
Approximating Ground State Energies and Wave Functions of Physical Systems with Neural Networks

Cesar Lema *

Department of Physics
New York University
c14393@nyu.edu

Anna Choromanska

Department of Electrical and Computer Engineering
New York University
ac5455@nyu.edu

Abstract

Quantum theory has been remarkably successful in providing an understanding of physical systems at foundational scales. Solving the Schrödinger equation provides full knowledge of all dynamical quantities of the physical system. However closed form solutions to this equation are only available for a few systems and approximation methods are typically used to find solutions. In this paper we address the problem of solving the time independent Schrödinger equation for the ground state solution of physical systems. We propose using end-to-end deep learning approach in a variational optimization scheme for approximating the ground state energies and wave functions of these systems. A neural network realizes a universal trial wave function and is trained in an unsupervised learning framework by optimizing the expectation value of the Hamiltonian of a physical system. The proposed approach is evaluated on physical systems consisting of a particle in a box with and without a perturbation. We demonstrate that our approach obtains approximations of ground state energies and wave functions that are highly accurate, which makes it a potentially plausible candidate for solving more complex physical systems for which analytical solutions are beyond reach.

1 Introduction

In the field of quantum mechanics the state of a physical system is given by a wave function. The wave function is mathematically represented in Dirac notation as an abstract vector $|\Psi\rangle$ in a complex vector space [1, 2]. In contrast, the state of a physical system in classical mechanics is given by a point in a phase space [3]. The usefulness of the wave function is confirmed by Born's statistical interpretation [4], which has been accepted as a postulate in foundations of quantum mechanics and interprets the modulus square of a wave function as the probability of a measurement of an observable at any given time (i.e., the probability for a specific result in the measurement of an observable X at any given time t is given by the modulus square of the wave function in the eigenbasis of X i.e. $|\langle x|\Psi(t)\rangle|^2$). Observables are quantities of a physical system that can be measured and in quantum theory are mathematically represented by Hermitian operators [2]. As an example, the knowledge of the wave function allows computing the probability of finding a particle at any given position or a particle having a specific momentum.

Access to the wave function provides full knowledge of all dynamical quantities of a physical system. In quantum mechanics, dynamical variables, such as the kinetic or angular momentum of a system, are represented by a combination of position and momentum operators. Their values are computed as expectation values of the corresponding operator given a wave function. This is analogous to classical mechanics, where all dynamical quantities can be represented as a function of position and

*undergraduate student

momentum. Thus, the wave function ultimately enables to analyze the behavior of physical systems, including those that cannot be easily studied experimentally.

The energy states of a non-relativistic quantum systems [5, 6], which involve particles with speeds much less than the speed of light, are found by solving the time independent Schrödinger equation (TISE) given as

$$H|\Psi_n\rangle = E_n|\Psi_n\rangle, \tag{1}$$

where H is the Hamiltonian of a quantum system (Hamiltonian is a Hermitian operator and is equal to the sum of the kinetic energy T and potential energy V operators), and Ψ_n and E_n are the n^{th} eigen function and eigen value of H respectively. TISE highlights that the variables characterizing quantum physical systems, such as momentum, spin, energy, etc., are quantized. TISE is an eigenvalue problem for the Hamiltonian operator H and its solution gives the eigenbasis and stationary states for the energy of a system. TISE is presently the major basic tool of many branches of modern physics [7] as well as chemistry [8]. For many physical systems, such as electron configuration of atoms, molecular and strongly interacting many body systems, there are no closed form solutions to TISE [6, 8, 9]. This work offers an accurate numerical way of solving this equation and presents a deep learning based optimization approach for approximating ground state solutions to TISE.

