PROCEEDINGS OF SCIENCE

Deep learning approach to high dimensional problems of quantum mechanics

V.A. Roudnev[a,*] **and M.M. Stepanova**[a]

[a] *Department of Computational Physics, St. Petersburg State University, 1 Ulyanovskaya Street, St-Petersburg, Russia*

E-mail: v.rudnev@spbu.ru, m.stepanova@spbu.ru

Traditional linear approximation of quantum mechanical wave functions are not practically applicable for systems with more than 3 degrees of freedom due to the "the curse of dimensionality". Indeed,the number of parameters required to describe a wave function in high-dimensional space grows exponentially with the number of degrees of freedom. Inevitably, strong model assumptions should be used when studying such systems numerically. There are, however, estimates of the complexity of a function reproduced by a deep neural network (DNN) that demonstrate the same exponential growth with respect to the number of the network layers. The number of parameters for DNN grows only linearly with the number of layers. This gives us a hope that application of DNN as an approximant for a wave function in high-dimensional space might moderate the computational requirements for reproducing such systems and make 4- or higher-dimensional systems feasible for direct numerical modeling. We present a study of DNN approximation properties for a multi-dimensional quantum harmonic oscillator. We demonstrate that the computational resources required to reproduce the wave function depend on the dimensionality of the problem and the quantum numbers of the state. Increasing the number of hidden layers in a fully-connected feed-forward DNN we can reproduce some excited states of a multidimensional system with computational resources comparable to low-dimensional cases. Using the DNN as an approximant for a wave function paves a way to developing a new class of computational schemes for solving the Schroedinger equation for high-dimensional systems.

**** The 6th International Workshop on Deep Learning in Computational Physics (DLCP2022) ****
**** 6-8 July 2022 ****
**** JINR, Dubna, Russia ****

[*] Speaker

PoS(DLCP2022)013

1. Introduction

Numerical modeling of complex quantum mechanical systems is an important branch of modern computational physics. Even though quite complex quantum systems – including many-electron systems – can be modeled on the base of the existing numerical algorithms, many important quantum systems can only be treated within model approaches that force some non-existent symmetry into the solution or the interaction model. The typical complexity barrier which prevents us from applying more realistic models arises from the curse of dimensionality. The quantum few-body problem provides a natural illustration of the difficulties we meet. Consider a system of n quantum particles. The dimension of the configuration space scales as dn, where d is the number of single particle degrees of freedom. In practice, it means that we have to solve the partial differential equations of the space with dimensionality more than 3 even for a two-body system, if the interaction between the particle is non-central and the particles are captured in a non-symmetric field, such as an optical trap or a semiconductor heterostructure. In practice, we often have to explore even more complex configurations that lead to 6- or higher-dimensional problem. It is easy to see, that traditional linear approximation approaches that can efficiently be applied to three-dimensional systems fail in higher dimensions just because of the prohibitive size of the linear basis required to reproduce the solutions accurately. It means that alternative approaches that do not rely on traditional linear techniques might be very beneficial when solving quantum mechanical problems in high dimension.

In this work we explore fully connected feed-forward deep neural network as an approximant for a multi-dimensional quantum mechanical problem. As a test ground we use the wave function of a multi-dimensional harmonic oscillator which we approximate with deep and shallow networks of varying depth and width.

2. Neural networks for Schroedinger equation

2.1 Early approach

The idea to use a neural network to represent an approximate solution of the Schrödinger equation – and some other partial differential equations – was introduced and used by Lagaris [1] more than 20 years ago. Recently, the idea of training a network to satisfy some differrential equations started being presented under a specially coined term "physics-informed neural networks" [2]. Here, in this section, we outline the original scheme proposed by Lagaris for solving the Schrödinger equation and illustrate it with the case of a multidimensional harmonic oscillator.

