QuYBE - An Algebraic Compiler for Quantum **Circuit Compression**

Sahil Gulania[†]

Mathematics and Computer Science Division Argonne National Laboratory Lemont, Illinois, 60439, United States sgulania@anl.gov

Bo Peng[†]

Physical and Computational Sciences Directorate Pacific Northwest National Laboratory Richland, Washington, 99352, United States peng398@pnnl.gov

Zichang He[†]

Department of Electrical and Computer Engineering University of California, Santa Barbara Santa Barbara, California, 93106, United States zichanghe@ucsb.edu

Niranjan Govind

Physical and Computational Sciences Directorate Pacific Northwest National Laboratory Richland, Washington, 99352, United States niri.govind@pnnl.gov

Yuri Alexeev **Computational Science Division** Argonne National Laboratory Lemont, Illinois 60439, United States yuri@anl.gov

Abstract-QuYBE is an open-source algebraic compiler for the compression of quantum circuits. It has been applied for the efficient simulation of the Heisenberg Hamiltonian on quantum computers. Currently, it can simulate the time dynamics of one-dimensional chains. It includes modules to generate the quantum circuits for the above as well as produce the compressed circuits, which are independent of the time step. It utilizes the Yang-Baxter equation (YBE) to perform the compression. QuYBE enables users to seamlessly design, execute, and analyze the time dynamics of the Heisenberg Hamiltonian on quantum computers. QuYBE is the first step toward making the YBE technique available to a broader community of scientists from multiple domains. The QuYBE compiler is available at https://github.com/ZichangHe/QuYBE.

I. INTRODUCTION

Current major efforts towards demonstrating a real quantum advantage are focused on designing practical large-scale quantum computer architectures. To accommodate these efforts with the supporting software, robust quantum compilation processes are much needed. A typical quantum compilation includes three-level circuit generation or optimization, namely, the circuit generation from a reduced set of universal quantum gates, the fault-tolerant design of the original circuit, and the circuit compilation to hardware-specific instructions. There is extensive literature devoted to employing classical approaches for quantum compiling (see Refs. 1]-[11 for past examples using temporal planning, machine learning, and other techniques). Recent exciting ideas include utilizing algebraic relations to compress and optimize the original circuit for more fault-resilient performance on NISQ devices [12]-[19]. The Yang-Baxter equation (YBE) or star-triangle relation has recently attracted much attention in quantum information science. This equation, which was originally introduced in theoretical physics [20] and statistical mechanics [21], is a consistency condition which arises in systems where the dynamics are two-body factorizable. The YBE link to quantum computing arises from the connection to topological entanglement, quantum entanglement, and quantum computational universality [22]–[28].

Briefly, the quantum YBE can be written as

$$(\mathcal{R} \otimes \mathbb{1})(\mathbb{1} \otimes \mathcal{R})(\mathcal{R} \otimes \mathbb{1}) = (\mathbb{1} \otimes \mathcal{R})(\mathcal{R} \otimes \mathbb{1})(\mathbb{1} \otimes \mathcal{R}) \quad (1)$$

where the \mathcal{R} operator is a linear mapping $\mathcal{R}: V \otimes V \to V \otimes V$ of vector space V. In the present discussion, \mathcal{R} represents a parameterized unitary gate and is limited as a two-qubit gate parameterized by a phase factor and a rotation. Since the YBE connects topological concepts like knots and links to entangled quantum states, the CNOT gate can be replaced by some unitary gate R via the YBE to maintain the universality of quantum computation. [21] Here, the unitary R gate serves as the solution for the condition of topological braiding, as well as the unitary solution to the YBE.

In our previous work [28], we have proved that for some model systems it is feasible to compress their time evolution circuit based on the YBE to a depth that scales linearly with respect to the number of qubits N. Furthermore, the YBE compilation in the classical pre-processing step scales as $\mathcal{O}(N^3)$. Following our previous conceptual work, and with an eye toward large-scale applications, here we report the design and execution of a quantum compiler (QuYBE) based on our previous conceptual YBE analysis. We demonstrate that our QuYBE compiler can successfully perform efficient YBE circuit compression for certain classes of the Heisenberg model.

