Department of Chemistry, Biochemistry, and Physics

Many-Body Systems: Integration of Computational & Neural Network Techniques

Leonel Sanchez Torres
PHY 461: Senior Research
Dr. Robert Ekey
4/30/2025

I. Introduction

Gaining knowledge about multi-electron atomic structures will increase the understanding of atoms' behavior in terms of their electronic structure and potential for chemical bonding and fundamental physics beyond the standard model[1], which is the theoretical framework that describes electromagnetic, weak, and strong nuclear interactions among elementary particles. This data is derived by solving the Schrödinger equation for a system with multiple electrons describing the quantum behavior of electrons within an atom paired with the configuration interaction (CI) approach[2], which involves expressing the total wavefunction as a linear combination of multiple electron configurations. Considering electron-electron repulsion and the shielding effect, it becomes highly complex for systems with a large number of electrons.

The Hartree-Fock (HF) and the multiconfiguration Dirac-Hartree-Fock (MCDHF) methods are computational techniques for calculating the electronic structures of atoms and molecules. HF provides an average-field approximation to account for electron-electron repulsion but does not consider relativistic interactions[3]. The MCDHF implements HF as the base and builds on top of the approximation by including electron configurations and the relativistic corrections, achieving a higher accuracy[4] in addition to a higher computational cost. For multi-electron structures, the number of iterations needed increases to hundreds of millions of terms for the multiple configurations, which increases the computational requirements for performing the calculations. The ability to pre-select basis states with relative weights would enable us to create wave functions that provide precise observables without requiring computational effort on the whole basis.

Selected CI approaches were applied to atomic and molecular systems according to perturbation theory[5] or the Monte Carlo methodology[6], both of which aim to approximate the contributions of configuration states without exhaustively sampling the full space. However, these methods become inefficient for large basis sets such as heavy metals due to the computations required over the entire basis in the perturbation theory, and the random selection in the Monte Carlo method disregards the properties of its states. In response to the limitations of traditional configuration interaction methods particularly for large systems, a recent shift has been made into integrating machine learning models to reduce the computational requirements needed for complex systems calculations. One of the models that has so far proven its versatility for active learning has been Neural networks (NN)[7].

Inspired by this approach, an efficient deep-learning approach was developed to construct an approximate wave function that delivers atomic-level energies for high-Z atoms and ions with many electrons. The key step is the application of the NN after the appropriate encoding of the basis states. This approach significantly improves the performance and reduces the memory requirements without hindering the accuracy of the results for atomic systems. The NN is trained during runtime and asked to select the important states for specific electronic levels. It is designed to work alongside the General Relativistic Atomic Structure Package GRASP2018[8] which also addresses the atomic structure calculations employing the multiconfiguration Dirac-Hartree-Fock (MCDHF) method[9]. The NN will be able to select the most important states , which are then calculated through GRASP2018 significantly reducing the number of states needed to be calculated.

To evaluate this approach, ground state energies for neutral ^{74}W (tungsten) atoms were calculated using GRASP2018 to be used as benchmarks for assessing the model's accuracy.

II. Theory

One of the key concepts in many-body physics relies on the Electronic structure theory[10]. Electronic structure theory deals with the quantum states of the electrons and addresses the forces that the electrons' presence creates on the nuclei; it is these forces that determine the geometries and energies of various stable structures of the molecule as well as transition states connecting these stable structures. Because there are ground and excited electronic states, each of which has different electronic properties, there are different stable-structure and transition-state geometries for each such electronic state. Electronic structure theory deals with all of these states, their nuclear structures, and the spectroscopies connecting them[11]. This theory focuses on the principles of quantum mechanics to model the behavior of electrons, which are treated as wave-like particles. Describing the molecule in quantum mechanics is done through equation 1, the time-independent Schrodinger equation.

$$\widehat{H}\Psi = \mathrm{E}\Psi \tag{1}$$

Where the wave function Ψ describes the quantum state of the particle, while the Hamiltonian $\widehat{H}$ captures the kinetic and potential energies of all the particles. For a simple molecule when H is expanded into the following components.

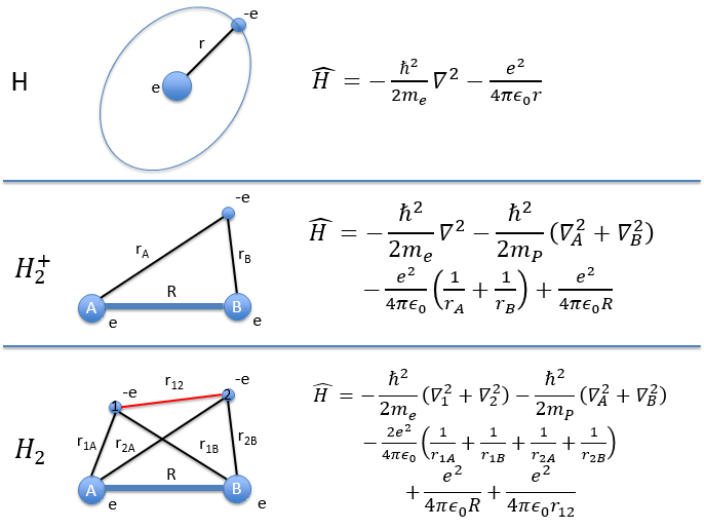

Figure 1 Schematic representations and corresponding Hamiltonians for the hydrogen atom.[12]

