Multibasis Encodings in Recurrent Neural Network Wave Functions for Variational Optimization

Wirawat Kokaew
Perimeter Institute for Theoretical Physics
Department of Physics and Astronomy
University of Waterloo
Ontario, N2L 3G1, Canada
wkokaew@perimeterinstitute.ca

Mohamed Hibat-Allah
Perimeter Institute for Theoretical Physics
Department of Physics and Astronomy
University of Waterloo
Ontario, N2L 3G1, Canada
mohamed.hibat.allah@uwaterloo.ca

Lauren E. Hayward
Perimeter Institute for Theoretical Physics
Department of Physics and Astronomy
University of Waterloo
Ontario, N2L 3G1, Canada
lhayward@perimeterinstitute.ca

Juan Carrasquilla
Vector Institute
MaRS Centre, Toronto, Ontario, M5G 1M1, Canada
Department of Physics and Astronomy
University of Waterloo
Ontario N2L 3G1, Canada
carrasqu@vectorinstitute.ai

Abstract

Solving optimization problems via neural networks has proven to be a promising approach in yielding better solutions. However, the full potential of parameterized models has yet to be fully explored. Motivated by the success of variational quantum optimization with multibasis encodings, we propose a quantum-inspired machine learning algorithm that integrates both approaches to reduce the system size as well as parameters in a neural network ansatz. We demonstrate the performance of the proposed algorithm by solving Ising chain systems, resulting in faster convergence towards the ground state energy. This study holds the potential for widespread applications across various fields that require efficient optimization for large-scale problems.

1 Introduction

Solving optimization problems plays a crucial role as a stepping stone in many practical fields, such as artificial intelligence, bioinformatics, and cloud computing [1]. Unfortunately, most of these problems are hard to solve in an acceptable time. Over the past few decades, numerous heuristic techniques have been proposed to find approximate solutions to these problems. One popular method is Simulated Annealing (SA), which can avoid the slow sampling of the Monte Carlo approach [2, 3]. By substituting the conventional simulated annealing with the annealing of a parameterized model, the Variational Neural Annealing (VNA) approach was introduced [4], and they have shown that VNA offers a powerful alternative to solve optimization problems.

Machine Learning and the Physical Sciences Workshop, NeurIPS 2023.

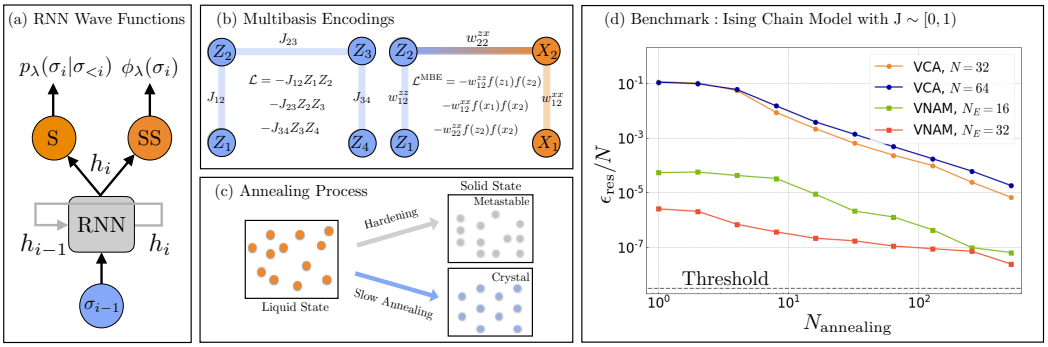

Figure 1: (a) An illustration of RNN Wave functions. At each step, the input σ_{i-1} (blue) is fed into the Recurrent Neural Network cell (RNN) along with h_{i-1} to generate the next hidden state h_i. The hidden state h_i is used to generate conditional probabilities, $p_\lambda(\sigma_i|\sigma_{<i})$, via the Softmax activation (S) and individual phases, $\phi_\lambda(\sigma_i)$, via the SoftSign activation (SS). The procedure is then iterated until σ_N has been obtained. (b) A comparison between traditional and multibasis encodings. For fully connected a 4-node problem, traditional encodings use only Z Pauli operators (blue) and optimize via a cost function $\mathcal{L}$, but multibasis encodings exploit both Z and X Pauli operators (blue and orange, respectively) and optimize via a cost function $\mathcal{L}^{\rm MBE}$ (see definition in text). (c) The annealing process can be adapted to improve the performance of the training algorithm. The global minimum (Crystal state) is expected if the decrease of pseudo temperature is slow enough compared to fast change, leading to getting stuck at a local minimum (Metastable state). (d) Benchmark of our algorithms for solving the Ising chain model with positive random coupling $[0, 1)$. The traditional encodings (VCA) for problem size $N = 32$ and 64 can be encoded with $N_E = 16$ and 32 for multibasis encodings (VNAM), respectively. Residual energy per spin ($\epsilon_{\rm res}/N$) obtained by VNAM reaches 10^{-8} within only 10^3 annealing steps ($N_{\rm annealing}$).

