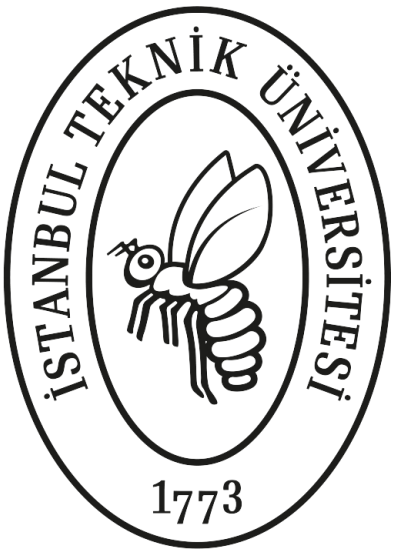
ISTANBUL TECHNICAL UNIVERSITY

FACULTY OF SCIENCE AND LETTERS

Graduation Project



Machine Learning and Non-linear Schrödinger Equation

Hüseyin Talha Şenyaşa

Department : Physics Engineering
Student ID : 090120132
Advisor : Assoc. Prof. A. Levent Subaşı

FALL 2017

Summary

We train an artificial neural network to estimate the ground state energy of a one-dimensional Bose-Einstein condensate in different type of potentials including random. Such a system can be described by the solution of a non-linear Schrödinger equation also called a Gross-Pitaevskii equation. We also use the method for the inverse problem of predicting the non-linearity parameter using the ground state density profile for a given harmonic trapping potential.

Contents

1 Introduction and Motivation

Machine learning.
General usage area.
ML in physics and Phyiscs in ML.
ML&SE article.
Ours difference.

2 Gross Pitaevskii Equation

A Bose-Einstein Condensate (BEC) at zero temperature is described by Gross Pitaevskii Equation (GPE) also known as non-linear Schrodinger Equation (NLSE). It is mean field approximation of a quantum many body system which the hamiltonian of the system is given by;

$$\hat{H} = \sum_{i=1}^{N} \left(\frac{\boldsymbol{p}_i^2}{2m} + V(\boldsymbol{r}_i) \right) + \frac{1}{2} \sum_{i=1}^{N} \sum_{j \neq i}^{N} U(|\boldsymbol{r}_i - \boldsymbol{r}_j|) \tag{1}$$

where $\mathbf{p}_i$ i^{th} atom's momentum, $\boldsymbol{r}$ is position vector, m is mass, V is external potential and U is interaction between i^{th} and j^{th} atoms. Since the ground state energy is the only possible minimum energy, one can minimize this hamiltonian in order to obtain ground state energy. To do that, the mean field approximation is utilized to represent all bosons with the same wave function since the condensate is at zero temperature which allows to assume that all bosons are at the ground state [1]. Also the free energy $F = E - \mu N$ which is a thermodynamic potential is introduced since we try to minimize energy while still satisfying the normalization condition that is given as,

$$\int |\Psi(\boldsymbol{r})|^2 \, \mathrm{d}^3\boldsymbol{r} = N \tag{2}$$

where N is the number of particles in the system. This condition equals to minimizing free energy such that $\delta E - \mu \delta N = 0$ [1]. If we plug in the corresponding equations to this expression it becomes,

$$F(\Psi) = \int \Psi^* \hat{H} \Psi \, \mathrm{d}^3\boldsymbol{r} - \mu \int |\Psi|^2 \, \mathrm{d}^3\boldsymbol{r} \tag{3}$$

Therefore, the problem can be stated as minimization of the above equation [2]. Applying variation method to this equation while treating Ψ^* and Ψ as independent objects and using $U(|\mathbf{r}_i - \mathbf{r}_j|) = g\delta(\mathbf{r}_i - \mathbf{r}_j)$ one can obtain Gross Pitaevskii Equation (GPE) in stationary form as,

$$\frac{-\hbar^2}{2m}\nabla^2\psi + V(\boldsymbol{r})\psi + g|\psi|^2\psi = \mu\psi \tag{4}$$

where $\hbar$ is Planck constant, t is time, $\Psi(\boldsymbol{r}, t)$ is the wave function, m is mass, ∇^2 is the Laplacian operator, V is the potential, g is interaction parameter and it is defined as

$$g = \frac{4\pi\hbar^2 a_s}{m} \tag{5}$$

where, a_s is the s wave scattering length. Nonlinearity of the equation is caused by the cubic term $g|\Psi|^2$ which represents the interactions between bosons. The $|\Psi|^2$ term is interpreted as density, thus; there occurs an energy contribution caused by the mean field interactions. If there is no interaction GPE reduces to the Schrödinger Equation (SE) and becomes a linear equation. By definition, g can be positive or negative. If $g > 0$, it represents repulsive interaction, and if $g < 0$, it represents attractive interactions[3].

The ground state energy of the system can be obtained from the following expression,

$$\langle E\rangle = \int \Psi^* \hat{H}\Psi \, \mathrm{d}^3\boldsymbol{r} \tag{6}$$

$$E = \int \left(\frac{\hbar^2}{2m}|\nabla\Psi|^2 + V|\Psi|^2 + \frac{g}{2}|\Psi|^4\right) d^3\boldsymbol{r} \tag{7}$$

Here, the terms represent kinetic, potential and interaction energy respectively.[1] If the potential does not depend on time then the total energy of the system is conversed and a stationary form of GPE can be obtained by factorizing the wave function into spatial and time parts such that,

$$\Psi(\boldsymbol{r}, t) = \psi(\boldsymbol{r})e^{-i\mu t/\hbar} \tag{8}$$

where μ is chemical potential. In this case, by invoking separation of variables time independent GPE can be written as;

$$\frac{-\hbar^2}{2m}\nabla^2\psi + V(\boldsymbol{r})\psi + g|\psi|^2\psi = \mu\psi \tag{9}$$

[1]Contribition from the interaction energy is divided by two to eliminate double counting while pairing bosons.

