Machine Learning for Prediction of Quantum Dynamics

Pritipriya Dasbehera
School of Physical Sciences
NISER, HBNI
Bhubaneswar
pritipriya.dasbehera@niser.ac.in

Abhishek Singh
School of Physical Sciences
NISER, HBNI
Bhubaneswar
abhishek.singh21@niser.ac.in

Abstract

We investigate data-driven approaches to reconstruct and predict quantum probability density evolutions from finite measurements sampled from the underlying probability density. Physics-Informed neural networks (PINNs) based on the continuity equation and Schrödinger equation have exhibited smearing effect beyond the training data. To overcome this, we employ the analytical expression to perform regression. For $V = 0$, a sequential optimizer strategy enabled accurate regressions ($MSE < 10^{-10}$), outperforming PINNs in interpolation and extrapolation with significantly small amount of input data. For $V \neq 0$, while analytical expressions were derived, the optimization techniques faced limitations.

1 Introduction

Quantum mechanics provides a foundational framework for understanding the behavior of microscopic systems. In classical mechanics, a state is represented by its position in space, x. We can predict the future values of x using Newton's laws. In contrast to classical mechanics, quantum mechanics characterizes the state of a system using a wave-function. This state of the system represented by the wave-function $\psi(x,t)$, offers insights into the probability of finding a particle at position x at time t, as encapsulated by $|\psi(x,t)|^2$, functioning essentially as a probability density. Consequently, measurements of position involve sampling from this distribution randomly. The time-evolution of this wave-function is governed by Schrödinger equation:

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

where:

- $\psi(x,t)$ is the wave function depending on position $\mathbf{r}$ and time t,
- $\hat{H}$ is the Hamiltonian representing the total energy of the system,
- $\hbar$ is the reduced Planck constant.

There are many systems like infinite potential well and Harmonic oscillator for which the Hamiltonian is known and analytical solutions can be derived. But, often times, the Hamiltonian of the system is itself unknown for eg. see Figure 1.

Our project is motivated by the work of M Casas et. al.[1] where the authors have used the concept of Fisher information to predict the pure state's wave-function from limited measurements of expectation values of an operator. We explore a central question: How can we infer or reconstruct wave-functions at discrete time instants, enabling the prediction of the time-evolution of such systems with limited measurement data.

Preprint. Under review.

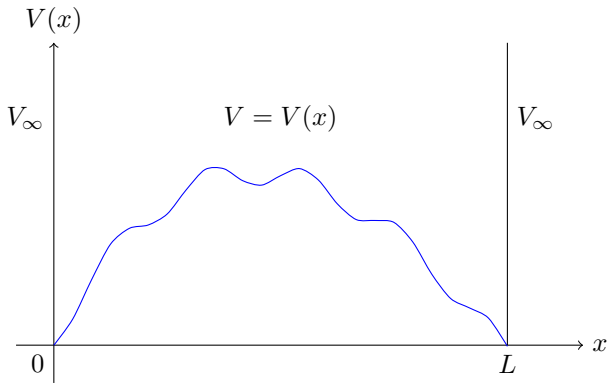

Figure 1: A complex potential for which Hamiltonian is unknown.

1.1 Formulating the problem statement

Given a set of 'n' finite measurements of an observable quantity at discrete time instants labeled as 1, 2, ... T, where each measurement is in theory a result of sampling from an underlying probability density represented by $|\psi(x,t)|^2 = \psi^*\psi$, the objectives are as follows:

- **Interpolation**: Reconstruct the underlying probability density at any intermediate time instant 't' that lies between any two consecutive measurement time instants. This involves developing a technique to interpolate the probability density based on the available measurements.
- **Extrapolation**: Extend the techniques to predict the evolution of the probability density beyond the final measurement time instant 'T'. In other words, extrapolate the wave-function to future time instants for which no measurements are available.

The ability to accurately reconstruct and predict continuous functions or probability density function from discrete measurements has broad applications in various scientific and engineering domains where continuous monitoring or high-resolution data acquisition may be impractical or infeasible.

