Deep Learning Shed Lights On Computational Condensed Matter Physics

Baiyu Zhu
Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
(Dated: May 17, 2025)

Condensed matter systems can be completely solved by solving the Schrodinger's Equation, but finding the exact solution is an NP hard problem. Thus, it's appealing to look for an approximation that's accurate, computational efficient and cheap. Methods such as Hartree Fock, Density Functional Theory(DFT), and Couples Cluster are used to approximate the wavefunction, but in most cases they are not accurate enough for practical use and are not computational efficient and scalable, especially for large systems that Condensed Matter physicists are interested in. Drawn by the success of Deep Learning, where neural network ansatz could be highly expressive and accurate while being scalable, we introduce Fermionic Neural Network ansatz to solve problems in Condensed Matter.

I. INTRODUCTION

Having an accurate approximation of wavefunctions for systems such as molecules and solids is crucial to the understanding of our physical world. It can push our understanding of superconductivity, light-atom interaction, chemical compounds and much more. Methods such as Coupled Cluster have been used to approximate the wavefunctions. However, they are either computationally expensive or of low accuracy, making them unusable for practical purposes. We are motivated to find a better method to represent our wavefunction using a more accurate and scalable method. As of today, Deep Learning has made significant progress in both Computer Vision and Natural Language Processing, where neural networks have the ability to generate realistic images and semantic texts by processing the input through layers of networks. Drawing by its tremendous capabilities, we are motivated to represent wavefunctions using neural network ansatz. In this paper, we offer an overview of PsiFormer and its application to systems such as solids, Moires materials, Fractional Quantum Hall, and Chern Insulator. Unfortunately, due to computational limitations, we are unable to carry out the full computation as it is done by [1][2][3][4][5] which requires multiple GPUs. At our best, we are only able to carry out systems around 10 electrons such as CH_4 molecules on a local computer, while to simulate a real solid, for example, a typical 2x2x2 LiH cell requires 32 electrons. We mostly quote the results from the papers referenced above.

II. NEURAL NETWORK ANSATZ

The Hamiltonian of many body system takes the form of

$$\mathcal{H} = -\frac{1}{2}\sum_i \nabla_i^2 + \sum_{i<j}\frac{1}{|r_i - r_j|} - \sum_{iI}\frac{Z_I}{|r_i - R_I|} + \sum_{IJ}\frac{Z_I Z_J}{|R_I - R_J|} \tag{1}$$

with r_i/R_I to be the positions of the electron/nucleus, and Z_I to be the charge of the nucleus. To represent an antisymmetric wavefunction of a system with n electrons, many approaches, such as Hartree Fock method, start with a collection of basis $\{\phi_i\}$ and choose a subset of n electron orbitals to construct the Ansatz as

$$\psi(x_1, ..., x_n) = \frac{1}{N}\sum_{\mathcal{P}}(-1)^{\mathrm{sng}(\mathcal{P})}\prod_i^n \phi_{\mathcal{P}(i)}(x_i) \tag{2}$$

with $\mathcal{P}$ being the permutation of the n indices that is to be summed over, and minimizes the Hamiltonian with respect to the set of Ansatzs in order to obtain the approximation for ground state wavefunction. The sign $(-1)^{\mathrm{sng}(\mathcal{P})}$ is used to take the fermionic nature of the system into consideration. However, such Ansatz doesn't capture the interactions between different electrons, while the set of basis grows exponentially with the number of electrons, making these methods untractable for large systems.