2.1 General information about GPE

2.2 Why and how nonlinearty is introduced.

2.3 Physical and mathematical interpretation of interaction parameter. (phy: attractive, repulsive math:dominance of the terms)

2.4 Stationary form.

2.5 Potential, kinetic and interaction energy expressions.

2.6 Reduction of dimension.

2.7 Analytic solution and approximation.

There is no general solution for GPE and known analytic solutions exist only for few cases. In our study, infinite well and harmonic potential with zero interaction parameter has analytic solutions. The ground state energies are given by

$$\mu = \frac{1}{2}\hbar\omega \tag{10}$$

GPE generally solved by numerically or by approximation such as variational calculation or Thomas-Fermi approximation. **dilute**

2.7.1 Thomas Fermi Approximation

It is said that the second derivative term with respect to position represents the kinetic energy. When the potential and the interaction energy are dominant compared to the kinetic energy, the kinetic term can be neglected. This situation occurs when the condensate is large adequately and the interaction between bosons is repulsive. **REF dilute gases**. When the kinetic term is dropped, the new equation can be written as;

$$V(z)\psi(z) + g|\psi(z)|^2\psi(z) = \mu\psi(z) \tag{11}$$

this equation is analytically solvable and the solution is given by,

$$n(z) = |\psi(z)|^2 = \begin{cases} (\mu_{TF} - V)/g & \text{if } |z| \leq z_{TF} \\ 0 & otherwise \end{cases} \tag{12}$$

where z_{TF} is called Thomas-Fermi Radius. The density defined in this range and cannot be negative, therefore; if the normalization condition **EQREF** is applied

then the equation reads,

$$\int_{-\infty}^{\infty} |\psi|^2 \, \mathrm{d}z = \int_{-z_{TF}}^{z_{TF}} \frac{(\mu - V)}{g} \, \mathrm{d}z = N \tag{13}$$

For an exact analytic expression, one dimensional harmonic potential can be used,

$$\frac{1}{g} \left[\int_{-z_{TF}}^{z_{TF}} \mu \, \mathrm{d}z - \int_{-z_{TF}}^{z_{TF}} \frac{1}{2} m\omega^2 z^2 \, \mathrm{d}z \right] = N \tag{14}$$

$$\frac{2\mu z_{TF}}{g} - \frac{m\omega^2 z_{TF}^3}{3g} = N \tag{15}$$

From boundary conditions,

$$\mu = V(z_{TF}) = \frac{1}{2} m\omega^2 z_{TF}^2 \tag{16}$$

If we plug Eq. (16) the equations reads,

$$\frac{m\omega^2 z_{TF}^3}{g} - \frac{m\omega^2 z_{TF}^3}{3g} = N \tag{17}$$

$$\frac{4}{3} \left(\frac{2\mu}{m\omega^2} \right)^{1/2} \frac{\mu}{g} = N \tag{18}$$

$$\mu = \left(\frac{9}{32} (N\omega g)^2 m \right)^{1/3} \tag{19}$$

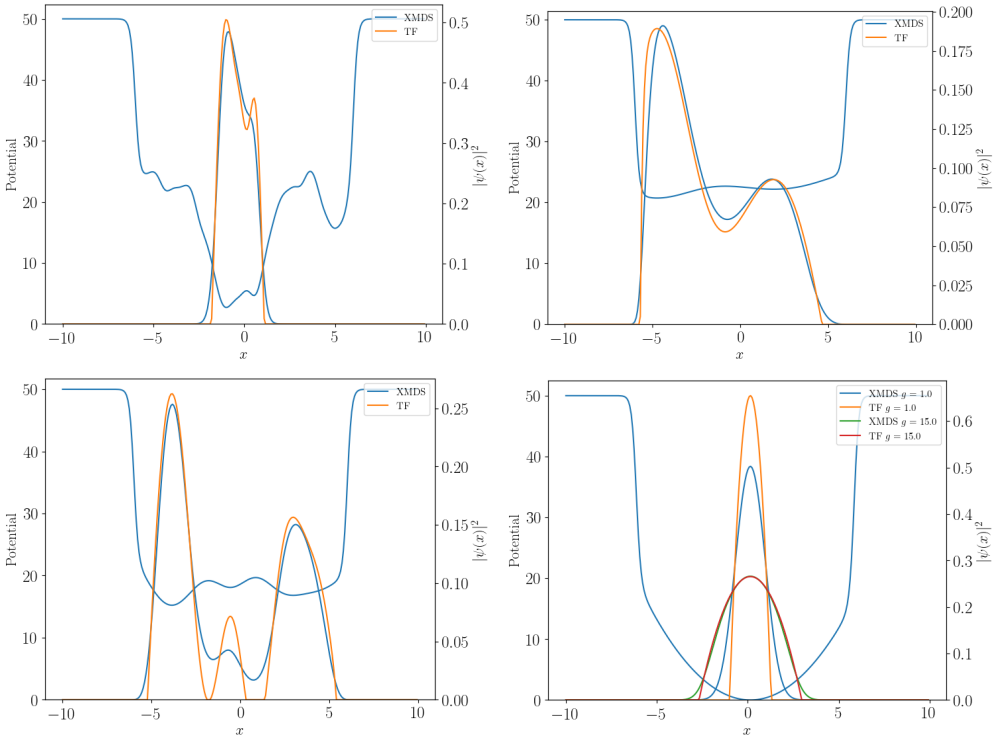


Figure 1: Here, the error between the numerical solution and approximation shrinks as the interaction parameter increases which corresponds to increment in interaction energy.