provides a visual representation of the interactions described in Equation 2. Each panel illustrates the components of the many-body Hamiltonian for increasingly complex systems: the hydrogen atom, the hydrogen molecular ion $H2^+$, and the hydrogen molecule H2. These correspond directly to the terms in Equation 2, the kinetic energy of nuclei $\sum_A$, kinetic energy of electrons $\sum_i$, electron-nucleus attraction rAi, nucleus-nucleus repulsion RAB, and electron-electron repulsion $\mathbf{r_{ij}}$. The diagram helps clarify how these interactions arise geometrically in atomic and molecular systems and illustrates the physical distances used to define each energy contribution in the Hamiltonian.

$$\widehat{H} = -\sum_{A}^{nuc}\frac{\hbar}{2M_A}\nabla_A^2 + \frac{\hbar^2}{2}\sum_{i}^{elec}\nabla_i^2 - \sum_{A}^{nuc}\sum_{i}^{elec}\frac{(Z_A)(e^2)}{4\pi\varepsilon_0 r_{Ai}} + \sum_{A>B}^{nuc}\frac{Z_A Z_0 e^2}{4\pi\varepsilon_0 R_{AB}} + \sum_{i>j}^{elec}\frac{e^2}{4\pi\varepsilon_0 r_{ij}} \quad (2)$$

Where the first summation describes the nuclear kinetic energy, the second the electron kinetic, the third the electron-nuclear attraction, the fourth the nuclear-nuclear repulsion,

and the fifth the electron-electron repulsion[13]. Since $\widehat{H}$ captures the kinetic and potential energies, it can be rewritten into a short-hand notation.

$$\widehat{H} = T_N(R) + T_e(r) + V_{(eN)}(r, R) + V_{(NN)}(R) + V_{(ee)}(R) \quad (3)$$

$T_N(R)$ is the kinetic energy of the nuclei, $T_e(r)$ is the kinetic energy of the electron, $V_{(eN)}(r, R)$ is the potential energy between the electron and nuclei, $V_{(NN)}(R)$ is the potential energy of the nuclear-nuclear repulsion, and $V_{(ee)}(R)$ is the potential energy of the electron-electron repulsion. Over time multiple methods to approximate the energy of the system, one of them being the Born-Oppenheimer Approximation[14] which has also become the starting point for other approximations. This method states that since the electrons around the nuclei are significantly faster and are extremely smaller in comparison, it can be assumed that this equivalates to the nuclei being frozen in a fixed position. This assumption allows us to ignore the $T_N(R)$ change the parameters of $V_{(eN)}(r, R)$ to $V_{(eN)}(r; R)$ where switching the separator indicates that R is fixed. To get the electronic Hamiltonian H_{el}, transforming the original Schrodinger equation into the electronic Schrodinger.

$$H_{el}(r; R)\psi(r; R) = E_{el}(\mathrm{R})\psi(r; R) \quad (4)$$

Now having the electronic Schrodinger equation, it is needed to solve for the electronic wave function ψ, this can be done through the Hartree-Fock Approximation[15]. This method states that for the electronic Schrodinger equation to be able to analytically solvable, the variables must be separable which in this system are the coordinates of the electrons, but the separation still requires retaining information about the electron-electron

interactions. Therefore, the ψ is approximated so it only depends on the coordinates of one electron.

$$\psi(r;R) \rightarrow \psi(r) \tag{5}$$

The change in ψ affects the H_{el}, meaning the interaction between electrons is going to be the average.

$$H_{el}(r)\psi(r) = E_{el}(\mathrm{R})\psi(r) \tag{6}$$

Then $H_{el}(r)$ becomes.

$$H_{el}(r) = T_e + V_{en} + V_{ee} + V_{nn} \tag{7}$$

The kinetic energy of the nuclei is ignored because of its dependency on R, which is fixed under the Born-Oppenheimer approximation. These simplifications reduce H to depend only on electronic coordinates. With this form, the energy expectation value of H_{el} can be calculated using a slater determinant[16] which encodes an antisymmetric wavefunction for a multi-electron system. A slater determinant is constructed by arranging spin orbitals into a determinant form, ensuring that the total wavefunction changes sign when any two electrons are exchanged satisfying the Pauli exclusion principle, following the slater-Condon rules [17]. This rules provide a systematic way to evaluate expectation values of operators over these determinants, they simplify the computation of kinetic, potential, and exchange integrals by reducing them to a sum over one-electron and two-electron terms.