2 Prior work

Solving TISE typically requires numerical approximations. Among approximation methods, the most popular are Hartree–Fock, configuration interaction, and coupled cluster [10] methods, conveniently reviewed in [6], as well as multideterminant quantum Monte Carlo technique [11]. The variational principle is the foundation for some commonly used methods [6] and provides a simple and flexible scheme for finding the ground state solution to TISE. The variational principle gives an upper bound for the ground state energy E_{gs} of a given (Hermitian) Hamiltonian operator H . The expectation value of the Hamiltonian $\langle H \rangle = \langle \Psi | H | \Psi \rangle$ (Eq. 2) with respect to an arbitrary wave function $|\Psi\rangle$ (from a complex Hilbert space) is the energy $\langle E \rangle$ of the system in the state $|\Psi\rangle$. By definition E_{gs} (the lowest eigenvalue of H) is the lowest possible energy of the Hamiltonian, therefore $\langle E \rangle \geq E_{\text{gs}}$. The choice of the trial wave function $|\Psi\rangle$ determines the tightness of the upper bound on E_{gs} for a given Hamiltonian and consequently its quality. Limited prior knowledge of the wave function, furthermore, affects the accuracy of this method.

Deep learning models have shown incredible aptitude in solving difficult learning problems and could be used to represent even highly complicated trial wave functions. Therefore deep learning can be perceived as a plausible candidate method for recovering the wave function that leads to tight variational bound and accurate approximation of the ground state of a physical system. Integrating artificial neural networks in a variational optimization method to solve for the ground state solution of TISE was explored only recently. For spin lattice many-body systems a restricted Boltzmann machine was used to realize the wave function ansatz in the variational Monte Carlo (VMC) method [9]. For the electronic Schrödinger equation, Deep Neural Networks (DNN) integrated into a wave function ansatz was also explored in the VMC setting [12, 13, 8]. More recently, [8] presented a deep learning wave function ansatz that achieved state-of-the-art results for solutions of the electronic Schrödinger equation (authors also conveniently summarize the existing methods for solving TISE problem in an online video [14]). For simple Hamiltonians, the radial basis function network [15] was used to realize the wave function ansatz also in the VMC method.

For arbitrary Hamiltonians in the VMC setting, the optimal wave function ansatz are not generally known and there exists a trade off between complexity of modeling the wave function and its accuracy [16, 17, 18]. The work presented here shows how to learn the wave function approximation in the end-to-end deep learning framework. We differ from the aforementioned techniques in a couple of ways: 1) we do not use the fixed expression for the wave function ansatz but instead model it directly via DNN, 2) our method is not constrained to a specific physical system, 3) our technique is simple and does not require sampling steps as opposed to VMC. Due to the high expressiveness of neural networks and the simplicity of our approach, our work has the potential to solve TISE for arbitrary and previously inaccessible systems. It also provides a general framework for solving TISE of discrete and continuous systems (i.e. solving TISE in finite and infinite dimensional Hilbert spaces).

3 Approach

We propose to approximate the ground state wave function $\Psi : \mathbb{R}^n \rightarrow \mathbb{C}$ (n is the number of particles) that constitutes the ground state solution to the TISE using a DNN. The expectation value of the given Hamiltonian (Eq. 2) with respect to the wave function approximated with DNN was used as the objective function to train DNN. Once DNN approximated the ground state wave function with high fidelity, this expectation value approached the E_{gs} of the given system. This gives us the ground state solution of TISE we are looking for. Note that using DNN allows operating in the regime where the wave function ansatz is universal and unconstrained. In the context of the variational method, the neural network serves as the trial wave function $|\Psi\rangle$. Using neural networks to represent the trial wave function allows to explore a rich set of parameterized wave functions. The approach, therefore, heavily hinges on the nature of neural networks as universal function approximators.