2.8 Numerical Solution and Dataset Generation

The dataset generation step is divided into two main parts and implemented independently. The first one is generating desired potential and the second one is numerical solution of the GPE under this potential by giving the generated potential to the numerical solution framework. The detailed description of the potential generation process is given in section 2.8.4.

In numerical solution part, we use a framework called XMDS [4], implemented specifically to solve differential equation systems with well optimized numerical methods. In this framework partial differential equation systems can be described by a markup language called XML. When equation system is declared properly, XMDS produces a source code written in C++ that solves the equation with specified numerical method. Because of modularity in our implementation, the framework only solves the equation by supplied parameters, the framework does not generate anything internally such as potential, even the scaling factors are supplied externally. Such a modularity has an huge advantage such that changing numerical solution framework does not effect potential generation step, therefore; the work to change numerical solution framework is minimized. The only require-

ment is implementation of file input output operations. By using this advantage, we also use another numerical solution framework called GPELab **GPELab Ref** implemented in Matlab to compare solutions' consistency and effect of scaling.

2.8.1 Scaling

The scaling of GPE is generally done according to potential type and there are more than one scaling conventions **REFS**. In our study, we use a more general scaling to investigate how different scalings affect the precision. To do that, we compare numerical solutions by representing same physical system with different scaling coefficients in both framework. Then we compare these results first internally and then we do a cross check.
We also compare numerical solutions of two different solvers implemented in different frameworks to show that the numerical solutions are **consistent**.

GPE is given as,

$$\frac{-\hbar^2}{2m}\frac{d^2\psi}{dz^2} + V(z)\psi + g|\psi|^2\psi = \mu\psi \tag{20}$$

First we define dimensionless potential, and then we make the length dimensionless,

$$\overline{V}(z) \equiv \frac{V(z)}{\gamma E_0}, \qquad \widetilde{z} \equiv \frac{z}{\beta L}$$

Here γ and β positive real numbers. E_0 is in energy unit and L is in length and they are defined respectively as;

$$E_0 = \frac{\hbar^2}{2m}$$

$$\widetilde{V}(\widetilde{z}) \equiv \overline{V}(\beta L z)$$

If these transformations are pluged into the Eq. (20) it becomes,

$$\frac{-\hbar^2}{2m\gamma E_0}\frac{1}{\beta^2 L^2}\frac{d^2\psi}{d\widetilde{z}^2} + \widetilde{V}(\widetilde{z})\psi + \frac{g}{\gamma E_0}|\psi|^2\psi = \frac{\mu}{\gamma E_0}\psi \tag{21}$$

To obtain final form, we define dimensionless energy, wave function and interaction parameter respectively.

$$\widetilde{\mu} \equiv \frac{\mu}{\gamma E_0}, \qquad \widetilde{\psi} \equiv \psi\sqrt{\frac{\beta L}{N}}, \qquad \widetilde{g} \equiv \frac{gN}{\gamma E_0 \beta L}$$

To control scaling coefficients we set the coefficient of the kinetic term to an

arbitrary positive real number α

$$\alpha = \frac{\hbar^2}{2m\gamma E_0}\frac{1}{\beta^2 L^2},$$

and set $E_0 = \hbar^2/2m$. Now the scaling of GPE can be controlled by α and β only.

$$-\alpha\frac{d^2\widetilde{\psi}}{d\widetilde{z}^2} + \widetilde{V}(\widetilde{z})\widetilde{\psi} + \widetilde{g}|\widetilde{\psi}|^2\widetilde{\psi} = \widetilde{\mu}\widetilde{\psi} \tag{22}$$

We are going to change these scaling coefficients and see their effects by comparing results. An example of setting the scale coefficients is shown in appendix A.