Consider the Hamiltonian for a harmonic oscillator in d dimensions.

$$H = -\Delta_x + \|x\|^2 \tag{1}$$

Here $x \in \mathbb{R}^d$ is the position of the particle in the d-dimensional configuration space. We shall look for the solutions of the Schrödinger equation

$$(H - E)\Psi(x) = 0\,, \tag{2}$$

$\Psi \in L_2(\mathbb{R}^d)$. The variables in this model system, obviously, separate, and the solutions read

$$\Psi(x) \equiv |n_1+1, n_2+1, \ldots, n_d+1\rangle = \frac{1}{\sqrt{\pi^{d/2} 2^{\sum_{k=1}^d n_k} \prod_{k=1}^d n_k!}} e^{-\frac{|x|^2}{2}} \prod_{k=1}^d H_{n_k}(x_k) \ . \tag{3}$$

Here $H_{n_k}(x_k)$ stands for the Hermit polynomials, n_k+1 are the quantum numbers corresponding to the index of the excitation in k-th coordinate. The energies of the corresponding states read $E_n = 2\sum_{k=1}^d n_k - d$.

In [1] it has been proposed to seek for a solution of the Schroedinger equation in the form of a neural network with one hidden layer and an asymptotic normalizing factor

$$\Psi(x; \mathbf{u}, \mathbf{v}, \mathbf{w}, \beta) = e^{-\beta x^2} \sum_{j=1}^m v_j \sigma(w_j x + u_j) \ . \tag{4}$$

Similar representations have been used also for two- and three-dimensional problems. The weights and biases have been obtained by minimization of the following functional

$$\mathcal{L}(\mathbf{u}, \mathbf{v}, \mathbf{w}, \beta) = \frac{\sum_i [H\Psi(x_i; \mathbf{u}, \mathbf{v}, \mathbf{w}, \beta) - \epsilon \Psi(x_i; \mathbf{u}, \mathbf{v}, \mathbf{w}, \beta)]^2}{\int dx |\Psi(x; \mathbf{u}, \mathbf{v}, \mathbf{w}, \beta)|^2} \tag{5}$$

where the energy has been evaluated as

$$\epsilon = \frac{\int dx \Psi^*(x_i; \mathbf{u}, \mathbf{v}, \mathbf{w}, \beta) H\Psi(x_i; \mathbf{u}, \mathbf{v}, \mathbf{w}, \beta)}{\int dx |\Psi(x; \mathbf{u}, \mathbf{v}, \mathbf{w}, \beta)|^2}$$

For the excited states, it has been suggested to use the same functional, but the trial function had to be explicitly orthogonalized with respect to the already calculated states.

This approach demonstrated its applicability to problems of small dimensions, but suffered from a few drawbacks. First of all, the approach does not scale up well as the dimensionality of the problem grows. The fixed set of collocation points had been chosen as the nodes of qubature formulas used for calculations of the integrals. The optimization procedure relied on local optimization algorithms which required very refined handwork to avoid non-optimal solutions to the problem. Even though applicability of neural networks for solving the Schrödingr equation has been demonstrated, practicality of the algorithm remained questionable. Indeed, the fixed set of collocation points requires a lot of computational resources at each step of the optimization algorithm. The single hidden layer – as we shall demonstrate in the subsecuent sections – is not flexible enough to reproduce strongly oscillating solutions, especially in higher dimensions.

The techniques developed for engineering deep learning applications, however, can compensate for the drawbacks of the earlier "demonstration of principle" approach.

2.2 Deep neural network approach advantages

The main disadvantage of the early attempts to apply neural networks for solving the Schrödinger equation [1] comes, probably, from the fixed set of collocation points employed in the calculation process. The other limitation of the original approach comes from the application of a single hidden layer networks as an approximant. Application of deep neural network (DNN) as an approximant