$$E[\psi_{HF}] = \langle \psi_{HF} | H_{ei} | \psi_{HF} \rangle \tag{8}$$

To derive the HF equation, it is necessary to minimize the energy functional[18] for N electrons with orthonormal constraints and introduce the Fock operator. A functional is a mathematical object that maps a function to a scalar value. The general functional takes the

set of orbitals as inputs and returns the corresponding total energy of the system. The introduction of this operator allows us to write the equation.

$$F(X_k)\phi_k(X_k) \equiv [h(X_k) + J(X_k) - K(X_k)]\phi_k(X_k) = \varepsilon_k\phi_k(X_k) \quad (9)$$

This is the Hartree-Fock equation, where F is the Fock operator, acting on the spin orbital ϕ_k to yield the orbital energy ε_k. Where $h(X_k)$ is the one-electron operator, representing the kinetic energy of an electron and its attraction to the nuclei. $J(X_k)$ a Coulomb operator[19] representing the average electrostatic repulsion between one electron and the others in the system. and $K(X_k)$ an exchange operator[20], a non-classical term arising from the antisymmetric of the wavefunction, accounting for exchange interactions between electrons with the same spin. The optimal total energy E_{HF} can be written in terms of molecular orbitals.

$$E_{HF} = \sum_{i=1}^{N} h_{ii} + \sum_{i=1}^{N}\sum_{j=1}^{N/2}\left[2J_{ii} - K_{ij}\right] + V_{nucl} \quad (10)$$

Where E_{HF}'s only dependency is on a single determinant and non-relativistic meaning that with the two approximations and the multiple manipulations, it has made a good approximation for solving a simple system at a classical level. It neglects the electron correlations due to assuming that each electron moves independently in the average field created by other electrons which leads to inaccuracies when dealing with many-electron systems. To address those limitations, the Multiconfiguration Dirac-Hartree-Fock (MCDHF) method implements the HF approximation and incorporates the relativistic effects by considering multiple electron configurations and multi-determinant which leads to the double summation over all configuration state functions (CSFs). The MCDHF is a

computational method that has been used over the last couple of decades, it provides a higher accuracy of atomic structure calculations, and it is most commonly used through the General Relativistic Atomic Structure Package (GRASP2018)[21].

To derive the MCDHF for the small system its needed to look back to equation (7) where the electronic kinetic energy needs to be substituted by the Dirca-Hamiltonian[22]

$$T_e \quad \rightarrow \quad \sum_i (c\alpha \cdot p_i + \beta mc^2) \tag{11}$$

$$H_{\mathrm{DC}} = \sum_i (c\alpha \cdot pi + \beta mc^2) + Ven + V_{ee} + V_{nn} \tag{12}$$

$$H_{\mathrm{DC}} = \sum_{i=1}^{N} \left(c\alpha i \cdot pi + (\beta_i - 1)c^2 + V\mathrm{nuc}(r_i) \right) + \sum i > j^N \frac{1}{r_{ij}} \tag{13}$$

Where $V_{\mathrm{nuc}}(r_i)$ is the monopole part of the electron-nucleus Coulomb interaction, r_{ij} is the distance between electrons i and j, $\boldsymbol{\alpha}$ and β are the 4-by-4 Dirac matrices, and c is the speed of light. The approximate solutions to the wave equations are referred to as atomic state functions (ASFs). An atomic state function, $\Psi(\gamma PJ)$, is in our approach represented by a linear combination of CSFs.

$$\Psi(\gamma PJ) = \sum_{j=1}^{N_{\mathrm{CSFs}}} c_j \Phi\left(\gamma_j PJ\right) \tag{14}$$

Where P is the parity and J is the total angular momentum. γ_j represents all necessary quantum numbers and the orbital occupancy to define the CSF, while c_j are the mixing coefficients. The γ is usually selected as the γ_j corresponding to the largest weight $|c_j|^2$. The CSFs are in turn constructed as angular-momentum-coupled, anti-symmetrized products of one-electron Dirac-orbitals of the form.

$$\psi(r) = \psi_{n\kappa,m}(r) = \frac{1}{r}(P_{n\kappa}(r)\chi_{\kappa,m}(\theta,\varphi) i Q_{n\kappa}(r)\chi_{-\kappa,m}(\theta,\varphi)) \quad (15)$$

Where $P_{n\kappa}(r)$ and $Q_{n\kappa}(r)$ are the large and small components of the radial wave function, and $\chi_{\pm\kappa,m}(\theta,\varphi)$ are two-component spin-orbit functions. The extended optimal level (EOL) scheme is used to determine energies and wave functions. In the EOL scheme, the radial parts of the Dirac orbitals and the expansion coefficients of the targeted states are optimized to self-consistency by solving the MCDHF equations, which are derived using the variational approach[4]. In subsequent relativistic configuration interaction (RCI) calculations the transverse photon interaction (Breit interaction).