To address this problem, PsiFormer built wavefunctions from set of basis that are constructed from neural network in a permutation equivariant way, so that they are invariant under exchange of coordinates. We denote the basis by $\{\psi_i(x_j; x_{/j})\}$ with the notation $x_{/j}$ being the set of coordinates that excludes x_j. We then feed such set to Slater Determinant and obtain

$$\psi(x_1, ..., x_n) = \exp(\mathcal{J}_\theta(x))\begin{vmatrix} \psi_1(x_1; x_{/1}) & \cdots & \psi_1(x_n; x_{/n}) \\ \vdots & & \vdots \\ \psi_n(x_1; x_{/1}) & \cdots & \psi_n(x_n; x_{/n}) \end{vmatrix} \tag{3}$$

which we have also included the Jastrow factor $\exp(\mathcal{J}_\theta(x))$ to ensure the Ansatz describes cusp condition, which is the non-analytic behavior of wavefunction when two electrons overlap with each other, correctly. It can be proven that a single determinant is enough to express any wavefunction [1]. However, experiments show that it's better to construct Ansatzs from multiple determinants[1] in the form

$$\psi(x_1, ..., x_n) = \exp(\mathcal{J}_\theta(x))\sum_k w_k \det \Psi^k \tag{4}$$

with w_k being the weights. We now describe the architecture of PsiFormer briefly below. Let α, β to be the indices for spins, l to be the index for layer, and L to be the

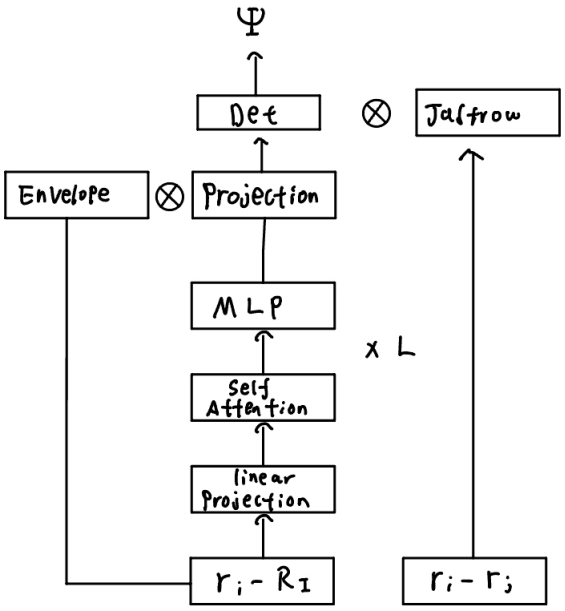

FIG. 1. Sketch for the PsiFormer architecture[2]

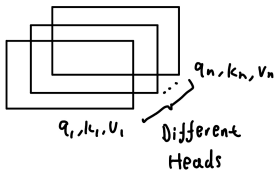

FIG. 2. Sketch for the MultiHead Self Attention Block[6]

depth of the network. For each electron with index i, we denote the corresponding feature to be f_i^0, which is constructed to be the concatenation of $\log(r_i - R_I)/|r_i - R_I|$ with $I \in \{1, 2..., N_{nuc}\}$. We then map the set of features linearly to the latent space by

$$h_i^0 = W^0 f_i^0 \tag{5}$$

Then we pass in h_i^0 to a series of blocks called the Multihead Self Attention as follows

$$f_i^l = \text{concat}[\text{SelfAttn}_h[h_1^{l-1}, ..., h_N^{l-1}]; W_q^{lh}, W_k^{lh}, W_v^{lh}] \tag{6}$$

$$h_i^l = \tanh\big(W^l f_i^l\big) + h_i^{l-1} \tag{7}$$

$$\text{SelfAttn}_i[h_1, ...h_N; W_q, W_k, W_v] = \frac{1}{\sqrt{d}} \sum_j \sigma(q_j^T k_i) v_j \tag{8}$$

