Optimization of Self-Consistent Equation with Machine Learning
Research Report PIBIC UFRGS 2022/2023*

Student: Giovanni Benedetti da Rosa
Advisor: David Möckli

I. INTRODUCTION

Mean-field or self-consistent approximations are commonly employed in various areas of statistical physics that involve interactions among multiple agents, such as magnetism[1], epidemic models[2], and superconductivity[3]. These approximations are useful in determining numerical values for order parameters in phase transitions of these systems. The method involves approximating these interactions by the average contributions of the agents (degrees of freedom), thereby neglecting terms of smaller magnitude than the fluctuations associated with these averages. However, such a formalism **always** leads to solving self-consistent (implicit) equations for the order parameter, which sometimes only have numerical solutions requiring extremely high computational time and storage.

Thus, this work aims to reduce the computational time required for these solutions and, consequently, enable the exploration of more complex systems with higher dimensions. To achieve this, we chose the Bogoliubov–de Gennes (BdG) equations for superconductivity and decided to test machine learning methods to solve the problem of finding the superconducting gap (Δ).

II. THE EQUATIONS

In the continuous space, the BdG equation is represented as follows:

$$\Delta = \lambda \int_{-E_c}^{E_c} \Delta \frac{\tanh\left(\frac{\sqrt{\xi^2+|\Delta|^2}}{2K_bT}\right)}{2\sqrt{\xi^2+|\Delta|^2}} d\xi \tag{1}$$

Where: - Δ is the superconducting gap, - λ represents a parameter, - E_c is a cutoff energy, - ξ is the energy variable, - K_b is the Boltzmann constant, - T is the temperature.

In the discrete space, a Hamiltonian is chosen to represent a one-dimensional chain with nearest-neighbor hopping and periodic boundary conditions:

$$H = \sum_{\langle i,j \rangle} t_{ij} c_i^\dagger c_j + \sum_{\langle i,j \rangle} \Delta_{ij} c_i^\dagger c_j^\dagger + \sum_{\langle i,j \rangle} \Delta_{ij} c_i c_j \tag{2}$$

Where: - H is the Hamiltonian operator, - $\langle i,j \rangle$ denotes summation over nearest-neighbor pairs, - t_{ij} represents the hopping amplitude between sites i and j, - Δ_{ij} represents the pairing potential between sites i and j, - $c_i^\dagger$ and c_i are the creation and annihilation operators, respectively, for particles at site i.

III. METHODS

In this section, we will present the existing numerical methods for solving the self-consistent equation and machine learning algorithms useful for addressing the problem.

A. Brute Force Method

To solve the equations using the brute force method, the following algorithm is followed, where f is a function that depends on Δ, and ϵ is the estimated value for the convergence criterion:

$\Delta_{\text{in}} \leftarrow$ random value
while $|f(\Delta_{\text{in}}) - \Delta_{\text{in}}| < \epsilon$ **do**
    $\Delta_{\text{in}} \leftarrow f(\Delta_{\text{in}})$
end while

The brute force method involves iteratively updating the value of Δ_{in} with $f(\Delta_{\text{in}})$ until the condition $|f(\Delta_{\text{in}}) - \Delta_{\text{in}}| < \epsilon$ is met, where ϵ represents the desired accuracy or convergence threshold. This approach can be computationally intensive, especially for complex systems with high-dimensional parameter spaces. Notice that the calculation process of the "function" f is performed repeatedly. For the equation 1, solved n times using the Trapezoidal method, its complexity would be $\mathcal{O}(n.N^3)$, where N is the number of subintervals chosen.

In the discrete basis, the process becomes even more costly, as it requires diagonalizing a matrix of size N and computing the value of Δ according to the Bogoliubov–de Gennes equations. Below, the diagonalization of the Hamiltonian is represented after a Bogoliubov transformation and mean-field approximation. The result is visualized in a $2N \times 2N$ space, where N is the number of sites in the system [3].

* I was a scholarship holder for this project for one month and continued as a volunteer. I presented these results at a meeting of the condensed matter physics group of UFRGS. The presentation can be viewed at this link: https://docs.google.com/presentation/d/14N4MZr7nkdCyMEPjoDZhc2xSragQ9VZS_V1l0fhb8NA/edit?usp=sharing

$$\begin{bmatrix} K & \Delta \\ \Delta^\dagger & -K^T \end{bmatrix} \begin{bmatrix} u_n(i) \\ v_n(i) \end{bmatrix} = E_n \begin{bmatrix} u_n(i) \\ v_n(i) \end{bmatrix} \tag{3}$$

After this step, we can calculate the value of Δ_{ij}:

$$\Delta_{ij} = V \sum_n u_n(i) v^* n(j) \tanh\left(\frac{\beta E n}{2}\right) \tag{4}$$

The entire process would result in a complexity of $\mathcal{O}(n.N^3)$, where N, in this case, is the dimension of the Hamiltonian.