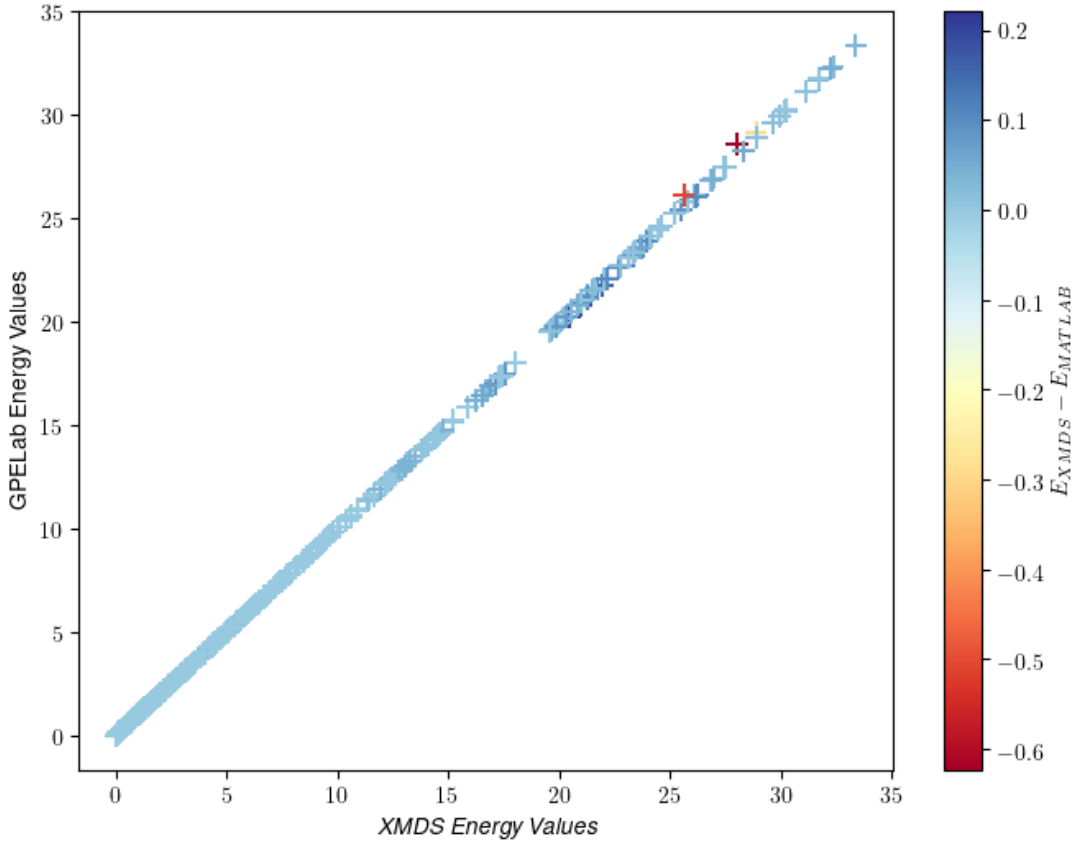


Figure 2: Error in density and energy

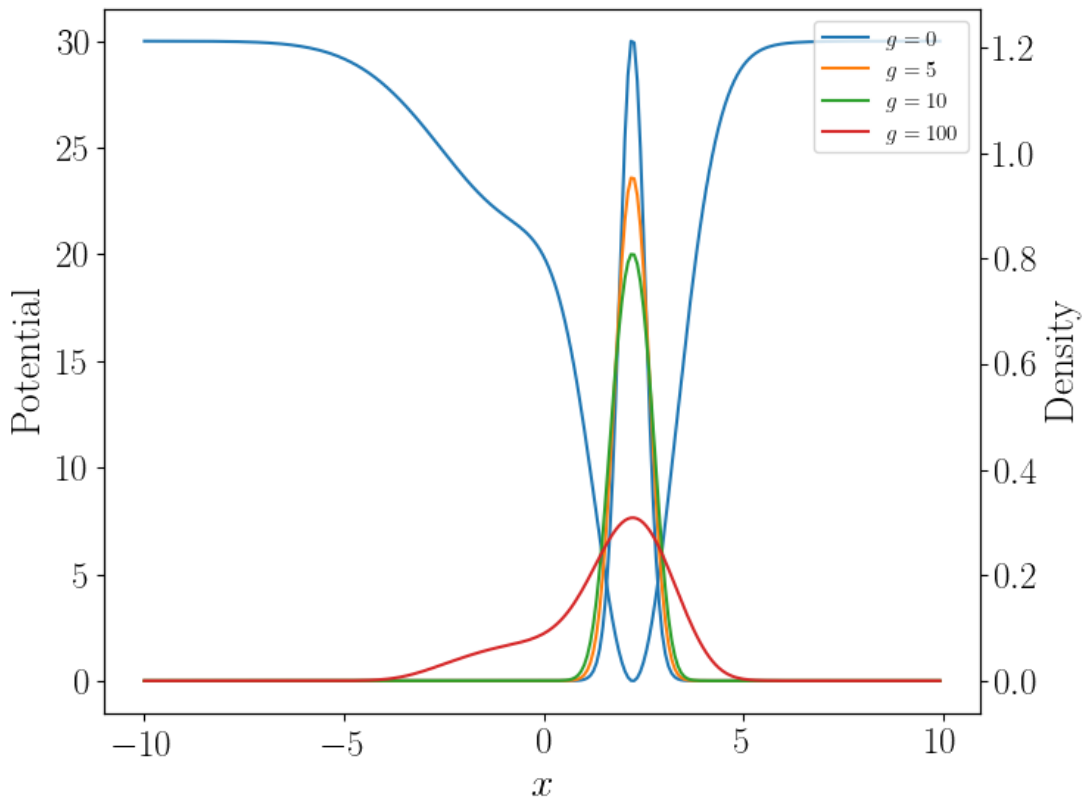


Figure 3: Density under Double Inverted Gaussian Potential with different interaction parameter.

2.8.2 Brief info about imaginary time evolution. (detailed in APPENDIX)

2.8.3 XMDS framerwork and other programs.

2.8.4 Potential generation

In potential generation, we built a modular structure such that the algorithms to generate potentials are independent from the restrictions forced by numerical techniques such as boundary conditions or scaling. The generated potentials are sent to another method supplied by module. This method rescales, and applies an envelope function to ensure that potential goes to numerical limit at boundaries. The given potential must satisfy the following conditions;

$$\begin{aligned} V(x < x_l) &= V_0 \\ V(x > x_r) &= V_0 \end{aligned} \tag{23}$$

To handle these two conditions, we define two envelope functions given by,

$$\mathrm{Env}_{LR}(x) = [(1 + \tanh(\beta(x + x_L))) + (1 - \tanh(\beta(x + x_R)))]/2 \tag{24}$$

$$\text{Env}_M(x) = 1 - \text{Env}_{LR}(x) \tag{25}$$

where x_L and x_R are bounds given in the Table (1) and the plot is given in Fig. 4.

