Learning Potentials of Quantum Systems using Deep Neural Networks

Arijit Sehanobish,[*1] **Hector H. Corzo,**[*2] **Onur Kara,**[3] **David van Dijk**[1]

[1] Internal Medicine (Cardiology) and Computer Science, Yale University
{arijit.sehanobish, david.vandijk}@yale.edu
[2] Center for Chemical Computation and Theory
University of California, Merced
hhcorzo@gmail.com
[3] Hindsight Technology Solutions
okara83@gmail.com

Abstract

Attempts to apply Neural Networks (NN) to a wide range of research problems have been ubiquitous and plentiful in recent literature. Particularly, the use of deep NNs for understanding complex physical and chemical phenomena has opened a new niche of science where the analysis tools from Machine Learning (ML) are combined with the computational concepts of the natural sciences. Reports from this unification of ML have presented evidence that NNs can learn classical Hamiltonian mechanics. This application of NNs to classical physics and its results motivate the following question: Can NNs be endowed with inductive biases through observation as means to provide insights into quantum phenomena? In this work, this question is addressed by investigating possible approximations for reconstructing the Hamiltonian of a quantum system in an unsupervised manner by using only limited information obtained from the system's probability distribution.

1 Introduction

In the Machine Learning (ML) realm, Neural Networks (NNs) are among the most used and exceptionally efficient models to learn and generalize information from data. These data interpretative capabilities have provoked the widespread use of NNs in Natural Language Processing (Torfi et al. 2020), Image Classification (Kolesnikov et al. 2019), Video Captioning (Sun et al. 2019) and Reinforcement Learning (Du and Narasimhan 2019; Higgins et al. 2016); furthermore, recent works have shown the capabilities of NNs in symbolic reasoning and mathematical problem solving (Lample and Charton 2019). In the case of natural sciences, applying ML to physics is not new, several works have been reported (Toth et al. 2019; Greydanus, Dzamba, and Yosinski 2019; Cranmer et al. 2020; Tong et al. 2020) where different authors have combined ML with Hamilton's equations of motion to generate trajectories that obey energy conservation principles and classical physical laws. In material sciences, on the other hand, ML has proven to be an important interpretative tool for the computational prediction of new materials (Schleder et al. 2019). The encouraging results reported by the different applications of ML have motivated the use of NNs as powerful tools to gain insight into the laws of physics that govern the behavior of complicated natural classical phenomena. Unlike classical physics, in quantum physics, objects have characteristics of both particles and waves (wave–particle duality) for which the concept of trajectory is no longer defined nor can their position and momentum, both, be measured simultaneously (Sakurai and Commins 1995; Robinett 1997; Feynman, Leighton, and Sands 1965; Robinett and Robinett 2006). Because of this wave–particle duality, the state of a quantum mechanical system is fully specified by its wave–function, which is typically obtained by solving the Schrödinger equation (Sakurai and Commins 1995; Robinett 1997; Feynman, Leighton, and Sands 1965; Robinett and Robinett 2006). In many cases, however, not only solving this equation is difficult, but also its correct formulation requires knowledge about the form of the potential energy operator, which often may not be completely known. In contrast, the inverse form of the Schrödinger equation (Nakatsuji 2002; Chadan and Sabatier 2012; Zakhariev and Suzko 2012; Jensen and Wasserman 2018) presents an alternative for describing quantum phenomena by reformulating the description of quantum mechanical systems as solutions of inverse problems (Aster, Borchers, and Thurber 2018; Groetsch and Groetsch 1993; Vogel 2002). Inverse problems are central to the study of quantum mechanical systems, insomuch that much of what is known about the electronic structure of matter has been mathematically characterized by solutions of inverse problems (Beals and Greiner 2009; Zakhariev and Suzko 2012; Jensen and Wasserman 2018; Vogel 2002). Thus, numerical algorithms for the inversion of the Schrödinger equation are important predictive tools for the further development of approximate quantum mechanical methodologies such as scattering approximated models (Zakhariev and Suzko 2012), Density Functional Theory (DFT) (Jensen and Wasserman 2018), etc.

In this work, rather than handcrafting numerical solutions for the inverse Schrödinger equation (Aster, Borchers, and Thurber 2018; Jensen and Wasserman 2018; Vogel 2002; Beals and Greiner 2009) to define a potential function and describe a quantum phenomena, a neural network, termed Quantum Potential Neural Network (QPNN), is instead designed to learn potential functions directly from observables

[*]Equal Contribution

in an unsupervised manner. This proposed QPNN for learning potential functions was developed based on the underlying formalism for the inverse solution of the Schrödinger equation. Thus, the proposed QPNN opens the possibility for generating simpler and succinct functions that can be used to construct effective Hamiltonians for the description of a variety of quantum systems using only a small portion of the available information known about the system. Since these effective Hamiltonians can be generalized to obtain other observables, QPNN may provide unique insights into complex quantum phenomena were only a small amount of information is available.

2 Theory

The mathematical description of a quantum particle typically takes the form of a complex function of spatial coordinates $\vec{x}$ and time coordinates t called wave–function, $\Psi(\vec{x},t)$ (Sakurai and Commins 1995; Robinett 1997; Feynman, Leighton, and Sands 1965; Robinett and Robinett 2006). $\Psi(\vec{x},t)$ is a complex–valued probability amplitude whose square modulus ($|\Psi(\vec{x},t)|^2$) correspond to the probability of finding the particle described by the wave–function at that given $\vec{x}$ and t. The classically measured value of a physical observable, however, is not given directly by $\Psi(\vec{x},t)$ but by the expectation values of the operators that represent the desired measurement acting on $\Psi(\vec{x},t)$. In many scenarios, wave–functions are obtained as direct solutions of the time–dependent Schrödinger equation,

$$i\hbar\frac{\partial\Psi(\vec{x},t)}{\partial t} = \hat{\mathrm{H}}\Psi(\vec{x},t), \tag{1}$$

where $\hbar$ is Planck's constant and $\hat{\mathrm{H}}$ is the Hamiltonian operator of the system, which is an Hermitian operator acting on an infinite dimensional space of L^2 functions. Thus, $\hat{\mathrm{H}}$ needs not be compact and as much may not have any eigenvalues. When $\hat{\mathrm{H}}$ is time–independent, equation 1 can be reduced to the following equation

$$\hat{\mathrm{H}}\psi_n(\vec{x}) = E\psi_n(\vec{x}), \tag{2}$$

where n indicates the quantum state of the system. In many cases, the physical information contained in the time–independent wave–function $\psi_n(\vec{x})$ may be enough for the characterization of the system under study.

2.1 Hamiltonian

The Hamiltonian operator, $\hat{\mathrm{H}}$, is fundamental in many formulations of quantum theory. This operator is expressed as the sum of the kinetic ($\hat{\mathrm{T}}$) and potential energy operators ($\hat{\mathrm{V}}$) for all particles in the quantum system,

$$\hat{\mathrm{H}} = \hat{\mathrm{T}} + \hat{\mathrm{V}}. \tag{3}$$

Generally, the kinetic energy operator contained in $\hat{\mathrm{H}}$ only depends on the second derivatives of the wave–function, with respect to its spatial coordinates. The potential energy operator, however, depends on the physical circumstances imposed onto the system, and varies from system to system. Thus equation 3 may be expressed as

$$\hat{\mathrm{H}} = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial\vec{x}^2} + \hat{\mathrm{V}}(\vec{x},t) \equiv -\frac{\hbar^2}{2m}\nabla^2_{\vec{x}} + \hat{\mathrm{V}}(\vec{x},t). \tag{4}$$

In quantum mechanics, the problem of finding the $\hat{\mathrm{H}}$ that characterizes a given phenomenon could be reduced to formulating the potential operator that contains all the governing physical descriptors.

2.2 Predicting potentials

The usual method for describing systems in quantum mechanics is by obtaining the wave–function of the system as a solution of the Schrödinger equation. These wave–functions strongly depend on the Hamiltonian, and in particular, the definition of the potential used to describe the system. However, one could also describe a quantum phenomena through the solution of the inverse problem, i.e. by finding an effective potential or function that contains all the important physical constrains that generated the observed outcomes. Inverse problems like this one are common in quantum mechanics and electronic structure, for example, the field of DFT (Jensen and Wasserman 2018; Parr and Yang 1995; Burke, Werschnik, and Gross 2005) has, at its core, this type of inverse problems. As discussed before, solutions for the full Schrödinger's equation, and thus wave–functions, are difficult to obtain except for some simple models. The probability density $|\psi(\vec{x})|^2$, on the other hand, may be experimentally inferred for several quantum systems. Thus, an approximated wave–function, $|\psi| = \sqrt{|\psi(\vec{x})|^2}$, may be defined for the construction of the effective potential.

