fluctuations in the field, including the number of particles.

Let us set $m = c = \hbar = 1$, $e^2 = \alpha$. For the stationary case the problem is formulated in three-dimensional space. Since the potential has spherical symmetry, then the variables in the equation are split and, as a result, the separated parts of the original equation can be solved independently. Therefore, we can write down the wavefunction as $\Psi(r, \tilde{\theta}, \varphi) = R(r)Y_l^m(\tilde{\theta}, \varphi)$, where $R(r)$ is the radial part, and $Y_l^m(\tilde{\theta}, \varphi)$ is the spherical harmonic function of degree l and order m, r is the radius, $\tilde{\theta}$ is the polar angle, φ is the azimuthal angle. For the radial part $R_{nl}(r)$ we introduce $\psi_{nl}(r) = rR_{nl}(r)$. Now the Klein–Gordon–Fock equation can be reduced to a one-dimensional equation for ψ_{nl} [26] with

$$\hat{H}_{\text{eff}} = -\frac{1}{2}\frac{d^2}{dr^2} + \frac{1}{2r^2}l(l+1) + V_{\text{eff}}, \quad V(r) = -\frac{Z\alpha}{r}, \tag{12}$$

where V_{eff} is defined by (6).

Since we are working in spherical coordinates and with the Coulomb interaction, the above integrals

$$\langle \psi_{nl,\theta} \mid \hat{H}_{\text{eff}} \mid \psi_{nl,\theta} \rangle = \int_0^\infty \left\{ \frac{1}{2}\left(\frac{d\psi_{nl,\theta}(r)}{dr} \right)^2 + \psi_{nl,\theta}^2(r)\left(\frac{1}{2r^2}l(l+1) + V_{\text{eff}} \right) \right\} dr, \tag{13}$$

where we have integrated the first term by parts, θ is the neural network parameters.

Applying discretization, we get

$$\langle \psi_{nl,\theta} \mid \hat{H}_{\text{eff}} \mid \psi_{nl,\theta} \rangle \approx \sum_{i=1}^{N_r} w_i \left[\frac{1}{2}\left\{ \frac{d\psi_{nl,\theta}(r)}{dr} \right\}_{r=r_i}^2 + \psi_{nl,\theta}^2(r_i)\left(\frac{1}{2r_i^2}l(l+1) + V_{\text{eff}}(r_i) \right) \right], \tag{14}$$

$$\langle \psi_{nl,\theta} \mid \psi_{nl,\theta} \rangle \approx \sum_{i=1}^{N_r} w_i \psi_{nl,\theta}^2(r_i) \equiv N, \tag{15}$$

$$Loss = \frac{\langle \psi \mid \hat{H}_{\text{eff}} \mid \psi \rangle}{\langle \psi \mid \psi \rangle} + L_{2,\,\text{regul}}, \tag{16}$$

w_i is the mesh weights, N_r is the number of nodes, N is the normalization factor, and $L_{2,\,\text{regul}}$ is the regularization from (9).

At this stage, we take an additional step and add the envelope to the NNQS when calculating losses. The envelope is a function that multiplies NN, and its purpose is to satisfy boundary conditions. In our case, we should require $\psi_{nl}(0) = 0$. This can be achieved by specifying the asymptotics at zero in the standard way:

$$\psi_{nl}(r) \longrightarrow \psi_{nl}(r)r^{l+1}.$$

We denote $U_{\text{eff}} = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N_r} w_i \psi_{nl,\theta}^2(r_i)\left(\frac{1}{2r_i^2}l(l+1) + V_{\text{eff}}(r_i) \right)$, $\quad K = \dfrac{1}{2N}\displaystyle\sum_{i=1}^{N_r} w_i \left\{ \frac{d\psi_{nl,\theta}(r)}{dr} \right\}_{r=r_i}^2$.

We use the ansatz $\psi_{nl,\theta}(r) = network(r) \cdot r^{l+1}$, analyze the behavior of the NNQS wavefunctions after training and compare the obtained energies with the exact formula (Fig. 2, red dotted line) for $Z = 50$. In Fig. 2 a the normalized NNQS wavefunction for the $n = 1, l = 0$ state is presented, and the difference between the relativistic ψ_{Norm} and non-relativistic ψ_{Sh} cases is noticeable. In Fig. 3, the loss function during training is shown. In Fig. 3 we are observing a break. It is well known that there is not a single pattern for the behavior of the loss function. This pattern (break) can be caused by the presence of a strong local minimum in a relatively simple solution. As a result, our model becomes stuck in this area for some time before it finds a more complex solution. From Fig. 2 c, one can see that the energy value is in good agreement with the exact formulas (10), (11) without perturbation theory and approximations. If we expand the expression (10) in small fine structure constant, we get

$$E_{nl} \approx -\frac{mc^2}{2n^2}Z^2\alpha^2\left\{ 1 + \frac{Z^2\alpha^2}{n^2}\left(\frac{2n}{2l+1} - \frac{3}{4} \right) \right\}. \tag{17}$$

