# Equation-driven Neural Networks for Periodic Quantum Systems

**Circe Hsu**
Department of Mathematics
Northeastern University
hsu.circe@northeastern.edu

**Marios Mattheakis**
School of Engineering and Applied Sciences
Harvard University
mariosmat@gmail.com

**Gabriel R. Schleder***
School of Engineering and Applied Sciences
Harvard University
gabriel.schleder@lnnano.cnpem.br

**Daniel T. Larson**
Department of Physics
Harvard University
dtlarson@g.harvard.edu

## Abstract

Deep learning equation-driven approaches, also known as physics-informed neural networks (PINNs), have seen a wave of success in modeling physical systems governed by differential equations. However, these techniques have rarely been applied to quantum systems, where traditional numerical methods provide accurate solutions but become computationally expensive for large-scale systems. We explore a neural network approach capable of solving the Schrödinger equation for two-dimensional systems of periodic potentials. Applying efficient sampling and normalization constraints allows the simultaneous discovery of the energy band structure and the associated wavefunctions, that are crucial, for example, to determine the properties of electronic, photonic, and metamaterials systems.

## 1 Introduction

Since their introduction [1], physics-informed neural networks (PINNs) have been investigated as a viable alternative to traditional numerical techniques to solve partial differential equations (PDEs) [2, 3] for various applications in mathematics [4, 5] and the physical sciences [6–9]. PINNs techniques leverage the universal approximation property of neural networks [10] to approximate solutions to PDEs. PINNs have been investigated as a solution to dimensionality scaling issues present with traditional numerical solvers [11, 12], and possess desirable interpolation and differentiability properties which allow for efficient error computation and optimization with gradient-based methods [13].

**Main contributions:** We adapt the PINN approach to solve the time-independent Schrödinger equation for a particle in a periodic potential, which is an eigenvalue problem relevant for many physical systems such as electrons in bulk and nanoscale crystalline structures. We define a neural network architecture that simultaneously learns the complex-valued wavefunctions and their associated real eigenvalues as a function of both the dual real space position and its reciprocal crystal momentum. By selectively sampling disjoint regions in momentum space a single network is able to learn multiple wavefunctions and eigenvalues corresponding to each irreducible $\mathbf{k}$ value. We demonstrate our approach with two example potentials.

*Additionally, Brazilian Nanotechnology National Laboratory (LNNano/CNPEM)

Machine Learning and the Physical Sciences Workshop, NeurIPS 2024.

## 2 Background

The time-independent SE for the wavefunction describing a single particle moving in an external potential $\mathcal{V}(r)$ is written:

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + \mathcal{V}(\mathbf{r})\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}). \tag{1}$$

Below we choose units where $\hbar = m = 1$. Our goal is to extend previous PINN methods [14, 15] to periodic systems. Specifically, we want to solve Eq. (1) for a potential which satisfies the following periodicity condition:

$$\mathcal{V}(\mathbf{r}) = \mathcal{V}(\mathbf{r} + \mathbf{R}), \quad \forall\, \mathbf{R} = \sum_i n_i \mathbf{a}_i, \quad \text{with } n_i \in \mathbb{N}, \tag{2}$$

where $\{\mathbf{a}_i\}$ are the set of basis vectors that define the periodicity of the potential in real space. Bloch's theorem [16] states that for such a potential, $\psi$ can always be written:

$$\psi_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{\mathbf{k}}(\mathbf{r}), \tag{3}$$

as the product of a plane wave and a *Bloch function*, $u_{\mathbf{k}}(\mathbf{r})$, which has the same periodicity as the potential $\mathcal{V}(\mathbf{r})$. The *crystal momentum* $\mathbf{k}$ is a parameter in the reciprocal space to the periodic real-space lattice. The "unit cell" of the reciprocal space is known as the Brillouin zone. Substituting Eq. (3) into Eq. (1) we obtain a modified SE that depends parametrically on $\mathbf{k}$:

$$\left[-\frac{1}{2}\nabla_{\mathbf{r}}^2 + i\mathbf{k}\cdot\nabla_{\mathbf{r}} - \frac{1}{2}\mathbf{k}^2 - \mathcal{V}(\mathbf{r})\right] u_{\mathbf{k}}(\mathbf{r}) = E_{\mathbf{k}} u_{\mathbf{k}}(\mathbf{r}). \tag{4}$$