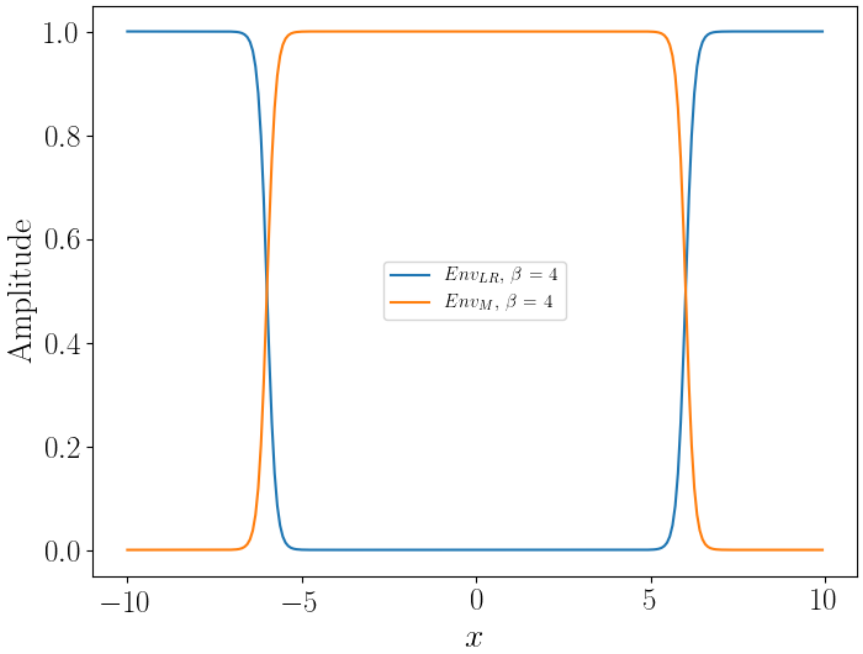
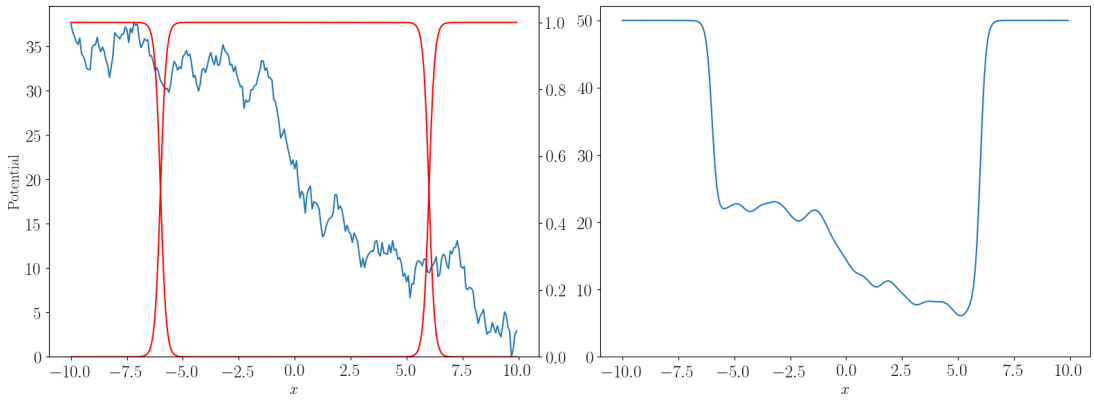


Figure 4: Envelope Function

The two potential in Figure 4 are combined and a gaussian filter applied.



2.8.5 Potential types (with analytic forms etc.)

We use **six** different types of potential in our study. The first three potential which have analytic form given in Table 1. are 1D version of potentials studied in [5]. We implement two more random potential generator to study effect of generation process. The generation processes of the random potentials are described in Section 2.8.6.

Table 1: Potentials

Potential	Analytic Form / Explanation	Parameters	Min	Max	Distribution
Envelope	$\text{Env}_{LR}(x) = [(1 + \tanh(\beta(x + x_L))) + (1 - \tanh(\beta(x + x_R)))]/2$ $\text{Env}_M(x) = 1 - \text{Env}_{LR}(x)$	x_L x_R β	-9 -4.5 -4.5	4.5 9 9	
Infinite Well	$V(x) = \begin{cases} 0 & \text{if } x_l < x < x_r \\ \infty & \text{if } \textit{otherwise} \end{cases}$	x_l x_r $x_r - x_l$	-9 -4.5 1	4.5 9 8	
Harmonic	$V(x) = \frac{1}{2}m\omega^2(x - x_0)^2$	ω x_0	0.01 -5	3 5	
DI Gaussian	$V(x) = -A_1 \exp(\frac{(x-\mu_1)^2}{\sigma_1^2}) - A_2 \exp(\frac{(x-\mu_2)^2}{\sigma_2^2})$	A_1, A_2 μ_1, μ_2 σ_1, σ_2	1 -5 0.5	10 5 4	
Random#1	$V(x_{i+1}) = V(x_i) + [X \sim \mathcal{N}(\mu, \sigma)]$	μ σ	-4 0.5	4 4	
Random#2	Summation of sines and cosines with random coefficients	Number of Terms A_1, A_2 n_1, n_2 σ	1 -4 -6.30 0.1	100 4 6.30 10	
Random#3	Substraction of two binary grid	Scale Factor	8	8	

2.8.6 Random potential generations with different method. (Reason)

To be able to observe the effect of random potential generation to the results we use three different random potential generation algorithms. The first one is random walk with random step size such that the first value of the potential array is initialized with a random number. After that, another random number is added to this value to obtain the next element of the array and so on. The distribution of the random numbers is gaussian in this process. The resultant array is not guaranteed to be smooth. Gaussian filter is applied with a random sigma value to the potential array to make it smooth.