2024, **25** (4), 464–475. doi 10.26089/NumMet.v25r435

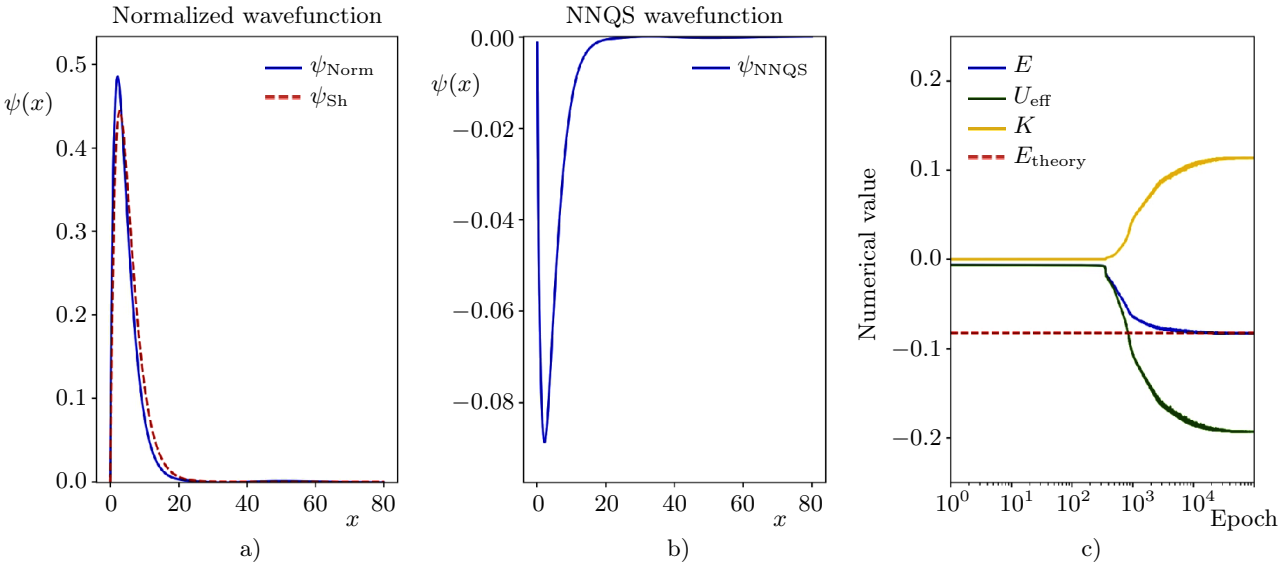

Figure 2. The NNQS wavefunction and the training process for $n = 1$, $l = 0$, $Z = 50$: a) the normalized NNQS wavefunction ψ_{Norm} (blue line) and the non-relativistic wavefunction ψ_{Sh} (red dotted line); b) the NNQS wavefunction; c) the energy E (blue line), the "kinetic" energy K (yellow line), U_{eff} (green line), and E_{theory} (red dotted line, see (10)).

For $n = 1$, $l = 0$, $Z = 50$ the comparison with exact formula (10) gives us 6% accuracy. This is essential when considering more subtle effects.

In Fig. 4 a, b the NNQS wavefunction for $n = 2$, $l = 1$ is presented. Relativistic corrections are now less noticeable, which is consistent with the analytical solution. The KGF equation is invariant with respect to the replacement of the sign of the wavefunction $\pm\psi(x)$, so arbitrariness in the sign is acceptable.

Now let's move on to the first excited state with $n = 2$, $l = 0$. The approach is almost identical to the procedure of finding the ground state wavefunction, but now we want to find an independent orthogonal solution

$$\langle \psi_{10,\theta} \mid \psi_{20,\theta'} \rangle = 0, \tag{18}$$

where $\psi_{10,\theta}$ and $\psi_{20,\theta'}$ are the wavefunctions of the ground and first excited states respectively. The obvious way to impose orthogonality is to add (18) to the loss function using a Lagrange multiplier:

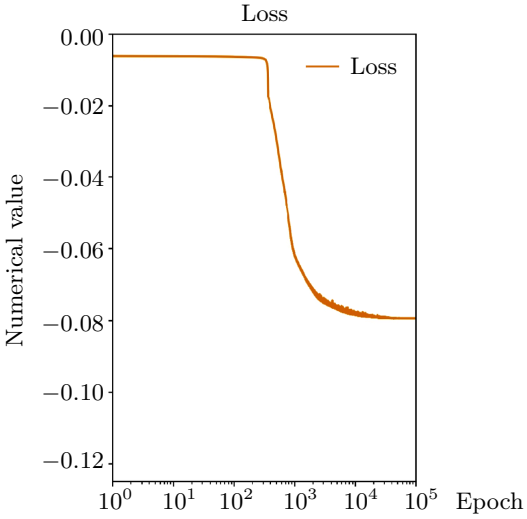

Figure 3. The loss function during training for $n = 1$, $l = 0$, $Z = 50$.

$$Loss = \frac{\langle \psi_{20,\theta'} \mid \hat{H}_{\text{eff}} \mid \psi_{20,\theta'} \rangle}{\langle \psi_{20,\theta'} \mid \psi_{20,\theta'} \rangle} + \lambda \frac{\langle \psi_{10,\theta} \mid \psi_{20,\theta'} \rangle^2}{\langle \psi_{20,\theta'} \mid \psi_{20,\theta'} \rangle \langle \psi_{10,\theta} \mid \psi_{10,\theta} \rangle} + L_{2,\,\text{regul}}, \tag{19}$$

where λ is the Lagrange multiplier.

In Fig. 5 the NNQS wavefunction for the $n = 2$, $l = 0$ state is presented, relativistic corrections are more noticeable compared with the case of $n = 2$, $l = 1$.

As one can see from Fig. 4 c and Fig. 5 c, the energy levels of the states ψ_{21} and ψ_{20} differ. This is a result of the influence of the centrifugal potential.

In Fig. 6 and 7, the wavefunctions and energy levels for two states ($n = 1$, $l = 0$ and $n = 2$, $l = 1$ correspondingly) with $Z = 1$ are given. The difference from Schrödinger's wavefunctions is tiny, which is expected, see the equation (17).

Levels with large numbers may be investigated similarly.