This is an eigenvalue equation for $u_{\mathbf{k}}(\mathbf{r})$ with eigenvalue $E_k$. To solve this PDE using a neural network approach we define the following feed-forward networks:

$$u_{\mathbf{k}}(\mathbf{r}) = \mathcal{N}_u : (r, k) \to \mathbb{C} \tag{5}$$

$$E_{\mathbf{k}} = \mathcal{N}_E : (k) \to \mathbb{R}. \tag{6}$$

Note that Eq. (4) is only satisfied when $\mathcal{N}_u$ and $\mathcal{N}_E$ are exact solutions, so we can treat it as a loss term and utilize gradient descent to find the weights that minimize the following expression:

$$L_{eq} = \left\langle \left( \left[-\frac{1}{2}\nabla_{\mathbf{r}}^2 + i\mathbf{k}\cdot\nabla_{\mathbf{r}} - \frac{1}{2}\mathbf{k}^2 - \mathcal{V}(\mathbf{r}) - \mathcal{N}_E(k)\right] \mathcal{N}_u(r, k) \right)^2 \right\rangle, \tag{7}$$

where $\langle \cdots \rangle$ denotes averaging with respect to the inputs $\mathbf{r}$ and $\mathbf{k}$, and gradients are taken using automatic-differentiation provided by PyTorch. We introduce additional loss terms to enforce relevant constraints, as described in the next section. We additionally incorporate weighted sampling of the reciprocal space to sample points more efficiently along "high symmetry" boundaries, which is further discussed in the appendix B. Finally, we show that an equation-driven approach may solve for multiple electronic states simultaneously, producing accurate results even at higher energy levels.

## 3 Methods

We refer to the two feed-forward networks, $\mathcal{N}_u$ and $\mathcal{N}_E$, as the eigenfunction unit and energy unit. The role of each unit is to learn part of the solution to the SE for a given potential. In practice, we define the eigenfunction unit to have two real-valued outputs, representing the real and imaginary components of the complex-valued $u_{\mathbf{k}}(\mathbf{r})$.

### 3.1 Loss Formulation

In addition to the loss term of Eq. (7) based on the differential equation, we also define boundary conditions (BC) and normalization (Norm) constraints which are important for a physically meaningful solution.

$$L_{bc} = \left|\mathcal{N}_u(\mathbf{r} + \mathbf{a}, k) - \mathcal{N}_u(\mathbf{r}, k)\right|^2 + \left|\nabla_r \mathcal{N}_u(\mathbf{r} + \mathbf{a}, k) - \nabla_r \mathcal{N}_u(\mathbf{r}, k)\right|^2, \tag{8}$$

where $a$ is any lattice vector. The assumption that Bloch states' first derivatives must match along cell boundaries stems from our assumption that the potential $\mathcal{V}(\mathbf{r})$ is finite. To avoid the trivial solution, $u_{\mathbf{k}}(\mathbf{r} = 0)$ impose a normalization constraint on the Bloch states across a unit cell:

$$L_{norm} = \left\langle \left(1 - \int_{Cell} \mathcal{N}_u^*(r,k)\mathcal{N}_u(r,k)\right)^2 \right\rangle. \tag{9}$$

This has the added benefit that the solution can be immediately interpreted as a probability density. Both loss terms are evaluated with fixed $k$, and in practice it is necessary to compute these regularization terms for many values of $k$ to effectively impose these constraints across all Bloch states throughout training.