3 Quantum Potential Neural Networks

In this section the proposed NN and the loss function that were used to compute the effective potentials of various quantum systems are described. This section is divided into two parts: (**i**) Time–independent systems and (**ii**) Time–dependent systems

3.1 Time–independent Systems

In order to obtain the effective potential, a new parametric function U_θ is learned in an unsupervised manner. This parametric function corresponds to the effective potential of the quantum system and was obtained by implementing a loss function that obeys the time–independent Schrödinger (TISE) equation (Eq. 2),

$$L_{TISE}(\theta) = \left|\left| D\left(-\frac{\hbar^2}{2m}\frac{\Delta_{\vec{x}}|\psi|}{|\psi|} + U_\theta(\vec{x})\right)\right|\right|_2^2, \tag{5}$$

where D is the total derivative operator acting on multivariate function $-\frac{\hbar^2}{2m}\frac{\Delta_{\vec{x}}|\psi|}{|\psi|} + U_\theta(\vec{x})$ and $||\cdot||_2$ is the Frobenius norm. Because of the definition of this loss function, energy conservation is effectively demanded for time–independent systems. Since the U_θ function is given by a differential equation, an initial condition was added to ensure that a unique function is learned. The initial conditions used for the different quantum systems are based on the inherent nature of each of the systems, more information and explanations about some of these conditions can be found in the literature (Romanowski 2007). Finally, the loss function after the consideration of the initial condition reads

$$L(\theta) = L_{TISE}(\theta) + (U_\theta(\vec{x}) - y)^2 \tag{6}$$

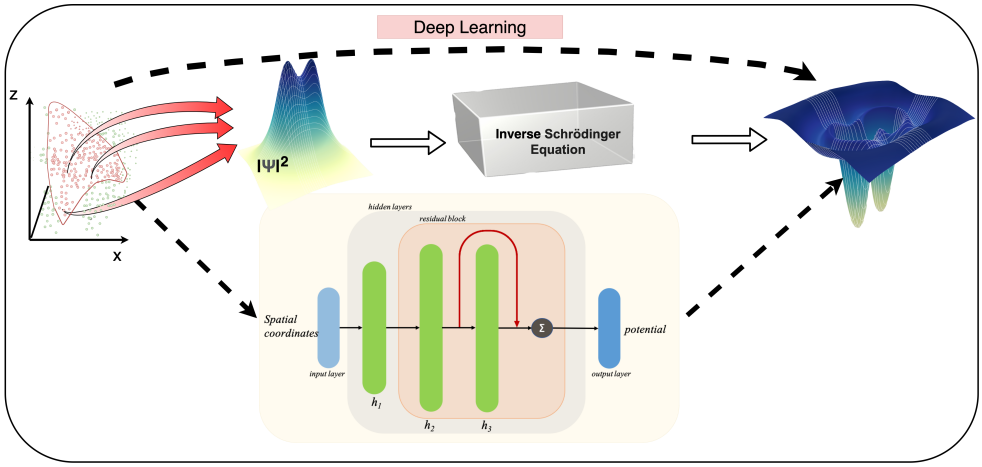

Figure 1: Graphical representation of our framework. The input coordinates (left) are fed into our model QPNN (bottom) which is trained by constraining the network to obey the Schrödinger's equation.

where $\vec{x}$ is some point in the domain of the function and y is the expected ground truth value of the true potential at that point.
The main observation here is that using $|\psi|$ (instead of ψ) to solve for the potential leads to the correct potential except, possibly, at finitely many points where ψ changes signs. However, this does not create any difficulties for training the proposed model.

3.2 Time–dependent Systems

For time–dependent systems, the formulation for the time–dependent Schrödinger equation (TDSE) loss reads

$$L_{TDSE}(\theta) = \left|\left| Re\left(\frac{i\frac{\partial\psi}{\partial t} + \frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial\vec{x}^2}}{\psi} \right) - U_\theta \right|\right|_2^2. \quad (7)$$

It is important to mention that complex numbers may be more common to appear in the time–dependent solution of the Schrödinger equation; however, for the current study, the probability density of the considered systems are described with an Hermitian Hamiltonian, and thus only real observables were considered to avoid the handling of complex values. For the time–dependent results presented in this report, the QPNN was trained with the full wave–function instead of just the probability density. In a future work, the density–to–potential results for time–dependent systems will be explored and discussed in detail.

3.3 Model Architecture

For the construction of the NN, a 4 layer feedforward network with hidden sizes of $32, 128$ and 128 with a residual connection between second and third layers was used. The residual layers help in faster training, stable gradients and also results in a smooth loss landscape (Li et al. 2017). The inputs to the model are the $\vec{x}$ spatial coordinates for time independent systems and $(\vec{x}, t)$ for time dependent systems. For the network training, 3000 of these coordinates were randomly selected from the domain of definition of each particular system. The model was trained for 500 epochs with Adam optimizer (Kingma and Ba 2017).

4 Related Work

The use of deep learning for understating physical phenomena has been an active field of development. In particular, there is a considerable amount of literature where authors have endowed neural networks with classical Hamiltonian mechanics (Toth et al. 2019; Greydanus, Dzamba, and Yosinski 2019; Tong et al. 2020; Iten et al. 2020; Bondesan and Lamacraft 2019; Zhong, Dey, and Chakraborty 2019; Chmiela et al. 2017); conservation of energy and irreversibility in time are the key features of such networks. There are recent reports extending these results in cases of damped pendula, i.e., systems where there is dissipation of energy (Zhong, Dey, and Chakraborty 2020). In computational quantum mechanics, deep neural networks have been implemented to learn representations and extract the necessary features to predict desired properties from raw unprocessed data (Goh, Hodas, and Vishnu 2017). Recently, two methods for estimating the density matrix for a quantum system, the Quantum Maximum Likelihood (QML) and Quantum Variational Inference (QVI) method, were introduced (Cranmer, Golkar, and Pappadopulo 2019). For these methods, the authors used a flow based method (Toth et al. 2019; Jimenez Rezende and Mohamed 2015) to increase the expressivity of their variational family of density matrices. The applicability of these methods, however, has been only validated for the harmonic and anharmonic quantum oscillator models. Application of deep learning to quantum mechanics is still in its early stages (Torfi et al. 2020; Raissi, Perdikaris, and Karniadakis 2017a,b, 2019; Jasinski et al. 2020; Carleo et al. 2019; Amabilino et al. 2019; Unke and Markus 2019; Schmitz, Godtliebsen, and Christiansen 2019; Schmidt et al. 2017; Hibat-Allah et al. 2020; Nakajima, Tanaka, and Hashimoto 2020; Pu, Li, and Chen 2020; Mills, Spanner, and Tamblyn 2017; Manzhos 2020). Most of the deep learning quantum mechanic frameworks introduced so

Table 1: A quantitative analysis for the QPNN

System	RMSE between True and Learned Potentials (QPNN)	RMSE between True and Learned Potentials (using RK4)	RMSE between True and Learned Energies (QPNN)	RMSE between True and Learned Energies (RK4)
Harmonic Oscillator	$1 \times 10^{-2} \pm 5 \times 10^{-3}$	$9 \times 10^{-3} \pm 4 \times 10^{-4}$	$1 \times 10^{-2} \pm 2 \times 10^{-3}$	$5 \times 10^{-2} \pm 7 \times 10^{-3}$
Pöschl–Teller potential	$1 \times 10^{-4} \pm 6 \times 10^{-5}$	$2 \times 10^{-4} \pm 3 \times 10^{-5}$	$8 \times 10^{-4} \pm 6 \times 10^{-5}$	$7 \times 10^{-3} \pm 8 \times 10^{-4}$
H_2 molecule	$2 \times 10^{-3} \pm 4 \times 10^{-4}$	$3 \times 10^{-3} \pm 2 \times 10^{-4}$	$9 \times 10^{-3} \pm 7 \times 10^{-4}$	$4 \times 10^{-3} \pm 2 \times 10^{-4}$
Soliton	$3 \times 10^{-2} \pm 4 \times 10^{-3}$	-	-	-

far are focused on either solving the Schrödinger equation or predicting the trends of specific observables such as the system's energy. Concerning inverse problems, Raissi et al, introduced the physics–informed neural network for solving forward and inverse problems involving nonlinear partial differential equations (Raissi, Perdikaris, and Karniadakis 2019). Although the impressive results reported in this work, in terms of inverse problem solutions, the deep learning framework reported by Raissi et al is focused only on the solution of a partial differential equation for the prediction of pressure profiles in an classical system. On the other hand, in quantum mechanics observables may be inferred when a valid effective potential is known for a given quantum system; thus, the solution of the density–to–potential inversion problem to predict effective potential functions (Jensen and Wasserman 2018) play an important role in the understanding of the quantum phenomena, and in particular in the elucidation of the electronic structure of molecules from a density functional theory perspective. In this regard, Nagai and coworkers have proposed the Neural–network Kohn–Sham exchange–correlation potential (Nagai et al. 2018), which propose a supervised training scheme that uses information from well defined potentials and probability densities to train a NN. However, to the best of our knowledge, there are no reported works that use deep learning to solve the density–to–potential *inverse problem* to systematically estimate potentials from observations in a completely *unsupervised* manner.

