

Zero and Finite Temperature Quantum Simulations Powered by Quantum Magic

Andi Gu¹, Hong-Ye Hu¹, Di Luo^{1,2,3}, Taylor L. Patti⁴, Nicholas C. Rubin⁵, and Susanne F. Yelin^{1,3}

¹Department of Physics, Harvard University, 17 Oxford Street, Cambridge, MA 02138, USA

²Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

³The NSF AI Institute for Artificial Intelligence and Fundamental Interactions

⁴NVIDIA, Santa Clara, CA 95051, USA

⁵Google Quantum AI, Venice, CA 90291, United States

We introduce a quantum information theory-inspired method to improve the characterization of many-body Hamiltonians on near-term quantum devices. We design a new class of similarity transformations that, when applied as a preprocessing step, can substantially simplify a Hamiltonian for subsequent analysis on quantum hardware. By design, these transformations can be identified and applied efficiently using purely classical resources. In practice, these transformations allow us to shorten requisite physical circuit-depths, overcoming constraints imposed by imperfect near-term hardware. Importantly, the quality of our transformations is *tunable*: we define a ‘ladder’ of transformations that yields increasingly simple Hamiltonians at the cost of more classical computation. Using quantum chemistry as a benchmark application, we demonstrate that our protocol leads to significant performance improvements for zero and finite temperature free energy calculations on both digital and analog quantum hardware. Specifically, our energy estimates not only outperform traditional Hartree-Fock solutions, but this performance gap also consistently widens as we tune up the quality of our transformations. In short, our quantum information-based approach opens promising new pathways to realizing useful and feasible quantum chemistry algorithms on near-term hardware.

A central task in quantum chemistry is determining ground state energies and finite temperature free energies for electronic Hamiltonians. While numerous algorithms aim to leverage quantum hardware for solutions [21, 31, 47, 48], the constraints of near-term hardware, particularly limited circuit depth, pose challenges. One approach to address this difficulty is

classically assisted quantum algorithms, wherein computation is offloaded to a classical computer where possible, and the quantum hardware only supplies the quantum resources which are absolutely necessary to identify an accurate solution. A natural way in which a classical computer might assist a quantum algorithm is to preprocess the electronic Hamiltonian into a more favorable form before it is analyzed on quantum hardware. For instance, one can use a classical algorithm to partition the terms of a Hamiltonian H into commuting groups which can be measured simultaneously, so that calculating the expectation $\langle H \rangle$ can be done with fewer samples [42].

In this work, we consider another preprocessing method: *similarity transformations*, wherein our aim is to identify a unitary change of basis U such that the transformed Hamiltonian $H_{\text{eff}} \equiv U^\dagger H U$ is more amenable to analysis on near-term quantum hardware. Indeed, similarity transforms have long been used to simplify Hamiltonians [4, 8, 13, 22, 24, 33, 37, 44, 46]. In this work, we take a novel quantum information theoretic approach to similarity transforms: we center our techniques around two fundamental measures of ‘quantumness’ – namely, entanglement and nonstabilizerness (popularly termed ‘quantum magic’). These notions of quantumness formalize the intuition that the degree to which a state is truly quantum can be measured by how difficult it is to simulate classically. For instance, low entanglement states can be simulated by matrix product states [10]. However, entanglement alone is insufficient to completely characterize quantumness – classically simulable (i.e., using classical resources that scales polynomially in system size) circuits [1], known as Clifford circuits can generate an unbounded amount of entanglement. Indeed, it has been shown that random Clifford circuits will almost always produce volume-law entangled states [43]. Yet, these circuits are unable to express a different quantum resource: quantum magic. More precisely, Clifford circuits are generated by a non-universal gateset (composed of the Hadamard, phase, and CNOT gates); this has motivated the formal definition of magic as a measure

Andi Gu: Equal contributions (Alphabetical order)

Hong-Ye Hu: Equal contributions (Alphabetical order)

Di Luo: Equal contributions (Alphabetical order)

Susanne F. Yelin: syelin@g.harvard.edu

of the amount of non-Clifford resources required to prepare a particular state [18, 35]. Like entanglement, quantum magic has also been shown to be directly connected with the hardness of classical simulation [5].

Our main insight in this work is that similarity transforms can be used to eliminate spurious quantumness (i.e., classically simulable entanglement and magic) in the Hamiltonian’s eigenbasis before it is analyzed on quantum hardware. Then, the real hardware only needs to supply the fundamentally necessary quantumness for the problem, hence helping shorten requisite circuit depths. We introduce a class of similarity transformations, represented as Clifford circuits augmented with a few (non-Clifford) single-qubit Z -rotation gates. For brevity, we call these circuits Clifford + k Rz circuits. Conveniently, this design allows good similarity transforms to be both identified and applied *efficiently using purely classical resources*. The Clifford component of the circuit can be interpreted as eliminating spurious entanglement from H_{eff} , while the non-Clifford gates eliminate spurious magic. Importantly, by varying the number of non-Clifford gates, we show that the power of these transformations is *tunable*, so that we can identify increasingly simple eigenbases (i.e., exhibiting less entanglement and magic) at the cost of increased classical computational resources. Since this tunability comes in discrete steps, we term this the “quantum magic ladder” (see Fig. 1). By applying our approach to zero and finite temperature free energy calculations, we demonstrate that this reduction in quantumness can have significant practical benefits. Specifically, we show that our similarity transforms find energies which are significantly more accurate than other popular classical solutions, and moreover, the performance gap only improves as we climb the quantum magic ladder. Furthermore, we present a classical algorithm which outputs a new (simplified) effective Hamiltonian H_{eff} , related to the original Hamiltonian via a similarity transform. We show that the resultant simplified Hamiltonians H_{eff} produced by our method enables significantly reduced circuit depths on quantum hardware. Under this quantum information theoretic lens, these results demonstrate a promising pathway for designing practical quantum chemistry algorithms in the NISQ era.

Quantum Chemistry. Quantum chemistry is a natural domain for assessing the efficacy of our protocol. The Hamiltonians produced by quantum chemistry systems have immediate physical interpretations, allowing us to apply physical intuition to better understand how our protocol modulates the ‘quantumness’ of a problem. Furthermore, there is a wealth of established quantum chemistry methods against which we can benchmark our protocol, such as the Hartree-Fock (HF) approximation. In short, quantum chemistry allows us to demonstrate the advantages of our

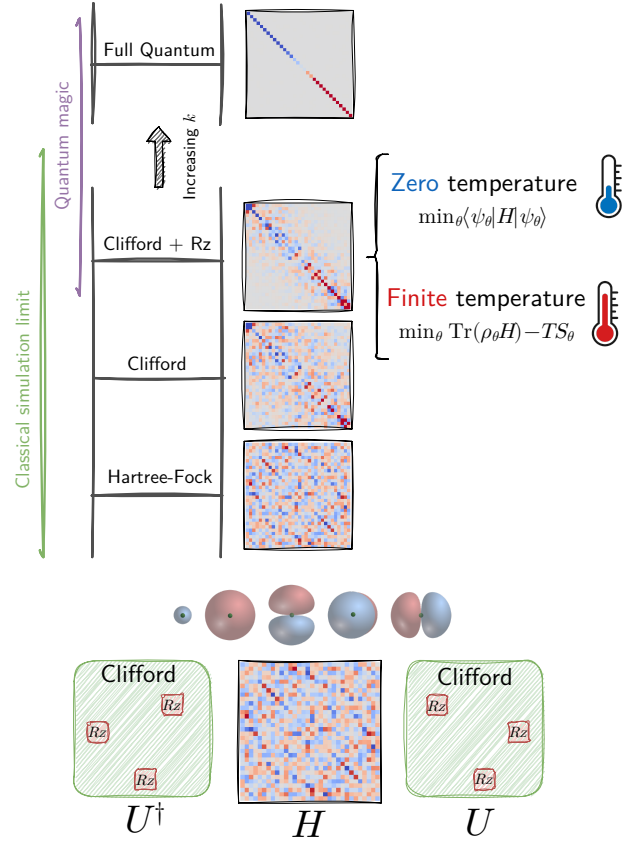


Figure 1: A general approach for similarity transformations using the quantum magic ladder. Starting with a second quantized Hamiltonian H , we use classically simulable quantum circuits (e.g., Clifford + k Rz circuits) to reduce the quantumness of the Hamiltonian by the basis transformation $H_{\text{eff}} = U^\dagger H U$. We climb up the quantum magic ladder by increasing k , which increases the degree to which U can simplify the Hamiltonian (by reducing the quantum magic of the H_{eff} eigenbasis). Note that for finite temperature calculations, when working in the grand canonical ensemble, we replace H with $H - \mu N$.

approach in a well-understood physical context.