## 4 Experiments

We evaluate our method on two different 2-dimensional potentials defined on a square unit cell in real space. The momentum parameter $\mathbf{k}$ is thus also defined on a 2-dimensional reciprocal space. The eigenfunction unit is initialized as a neural network with 4 layers of 384 neurons, and the energy unit is initialized as a 4-layer 128-neuron network; all the hidden neurons use the SiLU activation function. The first experiment explores a constant potential $\mathcal{V}(\mathbf{r}) = c$. Our second experiment evaluates our model on a pseudopotential approximating an atomic well (see C for details regarding the construction of the atomic potential). We report the training loss for both experiments in Table 1 and compare them to the ground truth defined by a well-converged numerical solution based on a Fourier series expansion of the potential and eigenfunctions. All training was performed on a NVIDIA A100 hardware utilizing the AdamW optimizer. Further implementation details and examples are available at https://github.com/circee/blochnet.

Table 1: Model performance for 2D periodic potentials

| Potential | Total Loss (MSE) | Equation Loss | Norm Loss | BC Loss |
|---|---|---|---|---|
| Constant Potential | $2.9e$ - 4 | $2.6e$ - 4 | $2e$ - 5 | $5e$ - 6 |
| Atomic Potential | $2.7e$ - 3 | $1.2e$ - 3 | $2.4e$ - 4 | $1.2e$ - 3 |

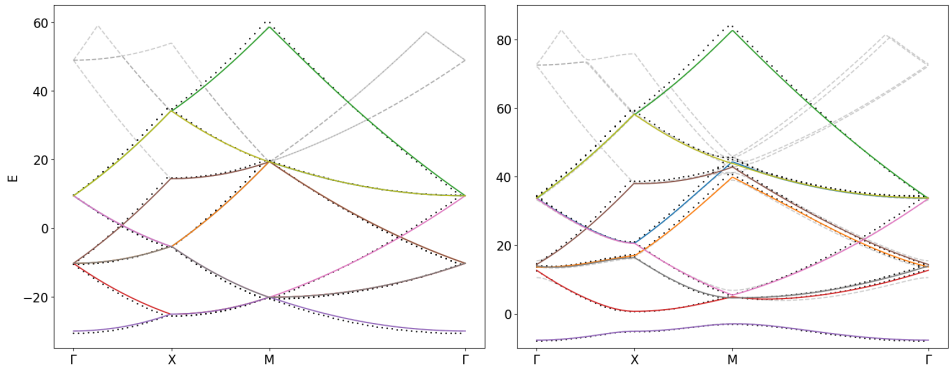


Figure 1: Energy eigenvalues $E_{\mathbf{k}}$ as a function of reciprocal space momentum (electronic band structure) for a constant potential (left) and atomic potential (right). The energy bands for the lowest state and next eight electronic states are plotted for values of $\mathbf{k}$ along a conventional high symmetry path in the reciprocal space. $\Gamma$ refers to $\mathbf{k} = 0$, while $M$ is at a corner of the square reciprocal space unit cell (Brillouin zone, see Fig. 4) and $X$ is the midpoint of one of the edges. Solid colored lines represent network energy values $\mathcal{N}_E$; dashed grey lines indicate the numerical ground truth; and black dotted lines show the expectation value of the energy eigenvalue computed from $\mathcal{N}_u$, defined in Appendix A.

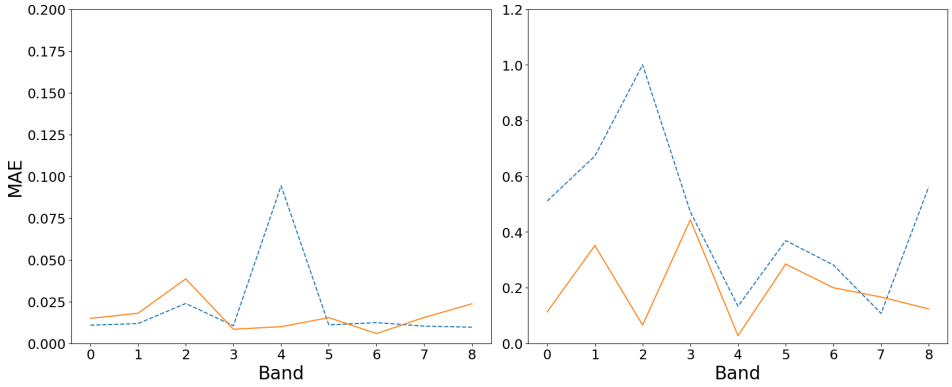


Figure 2: Mean absolute error for each energy eigenvalue (band) averaged over $\mathbf{k}$, for the constant (left) and atomic (right) potentials. Solid lines show the MAE for $\mathcal{N}_E$ while dashed lines represent the MAE of the energy expectation value.

