Jan Jakob
jan-jakob.tech
jan.jakob@campus.lmu.de

QFORMER: An Efficient Foundational PDE Model for the Time-Dependent Schrödinger Equation

Research Proposal

The Schrödinger equation describes the quantum-mechanical behavior of particles, making it the most fundamental equation in chemistry. It is too difficult to solve analytically for all but the simplest systems. However, accurately solving the equation for various quantum physical systems, particularly those with complex potential landscapes, is crucial for advancements in quantum chemistry, condensed matter physics and material science, as the wave function of a given system allows the computation of any of its properties [Scherbela, Reisenhofer, et al. 2022]. In recent years, tools from deep learning have been used to great effect to improve the quality of computational quantum physics [Carleo and Troyer 2017]. For the study of chemistry in particular, it is the quantum behavior of electrons that matters, which imposes certain constraints on the possible solutions of the Schrödinger equation. The use of deep neural networks for successfully computing the quantum behavior of molecules was introduced almost simultaneously by several groups ([Pfau et al. 2020], [J. Hermann, Schätzle, and Noé 2020], [Choo et al. 2020]), and has since led to a variety of extensions and improvements [H. Hermann et al. 2022]. However, follow-up work has mostly focused on applications and iterative improvements to the neural network architectures introduced in the first set of papers.

At the same time, neural networks using self-attention layers, like the Transformer [Vaswani et al. 2017], have had a profound impact on much of machine learning. They have led to breakthroughs in natural language processing [Devlin et al. 2018], language modeling [Brown et al. 2020], image recognition [Dosovitskiy et al. 2020], and protein folding [Jumper et al. 2021]. The basic self attention layer is also permutation equivariant, a useful property for applications to chemistry, where physical quantities should be invariant to the ordering of atoms and electrons [Fuchs et al. 2020]. In [Glehn, Spencer, and Pfau 2022] presented a novel neural network architecture using self-attention called Wavefunction Transformer (Psiformer), that can be used as an approximation (or Ansatz) for solving the steady state many-electron Schrödinger equation. Psiformer is an ab initio method that learns directly from the fundamental equations governing quantum systems without needing pre-existing datasets. It was the first work demonstrating that self-attention networks can learn complex quantum mechanical correlations between electrons. However, the method only determines the ground state of quantum mechanical systems using Transformers but does not solve the time-dependent case. Also, Psiformer is trained individually for each specific electron configuration. Thus, recent advancements have introduced the concept of transferability in neural wavefunction models. In [Scherbela, Gerard, and Grohs 2024], the authors propose to transform uncorrelated, computationally inexpensive Hartree-Fock orbitals into correlated, high-accuracy neural network orbitals. This method allows the model to generalize and adapt to various molecular systems without retraining from scratch for each new configuration. This method enhances the precision of quantum chemical calculations by more effectively capturing molecule electron correlations.

On the other hand, frameworks like PINNsFormer [Zhao, Ding, and Prakash 2023] can accurately approximate PDE solutions by utilizing multi-head attention mechanisms to capture temporal dependencies. Empirical results for the Navier-Stokes equations demonstrate that frameworks like PINNsFormer achieves superior generalization ability and accuracy across various scenarios, including Physics Informed Neural Networks (PINNs) failure modes and high-dimensional PDEs. In 2024, Herde et al. introduced POSEIDON, a Multiscale Operator Transformer (MOT) with time-conditioned layer norms, enabling continuous-in-time evaluations [Herde et al. 2024]. Additionally, it leverages the semi-group property of time-dependent PDEs to significantly scale up training data. The benefits of this approach include enhanced sample efficiency and accuracy in learning solution operators for PDEs, as well as improved generalization to new, unseen physics not encountered during pretraining. POSEIDON was pretrained on a diverse, large scale dataset for the governing equations of fluid dynamics.

Jan Jakob
jan-jakob.tech
jan.jakob@campus.lmu.de

In this research we are going to develop and pretrain a Transformer-based model QFORMER tailored for solving the time-dependent Schrödinger equation for arbitrary electron configurations and initial states. Similar to POSEIDON, we use a MOT to handle multiscale problems by capturing interactions across different scales, which is essential for accurately solving the Schrödinger equation that exhibit behaviors at various levels of granularity [Liu, Xu, and Zhang 2022]. We add PINN mechanisms to directly incorporate physical constraints of the wave function, like the normalization of the wave function to the loss function. Our approach involves integrating Poseidon's multiscale operator transformer architecture with PINNsFormer's sequential loss functions and Wavelet activation function. This integration aims to enhance the model's capability to capture both multiscale phenomena and temporal dependencies of the wave function. Additionally, by incorporating the semi-group property of the time-dependent Schrödinger equation into our training scheme, we seek to further improve efficiency and accuracy. We pretrain the model on various analytically solvable systems like the Harmonic Oscillator with time-dependent frequency, the Rabi Model, the Landau-Zener problem, systems with periodic Hamiltonians and electrons in a time-dependent magnetic field. In addition, we use as training data also systems with time independent Hamiltonians to enable QFORMER to calculate the energy levels of our system. Therefore we take datasets used in previous works like [Scherbela, Gerard, and Grohs 2024] into considerations. For all systems, we consider multiple different initial and boundary conditions for each system if they satisfy the mathematical and physical properties of the wave function.