In quantum chemistry, one key goal of the electronic structure problem is to solve the energy of many-body electronic Hamiltonians at zero and finite temperature. These electronic Hamiltonians are often expressed in a molecular orbital basis:

$$H = \sum_{p,q} h_{pq} \hat{c}_p^\dagger \hat{c}_q + \frac{1}{2} \sum_{p,q,r,s} V_{pqrs} \hat{c}_p^\dagger \hat{c}_q^\dagger \hat{c}_s \hat{c}_r, \quad (1)$$

where p, q, r, s label the spin-orbitals and $\hat{c}_\alpha, \hat{c}_\alpha^\dagger$ are the annihilation and creation fermionic operators. Anticipating analysis on qubit-based quantum computers, we apply a fermion-qubit mapping (e.g., parity mapping [36], Jordan-Wigner [20], Bravyi-Kitaev [41]) to map the fermionic Hamiltonian to a qubitized Hamiltonian $H = \sum_i \alpha_i P_i$, where P_i is an n -qubit Pauli operator. For the molecules considered in this work (LiH and H_2O), the qubitized Hamiltonians contain $O(10^2)$ Pauli strings. The explicit construc-

tion of these Hamiltonians was performed using the `OpenFermion` library [28]. In this work, we focus on two paradigmatic problems involving these electronic Hamiltonians, which are to calculate the ground state energy of H , and the free energy of the Gibbs state at finite temperature [9]. In this work, we work in the grand canonical ensemble, so that the Gibbs state is $\rho \propto \exp(-\beta(H - \mu N))$, with inverse temperature β , chemical potential μ , and particle number N .

The Quantum Magic Ladder. Similarity transformations have historically occupied a key role in the development of quantum mechanics. An early example is the Schrieffer-Wolff method, which generates low-energy effective Hamiltonians by constructing similarity transformations that separate high-energy and low-energy subspaces [4, 8, 33, 44]. In keeping with this tradition, we argue that similarity transforms can also play a central role in bringing quantum chemistry algorithms closer to practical feasibility in the NISQ era. One way of understanding the role of a similarity transformation U , particularly in the context of near-term hardware, is as a virtual circuit that is appended to the end of a ground state preparation circuit V . That is, if we define a transformed Hamiltonian $H_{\text{eff}} = U^\dagger H U$, we observe that $\langle V^\dagger H_{\text{eff}} V \rangle = \langle (UV)^\dagger H (UV) \rangle$, so that, in effect, the applied circuit is UV . This is especially important for NISQ devices, which face limited circuit depths (hence limited expressive power of V). An appropriate similarity transformation U can act like an appended circuit that significantly eases the burden of the real quantum circuit V .

In what follows, we describe a unified framework to programatically construct a new class of similarity transformations for quantum chemistry algorithms, a system we term “the quantum magic ladder.” These transformations are comprised of Clifford circuits augmented by a few Rz gates. There are several reasons for the attractiveness of the Clifford + k Rz approach. Firstly, Clifford circuits are classically simulable with complexity $O(n^3)$, where n is the system size [1]. As we discuss below, these Clifford circuits can be dramatically enhanced by doping them with just a few (k) non-Clifford gates. While this enhancement comes at the cost of an increase in simulation complexity that is exponential in k , Clifford + k Rz transforms with $k \leq O(\log n)$ can still be simulated with polynomial classical resources [1, 3, 6]. Secondly our circuits strictly generalize more conventional circuits such as Clifford circuits with a few T gates. Indeed, we find that the additional continuous degrees of freedom introduced by Clifford + k Rz circuits (i.e., the rotation angle of each Rz gate) is both easy to optimize and yields significant performance improvements. Furthermore, these improvements are free in the sense that the additional simulation cost for k Rz gates is exactly the same as that of k T gates. The simulation of these Clifford + k Rz circuits was performed

using the `PyClifford` package [19] (for details, see Appendix A). Finally, as illustrated in Fig. 1, ‘climbing’ the quantum magic ladder provides a *systematic* way to increase the expressive power of our similarity transformation: one simply needs to increase k .

The enhanced expressivity of ansatzes higher on this ladder can be understood through the lens of quantum magic. Much like entanglement defines one notion of quantumness, quantum magic, or nonstabilizerness, is another metric which measures the degree to which a state can be considered truly quantum [18, 35]. Our main insight in this work is that the optimal Clifford + k Rz transform can model the quantumness of H along *both* axes of quantumness, capturing its entanglement via the Clifford component of the circuit, and its quantum magic via the non-Clifford components. As we climb the magic ladder, we increase the degree to which our transform can model the quantum magic of the energy eigenbasis (hence reducing the magic of the transformed Hamiltonian’s eigenbasis). In the limit as $k \rightarrow \infty$, our ansatz can model any unitary [29], which ultimately leads to a full reduction in the Hamiltonian’s quantumness (i.e., complete diagonalization).

Our Clifford + k Rz circuits are parameterized as $U(\Theta)$, where Θ is a combination of discrete and continuous parameters, corresponding to the Clifford and Rz components respectively. To identify good transformations $U(\Theta)$, we make use of the variational principle for both ground state energies and finite temperature free energies. Specifically, consider the zero temperature case: $U(\Theta)$ defines a parameterized effective Hamiltonian $H_{\text{eff}}(\Theta)$, and our aim is then to minimize $\langle 0 | H_{\text{eff}}(\Theta) | 0 \rangle$ (which is equivalent to minimizing the energy, with respect to H , of $U(\Theta) | 0 \rangle$). Optimizing Θ is a hybrid discrete-continuous problem with a large search space, so we develop a three-step approach. First, we observe that qubitized electronic Hamiltonians strongly resemble disordered Hamiltonians, whose energy coefficients span a wide range of scales. We leverage this formal similarity by using the spectrum bifurcation renormalization group [13, 37, 46] (which produces Clifford similarity transformations effective for modelling strongly-disordered quantum materials) to find an initialization point Θ_0 for our Clifford + k Rz transformation. We then perform a combination of greedy search and gradient descent to further optimize Θ (see Appendix B for details).

Results. We apply our Clifford+ k Rz transformation for both ground state and finite temperature quantum chemistry calculations [7, 12, 16, 32, 39] on LiH ($n = 8$) and H₂O ($n = 12$). While these system sizes are modest compared to the capabilities of state-of-the-art classical numerical methods, they serve as valuable proof-of-concept demonstrations for our quantum-classical hybrid approach, and are representative of the system scales that are expected to be tractable on near-term quantum devices. We start

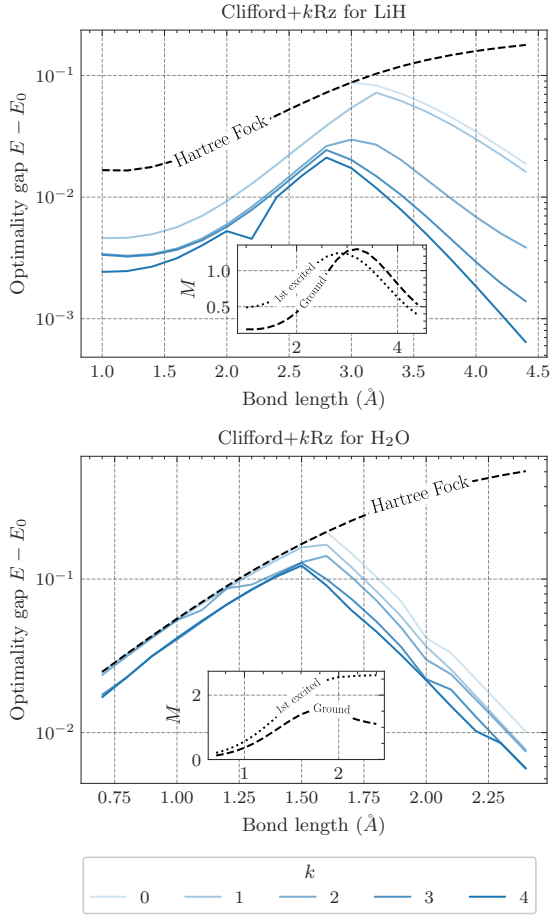


Figure 2: $E - E_0$ is the gap between E , the energy calculated by the optimized Clifford + k Rz circuit, and E_0 , the true ground state energy. We illustrate this gap for LiH and H₂O as a function of their bond lengths. Note that the HF solution is consistently higher in energy, as it is strictly less expressive than the Clifford ansatz, even when $k = 0$, underscoring the enhanced representational power of the Clifford circuits in capturing the ground state properties. Furthermore, the accuracy (as measured by $E - E_0$) improves as we climb the quantum magic ladder. The inset shows the stabilizer Rényi entropy M , which quantifies the magic of the ground and first excited states at different bond lengths.

with the zero temperature case, where we aim to identify a ground state energy $E = \langle 0 | H_{\text{eff}}(\Theta) | 0 \rangle$ that closely approximates the true ground state energy E_0 . In Fig. 2, the energy gap $E - E_0$ is plotted as a function of different bond lengths between atoms, where E is the energy predicted by an optimized Clifford + k Rz transformation. First, compare the accuracy to the accuracy of the Hartree-Fock (HF) solution: since the HF solution is a strict subset of the Clifford ansatz [38], the energy gap of a pure Clifford transformation ($k = 0$) is upper bounded by that of the HF solution. However, a pure Clifford circuit only surpasses the HF solution when the bond length exceeds the equilibrium bond length. In contrast, as we climb the quantum magic ladder, the energy gap is significantly reduced for *all* the bond lengths,

including bond lengths shorter than the equilibrium bond length. Moreover, as the number of non-Clifford gates increases, the energy gap consistently decreases. This is clear evidence that tuning up the power of the Clifford + k Rz ansatz can significantly improve accuracy for ground state calculations. It is important to note that the number of non-Clifford gates k should be considered in relation to the total number of qubits n . If we allow $k = \exp(\Theta(n))$, the doped circuit can approximate an arbitrary unitary gate. In contrast, in our study, we focus on the regime where $k \ll n$ to maintain the classical simulability of the circuits while still enhancing their expressibility. We observe that the convergence of our method improves as k increases, with significant improvements seen for $k = 1$ and $k = 2$. However, the improvement becomes marginal for higher values of k . There is a trade-off between the expressibility gained by increasing k and the computational cost associated with it; it appears that there are sharply diminishing returns beyond $k = 2$. Future work could investigate this trade-off in more detail and develop strategies for optimally selecting k based on the problem at hand. We also note the presence of kinks in some curves of Fig. 2, which are artifacts likely due to numerical precision limitations and minor optimization errors, rather than physical phenomena.