5 Experiments

The performance of the proposed Quantum Potential Neural Network is validated on four different quantum systems, one of these systems describes the temporal evolution of a quantum wave whereas the other three are examples of time–independent systems. Among the time–independent systems, exact analytical solutions for the time–independent Schrödinger equation can be obtained only for the harmonic oscillator and the Pöschl–Teller (PT) potentials whereas for the H_2 molecule, only approximate solutions are attained. Details about the solutions, physical implications and interpretations of these systems can be found in any standard book on quantum mechanics (Sakurai and Commins 1995; Robinett 1997; Pronchik and Williams 2003). Finally, the potentials learned by the QPNN were used to compute the total energy of each of the systems. The quantitative results for all the reported experiments are summarized in table 1. In order to compare the results obtained (in the form of solutions to a differential equation) via the NN techniques to those obtained through well established approaches, the differential equations were solved numerically using the well–known and standard Runge–Kutta 4th Order (or RK4) integrator implemented in the standard python libraries. The RK4 algorithm provides means of solving various kinds of differential equations and is generally considered as a robust workhorse to bench mark new computational techniques (Landau, Paez, and Bordeianu 2015). The differences in accuracy of the values obtained by both the proposed QPNN method and the RK4 numerical integrator were quantified through their root mean square error (RMSE) values. The code is available at https://github.com/arijitthegame/Quantum-Hamiltonians.

5.1 Density–to–Potential Experiments

For the density–to–potential experiments, the exact wave–functions, $\psi(\vec{x})$, for the quantum Harmonic oscillator and the PT potential were obtained by solving the time–independent Schrödinger equation. These wave–functions were later used to define the probability distribution, $|\psi(\vec{x})|^2$, for each of the systems. In the case of the Hydrogen molecule, the probability density was defined according to the one–electron $1s$ orbital function that delineate the approximated electronic density for the Hydrogen molecule in a Born–Oppenheimer approximation. These probability densities were used to define, for each of the quantum systems, an approximated probability amplitude function, $|\psi| = \sqrt{|\psi(\vec{x})|^2}$. The effective potential function for each of the systems was obtained by training the QPNN with information provided by randomly selected coordinates evaluated onto the approximated probability amplitudes and into the loss function.

Quantum Harmonic Oscillator (QHO): The motion of the the one–dimension QHO is, perhaps, the simplest quantum mechanical system whose motion follows a linear differential equation with constant coefficients. In the QPNN framework, for the prediction of the QHO potential, the coordinate variable $\vec{x}$, was randomly sampled from $[-5, 5]$, and was used as the input to the model. In the analytical solution of the time–independent Schrödinger equation, the wave–functions for the different states of the QHO are given by Hermite polynomials $H_n, n = 0, 1, \cdots$, whereas the energies corresponding to these states depend on the force constant w and are given by $E_n = \hbar w(n + 1/2)$. The analytical wave–functions for the different states of the QHO defined the probability densities used to train the QPNN. In this case, the initial condition imposed is the fact that at the zero point of the reference coordinates, all the energy in the system is kinetic, and thus the potential energy at this point is zero, i.e. $V(0) = 0$. Fig 2 shows the used probability density, the

learned potential and the energy computed by the QPNN.

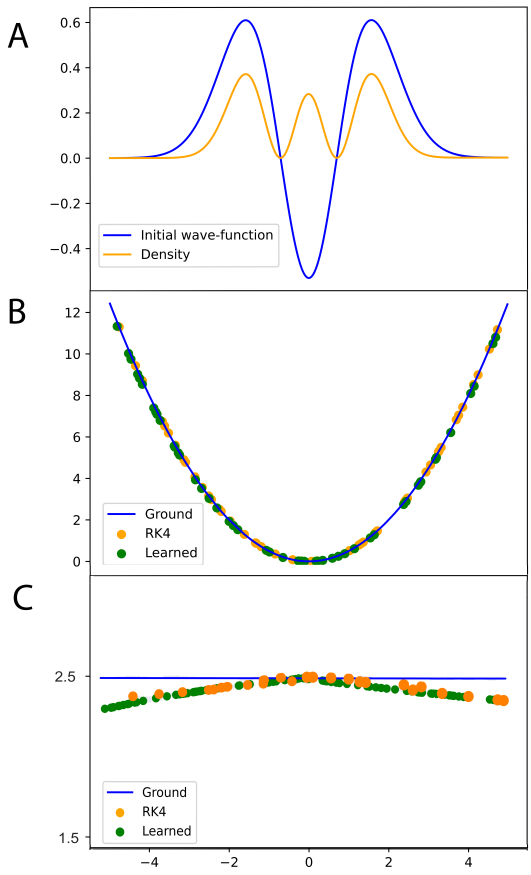

Figure 2: Harmonic Oscillator system. (A) Wave–function and probability Density used in the QPNN, (B) Ground and Learned Potential, (C) Total Energy.

Pöschl–Teller potential: The PT potential is a special class of anharmonic potential for which the one–dimensional Schrödinger equation can be solved in terms of special functions. This potential may be used to model vibrational molecular potentials with out–of–plane bending vibrations (Senn 1986; Jia, Zhang, and Peng 2017) and observables of diatomic potentials (Pronchik and Williams 2003). For the PT potential, the wave–functions used to define the approximated probability amplitude function are the Legendre functions P_λ^μ (Riley 1974) with energy eigenvalues E_μ and potential depth V_0 (Hernández de la Peña 2018),

$$\left\{ P_\lambda^\mu(\tanh x) \;\middle|\; E_\mu = \frac{-\mu^2}{2}, V_0 = \frac{-\lambda(\lambda+1)}{2}, \; \begin{matrix} \lambda=1,2,\cdots \\ \mu=1,2,\cdots,\lambda \end{matrix} \right\}. \tag{8}$$

Details about several suitable boundary terms and initial conditions for this type of potentials are formulated and reviewed in the literature (Agboola 2010; Hernández de la Peña 2018). The input for this experiment is the spatial coordinate $\overrightarrow{x}$ randomly sampled from $[-3, 3]$. For this experiment, the wave–function defined by the Legendre function with $\lambda = 2$ and $\mu = 1$ was employed. From the density, the initial wave–function takes the form:

$$\psi_2^1(x) = |\tanh x| \text{ sech } x. \tag{9}$$

Figure 3 shows the probability density used to train the system, as well as the learned potential and energy computed for the system.

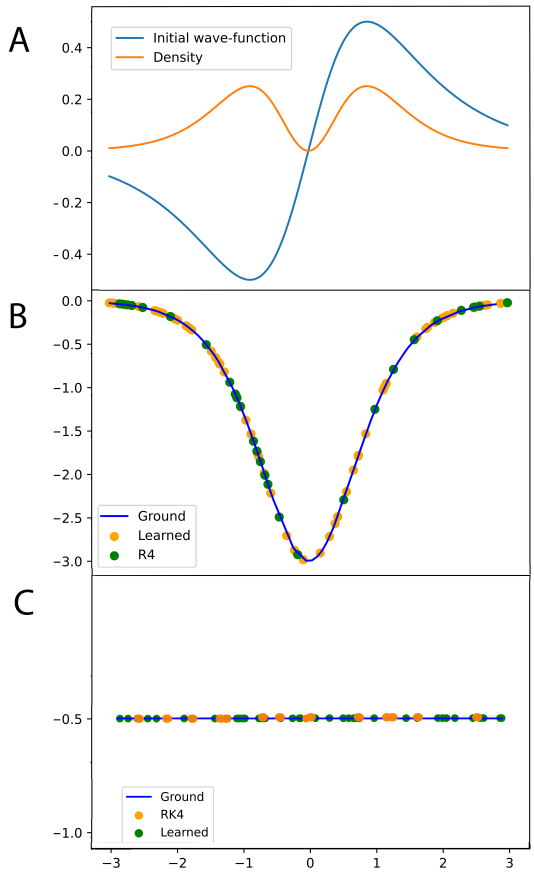

Figure 3: Pöschl–Teller system. (A) Wave–function and Density, (B) Ground and Learned Potential, (C) Total Energy.

Hydrogen molecule: The H_2 molecule is the first multi-electronic system with approximated probability density considered. For the training of the QPNN, an ab initio electron density for the H_2 molecule with an equilibrium bond length (x_{r_e}) of 1.346Å and total energy of -0.958470046928 a.u. was approximated by using a fast and systematic self–consistent field method (Helgaker, Jorgensen, and Olsen 2014). This density was computed using three Gaussian primitive functions for each H atom, where the $\vec{x}$ coordinate defined on $[-3, 3]$ was chosen as the reference internal coordinate. The initial conditions for the system were defined following the same lines as in (Rafi et al. 1995); specifically, the fact that $V(x_{r_e}) = 0$. Figure 4 shows the probability density used to train the QPNN as well as the learned potential and the energy computed using this potential.