$$q_j = W_q h_j,\ v_j = W_v h_j,\ k_i = W_k h_i \tag{9}$$

$$\sigma(x) = \frac{1}{1 + \exp(-x)} \tag{10}$$

where d denotes the dimension of the output, h denotes the different heads that are used to capture and communicate different information. Note that the different heads have different parameters for $\{W_q^h, W_k^h, W_v^h\}$, and different layers also have different parameters. We concatenate all of the outputs from different heads as done in (6). q,k,v stands for query, key, and value vectors. The query vector is matched against the key vector to determine the importance of the value vector, and at the end, for each position i, the attention block produces the sum of the value vectors weighted by $\sigma(q_j^T k_i)$. Equation (7), which is known as the MLP layer, is used to further mix information between different electrons. To make sure that the wavefunction has the correct far distance description, we introduce exponential decay envelope so that the wavefunction matrix reads

$$\Psi_{ij} = \sum_I \eta_{iI} \exp(-\sigma_{iI} |r_i - R_I|) w_i^T h_j^L \tag{11}$$

with w_i being the linear projection and η_{iI}, σ_{iI} being tunable parameters. Then, we take the determinant of the matrix. However, since we have not passed in the feature $r_i - r_j$, the network may have difficult time learning the cusp condition of pair-wise electron interaction. This is the reason why the Jastrow factor $\exp(\mathcal{J}_\theta(x))$ is introduced and we take it to have the form of

$$\mathcal{J}_\theta(x) = \sum_{s_i = s_j, i<j} \beta_1 \frac{\alpha_1^2}{\alpha_1 + |r_i - r_j|} + \sum_{s_i \neq s_j, i, j} \beta_2 \frac{\alpha_2^2}{\alpha_2 + |r_i - r_j|} \tag{12}$$

where s_i stands for the spin of the electron and $\{\alpha_i, \beta_i\}$ are tunable parameters. We then multiply the Jastrow factor with the determinant of the wavefunction matrix to obtain the ansatz. More insights and interpretation of this network is offered in the Appendix

To optimize the Network, we use Quantum Monte Carlo with the loss function defined as

$$\mathcal{L}(\theta) = \frac{\int d^3x \Psi^* \hat{H} \Psi}{\int d^3x \Psi^* \Psi} \tag{13}$$

If we introduce local energy $E_L = \Psi^{-1} \hat{H} \Psi$, we find that

$$E_L = -\frac{1}{2} \sum_i [(\partial_i \log |\Psi|)^2 + \partial_i^2 \log |\Psi|] + V(X) \tag{14}$$

$$\mathcal{L}_\theta = \mathbb{E}_{x \sim p_\theta}[(E_L - \mathbb{E}_{x \sim p_\theta[E_L]}) \nabla \log p_\theta] + \mathbb{E}_{x \sim p_\theta}[\nabla_\theta E_L] \tag{15}$$

where we can refer to the Appendix for the full derivation. Quantum Monte Carlo is realized by the Metropolis Hasting Algorithm. A comprehensive overview of QMC is offered in [9]

III. APPLICATIONS

A. Solid

To account for periodic boundary condition, the simpliest way to modify the architecture is to define a metric

$d(r)$ that is periodic with respect to the lattice vectors. It is proposed in [14] that we build such metric using real and reciprocal lattice (a_i, b_j) with

$$d(r) = \sqrt{APA^T} \tag{16}$$

$$P_{ij} = f^2(w_i) + g(w_i)(1 - \delta_{ij}),\ w_i = r \cdot b_i \tag{17}$$

and f and g being periodic functions. To also account for cusps, continuity at the boundary, and periodicity, we choose

$$f(w) = |w|(1 - \frac{|w/\pi|^3}{4}) \tag{18}$$

$$g(w) = w(1 - \frac{3}{2}|w/\pi| + \frac{1}{2}|w/\pi|^2) \tag{19}$$

where w is reduced to $[-\pi, \pi]$ by modding out by $2\pi n, n \in \mathbb{Z}$. For the Hamiltonian, we reduce the problem to a supercell that consists of multiple primitive cells that is interacting with its spatial images, so the Hamiltonian becomes