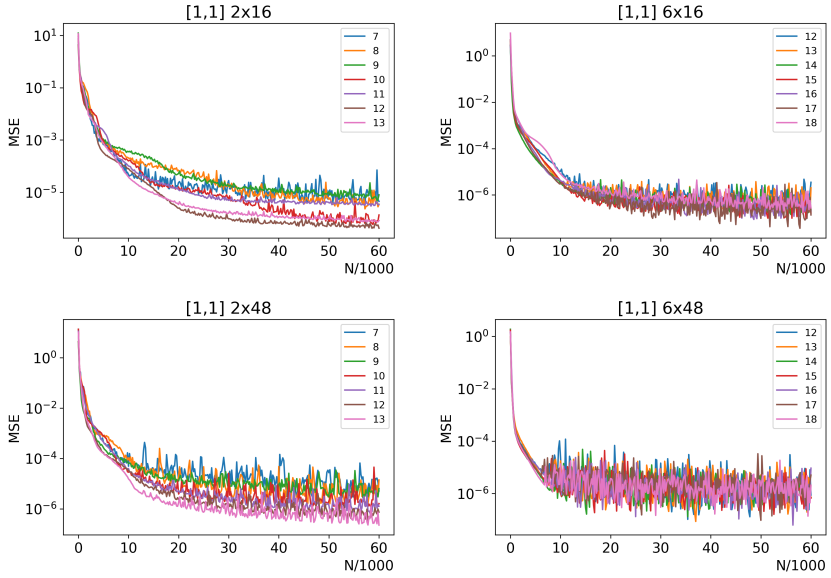

Figure 1: Learning curves observed for the two-dimensional ground state for different network architectures and varying batch sizes. The numbers n_b in the figure legends indicate the batch sizes of 2^{n_b}.

promises serious advantages. Let us mention a few properties of DNN that can be helpful when solving the Schroödinger equation in higher dimensions.

The complexity of a neural network approximant grows geometrically in the number of layers, but the number of parameters scales only linearly[3]. This gives us a hope that quite sophisticated functions can be approximated well even with a moderate number of the network parameters.

The number of required parameters grows linearly with the dimension of the problem [4]. This observation suggests that calculations in higher dimensions will become feasible.

There is a belief that all the local minima of a deep network are equivalent [5], which lowers the risks of hitting a local minimum.

The training procedure relies on the mini-batch approach which allows us to draw the collocation points from an infinite, and, moreover, continuous set of points. This minimizes the risks of not representing some essential features of the solution while keeping the computational costs moderate.

In the next section we shall describe our preliminary computational experiments with DNN as an approximant for the wave function. We shall see, that DNN have advantages over the shallow networks when approximating complex wave functions in higher dimension, which does promise to lift the curse of dimensionality from the computations of high-dimensional quantum problems.

3. Numerical experiments

In this work we mostly concentrate on the approximative properties of a DNN. Instead of solving the Schrödinger equation, we resort to a computationally simpler problem of approximating its solutions (3) in different dimensions. We shall use the following notation. A multiindex $[n_1, n_2, ..., n_d]$ corresponds to the quantum numbers that define a $d-$dimensional state of the quantum harmonic oscillator (3). The architecture of the neural network will be marked as $D \times W$, which