1.2 Relevant works

In our understanding and formulation of our problem statement, we came across many seminal works that resonate with our work. Greydanus et. al. in paper on HNN [2] explains how one can extract Hamiltonian of a classical system using neural networks and use them to predict the evolution of states for longer period of time. Huang et al. in [3] explains how even if the quantum process involved is highly complex, there exist a low-dimensional effective Hamiltonian that can capture the dynamics effectively. This work tries to argue how a NN might be able to approximate Hamiltonian of a complex process. Yu Yao, Chao Cao et. al. in [7] explores the idea of training a neural network using easily generated physics-rich examples and applying the extracted knowledge to solve more complex cases not explicitly represented during training. It aims at generalizability of learning through various potential landscapes and use the knowledge to predict time-evolution in new landscape. Secor et.al. in [5] discuss the training of ANN as propagators for specific time-dependent potentials and time-evolution of states.

2 Data to Probability density function

In our problem setting, we are provided with a set of finite measurements of an observable quantity at discrete time instants labeled as $1, 2, \ldots, T$. As described earlier, these measurements are obtained by sampling from the underlying probability density function given by the square of the wave-function at each time instant, i.e., $|\psi(x,t)|^2$.

To estimate the approximate probability density from the available data, we employ a non-parametric density estimation technique known as kernel density estimation (KDE). KDE is a widely used

method for constructing a smooth, continuous probability density function from a finite set of sample points. In our implementation, we choose the Gaussian kernel function:

$$K(u) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}u^2\right) \tag{2}$$

It is important to note that while the kernel density estimation provides a smooth approximation of the underlying probability density, it may introduce some inherent bias and variance due to the choice of kernel function. However, this non-parametric approach offers flexibility and avoids assumptions about the specific form of the probability density, making it a suitable choice for our problem where the true distribution is unknown.

With the approximate probability density obtained from the experimental data, we can proceed to develop models and techniques for reconstructing and predicting the wave-function evolution, leveraging both the data-driven and physics-based components of our approach.

3 Baseline Approach: Physics-Informed Neural Networks (PINNs)

3.1 Background and Motivation

Physics-Informed Neural Networks (PINNs), introduced by Raissi and Karniadakis [4], have emerged as a powerful technique for solving ordinary and partial differential equations using neural networks by leveraging initial or boundary conditions. PINNs combine the ability of neural networks to learn complex mappings from data with the incorporation of known physical laws and constraints encoded as differential equations. This approach has shown promising results in various domains, including fluid dynamics, heat transfer, and classical mechanics.

Motivated by the success of PINNs in classical physics, we sought to explore their applicability in the realm of quantum mechanics, specifically for the reconstruction and prediction of wave-function evolutions. The overarching goal was to develop a data-driven approach that could accurately estimate the underlying probability density at any desired time instant, given a finite set of measurements. This would enable us to leverage the available data and physical principles to make predictions beyond the measurement range, a crucial requirement in scenarios where continuous monitoring or high-resolution data acquisition is impractical or infeasible.

3.2 Initial Attempts with Governing Equations

3.2.1 Schrödinger Equation

Our initial attempt involved using the time-dependent Schrödinger equation as the governing differential equation within the PINN framework. However, this approach required prior knowledge of the potential function $V(\mathbf{x})$ in the Hamiltonian, which was not available in our problem formulation from the outset. Without this information, it was not possible to fully specify the Schrödinger equation, prompting us to explore alternative governing equations.

3.2.2 Continuity Equation for Probability Current Conservation

As an alternative, we explored the use of the continuity equation for the conservation of probability current, given by:

$$\frac{\partial|\psi(x)|^2}{\partial t} + \frac{\partial J}{\partial x} = 0 \qquad \text{where,} \qquad J = \frac{i\hbar}{2m}\left[\psi\frac{\partial\psi^*}{\partial x} - \psi^*\frac{\partial\psi}{\partial x}\right] \tag{3}$$

This differential equation, derived from the Schrödinger equation, governs the conservation of probability density $|\psi(x,t)|^2$ and the probability current density $J(x,t)$. It is satisfied by the wave-function evolution for any potential energy function $V(\mathbf{r},t)$.

While this approach seemed promising initially, as it did not require explicit knowledge of the potential function, we soon realized a critical limitation: the continuity equation alone does not provide enough constraints to uniquely identify the wave-function for specific potential. Consequently,

the PINN could predict probability density evolutions that satisfied the continuity equation but did not necessarily correspond to the desired potential.

