

STUDY OF NANOSTRUCTURES COUPLED WITH A SUPERCONDUCTING LEAD VIA THE NEURAL-NETWORK QUANTUM STATES METHODS

^{1,2}Jana KODRLOVÁ, ¹Martin ŽONDA, ²Martin FRIÁK

¹Charles University, Faculty of Mathematics and Physics, Prague, Department of Condensed Matter Physics, Ke Karlovu 2026/5, Prague, 120 00, Czech Republic, EU, kodrlava@ipm.cz, martin.zonda@matfyz.cuni.cz

²Institute of Physics of Materials, v.v.i., Czech Academy of Sciences, Žitkova 22, 616 00, Brno, Czech Republic, EU, friak@ipm.cz

<https://doi.org/10.37904/metal.2025.5156>

Abstract

The Neural-Network Quantum States (NNQS) method is rapidly emerging as a powerful tool for investigating quantum many-body physics. By combining variational Monte Carlo techniques with neural network-based variational functions, this approach leverages the remarkable advancements in deep learning achieved in recent years. While substantial progress has been made in simulating magnetic systems on lattices, simple molecules, and even continuous systems, the ab initio simulation of complex strongly correlated electron systems continues to pose significant challenges. With the help of the generalized atomic limit (GAL) – a recently developed model describing a system of quantum dots coupled to a superconducting lead – we attempt to show the efficiency and accuracy of NNQS, mainly the Restricted Boltzmann Machine, by comparing the acquired results to other available methods such as exact diagonalization. The simultaneous study of the system's properties such as the energy spectrum and quantum phase transitions could bring advancements in electronics, sensors or the design of high-quality qubits used in quantum computers.

Keywords: generalized atomic limit, quantum dot, superconducting lead, neural-network quantum states, restricted Boltzmann machine, Andreev bound states

1. INTRODUCTION

The hybrid systems made of quantum dots embedded onto the surface of a superconducting lead have garnered significant attention due to their potential applications and ability to probe fundamental physical phenomena [1-3]. A quantum dot, which can be thought of as an artificial atom, is a nanostructure capable of carrying a small number of electrons, effectively capturing them inside a potential well on discrete energy levels. Coupled to a superconductor, a condensate state of boson-like electron pairs, this strongly correlated compound system is difficult to accurately model.

We employ GAL – a recently proposed extension of the superconducting atomic limit of the superconducting impurity Anderson model. GAL can faithfully capture the Andreev bound states, i.e. sharp bound states within the superconducting gap, boundaries of quantum phase transitions, as well as other measurables.

Even though traditional numerical approaches have been crucial in advancements in strongly correlated many-body problems, recent progress in machine learning has shown a promising implementation in quantum physics and many other fields of all natural sciences. By combining the variational Monte Carlo methods with neural network-based approaches, the use of neural-network quantum states could be advantageous in approximating the complexity of many-body wavefunctions necessary for the study of highly correlated systems.

The restricted Boltzmann machine (RBM), a shallow neural-network model with rather simple architecture, is especially of interest due to its flexibility and simple implementation and it has been successfully applied to systems of quantum spins or correlated fermions in non-superconducting environments [4].

Leveraging a notable advantage of GAL, that is its straightforward implementation to complex quantum dot systems, we focus our study on linear chains of quantum dots through the method of exact diagonalization (ED) based on the Lanczos method and then use these results as a benchmark for the RBM model.

2. THEORY OVERVIEW

The basic structure of the model is based on the impurity Anderson model and the Bardeen-Cooper-Schrieffer theory of superconductivity. The atomic limit Hamiltonian can be written as

$$\begin{aligned} \mathcal{H} = & \sum_{j=1}^N \sum_{\sigma} \epsilon_j d_{j\sigma}^{\dagger} d_{j\sigma} + \sum_{j=1}^N U_j (d_{j\uparrow}^{\dagger} d_{j\uparrow} d_{j\downarrow}^{\dagger} d_{j\downarrow}) - t \sum_{\sigma} \sum_{\langle j,i \rangle} (d_{j\sigma}^{\dagger} d_{i\sigma} + d_{i\sigma}^{\dagger} d_{j\sigma}) \\ & - \Gamma \sum_{j=1}^N (d_{j\uparrow}^{\dagger} d_{j\downarrow}^{\dagger} + d_{j\downarrow} d_{j\uparrow}) - \Gamma \sum_{\langle j,i \rangle} \zeta_{\langle j,i \rangle} (d_{j\uparrow}^{\dagger} d_{i\downarrow}^{\dagger} + d_{i\uparrow}^{\dagger} d_{j\downarrow}^{\dagger} + \text{H.c.}), \end{aligned} \quad (1)$$