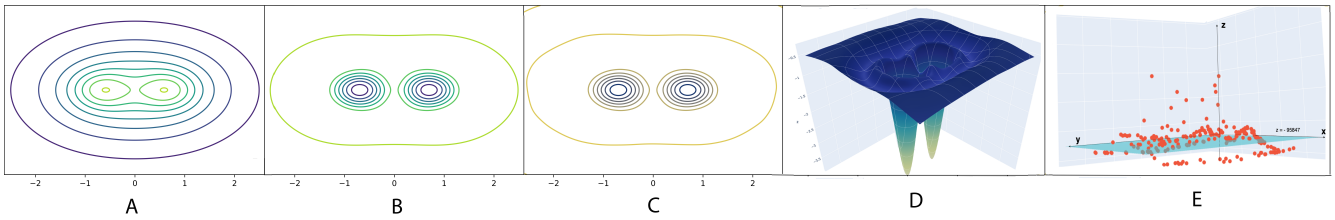

Figure 4: H_2 molecule system: (A) Density, (B) Ground Truth Potential, (C,D) Learned Potential, (E) Computed Energy.

5.2 Exploration of a Time–Dependent System

In order to explore the behavior of the QPNN in systems with dependence on time, the effective potential for a soliton model was computed. Solitons represent solitary waves propagating without any temporal evolution in shape or size when viewed in the reference frame moving with the group velocity of the waves (Wazwaz 2009). This type of solitary waves are particularly important in the Bose–Einstein condensation theory and arise in many contexts such as the elevation of the surface of water and the intensity of light in optical fibers. Solitons form a special class of solutions of model equations, including the Korteweg de–Vries (KdV) and the Nonlinear Schrödinger (NLS) equations. In this particular experiment, the one–dimensional soliton satisfies the following differential equation:

$$i\frac{\partial\psi}{\partial t} + \frac{\partial^2\psi}{\partial x^2} + U(x,t)\psi = 0. \tag{10}$$

Thus, for this system, the loss function used to train the QPNN is given by equation 7 where $\psi = 2\text{sech}(\sqrt{2}(x - 2t))e^{i(x+t)}$ and $U(x,t)$ is $|\psi|^2$. For this experiment, the coordinates for the QPNN input were defined on $\vec{x}, \vec{t} \in [0,1]$. Figure 5 shows the potential obtained by the QPNN for this system.

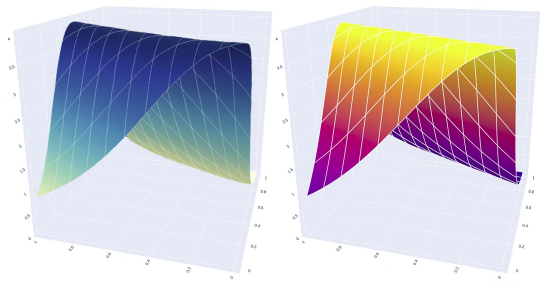

Figure 5: Left: Soliton Ground Truth Potential, Right: Soliton Learned Potential.

6 Discussion

Table 1 reports the RMSE values for the different models studied in this work. In the case of the time–independent models, for the QPNN, the harmonic oscillator presents the largest RMSE value between the true and learned potential whereas the PT potential has the lowest RMSE among all the time–independent systems. For all the systems but the harmoninc oscillator, the RMSE values for the learned potentials are comparable in magnitude with those obtained with the RK4 method. In the case of the energies, when compared against the exact energy, the RMSE values for both the QPNN and the RK4 method are around the same magnitude for all systems but the PT potential, where the RMSE for the QPNN is one order of magnitude lower than for the RK4 method. In terms of the energy values computed for each of the systems, when referenced to the exact energy (blue line) a similar trend can be observed for the energies computed with the RK4 method and the QPNN. If the RK4 method is regarded as a more robust and mathematically superior method for the calculation of the effective potential functions, this trend may be interpreted as an indicator of the reliability of the QPNN. In the case of the soliton model, although solutions using the RK4 were not feasible due to the nature of the system, the magnitude of RMSE values between the true potential and the learned potential suggests the same qualitative behaviour as the one obtained with the time–independent models.

7 Conclusion

In this work, the QPNN for learning the effective potential functions of different quantum systems was presented. This new neural network is capable of learning the effective potential functions of a variety of systems in a completely unsupervised manner. The results obtained for the different studied systems suggest that the QPNN has an accuracy comparable to the RK4 integrator. The potentials learned with this new QPNN can be used to calculate observables like the energy of the system.

The QPNN formalism presented in this work provide a foundation for the calculation of potentials of more complicated N–body systems. In future work, we will investigate the calculation of potentials for more complicated many electron systems that arise from the use of full configuration interaction densities and wave–functions. The use of the QPNN was also extended for time–dependent systems where the full wave–function was used. In the time–dependent case, $|\psi|$ can not simply be taken as a proxy for a wave–function as one needs to take into account some phase information, i.e., the wave–function in that case can not be a real valued function. In future work, the density–to–potential problem will be analysed by incorporating phase information to create a suitable proxy wave–function.

References

Agboola, D. 2010. Solutions to the Modified Pöschl–Teller Potential in D-Dimensions. *Chinese Physics Letters* 27(4): 040301.

Amabilino, S.; Bratholm, L. A.; Bennie, S. J.; Vaucher, A. C.; Reiher, M.; and Glowacki, D. R. 2019. Training neural nets to learn reactive potential energy surfaces using interactive quantum chemistry in virtual reality. *The Journal of Physical Chemistry A* 123(20): 4486–4499.

Aster, R. C.; Borchers, B.; and Thurber, C. H. 2018. *Parameter estimation and inverse problems*. Elsevier.

Athanassoulis, A. G. 2008. Exact equations for smoothed Wigner transforms and homogenization of wave propagation. *Applied and Computational Harmonic Analysis* 24(3): 378–392.

Beals, R.; and Greiner, P. C. 2009. Strings, waves, drums: spectra and inverse problems. *Analysis and Applications* 7(02): 131–183.

Bondesan, R.; and Lamacraft, A. 2019. Learning Symmetries of Classical Integrable Systems. *arXiv preprint arXiv:1906.04645* .

Burke, K.; Werschnik, J.; and Gross, E. K. U. 2005. Time-dependent density functional theory: Past, present, and future. *The Journal of Chemical Physics* 123(6): 062206.

Carleo, G.; Cirac, I.; Cranmer, K.; Daudet, L.; Schuld, M.; Tishby, N.; Vogt-Maranto, L.; and Zdeborová, L. 2019. Machine learning and the physical sciences. *Reviews of Modern Physics* 91(4): 045002.

Case, W. B. 2008. Wigner functions and Weyl transforms for pedestrians. *American Journal of Physics* 76(10): 937–946.

Chadan, K.; and Sabatier, P. C. 2012. *Inverse problems in quantum scattering theory*. Springer Science & Business Media.

Chen, Z.; Xiong, Y.; and Shao, S. 2019. Numerical methods for the Wigner equation with unbounded potential. *Journal of Scientific Computing* 79(1): 345–368.

Chmiela, S.; Tkatchenko, A.; Sauceda, H. E.; Poltavsky, I.; Schütt, K. T.; and Müller, K.-R. 2017. Machine learning of accurate energy-conserving molecular force fields. *Science advances* 3(5): e1603015.

Cranmer, K.; Golkar, S.; and Pappadopulo, D. 2019. Inferring the quantum density matrix with machine learning. *arXiv preprint arXiv:1904.05903* .

Cranmer, M.; Greydanus, S.; Hoyer, S.; Battaglia, P.; Spergel, D.; and Ho, S. 2020. Lagrangian Neural Networks. *arXiv preprint arXiv:2003.04630* .

Curtright, T.; Fairlie, D.; and Zachos, C. 1998. Features of time-independent Wigner functions. *Physical Review D* 58(2): 025002.

Du, Y.; and Narasimhan, K. 2019. Task-agnostic dynamics priors for deep reinforcement learning. In *International Conference on Machine Learning*, 1696–1705. PMLR.

Feynman, R.; Leighton, R.; and Sands, M. 1965. The Feynman Lectures on Physics Vol. III, chap. 21, sec. 21-9.

Galleani, L.; and Cohen, L. 2002. Approximation of the Wigner distribution for dynamical systems governed by differential equations. *EURASIP Journal on Advances in Signal Processing* 2002(1): 514609.

Goh, G. B.; Hodas, N. O.; and Vishnu, A. 2017. Deep learning for computational chemistry. *Journal of computational chemistry* 38(16): 1291–1307.