3.2.3 Incorporating Data Constraints

To address the limitations of the continuity equation approach, we introduced additional constraints by incorporating the available measurement data into the loss function. Specifically, we provided the PINN with measurement data up to 0.6T, where T is the final measurement time instant. By combining the differential equation loss with a data loss term, we aimed to guide the PINN towards reconstructing the probability density evolution consistent with both the physical laws (governed by the continuity equation) and the measured data.

The data loss term should ensure that the PINN's predictions adhered to the provided measurements within the training range, while the differential equation loss encouraged the model to learn solutions that satisfied the continuity equation. By combining these two loss components, we hoped to leverage the strengths of both data-driven and physics-based approaches, ultimately leading to accurate reconstructions and predictions of the probability density evolution.

3.3 Re-investigation and Validation

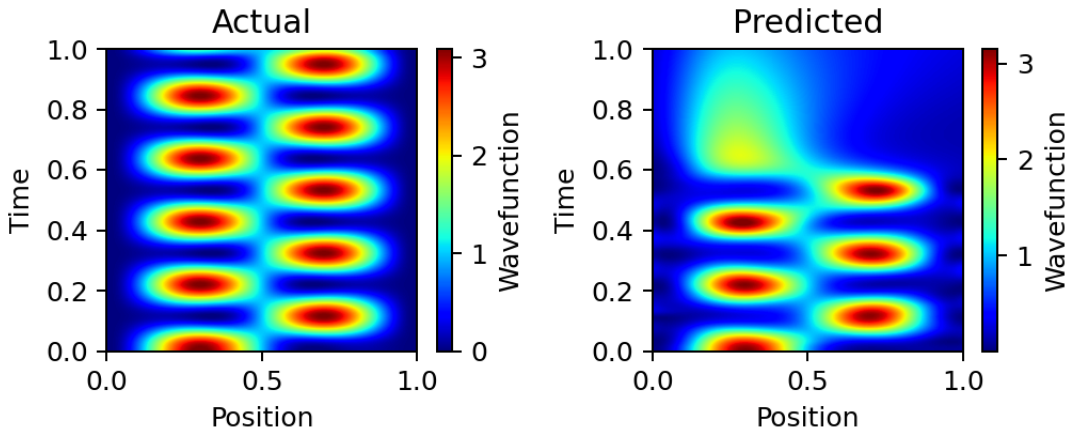

Figure 2: **Failure of PINN:** The model was given data of probability density upto time instant 0.6 T. Notice how the model interpolates well till 0.6 T. But, beyond the training region, we see an unexpected smearing of probability density.

Despite incorporating the data constraints, the PINN approach with the continuity equation exhibited an unexpected smearing behavior in the predicted probability density evolution beyond the training data range as shown in Figure 2. This behavior was surprising, as the incorporation of the physics-based differential equation loss should have contributed to better generalization and physically consistent predictions.

In an attempt to address this issue, we did an hyperparameter tuning process. We explored various strategies, including adjusting the relative weights of the data and equation loss terms, modifying the neural network architecture by varying the number of layers and neurons, and experimenting with different optimization algorithms and learning rates. Despite these efforts, the smearing behavior persisted, and we were unable to find a satisfactory solution within the PINN framework.

3.3.1 Verifying the Continuity Equation Solution

To understand the root cause of this issue, we first aimed to verify whether the smeared probability density solutions obtained from the continuity equation approach were indeed satisfying the governing differential equation. Surprisingly, upon further investigation, we found that the smeared wave-function predictions still satisfied the continuity equation for the conservation of probability current, albeit with undesirable characteristics such as the loss of localization and spreading of the probability density.

3.3.2 Validating the PINN Framework

This observation led us to question the effectiveness of the PINN framework itself in handling the specific problem at hand. To validate the PINN's capabilities, we implemented it for the Schrödinger equation with a (specified) zero potential (V = 0), where the expected solution should exhibit oscillatory behavior within the region of interest.

However, in this case, the PINN struggled to form travelling oscillatory solutions even with use of sinusoidal activation function [6], further raising doubts about its ability to handle second-order differential equations effectively, particularly in the context of quantum mechanics.

3.3.3 Testing on Classical Equations

As a validation step, we applied the PINN framework to classical equations like the advection equation, where it performed as expected, accurately reproducing the known solutions. This suggested that the PINN might have inherent limitations in learning solutions to specific types of differential equations encountered in the quantum mechanics domain, such as the Schrödinger equation or the continuity equation for probability current conservation.