In addition to classical algorithms, there is also significant interest in developing variational quantum algorithms to solve NP-hard optimization problems, such as the Variational Quantum Eigensolver [5, 6] and the Quantum Approximate Optimization Algorithm [7, 8]. However, many quantum algorithms are limited by their high requisite circuit depth and non-convexity. To address these limitations, variational quantum optimization with MultiBasis Encodings (MBE) was introduced in [9] successfully optimizing of the MaxCut problem while requiring fewer quantum resources.

In this work, we propose a new quantum-inspired machine learning algorithm as an alternative approach to solving large-scale combinatorial optimization problems. The main contribution of our work is combining two ideas from a different field: Recurrent Neural Networks (RNN) wave functions from machine learning algorithms and the multibasis encoding techniques from a quantum algorithm as shown in Fig. 1(a) and (b), respectively. For more details please refer to [10] and [9]. Specifically, we substitute the positive RNN in traditional VNA with a complex RNN. Additionally, the algorithm is enhanced using fewer parameters through multibasis encoding, enabling the computation of more expectation values from a smaller RNN ansatz. In addition to the multibasis encoding feature, we enhanced the algorithms by annealing techniques as shown in Fig. 1(c), which can be used to avoid local minimum [4]. This combined algorithm offers a new way to solve large-scale problems efficiently in a wide array of fields [11, 12, 13]

2 Multibasis Encodings in Recurrent Neural Network Wave Functions

Multibasis encodings (MBE) were proposed as a quantum optimization algorithm aimed at reducing susceptibility to local minima in the training landscape and it has demonstrated the ability to solve MaxCut optimization problems for various graphs [9]. The key ingredient of MBE is encoding variables into two (or more) qubit axes to modify the loss function and evaluate the product of expectation values of each single-qubit term. As a result, MBE halves the number of qubits required for a given optimization problem. To construct the MBE loss function, the Hamiltonian formed by the product of the operators would not be suitable since the quantum ground state it encodes does not correspond to the true ground state. Instead, we can consider the products of single-qubit

measurements $\langle X_i \rangle$ and $\langle Z_i \rangle$, where X_i and Z_i are Pauli operators for qubit i, in which $\langle X_i \rangle$ and $\langle Z_i \rangle$ operators are simultaneously optimized. This approach yields the MBE loss function denoted as

$$\mathcal{L}^{\mathrm{MBE}} = \sum_{j<i}^{n/2} w_{ij}^{zz} f(z_i) f(z_j) + \sum_{j<i}^{n/2} w_{ij}^{xx} f(x_i) f(x_j) + \sum_{j,i}^{n/2} w_{ij}^{zx} f(z_i) f(x_j), \tag{1}$$

where w_{ij}^{ab} is the coupling between site i and j with Pauli operators a and b, respectively. We define x_i and z_i as the shorthand of single-qubit measurement of X_i and Z_i for qubit i, and $f(x)$ is an activation function, which plays a role as a nonlinear function to enhance the expressibility of the algorithm. It is used to introduce non-linearity and capture more complex relationships between variables. This procedure is graphically depicted in Fig. 1(b) for the four-qubit encoding.

Motivated by the success of MBE, we leverage it to enhance the utility of RNN wave functions. Our approach involves encoding multiple variables into a single qubit and performing simultaneous optimization through variational neural optimization. Instead of using variational neural annealing to find the ground state from minimizing the expectation value of the target Hamiltonian, we substitute it with parameterized loss function with multibasis encodings $\mathcal{L}_\lambda^{\mathrm{MBE}}$ from Eq. (1) and the gradients can be computed by

$$\partial_\lambda \mathcal{L}_\lambda^{\mathrm{MBE}} = \sum_{j<i}^{n/2} w_{ij}^{zz} f(z_i) f'(z_j) \partial_\lambda z_j + z_i \leftrightarrow z_j + \dots \tag{2}$$

3 Incorporating classical annealing

The key components of the conventional annealing formulation include the Metropolis-Hasting algorithm, which enables the search for equilibrium by slowly decreasing the temperature to zero, mimicking the physical annealing process as shown in Fig. 1(c). The procedure terminates when the temperature reaches zero, at which point the solution is expected to be found. However, this formulation can only provide a quasi-distribution due to finite-time issues [14, 3] and may suffer from slow sampling when the optimization landscape is rough.