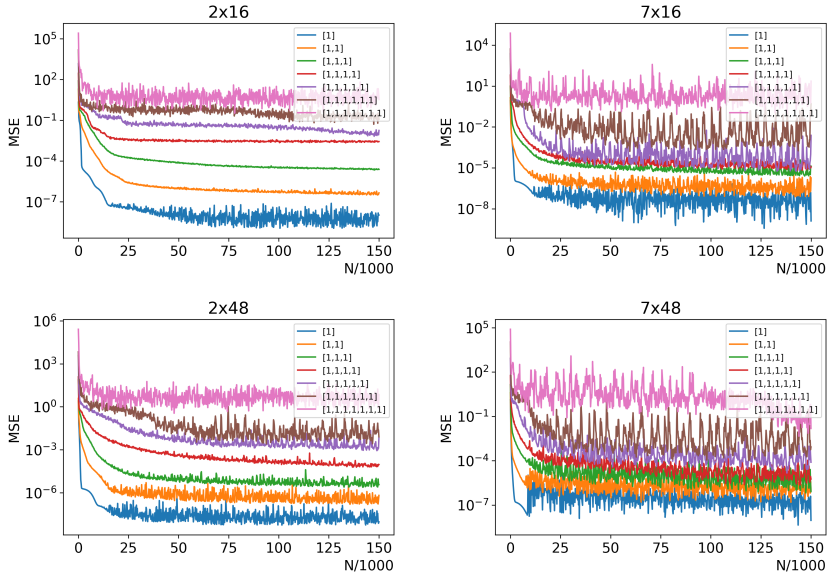

Figure 2: Convergence of the neural network approximants of different architectures to the oscillator ground state in different dimensions.

corresponds to D hidden layers of the width W each. In order to evaluate the convergence we shall use two different functionals which we shall evaluate on a fixed test set of collocation points $x = \{x_1, x_2, \ldots, x_N\}$. The first is the mean square deviation from the target function Ψ

$$MSE = \frac{1}{N}\sum_{k=1}^{N}(NN(x_k; w) - \Psi(x_k))^2\,, \tag{6}$$

where $NN(x_k; w)$ stands for the neural network approximant. The second one is the energy estimate of the state

$$E_e = \frac{\sum_k NN(x_k; w)(-\Delta + |x_k|^2)NN(x_k; w)}{\sum_k NN(x_k; w)^2}. \tag{7}$$

The energy estimate is important, as the convergence of the energy to the proper value indicates not only the point-wise convergence of the solutions, but also the proper behavior of its second derivatives. In all our runs we observe that well approximated states do correspond to its proper energy with a reasonable accuracy.

In Figure 1 we show the learning curves observed for the ground states of the 2d oscillator approximated by shallow $(2 \times W)$ and deep $6 \times W$ neural networks. The batch sizes about $2^1 1$–$2^1 3$ look sufficient in order to obtain stable converged results, and we use the batch sizes of that order in all the following numerical experiments.

First, let us take a look at the learning curves for the problems of similar complexity in different dimension. In Figure 2 we show the learning curves for two shallow and two deep neural networks with the task of approximation the oscillator ground state in dimensions from 1 to 7. In all the examples we observe a trend of slowing the convergence with the dimensionality of the problem. The shallow networks, however, practically fail to converge at dimension 4 or higher, whereas the

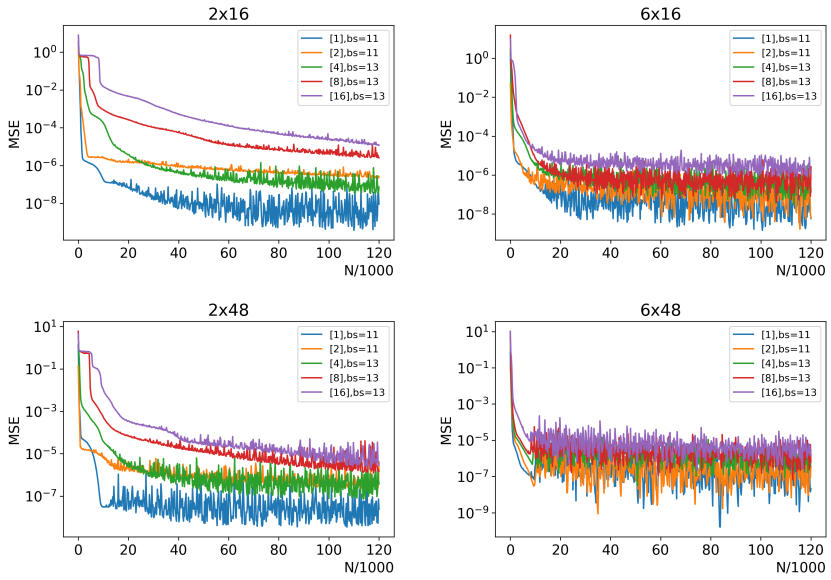

Figure 3: Learning curves for shallow and deeper neural networks approximating the one-dimensional oscillator states of different complexity.

wider and deeper networks keep improving even at 7-dimensional case, even though the convergence rate seems to be rather slow.