Interestingly, the energy gap peaks around a particular bond length, indicating a region of maximal quantum magic. To show this, we employ a popular metric specifically designed to quantify the magic of a state $|\psi\rangle$. This metric $M(|\psi\rangle)$ is known as the stabilizer entropy [23], and is defined as

$$M(\psi) = -\log_2 \left(\sum_{P \in \mathbb{P}_n} \frac{\text{tr}^4(P\psi)}{2^n} \right), \quad (2)$$

where $\mathbb{P}_n$ is the n -qubit Pauli group. Notably, it obeys $M(|\psi\rangle) = 0$ if and only if $|\psi\rangle$ is a stabilizer state, and it is strictly positive otherwise. In general, a higher value of $M(|\psi\rangle)$ indicates a greater degree of nonstabilizerness [23]. We see in the inset of Fig. 2 that quantum magic (as measured by M) and the energy gap peak near the same bond length, indicating that regions where the ground state has high quantum magic cannot accurately be reduced by basis transformations with insufficient non-Clifford gates. Interestingly, we observe that the magic of the ground state is typically very small at extreme bond lengths, and peaks at intermediate bond lengths. This peak in magic and energy gap likely corresponds to the Coulson-Fisher point [11], where the electronic structure transitions from single-reference to multi-reference character. This transition point is known to be challenging for computational methods [25], aligning with our observation of increased quantum resources being necessary in this region. Our quantum information theoretic approach thus provides a new perspective on this well-established concept in quan-

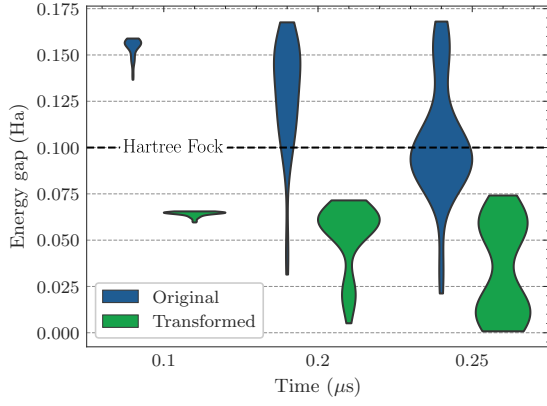


Figure 3: Pulse time reduction with the transformed Hamiltonian, depicted as a violin plot. This displays the probability density of a solution having a particular energy gap, with the width representing the solution density at each energy level. We model an analog quantum machine with a ^{87}Rb atomic chain for LiH ground state preparation using fixed lattice spacing and total evolution time. The Rabi frequency and the local detuning of each atom are optimized. For each evolution time, twenty randomly initialized variational ground states are optimized.

tum chemistry. This suggests that in practice, hyperparameters like k ought to be set accordingly — for extreme bond lengths, a small k should suffice, while intermediate bond lengths will require a larger k to accurately capture the increased magic. Near these intermediate bond length regions, accurate results demand either more powerful classical similarity transformations (i.e., larger k) or further analysis on real quantum devices.

We demonstrate this by applying our method on analog quantum computers. Specifically, we classically simulate the dynamics of a one-dimensional Rydberg array, which serves as a programmable quantum simulator (PQS) for preparing a ground state of LiH at a bond length $3.4\text{\AA}$. Note that this bond length exhibits high magic (Fig. 2), indicating that additional quantumness, supplied by the PQS, is necessary to prepare the ground state. To prepare this state, we fix the total PQS evolution time T and evolve the initial state $|0\rangle$ under a time-dependent Rydberg Hamiltonian [2] with variationally optimized site-dependent Rabi frequencies $\Omega_i(t)$ and local detunings $\Delta_i(t)$. For each fixed total evolution time T , we variationally optimize 20 randomly initialized ground states for both the bare Hamiltonian H , and for the Clifford-transformed Hamiltonian H_{eff} , using automatic differentiation with Jax [15] and Tensorcircuit [50] to directly optimize the pulse shapes $\Omega_i(t), \Delta_i(t)$. The results are shown in Fig. 3. Consider just the case of $T = 0.1\mu\text{s}$, for which dramatic differences are already clear. The optimized energy with the original Hamiltonian is concentrated at levels significantly above the HF energy, because the entanglement generated by the PQS within this small timeframe is limited. How-

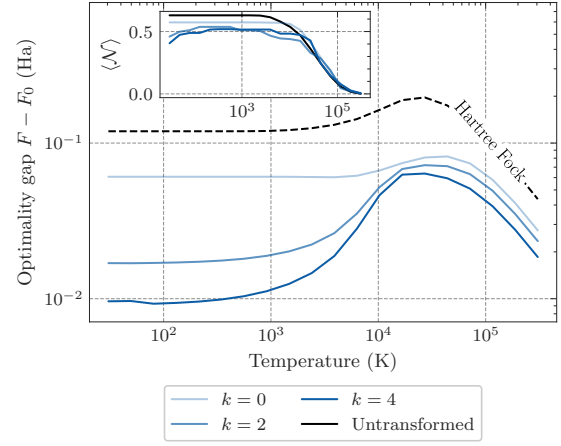


Figure 4: Free energy gap for LiH at a bond length $3.4\text{\AA}$. The inset shows the entanglement negativity $\mathcal{N}$ averaged over each site of the Gibbs state for the transformed Hamiltonian.

ever, the Clifford + $k\text{Rz}$ similarity transform already significantly reduces the entanglement of the ground state before evolution on the PQS; hence, even under the limited evolution time $T = 0.1\mu\text{s}$, the PQS can quickly identify a solution that vastly surpasses the accuracy of the Hartree Fock solution. This demonstrates that Clifford + $k\text{Rz}$ transformations effectively reduce requisite entanglement and pulse times, a significant advantage for near-term quantum devices.

For finite temperature calculations, we apply our approach to the LiH molecule for grand free energy estimation. Our ansatz for the density matrix is

$$\rho(\Theta) = \sum_{x \in \{0,1\}^n} p(x) U(\Theta) |x\rangle\langle x| U^\dagger(\Theta), \quad (3)$$

where $U(\Theta)$ is a Clifford + $k\text{Rz}$ circuit. Again applying the variational principle, we solve for Θ by minimizing the free energy $F = \langle H \rangle - \mu \langle N \rangle - TS$ of $\rho(\Theta)$ (see Appendix B for details). The results in Fig. 4 show that the Clifford + $k\text{Rz}$ ansatz consistently reduces the free energy gap $F - F_0$ (where F is the free energy of the optimized density matrix $\rho(\Theta)$, and F_0 is the free energy of the true Gibbs state) as we climb the quantum magic ladder. Similar to the zero temperature case, all Clifford transformations outperform the finite temperature HF method, including the pure Clifford transformation ($k = 0$), which reduces the free energy gap by $\sim 50\%$.

To quantify the extent to which our transforms reduce the quantumness of the transformed Hamiltonian, we examine the entanglement features and off-diagonal terms of the exact Gibbs state for the transformed Hamiltonian H_{eff} . We quantify the entanglement of a mixed state with its entanglement negativity $\mathcal{N}(\rho) = (\|\rho^{\Gamma_A}\|_1 - 1)/2$, where ρ^{Γ_A} is the partial transpose of ρ with respect to subsystem A [14, 17, 30]. Using this metric, in the inset of Fig. 4, we show that our similarity transformations indeed consistently reduce the entanglement of the

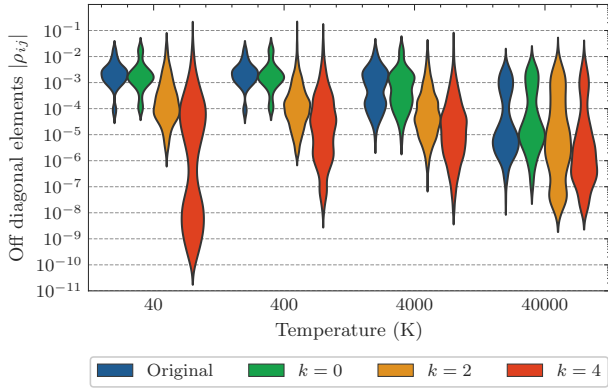


Figure 5: Density distribution of the off-diagonal term for the Gibbs state ρ of LiH across different temperatures for bond length 3.4 Å. The width of features is proportional to solution density.

Gibbs state. Furthermore, Fig. 5 shows that, across a wide temperature range, our transformation’s ability to suppress off-diagonal elements in the exact Gibbs state ρ grows with k . This provides further evidence that Clifford + k Rz transformations provide an effective method to reduce the quantumness of a Hamiltonian, and moreover, that this reduction is tunable by varying k .