### 4.1 Band-structure

**Constant potential:** Training loss results from Table 1 indicate that our model achieved good accuracy on the constant potential tasks, confirmed by the band-structure diagram results in Fig. 1 and computed mean-absolute error per-band in Fig. 2.

**Atomic Potential:** We chose $\mathcal{V}_{\text{atom}}(\mathbf{r})$ to approximate the Coulomb potential, $\mathcal{V}_{\text{Coulomb}}(\mathbf{r}) = \frac{1}{\mathbf{r}}$, but without the singularity at $\mathbf{r} = 0$. Compared to the constant potential, the atomic potential loss is roughly one order of magnitude larger. Notably, the boundary condition loss term is significantly larger for the atomic potential, attributable to the nontrivial Bloch states that the network must solve for. The mean absolute error (MAE) for the network energy, $\mathcal{N}_E$, is significantly lower than the energy expectation value across almost all bands.

### 4.2 Visualizing Bloch States

To visualize the Bloch states we show contour plots of $\mathcal{N}_u$ for the atomic potential with $\mathbf{k} = \Gamma$, $X$, and $M$, showing the for the five lowest energy eigenstates as a function of $\mathbf{r} = (x, y)$. Figure 3 demonstrates how our network solves for Bloch states across multiple values of $k$ simultaneously, while imposing appropriate boundary and normalization constraints from equations 8 and 9.

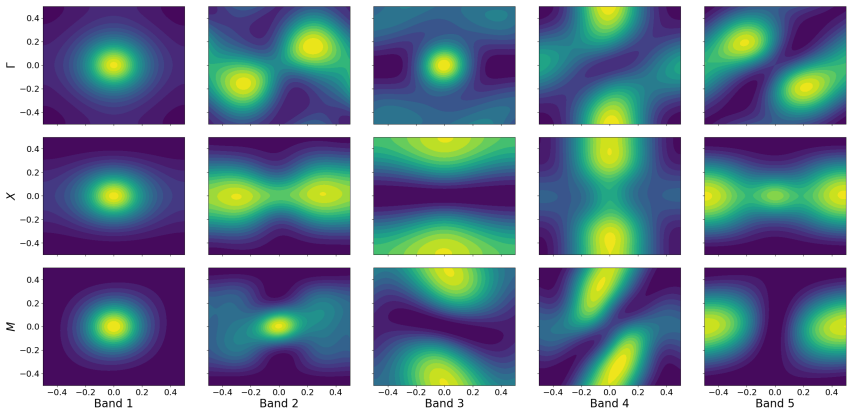


Figure 3: The probability density, $|\mathcal{N}_u(\mathbf{r})|^2$, at specific high symmetry points $\mathbf{k} = \Gamma$, $X$, and $M$, corresponding to the first five energy eigenstates, with increasing energy from left to right. $\mathcal{N}_u$ are clearly periodic in both horizontal and vertical directions.

## 5 Conclusion

We demonstrate how a neural network approach can be used to efficiently learn solutions to the Schrödinger equation with a periodic potential such as those that are ubiquitous in solid state physics. We show that theoretical assumptions may be included as soft constraints imposed through loss terms to enforce physically valid wavefunction and energy solutions to the SE, as demonstrated in Figs. 1 and 3. Beyond previous works that solve periodic systems in real space [17], our self-consistent neural network solutions provide continuous and differentiable representations of multiple bands (eigenstates) as a function of both position, $\mathbf{r}$, and crystal momentum, $\mathbf{k}$, and agree well with numerically computed ground truth values.