$$\begin{aligned} H = &-\frac{1}{2}\sum_i \nabla_i^2 + \frac{1}{2}{\sum_{a_s,i,j}}' \frac{1}{|r_i - r_j + a_s|} \\ &- \sum_{a_s,i,I} \frac{Z_I}{|r_i - R_I + a_s|} + \frac{1}{2}{\sum_{a_s,I,J}}' \frac{Z_I Z_J}{|R_I - R_J + a_s|} \end{aligned} \tag{20}$$

with $\{a_s\}$ being the set of supercell lattice vectors that is a subset of $\{a_p\}$ of primitive cell lattice vectors, $\{r_i\}$ to be the positions of the electrons in the supercell we consider, and $\sum'$ to denote the summation that excludes the case when $r_i = r_j$ and $R_I = R_J$. Under thermodynamic limit, as the size of the supercell increases, this wavefunction will be that of the real solid. Now, it's clear that the Hamiltonian is invariant under the transition of shifting all the electrons by some primitive lattice vector in $\{a_p\}$ or shift one electron by some vector in super lattice vector in $\{a_s\}$

$$\psi(r_1 + a_p, ..., r_n + a_p) = \exp(ik_p \cdot a_p)\psi(r_1, ..., r_n) \tag{21}$$

$$\psi(r_1 + a_s, ..., r_n) = \exp(ik_s \cdot a_s)\psi(r_1, ..., r_n) \tag{22}$$

with k_p and k_s being the respective momentum vector in the first Brillouin Zone. Similar to the Ansatz of traditional Ansatz for solid, DeepSolid[3] proposed the following Ansatz

$$\psi(r) = \det\{(e^{ik\cdot r_\uparrow}u_\uparrow(r))\}\det\{(e^{ik\cdot r_\downarrow}u_\downarrow(r))\} \tag{23}$$

where k's are momentum vectors that lie on the reciprocal lattice vectors of supercell $\{K_s\}$ that is also offset by k_s that lies in the first Brillouin Zone, and they are chosen to minimize the energy of the system. The outputs $u(r_i; r_{/i})$ are permutation equivariant with respects to coordinates $\{r_i\}$. The architecture for DeepSolid is not exactly PsiFormer but has the same essence, as shown in [1][3]

B. Fractional Quantum Hall Effect

Most of the studies are done by using Exact Diagonalization in this area, but only small systems are studied since ED is limited by the exponential increase of computational cost as the system scales. Furthermore, in the strong Landau Level mixing regime the traditional methods are not capable of capturing the strong correlation effects. In [3], the author used the PsiFormer to study the two dimensional electron gas in magnetic field, where they confined the system to two dimensional disk of radius a. They have demonstrated that the PsiFormer architecture is able to learn a wavefunction that has energy lower than the LL projected ED. For strong Landau Level mixing, they have also discovered a phase transition from Fractional Quantum Hall liquid to crystal state, indicating the power of PsiFormer to solve problems in condensed matter. We describe the setup from [5]. The electrons are constraint to move in an infinite two dimensional plane, and above the plan the authors proposed to place a positive charged uniform plate of radius a and total charge of Ne a distance d above the plane. Taking the circular gauge, the overall Hamiltonian takes the form of

$$H = H_0 + H_1 \tag{24}$$

$$H_0 = \sum_i \frac{1}{2}(-\nabla_i + \frac{1}{2}B \times r_i)^2 + \sum_{i\neq j}\frac{1}{2\epsilon}\frac{1}{|r_i - r_j|} \tag{25}$$

$$H_1 = \sum_i V_c(r_i) + V_b \tag{26}$$

where V_c stands for the potential energy between the electrons and the plate, and V_b stands for the potential energy of the disk, which is given as $V_b = \frac{8N^2}{3\pi\epsilon a}$[8]. The authors first compared the energy obtained by PsiFormer and ED with results obtained in the following table