2024, **25** (4), 464–475. doi 10.26089/NumMet.v25r435

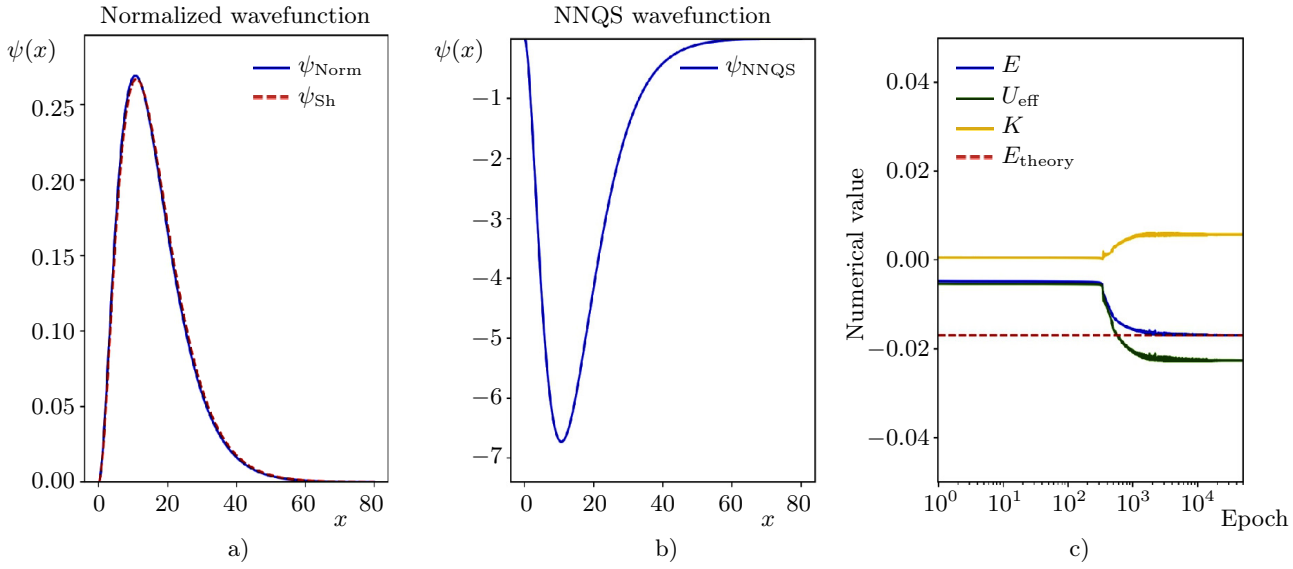

Figure 4. The NNQS wavefunction and the training process for $n = 2$, $l = 1$, $Z = 50$: a) the normalized NNQS wavefunction ψ_{Norm} (blue line) and the non-relativistic wavefunction ψ_{Sh} (red dotted line); b) the NNQS wavefunction; c) the energy E (blue line), the "kinetic" energy K (yellow line), U_{eff} including a centrifugal potential $l(l+1)/(2r^2)$ (green line), and E_{theory} (red dotted line, see (10)).

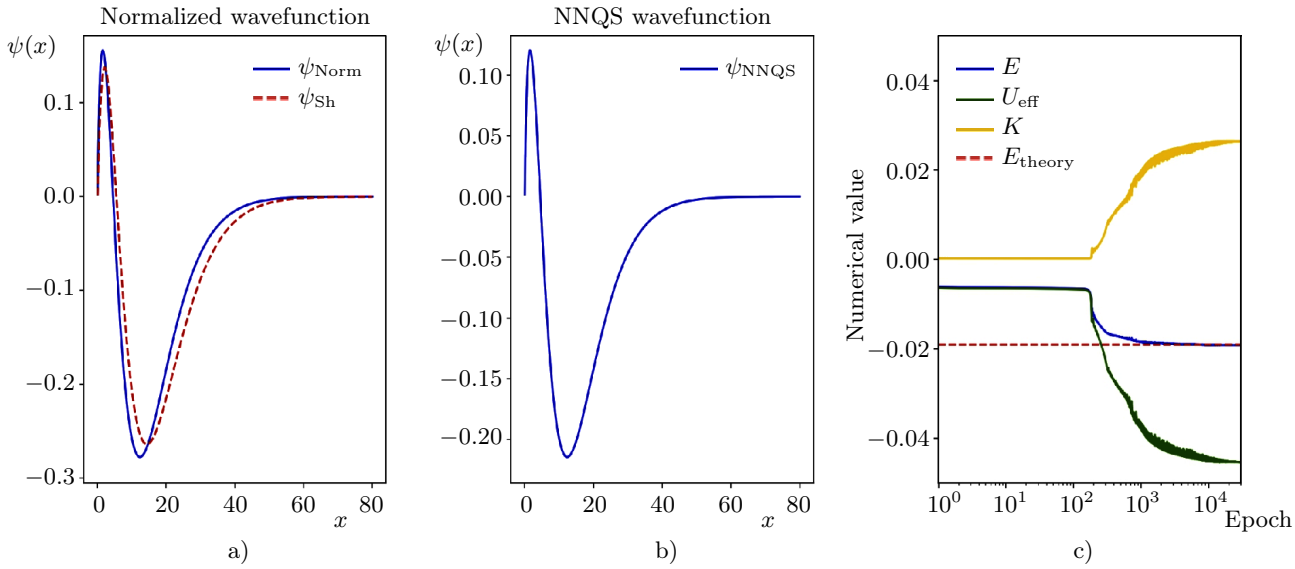

Figure 5. The NNQS wavefunction and the training process for $n = 2$, $l = 0$, $Z = 50$: a) the normalized NNQS wavefunction ψ_{Norm} (blue line) and the non-relativistic wavefunction ψ_{Sh} (red dotted line); b) the NNQS wavefunction; c) the energy E (blue line), the "kinetic" energy K (yellow line), U_{eff} including a centrifugal potential $l(l+1)/(2r^2)$ (green line), and E_{theory} (red dotted line, see (10)).

3.2. Relativistic 1D harmonic oscillator. Here we solve the Klein–Gordon–Fock equation for the harmonic oscillator (HO) in 1D for a single particle using the NNQS method. The non-relativistic Hamiltonian is

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{1}{2}m\omega^2 x^2, \tag{20}$$

where m is the particle mass, and ω is the frequency.

