Artificial neural networks for density-functional optimizations in fermionic systems

Caio A. Custódio¹, Érica R. Filletti¹, and Vivian V. França¹*

¹Institute of Chemistry, Sãpão Paulo State University, UNESP, 14800-090, Araraquara, São Paulo, Brazil
¹vivian.franca@unesp.br

ABSTRACT

In this work we propose an artificial neural network functional to the ground-state energy of fermionic interacting particles in homogeneous chains described by the Hubbard model. Our neural network functional was proven to have an excellent performance: it deviates from numerically exact calculations by less than 0.15% for all the regimes of filling factors and magnetizations, and for a vast regime of interactions. When compared to analytical functionals, the neural functional found to be more precise for all the regimes of parameters, being particularly superior at the weakly interacting regime: where the analytical parametrization fails the most, ~7%, against only ~0.1% for the neural functional. We have also applied our homogeneous functional to finite, localized impurities and harmonically confined systems within density-functional theory (DFT) methods. The results show that while our artificial neural network approach is substantially more accurate than other equivalently simple and fast DFT treatments, it has similar performance than more costly DFT calculations and other independent many-body calculations, at a fraction of the computational cost.

Introduction

Theoretical approaches in the quantum mechanical level are essential to the full understanding of the matter and its properties. Although the wavefunction of a given system — and thus any property of that system — may in principle be obtained by solving the Schrödinger equations, this exact method is extremely costly for many-particle interacting systems, despite the availability of high computational resources. This difficulty is enhanced further when the theoretical simulations incorporate also spatial inhomogeneities — such as borders, impurities, confining potentials and disorder — which are sometimes desired but in most of the cases unavoidable in realistic systems and experimental setups.

In this context the density-functional theory (DFT)²–⁴ appears as an interesting and powerful tool. Within DFT the particle density of a system with N interacting particles is used as the central quantity, instead of the wavefunction. Thus the task of obtaining \( n(r_1, r_2, \ldots, r_N) \) — a 3N-dimensional function — is reduced to obtain \( n(r) \) — a 3-dimensional function. This illustrates how much simpler DFT calculations are in comparison to wavefunction-based quantum calculations. The one-to-one mapping between \( n \) and \( \Psi \) is guaranteed by the Hohenberg-Kohn (HK) theorem², which has been recently formally proved to be valid also for lattice systems⁵,⁶. As a consequence of the HK theorem all the observables are a density functional and therefore may in principle be obtained via DFT calculations. The theorem does not provide however any hint of how to obtain both, density of the system and density functional of any desired property.

The Kohn-Sham (KS) scheme⁷ is a clever iterative mapping commonly used to obtain the density of an interacting system. It maps the many-body interacting system into a fictitious non-interacting one via the so-called KS potential \( v_{KS} \). Here \( v_{KS} \) is constructed to reproduce the density \( n(r) \) of the interacting system. Since \( v_{KS} \) depends on the exchange-correlation energy and the latter is a density functional typically unknown exactly, in practical calculations one needs to make use of approximations to obtain the particle density. Hence, the performance of the DFT results depends on the approximations and functionals used.

Therefore there is a very active community in DFT dedicated to the development and the optimization of density functionals for practical DFT calculations. Less exploited tools in this research line are however the artificial neural networks (ANN)⁸. ANN are computational models that refer to the human brain functioning and are capable of calculating non-linear mathematical functions. It resembles the biological brain: it is able to be trained with a dataset and afterwards generalize the learning to other data of the same problem, being a type of machine learning used to estimate parameters and to recognize patterns.

Among the advantages of ANN one can mention the precision of the results, the easy implementation, the fast calculations and the ability of generalizing from examples⁸. Once an ANN has been trained and validated satisfactorily, it is able to predict new outputs from different input data (in the same domain, but different from those used in the training). ANN are used in many areas including neurocomputation, chemical engineering, industrial applications, medicine, chemistry and physics¹⁰–¹⁷.