3.4 Outcome and Next Steps

Despite our efforts to resolve the smearing issue and validate the PINN framework for our problem, we were ultimately unable to find a satisfactory solution within the given time constraints. Consequently, we decided to explore alternative strategies and approaches, which will be discussed in the following section.

4 Regression

We started the new approach of regression with the restriction that potential inside the well is zero. Then, the problem is simply solving the time independent Schrodinger equation for the same,

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) = E\psi(x) \tag{4}$$

and then writing the final solution to time dependent Schrodinger equation 1 as,

$$\psi(x,t) = \psi(x)e^{-iEt/\hbar} \tag{5}$$

This is a well known problem with solution,

$$\psi(x,t) = \sum_{n=1}^{N} a_n \psi_n^\circ(x) e^{-i(E_n^\circ t + \phi_n)} \tag{6}$$

$$\text{where,} \qquad E_n^\circ = \frac{1}{2m}\left(\frac{n\pi\hbar}{L}\right)^2 \qquad \text{and} \qquad \psi_n^\circ(x) = \sqrt{\frac{2}{L}}\sin\left(\frac{n\pi x}{L}\right) \tag{7}$$

The known data for regression is a grid of $|\psi(x,t)|^2$ values, so we square the analytic solution to arrive at[1],

$$|\psi(x,t)|^2 = \frac{2}{L}\sum_{m,n=1}^{N} a_n a_m \sin\left(\frac{n\pi x}{L}\right)\sin\left(\frac{m\pi x}{L}\right)\cos\left((E_n^\circ - E_m^\circ)t + (\phi_n - \phi_m)\right) \tag{8}$$

Now the task at hand is to determine the parameters a_n and ϕ_n, of the order $\mathcal{O}(2N)$, from data. The problem has been reduced to that of regression. We implemented the above using PyTorch and used common optimisation techniques to minimise the Mean Squared Error(MSE) between prediction from analytic solution and the given data.

[1] derivation is provided at the appendix

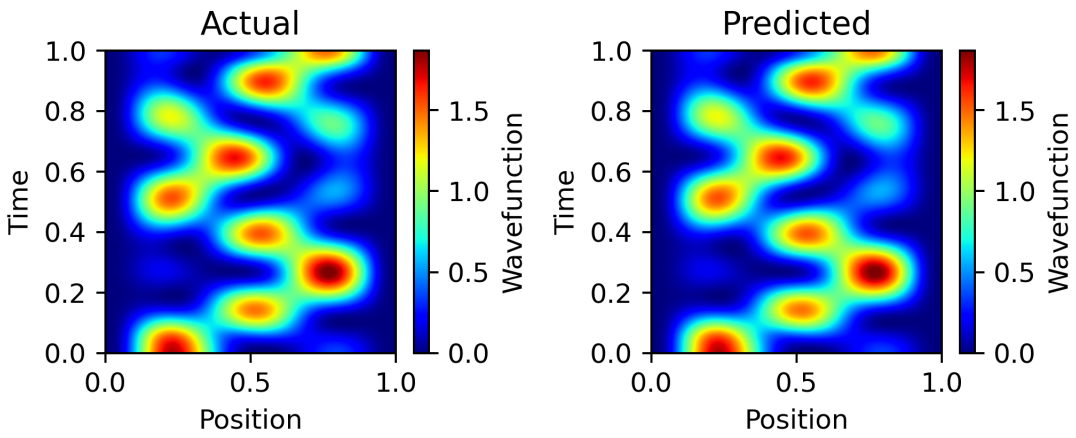

Figure 3: **Success of Regression:** Here the model was given data upto just 0.2T. Notice the striking similarity between the actual and predicted dynamics of probability density.

- **LBFGS:** This uses ... It is good at finding the descent path efficiently but can't jump out of local minima. In our case sometime it worked (bringing down the MSE to 1e-10) but there were times when it got stuck at a local minima (with MSE 1e-7).
- **Adam:** It has a very slow convergence rate, but it doesn't normally get stuck at local minima. In our case it took hours to bring the MSE down to 1e-4.