Fortunately, RNN wave functions have proven to be suitable candidates as variational wave functions for approximating ground state energies, correlation functions, and other properties [10]. To incorporate annealing for RNN ansatz, we first encode the problem in a target Hamiltonian $\hat{H}_{\mathrm{target}}$ and add an entropy term to form a new cost function:

$$F_\lambda(T) = \langle \hat{H}_{\mathrm{target}} \rangle_\lambda + T \sum_\sigma p_\lambda(\sigma) \log(p_\lambda(\sigma)), \tag{3}$$

where the second term is related to pseudo temperature, T, and probability of obtaining some configuration, $p(\sigma)$, that is analogous to thermal fluctuation. By the end of the annealing process ($T = 0$), we expect to obtain the ground state of the target Hamiltonian.

Here, we can exploit an annealing procedure similar to Eq. (3) to improve the performance of finding the ground state of the MBE loss function. Following the essence of the annealing process, we modify the pseudo-free energy loss function as

$$\mathcal{L}_\lambda^F(T) = \mathcal{L}_\lambda^{\mathrm{MBE}} + T \sum_\sigma p_\lambda(\sigma) \log(p_\lambda(\sigma)) \tag{4}$$

We note that using the activation function as the $f(x) = \tanh(x)$ in MBE as in the original text provides a cost function where local and global minima are close to each other, resulting in fluctuations during annealing that can lead to getting stuck in local minima. To address this issue, we introduce the new activation function $f(x) = \sin(\pi x/\sqrt{2})$. This activation function helps increase the difference between local and global values by extending the expectation values to be 1 in both axes. Additionally, the lowest value of the cost function with the sin activation corresponds to the global minima of the original target Hamiltonian.

After the training process, we sample millions of samples from the RNN wave function to evaluate the expectation value for each Pauli operator and each spin. We obtain the ground state configurations by performing a rounding procedure (interpreting positive expectation values as spin-up and negative values as spin-down). The ground state energy can be computed by substituting the rounded expectation values ($x_i, z_i \in \{-1, 1\}$) in Eq. (1).

4 Numerical experiment: Ising chain with positive random coupling

As benchmarking, we consider the simple Ising model with positive random coupling $J_{i,i+1}$ sampled from a uniform distribution in $[0, 1)$. The Hamiltonian can be written as

$$H_{\text{target}} = -\sum_{i}^{N-1} J_{i,i+1}\sigma_i\sigma_{i+1}. \tag{5}$$

Notice that the ground state configuration is given by either all spin up or down, and then the ground state can be calculated by summing all couplings, i.e., $E_G = -\sum J_{i,i+1}$. To quantify the performance of our algorithms, we use the residual energy as used in [4],

$$\epsilon_{\text{res}} = \left[\langle H_{\text{target}}\rangle_\lambda - E_G\right]_{\text{dis}}, \tag{6}$$

where $\langle \dots \rangle_\lambda$ is the arithmetic mean of samples from RNN wave function and $[\dots]_{\text{dis}} \equiv \exp(\ln(\langle \dots \rangle_\lambda))$ is the average of ϵ_{res} over different realizations and initialization.

In Fig. 1 (d), we investigate the Ising chain model for the different number of sites (N) with $N = 32$ and $N = 64$ which can be encoded with $N_E = 16$ and $N_E = 32$, respectively. We conduct 25 realizations of coupling. The RNN wave function is then utilized to optimize the system and determine the ground state energy and the threshold is set at $\epsilon_{\text{res}} = 10^{-7}$.

To benchmark our proposed algorithms, we compare our algorithms against the Variational Classical Annealing (VCA), as indicated by the yellow and blue lines for 32 and 64 sites, respectively, which has been shown to yield the best results in [4]. For $N = 32$ (or $N_{\text{E}} = 16$), within 10^3 steps of annealing, VCA achieves residual energy levels of approximately 10^{-5}, as shown in the orange line in Fig. 1 (d). The introduction of the multibasis encodings to the algorithm successfully offers fast convergence allowing us to find the ground state energy more effectively. Variational Neural Annealing with Multibasis encodings (VNAM) can reach 10^{-7} within 10^3 steps of annealing (green line). Interestingly, in the case of $N = 64$ (or $N_{\text{E}} = 32$), VCA algorithm provides similar trend as $N = 32$ (blue line), but VNAM (red line), and we observe rapid convergence to small residual energy (10^{-7}) within just 10 annealing steps, which is contrast to the trend observed in Ref. [4]. This outcome indicates that the combination of multibasis encodings and annealing significantly outperforms VCA when dealing with large-scale problems.

5 Conclusion and outlook

We have proposed a quantum-inspired machine learning algorithm that combines variational neural optimization, multibasis encodings, and the annealing method. By leveraging multibasis encodings, we are able to reduce the number of parameters in the parameterized model. Additionally, the annealing method ensures that the algorithm approaches the approximate ground state.