Here we apply ANN’s concepts to estimate the ground-state energy of fermionic interacting particles in homogeneous chains as described by the one-dimensional Hubbard model. We designed an ANN functional which is able to recover the numerically
We consider one-dimensional homogeneous chains described by the fermionic Hubbard model

\[ \hat{H} = -t \sum_{<i,j>\sigma} (\hat{c}_i^\dagger \hat{c}_j + \text{h.c.}) + U \sum_i \hat{n}_i \hat{n}_i, \]

where \( U \) is the on-site interaction and \( t \) is the nearest-neighbor hopping term, while \( \hat{c}_i^\dagger \) (\( \hat{c}_i \)) is the creation (annihilation) operator of fermions at site \( i \) with spin \( \sigma = \uparrow, \downarrow \), and \( \hat{n}_i = \hat{c}_i^\dagger \hat{c}_i \) the number operator. The average charge density or filling factor is given by \( n = N/L = \langle \hat{n}_i \rangle + \langle \hat{n}_i \rangle \), while the magnetization or spin density is \( m = \langle \hat{n}_i \rangle - \langle \hat{n}_i \rangle \), where \( N = N_\uparrow + N_\downarrow \) is the total number of particles and \( L \) the chain size.

This is the simplest model to describe itinerant and interacting electrons in a chain, such as in solids and in nanostructures, or cold atoms in an optical lattice. Despite its simplicity the Hubbard model describes important phenomena, such as the Mott transition from metal to insulator (for \( U > 0 \), at \( n = 1 \)) \( ^{30} \), the transition from Bardeen-Cooper-Schrieffer superfluid to Bose-Einstein condensate (for \( U < 0 \), at \( m = 0 \)) \( ^{30} \) and exotic superfluidity (for \( U < 0 \) and \( m \neq 0 \)) \( ^{28,31,32} \).

Nevertheless the exact analytical solution of the model is in general unknown \( ^{-} \) except for some specific parameters \( ^{33} \) – thus one has to perform numerical calculations. For spatially homogeneous chains, \( L = \infty \), one can obtain an exact fully numerical (FN) solution by solving the Lieb-Wu integrals \( ^{34} \). For finite chains, using exact diagonalization we are limited to small chains, \( L \lesssim 15 \), while with density-matrix renormalization group (DMRG) \( ^{35} \), which is almost exact, one can solve chains of \( L \lesssim 200 \) sites. DMRG are however computationally expensive calculations: a single calculation might take several hours in a high-performance computing cluster.

Thus DFT appears as a powerful tool to solve larger chains at a lower computational cost: a typical DFT calculation for the Hubbard model takes no longer than 1 minute. DFT may also be applied to solve inhomogeneous chains, i.e., in the presence not only of boundaries, but also impurities, interfaces, superlattices, harmonic confinement and disorder. This is possible thanks to \( ^{-} \) local or non-local \( ^{-} \) approximations to the exchange-correlation functional, which require, as an input, a functional for the per-site ground-state energy of the homogeneous system.

In particular, the local-spin-density approximation (LSDA) for the ground-state energy is given by

\[ e_0^{\text{inh}} \approx e_0^{\text{LSDA}} = \frac{1}{L} \sum_{i=1}^{L} e_0^{\text{hom}}[n,m,U] \bigg|_{n=m_i} \]

where \( e_0^{\text{hom}} \) is the homogeneous functional for the ground-state energy density and \( (n_i,m_i) \) are local densities at site \( i \) of the inhomogeneous system. Thus Eq.\( (2) \) approximates the energy of inhomogeneous chains by evaluating the energy of the homogeneous system with the site-by-site replacement \( n \rightarrow n_i \) and \( m \rightarrow m_i \). In practical DFT, using the self-consistent KS scheme \( ^{7} \), it implies the evaluation of partial derivatives of the energy with respect to \( n_i \) and \( m_i \) (to obtain \( v_{KS} \)) in each of the iterations of the self-consistency cycle. Therefore, in order to solve inhomogeneous Hubbard chains within DFT-LSDA, the first step is to obtain a simple and reliable functional \( e_0^{\text{hom}}[n,m,U] \) for the homogeneous limit.

Although the fully numerical (FN) solutions of the Lieb-Wu integrals \( ^{34} \) (exact only for infinite chains) are certainly a trusty functional for the DFT-LSDA treatment of inhomogeneous systems, they are computationally expensive. Besides, the integrals are not direct functions of \( (n,m) \) and hence we do not have total freedom to choose specific values to be used site by site in each iteration of the KS cycles. Therefore, for each interaction \( U \), one has to produce a huge amount of FN data to cover all the regimes of \( n_i \) and \( m_i \) in order to obtain the \( v_{KS} \) numerical derivatives with high precision.