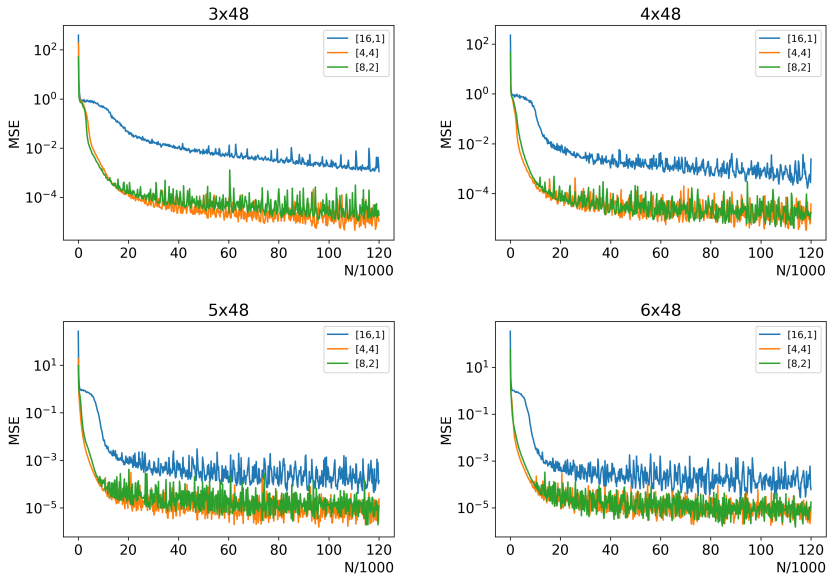

Figure 4: Learning curves for neural networks of varying depth approximating the two-dimensional oscillator states of different complexity.

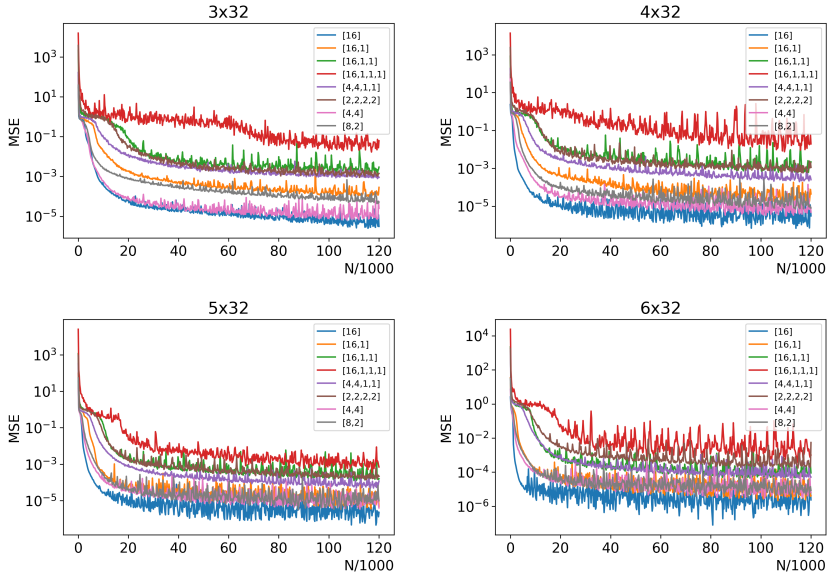

Figure 5: Learning curves for neural networks of varying depth approximating the oscillator states of different complexity in varying dimensions. The higher-dimensional states with high excitations in one single degree of freedom are the most difficult to approximate.

The complexity of the approximated state also affects the convergence rate for the DNN approximants. We illustrate it in Figures 3, 4 and 5. For 1-dimensional problems all the network architectures demonstrate rather good convergence (Figure 3). We can observe, however, that as the number of the approximated state oscillations – the state complexity – grows, the convergence slows down. This effect is much less pronounced for the networks with multiple hidden layers.