Conclusions. We have developed a general approach for zero and finite temperature quantum calculations from a quantum information theoretic perspective, facilitated by the Clifford + k Rz transformation. One of the primary benefits of our approach is that it is fully classically simulable for small k , providing an convenient method to conserve costly quantum resources. Furthermore, by tuning k , we provide a means of customizing the similarity transforms produced by our technique: larger k enables greatly increased expressive power and accuracy, at the cost of increased classical computation. This tunability forms the basis of what we term the quantum magic ladder. We apply these methods in the context of ground state and free energy calculations for quantum chemistry, demonstrating superior performance to HF techniques and consistent improvement as we ascend the quantum magic ladder, for both digital and analog quantum computers. We examine how the Clifford + k Rz transformation reduces quantumness (as measured by entanglement and magic) before any algorithms are run on quantum hardware, and demonstrated how this reduction directly translates to reduced requirements for quantum resources. While our approach shows promise for small-scale systems like LiH and H₂O, it is necessary to investigate its scalability to larger systems. The classical simulation cost of the Clifford+ k Rz circuits grows exponentially with k , which could limit the applicability of our method to larger systems. However, it is encouraging to note that we are already able to find significant improvements for $k \leq 2$, which suggests that even

a small number of strategically placed non-Clifford gates can provide significant benefits. As quantum hardware continues to improve, we expect that the insights and techniques developed in this work, such as the use of classically-optimized similarity transformations to reduce the “quantumness” of the problem, will be valuable for scaling up to larger system sizes.

Having established this framework, we also identify a number of remaining open questions and future directions. Algorithmically, in addition to the greedy search algorithms used in this work, powerful methods such as simulated annealing and differentiable quantum architecture search [49] could be considered. Our similarity transformation based on doped Clifford circuits could also potentially augment well-known numerical approaches such as density matrix renormalization group (DMRG) [34]. Since our method identifies effective Hamiltonians whose ground states have significantly lower entanglement compared to the original problem, it could potentially aid in the construction of more efficient tensor network representations. Moreover, recent work [27] has shown that the stabilizer formalism can be integrated with tensor networks. By transforming the problem into a basis where the ground state is closer to a stabilizer state, our approach could enable the use of these compact representations. Investigating the interplay between our similarity transformations and tensor network methods could lead to powerful new tools for simulating quantum many-body systems. Finally, to scale finite temperature calculations to larger systems, high-dimensional classical probability distributions of the Gibbs state can be simulated using generative models in the spirit of β -VQE [26]. Our approach can also be applied to the finite temperature calculation based on the thermal double field framework [40, 45]. This offers an opportunity to tailor our transformations to specific applications, opening new pathways for near-term quantum computers to address a wide array of quantum chemistry problems.

Acknowledgments

The authors would like to thank Chong Sun, Yi Tan, Linqing Peng, Yuan Liu, Shi-Xin Zhang, Joonho Lee, James Whitfield, and Ryan Babbush for the helpful discussions. DL acknowledges support from the NSF AI Institute for Artificial Intelligence and Fundamental Interactions (IAIFI). This material is based upon work supported by the U.S. Department of Energy, Office of Science, National Quantum Information Science Research Centers, Co-design Center for Quantum Advantage (C2QA) under contract number DE-SC0012704. HYH is grateful for the support from Harvard Quantum Initiative Fellowship. SFY would like to thank the NSF through a Cornell HDR and the CUA PFC grant.

References

- [1] S. Aaronson and D. Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70:052328, Nov. 2004. DOI: [10.1103/PhysRevA.70.052328](https://doi.org/10.1103/PhysRevA.70.052328).
- [2] C. S. Adams, J. D. Pritchard, and J. P. Schaffer. Rydberg atom quantum technologies. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 53(1):012002, Dec. 2019. DOI: [10.1088/1361-6455/ab52ef](https://doi.org/10.1088/1361-6455/ab52ef).
- [3] S. Bravyi and D. Gosset. Improved Classical Simulation of Quantum Circuits Dominated by Clifford Gates. *Phys. Rev. Lett.*, 116:250501, June 2016. DOI: [10.1103/PhysRevLett.116.250501](https://doi.org/10.1103/PhysRevLett.116.250501).
- [4] S. Bravyi, D. P. DiVincenzo, and D. Loss. Schrieffer–Wolff transformation for quantum many-body systems. *Annals of Physics*, 326(10):2793–2826, Oct. 2011. ISSN 00034916. DOI: [10.1016/j.aop.2011.06.004](https://doi.org/10.1016/j.aop.2011.06.004).
- [5] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard. Simulation of quantum circuits by low-rank stabilizer decompositions. *Quantum*, 3:181, Sept. 2019. DOI: [10.22331/q-2019-09-02-181](https://doi.org/10.22331/q-2019-09-02-181).
- [6] S. Bravyi, D. Gosset, and Y. Liu. How to Simulate Quantum Measurement without Computing Marginals. *Phys. Rev. Lett.*, 128:220503, June 2022. DOI: [10.1103/PhysRevLett.128.220503](https://doi.org/10.1103/PhysRevLett.128.220503).
- [7] B. J. Brown, D. Loss, J. K. Pachos, C. N. Self, and J. R. Wootton. Quantum memories at finite temperature. *Rev. Mod. Phys.*, 88:045005, Nov. 2016. DOI: [10.1103/RevModPhys.88.045005](https://doi.org/10.1103/RevModPhys.88.045005).
- [8] L. S. Cederbaum, J. Schirmer, and H. D. Meyer. Block diagonalisation of Hermitian matrices. *Journal of Physics A: Mathematical and General*, 22(13):2427–2439, July 1989. ISSN 0305-4470, 1361-6447. DOI: [10.1088/0305-4470/22/13/035](https://doi.org/10.1088/0305-4470/22/13/035).
- [9] A. N. Chowdhury And and R. D. Somma. Quantum algorithms for Gibbs sampling and hitting-time estimation. *Quantum Information and Computation*, 17(1&2):41–64, Jan. 2017. ISSN 15337146, 15337146. DOI: [10.26421/QIC17.1-2-3](https://doi.org/10.26421/QIC17.1-2-3).
- [10] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete. Matrix product states and projected entangled pair states: Concepts, symmetries, theorems. *Reviews of Modern Physics*, 93(4), Dec. 2021. DOI: [10.1103/revmodphys.93.045003](https://doi.org/10.1103/revmodphys.93.045003).
- [11] C. Coulson and I. Fischer. Xxxiv. notes on the molecular orbital treatment of the hydrogen molecule. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 40(303):386–393, 1949. DOI: [10.1080/14786444908521726](https://doi.org/10.1080/14786444908521726). URL <https://doi.org/10.1080/14786444908521726>.
- [12] J. F. Donoghue, B. R. Holstein, and R. Robinett. Quantum electrodynamics at finite temperature. *Annals of Physics*, 164(2):233–276, 1985. ISSN 0003-4916. DOI: [https://doi.org/10.1016/0003-4916\(85\)90016-8](https://doi.org/10.1016/0003-4916(85)90016-8).
- [13] C. M. Duque, H.-Y. Hu, Y.-Z. You, V. Khemani, R. Verresen, and R. Vasseur. Topological and symmetry-enriched random quantum critical points. *Phys. Rev. B*, 103:L100207, Mar. 2021. DOI: [10.1103/PhysRevB.103.L100207](https://doi.org/10.1103/PhysRevB.103.L100207).
- [14] J. Eisert and M. B. Plenio. A comparison of entanglement measures. *Journal of Modern Optics*, 46(1):145–154, Jan. 1999. ISSN 1362-3044. DOI: [10.1080/09500349908231260](https://doi.org/10.1080/09500349908231260).
- [15] R. Frostig, M. Johnson, and C. Leary. Compiling machine learning programs via high-level tracing. 2018. URL <https://mlsys.org/Conferences/doc/2018/146.pdf>.
- [16] N. Haldolaarachchige, A. B. Karki, W. A. Phe-lan, Y. M. Xiong, R. Jin, J. Y. Chan, S. Stadler, and D. P. Young. Effect of chemical doping on the thermoelectric properties of FeGa3. *Journal of Applied Physics*, 109(10), May 2011. ISSN 1089-7550. DOI: [10.1063/1.3585843](https://doi.org/10.1063/1.3585843).
- [17] M. Horodecki, P. Horodecki, and R. Horodecki. Separability of mixed states: necessary and sufficient conditions. *Physics Letters A*, 223(1–2):1–8, Nov. 1996. ISSN 0375-9601. DOI: [10.1016/s0375-9601\(96\)00706-2](https://doi.org/10.1016/s0375-9601(96)00706-2). URL [http://dx.doi.org/10.1016/S0375-9601\(96\)00706-2](http://dx.doi.org/10.1016/S0375-9601(96)00706-2).
- [18] M. Howard and E. Campbell. Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing. *Physical Review Letters*, 118(9), Mar. 2017. ISSN 1079-7114. DOI: [10.1103/physrevlett.118.090501](https://doi.org/10.1103/physrevlett.118.090501).
- [19] H.-Y. Hu, C. Zhao, T. L. Patti, A. Gu, Y.-Z. You, and S. Yelin. PyClifford: Efficient Clifford Circuit Simulation in Python. *In preparation*, 2023.
- [20] P. Jordan and E. P. Wigner. About the Pauli exclusion principle. *Z. Phys.*, 47:631–651, 1928. DOI: [10.1007/BF01331938](https://doi.org/10.1007/BF01331938).
- [21] C. Kang, N. P. Bauman, S. Krishnamoorthy, and K. Kowalski. Optimized Quantum Phase Estimation for Simulating Electronic States in Various Energy Regimes. *Journal of Chemical Theory and Computation*, 18(11):6567–6576, Nov. 2022. DOI: [10.1021/acs.jctc.2c00577](https://doi.org/10.1021/acs.jctc.2c00577).
- [22] V. Khemani, F. Pollmann, and S. L. Sondhi. Obtaining Highly Excited Eigenstates of Many-Body Localized Hamiltonians by the Density Matrix Renormalization Group Approach. *Phys. Rev. Lett.*, 116:247204, June 2016. DOI: [10.1103/PhysRevLett.116.247204](https://doi.org/10.1103/PhysRevLett.116.247204).
- [23] L. Leone, S. F. Oliviero, and A. Hamma. Stabilizer Rényi Entropy. *Physical Review Letters*, 128(5), Feb. 2022. DOI: [10.1103/physrevlett.128.050402](https://doi.org/10.1103/physrevlett.128.050402).