where $d_{j\sigma}^{\dagger}$ and $d_{j\sigma}$ are the creation and annihilation operators of a fermion on site j with spin $\sigma = -1/2, 1/2$. The dots are thought of as having a single energy level ϵ and can carry up to two electrons to satisfy the Pauli exclusion principle. This energy can be changed with an applied gate voltage. Adding a second fermion to an already occupied dot is penalized by the charging energy U . In the case of multiple dots, fermions can hop directly between the dots, denoted by the hopping constants t . The coupling constant Γ quantifies the induced superconducting behavior resulting from the proximity effect of the superconducting lead. Lastly, the dots can also influence each other in a non-local fashion via a lead-mediated interaction. This is introduced through the coherence length ζ , which correlates with the spatial distance between the dots. The sum over $\langle j, i \rangle$ in Equation (1) expresses how the interaction is restricted to only neighboring dots and H.c. in the last term stands for the Hermitian conjugate of the rest of the term inside the brackets. The Hamiltonian (1) results from the superconducting atomic limit, that assumes that the superconducting gap $\Delta \rightarrow \infty$ [5]. Even though this limit does not realize in real systems where Δ is typically smaller than U and other energy scales, it preserves the key features of the system such as the non-local interdot interaction.

2.1 Generalized Atomic Limit

Reintroducing the superconducting gap Δ (which also represents the energy unit in what follows) back to the model is done by first shifting the first two terms in Equation (1) as $\epsilon_j = \epsilon + U_j/2$, which ensures that for $\epsilon_j = 0$, the average occupation per dot is one electron, corresponding to the half-filling regime. The rescaling coefficients are found through perturbation theory in U (the details can be found in [6]), while neglecting any contribution from the quasiparticle continuum above Δ and ensuring the correct behavior of Andreev bound states near quantum phase transitions (QPTs) observed at zero temperature. The new coefficients are given as

$$\tilde{\epsilon} = v\epsilon, \quad \tilde{U} = v^2 U, \quad \tilde{t} = vt, \quad \tilde{\Gamma} = v, \quad \text{where} \quad v = \frac{1}{1 + \frac{\Gamma}{\Delta}}.$$

We make additional assumptions about our system, such that there is no direct hopping between dots $t = 0$, the dots are identical $U_j = U$ and are connected to the same gate voltage $\epsilon_j = \epsilon = 0$. The distance between two adjacent dots is constant within the chain, thus $\zeta_{\langle j,i \rangle} = \zeta$. The resulting GAL Hamiltonian reads

$$\begin{aligned} \mathcal{H}^{\mathcal{GAL}} = & v\epsilon \sum_{j=1}^N \sum_{\sigma} \left(d_{j\sigma}^{\dagger} d_{j\sigma} - \frac{1}{2} \right) + v^2 \frac{U}{2} \sum_{j=1}^N \left(d_{j\uparrow}^{\dagger} d_{j\uparrow} + d_{j\downarrow}^{\dagger} d_{j\downarrow} - 1 \right)^2 - v\Gamma \sum_{j=1}^N \left(d_{j\uparrow}^{\dagger} d_{j\downarrow}^{\dagger} + d_{j\downarrow} d_{j\uparrow} \right) \\ & - \zeta v\Gamma \sum_{\langle j,i \rangle} \left(d_{j\uparrow}^{\dagger} d_{i\downarrow}^{\dagger} + d_{i\uparrow}^{\dagger} d_{j\downarrow}^{\dagger} + d_{i\downarrow} d_{j\uparrow} + d_{j\downarrow} d_{i\uparrow} \right). \end{aligned} \quad (2)$$

2.2 Variational Monte Carlo

As the system grows in size and complexity, solving the Schrödinger equation in order to find the eigenvalues and eigenstates of the system becomes intractable for exact methods. Advanced numerical methods such as variational Monte Carlo (VMC) are then used to overcome this problem. In its core, VMC leverages the standard quantum mechanics variational principle. It states that given an expected energy E_0 of the true ground state Ψ_0 , any other trial wavefunction Ψ' with energy E' will comply with the inequality $\langle \mathcal{H} \rangle_{\Psi'} = E' \geq E_0 = \langle \mathcal{H} \rangle_{\Psi_0}$. Thus, we can approximate the ground state wavefunction by suggesting a suitable parametrized wavefunction $\Psi(\theta)$ and find an optimized set of θ^* such that $E(\theta^*) \approx E_0$. Given that the expected energy can be computed as

$$E(\theta) = \frac{\langle \Psi(\theta) | \mathcal{H} | \Psi(\theta) \rangle}{\langle \Psi(\theta) | \Psi(\theta) \rangle}. \quad (3)$$

However, the evaluation of $E(\theta)$ is not straightforward and methods such as Markov chain Monte Carlo must be used [7].