Algorithm 1 RandomPotential1

procedure RANDOMPOTENTIAL1
    $Points = GaussianDistributedRandomPoints()$
    $Len = Length(Points)$
    $Potential[0] = Points[0]$
    for i $= 0$ **to** $Len - 1$ **do**
        $Potential[i+1] = Potential[i] + Points[i]$
    $Potential = GaussianFilter(Potential, \sigma)$

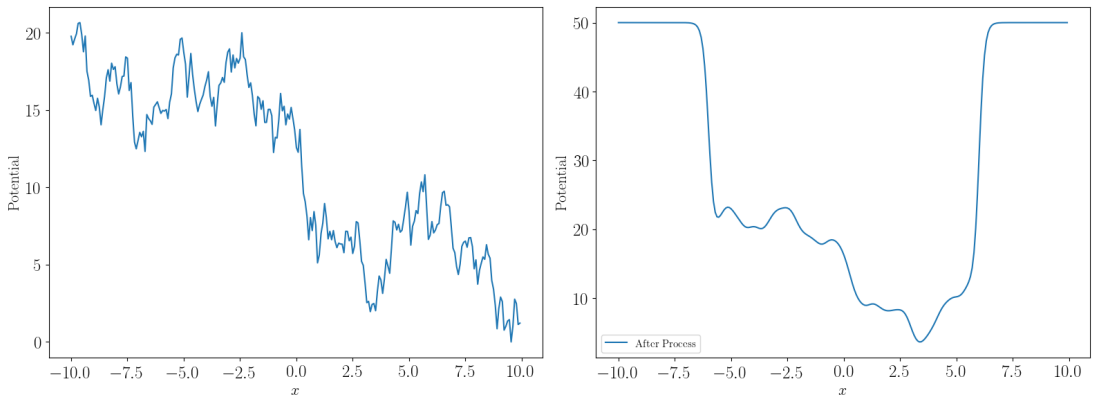


Figure 5: Random Potential 1 Before and After Process

The second one is summation of sines and cosines with random coefficients. Firstly, the number of terms are determined, after that each coefficient of the terms are assigned and summed iteratively.

Algorithm 2 RandomPotential2

procedure RANDOMPOTENTIAL2
    $Nterms = RandomInteger(1, 100)$
    for i $= 0$ **to** $Nterms$ **do**
        $A = GaussianDistributedRandomNumber()$
        $B = GaussianDistributedRandomNumber()$
        $n_1 = GaussianDistributedRandomNumber() * \sigma * \pi / width$
        $n_2 = GaussianDistributedRandomNumber() * \sigma * \pi / width$
        $Potential \mathrel{+}= A \sin(n_1 x) + B \cos(n_2 x)$
    $Potential = GaussianFilter(Potential, \sigma)$

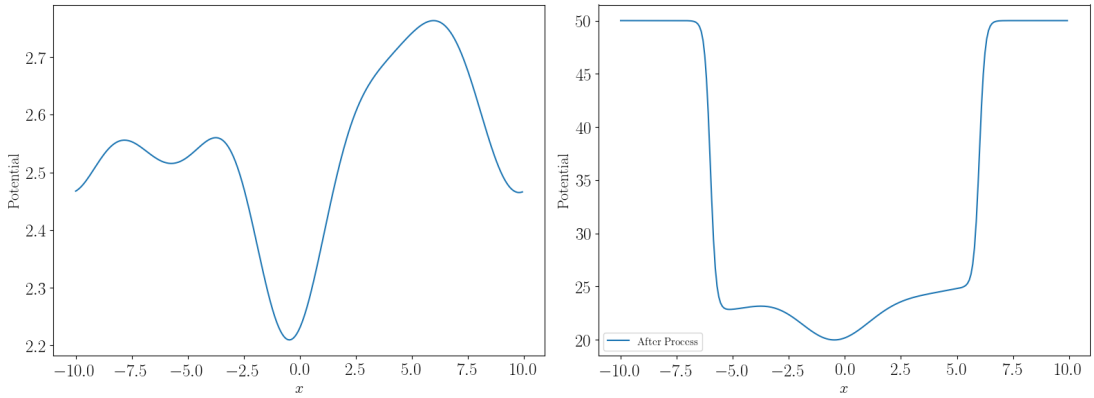


Figure 6: Random Potential 2 Before and After Process

The third one is one dimensional version of the algorithm described in [5] with slight modification. First a binary array of length 16 is generated by assigning random ones or zeros. Then the array is upscaled to 128 by repeating each element of the array 8 times. After that, another binary array of length 8 is generated with the same procedure and upscaled to 64 by repeating elements. Left and right paddings of length 32 is added to the second binary array to be able to do element wise substraction between two binary array. Then, the second array is substracted from the first one to obtain the potential in binary form. Finally, a gaussian filter is applied to make the potential smooth.