$$H_{\text{Breit}} = -\sum_{i<j}^{N}\left[\frac{\alpha i \cdot \alpha j \cos(wijrij/c)}{r_{ij}} + (\alpha i \cdot \nabla_i)(\alpha j \cdot \nabla_j)\frac{\cos(wijrij/c)-1}{w_{ij}^2 r_{ij}/c^2}\right] \quad (16)$$

The photon frequency w_{ij}, used by the RCI program in calculating the matrix elements of the transverse photon interaction, is taken as the difference of the diagonal Lagrange multipliers associated with the orbitals. The leading quantum electrodynamic (QED) corrections effects, in the form of self-energy and vacuum polarization, are also included. The CSFs are given in the *jj*-coupling scheme during this procedure, but to make a comparison with experiments more feasible, transforming the resulting wave function to the *LSJ*-coupling scheme. Once well-converged and effectively complete ASFs have been obtained radiative transition such as transition probabilities and weighted oscillator strengths can be determined. The transition parameters between two states γJM and $\gamma' J' M'$ are expressed in terms of reduced matrix elements.

$$\langle \Psi(\gamma PJ)||T||\Psi(\gamma' P' J')\rangle = \sum_{jk} c_j c_k{}' \langle \Phi(\gamma_j PJ)||T||\Phi(\gamma_k{}' P' J')\rangle \quad (17)$$

Where **T** is the transition operator. For electric multipole transitions, there are two forms of the transition operator, the length (Babushkin gauge) and velocity (Coulomb gauge) forms. Due to the definition of these two, the length form is more sensitive to the outer part of the wave functions, which are usually active in radiative transitions. Several studies have shown that the length form generally gives more reliable values at a given level of valence and core-valence electron correlation although the velocity form seems to be more stable for transitions including highly excited Rydberg states. It is common to use the agreement between transition rates A_l and A_v computed in two forms as an indicator of the accuracy of the wave functions. A possible measure of this is the quantity dT, characterizing the uncertainty of the calculated transition rates and defined as

$$dT = \frac{|A_1 - A_v|}{\max(A_1, A_v)} \tag{18}$$

where A_l and A_v are the transition rates in length and velocity forms respectively. The values of dT do not represent uncertainty estimates for each transition but should be considered as statistical indicators of uncertainties within given sets of transitions. Because of the extraordinarily complicated character of the MCDHF formalism—particularly the massive number of CSFs required to adequately represent electron correlation and relativistic effects—solving the ensuing equations analytically becomes impossible, even for comparatively light atoms.

The variational method to orbital and mixing coefficient optimization produces enormous coupled integro-differential equations that must be self-consistently solved. Incorporating Breit interaction and QED effects, as well as switching between coupling schemes (jj ↔

LSJ), necessitates managing huge matrix elements and analyzing multi-electron integrals across a broad basis set.

As a result, the MCDHF and RCI methods are computationally implemented using specialized atomic structure codes such as GRASP2018, which can efficiently generate and manage large CSF spaces, numerically solve Dirac equations for radial orbitals, perform self-consistent field (SCF) iterations, compute transition matrix elements in different gauges, and estimate uncertainties using gauge agreement indicators such as d_T. This computational tool is critical for practical applications of the theory, allowing for high-precision predictions of atomic energy levels, lifetimes, and transition probabilities that can be directly compared to actual data.

III. Experiment

While the theoretical underlying for MCDHF and RCI is well established, the computational expense of performing these calculations—especially for heavy components or systems requiring massive configuration expansions—is sometimes a considerable obstacle. Running GRASP2018 on large atomic systems with significant CSF expansions can take hours or days, even on high-performance clusters. Furthermore, not all CSFs make a substantial contribution to the final atomic state functions, particularly in weakly mixing configurations or highly excited orbitals.

To overcome this, a NN was integrated into the atomic structure workflow to intelligently decrease computational cost while maintaining accuracy. Specifically, a NN model that predicts the possibility that a specific CSF would significantly contribute to the overall

ASF was created. This enables us to exclude low-probability alternatives early in the process and concentrate computing effort on the most important wavefunctions. To do this, a unique Linux-based computing cluster made of recycled CPUs was built, which allows GRASP2018 to operate locally in parallel across many nodes. This architecture eliminates the need for external HPC resources such as OSC, which reduces wait times and gives us more control over task scheduling and data processing. The NN serves as a pre-selection tool for GRASP2018, generating a prioritized list of configurations based on expected significance. As a consequence, it may drastically reduce the overall number of CSFs analyzed, resulting in faster convergence and shorter calculation times, particularly during the SCF and RCI phases.