2.3 Restricted Boltzmann Machine

The basic idea behind the learning process in any machine learning task is finding an existing relationship, i.e. mapping, between the input and output data, denoted by $f(x)$, by setting a basic structure of that mapping $f'(x, \theta)$ and searching for the best set of parameters θ^* , so that $f(x) \approx f'(x, \theta^*)$. In the core of machine learning lies the loss function $\mathcal{L}$ and its minimization through an optimization process. The specific formula of $\mathcal{L}$ is chosen based on the task, nevertheless it provides a quantitative measure of how well our current model function $f'(x, \theta)$ agrees with the original input-output mapping. The new set of parameters is proposed according to an optimization method, usually based on gradient descent where the new set is chosen based on the steepest path to a minimum within the parameter space.

Troyer and Carleo were able to combine the variational principle mentioned above with techniques of machine learning in their 2017 paper [8] and introduced a new technique to tackle the many-body problem in quantum mechanics – neural-network quantum states. According to their approach, the loss function $\mathcal{L}$ takes on the form of Equation (3) and the trial wavefunction is represented with a neural network, specifically a model called the restricted Boltzmann machine. They were able to show that, despite its shallow and rather simple architecture, it is capable of giving accurate results for both the Ising and Heisenberg models.

Getting their inspiration from the human brain, the building blocks of a neural network are artificial neurons or nodes, that are composed into visible and hidden layers. The nodes in layer i receive an input from the previous layer $i - 1$ and give an output via an activation function $\mu(\sum_j w_{ij} v_j + b_i)$ which is evaluated based on the received input v_i from all nodes of the previous layer, the weight of the connection between the two nodes w_{ij} and the bias of the activated layer b_i . The activation function is a chosen nonlinear function to ensure its ability to express complex mappings. In the case of RBM, there is only one visible and one hidden layer. The visible layer acting as an input layer consists of N nodes v_i with a bias a_i , the hidden layer has M binary nodes. The

size of the network can be expressed through the ratio $\alpha = \frac{M}{N}$ – it defines the density of the network and correlates with its expressiveness. The resulting RBM variational function (machine learning model) is given as

$$\Psi(\theta) = \exp\left(\sum_i a_i v_i\right) \prod_i \left[2 \cosh\left(\sum_{i,j} w_{ij} v_j + b_i\right) \right] \quad (4)$$

3. RESULTS

3.1 Andreev Bound States via ED

The Andreev bound states (ABS) represent the many-body states observed in the single-particle spectral function as δ -peaks with energies within the interval $[-\Delta, \Delta]$ [9], hence they are called subgap states. Study of ABS is crucial in explanation of Majorana bound states [10], d -wave superconductors [11] and they can be also potentially implemented as quantum bits [12].

The spectrum of the GAL Hamiltonian, see Equation (2), changes with the charging energy U , the coupling constant Γ and the coherence length ζ . It is emphasized that not all energy levels found through ED necessarily correspond to ABS as they are subjected to certain selection rules and are not all observed in tunneling spectroscopy.