Algorithm 3 RandomPotentia3

procedure RANDOMPOTENTIAL3
    $ScaleFactor = 8$
    for i $= 0$ **to** $NumberOfPoints/ScaleFactor$ **do**
        $BinaryGrid[i] = RandomInteger(0, 1)$
    $BinaryGrid = RepeatElements(BinaryGrid, ScaleFactor)$
    for i $= 0$ **to** $NumberOfPoints/(ScaleFactor * 2)$ **do**
        $BinaryGrid2[i] = RandomInteger(0, 1)$
    $BinaryGrid2 = RepeatElements(BinaryGrid, ScaleFactor)$
    $Padding = Zeros((Length(BinaryGrid) - Length(BinaryGrid2))/2)$
    $BinaryGrid2 = Concatanate(Padding, BinaryGrid2, Padding)$
    $Potential = BinaryGrid - BinaryGrid2$
    $Potential = GaussianFilter(Potential, \sigma)$

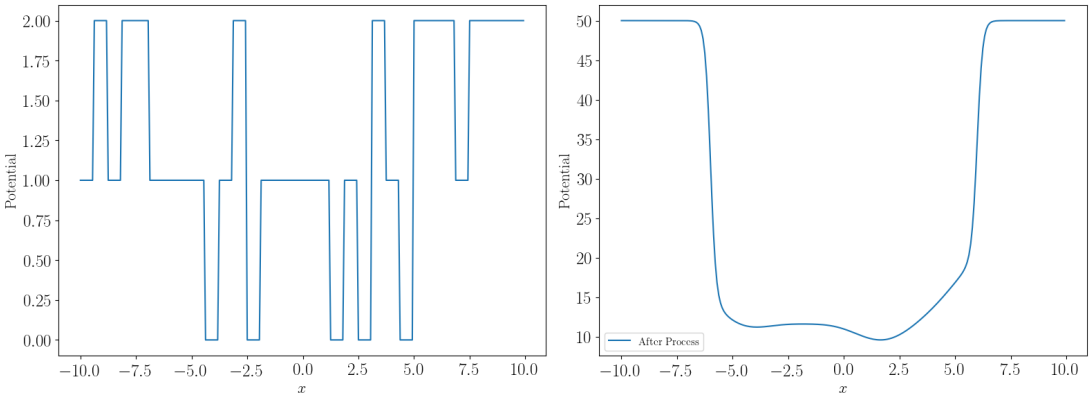


Figure 7: Random Potential 3 Before and After Process

2.8.7 Density and Ground State Energy

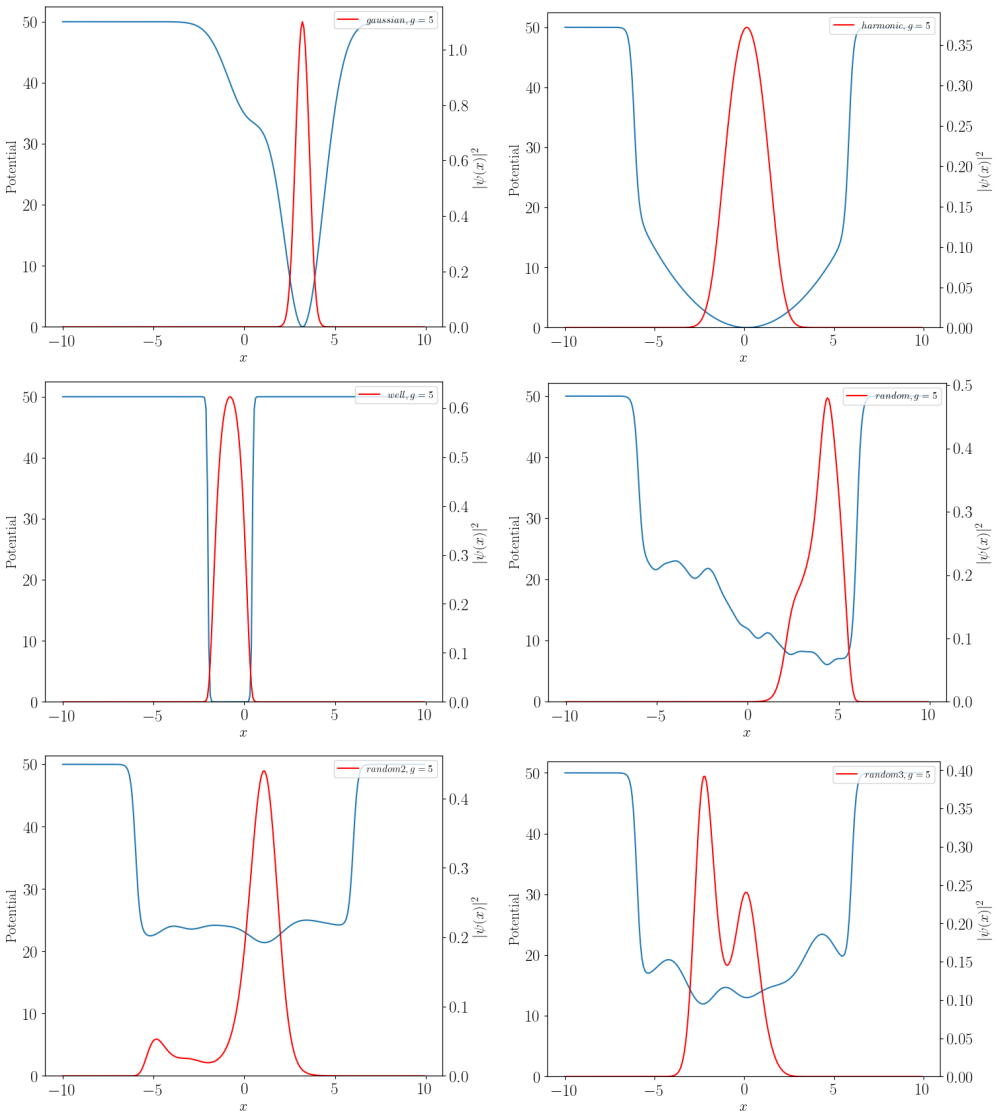


Figure 8: Potentials and Corresponding Ground State Densities

2.8.8 Boundaries. (Table)

2.8.9 Convergence (detailed in APPENDIX)

2.8.10 Dataset generation. (Total number of examples etc)

2.9 Dataset Features

2.9.1 Energy distribution

3 Machine Learning

3.1 Network architecture

Architecture of the network.
A general figure like in the ML&SE article that describes the work done.
Another figure about internals of the network such as number of layers, how interaction parameter is introduced to the network etc.
Hyperparameters.

3.2 Training

Detailed info about dataset (energy distribution etc).
Indicate that if there is any method to increase the number of examples in low and high energy values.

3.3 Results

4 Inverse Problem

5 Conclusion and Discussion

5.1 Conclusion

5.2 Discussion

5.3 Effects of random potential generation method

5.4 Are there problems in low and high energies compared to the mean

5.5 Inverse problem

References

[1] C. J. Pethick and H. Smith, *Bose-Einstein condensation in dilute gases*. Cambridge university press, 2002.

[2] J. Rogel-Salazar, "The gross–pitaevskii equation and bose–einstein condensates," *European Journal of Physics*, vol. 34, no. 2, p. 247, 2013.

[3] C. F. Barenghi and N. G. Parker, *A primer on quantum fluids*. No. arXiv: 1605.09580, Springer, 2016.

