GFlowNets for Hamiltonian decomposition in groups of compatible operators

Isaac L. Huidobro-Meezs
Department of Chemistry and Chemical Biology
McMaster University

Jun Dai
Mila & DIRO
Université de Montréal
Montréal, QC, Canada

Guillaume Rabusseau
Mila & DIRO
Université de Montréal
Montréal, QC, Canada

Rodrigo A. Vargas-Hernández*
Department of Chemistry and Chemical Biology
McMaster University
*vargashr@mcmaster.ca

Abstract

Quantum computing presents a promising alternative for the direct simulation of quantum systems with the potential to explore chemical problems beyond the capabilities of classical methods. However, current quantum algorithms are constrained by hardware limitations and the increased number of measurements required to achieve chemical accuracy. To address the measurement challenge, techniques for grouping commuting and anti-commuting terms, driven by heuristics, have been developed to reduce the number of measurements needed in quantum algorithms on near-term quantum devices. In this work, we propose a probabilistic framework using GFlowNets to group fully (FC) or qubit-wise commuting (QWC) terms within a given Hamiltonian. The significance of this approach is demonstrated by the reduced number of measurements for the found groupings; 51% and 67% reduction factors respectively for FC and QWC partitionings with respect to greedy coloring algorithms, highlighting the potential of GFlowNets for future applications in the measurement problem. Furthermore, the flexibility of our algorithm extends its applicability to other resource optimization problems in Hamiltonian simulation, such as circuit design.

1 Introduction

Quantum computing has gained considerable attention for its potential to solve the electronic structure problem (ESP), a fundamental challenge in computational chemistry and material science [1–5]. However, a major obstacle for quantum algorithms addressing the ESP, particularly on current noisy intermediate-scale quantum (NISQ) devices, is the measurement problem—the large number of measurements required to achieve chemical accuracy [6]. To mitigate this, two main approaches have been developed: (i) grouping commuting/anti-commuting terms and applying factorization/modification techniques to the Hamiltonian [7–15], and (ii) partial-tomography protocols inspired by shadow tomography [16–18]. A third approach integrates both frameworks [19]. While these methods have made substantial progress in tackling the measurement problem, their reliance on heuristics for optimization or the grouping per-se does not guarantee that the selected grouping minimizes the number of measurements.

In this work, we introduce the use of generative models for the measurement problem. Specifically, we utilize Generative Flow Networks (GFlowNets) to sample different valid groupings for a given Hamiltonian. GFlowNets are designed to sample from complex probability distributions by construct-

Machine Learning and the Physical Sciences Workshop, NeurIPS 2024.

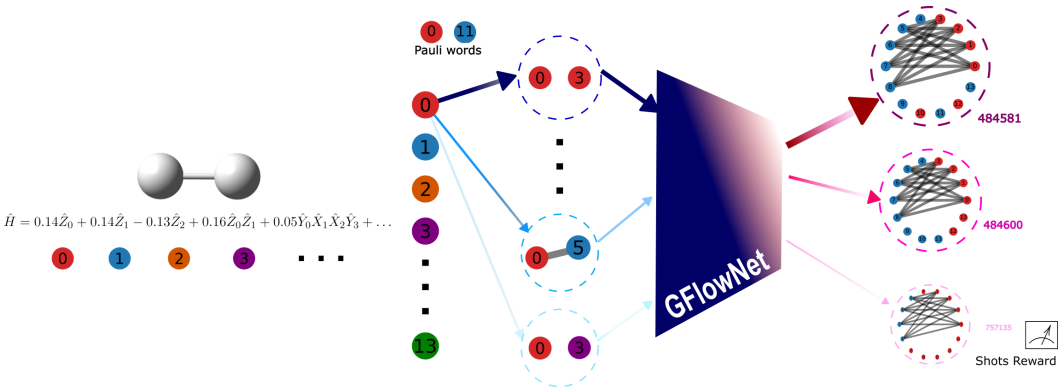

Figure 1: General scheme of GFlowNets for grouping Hamiltonian terms of H_2. Each node is a Pauli word from the Hamiltonian. The relative width of the arrows shows the probability of sampling a given state. The number of shots estimated to achieve chemical accuracy for each terminal state is shown in the last column.