So one possibility to overcome this numerical difficulty is then to adopt a reliable analytical expression for the homogeneous functional. Actually it has been done: the analytical FVC functional \( ^{33} \), which extended previous expressions \( ^{36} \) by incorporating...
the magnetization, has been currently used as a practical tool for DFT-LSDA calculations. The FVC parametrization properly recovers all the analytical known limiting cases (for $U = 0$, $U = \infty$ and for the combined set $n = 1$ and $m = 0$), correctly predicts a positive Mott gap at $n = 1$ for any $U > 0$, and is also more precise in general than the previous expressions. While for most of the parameters the FVC functional presents relatively low deviations from numerically exact results (~2%), it shows however inopportune larger deviations for the weakly interacting regime. Thus, the ANN functional proposed here is a simpler and faster numerical approach to inhomogeneous systems within DFT calculations, with superior accuracy than FVC and computationally cheaper than FN and DMRG.

**Results and discussion**

The training of our neural model was performed with the Levenberg-Marquardt algorithm. The process involves the random split of the exact data in three sets: 70% for the training itself, 15% for the validation and the remaining 15% for testing the ANN performance. The input parameters of our ANN model are the variables density $n$, magnetization $m$ and interaction $U$, while the output is the per-site ground-state energy $e_{0}^{\text{ANN}}(n,m,U)$. We consider 20,891 numerically exact Lieb-Wu results within the regimes $0 \leq U \leq 10t$, $0 \leq n \leq 1$ and $0 \leq m \leq n$. Thus the possible topologies of our ANN are of type: 3 input neurons $\{n,m,U\}$ — a certain number of hidden layers — a single output neuron $\{e_{0}^{\text{ANN}}(n,m,U)\}$. Several topologies were studied, e.g., with a single hidden layer — 3-5-1, 3-10-1 and 3-20-1 — and with two hidden layers, 3-12-12-1 and 3-10-10-1.

Our target performance was below $10^{-6}$ for the mean squared error, which was obtained after extensive training by the topology 3-20-1 with 1000 epochs (we choose to stop the training in 1000 epochs to avoid over-fitting by the network, which could lead to a poorer generalization performance), as shown in Figure 1. This architecture contains one hidden layer composed of twenty neurons, with tangent sigmoid transfer function, while the output layer has only one node with a linear transfer function. The trained neural network can be found online free of charge as Supplementary Material (ANN_to_Hubbard.mat and Read_me.pdf files).

Now we compare the performance of our ANN model with the FVC parametrization, for all the regimes of parameters. We start by considering non-magnetic systems. In Figure 2 we show the energy as a function of the density for a strongly correlated case, with $U = 6t$. While the overall behavior is well described by both, ANN and FVC, we find that the ANN deviations are below ~0.1% (see inset), smaller than the FVC ones, which in some cases are 10 times bigger. This higher performance of ANN over FVC for strongly interacting systems is surprising, since at this regime FVC is recognized as sufficiently precise.

**Figure 1.** Structure of our ANN model, 3-20-1, and its performance during training.
Figure 2. Per-site ground-state energy as a function of the density for the FVC and ANN functionals in comparison to the numerically exact FN result. The inset shows the percentage deviations, obtained via $100 \times |(e_0^{\text{approx}} - e_0^{\text{exact}})/e_0^{\text{exact}}|$. In all cases the interaction is $U = 6t$, the magnetization is $m \sim 0$ and $D$ is the average percentage deviation.

Figure 3. Per-site ground-state energy as a function of the on-site interaction $U$ (in units of $t$), for the FVC and ANN functionals in comparison to the numerically exact FN result. Here $n \sim 0.5$ and magnetization $m \sim 0$. The inset shows the percentage deviations, obtained via $100 \times |(e_0^{\text{approx}} - e_0^{\text{exact}})/e_0^{\text{exact}}|$, while $D$ is the average percentage deviation.