In Figures 4 and 5 we fix the complexity of the state by fixing the product of the quantum numbers of the state being approximated. The oscillations are distributed differently between different degrees of freedom. Again, we see that deeper networks demonstrate better convergence. The worst case scenario, however, corresponds to the states that strongly oscillate in one dimension, while behaving smoothly in all others.

We observed that deeper networks demonstrate better convergence, especially in the case of approximating the more complex functions in higher dimensions. The question remains, whether increasing the width of the network, i.e. the number of neurons in a hidden layer, can help us to solve higher-dimensional problems. In Figures 6 and 7 we applied the two alternative strategies of growing the network to approximate the states $[1,1,1,1]$ and $[4,4,1,1]$ of a four-dimensional harmonic oscillator. We should emphasize that these states are beyond the reach of the traditional linear mesh-based approximation methods. Figure 6 demonstrates that the single hidden layer network does not demonstrate any convergence for the excited four-dimensional state, whereas for the ground state the convergence is very slow even for the hidden layer as wide as 1024 neurons. The two hidden layer networks converge somewhat better, but the excited state remains very poorly reproduced. The computational cost scales quadratically with the layer width, and the computational cost rapidly becomes prohibitive.

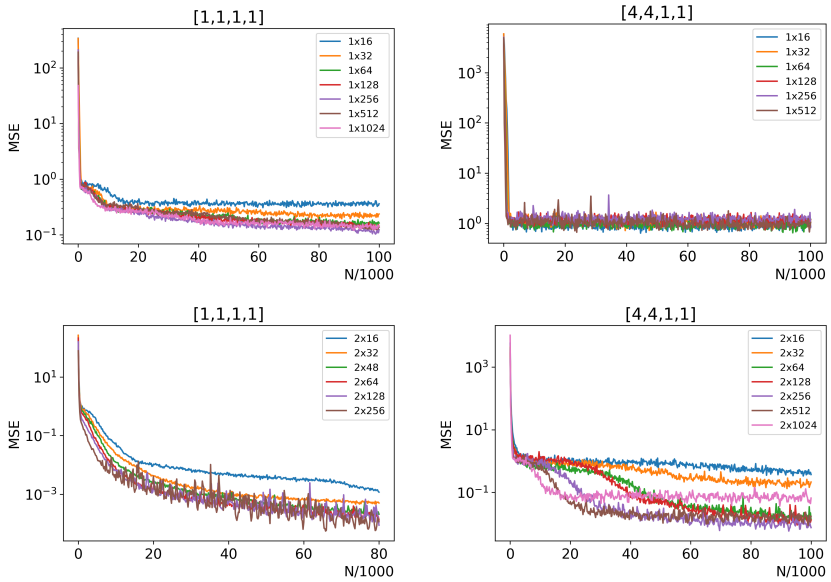

Figure 6: Learning curves for shallow networks of growing width approximating four-dimensional states [1, 1, 1, 1] and [4, 4, 1, 1]