Gomes, D. A.; and Silva, J. D. 2008. On the Wigner transform of solutions to the Schrodinger equation. *São Paulo Journal of Mathematical Sciences* 2(1): 85–97.

Greydanus, S.; Dzamba, M.; and Yosinski, J. 2019. Hamiltonian neural networks. In *Advances in Neural Information Processing Systems*, 15353–15363.

Groetsch, C. W.; and Groetsch, C. 1993. *Inverse problems in the mathematical sciences*, volume 52. Springer.

Helgaker, T.; Jorgensen, P.; and Olsen, J. 2014. *Molecular electronic-structure theory*. John Wiley & Sons.

Heller, E. J. 1976. Wigner phase space method: Analysis for semiclassical applications. *The Journal of Chemical Physics* 65(4): 1289–1298.

Hernández de la Peña, L. 2018. A Simplified Pöschl–Teller Potential: An Instructive Exercise for Introductory Quantum Mechanics. *Journal of Chemical Education* 95(11): 1989–1995.

Hibat-Allah, M.; Ganahl, M.; Hayward, L. E.; Melko, R. G.; and Carrasquilla, J. 2020. Recurrent neural network wave functions. *Physical Review Research* 2(2): 023358.

Higgins, I.; Matthey, L.; Glorot, X.; Pal, A.; Uria, B.; Blundell, C.; Mohamed, S.; and Lerchner, A. 2016. Early visual concept learning with unsupervised deep learning. *arXiv preprint arXiv:1606.05579* .

Iten, R.; Metger, T.; Wilming, H.; del Rio, L.; and Renner, R. 2020. Discovering Physical Concepts with Neural Networks. *Physical Review Letters* 124(1): 010508.

Jasinski, A.; Montaner, J.; Forrey, R.; Yang, B.; Stancil, P.; Balakrishnan, N.; Dai, J.; Vargas-Hernández, R.; and Krems, R. 2020. Machine learning corrected quantum dynamics calculations. *Physical Review Research* 2(3): 032051.

Jensen, D. S.; and Wasserman, A. 2018. Numerical methods for the inverse problem of density functional theory. *International Journal of Quantum Chemistry* 118(1): e25425.

Jia, C.-S.; Zhang, L.-H.; and Peng, X.-L. 2017. Improved Pöschl–Teller potential energy model for diatomic molecules. *International Journal of Quantum Chemistry* 117(14): e25383.

Jimenez Rezende, D.; and Mohamed, S. 2015. Variational Inference with Normalizing Flows. *arXiv e-prints* arXiv–1505.

Kingma, D. P.; and Ba, J. 2017. Adam: A Method for Stochastic Optimization. *arXiv preprint arXiv:1412.6980* .

Klimov, A.; Sainz, I.; and Romero, J. 2020. Truncated Wigner approximation as non-positive Kraus map. *Physica Scripta* 95(7): 074006.

Kolesnikov, A.; Beyer, L.; Zhai, X.; Puigcerver, J.; Yung, J.; Gelly, S.; and Houlsby, N. 2019. Big Transfer (BiT): General Visual Representation Learning. *arXiv preprint arXiv:1912.11370* 6(2): 8.

Lample, G.; and Charton, F. 2019. Deep Learning for Symbolic Mathematics. *arXiv preprint arXiv:1912.01412* .

Landau, R.; Paez, M. J.; and Bordeianu, C. 2015. *Computational Physics: Problem Solving with Python*. Wiley, 3rd edition.

Li, H.; Xu, Z.; Taylor, G.; Studer, C.; and Goldstein, T. 2017. Visualizing the Loss Landscape of Neural Nets. *arXiv preprint arXiv:1712.09913* .

Manzhos, S. 2020. Machine learning for the solution of the Schrödinger equation. *Machine Learning: Science and Technology* 1(1): 013002.

Mills, K.; Spanner, M.; and Tamblyn, I. 2017. Deep learning and the Schrödinger equation. *Physical Review A* 96(4): 042113.

Nagai, R.; Akashi, R.; Sasaki, S.; and Tsuneyuki, S. 2018. Neural-network Kohn-Sham exchange-correlation potential and its out-of-training transferability. *The Journal of chemical physics* 148(24): 241737.

Nakajima, M.; Tanaka, K.; and Hashimoto, T. 2020. Neural Schrödinger Equation: Physical Law as Neural Network. *arXiv preprint arXiv:2006.13541* .

Nakatsuji, H. 2002. Inverse Schrödinger equation and the exact wave function. *Physical Review A* 65(5): 052122.

Parr, R. G.; and Yang, W. 1995. Density-functional theory of the electronic structure of molecules. *Annual Review of Physical Chemistry* 46(1): 701–728.

Pronchik, J. N.; and Williams, B. W. 2003. Exactly Solvable Quantum Mechanical Potentials: An Alternative Approach. *Journal of Chemical Education* 80(8): 918.

Pu, J.; Li, J.; and Chen, Y. 2020. Soliton, Breather and Rogue Wave Solutions for Solving the Nonlinear Schrödinger Equation Using a Deep Learning Method with Physical Constraints. *arXiv preprint arXiv:2011.04949* .

Rafi, M.; et al. 1995. An empirical potential function of diatomic molecules. *Physics Letters A* 205(5-6): 383–387.

Raissi, M.; Perdikaris, P.; and Karniadakis, G. E. 2017a. Physics Informed Deep Learning (Part I): Data-driven Solutions of Nonlinear Partial Differential Equations. *arXiv preprint arXiv:1711.10561* .

Raissi, M.; Perdikaris, P.; and Karniadakis, G. E. 2017b. Physics Informed Deep Learning (Part II): Data-driven Discovery of Nonlinear Partial Differential Equations. *arXiv preprint arXiv:1711.10566* .

Raissi, M.; Perdikaris, P.; and Karniadakis, G. E. 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics* 378: 686–707.

Riley, K. F. 1974. *Mathematical Methods for the Physical Sciences: An Informal Treatment for Students of Physics and Engineering*. Cambridge University Press. doi:10.1017/CBO9781139167550.

Robinett, R.; and Robinett, R. W. 2006. *Quantum mechanics: Classical results, modern systems, and visualized examples*. Oxford University Press.

Robinett, R. W. 1997. *Quantum mechanics*. Oxford University Press, New York.

Romanowski, Z. 2007. Numerical Solution Of Kohn-Sham Equation For Atom. *Acta Physica Polonica B* 38(10).

Sakurai, J. J.; and Commins, E. D. 1995. *Modern Quantum Mechanics, Revised Edition*. American Association of Physics Teachers.

Schleder, G. R.; Padilha, A. C.; Acosta, C. M.; Costa, M.; and Fazzio, A. 2019. From DFT to machine learning: recent approaches to materials science–a review. *Journal of Physics: Materials* 2(3): 032001.

Schmidt, J.; Shi, J.; Borlido, P.; Chen, L.; Botti, S.; and Marques, M. A. 2017. Predicting the thermodynamic stability of solids combining density functional theory and machine learning. *Chemistry of Materials* 29(12): 5090–5103.

Schmitz, G.; Godtliebsen, I. H.; and Christiansen, O. 2019. Machine learning for potential energy surfaces: An extensive database and assessment of methods. *The Journal of chemical physics* 150(24): 244113.

Senn, P. 1986. The modified Poschl-Teller Oscillator. *Journal of Chemical Education* 63(1): 75.

Sun, C.; Myers, A.; Vondrick, C.; Murphy, K.; and Schmid, C. 2019. VideoBERT: A Joint Model for Video and Language Representation Learning. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 7464–7473.

Tong, Y.; Xiong, S.; He, X.; Pan, G.; and Zhu, B. 2020. Symplectic Neural Networks in Taylor Series Form for Hamiltonian Systems. *arXiv preprint arXiv:2005.04986* .

Torfi, A.; Shirvani, R. A.; Keneshloo, Y.; Tavvaf, N.; and Fox, E. A. 2020. Natural Language Processing Advancements By Deep Learning: A Survey. *arXiv preprint arXiv:2003.01200* .

Toth, P.; Rezende, D. J.; Jaegle, A.; Racanière, S.; Botev, A.; and Higgins, I. 2019. Hamiltonian Generative Networks. *arXiv preprint arXiv:1909.13789* .

Unke, O. T.; and Markus, M. 2019. Machine Learning Potential Energy Surfaces. *arXiv:1909.08027* .

Vogel, C. R. 2002. *Computational methods for inverse problems*, volume 23. Siam.

Wazwaz, A. 2009. *Partial Differential Equations and Solitary Waves Theory*. Springer Berlin Heidelberg.

Zakhariev, B. N.; and Suzko, A. A. 2012. *Direct and inverse problems: potentials in quantum scattering*. Springer Science & Business Media.

Zhong, Y. D.; Dey, B.; and Chakraborty, A. 2019. Symplectic ODE-Net: Learning Hamiltonian Dynamics with Control. *arXiv preprint arXiv:1909.12077* .