N_e	PsiFormer	ED	ED dimension
6	1.1372(4)	1.13836	1206
7	1.1357(4)	1.13687	8033
8	1.1347(2)	1.13588	55974
9	1.1346(1)	1.13545	403016
10	1.1342(4)	1.13496	2977866
11	1.1337(4)	NA	22464381
12	1.1334(3)	NA	172388026

TABLE I. PsiFormer Scales as $\mathcal{O}(N^\alpha)$ while ED scales exponentially

After plotting the density of electrons(refer to [5] for concrete diagram), the authors discovered that the neural network is capable of capturing the symmetric properties of the ground state, even though no circular symmetric information is given to the input. To investigate possible phase transition, the authors changed the LL mixing

parameter $\lambda = \frac{1}{\epsilon\sqrt{w_c}}$ and plotted the wavefunctions and calculated the corresponding value for angular momentum by considering

$$\langle L \rangle = \frac{\int d^3x \Psi^* \hat{L} \Psi}{\int d^3x \Psi^* \Psi} \tag{27}$$

It is expected that as the level of mixing increase, the system will transition from Fractional Quantum Hall droplets to Wigner crystal. For the nine-electron system and $\lambda \in \{\frac{1}{3}, 1, 3\}$, the angular momentum stayed at $L = 108$. While for λ, the angular momentum changed to $L = 100$. The spatial profile of the wavefunctions are also plotted and it is observed that for $\lambda \in \{\frac{1}{3}, 1, 3\}$ the electrons are quite diluted. In comparison, for $\lambda = 9$, there's an electron well localized at the origin, which is further confirmed by the plateau in spatial profile of electron density, and the rest of the electrons are separated at the outer boundary. This study demonstrates the capability and universality of PsiFormer architecture.

C. Topological Insulator

In recent years, a fascinating material called Moire material is discovered by experimentalists. Due to its versatility, experimentalists can tune Moire materials to produce novel quantum states such as QHE, Chern Insulators, Z_2 topology insulators and many more. In [4], the author studied a Moire material called $tMoTe_2$ which is a twisted bilayer of MoTe that is characterized by the twisted angle θ. Due to its strong spin valley locking, the top valence band is separated from others and give rise to the spinor representation of the wavefunction. Then, the effective Hamiltonian for spin up electrons after particle hole transformation takes the following form

$$H_\uparrow(r) = \begin{pmatrix} \frac{(-i\nabla - K_+)^2}{2m} + \Delta_b(r) & \Delta_T(r) \\ \Delta_T^*(r) & \frac{(-i\nabla - K_-)^2}{2m} + \Delta_t(r) \end{pmatrix} \tag{28}$$

$$\Delta_{b/t}(r) = -2V \sum_{i=1,3,5} \cos(h_i \cdot r + \varphi) \tag{29}$$

$$\Delta_T(r) = w(1 + \exp(ih_2 \cdot r) + \exp(ih_3 \cdot r)) \tag{30}$$

$$h_i = \frac{8\pi \sin(\theta/2)}{\sqrt{3}a_0}(\cos\left(\frac{\pi(i-1)}{3}\right), \sin\left(\frac{\pi(i-1)}{3}\right)) \tag{31}$$

$$K_+ = \frac{h_1 + h_3}{3},\ K_- = \frac{h_1 + h_0}{3} \tag{32}$$

where $\Delta_{b/t}$ stands for the interaction between the intra-layers(top and bottom) while Δ_T stands for interaction between inter-layers. $K_\pm$ denotes the high-symmetric points in Brillouin Zone. a_0 is the typical Moire length and θ is the twisted angle. V, w, φ are fitted from density functional results [10]. To study the topological properties of the system, Chern number is used, which is defined as

$$C = \frac{1}{2\pi i} \int_{\text{SBZ}} B(k_s) dk_s \tag{33}$$

$$B(k_s) = \nabla_{k_s} \times \langle \Psi_{k_s} | \nabla_{k_s} | \Psi_{k_s} \rangle \tag{34}$$

and can be calculated by the single point formula [11]. For filling at $n = -1$, ferroelectric states are observed at small twisted angle of $\theta \lesssim 1.75°$ with electrons accumulate at one layer, which is due to the Coulomb interaction being the dominate force at small angles. For $\theta \gtrsim 2.0°$, the system acquires Chern number of 1 and is identified as Chern Insulator. By further increasing the twisted angle, a transition from ferromagnetic to antiferromagnetic state is observed. This study highlights the tremendous potential of using neural networks to discover novel quantum phenomenons.