- [24] S. P. Lim and D. N. Sheng. Many-body localization and transition by density matrix renormalization group and exact diagonalization studies. *Phys. Rev. B*, 94:045111, July 2016. DOI: [10.1103/PhysRevB.94.045111](https://doi.org/10.1103/PhysRevB.94.045111).
- [25] P. A. Limacher, P. W. Ayers, P. A. Johnson, S. De Baerdemacker, D. Van Neck, and P. Bultinck. A new mean-field method suitable for strongly correlated electrons: Computationally facile antisymmetric products of nonorthogonal geminals. *Journal of chemical theory and computation*, 9(3):1394–1401, March 2013. ISSN 1549-9618. DOI: [10.1021/ct300902c](https://doi.org/10.1021/ct300902c). URL <https://doi.org/10.1021/ct300902c>.
- [26] J.-G. Liu, L. Mao, P. Zhang, and L. Wang. Solving quantum statistical mechanics with variational autoregressive networks and quantum circuits. *Machine Learning: Science and Technology*, 2(2):025011, Feb. 2021. DOI: [10.1088/2632-2153/aba19d](https://doi.org/10.1088/2632-2153/aba19d).
- [27] S. Masot-Llima and A. Garcia-Saez. Stabilizer tensor networks: universal quantum simulator on a basis of stabilizer states. *arXiv/quant-ph*, 2024. DOI: [10.48550/arXiv.2403.08724](https://doi.org/10.48550/arXiv.2403.08724).
- [28] J. R. McClean, N. C. Rubin, K. J. Sung, I. D. Kivlichan, X. Bonet-Monroig, Y. Cao, C. Dai, E. S. Fried, C. Gidney, B. Gimby, P. Gokhale, T. Häner, T. Hardikar, V. Havlíček, O. Higgott, C. Huang, J. Izaac, Z. Jiang, X. Liu, S. McArdle, M. Neeley, T. O’Brien, B. O’Gorman, I. Ozfidan, M. D. Radin, J. Romero, N. P. D. Sawaya, B. Senjean, K. Setia, S. Sim, D. S. Steiger, M. Steudtner, Q. Sun, W. Sun, D. Wang, F. Zhang, and R. Babbush. Openfermion: the electronic structure package for quantum computers. *Quantum Science and Technology*, 5(3):034014, jun 2020. DOI: [10.1088/2058-9565/ab8ebc](https://doi.org/10.1088/2058-9565/ab8ebc). URL <https://dx.doi.org/10.1088/2058-9565/ab8ebc>.
- [29] G. Nebe, E. M. Rains, and N. J. Sloane. *Real and Complex Clifford Groups*, pages 171–192. Springer Berlin Heidelberg, Berlin, Heidelberg, 2006. ISBN 978-3-540-30731-0. DOI: [10.1007/3-540-30731-1_6](https://doi.org/10.1007/3-540-30731-1_6).
- [30] A. Peres. Separability Criterion for Density Matrices. *Physical Review Letters*, 77(8):1413–1415, Aug. 1996. ISSN 1079-7114. DOI: [10.1103/physrevlett.77.1413](https://doi.org/10.1103/physrevlett.77.1413).
- [31] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O’Brien. A variational eigenvalue solver on a photonic quantum processor. *Nature Communications*, 5(1):4213, 2014. DOI: [10.1038/ncomms5213](https://doi.org/10.1038/ncomms5213).
- [32] C. Powers, L. Bassman Otfelie, D. Camps, and W. A. De Jong. Exploring finite temperature properties of materials with quantum computers. *Scientific Reports*, 13(1):1986, Feb. 2023. ISSN 2045-2322. DOI: [10.1038/s41598-023-28317-5](https://doi.org/10.1038/s41598-023-28317-5).
- [33] G. Roussy. An approximate block diagonalization for hermitian matrices as an improvement on successive Van Vleck transformations. *Molecular Physics*, 26(5):1085–1092, Nov. 1973. ISSN 0026-8976, 1362-3028. DOI: [10.1080/00268977300102311](https://doi.org/10.1080/00268977300102311).
- [34] U. Schollwöck. The density-matrix renormalization group. *Reviews of Modern Physics*, 77(1):259–315, Apr. 2005. ISSN 1539-0756. DOI: [10.1103/revmodphys.77.259](https://doi.org/10.1103/revmodphys.77.259).
- [35] J. R. Seddon and E. T. Campbell. Quantifying magic for multi-qubit operations. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 475(2227):20190251, July 2019. ISSN 1471-2946. DOI: [10.1098/rspa.2019.0251](https://doi.org/10.1098/rspa.2019.0251).
- [36] J. T. Seeley, M. J. Richard, and P. J. Love. The Bravyi-Kitaev transformation for quantum computation of electronic structure. *The Journal of Chemical Physics*, 137(22), Dec. 2012. ISSN 1089-7690. DOI: [10.1063/1.4768229](https://doi.org/10.1063/1.4768229).
- [37] K. Slagle, Y.-Z. You, and C. Xu. Disordered XYZ spin chain simulations using the spectrum bifurcation renormalization group. *Phys. Rev. B*, 94:014205, July 2016. DOI: [10.1103/PhysRevB.94.014205](https://doi.org/10.1103/PhysRevB.94.014205).
- [38] G. Subramanian Ravi, P. Gokhale, Y. Ding, W. M. Kirby, K. N. Smith, J. M. Baker, P. J. Love, H. Hoffmann, K. R. Brown, and F. T. Chong. CAFQA: A classical simulation bootstrap for variational quantum algorithms. *arXiv e-prints*, art. arXiv:2202.12924, Feb. 2022. DOI: [10.48550/arXiv.2202.12924](https://doi.org/10.48550/arXiv.2202.12924).
- [39] S.-N. Sun, M. Motta, R. N. Tazhigulov, A. T. Tan, G. K.-L. Chan, and A. J. Minnich. Quantum Computation of Finite-Temperature Static and Dynamical Properties of Spin Systems Using Quantum Imaginary Time Evolution. *PRX Quantum*, 2(1), Feb. 2021. ISSN 2691-3399. DOI: [10.1103/prxquantum.2.010317](https://doi.org/10.1103/prxquantum.2.010317).
- [40] Y. Takahashi and H. Umezawa. Thermo Field Dynamics. *International Journal of Modern Physics B*, 10(13n14):1755–1805, June 1996. ISSN 0217-9792, 1793-6578. DOI: [10.1142/S0217979296000817](https://doi.org/10.1142/S0217979296000817).
- [41] A. Tranter, S. Sofia, J. Seeley, M. Kaicher, J. McClean, R. Babbush, P. V. Coveney, F. Mintert, F. Wilhelm, and P. J. Love. The Bravyi-Kitaev transformation: Properties and applications. *International Journal of Quantum Chemistry*, 115(19):1431–1441, 2015. DOI: [10.1002/qua.24969](https://doi.org/10.1002/qua.24969).
- [42] V. Verteletskyi, T.-C. Yen, and A. F. Izmaylov. Measurement optimization in the variational quantum eigensolver using a minimum clique cover. *The Journal of Chemical Physics*, 152(12), Mar. 2020. DOI: [10.1063/1.5141458](https://doi.org/10.1063/1.5141458).