The Hamiltonian consists of operators which cannot be easily applied to a DNN wave function. This makes it difficult to directly compute the expectation value of the Hamiltonian using a DNN ansatz in the variational method. We propose to numerically compute the decomposition of the function approximated by the DNN onto a computational basis (that spans the corresponding Hilbert space) and then compute the expectation value of the Hamiltonian in that computational basis. The expectation value of the energy of a system (Hamiltonian) with respect to the DNN wave function can be computed using an arbitrary computational basis according to the following equation:

$$\langle H \rangle = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_{x, x'} \langle \Psi | x \rangle \langle x | H | x' \rangle \langle x' | \Psi \rangle}{\sum_x \langle \Psi | x \rangle \langle x | \Psi \rangle} = \frac{\sum_{x, x'} \Psi^*(x) H_{xx'} \Psi(x')}{\sum_x |\Psi(x)|^2}, \quad (2)$$

where x, x' denote arbitrary basis functions and the decomposition of the neural network trial wave function $|\Psi\rangle$ onto the computational basis is given by $|\Psi\rangle = \sum_x |x\rangle \langle x | \Psi \rangle$. The decomposition coefficients $\Psi(x') = \langle x' | \Psi \rangle$ are generally multidimensional integrals and were estimated using Riemann approximations (Riemann approximations rely on batch computations, i.e., they require wave function values computed over the entire domain which must be finite). The expectation value of the Hamiltonian $\langle x | H | x' \rangle$ computed for two arbitrary basis functions gives $H_{xx'}$. The computational basis in our experiments was chosen such that calculating $H_{xx'}$ was feasible analytically or computationally. In the experiments we show that despite the computational basis possibly being infinite, we can recover accurate solutions using its finite subset.

4 Experiments

DNN was trained in an unsupervised setting to approximate the ground state wave function for physical systems given below. We consider the unperturbed and perturbed particle in a box systems since they have known solutions that we can compare our results with. These systems consist of a single one dimensional particle in a real space and therefore the input of the DNN is a scalar value representing the position of the particle in its domain. The output of the network is the approximated value of the ground state wave function. The parameters of the DNN were optimized using batch gradient descent and automatic differentiation tools were used to compute gradients for backpropagation. This was implemented using Pytorch.

The first $N \in \mathbb{Z}^+$ (a hyperparameter of the framework) eigenfunctions of the Hamiltonian of the particle in a box systems were used as the computational basis. Real valued functions were used for $|\Psi\rangle$, but we emphasize that the framework described in this paper would be identical when using complex functions for $|\Psi\rangle$ instead.

4.1 A particle in a box

A particle in a box consists of a free particle constrained to a domain $x \in \{0, a\} : a \in \mathbb{R}$ by infinite potential outside its domain, i.e., infinite potential well:

$$V(x) = \begin{cases} 0 & 0 < x < a \\ \infty & \text{otherwise} \end{cases} \quad (3)$$

where a is the well width. The Hamiltonian of this system is $H = -\frac{\hbar^2}{2\mu} \frac{d^2}{dx^2}$, where μ is the particle mass and \hbar is the Planck constant. Its eigenvalues are $E_n = \frac{n^2 \pi^2 \hbar^2}{2\mu a^2}$ and its eigenbasis is:

$$b_n = \begin{cases} \sqrt{\frac{2}{a}} \sin\left(\frac{n\pi}{a}x\right) & 0 < x < a \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

where $n \in \mathbb{Z}^+$ indicates the n^{th} stationary state. We used the first N eigenfunctions of this system as the computational basis and computed matrix H_{nm} (m also indicates an arbitrary stationary state) as follows: $H_{nm} = \delta_{nm}E_n = \delta_{nm} \frac{n^2\pi^2\hbar^2}{2\mu a^2}$. Thus finally, the objective function in this experiment is:

$$\langle H \rangle = \frac{\sum_n |\langle b_n | \Psi \rangle|^2 H_{nn}}{\sum_n |\langle b_n | \Psi \rangle|^2} \quad (5)$$