2024, **25** (4), 464–475. doi 10.26089/NumMet.v25r435

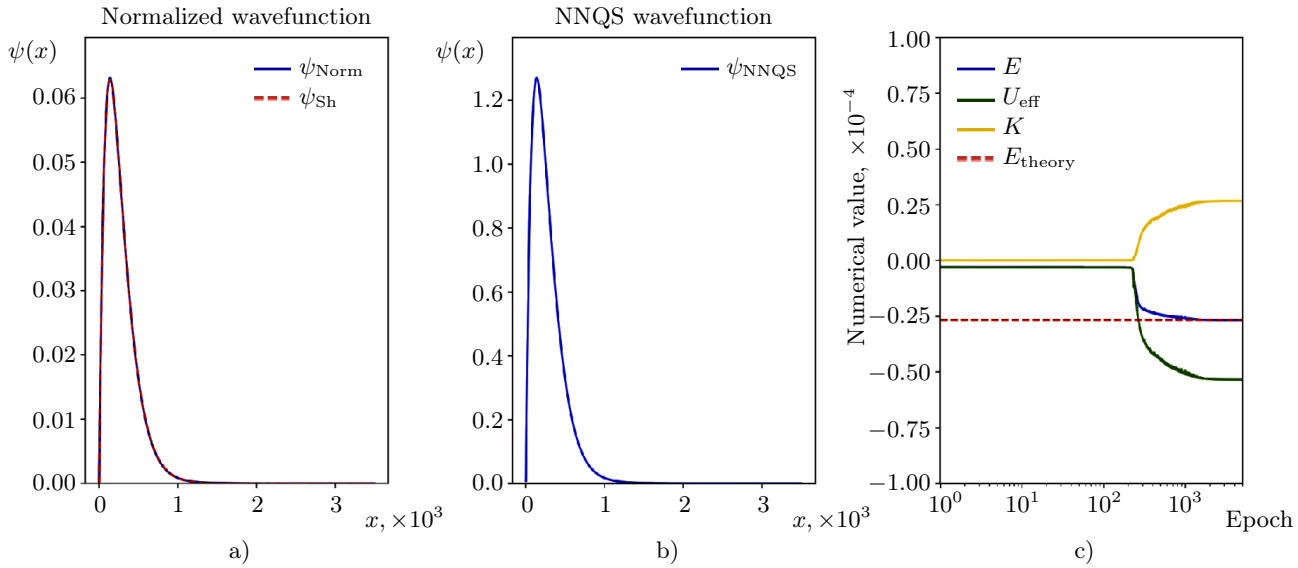

Figure 6. The NNQS wavefunction and the training process for $n = 1$, $l = 0$, $Z = 1$: a) the normalized NNQS wavefunction ψ_{Norm} (blue line) and the non-relativistic wavefunction ψ_{Sh} (red dotted line); b) the NNQS wavefunction; c) the energy E (blue line), the "kinetic" energy K (yellow line), U_{eff} (green line), and E_{theory} (red dotted line, see (10)).

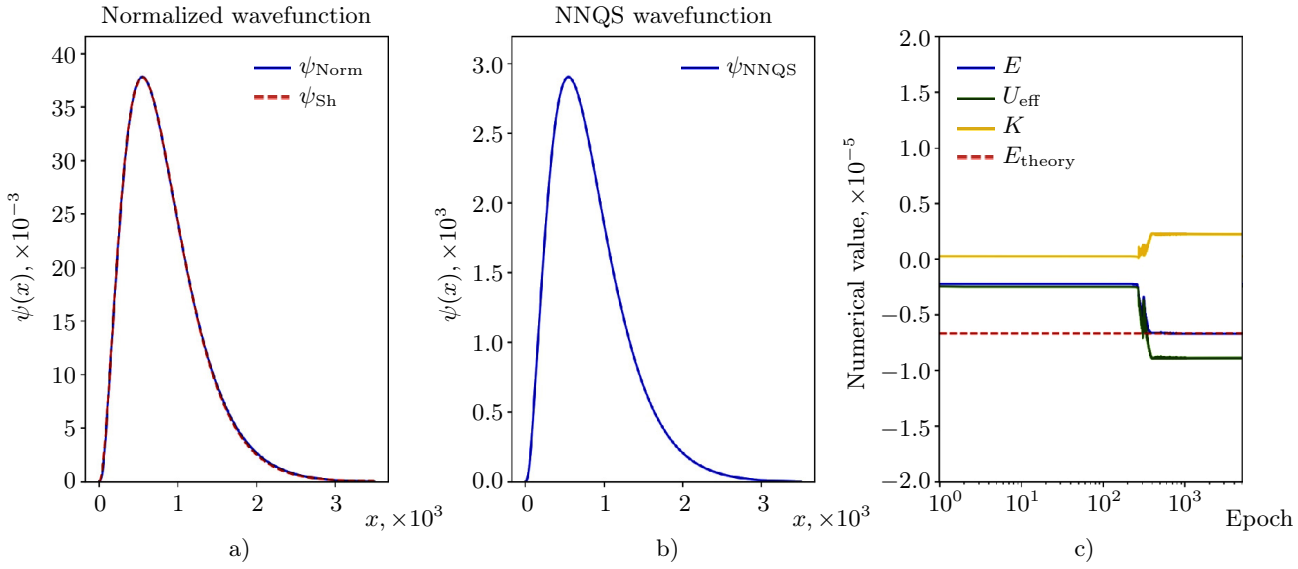

Figure 7. The NNQS wavefunction and the training process for $n = 2$, $l = 1$, $Z = 1$: a) the normalized NNQS wavefunction ψ_{Norm} (blue line) and the non-relativistic wavefunction ψ_{Sh} (red dotted line); b) the NNQS wavefunction; c) the energy E (blue line), the "kinetic" energy K (yellow line), U_{eff} including a centrifugal potential $l(l+1)/(2r^2)$ (green line), and E_{theory} (red dotted line, see (10)).

We work in a $m = c = \hbar = 1$ unit system. The loss function is

$$Loss = \frac{\langle \psi \mid \hat{H}_{\text{eff}} \mid \psi \rangle}{\langle \psi \mid \psi \rangle} + \mu(\langle \psi \mid \psi \rangle - 1)^2 + L_{2,\,\text{regul}}, \tag{21}$$

where μ is the hyperparameter (a machine learning parameter that is used to control the learning process) at the the normalization term. We denote $U_{\text{eff}} = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N_r} w_i \psi^2_{nl,\theta}(r_i) V_{\text{eff}}(r_i)$.

2024, **25** (4), 464–475. doi 10.26089/NumMet.v25r435

In the previous problem, we have not added normalization to simplify the learning process. Normalization during the learning process makes the search for the optimal function more difficult, but it can be applied after training to help improve convergence. It can also be incorporated into a loss function, as we have done. In Fig. 8–10 we compare the obtained wavefunctions and energies with non-relativistic cases.