Zhong, Y. D.; Dey, B.; and Chakraborty, A. 2020. Dissipative SymODEN: Encoding Hamiltonian Dynamics with Dissipation and Control into Deep Learning. *arXiv preprint arXiv:2002.08860* .

A Additional experiments using the full wave–function

In order to further explore the capabilities and accuracy of our NN, we start off by describing additional experiments using the QPNN and full wave–function as well as Wigner's functions. The formulation for the new loss functions are described in detail below.

A.1 Using the 1D time–independent Schrödinger equation

In this section we consider some simple one–dimensional time–independent systems. Wave–functions, potential energy and energy levels can be found in Table 2. The learned potentials are reported in figure 6, whereas in figure 7 we show that all proposed models obey energy conservation laws. The quantitative results for the experiments can be found in table 3. For the derivation of these wave–functions and the general properties of these systems, please see (Sakurai and Commins 1995; Feynman, Leighton, and Sands 1965; Robinett 1997). As before, $\hbar$, m and ω were set equal to 1. **Quantum Harmonic Oscillator:** The wave–

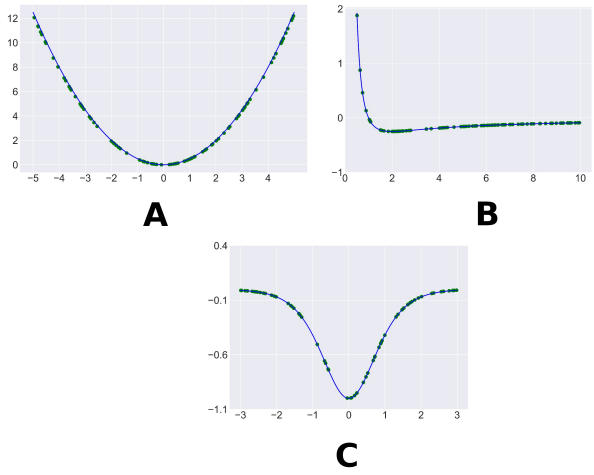

Figure 6: Position (x-axis) vs Potential (y-axis), Ground (Blue) and Learned (Green Dots) Potentials: A) Harmonic Oscillator; B) Hydrogen Atom (2p case); C) Pöschl-Teller (1,1 case).

functions in this case are given by Hermite polynomials $H_n, n = 0, 1, \cdots$. We chose $x \in [-5, 5]$ as input to our model. Since U_θ is given by a differential equation (equation 5), one needs to impose an initial condition to get a unique solution. However, when the output of $U_\theta \in [0, 12.5]$ is considered, the need for the initial conditions are removed. Figure 6 and figure 7 (A) show the learned potential and energy of the system.

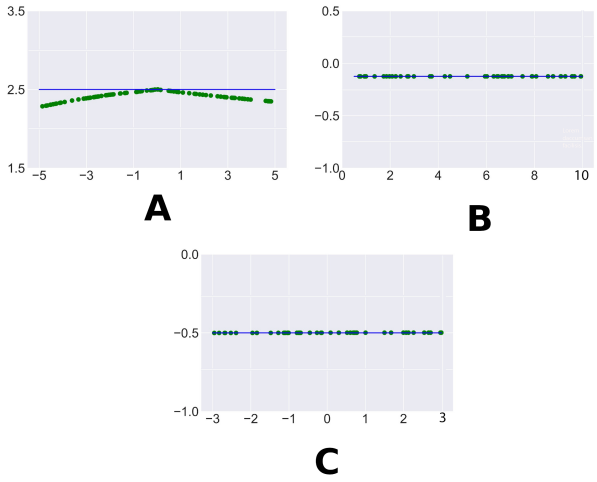

Figure 7: Position (x-axis) vs Energy (y-axis); Ground (Blue) and Learned (Green Dots) Energies: A) Harmonic Oscillator; B) Hydrogen Atom (2p case); C) Pöschl-Teller (1,1 case).

The Hydrogen Atom (2p orbital case) : The general radial wave–functions are given by generalized Laguerre polynomials $L_n^l, n = 1, 2, \cdots$ and $l = 0, 1, 2, \cdots, n-1$ but in this case simplifies to $\psi(r) = \frac{1}{8\sqrt{\pi}} r e^{\frac{-r}{2}}$. We used $r \in [0.5, 10]$ as input to our model and the initial condition $U(1) = 0$. For this system the loss function reads $\mathcal{L}(\theta) = L(\theta) + U_\theta(1)^2$ where $L(\theta)$ is given by equation 5. Figure 6 and figure 7 (B) show the learned potential and energy for this system.

Pöschl-Teller potential : The wave–function ψ generated by this potential is defined by Legendre functions $P_\lambda^\mu(\tanh(x)), \lambda = 1, 2, 3\cdot;\ \mu = 1, 2, \cdot, \lambda - 1, \lambda$. For simplicity, let $\mu = 1$. We chose $x \in [-3, 3]$ as input to our model. We imposed an initial condition $U_\theta(0) = -\frac{\lambda(\lambda+1)}{2}$ and used a similar auxiliary loss function as the one defined above. Figure 6 and figure 7 (C) show the learned potential and energy of the system.

A.2 Particle in a box (perturbed by some external potential)

Now, we turn our attention to a quantum system where the Schrödinger equation cannot be solved exactly, but can be formulated in an approximate manner using perturbation theory. A particle with no spin, of mass m, was placed in an one dimensional square box, $x \in [0, L]$, of length L. Later the particle was presented with the external potential $V(x) = 10x^2$ as perturbation. The wave–function for the perturbed system was approximated by considering first order corrections for the unperturbed particle in a box wave–

Table 2: Wave–functions, Energies, and Potentials of various time independent systems

System	Potential $V(x)$	wave–function $\psi(x)$	Energy
Harmonic Oscillator	$\frac{1}{2}kx^2$	$\frac{1}{\sqrt{2^n\, n!}} \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega x^2}{2\hbar}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right)$	$\hbar\omega(n+\frac{1}{2})$
Pöchl–Teller potential	$-\frac{\lambda(\lambda+1)}{2}\text{sech}^2(x)$	$P_\lambda^\mu(\tanh(x))$	$-\frac{\hbar^2}{2m}(\lambda-\mu)$
Radial Hydrogen atom	$\frac{l(l+1)}{2r^2}-\frac{1}{r}$	$e^{-r/n}(\frac{2r}{na_0})^l L_{n+l}^{2l+1}(\frac{2r}{na_0})$	$-\frac{R_H}{(n+l)^2}$
2D Harmonic Oscillator	$\frac{1}{2}k(x^2+y^2)$	$H_{n_x}\left(\sqrt{\frac{m\omega}{\hbar}}x\right) H_{n_y}\left(\sqrt{\frac{m\omega}{\hbar}}y\right) e^{-\frac{m\omega(x^2+y^2)}{2\hbar}}$	$\hbar\omega(n_x+n_y+1)$

function,

$$\psi_n = \psi_n^0 + \sum_{n\neq k} \frac{<\psi_n^0|V(x)|\psi_k^0>}{E_n^0 - E_k^0}\psi_k^0, \quad n,k = 1,2,3,\ldots, \tag{11}$$

where ψ_n^0 and E_n^0 are the particle in a box's unperturbed nth state wave–function and its energy, whereas, $<\psi_n^0|V(x)|\psi_k^0>$ indicates the following integral

$$<\psi_n^0|V(x)|\psi_k^0> = \int (\psi_n^0)^* V(x)\psi_k^0 dx. \tag{12}$$

For our computations, the wave–function, ψ_n^0, obtained as solution of the Schröedinger equation for the particle in a box model reads,

$$\psi_n^0 = \sqrt{\frac{2}{L}}\sin(\frac{n\pi}{L})x \quad n = 1,2,3,\ldots, \tag{13}$$

and the energy for the system is given by

$$E_n^0 = \frac{n^2\hbar^2\pi^2}{2mL^2} \quad n = 1,2,3,\ldots. \tag{14}$$

Here, we use the wave–function corrected only up to first order and $x \in [0,1]$. In this experiment we not only learned the potential, but also the perturbed wave–function based only in the systems initial conditions without perturbation. We use two neural networks, one to learn the potential and the other to learn the perturbed wave–function. The perturbed wave–function was learned in a supervised manner, whereas the potential was learned in an unsupervised manner. If W_θ is the neural network learning the perturbed wave–function ψ_{pert}, then our auxiliary loss function becomes

$$\mathcal{L}_\theta = ||W_\theta - \psi_{\text{pert}}||_2^2 + L_\theta, \tag{15}$$

where L_θ is the time–independent Schrödinger loss defined in the main text, which was used to learn the potential and calculated by the perturbed wave–function. Figure 8 shows the results for this system. It seems that energy is not conserved for this system, but that is merely due to the nature of the truncated, first–order perturbation approximation.