Figure 3 presents the energy as a function of the interaction for non-magnetic chains. It reveals that the excellent performance of our ANN model is not restricted to the strongly correlated systems: within $0 \leq U \leq 10t$ and $m \sim 0$ the ANN deviates by less than $\sim 0.08\%$ for all the interactions, while FVC deviates on average $\sim 0.5\%$. 
We also compare the performance of ANN and FVC for magnetic systems. Figure 4 presents the energy as a function of the magnetization for several interactions. We find that while FVC in some cases fails to recover even the qualitative behavior of the energy, our ANN model properly describes the exact trend. Quantitatively, while FVC deviations reach up to $\sim 7\%$ (for $U = 0.5t$), the ANN model remains reliable, with deviations inferior to $0.15\%$.

![Figure 4](image-url)

**Figure 4.** Per-site ground-state energy as a function of the magnetization $m$, for the FVC and ANN functionals in comparison to the numerically exact FN result. Here $n \sim 0.5$ and $D$ is the average percentage deviation.

Finally, we illustrate the use of our ANN functional within DFT-LSDA calculations by applying it to inhomogeneous i) unconfined finite chains, ii) localized impurity systems ($V_i = V \delta_{i,i_0}$) and iii) harmonically confined chains ($V_i = k(i - i_0)^2$). The later inhomogeneous potential is particularly interesting to simulate state-of-the-art experiments with cold fermionic atoms in optical lattices. We compare the performance of DFT-LSDA $-$ with each functional ANN, FVC and FN $-$ to DMRG calculations.

Figure 5 summarizes the essential features of our analysis to the several inhomogeneous test cases. For finite unconfined chains, Fig. 5a, the typical Friedel-like oscillations $^{42}$ on the local densities are in general well reproduced by all the three functionals. The average percentage deviations in this case are similar, slightly larger for FVC, because the chain is relatively large, $L = 40$. For smaller chains and/or with inhomogeneous external potentials $V_i$, as in Fig. 5b and Fig. 5c, we see that the FVC is clearly the less accurate functional. Fig. 5d, for the harmonically confined case, shows that the deviations on the ground-state energy follow similar trend: although the deviation oscillates with the harmonic strength $k$ for the approaches, the FVC deviations are systematically larger than the FN and ANN ones. Thus, our results demonstrate that the ANN approach has a very good balance between precision and computational cost: it is faster than FN and DMRG and significantly more precise than FVC.

**Conclusions**

In summary, we have developed a simple, fast and reliable artificial neural network functional for the ground-state energy of the homogeneous fermionic Hubbard chains. We have exemplified the powerfulness of the ANN functional into DFT methods by exploring finite, localized impurities and harmonically confined systems. The results show that our ANN approach not only properly captures the qualitative behavior of energy and density profiles for all the inhomogeneous systems, but also leads to more accurate results than previous parameterizations. This excellent performance is similar to other numerical solutions (FN and DMRG), but at a fraction of the computational cost. Therefore our ANN functional is a practical tool to be implemented in DFT calculations, such as in local and non-local density approximations, for inhomogeneous Hubbard chains. Our approach is particularly interesting to investigate disordered systems, where many realizations are needed for each randomly localized impurities scenario, what makes analyses via exact methods computationally prohibitive.
Figure 5. Density profile for (a) a finite chain of size $L = 40$ and $n = 0.8$, (b) a localized impurity system, with $L = 15$, impurity strength $V = 2t$ and $n = 0.4$, and (c) a harmonically confined chain with $L = 40$, harmonic curvature $k = 0.01$ and $n = 0.8$, while (d) shows the per-site ground-state energy as a function of the harmonic confinement curvature for chains of size $L = 40$ and $n = 0.8$. In all cases we adopted open boundary conditions, $U = 1t$ and $m = 0$. The results are either exact DMRG or approximated DFT-LSDA, with one of the functionals: FN, FVC or ANN. The percentage deviations are obtained via $100 \times \left| \frac{e_{0}^{\text{approx}} - e_{0}^{\text{exact}}}{e_{0}^{\text{exact}}} \right|$, and $D$ is the average percentage deviation.