For the ground state, we choose the following ansatz

$$\psi_{0,\theta}(x) = network(x) \cdot \exp(-\omega x^2/2).$$

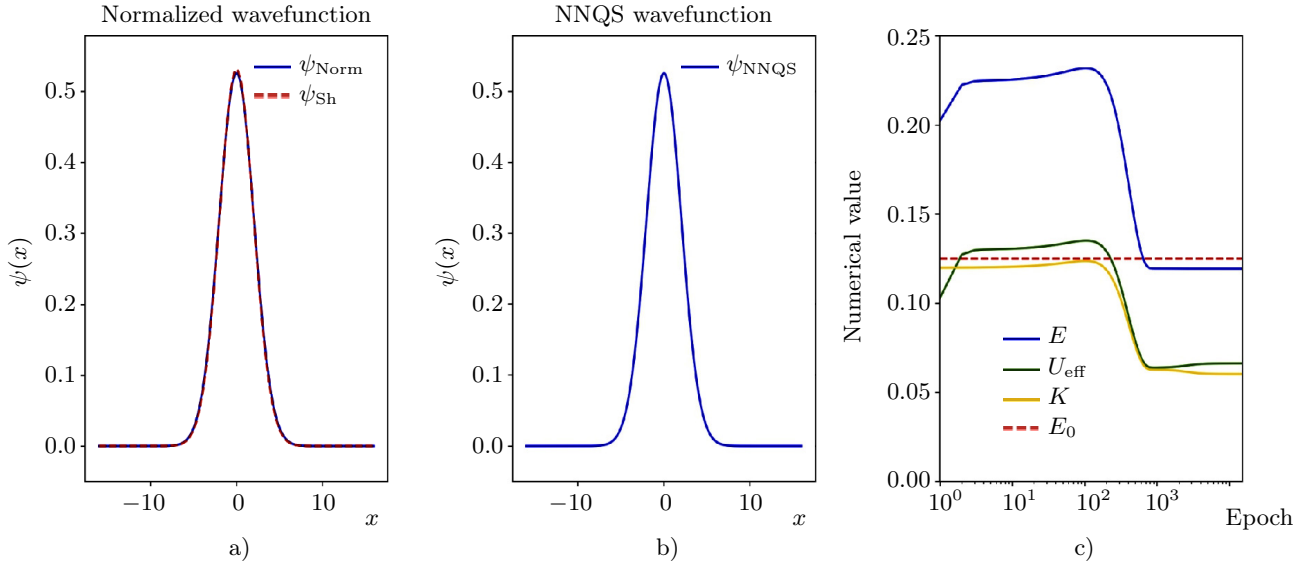

Figure 8. The NNQS wavefunction and the training process for $n = 0$, $\omega = 0.25$: a) the normalized NNQS wavefunction ψ_{Norm} (blue line) and the non-relativistic wavefunction ψ_{Sh} (red dotted line); b) the NNQS wavefunction (blue line); c) the energy E (blue line), the "kinetic" energy K (yellow line), U_{eff} (green line) and the energy E_0 in the non-relativistic case (red dotted line).

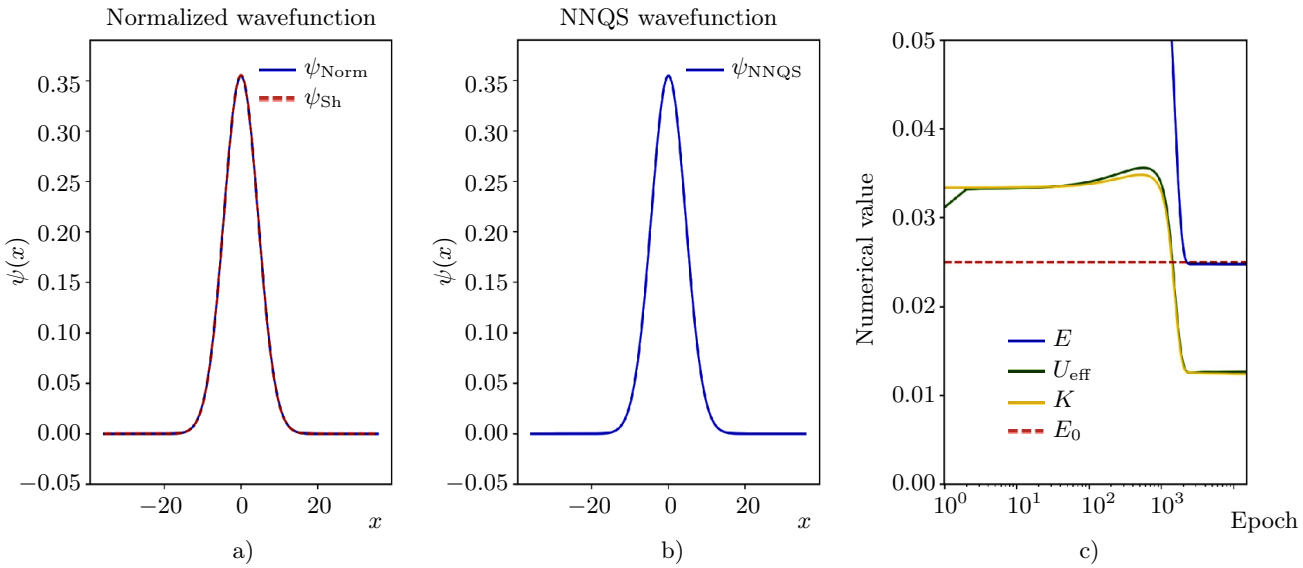

Figure 9. The NNQS wavefunction and the training process for $n = 0$, $\omega = 0.05$: a) the normalized NNQS wavefunction ψ_{Norm} (blue line) and the non-relativistic wavefunction ψ_{Sh} (red dotted line); b) the NNQS wavefunction (blue line); c) the energy E (blue line), the "kinetic" energy K (yellow line), U_{eff} (green line) and the energy E_0 in the non-relativistic case (red dotted line).

2024, **25** (4), 464–475. doi 10.26089/NumMet.v25r435

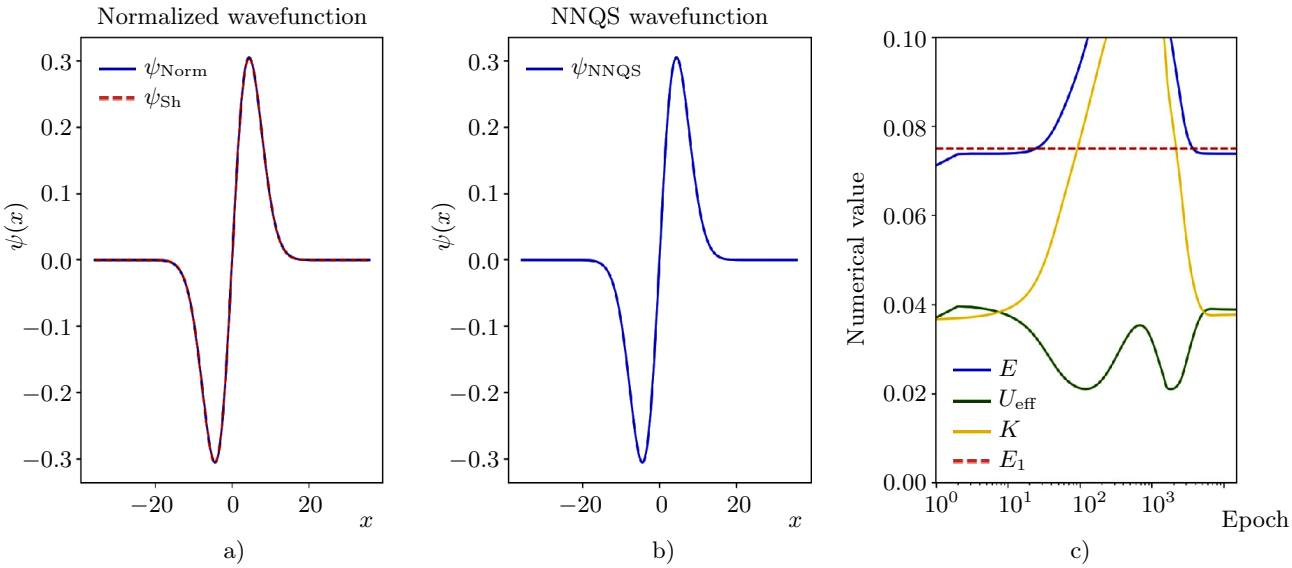

Figure 10. The NNQS wavefunction and the training process for $n = 1$, $\omega = 0.05$: a) the normalized NNQS wavefunction ψ_{Norm} (blue line) and the non-relativistic wavefunction ψ_{Sh} (red dotted line); b) the NNQS wavefunction (blue line); c) the energy E (blue line), the "kinetic" energy K (yellow line), U_{eff} (green line) and the energy E_1 in the non-relativistic case (red dotted line).