Our future work plans to address the limitations of our model expressivity by exploring alternatives to the feed-forward architectures, such as those developed in [18, 19]. Increasing expressive ability is necessary for our model to scale to higher dimensions, and will enable more accurate and generalizable neural network solutions to PDEs.

## 6 Acknowledgments

This work was supported by the STC Center for Integrated Quantum Materials, NSF Grant No. DMR-1231319. The computations in this paper were run on the FASRC Cannon cluster supported by the FAS Division of Science Research Computing Group at Harvard University. The authors would like to thank Efthimios Kaxiras and Robin Walters for helpful discussions.

## References


[1] Isaac E. Lagaris, Aristidis Likas, and Dimitrios I. Fotiadis. Artificial neural networks for solving ordinary and partial differential equations. *IEEE transactions on neural networks*, 9:987–1000, 1998. doi: 10.1109/72.712178.

[2] Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Physics informed deep learning (part i): Data-driven solutions of nonlinear partial differential equations, 2017.

[3] George Em Karniadakis, Ioannis G. Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Review Physics*, 2021. doi: https://doi.org/10.1038/s42254-021-00314-5.

[4] Michael Douglas, Subramanian Lakshminarasimhan, and Yidi Qi. Numerical calabi-yau metrics from holomorphic networks. In *Mathematical and Scientific Machine Learning*, pages 223–252. PMLR, 2022.

[5] Mustafa Hajij, Ghada Zamzmi, Matthew Dawson, and Greg Muller. Algebraically-informed deep networks (aidn): A deep learning approach to represent algebraic structures, 2021.

[6] Giuseppe Carleo, Ignacio Cirac, Kyle Cranmer, Laurent Daudet, Maria Schuld, Naftali Tishby, Leslie Vogt-Maranto, and Lenka Zdeborová. Machine learning and the physical sciences. *Rev. Mod. Phys.*, 91: 045002, Dec 2019. doi: 10.1103/RevModPhys.91.045002.

[7] Jan Hermann, James Spencer, Kenny Choo, Antonio Mezzacapo, W. M. C. Foulkes, David Pfau, Giuseppe Carleo, and Frank Noé. Ab initio quantum chemistry with neural-network wavefunctions. *Nature Reviews Chemistry*, 7(10):692–709, August 2023. ISSN 2397-3358. doi: 10.1038/s41570-023-00516-8. URL `http://dx.doi.org/10.1038/s41570-023-00516-8`.

[8] Jaideep Pathak, Shashank Subramanian, Peter Harrington, Sanjeev Raja, Ashesh Chattopadhyay, Morteza Mardani, Thorsten Kurth, David Hall, Zongyi Li, Kamyar Azizzadenesheli, Pedram Hassanzadeh, Karthik Kashinath, and Animashree Anandkumar. Fourcastnet: A global data-driven high-resolution weather model using adaptive fourier neural operators, 2022.

[9] Ariel Norambuena, Marios Mattheakis, Francisco J. González, and Raúl Coto. Physics-informed neural networks for quantum control. *Phys. Rev. Lett.*, 132:010801, Jan 2024. doi: 10.1103/PhysRevLett.132.010801. URL `https://link.aps.org/doi/10.1103/PhysRevLett.132.010801`.

[10] Allan Pinkus. Approximation theory of the mlp model in neural networks. *Acta Numerica*, 8:143–195, 1999. doi: 10.1017/S0962492900002919.

[11] Zheyuan Hu, Khemraj Shukla, George Em Karniadakis, and Kenji Kawaguchi. Tackling the curse of dimensionality with physics-informed neural networks. *Neural Networks*, 176:106369, 2024. ISSN 0893-6080. doi: https://doi.org/10.1016/j.neunet.2024.106369. URL `https://www.sciencedirect.com/science/article/pii/S0893608024002934`.