- [43] Z. Webb. The Clifford group forms a unitary 3-design. *Quantum Information and Computation*, 16(15&16):1379–1400, Nov. 2016. ISSN 15337146, 15337146. DOI: [10.26421/QIC16.15-16-8](https://doi.org/10.26421/QIC16.15-16-8).
- [44] S. R. White. Numerical canonical transformation approach to quantum many-body problems. *The Journal of Chemical Physics*, 117(16):7472–7482, Oct. 2002. ISSN 0021-9606, 1089-7690. DOI: [10.1063/1.1508370](https://doi.org/10.1063/1.1508370).
- [45] J. Wu and T. H. Hsieh. Variational Thermal Quantum Simulation via Thermofield Double States. *Physical Review Letters*, 123(22), Nov. 2019. DOI: [10.1103/physrevlett.123.220502](https://doi.org/10.1103/physrevlett.123.220502).
- [46] Y.-Z. You, X.-L. Qi, and C. Xu. Entanglement holographic mapping of many-body localized system by spectrum bifurcation renormalization group. *Phys. Rev. B*, 93:104205, Mar. 2016. DOI: [10.1103/PhysRevB.93.104205](https://doi.org/10.1103/PhysRevB.93.104205).
- [47] X. Yu, D. Pekker, and B. K. Clark. Bulk Geometry of the Many Body Localized Phase from Wilson-Wegner Flow. *arXiv e-prints*, art. arXiv:1909.11097, Sept. 2019. DOI: [10.48550/arXiv.1909.11097](https://doi.org/10.48550/arXiv.1909.11097).
- [48] X. Yuan, S. Endo, Q. Zhao, Y. Li, and S. C. Benjamin. Theory of variational quantum simulation. *Quantum*, 3:191, Oct. 2019. ISSN 2521-327X. DOI: [10.22331/q-2019-10-07-191](https://doi.org/10.22331/q-2019-10-07-191).
- [49] S.-X. Zhang, C.-Y. Hsieh, S. Zhang, and H. Yao. Differentiable quantum architecture search. *Quantum Science and Technology*, 7(4):045023, Aug. 2022. DOI: [10.1088/2058-9565/ac87cd](https://doi.org/10.1088/2058-9565/ac87cd).
- [50] S.-X. Zhang, J. Allcock, Z.-Q. Wan, S. Liu, J. Sun, H. Yu, X.-H. Yang, J. Qiu, Z. Ye, Y.-Q. Chen, C.-K. Lee, Y.-C. Zheng, S.-K. Jian, H. Yao, C.-Y. Hsieh, and S. Zhang. TensorCircuit: a Quantum Software Framework for the NISQ Era. *Quantum*, 7:912, Feb. 2023. ISSN 2521-327X. DOI: [10.22331/q-2023-02-02-912](https://doi.org/10.22331/q-2023-02-02-912).

A Classical simulation of Clifford-dominated quantum circuits

In this section, we review the basic idea of the generalized stabilizer representation and how to simulate non-Clifford gates and general quantum states with stabilizer frames. First, let us define *Pauli group* and *stabilizer group*:

Definition A.1 (Pauli group). *The Pauli group on n -qubits is the set of all tensor products of Pauli matrices, i.e.*

$$G_n = \{i^k \sigma^{i_1} \otimes \sigma^{i_2} \otimes \cdots \otimes \sigma^{i_n} | k, i_h \in \{0, 1, 2, 3\}\} \quad (4)$$

Definition A.2 (Stabilizer group). *Stabilizer group $\mathcal{S}$ is a subgroup of Pauli group satisfies the following properties:*

1. $\mathcal{S}$ is an abelian group. For $\forall s_i, s_j \in \mathcal{S}$, $[s_i, s_j] = 0$.
2. It doesn't contain negative identity, i.e. $-I \notin \mathcal{S}$.

The stabilizer group can always be represented by its generators. For example, a stabilizer group containing 2^r elements can be represented by r generators of the group: $\mathcal{S} = \langle s_1, s_2, \dots, s_r \rangle$. One should notice the generators are not unique. For example, $\mathcal{S} = \langle s_1, s_2, \dots, s_r \rangle = \langle s_1 s_2, s_2, \dots, s_r \rangle$. With the stabilizer group, one can define stabilizer states:

Theorem A.1 (Pure stabilizer state). *A full stabilizer (n generators for a n -qubit system) stabilizes a unique pure state $|\psi_S\rangle$*

One can check the density matrix can be represented as

$$\rho_S = \frac{1}{2^n} \prod_{i=1}^n (I + s_i) \quad (5)$$

Another concept that is useful for describing general quantum state is the *destabilizers*:

Definition A.3 (Destabilizer group). *A destabilizer group $\mathcal{D}$ is a subgroup of Pauli group associated with the stabilizer group $\mathcal{S} = \langle s_1, s_2, \dots, s_r \rangle$. It has the following properties:*

1. $\mathcal{D}$ has the same number of generators as $\mathcal{S}$: $\mathcal{D} = \langle d_1, d_2, \dots, d_r \rangle$.
2. For each generator d_i , it anti-commute with s_i , but commute with all the other generators $s_j, j \neq i$.

Definition A.4 (Full tableau). *A full stabilizer tableau $\mathcal{T} = (\mathcal{S}, \mathcal{D})$ is a pair of full stabilizer and destabilizer.*

Definition A.5 (Stabilizer basis). *The stabilizer basis with full tableau $\mathcal{T}$ is $\mathcal{B}(\mathcal{T}) = \{d |\psi_S\rangle | d \in \mathcal{D}\}$*

The stabilizer basis is an orthonormal and complete basis for n -qubit Hilbert space. Since the destabilizers are not unique, if $\mathcal{D}$ and $\mathcal{D}'$ are both destabilizers for $\mathcal{S}$, then $\mathcal{B}(\mathcal{S}, \mathcal{D})$ and $\mathcal{B}(\mathcal{S}, \mathcal{D}')$ are the same set of basis modulo phase difference. Since $\mathcal{B}(\mathcal{T})$ is a complete basis for Hilbert space, then any quantum state can be represented by the stabilizer basis. Therefore, we introduce the generalized stabilizer representation.

Definition A.6 (Generalized stabilizer representation). *Any pure quantum state can be represented with stabilizer basis. More particularly, the density matrix ρ of any pure quantum state can be written as*

$$\rho = \sum_{i=1}^{2^n} \sum_{j=1}^{2^n} \chi_{ij} d_i |\psi_S\rangle \langle \psi_S| d_j = \sum_{ij} \chi_{ij} d_i \rho_S d_j. \quad (6)$$

Therefore, we call the two-tuple $(\chi, \mathcal{B}(\mathcal{S}, \mathcal{D}))$ the generalized stabilizer representation.

For a general state, in the worst case, we need $\mathcal{O}(4^n)$ classical memory to store and simulate its dynamics. However, this formulation is particularly useful when this decomposition is sparse. For simplicity, we can use the zero norm of χ matrix to denote its complexity, i.e. $\Lambda(\chi) = \|\chi\|_0$. And the simulation complexity scales as $\mathcal{O}(\Lambda(\chi))$. Now we define the *Pauli channel* and see how to simulate general quantum states with the generalized stabilizer representation.

Definition A.7 (Pauli channel). *The Pauli channel of a n -qubit system can be written as*

$$\mathcal{E}(\rho) = \sum_{ij} \phi_{ij} P_i \rho P_j \quad (7)$$

where $P_i, P_j \in G_n$, and ϕ_{ij} are complex numbers. We use $\Lambda(\mathcal{E}) = \|\phi\|_0$ to denote the complexity of the Pauli channel.

A.1 General state evolution under Clifford gates

The evolution of the general state under a Clifford circuit C evolution can be written as

$$C\rho C^\dagger = C \left(\sum_{ij} \chi_{ij} d_i \rho_S d_j \right) C^\dagger = \sum_{ij} \chi_{ij} C d_j C^\dagger (C \rho_S C^\dagger) C d_j C^\dagger \quad (8)$$

Therefore, we only need to update the full tableau $\mathcal{T} = (\mathcal{S}, \mathcal{D})$, which can be done with $\mathcal{O}(n^2)$ time. Under Clifford gates, the general stabilizer frame will update as $(\chi, \mathcal{S}, \mathcal{D}) \rightarrow (\chi, \mathcal{S}', \mathcal{D}')$.

A.2 General state evolution under Pauli channels

First, we use the fact that any Pauli string can be decomposed with stabilizer and destabilizers, $P_m = \alpha_m d_{b_m} s_{c_m}$. An efficient implementation of this decomposition is in Algorithm 1. Suppose we want to simulate a general Pauli channel, $\mathcal{E}(\rho) = \sum_{m,n} \phi_{m,n} P_m \rho P_n^\dagger$. For each term in the Pauli channel:

$$\begin{aligned} \phi_{m,n} P_m \rho P_n^\dagger &= \sum_{ij} \phi_{m,n} P_m \chi_{ij} d_i \rho_S d_j P_n^\dagger = \sum_{ij} \phi_{m,n} \chi_{ij} P_m d_i \rho_S d_j P_n^\dagger \\ &= \sum_{ij} \phi_{m,n} \chi_{ij} \alpha_m \alpha_n^* d_{b_m} s_{c_m} d_i \rho_S d_j s_{c_n} d_{b_n} \\ &= \sum_{ij} \phi_{m,n} \chi_{ij} \alpha_m \alpha_n^* (-1)^{c_m \cdot i + c_n \cdot j} d_{b_m} d_i s_{c_m} \rho_S s_{c_n} d_j d_{b_n} \\ &= \sum_{ij} (\phi_{m,n} \chi_{ij} \alpha_m \alpha_n^* (-1)^{c_m \cdot i + c_n \cdot j}) d_{b_m+i} \rho_S d_{j+b_n} \\ &= \sum_{i'j'} \chi'_{i'j'} d_{i'} \rho_S d_{j'}. \end{aligned} \quad (9)$$

Therefore, for each term in Pauli channel, we need to update $(\chi, \mathcal{S}, \mathcal{D}) \rightarrow (\chi', \mathcal{S}, \mathcal{D})$. And we sum over all the terms in the Pauli channel, which contains $\Lambda(\phi)$ terms. Especially, T-gate can be written as the following channel:

$$T\rho T^\dagger = \left(\cos \frac{\pi}{8} I + i \sin \frac{\pi}{8} Z \right) \rho \left(\cos \frac{\pi}{8} I - i \sin \frac{\pi}{8} Z \right), \quad (10)$$

where $\Lambda(T) = 4$.