4.2 A particle in a box with a linear perturbation

For this experiment a linear potential $V(x) = \alpha x$, where $\alpha \in \mathbb{R}$, was added to the infinite square well potential of the particle in a box system. The Hamiltonian of this system is $H = H^0 + H' = -\frac{\hbar^2}{2\mu} \frac{d^2}{dx^2} + \alpha x$, where H^0 is the Hamiltonian of the unperturbed particle in the box system and H' is the perturbation. The objective function in this experiment was the expectation of the perturbed Hamiltonian given by

$$\langle H \rangle = \langle H^0 \rangle + \langle H' \rangle, \quad (6)$$

due to the linearity of the expectation (this property was also used to compute the matrix H_{nm} for this experiment). The expectation value of the perturbation is given by:

$$\langle H' \rangle = \alpha \langle x \rangle = \alpha \frac{\sum_n |\langle b_n | \Psi \rangle|^2 x_{loc}(b_n)}{\sum_n |\langle b_n | \Psi \rangle|^2}, \quad x_{loc}(b_n) = \sum_m x_{nm} \frac{\langle b_m | \Psi \rangle}{\langle b_n | \Psi \rangle} \quad (7)$$

The expectation value of the position operator in the computational basis x_{nm} was computed analytically:

$$x_{nm} = \begin{cases} \frac{a}{2} & n = m \\ \frac{a}{\pi^2} \left[\frac{\cos((n-m)\pi) - 1}{(n-m)^2} - \frac{\cos((n+m)\pi) - 1}{(n+m)^2} \right] & n \neq m \end{cases} \quad (8)$$

4.3 Results

In our experiment for the unperturbed system (well width $a=1$) we used a DNN with 1 hidden layer of width 1000 and ReLU activation functions. For the experiments with perturbed systems (first considered system, system A, had well width $a = 1$ and perturbation constant $\alpha = 8$ and the second one, system B, had well width $a = 10$ and perturbation constant $\alpha = 2$) we used DNNs with 2 hidden layers of width 500 and 100 and ReLU activation functions. For all experiments presented $N = 100$ basis functions were used. The normalized ground state wave function found by our neural networks is compared to the exact wave function in Figure 1. In Table 1 the ground state energies (in natural units) of each system are compared to the exact values as well as the values obtained using the VMC method of [15]. The exact values were obtained by solving TISE using a symbolic differential equation solver, i.e., Mathematica.

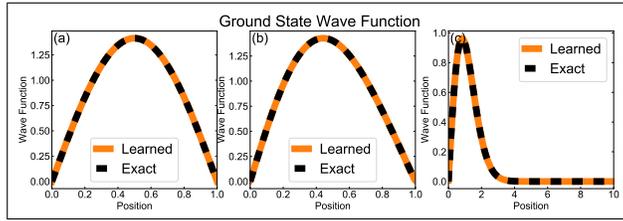


Figure 1: **Normalized ground state wave functions.** Results (solid line) for the unperturbed system (a), perturbed system A (b), and perturbed system B (c) are compared to the exact results (dashed line).

Note that the ground state wave functions obtained with our DNN model are practically identical to the analytically obtained target wave functions (Figure 1). Moreover, as can be observed in Table 1, the proposed approach finds highly accurate approximations of the ground state solutions, which are also comparable or better than the approximations obtained with the VMC method of [15]. We also accurately approximated system (perturbed system B) that corresponds to a more complicated wave function that was not analyzed in [15]. The obtained preliminary experiments confirm the plausibility of using DNNs to represent the trial wave function and effectiveness of the proposed variational method. Such approach can be especially useful in case of systems where there exist a limited prior knowledge to motivate a specific choice of the trial wave function. In future works we will generalize our approach to approximate the solution of TISE for more complex physical systems [8, 12].

Table 1: Ground state energies

System	$E_{\text{ground state}}$		
	Computed	VMC [15]	Exact
Unperturbed	4.93484	4.9348	4.93480
Perturbed A	8.79510	8.7960	8.79507
Perturbed B	2.94583	NA	2.94583

5 Conclusion

In this manuscript a new framework for approximating the ground state solution of the physical system described by a given Hamiltonian is presented. It is based on using a deep neural network to represent a trial wave function in a variational optimization scheme via an end-to-end deep learning framework. The DNN is trained by minimizing the expectation value of the Hamiltonian. Initial experimental results on particle in a box physical systems are presented and confirm the validity of the proposed approach.