[4] G. R. Dennis, J. J. Hope, and M. T. Johnsson, "Xmds2: Fast, scalable simulation of coupled stochastic partial differential equations," *Computer Physics Communications*, vol. 184, no. 1, pp. 201–208, 2013.

[5] K. Mills, M. Spanner, and I. Tamblyn, "Deep learning and the schrödinger equation," *Physical Review A*, vol. 96, no. 4, p. 042113, 2017.

A APPENDIX A

Obtaining harmonic trap potential scaling from Eq. (22)

$$V(z) \rightarrow V(\widetilde{z}) \rightarrow \widetilde{V}(\widetilde{z}) \tag{26}$$

$$\widetilde{V}(z) \equiv \frac{V(z)}{\gamma E_0} \tag{27}$$

$$\widetilde{z} \equiv \frac{z}{\beta L} \tag{28}$$

$$V(z) = \frac{1}{2} m\omega^2 (z - z_0)^2 \tag{29}$$

$$V(\widetilde{z}) = \frac{1}{2} m\omega^2 \beta^2 L^2 (\widetilde{z} - \widetilde{z}_0)^2 \tag{30}$$

$$\widetilde{V}(\widetilde{z}) = \frac{1}{2} m\omega^2 \frac{\beta^2 L^2}{\gamma E_0} (\widetilde{z} - \widetilde{z}_0)^2 \tag{31}$$

The coefficient of this equation must be dimensionless, therefore;

$$\frac{1}{2} m\omega^2 \frac{\beta^2 L^2}{\gamma E_0} = C \tag{32}$$

Where C is a positive constant. We know that $E_0 = \frac{\hbar^2}{2m}$ and the definition of α is given as,

$$\frac{\hbar^2}{2m\gamma E_0}\frac{1}{\beta^2 L^2} = \alpha \tag{33}$$

thus;

$$\frac{1}{\gamma} = \alpha\beta^2 L^2 \tag{34}$$

if we plug Eq. (34) in to Eq. (32), then equation becomes,

$$\frac{m^2\omega^2}{\hbar^2}\alpha\beta^4 L^4 = C \tag{35}$$

in the case of expressing same physical system, m and ω must be constant. In that case $\alpha = C(\beta L)^{-4}$.

In this case, α becomes

$$\alpha = \frac{1}{2}\left(\frac{\hbar\omega}{\gamma E_0}\right)^2$$

Conventionally, α is set to $1/2$, therefore;

$$\hbar\omega = \gamma E_0$$

$$\beta L = \sqrt{\frac{\hbar}{m\omega}}$$

βL is generally defined as harmonic oscillator length ℓ

$$\ell = \sqrt{\frac{\hbar}{m\omega}}$$

$$\widetilde{\mu} = \frac{\mu}{\hbar\omega}$$

$$\widetilde{g} = \frac{g}{\hbar\omega\ell}$$

Finally, dimensionless GPE scaled for harmonic trapping potential can be written as,

$$\widetilde{\mu}\widetilde{\psi} = -\frac{1}{2}\frac{d^2\widetilde{\psi}}{d\widetilde{z}^2} + \frac{1}{2}\widetilde{z}^2\widetilde{\psi} + \widetilde{g}|\widetilde{\psi}|^2\widetilde{\psi}$$