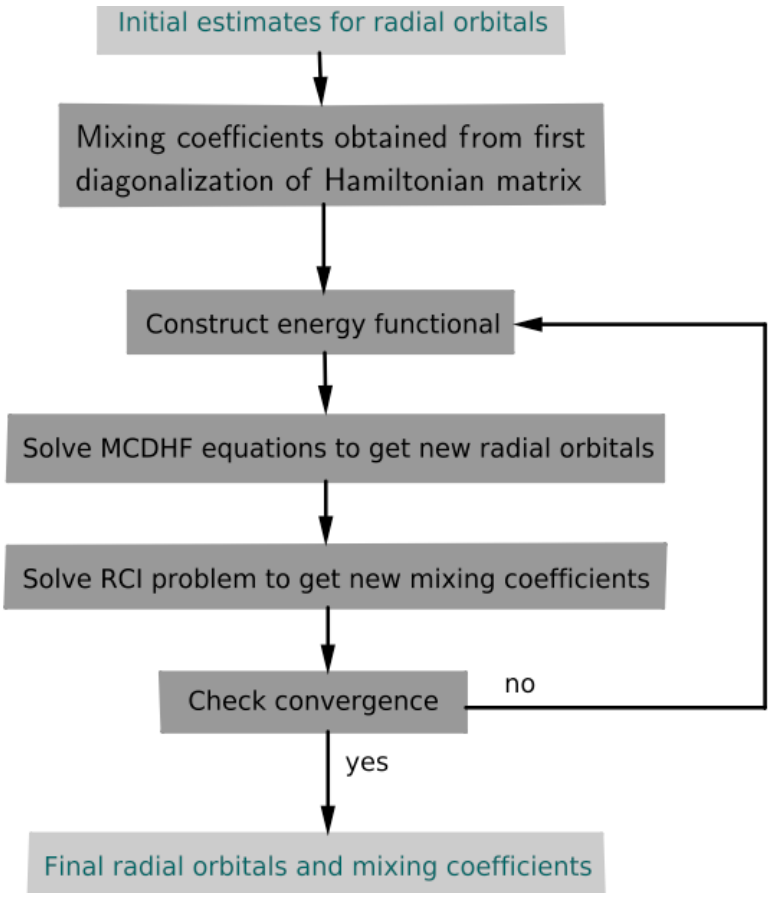

Figure 2 GRASP2018 Structure

Figure 2 visually represents the process of atomic structure calculations using GRASP2018. The process begins with an initial guess for the radial orbitals. Next, the self-consistent Dirac-Hartree-Fock (DHF) iterations are solved to determine the orbital wavefunctions. The radial part of the one-electron wavefunctions is then optimized by minimizing the system's total energy. After obtaining solutions, their quality is assessed. The optimization and calculation of residuals are repeated until convergence is achieved. Once convergence standards are met, various atomic properties are calculated, including transition probabilities, energy levels, and oscillator strengths. The NN would be integrated

into the structure before checking for convergence to discard the wave functions that would not affect the results.

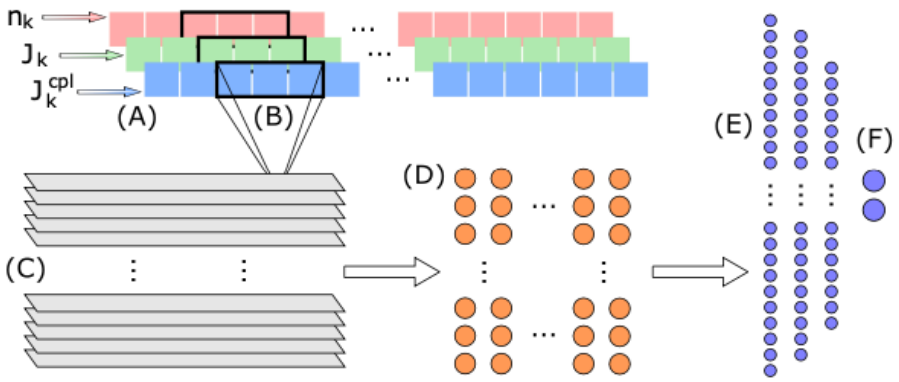

Figure 3 NN Architecture[23]

A network architecture that was found to work efficiently and which focuses on here is shown in Figure 3. The input layer (A) consists of 3 channels (see color encoding in Fig. 2) of size N. The input is processed with a filter kernel of size 3 (B) resulting in 96 feature maps (C)each of size N−2. A filter kernel is a small matrix of learnable weights that slides over the input data, performing a convolution operation. At each position, it extracts local patterns by computing a weighted sum over a small neighborhood of input values. The latter are mapped to 16 feature maps (D) of size N−2 by application of a filter kernel of size 1 (thus representing a purely local transformation). The obtained output of 16×(N−2) values is then flattened and forwarded to a network of 4 dense layers (E) with 150, 120, 90, and 2 neurons, respectively. The 96 feature maps are compressed into 16(D) using another filter of size 1, which performs a local transformation without looking at surrounding values.

The result is a set of 16 vectors, each of length N-2, which are flattened into a single vector and passed into a series of four fully connected layers (E).