IV. IMPROVING THE ARCHITECTURES

We offer few insights on how to improve the architectures. To improve the PsiFormer architecture, one could consider using Mixture of Expert that is heavily popularized recently by DeepSeek[13], which achieved significantly lower training and inference costs. This may potentially lead to the neural network architectures to work on larger systems. It is also proposed to introduce sparse attention, which we only consider query, key, and value evaluation for a subset of electrons, but to ensure permutation equivariant one must be careful about this implementation. Since the query, value, and key vectors are merely linear transformation of h_i, we could add non linearity such as the sigmoid function. However, this comes with the increase in computational costs. It may also seem that the Jastrow factor and envelope are artificially introduced by design choice, so there could be separate streams that learn these factors and multiply them at the end. Also, it is found that there's no significant improvement by including the Jastrow factor, and instead it slows down the GPU time per step by 20%[12]. Aside from using Slater determinant, one may also consider using Pfaffians that is introduced in [13] to take into account of the fermionic nature. For DeepSolid, the features could be potentially be reduced to merely $((\sin(g_1 \cdot r), \sin(g_2 \cdot r), \cos(g_1 \cdot r), \cos(g_2 \cdot r))$ with $g_{1,2}$ being the crystal momentum[12].

V. FURTHER APPLICATIONS

Further works could be done on superconductivity, excited state of molecules, fractional excitations and anyons.

VI. CONCLUSION

In this paper, we offered an overview of recent advancements in AI for Physics. With the advancements of Deep Learning and computational power, we have demonstrated the ability of neural networks to capture complex wavefunctions to an unprecedented accuracy while being scalable, while also offering valuable insights to physical systems.

VII. ACKNOWLEDGEMENTS

I'm grateful for Professor Mehran Kardar for an amazing introduction to Statistical Field Theory and showing me the wonders of our world. I would also like to show appreciation to Alex Gu, who has supported me throughout this difficult semester and has been a great friend. The author acknowledges that there are potential areas of improvement to be made, as it is the second paper written by the author who's a sophomore, and AI for Science is a relatively new field. However, the author is optimistic about the future of AI for Science and believes that we could advance our knowledge of Condense Matter greatly using powerful computational power and advanced neural network architectures.

VIII. APPENDIX

A. Self Attention Mechanism

self-attention[6] was originally designed to handle problems in Natural Language Processing. Each words of a sentence is represented by a token that is represented as a vector in $\mathbb{R}^N$ which we denote as f_i^0. Then we construct three vectors called query, key, and value, denoted as q_i^0, k_i^0, v_i^0 through linear transformation. Key encapsulate the information to be communicated. The query q_i^0 for each token f_i^0 is matched against the all the keys of the tokens through dot product to determine the importance of the information v_j^0 that is to be mixed together. In this way tokens are able to communicate with each other. Then the outputs of the attention layer are mixed by a layer that consists of a linear transformation and nonlinear function to form the tokens for the next layer. The term head means we have the same process to happen in parallel, which is to be interpreted as processing and capturing different information . This process is repeated a total of L times. For PsiFormer, we treat the whole whole to be processed as sentence and each electron's information is encapsulated in token similar to each words.