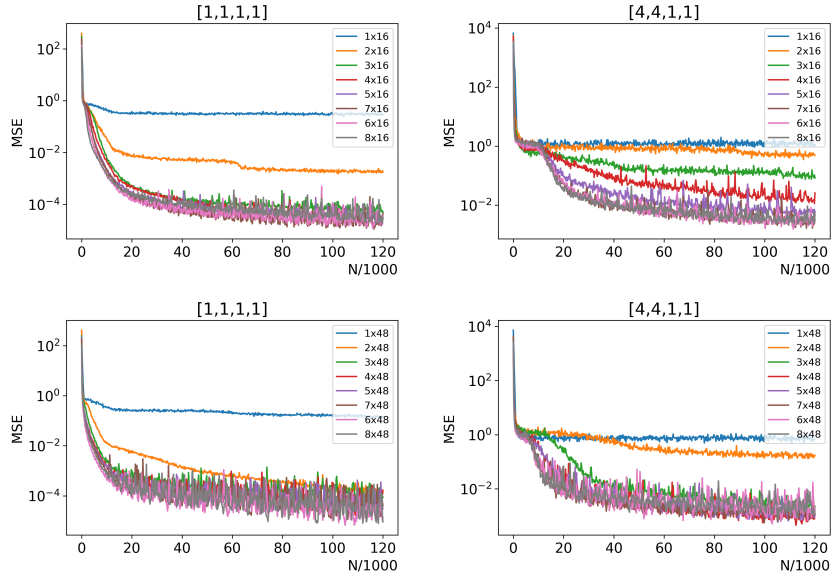

Figure 7: Learning curves for networks of growing depth approximating four-dimensional states [1, 1, 1, 1] and [4, 4, 1, 1]

The strategy of growing the number of layers demonstrates rather different results (Figure 7). The networks of more than 4 hidden layers do demonstrate convergence, even though it seems to be rather slow in the case of the excited state.

Finally, in Figure 8 we show the learning curves for the [16, 1, 1, 1] state. There we show not only the mean square error, but also the energy estimate of the state. As we have mentioned, the

convergence of the energy estimate to the correct value ensures that the state is being approximated correctly with – at least – its second derivatives.

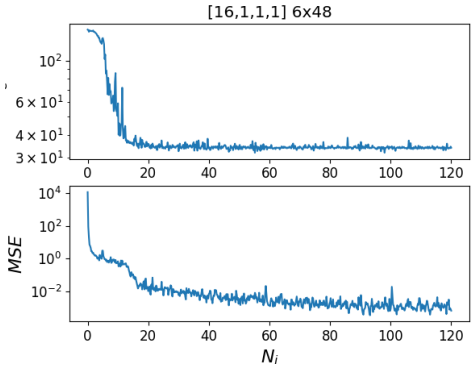

Figure 8: Learning curves for the energy of the [16, 1, 1, 1] state approximated by 6×48 neural network. Upper panel shows the energy estimate (7), the lower panel demonstrates the mean square error.

4. Conclusions

We have reported a series of numerical experiments that we have performed to check, whether a fully-connected multiple hidden layer neural network can be used as an approximant for a quantum mechanical state in high dimensions that remain beyond reach of more traditional computational methods. We can characterize our results as promising. Using rather moderate computational resources – each run usually takes no longer than 30 minutes on a single GPU – we have managed to reproduce four-dimensional states with reasonable accuracy. We have observed that increasing the depth of the network rather than its width does allow us to work with quantum systems with more than four degrees of freedom. We consider this as a first step towards developing a new computational technology that would make direct modeling of very complex quantum systems possible.

Many unsolved problems remain, however. In this work we only studied the approximation properties of the networks, but the problem of solving the Schrödinger equation directly will also require new studies of the most appropriate objective functions. In order to improve the convergence we should study optimal policy of drawing the collocation points from the corresponding domain in configuration space. We have also studied the problem with very simple boundary conditions that correspond to a bound state of the quantum system. The ways to implement appropriate boundary conditions for scattering states will also require extensive research.

References

[1] *Lagaris I. E. Likas A. Fotiadis D. I.* Artificial neurla network method in quantum mechanics // Computer Physics Communications. — 1997. — Vol. 104. — P. 1.

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

[3] *Raghu M. Poole B. Kleinberg J. Ganguli S. Sohl-Dickstein J.* On the Expressive Power of Deep Neural Networks // https://doi.org/10.48550/arXiv.1606.05336

[4] *Mhaskar H.Ṅ. Poggio T* Function approximation by deep networks.// Communications on Pure & Applied Analysis — 2020. — Vol. 19, No. 8. — P. 4085–4095.

[5] *Kawaguchi K.* Deep learning without poor local minima // 30th Conference on Neural Information Processing Systems (NIPS 2016), Barcelona, Spain