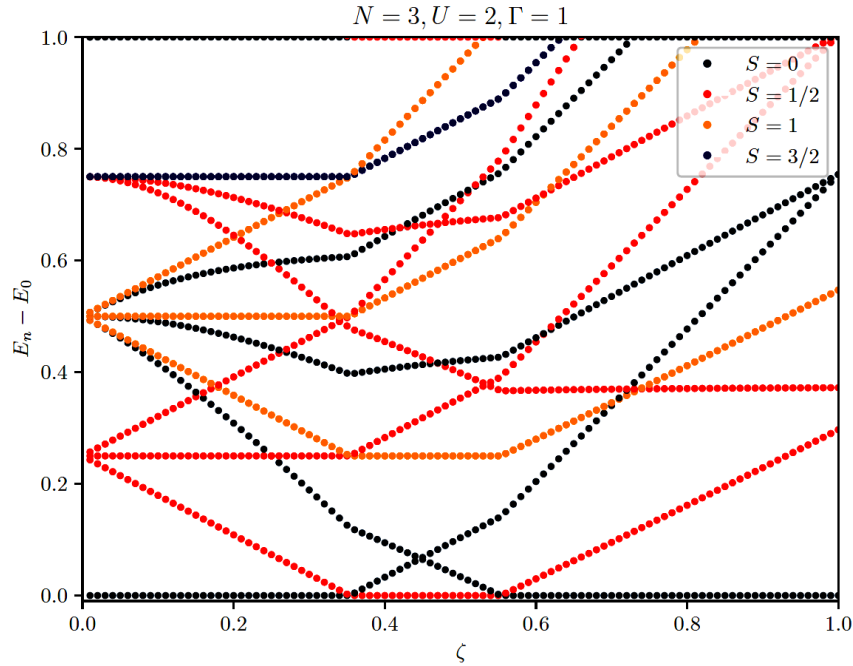


Figure 1 Subgap spectrum visualized through spin for a chain of $N = 3$ dots for $U = 2$ and $\Gamma = 1$

Figure 1 shows the subgap spectrum of $N = 3$ chain of dots. The y-axis shows the difference between n -th energy and the ground energy E_0 and its development as ζ changes from the serial regime of $\zeta = 0$ (where each dot has its own lead) to the parallel regime of $\zeta = 1$ (where two adjacent dots are connected to the same lead). Notable behavior of the ground state can be seen in **Figure 1** around $\zeta \approx 0.35$, where the first excited state touches the ground state. Since the found energies are color-coded according to the total spin S of their corresponding state, we can see that at the critical point $\zeta \approx 0.35$, the ground state goes from the singlet $S = 0$

to a doublet $S = 1/2$ state, as it seemingly exchanges place with the first excited state. At around $\zeta \approx 0.55$, we can see another phase transition, this time back to the more common singlet state.

3.2 Approximation of the Ground State via RBM

Whether the neural network has the ability to approximate the wavefunction is largely influenced by the optimal choice of hyperparameters. These are parameters that have to be set manually before the learning process begins. For this specific case, the hyperparameters (with basic explanation of their meaning) were set as follows: $\alpha = 5$ (the layer density), $N_{\text{chain}} = 3$ (number of Markov chains), $N_{\text{sample}} = 2048$ (total number of samples $s^{(i)}$ generated within each iteration across all chains) and $N_{\text{discard}} = 200$ (number of discarded samples as the chain was going through thermalization). The learning rate η was adjusted during the learning process – the initial value was $\eta = 0.08$ and it eventually decayed into $\eta = 0.001$.

Figure 2 shows the learning curve of RBM for $N = 3$ in the same setting as in **Figure 1**. We include the exact energy of the ground state E_{ED} , the expectation energy E_{RBM} and the expectation value of total spin squared S^2 of the wavefunction RBM converged to.

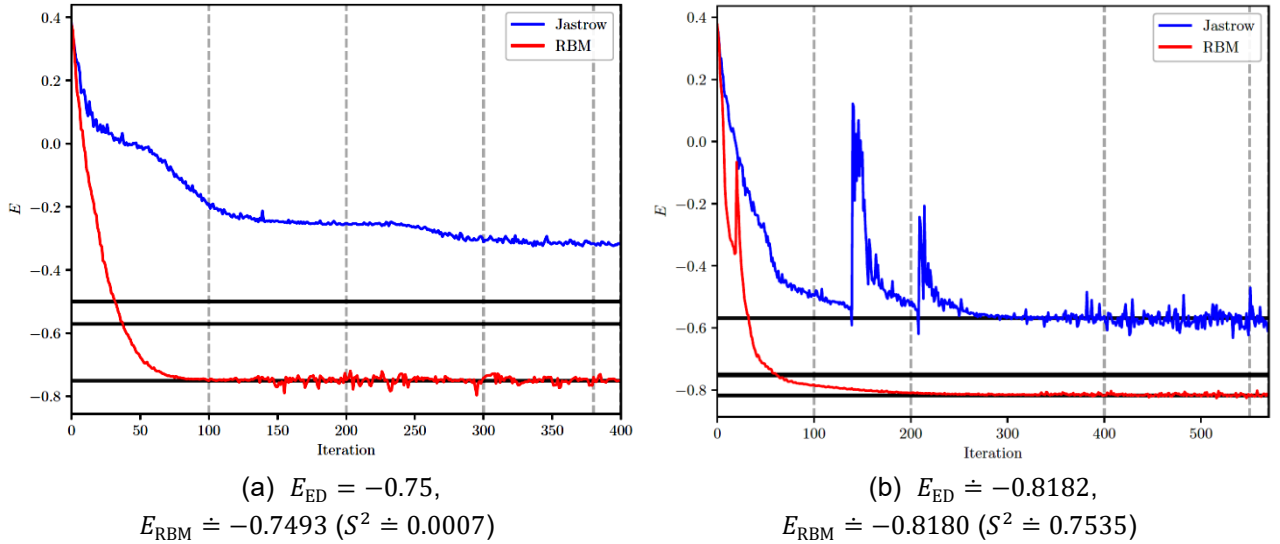


Figure 2 The learning curve of the RBM approximation of ground state energy for $N = 3$ dots, $U = 2$ and $\Gamma = 1$ at (a) $\zeta = 0.1$ (i.e. the singlet region) (b) $\zeta = 0.45$ (i.e. the doublet region)

The energy levels according to ED are shown as horizontal grey lines. Lastly, we provide the learning curve of the Jastrow model [13] initialized with the same parameters as RBM to emphasize its relative fast learning.