II. METHODOLOGY

In this section, as an illustration, we introduce how our QuYBE compiler compresses the quantum circuit corresponding to the time dynamics of the 1D-Heisenberg XY model Hamiltonian. The same method can also be applied to other Heisenberg model Hamiltonians with or without external field. For an N-spin 1D-Heisenberg XY model Hamiltonian, to simulate its time dynamics with T Trotterization steps, the corresponding circuit has M = 2T layers. For such a circuit with M > N, the QuYBE compiler is able to compress the full circuit to N layers analytically without loss of accuracy. The workflow is summarized in Fig. 1. In the following, we elaborate our algorithms that have been used in the QuYBE compiler. We will first be focused on a sequential compression scheme for a demonstration purpose, and then we will briefly discuss how it can be made parallel.

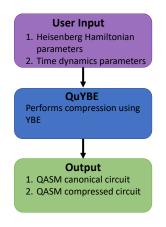


Fig. 1. Worflow of QuYBE compiler.

A. Sequential compression scheme

The overall algorithm is described in Algorithm. 1. It first transforms the first N-layer circuit to the reflection symmetric form. Then the (N + 1)-th layer can be merged into the reflected circuit, and the circuit depth is decreased by one. By repeating these two steps, we can compress the original circuit to a compact form with only N layers, i.e. circuit depth is a linear function of number of qubits N. As an example how YBE fascilitate the reflection symmetry, Fig 2 shows how to achieve reflection symmetry for four qubits where YBE is true.

Algorithm 1 simply describes how a sequential compression works for a given time evolution circuit of the Heisenberg XY model. Inside the compression algorithm, the key function CIRCUITREFLECTION is the multi-qubit reflection operation,

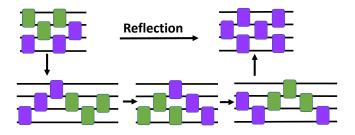


Fig. 2. Reflection symmetry can be achieved by using the YBE four times on four qubits. For each step, the triplets of qubits on which YBE act is shown by green color.

Algorithm	1	А	sequential	circuit	compression	employing
QuYBE						

Input: A N-qubit time evolution circuit with M(M > N) layers

Output: Functionally equivalent circuit with N layers 1: while M > N do

- 2: Execute the function CIRCUITREFLECTION on the first N layers.
- 3: Merge the adjacent gates on the same qubits
- 4: $M \leftarrow M 1$
- 5: end while

of which the detailed workflow is given in Algorithm 2. As can be seen, CIRCUITREFLECTION is a recursive function, which also reflects how we proved the existence of a general reflection symmetry of an *N*-layer *N*-qubit time evolution circuit for the studied models in our previous analysis. [28] Essentially, what CIRCUITREFLECTION does is to gradually offload the two-qubit gates in a given time evolution circuit from top left to lower right and restore the *untouched* two-qubit gate from lower right to top left to get the reflection done. To better illustrate how exactly the key function CIRCUITREFLECTION works, take a five-qubit time evolution circuit as an example, Fig. 3 shows how we can offload the restore the two-qubit gates to vertically reflect the first five layers of the circuite to enable the merge between the fifth and sixth layers.

Algorithm 2 Multi-qubit circuit reflection operation
1: function CIRCUITREFLECTION(circuit of <i>N</i> -qubit)
2: Offload the gate on the first two qubits to the side.
3: Perform CIRCUIT REFLECTION for the circuit of $(N - $
1)-qubit.
4: Restore the <i>untouched</i> two-qubit gates in Step 1 back
to the top row.
5: end function

Another key component in Algorithm 2, as can be seen from Fig. 3, is the three-qubit YBE (or reflection) operation which we have taken in our QuYBE compiler as a basic operation. As demonstrated in Fig. 4, the three-qubit YBE operation includes transforming an A-shape circuit to a V-shape circuit