A.3 2D Harmonic Oscillator

Unlike other NN, our QPNN scales easily and quickly to higher dimensions. For the 2–dimensional Harmonic Oscillator, the wave–function is defined as a product of two Hermite polynomials (such as the one defined in the main text).

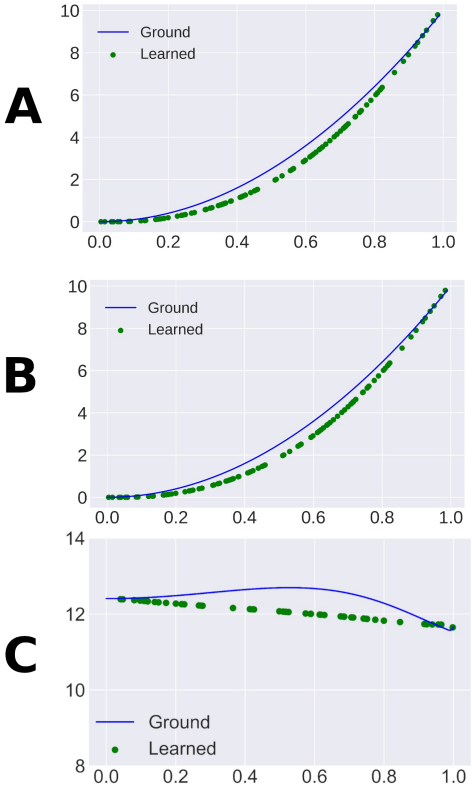

Figure 8: A) Wave–function of the particle in a box; B) Potential (y-axis) vs Position (x-axis) of the particle; C) Conservation of an "energy" like object.

We chose $x, y \in [0,1]$ as input to our model and constrained our output to $[0,1]$. The loss function in this case is exactly as the one defined for the 1D Harmonic Oscillator. Figure 9 shows the results for this system, as this figure shows, the learned energy is a good approximation to the total energy (z scale chosen from $[4.99, 5.01]$).

B Motivation behind the time–independent Schrödinger Loss

We now present a brief, yet complete, explanation for our time–independent Schrödinger loss function. The Hamiltonian $\hat{\text{H}}$ is the sum of the kinetic energy $\hat{\text{T}}$ and the potential

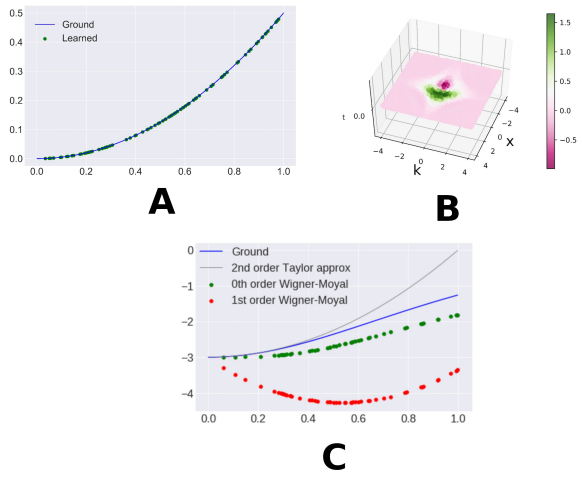

Figure 9: 2D Harmonic Oscillator. (A) Ground Truth Potential. (B) Learned Potential by our network. (C) Ground Truth Energy surface and Learned Energies by our model (z axis is between 4.99 and 5.005).

energy $\hat{\mathrm{V}}$, where the kinetic energy is given by the Laplacian operator,

$$\hat{\mathrm{H}} = -\frac{\hbar^2}{2m}\nabla_x^2 + \hat{V}(x). \tag{16}$$

For the time independent case, the Schrödinger's equation boils down to

$$\hat{\mathrm{H}}\psi = E\psi \tag{17}$$

where E is the energy of the system. For simplicity, let $\hbar = m = 1$. By using equation 16, equation 17 reads

$$(-\frac{1}{2}\nabla_x^2 + \hat{V}(x))\psi = E\psi. \tag{18}$$

Dividing the above equation by ψ yields

$$\frac{-\frac{1}{2}\nabla_x^2\psi}{\psi} + \hat{V}(x) = E. \tag{19}$$

Since the energy of a given system is a scalar quantity, the derivative with respect to x on the left hand side of equation 19 is zero, which defines the time–independent Schrödinger loss function.

C Wigner Functions

An alternative formulation of quantum dynamics may be given by the Wigner function (Curtright, Fairlie, and Zachos 1998; Chen, Xiong, and Shao 2019). The Wigner function, $W(x, p, t)$, is a phase space distribution function which behaves similarly to the position $|\psi(x)|^2$ and the momentum $|\psi(p)|^2$ distribution functions (Case 2008). Unlike wave–functions, Wigner functions are real valued and bounded. However, contrary to probability distributions, $W(x, p, t)$ can take negative values. Thus, the Wigner distribution is termed as a quasi–probability distribution and in a sense loses some of its classical appeal. Using the Schrödinger's equation (equation 1) and the Taylor expansion, the time evolution of the Wigner function is given by an infinite order partial differential equation called Wigner–Moyal equation (Case 2008).

$$\begin{aligned}\frac{\partial W(x,p,t)}{\partial t} &= -\frac{p}{m}\frac{\partial W(x,p,t)}{\partial x} \\ &+ \sum_{s=0}^{\infty}(-h^2)^s\frac{1}{(2s+1!)}(\frac{1}{2})^{2s}\frac{\partial^{2s+1}U(x)}{\partial x^{2s+1}} \\ &\times \frac{\partial^{2s+1}W(x,p,t)}{\partial p^{2s+1}}.\end{aligned} \tag{20}$$

C.1 Learning Potentials from Wigner Functions

In the case of the Wigner function, our Neural Network was trained by implementing a truncated Wigner–Moyal loss,

$$\begin{aligned}L_{Wigner}(\theta) = \Bigg|\Bigg|&\frac{\partial W(x,p,t)}{\partial t} + \frac{p}{m}\frac{\partial W(x,p,t)}{\partial x} \\ &- \sum_{s=0}^{k}(-h^2)^s\frac{1}{(2s+1!)}(1/2)^{2s}\frac{\partial^{2s+1}U_\theta(x)}{\partial x^{2s+1}} \\ &\times \frac{\partial^{2s+1}W(x,p,t)}{\partial p^{2s+1}}\Bigg|\Bigg|_2^2,\end{aligned} \tag{21}$$

where for all our experiments $k = 0, 1$. The case where $k = 0$ is known as the Liouville equation. However, we note that equation 21 determines U_θ up to a constant. Thus, an initial condition depending on each individual system was added.

C.2 Experiments with the Wigner functions

Harmonic Oscillator : The Wigner function for the harmonic oscillator has the following form (Case 2008):

$$W(x,p,t) = e^{-(x^2+p^2)}\left(x^2+p^2+\sqrt{2}x\cos t - \sqrt{2}p\sin t\right).$$

Since $\frac{\partial^n U}{\partial x^n} = 0,\ \forall\ n \geq 3$, the Moyal–Wigner equation in this case regresses to the classical Liouville equation. Let $x, p, t \in [0, 1]$ and x is the input to the model. The initial condition for this systems is $U_\theta(0) = 0$ and the loss function reads

$$\mathcal{L}(\theta) = L_{\text{Wigner}}(\theta) + U_\theta(0)^2,$$

where L_{Wigner} is given by equation 21. Figure 10 (A) shows the potential learned by the model.