Algorithm 1 Pauli decomposition algorithm.

This algorithm efficiently decomposes a Pauli string P into a product of stabilizer and destabilizer elements $\mathbf{d}_b, \mathbf{s}_c$, and a phase factor α . The decomposition is used in the classical simulation of the Clifford+ k Rz circuits.

```

1: procedure DECOMPOSE( $P$ )  $\rightarrow \alpha, \mathbf{d}_b, \mathbf{s}_c$ 
2:   Initialize zero vector  $\mathbf{b} = (0, \dots, 0)$ , and  $\mathbf{c} = (0, \dots, 0)$ 
3:   phase = 0
4:    $P_t = I$  ▷ Binary vector encoding, no phase information.
5:   for  $i = 0, \dots, N - 1$  do ▷ Loop over stabilizer generators
6:     if  $\{P, \mathbf{s}_i\} = 0$  then
7:        $b_i = 1$ 
8:       phase  $\leftarrow$  phase - ipow( $P_t, \mathbf{d}_i$ )
9:        $P_t = P_t * \mathbf{d}_i$ 
10:    end if
11:  end for
12:  for  $i = 0, \dots, N - 1$  do ▷ Loop over destabilizer generators
13:    if  $\{P, \mathbf{d}_i\} = 0$  then
14:       $c_i = 1$ 
15:      phase  $\leftarrow$  phase - ipow( $P_t, \mathbf{s}_i$ )
16:       $P_t = P_t * \mathbf{s}_i$ 
17:    end if
18:  end for
19:  return phase,  $\mathbf{b}$ , and  $\mathbf{c}$ .
20: end procedure

```

B Optimization of Clifford+ k Rz circuits

We use a brickwork circuit ansatz (see Fig. 6) with L layers. One layer is composed of n Clifford gates, each of which acts on neighboring qubits. Although a common gate set for the Clifford group is taken to be the Hadamard, phase, and CNOT gates, we instead use the gate set composed of $e^{i\pi P/4}$, where $P \in \{I, X, Y, Z\}^{\otimes 2}$. We dope the circuit with k Rz gates at a fixed layer, but a variable site index (which is optimized over). To summarize, our free parameters belong to the domain $\mathbb{Z}_{16}^{n \times L} \times \mathbb{Z}_n^k \times \mathbb{R}^k$. The first term corresponds to 16 possible choices for the Clifford gates (of which there are $n \times L$). The second corresponds to the n possible choices for the site index of the k Rz gates, and the last is simply the argument for the k Rz gates.

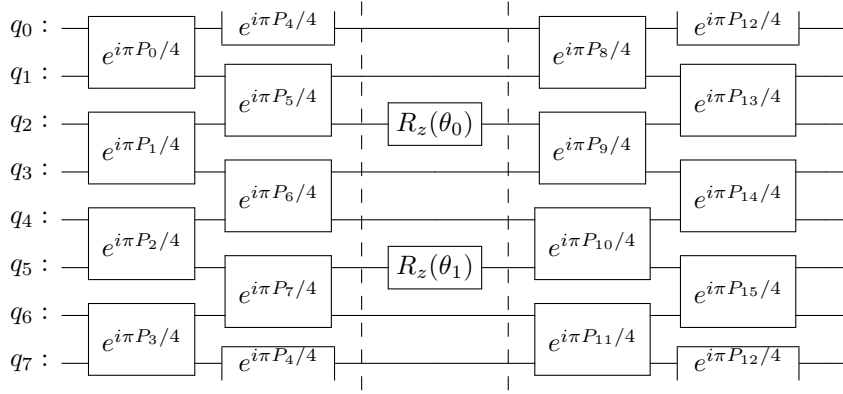


Figure 6: The brickwork circuit ansatz for $L = 2$ and $k = 2$.

To optimize our circuit, we treat the discrete and continuous parameters separately. We label $X \in \mathbb{Z}_{16}^{n \times L} \times \mathbb{Z}_n^k$ the discrete parameters and $\Theta \in \mathbb{R}^k$ the continuous ones. For any cost function $f(X, \Theta)$, we can define a ‘marginalized’ cost function by minimizing over Θ :

$$f(X) \equiv \min_{\Theta} f(X, \Theta). \quad (11)$$

This is a useful construction when it is easy to do this minimization; it turns out that for small k , this process is very easy. Since the cost is periodic in Θ , we only need to search the small volume $\Theta \in [0, 2\pi]^k$. This search is further simplified by the fact that $f(X, \Theta)$ is continuous in Θ , and does not have many local minima (see Fig. 7). By applying a simple gradient-based optimizer (e.g., gradient descent, conjugate gradient, etc.) to just a few (~ 10) randomly initialized choices of Θ , we can often find the global minimum within very few iterations. This is to say, the marginalized cost function $f(X)$ is easy to evaluate. With this, we optimize over the remaining parameters X with a simple greedy search. Each iteration, we pick one of the parameters in X to optimize over. For instance, if we choose to optimize P_1 (refer to Fig. 6), we would evaluate the cost with all other parameters held constant, varying P_1 over all 16 possibilities $\{I, X, Y, Z\}^{\otimes 2}$. Then, we simply replace P_1 with the choice that yielded the lowest cost $f(X)$. Finally, we do this discrete optimization in parallel for a large number ($n_{init} \sim 1000$) of randomly initialized parameters X , and simply take the best performing solution after a fixed number ($n_{iter} \sim 100$) iterations. Although we considered more advanced schemes for this discrete optimization, such as a simulated annealing process, we found that these schemes did not yield better results.

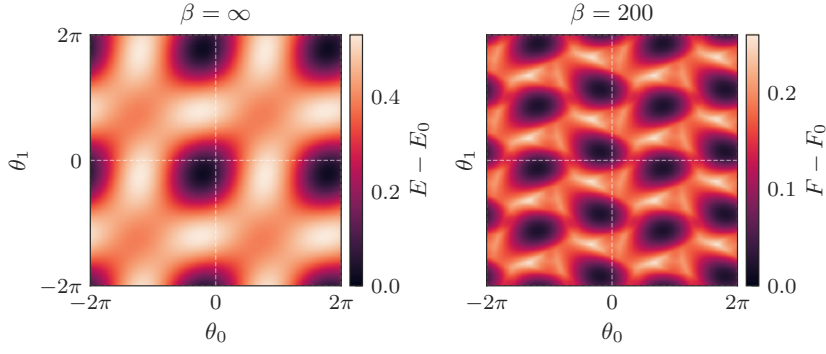


Figure 7: The landscape of $f(\Theta) \equiv f(X_{best}, \Theta)$ for LiH at a bond length $3.4\text{\AA}$, where X_{best} was the best performing solution for $k = 2$. On the left, we show the landscape for the ground state calculation, and on the right, we show the landscape at a finite temperature $\beta = 200$ ($T \approx 1500\text{K}$). Note that there are very few suboptimal local minima – applying gradient descent to a randomly initialized Θ will find the global minimum with high probability.

Algorithm 2 Optimizing Clifford + $k\text{Rz}$ Circuits

```

procedure OPTIMIZE( $f, L, k, n_{iter}$ )
   $X \leftarrow$  random initialization in  $\mathbb{Z}_{16}^{n \times L} \times \mathbb{Z}_n^k$ 
  for  $i = 1, \dots, n_{iter}$  do
     $j \leftarrow$  random index in  $1, \dots, n \cdot L + k$ 
    if  $j \leq n \cdot L$  then
       $params \leftarrow \mathbb{Z}_{16}$ 
    else
       $params \leftarrow \mathbb{Z}_n$ 
    end if
    for  $p \in params$  do
       $X[j] \leftarrow p$ 
       $cost[p] \leftarrow f(X)$ 
    end for
     $X[j] \leftarrow \text{argmin}_p cost[p]$ 
  end for
end procedure

```

C Symmetry of the Recovered Solutions

A noteworthy drawback of the Clifford + $k\text{Rz}$ ansatz (Fig. 6) is that it does not necessarily conserve any symmetries of the Hamiltonian. Despite this, we observe that the optimization process *prefers* to recover solutions that are both eigenstates of the number and spin operators, and furthermore, typically states that conserve symmetries in the Hamiltonian (e.g., total spin zero). Although observe that the latter does not hold for every solution (see Fig. 9), it holds in the vast majority of cases. Remarkably, we observe a greater tendency to preserve symmetries as more Rz gates are injected. However, we did not explicitly investigate the relationship between these preserved symmetries and the magic of the initial state. Understanding this connection could provide valuable insights into the role of symmetries and magic in our approach. Future work could explore this relationship in more detail, potentially leading to better ansatz designs and improved performance.