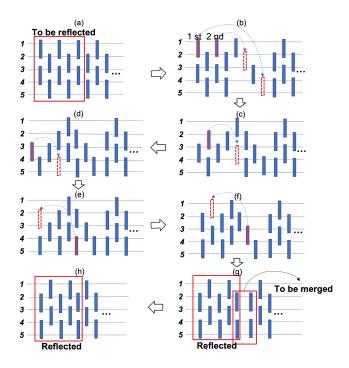


Fig. 3. Schematic demonstration of compressing one layer from the five-qubit time evolution circuit. (a) shows the original circuit. (b-g) converts the first five-layer circuit into its reflection symmetric form, where (c-e) reflects the four-layer four-qubit time evolution circuit to its reflection symmetric form. (g) shows the circuit that enables the merge of the fifth and six-th layers. (f) shows the compressed circuit.

(denoted as the A2V transformation), and its reverse operation (denoted as the V2A transformation). We have coded these two transformations, as well as a gate merge function, as core functions in our QuYBE compiler. The detailed analytic transformation rules are based on the Theorem I in Ref.28. Also, regarding the scaling of CIRCUITREFLECTION, the reflection of an N-qubit time evolution circuit with N layers requires $\mathcal{O}(N^3)$ number of 3-qubit YBE transformations (see Fig. 7 (red), also see the derivation in Appendix A of Ref.28).



Fig. 4. Three-qubit reflection operation is treated as a basic operation in our QuYBE compiler.

B. Parallel compression scheme

In comparison with some heuristic approaches, the QuYBE compiler offers a deterministic approach for compressing the time evolution circuits for some model systems. Nevertheless,

a sequential version would become a very time consuming preprocessing step given only one layer is compress after $\mathcal{O}(N^3)$ three-qubit YBE operations. To reduce the time-to-solution, we also explore the feasibility of having a parallel scheme.

1) Starting point of the circuit compression: From the above demonstration, one important observation is that the compression in Fig. 3 starts from the very beginning of the circuit structure where there are no gates prior to the circuit fragment that is about to get compressed. Therefore, we only achieve an (N + 1)-to-N compression. Alternatively, if we choose to start the compression from anywhere other than the beginning, we will get an (N+2)-to-N two-layer compression, i.e., one layer gets compressed at the beginning of the circuit fragment and the other gets compressed at the end of the circuit fragment, after performing one reflection operation (see the inset in Fig. 5).

2) Parallel compression: Another important observation is that the (N + 2)-to-N compression in the (N + 2)-layer is independent of other (N + 2)-layer circuit fragments. The independence between the circuit fragments then triggers a straightforward parallel scheme to reduce the time-to-solution. Specifically, as shown in Fig. 5, we only need to partition the original circuit to multiple circuit fragments, each of which is an (N + 2)-layer circuit. The compression of these circuit fragments can straightforwardly be distributed to classical computing resources. Suppose the entire time evolution circuit include M layers, the simple parallel scheme would require $\lceil \frac{M}{N+2} \rceil$ processors at the maximum to do the compression.

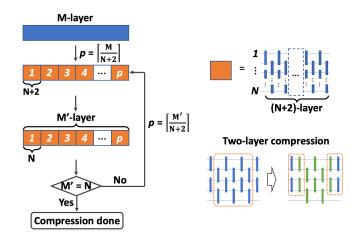


Fig. 5. A parallel execution of QuYBE compiler. p denotes the maximal number of processes. The inset at the lower right demonstrate a two-layer compression for a seven-layer five-qubit time evolution circuit if the reflection takes place at the layers #2-#6.

III. RESULTS

In this section, we discuss the impact of circuit compression by showing the reduction in two-qubits gate, i.e., CNOTs, and the growth of CNOTs with increasing system size for the time dynamics of 1D-XY Hamiltonian, a subclass of Heisenberg Hamiltonian. The Heisenberg Hamiltonian [29]–[31] is widely used to study magnetic systems, where the magnetic spins are treated quantum mechanically. The Hamiltonian (with nearest neighbor interaction) can be written as

$$\hat{H} = -\sum_{\alpha} \{ J_{\alpha} \sum_{i=1}^{N-1} \sigma_i^{\alpha} \otimes \sigma_{i+1}^{\alpha} \},$$
(2)

where α sums over $\{x, y, z\}$, the coupling parameter J_{α} denotes the exchange interaction between nearest-neighbour spins along the α -direction, and σ_i^{α} is the α -Pauli operator on the *i*th spin. For XY Hamiltonian, $J_z = 0$ in Eq. (III). We also show a real device (IBM-Manila) simulation using the same compression scheme.