4.1 Sequential optimisation strategy

After much trial we discovered that an reliable optimisation technique for our case was to use LBFGS and Adam optimiser in succession. This solved the problem for zero potential case. An example problem and the prediction is presented in Figure 3.

```
while MSE decreasing:
    Run LBFGS

# If MSE has decreaseed below 1e-10 then we can skip Adam else,
   perform Adam till convergence
while MSE > 1e-10:
    Run Adam
```

Listing 1: "Pseudo-code of Optimisation"

4.2 Extension to arbitrary potential

Although we don't know the Schrodinger equation for an arbitrary non-zero potential, if we assume it is independent of time then we can say that the solution has the form,

$$\Psi_n(x,t) = \sum_{n=1}^{N} a_n \psi_n(x) e^{-i(E_n t + \phi_n)} \tag{9}$$

Now using the completeness[2] of eigenstates of zero potential,

$$\psi_n(x) = \sum_{m=1}^{N} b_{nm} \psi_m^{\circ}(x) e^{i\theta_{nm}} \tag{10}$$

[2]Any function with domain 0 to L can be expanded as a linear combination of the functions $\psi_n^{\circ}(x)$.

Hence, the square of analytic solution is given by[3],

$$|\psi(x,t)|^2 = \frac{2}{L} \sum_{m,n,p,q=1}^{N} a_m b_{mn} a_p b_{pq} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{q\pi x}{L}\right) \cos\left((E_m - E_p)t + (\phi_m + \theta_{mn} - \phi_p - \theta_{pq})\right) \tag{11}$$

The problem has again reduced to determining the parameters a_n, b_{nm}, E_n, ϕ_n and θ_{nm}, of the order $\mathcal{O}(2N^2 + 3N)$, from data. The regression model for arbitrary potential was implemented as well but we were unable to find a reliable optimisation technique to find the parameters. We have discovered that if the energies (E_n) are fixed then LBFGS + Adam can converge to the true solution but without any such assumptions it generally doesn't work.

5 Future work

For future work, several promising directions can be explored. Firstly, resolving the issues encountered with the regression model for cases where V$\neq 0$ is a priority. Concurrently, efforts should be made to improve the PINN approach for cases where the potential is simply, V = 0. By drawing inspiration from the work on Hamiltonian Neural Networks (HNN) [2], the PINN framework can be extended to compute the non-zero potential energy functions more effectively.

If successful in addressing the above challenges, a natural extension would be to explore the applicability of the developed techniques to time-dependent potential. This would significantly broaden the scope of the project, enabling the reconstruction and prediction of wave-function evolutions in dynamic systems with time-varying potential landscapes.

6 Conclusion

In this work, we explored the application of machine learning techniques to predict the time-evolution of quantum wave-functions. We investigated two primary approaches: Physics-Informed Neural Networks (PINNs) and regression.

Our initial attempts with PINNs employed the continuity equation for probability current conservation as the governing differential equation. While this approach incorporated physics-based constraints, it resulted in unexpected smearing of the predicted wave-function evolution beyond the training data range. We also tried to use free particle Schrodinger equation (V=0), but that too didn't work. We suspect that PINN has limitations in solving second-order partial differential equations.

The regression approach, implemented for a zero potential case, achieved promising results. By leveraging the known analytical solutions for the free particle time-independent Schrödinger equation and employing a regression model to fit the data points, we were able to accurately predict the wave-function dynamics. We tried to extend this approach to account for arbitrary potentials by incorporating the completeness of known eigenstates for the zero-potential case to decompose the unknown eigenstates of unknown potential. However, further work is required to address the challenges of optimisation to find the parameters associated with non-zero potentials.

Acknowledgments and Disclosure of Funding

We thank the course instructor, Dr. S Mishra, for their guidance and support. We are grateful for the computing resources provided by the NISER, to conduct extensive experiments and simulations.

[3]The derivation is provided at the appendix

References

[1] M. Casas, A. Plastino, and A. Puente. Fisher information and the inference of wave functions for systems of unknown hamiltonian. *Physics Letters A*, 248(2):161–166, 1998. ISSN 0375-9601. doi: https://doi.org/10.1016/S0375-9601(98)00663-X. URL https://www.sciencedirect.com/science/article/pii/S037596019800663X.