[12] Junwoo Cho, Seungtae Nam, Hyunmo Yang, Seok-Bae Yun, Youngjoon Hong, and Eunbyung Park. Separable physics-informed neural networks. In *Proceedings of the 37th International Conference on Neural Information Processing Systems*, NIPS '23, Red Hook, NY, USA, 2024. Curran Associates Inc.

[13] Pratik Rathore, Weimu Lei, Zachary Frangella, Lu Lu, and Madeleine Udell. Challenges in training pinns: A loss landscape perspective, 2024. URL `https://arxiv.org/abs/2402.01868`.

[14] Marios Mattheakis, Gabriel R. Schleder, Daniel T. Larson, and Efthimios Kaxiras. First principles physics-informed neural network for quantum wavefunctions and eigenvalue surfaces, 2022. URL `https://arxiv.org/abs/2211.04607`.

[15] Henry Jin, Marios Mattheakis, and Pavlos Protopapas. Physics-informed neural networks for quantum eigenvalue problems, 2022. URL `https://arxiv.org/abs/2203.00451`.

[16] N. W. Ashcroft and N. D. Mermin. *Solid State Physics*. Holt-Saunders, 1976.

[17] Gabriel Pescia, Jiequn Han, Alessandro Lovato, Jianfeng Lu, and Giuseppe Carleo. Neural-network quantum states for periodic systems in continuous space. *Phys. Rev. Res.*, 4:023138, May 2022. doi: 10.1103/PhysRevResearch.4.023138. URL `https://link.aps.org/doi/10.1103/PhysRevResearch.4.023138`.

[18] Mauricio E. Tano Quincy A. Huhn and Jean C. Ragusa. Physics-informed neural network with fourier features for radiation transport in heterogeneous media. *Nuclear Science and Engineering*, 197(9):2484–2497, 2023. doi: 10.1080/00295639.2023.2184194. URL `https://doi.org/10.1080/00295639.2023.2184194`.

[19] Ziming Liu, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin Soljačić, Thomas Y. Hou, and Max Tegmark. Kan: Kolmogorov-arnold networks, 2024. URL https://arxiv.org/abs/2404.19756.

## Appendix A Computing the expectation energy

The expectation energy can be computed from a given wavefunction $\psi$ according to the following:

$$E = \int_{Cell} \psi^* \hat{H} \psi \, d\mathbf{r}, \tag{10}$$

where $\psi^*$ is the complex conjugate of $\psi$ and $\hat{H}$ is the Hamiltonian operator. From 4 the Hamiltonian is given by:

$$\hat{H} = \left[ -\frac{1}{2}\nabla_{\mathbf{r}}^2 + i\mathbf{k} \cdot \nabla_{\mathbf{r}} - \frac{1}{2}\mathbf{k}^2 - \mathcal{V}(\mathbf{r}) \right]. \tag{11}$$

For a value of $\mathbf{k}$, the expectation energy from a network-produced Bloch state can thus be computed by:

$$E = \int_{Cell} \mathcal{N}_u^*(\mathbf{r}, k) \left[ -\frac{1}{2}\nabla_{\mathbf{r}}^2 + i\mathbf{k} \cdot \nabla_{\mathbf{r}} - \frac{1}{2}\mathbf{k}^2 - \mathcal{V}(\mathbf{r}) \right] \mathcal{N}_u(\mathbf{r}, k) \, d\mathbf{r}. \tag{12}$$

In practice, numeric integration is used to compute the above quantity with a dense mesh grid.

## Appendix B Reciprocal space sampling

Because the constant and atomic potentials are fully radially symmetric within a square unit cell, we may leverage the underlying $D_4$ symmetry to reduce required sampling in reciprocal space to a fraction of the entire space, as shown below:

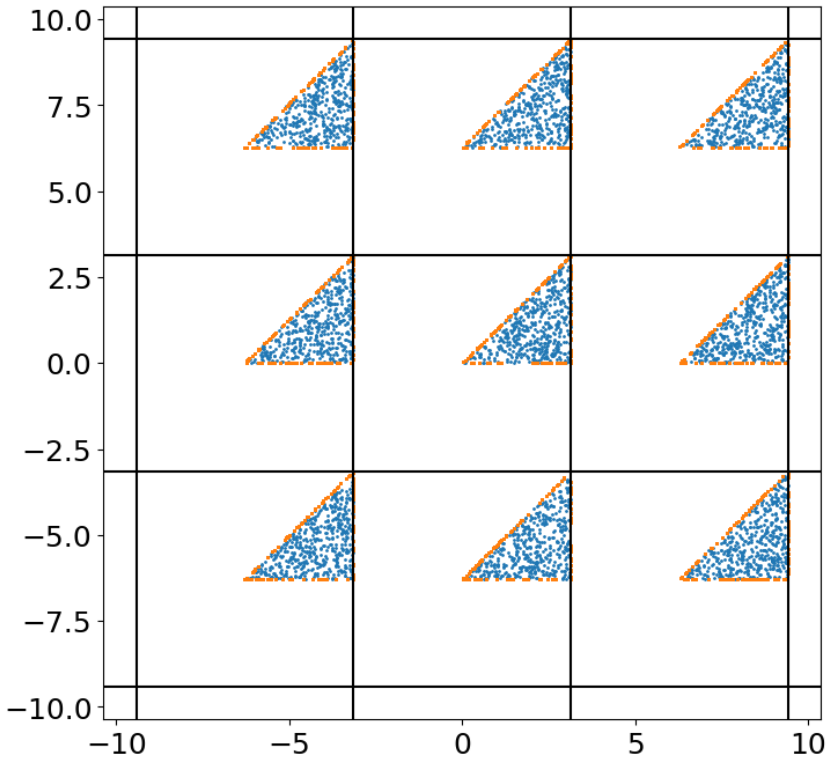


Figure 4: Reciprocal space sampling strategy. Each triangle represents an irreducible Brillouin zone (refer to [16]). The blue dots represent k-points sampled within the triangle boundary while the orange points are enforced to be directly sampled along the high-symmetry boundary path.

The central triangular cell is shifted by a linear combination of lattice basis vectors to find higher-energy states. This model of the reciprocal space is referred to as the extended zone scheme.

## Appendix C Construction of the pseudopotential well

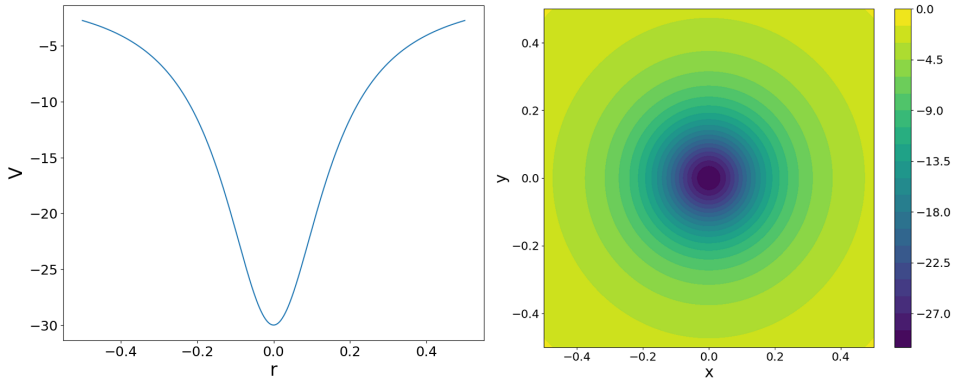


Figure 5: Radial cut view (left) and contour map (right) of the atomic pseudopotential. A depth of $-30$ was used during training.

We construct the atomic pseudopotential in two dimensions to mimic the behavior of the Coulomb potential, $\mathcal{V}_{\text{Coulomb}}(\mathbf{r}) = \frac{1}{\mathbf{r}}$, while avoiding the singular point at $\mathbf{r} = 0$. Mathematically, the atomic pseudopotential is given as follows:

$$\mathcal{V}_{\text{atom}}(\mathbf{r}) = \frac{V_0}{1 + \|\mathbf{r}\|^2}. \tag{13}$$