Fig. 6 shows the growth of CNOTs between the canonical and compressed circuits with increasing steps in the quantum time dynamics circuit. The dashed line represents the growth

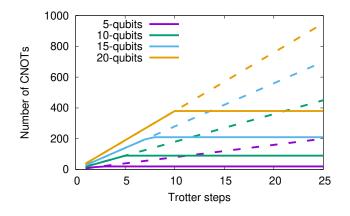


Fig. 6. Comparison between the canonical and compressed quantum circuits for the growth of CNOTs with Trotter steps. The dashed line represents the growth of CNOTs in the canonical quantum circuit, while the filled line represents the growth of CNOTs in the compressed quantum circuit.

of CNOTs with Trotter steps in the canonical circuit. On the other hand, after the compression, CNOTs grow until a step number and plateau to a constant after that, as shown by the filled line. Fig 7 (blue) shows the scaling of the maximum CNOTs required for arbitrary time dynamics in compressed circuits with the number of qubits. The growth of CNOTs is quadratic with the number of qubits. Fig 8, shows the evolution of the staggered magnetization for 5-spins for the Heisenberg XY Hamiltonian ($J_x = -0.8$, $J_y = -0.2$). Each Trotter step is 0.05 units of time. In order to confirm the validity of compression, we plot results from the canonical circuit and compressed circuit using the Qiskit simulator for every tenth step. It shows that the compressed circuit (composed of only 20 CNOTs) provides the same level of accuracy as the canonical circuit on the simulator.

The impact of the compression is clearly seen when the same simulation is run on a real quantum device. Fig. 8 shows the results of the canonical circuit and compressed circuit on the IBM-Manila quantum device. We computed the time-dependent staggered magnetization, $m_s(t)$, which can be

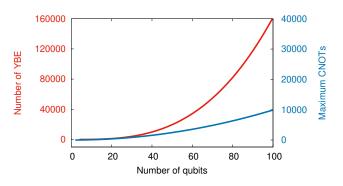


Fig. 7. Scaling of number of 3-qubit YBE transformations to finish one reflection operation (red). Scaling of maximum CNOTs is required in compressed quantum circuits for arbitrary time evolution with the number of qubits (blue).

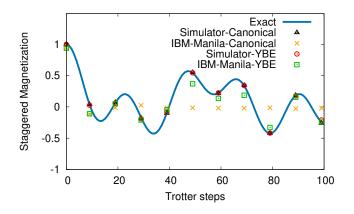


Fig. 8. Comparison of the staggered magnetization evolution for 5 spins with the XY Hamiltonian ($J_x = -0.8$, $J_y = -0.2$) between the canonical circuit and compressed circuit via YBE. Each simulation on IBM-Manila used 20000 shots.

connected to the antiferromagnetism and ferrimagnetism in materials as follows:

$$m_s(t) = \frac{1}{N} \sum_i (-1)^i \left\langle \sigma_z(t) \right\rangle.$$
(3)

The initial state is the ground state (Neél state) of the XY Hamiltonian, defined as $\Psi_0 = |\uparrow\downarrow\uparrow\downarrow\dots\uparrow\downarrow\rangle$. The staggered magnetization of the Neél state is one. We performed the time evolution for 5 units of time with a Trotter step size of 0.05 units.

The staggered magnetization using the canonical circuit quickly drops to zero and remains zero for future time steps. The increasing circuit depth and noise results in a maximally mixed state corresponding to zero magnetization. This does not happen in the compressed circuit, which only requires 20 CNOTs for arbitrary time evolution. In other words, the maximum error is capped, and we observe better qualitative and quantitative results on real quantum devices.