As far as we know, this will be the first work pertaining a Transformer model to solve the time-dependent Schrödinger equation. We expect our model to solve the Schrödinger equation for general complex systems efficiently and accurately. Combining multiscale operator transformers with physics-informed approaches is expected to enhance model performance beyond traditional numerical methods. Successfully developing such a model could revolutionize the computation of quantum mechanical systems by reducing computational costs and increasing scalability. This advancement would have far-reaching applications in fields like quantum chemistry, materials science, and physics.

Brown, Tom B. et al. (2020). "Language Models are Few-Shot Learners". In: *arXiv preprint arXiv:2005.14165*. URL: https://arxiv.org/abs/2005.14165.

Carleo, Giuseppe and Matthias Troyer (2017). "Solving the quantum many-body problem with artificial neural networks". In: *Science* 355.6325, pp. 602–606. DOI: 10.1126/science.aag2302. URL: https://www.science.org/doi/10.1126/science.aag2302.

Choo, Kenny et al. (2020). "Fermionic neural-network states for ab-initio electronic structure". In: *Nature Communications* 11, p. 2368. DOI: 10.1038/s41467-020-15724-9. URL: https://www.nature.com/articles/s41467-020-15724-9.

Devlin, Jacob et al. (2018). "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding". In: *arXiv preprint arXiv:1810.04805*. URL: https://arxiv.org/abs/1810.04805.

Dosovitskiy, Alexey et al. (2020). "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale". In: *arXiv preprint arXiv:2010.11929*. URL: https://arxiv.org/abs/2010.11929.

Fuchs, Fabian B. et al. (2020). "SE(3)-Transformers: 3D Roto-Translation Equivariant Attention Networks". In: *Advances in Neural Information Processing Systems* 33, pp. 1970–1981. URL: https://arxiv.org/abs/2006.10503.

Glehn, Ingrid von, James S. Spencer, and David Pfau (2022). "A Self-Attention Ansatz for Ab-initio Quantum Chemistry". In: *arXiv preprint arXiv:2211.13672*. URL: https://arxiv.org/abs/2211.13672.

Herde, Maximilian et al. (2024). "Poseidon: Efficient Foundation Models for PDEs". In: *arXiv preprint arXiv:2405.19101*. URL: https://arxiv.org/abs/2405.19101.

Hermann, Hauke et al. (2022). "Behavioural phenomena in persons with an intellectual developmental disorder according to the level of emotional development". In: *Journal of Intellectual Disability Research* 66.5, pp. 483–498. DOI: 10.1111/jir.12930. URL: https://onlinelibrary.wiley.com/doi/10.1111/jir.12930.

Hermann, Jan, Zeno Schätzle, and Frank Noé (2020). "Deep-neural-network solution of the electronic Schrödinger equation". In: *Nature Chemistry* 12, pp. 891–897. DOI: 10.1038/s41557-020-0544-y. URL: https://www.nature.com/articles/s41557-020-0544-y.

Jumper, John et al. (2021). "Highly accurate protein structure prediction with AlphaFold". In: *Nature* 596.7873, pp. 583–589. DOI: 10.1038/s41586-021-03819-2. URL: https://www.nature.com/articles/s41586-021-03819-2.

Liu, Xinliang, Bo Xu, and Lei Zhang (2022). "HT-Net: Hierarchical Transformer based Operator Learning Model for Multiscale PDEs". In: *arXiv preprint arXiv:2210.10890*. URL: https://arxiv.org/abs/2210.10890.

Pfau, David et al. (2020). "Ab initio solution of the many-electron Schrödinger equation with deep neural networks". In: *Physical Review Research* 2, p. 033429. DOI: 10.1103/PhysRevResearch.2.033429. URL: https://journals.aps.org/prresearch/abstract/10.1103/PhysRevResearch.2.033429.

Scherbela, Michael, Leon Gerard, and Philipp Grohs (2024). "Towards a Transferable Fermionic Neural Wavefunction for Molecules". In: *Nature Communications* 15.1, p. 120. DOI: 10.1038/s41467-023-44216-9. URL: https://www.nature.com/articles/s41467-023-44216-9.

Scherbela, Michael, Rafael Reisenhofer, et al. (2022). "Solving the electronic Schrödinger equation for multiple nuclear geometries with weight-sharing deep neural networks". In: *Nature Computational Science* 2, pp. 331–341. DOI: 10.1038/s43588-022-00228-x. URL: https://www.nature.com/articles/s43588-022-00228-x.

Vaswani, Ashish et al. (2017). "Attention Is All You Need". In: *arXiv preprint arXiv:1706.03762*. URL: https://arxiv.org/abs/1706.03762.

Zhao, Leo Zhiyuan, Xueying Ding, and B. Aditya Prakash (2023). "PINNsFormer: A Transformer-Based Framework For Physics-Informed Neural Networks". In: *arXiv preprint arXiv:2307.11833*. URL: https://arxiv.org/abs/2307.11833.