We additionally investigate other similar ansatzes, for example, $\psi_{0,\theta}(x) = network(x) \cdot \exp(-\omega x^2/3)$ and neural network based on the Shrödinger equation, and obtain the same results. The closer the ansatz is to true dependence, the faster the learning and more accurate the approximation of the neural network. It should also be pointed out that our problem has only quasi-stationary states, even in perturbation theory there is the same feature [30]. The potential $m\omega^2 x^2/2$ is not limited from above and we get strong fields, which leads to quasi-stationarity. Therefore, the selection of the ansatz and the training is not arbitrary.

For the case $\omega = 0.25$ we get the energy for the ground state $E_{NN,0} \approx 0.11940$. In its turn, the expression for energy levels, obtained in the first order approximation in the perturbation theory, with [16, 31]

$$E_{\text{pert},n} \approx \hbar\omega\left(n + \frac{1}{2}\right) - \frac{3\hbar^2\omega^2}{32mc^2}(2n^2 + 2n + 1), \tag{22}$$

gives $E_{\text{pert},0} \approx 0.11914$ vs $E_{NN,0} \approx 0.11940$ (see Fig. 8), $E_n = \hbar\omega(n+\frac{1}{2})$ is the energy levels of the non-relativistic harmonic oscillator. All red dotted lines in Fig. 8 correspond to the Shrödinger solution.

Now let us consider the first excited state. The approach is almost identical to the procedure of finding the ground state wavefunction, but we need to impose the orthogonality requirement

$$\langle \psi_{0,\theta} \mid \psi_{1,\theta'} \rangle = 0,$$

where $\psi_{0,\theta}, \psi_{1,\theta'}$ are the wavefunctions of the ground and first excited states, respectively. An easy way to impose this condition is to make ψ_1 antisymmetric (taking into account that ψ_0 is symmetric),

$$\psi_{1,\theta'}(x) \to \psi_{1,\theta'}(x) - \psi_{1,\theta'}(-x).$$

We choose the ansatz based on the antisymmetry of the wavefunction and the solution of the Schrödinger equation.

Setting $\omega = 0.05$ (see Fig. 9, 10), we obtain the following results: $E_{NN,0} \approx 0.024768$, $E_{NN,1} \approx 0.073865$. Comparison with the results calculated by the perturbation theory (22) gives:

$$E_{NN,0} \approx 0.024768 \text{ vs } E_{\text{pert},0} \approx 0.024766,$$
$$E_{NN,1} \approx 0.073865 \text{ vs } E_{\text{pert},1} \approx 0.073828.$$

2024, **25** (4), 464–475. doi 10.26089/NumMet.v25r435

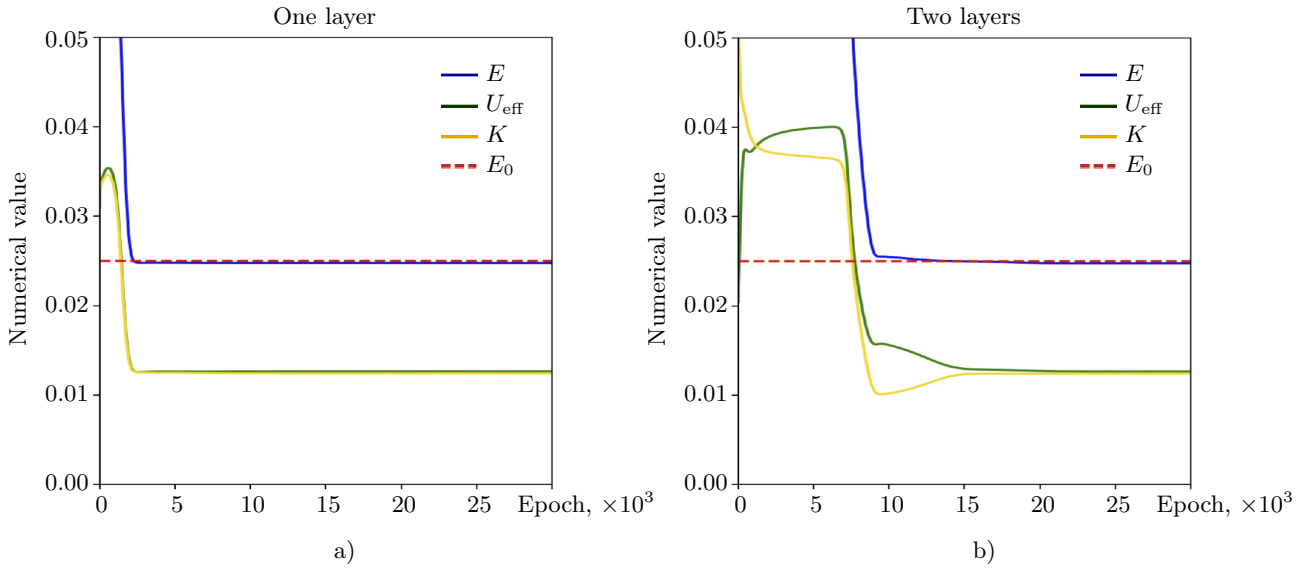

Figure 11. Comparison of the learning process from the epochs at $n = 0$, $\omega = 0.05$ for: a) a single-layer neural network and b) a two-layer neural network. Number of nodes in hidden layers is 100.