The Relu activation function was used throughout the NN apart from the two-neuron output layer (F), where the SoftMax function is applied yielding the probabilities of the CSF to be important or unimportant. The categorical cross-entropy was chosen as the loss function. The NN is trained on batches of size 32 using the Adam optimization algorithm with early stopping. The described NN functionality was implemented using the Python library Kera's[24] which is a high-level interface of the TensorFlow library[25].

This NN architecture, together with the strategy of CSF selection and iterative re-training described above, constitutes our new approach to atomic structure calculations. To benchmark, it started with a somewhat sized basis set that can be directly calculated using GRASP2018. Selecting SD^*(3p,9h) for the Re atom ground state (4,267,362 CSFs with 64 ×3 quantum numbers). D^* is a subset of D with the limitation that each virtual orbital can be doubly excited or empty.

A specialized neural network was designed to help predict which atomic configurations are most important in large quantum calculations. The network processes input data using small filters that detect patterns, then passes the results through several layers to decide whether each configuration matters. It uses a common training approach to improve its accuracy and was built using popular machine learning tools like Keras and TensorFlow. This model was tested on a large atomic system to see how well it could reduce the number of configurations needed, making the calculations much faster without losing accuracy.

IV. Discussion and Conclusion

The NN architecture was created to process features derived from electronic configurations, such as orbital quantum numbers, occupancy patterns, and energy estimates, and return a probability-like score indicating a given configuration's expected contribution to the total atomic state function (ASF). This technique was motivated by the methods described in *Deep-learning approach for the atomic configuration interaction problem on large basis sets[23]*, which used a deep learning framework to help reduce the configuration space in multiconfiguration Dirac-Hartree-Fock (MCDHF) computations. Their research shows that merging machine learning with standard atomic structure codes may drastically reduce the amount of configuration state functions (CSFs) evaluated while retaining excellent accuracy in computed observables. Our first goal was to duplicate this work with our custom-made computational infrastructure, which comprises a Linux-based CPU cluster created from recycled hardware and set to run GRASP2018 in a parallel environment. However, obstacles were faced during the replication process. The original study used high-performance computing (HPC) capabilities and curated datasets, which were either unavailable or incompatible with our solution.

Furthermore, the NN models employed in the study required a large amount of labeled training data, most of which was created under computing conditions that were above our capabilities. In light of these limitations, the project was redesigned with a more manageable scope. It was decided to shift efforts to create a supervised learning model capable of forecasting MCDHF total energies using inputs from simpler HF computations which are less computationally costly, making them ideal for quick data production in a

wide range of atomic systems. The revised approach reformulated the problem as a regression job, where the NN was trained to learn the nonlinear mapping between HF-level characteristics and the related MCDHF total energies. The major goal of this technique was to investigate how well machine learning may approximate the results of fully relativistic MCDHF computations, minimizing the requirement for extensive self-consistent field (SCF) operations in circumstances when predictions are sufficiently accurate.

As part of this revised method, now gathering training data from the Ohio Supercomputer Center (OSC) to create a huge dataset of HF and MCDHF total energies. The goal is to thoroughly investigate several neural network architectures and determine the best model structure for effectively forecasting MCDHF energies from HF-derived variables. For this revised method, data was created through Dr. Heiko Hergert approach covered on their graduate level Nuclear Physics course at Michigan State University[26]. Through various calculations their program is able to calculate the Couple Cluster, Reference and Perturbation energies for different molecules. For this method it was implemented to calculate the energies for full shell molecules up to 114 electrons and 7 shells with densities ranging from 0.0 to 1.

The revised method attempted to explore the relationship between the energies and the MCDHF method with the pattern recognition capabilities of NNs. By implementing a simple structure of a single layers, having 32 neurons and a single output layer with 5 input features (number of particles, density, number of shells, couple cluster energy and perturbation energy) that focuses on the relationship between $E_{HF} = E_{CCD} + (E_{CCD} -$

E_{Embpt2}) where E_{CCD} is the couple cluster energy, and E_{Embpt2} being the perturbation energy.

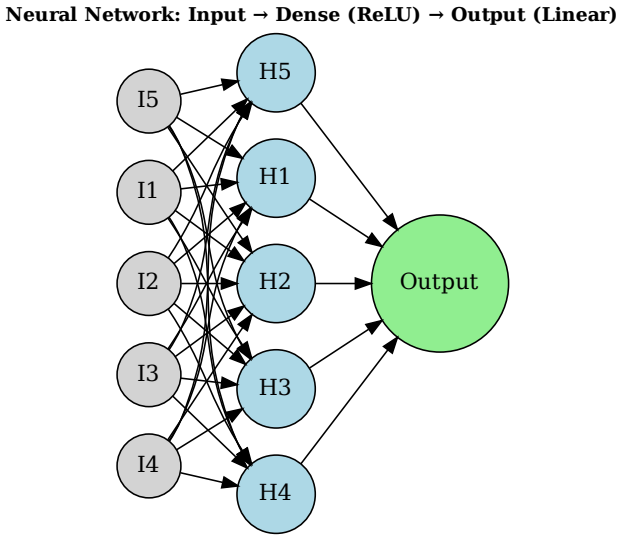

Figure 4 Visual Representation of NN Structure[27]

This method recorded a surprisingly high accuracy for predicting E_{HF} with an accuracy of about 98%. At a superficial level this model looks ideal for predicting the energy values for other molecules, in addition when plotting the prediction results on a plot to find the fit we find a linear fit for the model which can be seen in Figure 5.