Data Availability

The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

References

1. Bruenner, T., Runge, E., Buchleitner, A. & França, V. V. Entanglement enhancement in spatially inhomogeneous many-body systems. Phys. Rev. A 87, 3231 (2013).
2. Hohenberg, P. & Kohn, W. Inhomogeneous Electron Gas. Phys. Rev. 136, B864-B871 (1964).
3. Kohn, W. Nobel Lecture: Electronic structure of matter - wave functions and density functionals. Rev. Mod. Phys. 71, 1253-1266 (1999).
4. Capelle, K. A Bird’s-Eye View of Density-Functional Theory. *Braz. J. Phys.* **36**, 1318-1343 (2006).
5. Coe, J. P., D’Amico, I. & França, V. V. Uniqueness of density-to-potential mapping for fermionic lattice systems. *EPL Europhysics Lett.* **110**, 63001 (2015).
6. França, V. V., Coe, J. P., & D’Amico, I. Testing density-functional approximations on a lattice and the applicability of the related Hohenberg-Kohn-like theorem. *Sci. Rep.* **8**, 664 (2018).
7. Kohn, W. & Sham, L. J. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, A1133-A1138 (1965).
8. Bishop, C. M. *Neural networks for pattern recognition* (Clarendon Press, 1995).
9. Nafey, A. S. Neural network based correlation for critical heat flux in steam-water flows in pipes. *Int. J. Therm. Sci.* **48**, 2264-2270 (2009).
10. Ortega-Zamorano, F., Jerez, J. M., Juárez, G. E. & Franco, L. FPGA Implementation of Neurocomputational Models: Comparison Between Standard Back-Propagation and C-Mantec Constructive Algorithm. *Neural Process. Lett.* **46**, 1-16 (2017).
11. Filletti, E. R., Silva, J. M. & Ferreira, V. G. Predicting of the fibrous filters efficiency for the removal particles from gas stream by artificial neural network. *Adv. Chem. Eng. Sci.* **05**, 317-327 (2015).
12. Ramil, A., Lopez, A., Pozo-Antonio, J. & Rivas, T. A computer vision system for identification of granite-forming minerals based on rgb data and artificial neural networks. *Measurement* **117**, 90-95 (2018).
13. Walczak, S. & Velanovich, V. Improving prognosis and reducing decision regret for pancreatic cancer treatment using artificial neural networks. *Decis. Support. Syst.* **106**, 110-118 (2018).
14. Balabin, R. M. & Lomakina, E. I. Neural network approach to quantum-chemistry data: Accurate prediction of density functional theory energies. *J. Chem. Phys.* **131** (2009).
15. Brockherde, F. et al. Bypassing the Kohn-Sham equations with machine learning. *Nat. Commun.* **8**, 872 (2017).
16. McGibbon, R. T. & Pande, V. S. Learning Kinetic Distance Metrics for Markov State Models of Protein Conformational Dynamics. *J. Chem. Theory Comput.* **9**, 2900-2906 (2013).
17. Fletcher, T. L., Davie, S. J. & Popelier, P. L. A. Prediction of Intramolecular Polarization of Aromatic Amino Acids Using Kriging Machine Learning. *J. Chem. Theory Comput.* **10**, 3708-3719 (2014).
18. Zhang, L.-L., Huang, J., Duan, C.-B. & Wang, W.-Z. Incommensurate charge density wave modulated by period of Hubbard superlattices. *Mod. Phys. Lett. B* **29**, 1550208 (2015).
19. Wei-Zhong, Z. L.-L., Jin, H., Cheng-Bo, D. & Wang, W.-Z. Structure-dependent metal?insulator transition in one-dimensional Hubbard superlattices. *Chin. Phys. B* **24**, 77101 (2015).
20. Vetchchinkina, V., Kartsev, A., Karlsson, D., & Verdozzi, C. Interacting fermions in one-dimensional disordered lattices: Exploring localization and transport properties with lattice density-functional theories. *Phys. Rev. B* **87**, 115117 (2013).
21. França, V. V. & D’Amico, I. Entanglement from density measurements: Analytical density functional for the entanglement of strongly correlated fermions. *Phys. Rev. A* **83**, 42311 (2011).
22. Hubbard, J. Electron Correlations in Narrow Energy Bands. *Source: Proc. Royal Soc. London. Ser. A Math. Phys. Sci.* **276**, 238-257 (1963).
23. Coe, J. P., França, V. V. & D’Amico, I. Hubbard model as an approximation to the entanglement in nanostructures. *Phys. Rev. A* **81**, 052321 (2010).