We have shown a comparison with the VCA approach for solving the Ising chain model with positive random coupling. The VNAM algorithms outperform traditional methods VCA, exhibiting rapid convergence, leading to lower residual energy levels. Moreover, the utilization of the sin activation function further contributes to superior performance by effectively suppressing fluctuations. These findings highlight the potential of sin activation. We note that the results can be further improved by a more careful choice of the hyperparameters and vertex partitioning.

Future research directions in this field could involve the exploration of these techniques in other advanced generative models besides the RNN, such as Transformer architecture. The other next research direction is the investigation of the influence of different multibasis encodings on the accuracy of the solutions. More precisely, there is potential to investigate the advantages of encoding more basis states in a single qubit using high dimensional orthogonal bases to reduce system size. This approach can fully maximize the utility of RNN wave functions in solving large-scale problems and have widespread applications in various fields, providing a promising route toward efficient solutions to real-world optimization problems in science, industry, and technology.

Broader Impact

Our work focuses on improving the algorithm for solving large-scale combinatorial optimization problems. Any combinatorial problems can be solved by our method. Specifically, the problem with problem size N can be turned to and solved by $N/2$ parameters in a neural network. This method has fruitful applications for real-world optimization problems, such as traveling salesman problems, finding the ground state of quantum many-body systems, etc.

References

[1] Wenjun Li, Yang Ding, Yongjie Yang, R. Simon Sherratt, Jong Hyuk Park, and Jin Wang. Parameterized algorithms of fundamental NP-hard problems: a survey. *Human-centric Computing and Information Sciences*, 10(1):29, December 2020.

[2] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, 1983.

[3] Michel Gendreau and Jean-Yves Potvin, editors. *Handbook of Metaheuristics*, volume 272 of *International Series in Operations Research & Management Science*. Springer International Publishing, Cham, 2019.

[4] Mohamed Hibat-Allah, Estelle M. Inack, Roeland Wiersema, Roger G. Melko, and Juan Carrasquilla. Variational Neural Annealing, January 2021. arXiv:2101.10154 [cond-mat, physics:quant-ph].

[5] Alberto Peruzzo, Jarrod McClean, Peter Shadbolt, Man-Hong Yung, Xiao-Qi Zhou, Peter J. Love, Alán Aspuru-Guzik, and Jeremy L. O'Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1):4213, July 2014.

[6] Juneseo Lee, Alicia B. Magann, Herschel A. Rabitz, and Christian Arenz. Progress toward favorable landscapes in quantum combinatorial optimization. *Physical Review A*, 104(3):032401, September 2021.

[7] Edward Farhi, Jeffrey Goldstone, and Sam Gutmann. A Quantum Approximate Optimization Algorithm, November 2014. arXiv:1411.4028 [quant-ph].

[8] Guido Pagano, Aniruddha Bapat, Patrick Becker, Katherine S. Collins, Arinjoy De, Paul W. Hess, Harvey B. Kaplan, Antonis Kyprianidis, Wen Lin Tan, Christopher Baldwin, Lucas T. Brady, Abhinav Deshpande, Fangli Liu, Stephen Jordan, Alexey V. Gorshkov, and Christopher Monroe. Quantum approximate optimization of the long-range Ising model with a trapped-ion quantum simulator. *Proceedings of the National Academy of Sciences*, 117(41):25396–25401, October 2020.

[9] Taylor L. Patti, Jean Kossaifi, Anima Anandkumar, and Susanne F. Yelin. Variational Quantum Optimization with Multi-Basis Encodings, January 2022. arXiv:2106.13304 [quant-ph].

[10] Mohamed Hibat-Allah, Martin Ganahl, Lauren E. Hayward, Roger G. Melko, and Juan Carrasquilla. Recurrent Neural Network Wave Functions. *Physical Review Research*, 2(2):023358, June 2020. arXiv:2002.02973 [cond-mat, physics:physics, physics:quant-ph].

[11] C Koulamas, S Antony, and R Jaen. A survey of simulated annealing applications to operations research problems. *Omega*, 22(1):41–56, January 1994.

[12] D.I. Svergun. Restoring Low Resolution Structure of Biological Macromolecules from Solution Scattering Using Simulated Annealing. *Biophysical Journal*, 76(6):2879–2886, June 1999.

[13] Bruce Hajek. A tutorial survey of theory and applications of simulated annealing. In *1985 24th IEEE Conference on Decision and Control*, pages 755–760, Fort Lauderdale, FL, USA, December 1985. IEEE.

[14] Debasis Mitra, Fabio Romeo, and Alberto Sangiovanni-Vincentelli. Convergence and finite-time behavior of simulated annealing. In *1985 24th IEEE Conference on Decision and Control*, pages 761–767, 1985.