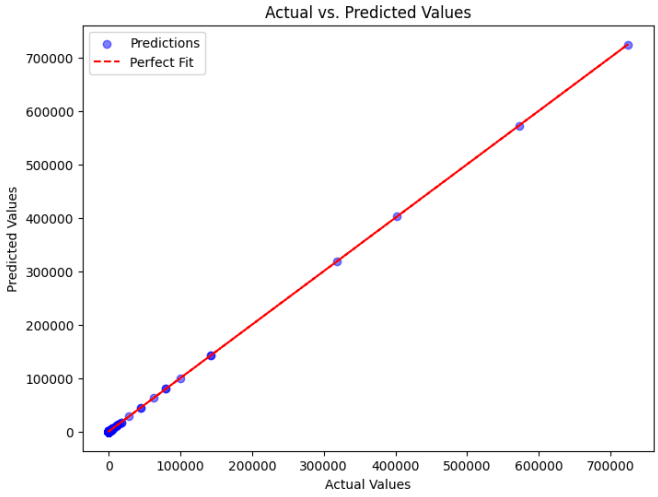

Figure 5 NN Prediction vs Actual Values

While the results from this method initially appeared promising, they raised an important question: if a simple neural network can achieve such high accuracy, why hasn't this approach been used before? Upon further analysis, a key limitation became evident. The training data selection was completely random, and because the energy values ranged from -0.02 to $1.2x10^6$, Eh (where Eh is the Hartree unit, equal to 27.211 eV), each training cycle used a very different set of energy scales. This variability made the prediction task inconsistent, as the neural network was forced to learn over widely differing energy ranges. As a result, the model struggled to predict the Hartree-Fock energy E_{HF} accurately across different systems.

Although the model could not fully reproduce the results of the original study due to computational and resource constraints—including limited access to high-performance computing, reduced configuration sampling, and smaller training datasets—this project still provided valuable insights into both neural network architecture and the challenges of atomic structure modeling. Designing a model that could learn meaningful

relationships between approximate and relativistic energy outputs required thoughtful experimentation with network depth, input encoding, and loss behavior across different training regimes.

In particular, this study emphasized the importance of consistent training data ranges, balanced input representations, and dataset diversity—factors that significantly affect the performance of regression models in physical systems. Additionally, the shift from a classification-based approach to an energy-based regression model revealed the potential of neural networks to bypass the need for exhaustive configuration interaction expansions, provided that input quality and model architecture are carefully optimized. Despite the obstacles, this work lays a strong foundation for further investigation into the integration of machine learning with multiconfiguration Dirac-Hartree-Fock (MCDHF) methods. Future studies can build on these findings by incorporating more scalable input pipelines, larger datasets, or physics-informed neural networks that embed symmetry and conservation laws directly into the model. With continued development, data-driven models may significantly accelerate relativistic electronic structure calculations and offer a new computational pathway for studying large, complex atoms and ions.

V. References

[1] C. F. Fischer, M. Godefroid, T. Brage, P. Jönsson, and G. Gaigalas, "Advanced multiconfiguration methods for complex atoms: I. Energies and wave functions," *J. Phys. B: At. Mol. Opt. Phys.*, vol. 49, no. 18, p. 182004, Sep. 2016, doi: 10.1088/0953-4075/49/18/182004.

[2] "Configuration interaction - Wikipedia." Accessed: Apr. 02, 2025. [Online]. Available: https://en.wikipedia.org/wiki/Configuration_interaction

[3] "The Hartree-Fock Approximation." Accessed: Apr. 02, 2025. [Online]. Available: https://web1.eng.famu.fsu.edu/~dommelen/quantum/style_a/hf.html

[4] B. Atalay, T. Brage, P. Jönsson, and H. Hartman, "MCDHF and RCI calculations of energy levels, lifetimes, and transition rates in Si III and Si IV," *A&A*, vol. 631, p. A29, Nov. 2019, doi: 10.1051/0004-6361/201935618.

[5] "Perturbation Theory," Chemistry LibreTexts. Accessed: Apr. 02, 2025. [Online]. Available: https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Advanced_Theoretical_Chemistry_(Simons)/04%3A__Some_Important_Tools_of_Theory/4.01%3A_Perturbation_Theory

[6] "Basics of the Monte Carlo method - Computational Physics." Accessed: Apr. 02, 2025. [Online]. Available: https://compphys.quantumtinkerer.tudelft.nl/proj2-monte-carlo/

[7] W. Jeong, C. A. Gaggioli, and L. Gagliardi, "Active Learning Configuration Interaction for Excited-State Calculations of Polycyclic Aromatic Hydrocarbons," *J. Chem. Theory Comput.*, vol. 17, no. 12, pp. 7518–7530, Dec. 2021, doi: 10.1021/acs.jctc.1c00769.