As we can see, the higher the energy level and frequency, the more the perturbative theory differs from the NNQS method. It should be noted that although the architecture of a neural network does affect the learning process, it does not significantly impact the final result. For example, Fig. 11 presents a comparison of the learning process with a single-layer neural network (a) and a two-layer neural network (b). A two-layer neural network reaches a solution at 20,000 epochs, while a single-layer network reaches it at 5,000 epochs. It is well known that with an increase in the number of layers, a neural network is able to describe more complex functions. However, the learning process becomes more expensive. The learning process also depends on the weight initialization and seed parameters.

4. Error analysis. We evaluate the accuracy and stability of our method and analyze the standard deviation of the normalized wavefunctions, the relative standard deviation of energy at different initial weights of the neural network, optimizer parameters, and seeds. The corrected sample standard deviation is

$$sd = \sqrt{\frac{1}{N_s - 1}\sum_{i=1}^{N_s}(x_i - \overline{x})^2}, \tag{23}$$

where N_s is the size of sample, x_i are the observed values of the sample items, and $\overline{x} = \sum\limits_{i=1}^{N_s}\dfrac{x_i}{N_s}$ is the mean value of these observations. The results for $N_s = 30$ samples are presented below.

In Fig. 12 a a the standard deviation of the normalized wavefunction for a relativistic spinless particle in a Coulomb potential with $n = 1, l = 0$ is shown. The relative standard deviation of sd_E/E_{mean} (use E instead of x in (23)) for this state is 0.07%. The analogous characteristics for a relativistic harmonic potential with $n = 0$, $\omega = 0.05$ are demonstrated in Fig. 12 b. The relative standard deviation of energy for this state is 0.0025%. We also obtain quite small deviations for other states.

5. Conclusion. In this work, the NNQS for the Klein–Gordon–Fock equation is proposed, which extends the previously proposed method based on solving the non-relativistic Schrödinger equation [6]. The paper is devoted to solving two problems. The first problem describes a relativistic spinless particle in a Coulomb potential. There is an exact analytical formula for energy levels. The analysis and calculations show a good agreement between the neural network and analytical expressions and, simultaneously identifying features of solutions associated with strong fields. The second problem describes a relativistic 1D harmonic oscillator. The neural network wavefunctions and energy levels of a relativistic 1D harmonic oscillator are compared with the results calculated by the perturbation theory. We also observe a good agreement with the various initial ansatzes.

2024, **25** (4), 464–475. doi 10.26089/NumMet.v25r435

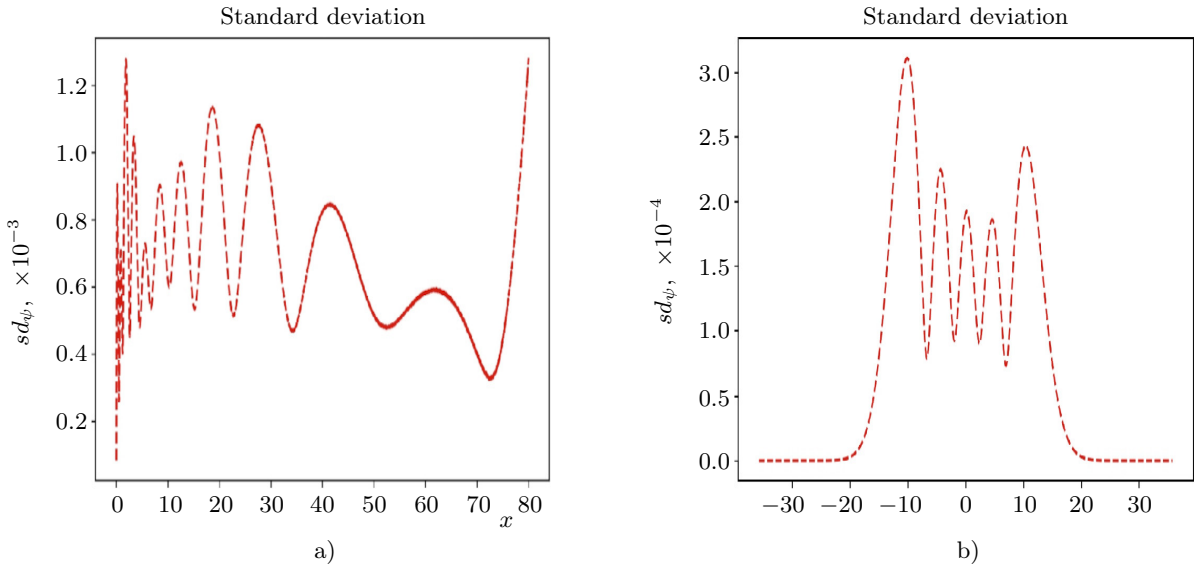

Figure 12. The standard deviation of the normalized wavefunctions for:
a) a relativistic spinless particle in a Coulomb potential with $n = 1$, $l = 0$, $Z = 50$, $N_s = 30$;
b) a relativistic particle in a harmonic potential with $n = 0$, $\omega = 0.05$, $N_s = 30$.

The method developed by us allows to find the "exact" expressions for wavefunctions and energy levels of relativistic particles. It is not necessary to use the perturbation theory, which simplifies the calculations in quantum chemistry [32] and other related fields. This method can also detect the energy level instabilities and can be applied to other types of potentials and systems. The program code is available on GitHub (https://github.com/alexkalitenko125/NNQS_for_Klein-Gordon-Fock_equation).

References

1. J. W. T. Keeble and A. Rios, "Machine Learning the Deuteron," Phys. Lett. B **809**, Article Number 135743 (2020). doi 10.1016/j.physletb.2020.135743.
2. Z. Liu, P. O. Sturm, S. Bharadwaj, et al., "Interpretable Conservation Laws as Sparse Invariants," Phys. Rev. E **109**, Article Number L023301 (2024). doi 10.1103/PhysRevE.109.L023301.
3. A. M. Kalitenko, "Phenomenological Model of a Free-Electron Laser Using Machine Learning," Phys. Scr. **98** (10), Article Number 106003 (2023). doi 10.1088/1402-4896/acf814.
4. A. M. Aleshin, V. V. Nikitin, and P. I. Pronin, "Measurement Procedure in the de Broglie–Bohm Theory," Memoirs of the Faculty of Phys. N 4, Article Number 2341511 (2023).
5. A. M. Kalitenko, "Three-Dimensional, Time-Dependent Simulation of Tapered EUV FELs with Phase Shifters," Phys. Scr. **99** (4), Article Number 045514 (2024). doi 10.1088/1402-4896/ad2f8d.
6. G. Carleo and M. Troyer, "Solving the Quantum Many-Body Problem with Artificial Neural Networks," Science **355** (6325), 602–606 (2017). doi 10.1126/science.aag2302.
7. R. G. Melko, G. Carleo, J. Carrasquilla, and J. I. Cirac, "Restricted Boltzmann Machines in Quantum Physics," Nat. Phys. **15**, 887–892 (2019). doi 10.1038/s41567-019-0545-1.
8. G. Passetti, D. Hofmann, P. Neitemeier, et al., "Can Neural Quantum States Learn Volume-Law Ground States?" Phys. Rev. Lett. **131**, Article Number 036502 (2023). doi 10.1103/PhysRevLett.131.036502.
9. M. Reh, M. Schmitt, and M. Gärttner, "Optimizing Design Choices for Neural Quantum States," Phys. Rev. B **107**, Article Number 195115 (2023). doi 10.1103/PhysRevB.107.195115.
10. D. Luo and J. Halverson, "Infinite Neural Network Quantum States: Entanglement and Training Dynamics," Mach. Learn.: Sci. Technol. **4** (2), Article Number 025038 (2023). doi 10.1088/2632-2153/ace02f.
11. Y. Zhu, Y.-D. Wu, G. Bai, et al., "Flexible Learning of Quantum States with Generative Query Neural Networks," Nat. Commun. **13**, Article Number 6222 (2022). doi 10.1038/s41467-022-33928-z.