24. Coe, J. P., França, V. V. & D’Amico, I. Approximation of the entanglement in quantum dot chains using Hubbard models. *J. Physics: Conf. Ser.* **286**, 12048 (2011).
25. Coe, J. P., França, V. V. & D’Amico, Feasibility of approximating spatial and local entanglement in long-range interacting systems using the extended Hubbard model. *EPL 93*, 10001 (2011).
26. Johnson, T. et al. Hubbard Model for Atomic Impurities Bound by the Vortex Lattice of a Rotating Bose-Einstein Condensate. *Phys. Rev. Lett.* **116**, 240402 (2016).
27. França, V. V. & Capelle, K. Entanglement in spatially inhomogeneous many-fermion systems. *Phys. Rev. Lett.* **100**, 1-4 (2008).
28. França, V. V., Hörndlein, D. & Buchleitner, A. Fulde-Ferrell-Larkin-Ovchinnikov critical polarization in one-dimensional fermionic optical lattices. *Phys. Rev. A* **86**, 033622 (2012).
29. Capelle, K. & Campo, V. L. Density functionals and model Hamiltonians: Pillars of many-particle physics. Phys. Reports 528, 91-159 (2013).
30. França, V. V. & Capelle, K. Entanglement of strongly interacting low-dimensional fermions in metallic, superfluid, and antiferromagnetic insulating systems. Phys. Rev. A 74, 42325 (2006).
31. França, V. V. Entanglement and exotic superfluidity in spin-imbalanced lattices. Phys. A: Stat. Mech. its Appl. 475, 82-87 (2017).
32. De Picoli, T., D’Amico, I. & França, V. V. Metric-space approach for distinguishing quantum phase transitions in spin-imbalanced systems. Braz. J. Phys. 48, 472 (2018).
33. França, V. V., Vieira, D. & Capelle, K. Simple parameterization for the ground-state energy of the infinite Hubbard chain incorporating Mott physics, spin-dependent phenomena and spatial inhomogeneity. New J. Phys. 14 (2012).
34. Lieb, E. H. & Wu, F. Y. Absence of Mott Transition in an Exact Solution of the Short-Range, One-Band Model in One Dimension. Phys. Rev. Lett. 20, 1445-1448 (1968).
35. Schollwöck, U. The density-matrix renormalization group. Rev. Mod. Phys. 77, 259-315 (2005).
36. Lima, N. A., Silva, M. E., Oliveira, L. N. & Capelle, K. Density functionals not based on the electron gas: Local-density approximation for a luttinger liquid. Phys. Rev. Lett. 90, 146402 (2003).
37. Xianlong, G., Chen, A.-H., Tokatly, I. V., & Kurth, S. Lattice density functional theory at finite temperature with strongly density-dependent exchange-correlation potentials. Phys. Rev. B 86, 235139 (2012).
38. Mori-Sanchez, P., & Cohen, A. J. The derivative discontinuity of the exchange?correlation functional. Phys. Chem. Chem. Phys. 16, 14378 (2014).
39. Mitxelena, I., Piris, M., & Rodriguez-Mayorga, M. On the performance of natural orbital functional approximations in the Hubbard model. J. Phys.: Condens. Matter 29, 425602 (2017).
40. Ulrich, C. A. Density-functional theory for systems with noncollinear spin: Orbital-dependent exchange-correlation functionals and their application to the Hubbard dimer, Phys. Rev. B 98, 035140 (2018).
41. Hagan, M. T. & Menhaj, M. B. Training Feedforward Networks with the Marquardt Algorithm. IEEE Transactions on Neural Networks 5, 989-993 (1994).
42. Vieira, D., Freire, H. J. P., Campo Jr., V. L. & Capelle, K. Friedel oscillations in one-dimensional metals: From Luttinger’s theorem to the Luttinger liquid. J. Magn. Magn. Mater. 320, E421-E424 (2008).

Acknowledgements
VVF acknowledges FAPESP (Grant: 2013/15982-3), VVF and ERF acknowledge CNPq (Grant: 448220/2014-8). This research was supported by resources supplied by the Center for Scientific Computing (NCC/GridUNESP) of the São Paulo State University (UNESP).

Contributions
All authors conceived the idea, C.A.C. performed the calculations. All authors discussed the results and contributed to the writing of the manuscript.

Additional information
Competing financial interests
The authors declare that they have no competing interests.