B. Derivation Of The Equations

For kinectic energy we consider

$$\begin{aligned}\Psi^{-1}\partial_i(\partial_i\Psi) &= \Psi^{-1}\partial_i(\Psi\Psi^{-1}\partial_i\Psi)\\ &= \Psi^{-1}\partial_i(\Psi\partial_i\log|\Psi|)\\ &= (\partial_i\log|\Psi|)^2+\partial_i^2\log|\Psi|\end{aligned} \tag{35}$$

and this implies that

$$E_L(X) = -\frac{1}{2}[(\partial_i\log|\Psi|)^2+\partial_i^2\log|\Psi|] + V(X) \tag{36}$$

If we denote $p_\theta \equiv |\Psi|^2$, then we have

$$\begin{aligned}\nabla_\theta\mathcal{L} &= \frac{[\int p_\theta(p_\theta^{-1}\nabla p_\theta E_L(X)+\nabla E_L(X)]\int p_\theta - \int\nabla p_\theta\int p_\theta E_L}{\int p_\theta^2}\\ &= \frac{\int p_\theta[E_L\nabla\log p_\theta(X)+\nabla E_L(X)]}{\int p_\theta} - \frac{\int\nabla\log p_\theta\int p_\theta E_L}{\int p_\theta}\\ &= \mathbb{E}_{x\sim p_\theta}[(E_L-\mathbb{E}_{x\sim p_\theta}[E_L])\nabla\log p_\theta] + \mathbb{E}_{x\sim p_\theta}[\nabla E_L]\end{aligned} \tag{37}$$

[1] D. Pfau, J. S. Spencer, A. G. D. G. Matthews, and W. M. C. Foulkes, Phys. Rev. Res. 2, 033429 (2020). https://doi.org/10.1103/PhysRevResearch.2.033429.

[2] I. von Glehn, J. S. Spencer, and D. Pfau, arXiv:2211.13672 [physics.chem-ph] (2022).

[3] X. Li, Z. Li, and J. Chen, Nat. Commun. 13, 7895 (2022). DOI: 10.1038/s41467-022-35627-1mechanics. Oxford,: Clarendon Press.

[4] X. Li, Y. Chen, B. Li, H. Chen, F. Wu, J. Chen, and W. Ren, arXiv:2503.11756 [cond-mat.str-el] (2025). DOI: 10.48550/arXiv.2503.11756

[5] Y. Teng, D. D. Dai, and L. Fu, arXiv:2412.00618 [cond-mat.str-el] (2024). DOI: 10.48550/arXiv.2412.00618

[6] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin, arXiv:1706.03762 [cs.CL] (2017).

[7] D. Luo, D. D. Dai, and L. Fu, arXiv:2406.17645 [cond-mat.str-el] (2024)

[8] O. Ciftja, Results in Physics 7, 1674–1675 (2017) DOI: 10.48550/arXiv.1706.03762

[9] J. Toulouse, R. Assaraf, and C. J. Umrigar, arXiv:1508.02989 [physics.chem-ph] (2015)

[10] A. P. Reddy, F. Alsallom, Y. Zhang, T. Devakul, and L. Fu, Fractional quantum anomalous Hall states in twisted bilayer MoTe2 and WSe2, Physical Review B 108, 085117 (2023)

[11] I. Gilardoni, F. Becca, A. Marrazzo, and A. Parola, Real-space many-body marker for correlated Z2 topological insulators, Physical Review B 106, L161106 (2022)

[12] M. Geier, K. Nazaryan, T. Zaklama, and L. Fu, arXiv:2502.05383 [cond-mat.str-el] (2025)

[13] eepSeek-AI, A. Liu, B. Feng, B. Wang, B. Wang, B. Liu, C. Zhao, C. Deng, C. Ruan, D. Dai, D. Guo, et al.,

arXiv:2405.04434 [cs.CL] (2024)

[14] Whitehead, T. M., Michael, M. H. Conduit, G. J. Jastrow correlation factor for periodic systems. Phys. Rev. B 94, 035157 (2016).