IV. CONCLUSIONS

In this paper, we have presented the QuYBE, which is an open-source algebraic compiler for the compression of quantum circuits. It has been applied for the efficient simulation of the time dynamics of the Heisenberg Hamiltonian on quantum computers. The compiler is based on a circuit compression algorithm based on the quantum Yang-Baxter equation. We demonstrate the efficacy of this approach by comparing time dynamics simulations with canonical and compressed circuits of the 1D-Heisenberg XY model on a real quantum device.

The QuYBE compiler is available at https://github.com/ZichangHe/QuYBE. One of our goals is to make the QuYBE compiler easy to use with existing popular quantum compilers. It is achieved by using QuYBE as an intermediate step in the compilation pipeline, where YBE takes a QASM circuit, compresses it, and generates a QASM circuit in the target format.

The proposed YBE technique is a general approach that can be applied to any quantum circuit compliant partially or fully with YBE. For example, other types of Ising Hamiltonians are obvious targets. This work is important because it paves the way for large-scale applications and investigations of broader classes of models.

ACKNOWLEDGMENT

This material is based upon work supported by the U.S. Department of Energy, Office of Science, and National Quantum Information Science Research Centers. Y.A. acknowledges support from the U.S. Department of Energy, Office of Science, under contract DE-AC02-06CH11357 at Argonne National Laboratory.

REFERENCES

- A. Oddi and R. Rasconi, "Greedy randomized search for scalable compilation of quantum circuits," in *Integration of Constraint Programming*, *Artificial Intelligence, and Operations Research*, W.-J. van Hoeve, Ed. Cham: Springer International Publishing, 2018, pp. 446–461.
- [2] T. Häner, D. S. Steiger, K. Svore, and M. Troyer, "A software methodology for compiling quantum programs," *Quantum Science and Technology*, vol. 3, no. 2, p. 020501, feb 2018.
- [3] L. E. Heyfron and E. T. Campbell, "An efficient quantum compiler that reduces t count," *Quantum Science and Technology*, vol. 4, no. 1, p. 015004, sep 2018.
- [4] L. Cincio, Y. Subaşı, A. T. Sornborger, and P. J. Coles, "Learning the quantum algorithm for state overlap," *New Journal of Physics*, vol. 20, no. 11, p. 113022, nov 2018.
- [5] D. Venturelli, M. Do, E. Rieffel, and J. Frank, "Compiling quantum circuits to realistic hardware architectures using temporal planners," *Quantum Science and Technology*, vol. 3, no. 2, p. 025004, feb 2018.
- [6] F. Chong, D. Franklin, and M. Martonosi, "Programming languages and compiler design for realistic quantum hardware," *Nature*, vol. 549, p. 180–187, 2017.
- [7] Y. Nam, N. Ross, Y. Su, A. Childs, and D. Maslov, "Automated optimization of large quantum circuits with continuous parameters," *npj Quantum Inf.*, vol. 4, p. 23, 2018.
- [8] J. Booth, "Quantum compiler optimizations," 2012.
- [9] A. G. Fowler, "Constructing arbitrary steane code single logical qubit fault-tolerant gates," *Quantum Info. Comput.*, vol. 11, no. 9–10, p. 867–873, sep 2011.
- [10] D. Maslov, G. W. Dueck, D. M. Miller, and C. Negrevergne, "Quantum circuit simplification and level compaction," *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, vol. 27, no. 3, pp. 436–444, 2008.