[2] S. Greydanus, M. Dzamba, and J. Yosinski. Hamiltonian neural networks. *Advances in neural information processing systems*, 32, 2019. doi: https://doi.org/10.48550/arXiv.1906.01563. URL https://arxiv.org/abs/1906.01563.

[3] H.-Y. Huang, S. Chen, and J. Preskill. Learning to predict arbitrary quantum processes. *PRX Quantum*, 4: 040337, Dec 2023. doi: 10.1103/PRXQuantum.4.040337. URL https://link.aps.org/doi/10.1103/PRXQuantum.4.040337.

[4] M. Raissi, P. Perdikaris, and G. E. Karniadakis. Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations, 2017. URL https://arxiv.org/abs/1711.10561.

[5] M. Secor, A. V. Soudackov, and S. Hammes-Schiffer. Artificial neural networks as propagators in quantum dynamics. *The Journal of Physical Chemistry Letters*, 12(43):10654–10662, Oct. 2021. ISSN 1948-7185. doi: 10.1021/acs.jpclett.1c03117. URL http://dx.doi.org/10.1021/acs.jpclett.1c03117.

[6] V. Sitzmann, J. N. P. Martel, A. W. Bergman, D. B. Lindell, and G. Wetzstein. Implicit neural representations with periodic activation functions, 2020. URL https://arxiv.org/abs/2006.09661.

[7] Y. Yao, C. Cao, S. Haas, M. Agarwal, D. Khanna, and M. Abram. Emulating quantum dynamics with neural networks via knowledge distillation. *Frontiers in Materials*, 9:1060744, 2023. doi: https://doi.org/10.3389/fmats.2022.1060744. URL https://www.frontiersin.org/articles/10.3389/fmats.2022.1060744/full.

A Supplementary Material

The code pertaining to results in the paper can be found at GitHub

A.1 Derivation of Equation 8

The solution of Schrodinger equation for $V = 0$ case is,

$$\psi(x,t) = \sqrt{\frac{2}{L}} \sum_{n=1}^{N} a_n \sin\left(\frac{n\pi x}{L}\right) e^{-i(E_n^\circ t + \phi_n)} \tag{12}$$

Then,

$$|\psi(x,t)|^2 = \frac{2}{L} \sum_{m,n=1}^{N} a_n a_m \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) e^{-i((E_n^\circ - E_m^\circ)t + (\phi_n - \phi_m))} \tag{13}$$

Since m,n are summation indices one can exchange them and then add the two equations to arrive at, (let $\Delta = (E_n^\circ - E_m^\circ)t + (\phi_n - \phi_m)$)

$$2 \times |\psi(x,t)|^2 = \frac{2}{L} \sum_{m,n=1}^{N} a_n a_m \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) \left\{e^{i\Delta} + e^{-i\Delta}\right\} \tag{14}$$

$$= 2 \times \frac{2}{L} \sum_{m,n=1}^{N} a_n a_m \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) \left\{\frac{e^{i\Delta} + e^{-i\Delta}}{2}\right\} \tag{15}$$

$$= 2 \times \frac{2}{L} \sum_{m,n=1}^{N} a_n a_m \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{m\pi x}{L}\right) \cos\left(\Delta\right) \tag{16}$$

A.2 Derivation of Equation 11

Similarly solution for $V \neq 0$ is, (here $\Delta = (E_m - E_p)t + (\phi_m + \theta_{mn} - \phi_p - \theta_{pq})$)

$$\psi_n(x) = \sum_{m=1}^{N} b_{nm} \psi_m^\circ(x) e^{i\theta_{nm}} \tag{17}$$

$$|\psi(x,t)|^2 = \frac{2}{L} \sum_{m,n,p,q=1}^{N} a_m b_{mn} a_p b_{pq} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{q\pi x}{L}\right) e^{i\Delta} \tag{18}$$

$$= \frac{2}{L} \sum_{m,n,p,q=1}^{N} a_m b_{mn} a_p b_{pq} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{q\pi x}{L}\right) \left\{\frac{e^{i\Delta} + e^{-i\Delta}}{2}\right\} \tag{19}$$

$$= \frac{2}{L} \sum_{m,n,p,q=1}^{N} a_m b_{mn} a_p b_{pq} \sin\left(\frac{n\pi x}{L}\right) \sin\left(\frac{q\pi x}{L}\right) \cos\Delta \tag{20}$$

$$\tag{21}$$