References

- [1] David J. Griffiths and Darrell F. Schroeter. *Introduction to Quantum Mechanics*. Cambridge University Press, 3 edition, 2018.
- [2] P. A. M. Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, 4 edition, 1967.
- [3] John R. Taylor. *Classical Mechanics*. University Science Book, 2005.
- [4] Max Born. Zur quantenmechanik der stoßvorgänge. *Zeitschrift für Physik*, 37(12):863–867, 1926.
- [5] L.D. Landau and E.M. Lifshitz. *The Classical Theory of Fields*. Butterworth-Heinemann, 4 edition, 1980.
- [6] Lucjan Piela. *Ideas of Quantum Chemistry*. Elsevier, 3 edition, 2020.
- [7] Franck Laloë. *Do We Really Understand Quantum Mechanics*. Cambridge University Press, 2 edition, 2019.
- [8] Jan Hermann, Zeno Schätzle, and Frank Noé. Deep neural network solution of the electronic schrödinger equation, 2020.
- [9] Giuseppe Carleo and Troyer Matthias. Solving the quantum many-body problem with artificial neural networks. *Science*, 355(602), 2017.
- [10] Fritz Coester. Bound states of a many-particle system. *Nuclear Physics*, 7:421–424, 1958.
- [11] M. A. Morales, J. McMinis, B. K. Clark, J. Kim, and G. E. Scuseria. Multideterminant wave functions in quantum monte carlo. *J. Chem. Theory Comput.*, 8(7), 2012.
- [12] Jiequn Han, Linfeng Zhang, and Weinan E. Solving many-electron schrödinger equation using deep neural networks. *Journal of Computational Physics*, 399(108929), 2019.
- [13] David Pfau, James S. Spencer, Alexander G. D. G. Matthews, and W. M. C. Foulkes. Ab initio solution of the many-electron schrödinger equation with deep neural networks. *Phys. Rev. Research*, 2(033429), 2020.
- [14] Deep neural network solution of the electronic Schrödinger equation. https://www.youtube.com/watch?v=_bdvpmleAgw&feature=youtu.be.
- [15] Teng Peiyuan. Machine learning quantum mechanics: solving quantum mechanics problems using radial basis function network. *Phys. Rev. E*, 98(033305), 2018.
- [16] Jeongnim Kim et al. QMCPACK: an open source ab initio quantum monte carlo package for the electronic structure of atoms, molecules and solids. *Journal of Physics: Condensed Matter*, 30(19), 2018.
- [17] Tao Pang. *An Introduction to Quantum Monte Carlo Methods*. Morgan Claypool Publishers, 2016.
- [18] Federico Becca and Sandro Sorella. *Quantum Monte Carlo Approaches for Correlated Systems*. Cambridge University Press, 2017.

Broader Impacts

The authors of this paper believe that this work helps to bridge the gap between two fields: machine learning and physics, and contributes to the popularization of machine learning tools beyond computer science community. The authors of this work commit to the following:

- They will organize a research camp for undergraduate students of machine learning and physics to popularize the outcomes of the presented research and further nurture cross-disciplinary education at early stages (one of the authors of this paper is an undergraduate student). The authors will also organize a similar research camp targeting graduate students.
- The authors will ensure the participation of students from underrepresented and minority groups and disadvantaged backgrounds in the above mentioned camps. Similarly the authors will promote their camps among racially and ethnically diverse communities and ensure their participation. The authors will have a support of their home institutions in conducting this outreach.
- The elements of the proposed work will be included in the K12 program run by one of the authors. The program aims at offering university-level training to high-school students.
- The authors will organize a Workshop on Machine Learning for Physics in their home institution, with the participation of industry and academia, to further facilitate the popularization of cross-disciplinary research in machine learning and physics.

The authors hope that the above described outreach activities will help to strengthen the field of computational physics and will carry high societal impact by propagating the knowledge of machine learning, deep learning, and physics to the next generation of engineers and researchers.