- [11] K. E. C. Booth, M. Do, J. C. Beck, E. Rieffel, D. Venturelli, and J. Frank, "Comparing and integrating constraint programming and temporal planning for quantum circuit compilation," 2018.
- [12] L. Bassman, R. V. Beeumen, E. Younis, E. Smith, C. Iancu, and W. A. de Jong, "Constant-depth circuits for dynamic simulations of materials on quantum computers," 2021.
- [13] E. Kökcü, D. Camps, L. Bassman, J. K. Freericks, W. A. de Jong, R. V. Beeumen, and A. F. Kemper, "Algebraic compression of quantum circuits for Hamiltonian evolution," *Physical Review A*, vol. 105, no. 3, mar 2022. [Online]. Available: https://doi.org/10.1103%2Fphysreva. 105.032420
- [14] D. Camps, E. Kökcü, L. Bassman, W. A. de Jong, A. F. Kemper, and R. Van Beeumen, "An algebraic quantum circuit compression algorithm for hamiltonian simulation," 2021. [Online]. Available: https://arxiv.org/abs/2108.03283
- [15] S.-H. Lin, R. Dilip, A. G. Green, A. Smith, and F. Pollmann, "Realand imaginary-time evolution with compressed quantum circuits," *PRX Quantum*, vol. 2, no. 1, p. 010342, 2021.
- [16] C. Cirstoiu, Z. Holmes, J. Iosue, L. Cincio, P. J. Coles, and A. Sornborger, "Variational fast forwarding for quantum simulation beyond the coherence time," *npj Quantum Information*, vol. 6, no. 1, pp. 1–10, 2020.
- [17] Y. Atia and D. Aharonov, "Fast-forwarding of hamiltonians and exponentially precise measurements," *Nature Communications*, vol. 8, no. 1, pp. 1–9, 2017.
- [18] N. F. Berthusen, T. V. Trevisan, T. Iadecola, and P. P. Orth, "Quantum dynamics simulations beyond the coherence time on NISQ hardware by variational Trotter compression," 2021. [Online]. Available: https://arxiv.org/abs/2112.12654
- [19] F. Barratt, J. Dborin, M. Bal, V. Stojevic, F. Pollmann, and A. G. Green, "Parallel quantum simulation of large systems on small nisq computers," *npj Quantum Information*, vol. 7, no. 1, pp. 1–7, 2021.
- [20] C.-N. Yang, "Some exact results for the many-body problem in one dimension with repulsive delta-function interaction," *Physical Review Letters*, vol. 19, no. 23, p. 1312, 1967.
- [21] R. J. Baxter, Exactly solved models in statistical mechanics. Elsevier, 2016.
- [22] M.-L. Ge, K. Xue, R.-Y. Zhang, and Q. Zhao, "Yang–Baxter equations and quantum entanglements," *Quantum Information Processing*, vol. 15, no. 12, pp. 5211–5242, 2016.
- [23] C. Nayak, S. H. Simon, A. Stern, M. Freedman, and S. Das Sarma, "Non-Abelian anyons and topological quantum computation," *Rev. Mod. Phys.*, vol. 80, pp. 1083–1159, Sep 2008. [Online]. Available: https://link.aps.org/doi/10.1103/RevModPhys.80.1083
- [24] L. H. Kauffman and S. J. J. Lomonaco, "Topological quantum information theory," in *Proceedings of Symposia in Applied Mathematics*, ser. Vol. 68, S. J. Lomonaco, Ed. Washington, DC: AMS, 2010.
- [25] Y. Zhang, "Integrable quantum computation," *Quantum Inf. Process.*, vol. 12, pp. 631–639, 2013.
- [26] F. A. Vind, A. Foerster, I. S. Oliveira, R. S. Sarthour, D. d. O. Soares-Pinto, A. M. d. Souza, and I. Roditi, "Experimental realization of the Yang-Baxter equation via NMR interferometry," *Scientific reports*, vol. 6, no. 1, pp. 1–8, 2016.
- [27] M. T. Batchelor and A. Foerster, "Yang-Baxter integrable models in experiments: from condensed matter to ultracold atoms," *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 17, p. 173001, 2016.
- [28] B. Peng, S. Gulania, Y. Alexeev, and N. Govind, "Quantum time dynamics employing the yang-baxter equation for circuit compression," *Phys. Rev. A*, vol. 106, p. 012412, Jul 2022.
- [29] P. Fazekas, Lecture notes on electron correlation and magnetism. World scientific, 1999, vol. 5.
- [30] R. Skomski, Simple models of magnetism. Oxford University Press, 2008.
- [31] A. S. T. Pires, *Theoretical Tools for Spin Models in Magnetic Systems*. IOP Publishing, 2021.