12. J. Hermann, Z. Schätzle, and F. Noé, "Deep-Neural-Network Solution of the Electronic Schrödinger Equation," Nat. Chem. **12**, 891–897 (2020). doi 10.1038/s41557-020-0544-y.
13. J. R. Sarmiento, J. W. T. Keeble, and A. Rios, "Machine Learning the Deuteron: New Architectures and Uncertainty Quantification," Eur. Phys. J. Plus **139**, Article Number 189 (2024). doi 10.1140/epjp/s13360-024-04983-w.
14. A. Radu and C. A. Duque, "Neural Network Approaches for Solving Schrödinger Equation in Arbitrary Quantum Wells," Sci. Rep. **12**, Article Number 2535 (2022). doi 10.1038/s41598-022-06442-x.
15. J. M. Martyn, Kh. Najafi, and D. Luo, "Variational Neural-Network Ansatz for Continuum Quantum Field Theory," Phys. Rev. Lett. **131** (8), Article Number 081601 (2023). doi 10.1103/PhysRevLett.131.081601.
16. D. J. Griffiths, *Introduction to Quantum Mechanics* (Prentice Hall, Upper Saddle River, 1995).
17. J. Han, L. Zhang, and E. Weinan, "Solving Many-Electron Schrödinger Equation Using Deep Neural Networks," J. Comput. Phys. **399**, Article Number 108929 (2019). doi 10.1016/j.jcp.2019.108929.
18. W. Gordon, "Der Comptoneffekt nach der Schrödingerschen Theorie," Z. Physik **40**, 117–133 (1926). doi 10.1007/BF01390840.
19. O. Klein, "Quantentheorie und fünfdimensionale Relativitätstheorie," Z. Physik **37**, 895–906 (1926). doi 10.1007/BF01397481.
20. V. Fock, "Über die invariante Form der Wellen- und der Bewegungsgleichungen für einen geladenen Massenpunkt," Z. Physik **39**, 226–232 (1926). doi 10.1007/BF01321989.
21. D. Baye, "Klein–Gordon Equation on a Lagrange Mesh," Phys. Rev. E **109**, Article Number 045303 (2024). doi 10.1103/PhysRevE.109.045303.
22. Y. Qian, Y. Zhang, Yu. Huang, and S. Dong, "Physics-Informed Neural Networks for approximating Dynamic (Hyperbolic) PDEs of Second Order in Time: Error Analysis and Algorithms," J. Comput. Phys. **495**, Article Number 112527 (2023). doi 10.1016/j.jcp.2023.112527.
23. F. J. Dyson, *Lectures on Advanced Quantum Mechanics.* https://arxiv.org/abs/quant-ph/0608140. Cited November 3, 2024.
24. R. L. Hall and M. D. Aliyu, "Comparison Theorems for the Klein–Gordon Equation in d Dimensions," Phys. Rev. A **78**, Article Number 052115 (2008). doi 10.1103/PhysRevA.78.052115.
25. T. Shivalingaswamy and B. A. Kagali, "Rayleigh–Ritz Variational Method for Spin-Less Relativistic Particles," Int. J. Sci. Res. **3** (8), 634–638 (2014). https://www.ijsr.net/archive/v3i8/MDIwMTU0MjY=.pdf. Cited November 3, 2024.
26. V. V. Kisilev, *Quantum Mechanics* (Mosk. Tsentr Nepreryv. Matem. Obraz., Moscow, 2009) [in Russian].
27. A. N. Tikhonov, "Solution of Incorrectly Formulated Problems and the Regularization Method," Sov. Math. Dokl. **4** (4), 1035–1038 (1963).
28. A. N. Tikhonov, A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola, *Numerical Methods for the Solution of Ill-Posed Problems* (Kluwer, Dordrecht, 1995).
29. I. Loshchilov and F. Hutter, "Decoupled Weight Decay Regularization," https://arxiv.org/pdf/1711.05101. Cited November 3, 2024.
30. A. S. Davydov (Ed.), *Quantum Mechanics (Second Edition), pp. 185–206* (Pergamon, New York, 1976; Nauka, Moscow, 1973). doi 10.1016/B978-0-08-020438-3.50013-6. doi 10.1016/B978-0-08-020438-3.50003-3.
31. S. Flügge, *Practical Quantum Mechanics* (Springer, Berlin, 1994).
32. X. Chang, E. O. Dobrolyubov, and S. V. Krasnoshchekov, "Fundamental Studies of Vibrational Resonance Phenomena by Multivalued Resummation of the Divergent Rayleigh–Schrödinger Perturbation Theory Series: Deciphering Polyad Structures of Three $H_2^{16}O$ Isotopologues," Phys. Chem. Chem. Phys. **24** (11), 6655–6675 (2022). doi 10.1039/D1CP04279C.

Received
September 17, 2024

Accepted for publication
November 2, 2024

Information about the authors

Aleksandr M. Kalitenko — Ph.D. student, Lomonosov Moscow State University, Faculty of Physics, Leninskie Gory, 1, building 2, 119991, Moscow, Russia.

Petr I. Pronin — Ph.D., Associate Professor; Lomonosov Moscow State University, Faculty of Physics, Leninskie Gory, 1, building 2, 119991, Moscow, Russia.