D Pulse optimization of programmable quantum simulators

The Rydberg Hamiltonian can be modeled as

$$H/\hbar = \sum_j \frac{\Omega_j(t)}{2} \left(e^{i\phi_j(t)} |g\rangle \langle r| + e^{-i\phi_j(t)} |r\rangle \langle g| \right) - \sum_j \Delta_j(t) \hat{n}_j + V_{ij} \hat{n}_i \hat{n}_j, \quad (12)$$

where $\Omega_j(t)$, $\phi_j(t)$, $\Delta_j(t)$ are Rabi frequency, laser phase, and local detuning of the driving field for each atom, and $|g\rangle$ is the ground state, $|r\rangle$ is the Rydberg state, $\hat{n}_j = |r\rangle_j \langle r|_j$ is the number operator. We use

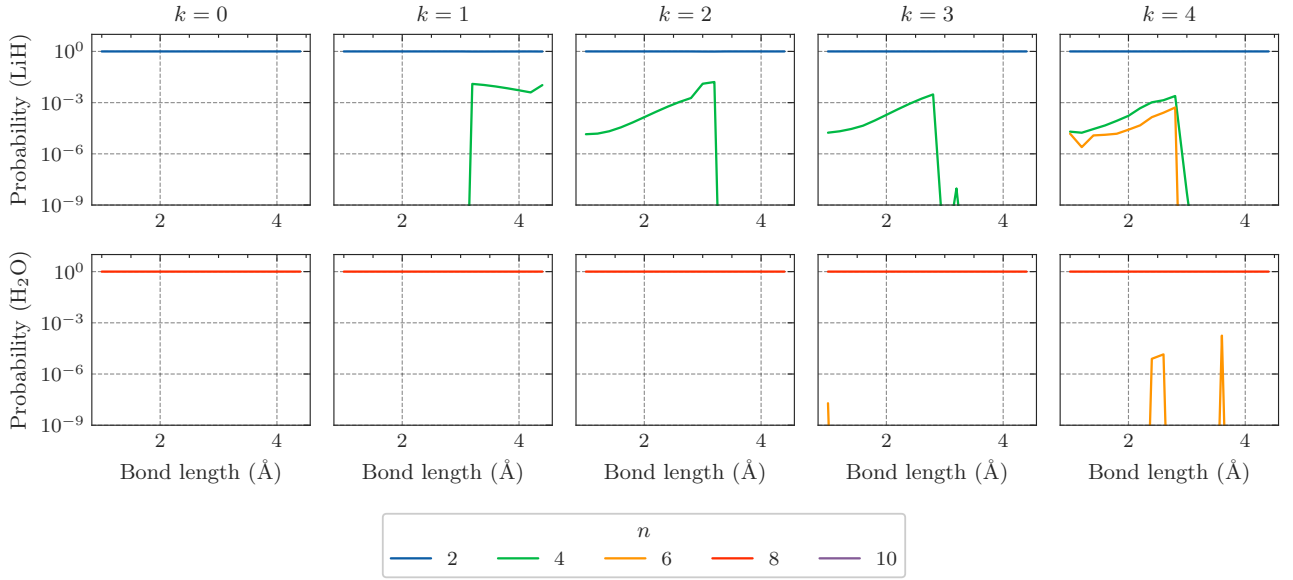


Figure 8: We analyze optimized Clifford + k Rz solution in the eigenbasis for the qubitized number operator N for both LiH (top) and H_2O (bottom), demonstrating that the recovered solution approximately conserves of N .

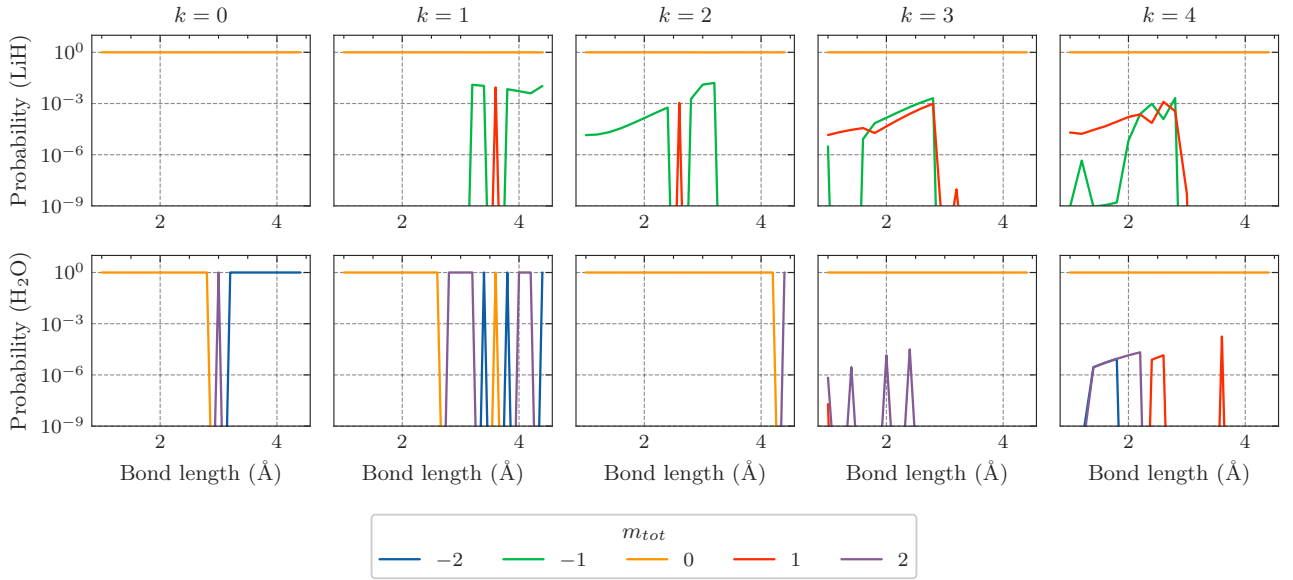


Figure 9: We analyze optimized Clifford + k Rz solution in the eigenbasis for the qubitized spin operator S for both LiH (top) and H_2O (bottom), demonstrating that the recovered solution, for most cases, approximately conserves of S .

$V_{ij} = C_6/|x_i - x_j|^6$ with $C_6 = 2\pi * 8.627 * 10^5 \text{ MHz } \mu m^6$ to describe the Rydberg interactions. In our optimization, we translate the Hamiltonian into Pauli basis, which gives

$$H/\hbar = \sum_j \left(\frac{\Omega_j}{2} \cos \phi_j X_j - \frac{\Omega_j}{2} \sin \phi_j Y_j \right) + \sum_j \left(\frac{\Delta_j}{2} - V_{ij} \right) Z_j + \sum_{i,j} \frac{V_{ij}}{4} Z_i Z_j + H_{\text{const}}, \quad (13)$$

where $H_{\text{const}} = \sum_{i,j} \frac{V_{ij}}{4} I - \sum_j \frac{\Delta_j}{2} I$. Then we optimize the pulse in the Pauli basis, such as $\frac{\Omega_j}{2} \cos \phi_j$. We parameterize each pulse as a piece-wise constant function with ten pieces, and initialize them randomly. The atom positions are fixed to be separated with $9.37 \mu m$. Since the interaction strength of the next nearest neighbors are 64 times smaller than the interaction strength of the nearest neighbors, we only consider the leading nearest neighbor interactions for our model.

E Finite temperature calculation

The variational principle extends beyond ground state calculations; indeed, it can be shown that the free energy satisfies a variational principle. That is, defining

$$F_\beta(\rho) \equiv \text{Tr}(H\rho) - S(\rho)/\beta, \quad (14)$$

where $S(\rho)$ is the von Neumann entropy, it can be shown that for all β ,

$$F_\beta(\rho) \geq F_\beta \left(\frac{e^{-\beta H}}{\text{Tr}(e^{-\beta H})} \right) \equiv F_0 \quad (15)$$

for all ρ . We take advantage of this variational principle to estimate the free energy at finite temperatures. The procedure is simple: we fix an orthogonal basis of stabilizer states $\{|\psi_k\rangle \mid k = 1, \dots, 2^n\}$ (typically derived from SBRG), so that the initial state is

$$\rho_0 \equiv \sum_k p_k |\psi_k\rangle\langle\psi_k|, \quad (16)$$

where the $p_k \in [0, 1]$ form a probability distribution (i.e., $\sum_k p_k = 1$). We construct a Clifford+KRz unitary U from Clifford + kT gates, and evolve each of the under this unitary. The final state is therefore

$$\rho = \sum_k p_k U |\psi_k\rangle\langle\psi_k| U^\dagger. \quad (17)$$

Defining $E_k \equiv \text{Tr}((H - \mu N)U |\psi_k\rangle\langle\psi_k| U^\dagger)$, we observe that

$$F_\beta(\rho) = \sum_k p_k \left(E_k - \frac{\log p_k}{\beta} \right). \quad (18)$$

The optimal p_k can be solved for in closed form (i.e., they do not need to be optimized numerically), yielding a free energy estimate

$$F_\beta(\rho) = -\frac{1}{\beta} \log \left(\sum_k \exp(-\beta E_k) \right). \quad (19)$$

Since calculating E_k as a function of the unitary U is efficient, we can then apply our optimization techniques over the unitary U to minimize $F_\beta(\rho)$.