[8] C. Froese Fischer, G. Gaigalas, P. Jönsson, and J. Bieroń, "GRASP2018—A Fortran 95 version of the General Relativistic Atomic Structure Package," *Computer Physics Communications*, vol. 237, pp. 184–187, Apr. 2019, doi: 10.1016/j.cpc.2018.10.032.

[9] D. E. Salhi, S. B. Nasr, S. Manai, and H. Jelassi, "Multiconfiguration Dirac–Hartree–Fock energy levels, weighted oscillator strengths, transitions probabilities, lifetimes, hyperfine interaction constants, Landé g-factors and isotope shifts of *O VII*," *Results in Physics*, vol. 23, p. 103960, Apr. 2021, doi: 10.1016/j.rinp.2021.103960.

[10] David Sherrill, *Intro to Electronic Structure Theory Part 3*, (Feb. 04, 2021). Accessed: Apr. 02, 2025. [Online Video]. Available: https://www.youtube.com/watch?v=PyLrTJhrkVE

[11] "6: Electronic Structure," Chemistry LibreTexts. Accessed: Apr. 02, 2025. [Online]. Available: https://chem.libretexts.org/Bookshelves/Physical_and_Theoretical_Chemistry_Textbook_Maps/Advanced_Theoretical_Chemistry_(Simons)/06%3A_Electronic_Structure

[12] "molecular physical chemistry - Molecules and Born-Oppenheimer," BORZUYA UNIVERSITY. Accessed: May 01, 2025. [Online]. Available: http://brussels-scientific.com/?p=6431

[13] David Sherrill, *Intro to Electronic Structure Theory Part 1*, (Jan. 30, 2021). Accessed: Apr. 02, 2025. [Online Video]. Available: https://www.youtube.com/watch?v=srL6i7zwzu4

[14] David Sherrill, *Intro to Electronic Structure Theory Part 2*, (Feb. 02, 2021). Accessed: Apr. 02, 2025. [Online Video]. Available: https://www.youtube.com/watch?v=qrFE1Uubi4c

[15] David Sherrill, *Introduction to Hartree-Fock Molecular Orbital Theory Part 1*, (Feb. 07, 2021). Accessed: Apr. 02, 2025. [Online Video]. Available: https://www.youtube.com/watch?v=qcYxyP_SDLU

[16] "Slater determinant," *Wikipedia*. Jan. 29, 2025. Accessed: Apr. 02, 2025. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Slater_determinant&oldid=1272612105

[17] "Slater–Condon rules," *Wikipedia*. Dec. 03, 2024. Accessed: Apr. 02, 2025. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Slater%E2%80%93Condon_rules&oldid=1260975432

[18] "Functional (mathematics)," *Wikipedia*. Nov. 05, 2024. Accessed: May 01, 2025. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Functional_(mathematics)&oldid=1255507319

[19] "Coulomb operator," *Wikipedia*. Jan. 21, 2024. Accessed: Apr. 02, 2025. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Coulomb_operator&oldid=1197786438

[20] "Exchange operator," *Wikipedia*. Sep. 02, 2024. Accessed: Apr. 02, 2025. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Exchange_operator&oldid=1243662773

[21] C. Froese Fischer, G. Gaigalas, P. Jönsson, and J. Bieroń, "GRASP2018—A Fortran 95 version of the General Relativistic Atomic Structure Package," *Computer Physics Communications*, vol. 237, pp. 184–187, Apr. 2019, doi: 10.1016/j.cpc.2018.10.032.

[22] "Dirac equation," *Wikipedia*. Feb. 23, 2025. Accessed: Apr. 03, 2025. [Online]. Available: https://en.wikipedia.org/w/index.php?title=Dirac_equation&oldid=1277286300

[23] P. Bilous, A. Pálffy, and F. Marquardt, "Deep-learning approach for the atomic configuration interaction problem on large basis sets," Jun. 21, 2023, *arXiv*: arXiv:2209.05867. doi: 10.48550/arXiv.2209.05867.

[24] "Keras: Deep Learning for humans." Accessed: Apr. 03, 2025. [Online]. Available: https://keras.io/

[25] "TensorFlow." Accessed: Apr. 03, 2025. [Online]. Available: https://www.tensorflow.org/

[26] *ManyBodyPhysics/LectureNotesPhysics*. (Jan. 28, 2025). C++. ManyBodyPhysics. Accessed: May 01, 2025. [Online]. Available: https://github.com/ManyBodyPhysics/LectureNotesPhysics

[27] OpenAI ChatGPT, *AI-Generated Diagram of Molecular Hamiltonian and Orbital Radii*. 2025.