Figure 2 shows how RBM is able to approximate the wavefunction and not only give a precise energy of the ground state but also its correct spin S (compare with **Figure 1**), as in **Figure 2(a)** the total spin squared is $S^2 = 0$, corresponding to $S = 0$ and in **Figure 2(b)** the state has $S^2 = 3/4$, corresponding to $S = 1/2$, according to $S^2 = S(S + 1)$.

4. CONCLUSION

We used the recently developed GAL model to study a system of quantum dots coupled to a superconducting lead. We showed that GAL is capable of capturing the subgap spectrum consisting of Andreev bound states as well as the quantum phase transitions for an $N = 3$ quantum dot chain. We observed that in the regime $U = 2, \Gamma = 1$, the bound state of this system consisted of singlet and doublet states and showed two quantum phase transitions at $\zeta \approx 0.35$ and $\zeta \approx 0.55$. Having acquired the results through ED and using them as benchmark

value, we were able to demonstrate that RBM far outperforms the Jastrow model both qualitatively and quantitatively, is indeed able to approximate the complex ground state wavefunction of such correlated systems and can accurately find its energy and spin.

ACKNOWLEDGEMENTS

Financial supports from the Czech Academy of Science are gratefully acknowledged, in particular the Praemium Academiae awarded to Mgr. Martin Friák, Ph.D., and the Strategy AV21 program "AI: Artificial Intelligence for Science and Society".

REFERENCES

- [1] DE FRANCESCHI, KOUWENHOVEN, S.L., SCHÖNENBERGER, C., WERNSDORFER, W.. *Nat. Nanotechnol.* 2010, vol. 5, p. 703.
- [2] BENITO, Mónica, BURKARD, Guido. Hybrid superconductor-semiconductor systems for quantum technology. *Applied Physics Letters*. 2020, vol. 116, p.19.
- [3] LI, Chao, et al. Individual Assembly of Radical Molecules on Superconductors: Demonstrating Quantum Spin Behavior and Bistable Charge Rearrangement. *ACS nano*, 2025.
- [4] MEDVIDOVIĆ, Matija, MORENO, Javier Robledo. Neural-network quantum states for many-body physics. *The European Physical Journal Plus.*, 2024, vol. 139, no. 7, pp. 1-26.
- [5] HALDANE, F.D.M. Theory of the atomic limit of the Anderson model. I. Perturbation expansions re-examined. *Journal of Physics C: Solid State Physics*. 1978, vol. 11, no. 24, p. 5015.
- [6] ŽONDA, Martin, et al. Generalized atomic limit of a double quantum dot coupled to superconducting leads. *Physical Review B*. 2023, vol. 107, no. 11, p. 115407.
- [7] BECCA, Federico, SORELLA, Sandro. *Quantum Monte Carlo approaches for correlated systems*. Cambridge University Press, 2017.
- [8] CARLEO, Giuseppe, TROYER, Matthias. Solving the quantum many-body problem with artificial neural networks. *Science*. 2017, vol. 355, no. 6325, pp. 602-606.
- [9] MEDEN, V. The Anderson–Josephson quantum dot—a theory perspective. *Journal of Physics: Condensed Matter*. 2019, vol. 31, no. 16, p. 163001.
- [10] NADJ-PERGE, Stevan et al. Observation of Majorana fermions in ferromagnetic atomic chains on a superconductor. *Science*. 2014, vol. 346, no. 6209, pp. 602-607.
- [11] BIBORSKI, A., NOWAK, M.P., ZEGRODNIK, M. Correlation-induced d-wave pairing in a quantum dot square lattice. *Physical Review B*. 2021, vol. 104, no. 24, pp. 245430.
- [12] ZAZUNOV, A. et al. Andreev level qubit. *Physical Review Letters*, 2003, vol. 90, no. 8, p. 087003.
- [13] JASTROW, Robert. Many-body problem with strong forces. *Physical Review*. 1955, vol. 98, no. 5, p. 1479.