The high computational complexity of both the continuous and discrete methods highlights the necessity of exploring more efficient approaches, such as the machine learning algorithms previously mentioned, to speed up the calculations and enable the study of larger and more complex systems in condensed matter physics.

B. Decision Trees and Random Forests

Decision Trees work through sequences of hierarchical binary decisions. These decision criteria are based on splits in the vector space containing the data represented by their potential predictors. These splits are made to minimize an error metric such as Shannon's entropy or Mean Squared Error (MSE) within a particular subspace. Once this region is found, the process continues iteratively based on the preceding splits until a stopping criterion is met. When using a set of decision trees, a random sample of the original dataset is taken to create N independent trees. The output for the problem is determined by the set that, on average, provides the best classification across all the trees. This procedure is known as Random Forests.

Random Forests perform well for low-dimensional data and have a computational complexity of the order of $\mathcal{O}(D \cdot T)$, where D is the depth of the tree and T is the number of trees [6].

IV. PROPOSAL

Considering the problem and the computational complexities mentioned, one might wonder: could the machine learn the existing patterns in determining Δ and find a function f such that $\Delta \leftarrow f(T, V, N, E_c, \text{etc})$? By using only the physical parameters that determine the system, it could optimize the computational time.

To achieve this, the idea is to use Random Forests due to their low computational complexity and the low dimensionality of the physical parameters. The goal is to improve the initial "guess" for Δ, which is typically random. The machine would then learn this pattern and provide a more accurate guess, thus optimizing the computational time.

V. PRELIMINARY RESULTS

So far, tests have only been performed on the continuous basis. A dataset of 18,000 data points was generated, with values of $\lambda \in (0, 1)$, $E_c \in (1000, 10000)$, and $T \in (0, 90)$. Hyperparameter tuning and 5-fold cross-validation were used for training and testing the algorithm.

Métricas	Valores
R^2	0,98
RMSE(Root Mean Squared Error)	237,98
MAPE(Percent Mean Absolute Error)	3,49

TABLE I. Métricas Random Forest

On average, the Brute Force method took about $0.056s$ per convergence value, while the Random Forest took $8.85 \times 10^{-6}s$. This represents a speed-up of almost 1000 times. However, it is essential to consider that all machine learning methods have an associated intrinsic error (ϵ). Therefore, to achieve the desired result according to our proposal, the Random Forest method took an average of $0.035s$. Thus, the proposed method, Random Forest Bogoliubov-de Gennes (RF-BdG), seems to be around 2 times faster.

VI. CONCLUSION AND FUTURE WORK

Based on the tests conducted so far, there are indications that Machine Learning techniques can indeed optimize the computational time for calculating order parameters using the BdG formalism. The preliminary results with Random Forests show promising improvements in speed and accuracy compared to traditional brute force methods.

The preliminary tests indicate that Machine Learning techniques can optimize the computational time for calculating order parameters using the BdG formalism. The use of Random Forests shows promising improvements in speed and accuracy compared to traditional brute force methods. Future work includes extending the investigations to the discrete case, exploring higher-dimensional systems, and applying this approach to other self-consistent problems in condensed matter physics. By leveraging data-driven methods, researchers have the potential to overcome computational barriers and accelerate scientific discoveries in this field.

[1] SAKTHIVADIVEL, D. A. R. *Magnetisation and Mean Field Theory in the Ising Model.* SciPost Physics Lecture Notes, 19 jan. 2022.

[2] FRIESWIJK, K. et al.*A mean-field analysis of a network behavioural-epidemic model.* IEEE Control Systems Letters, v. 6, p. 2533–2538, 2022.

[3] FRANZ, M.; KALLIN, C.; BERLINSKY, A. J.*Impurity scattering and localization ind-wave superconductors.* Physical Review B, v. 54, n. 10, p. R6897–R6900, 1 set. 1996.

[4] NAGAI, Y. *N-independent Localized Krylov–Bogoliubov-de Gennes Method: Ultra-fast Numerical Approach to Large-scale Inhomogeneous Superconductors.* Journal of the Physical Society of Japan, v. 89, n. 7, p. 074703, 15 jul. 2020.

[5] NAGAI, Y. *N-independent Localized Krylov–Bogoliubov-de Gennes Method: Ultra-fast Numerical Approach to Large-scale Inhomogeneous Superconductors.* Journal of the Physical Society of Japan, v. 89, n. 7, p. 074703, 15 jul. 2020.

[6] SOLÉ, X.; RAMISA, A.; TORRAS, C. *Evaluation of Random Forests on large-scale classification problems using a Bag-of-Visual-Words representation. [s.l: s.n.].*