Pöschl-Teller potential : The Wigner function in this case (Chen, Xiong, and Shao 2019) is given by

$$\begin{aligned}&W_{2,1,0}(x,k,t): \\ &= \frac{3}{8}\int_{-\infty}^{\infty}\text{sech}^2(x+\frac{y}{2})\text{sech}^2(x-\frac{y}{2}) \\ &\times\Bigg[2\sinh(x+\frac{y}{2})\sinh(x-\frac{y}{2}) \\ &+\sqrt{2}\sinh(x-\frac{y}{2})e^{\frac{i3t}{2}} \\ &+\sqrt{2}\sinh(x+\frac{y}{2})e^{\frac{-i3t}{2}}+1\Bigg]e^{-iky}dy.\end{aligned} \tag{22}$$

The Wigner function is a real–valued bounded function. Thus by breaking the integral in equation 22 into real and complex parts, we only focus on the real part. Using Euler's formula, we get the following:

$$
\begin{aligned}
g_{2,1,0}(x,k,t) = \frac{3}{8}\int_{-\infty}^{\infty} &\operatorname{sech}^2(x+\frac{y}{2})\operatorname{sech}^2(x-\frac{y}{2}) \\
&\times \Big[2\sinh(x+\frac{y}{2})\sinh(x-\frac{y}{2})\cos(-ky) \\
&+\sqrt{2}\sinh(x-\frac{y}{2})\cos(\frac{3t}{2}-ky) \\
&+\sqrt{2}\sinh(x+\frac{y}{2})\cos(-\frac{3t}{2}-ky) \\
&+\cos(-ky)\Big].
\end{aligned}
\tag{23}
$$

Note that the integral in equation 23 is invariant under the change of variable $y \to -y$. This implies that in order to calculate $g_{2,1,0}(x,k,t)$, we only have to integrate from 0 to ∞ and multiply the integral by 2. Our final simplification comes from studying the decay properties of the Wigner functions. Using $\operatorname{sech}(x) = \frac{2}{e^x+e^{-x}}$ and $\sinh(x) = \frac{e^x-e^{-x}}{2}$, we found that the integrand in equation 23 behaves like $O(e^{-y})$ (resp. $O(e^{y})$) as $y \to \infty$ (resp. $y \to -\infty$). We established a threshold of 10^{-9} to truncate the integral from positive real axis to a bounded interval which gives the following form:

$$
\begin{aligned}
f_{2,1,0}(x,k,t) = \frac{3}{4}\int_{0}^{20} &\operatorname{sech}^2(x+\frac{y}{2})\operatorname{sech}^2(x-\frac{y}{2}) \\
&\times \Big[2\sinh(x+\frac{y}{2})\sinh(x-\frac{y}{2})\cos(-ky) \\
&+\sqrt{2}\sinh(x-\frac{y}{2})\cos(\frac{3t}{2}-k) \\
&+\sqrt{2}\sinh(x+\frac{y}{2})\cos(-\frac{3t}{2}-ky) \\
&+\cos(-ky)\Big].
\end{aligned}
\tag{24}
$$

The Wigner method to study the time–frequency properties of dynamical systems involves taking the partial derivatives with respect to time of the Wigner function. These derivatives on the Wigner function yield what is known as the Wigner–Moyal equation. The physical interpretations, numerical difficulties and approximations of the Wigner–Moyal equation have been widely discussed in the literature; for information about the mathematical challenges associated with the Wigner–Moyal equation, we recommend readers to consult these references (Chen, Xiong, and Shao 2019; Case 2008; Galleani and Cohen 2002; Heller 1976; Curtright, Fairlie, and Zachos 1998; Klimov, Sainz, and Romero 2020; Athanassoulis 2008; Gomes and Silva 2008). In this case, the potential is $U(x) = -3\operatorname{sech}(x)^2$ which is infinitely differentiable. For this experiment we chose $x \in [0,1]$ as input to our model and approximated the infinite order PDE (equation 20) by equation 21. Thus, one cannot assume that any non–steady state solution predicted by the truncated Wigner function is immediately valid, as it can be shown that higher order quantum corrections are responsible for quantum mechanical phase–space behavior. The 0th order truncation matches the potential in a small neighborhood of 0. Figure 10 summarizes some of our findings,

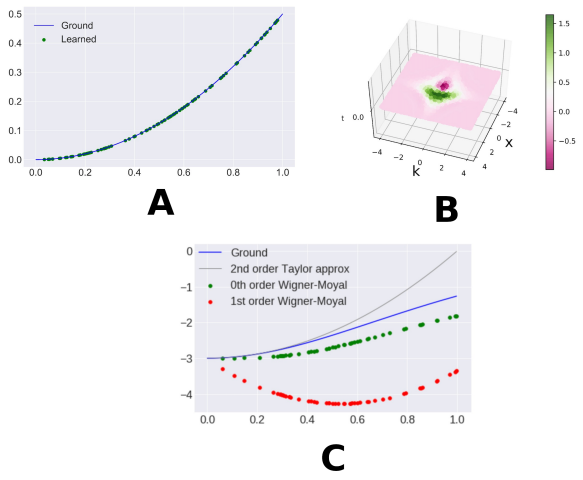

Figure 10: A) Position (x-axis) vs Potential (y-axis) of Harmonic oscillator using the Wigner function. B) Our approximation of the Wigner function of Pöschl–Teller potential. C) Various approximations of the potential using the Wigner–Moyal equation.

D Some training details and hyperparameters

Our Neural Network is a 4-layer feedforward network with a residual connection between the second and the third layers. The activation and the scaling in the final layers varied from experiment to experiment. Our main motivation for scaling and using different activation is to show that an appropriate architecture can perfectly learn the correct potential without an initial condition. All the models are trained for 1000 epochs. Table 4 shows the activation, scaling and the size of the training data for each of the studied systems. All the training data was randomly sampled from the appropriate domains and trained in a minibatch fashion with batch size of 32.

Table 3: A quantitative analysis of our model

System	RMSE between True and Learned Potentials	RMSE between True and Learned Energies
Harmonic Oscillator	$1.1 \times 10^{-1} \pm 5.0 \times 10^{-2}$	$1.0 \times 10^{-1} \pm 2.0 \times 10^{-2}$
Pöchl–Teller potential	$1.0 \times 10^{-4} \pm 6.0 \times 10^{-5}$	$8.0 \times 10^{-4} \pm 6.0 \times 10^{-5}$
Radial Hydrogen atom	$3.0 \times 10^{-4} \pm 8.0 \times 10^{-5}$	$3.0 \times 10^{-4} \pm 7.0 \times 10^{-5}$
2D Harmonic Oscillator	$3.0 \times 10^{-3} \pm 9.0 \times 10^{-4}$	$4.0 \times 10^{-3} \pm 8.0 \times 10^{-4}$
Particle in a Box	$4.3 \times 10^{-1} \pm 6.0 \times 10^{-2}$	$5.5 \times 10^{-1} \pm 8.0 \times 10^{-1}$
Soliton	$2.9 \times 10^{-1} \pm 4.0 \times 10^{-2}$	-
Harmonic Oscillator from Wigner	$4.0 \times 10^{-3} \pm 8.0 \times 10^{-5}$	-

Table 4: Some training details and model hyperparameters

System	Final Layer Activation	Final Layer Scaling	Size of training data
Harmonic Oscillator	Sigmoid	12.5	2500
Pöchl–Teller potential	None	None	2500
Radial Hydrogen atom	None	None	2500
2D Harmonic Oscillator	Sigmoid	None	5000
Potential for Particle in a Box	Sigmoid	10	4000
Perturbation for Particle in a Box	None	None	4000
Soliton	None	None	3000
Harmonic Oscillator from Wigner	None	Sigmoid	5000
Pöschl–Teller from Wigner	None	None	2000

E Additional Figures

Some additional figures of the wave–functions and Wigner functions used in our experiments.

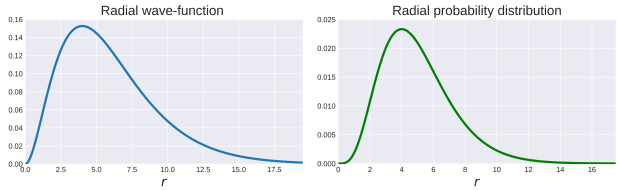

Figure 11: Left: $2p$ Radial wave–function for the Hydrogen atom. Right: $2p$ Radial probability distribution for the Hydrogen atom.

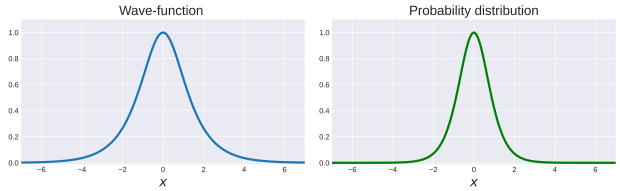

Figure 12: Wave–function (Left) and Probability distribution (Right) for the first bound state of the Pöschl–Teller potential.

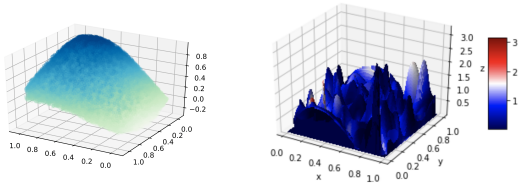

Figure 13: Left: Wigner quasi–probability distribution for the ground state of the Harmonic Oscillator. Right: Our approximation of the Wigner function for the Pöschl–Teller.

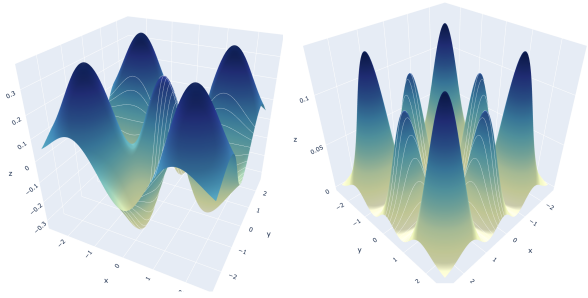

Figure 14: Left: 2D Harmonic Oscillator Wave–function. Right: 2D Harmonic Oscillator Wave–function.