ing flow-based policies to generate structured objects, such as graphs or sequences, through a series of intermediate steps [20, 21]. Unlike previous deterministic approaches, GFlowNets could learn a policy to sample the grouping configurations for a given Hamiltonian; see Fig. 1. We build on the relationship between the graph representation of optimal groupings through the minimum-clique cover (MCC) and the coloring of its complementary graph [11]. MCC is a known NP-hard problem with highly complex solution space which makes GFlowNets an attractive approach. GFlowNets-based approaches have also been introduced for graph combinatorial problems [22], like molecule generation with target properties [20]. Due to their stochastic nature, GFlowNets offer a diverse set of high-quality solutions leading to a broad exploration, making them a promising approach for grouping strategies in tackling the measurement problem.

2 Methods

Here, we briefly introduce the GFlowNets algorithm, Section 2.1, and grouping term decomposition used for Hamiltonians, Section 2.2. Finally, we give a description of our algorithm for the measurement problem.

2.1 GFlowNets

The key feature of GFlowNets is the incremental generation of each object x through a sequence of actions, which allows efficient sampling in complex, high-dimensional spaces [20–25]. For a discrete set $\mathcal{X}$, the probability $P(x)$ to sequentially build $\mathcal{X}$ is given by

$$P(x) = \frac{R(x)}{Z} = \frac{R(x)}{\sum_{x' \in \mathcal{X}} R\left(x'\right)}. \tag{1}$$

where $R(x)$ is the reward function associated with state x, and Z is the normalization constant. This formulation approximates the probability distribution $P(x)$ over the discrete set $\mathcal{X}$, proportional to the rewards of each state. Let $\mathcal{S}$ denote the set of states and $\mathcal{X} \subset \mathcal{S}$ denote the set of terminal states.

To sample from this probability distribution (Eq. 1), we define a trajectory $\boldsymbol{s} = (s_0, s_1, \ldots, s_n)$, where s_0 the initial state of any trajectory, progresses through intermediate states $(s_1 \cdots s_{n-1})$ and ends with a valid terminal state $(s_n \in \mathcal{X})$. We define $\mathcal{T}$ as the set of all possible trajectories, and $\mathcal{T}_x$ as the subset of trajectories that terminate at a specific terminal state $x \in \mathcal{X}$. We then introduce a flow function $F : \mathcal{T} \to \mathbb{R}^+$, associated with a normalized probability distribution over trajectories, $P(\boldsymbol{s}) = F(\boldsymbol{s})/Z, \quad Z = \sum_{\boldsymbol{s} \in \mathcal{T}} F(\boldsymbol{s})$. A flow function is considered valid if, for each terminal state x, the total flow into x matches its reward, $R(x) = \sum_{\boldsymbol{s} \in \mathcal{T}_x} F(\boldsymbol{s})$ [20, 21]. This implies that the probability $P(x)$, is proportional to its reward $P(x) = \sum_{\boldsymbol{s} \in \mathcal{T}_x} F(\boldsymbol{s})/Z \propto R(x)$.

If the trajectory is Markovian, $P(\boldsymbol{s})$ can be decomposed as $P\left(\boldsymbol{s}\right) = \prod_{i=1}^{n} P_F\left(s_i \mid s_{i-1}\right)$, where $P_F\left(s_i \mid s_{i-1}\right)$ is the forward policy corresponding to F. At each state, P_F can be computed

by $P_F(s' \mid s) = F(s \to s')/F(s)$, where $F(s) = \sum_{s'} F(s \to s')$. The flow matching constraint needs to be satisfied for all intermediate states [20, 21, 25], $\sum_{(s''\to s)\in\mathcal{A}} F(s'' \to s) = \sum_{(s\to s')\in\mathcal{A}} F(s \to s')$, where $\mathcal{A}$ denotes all possible transitions $(s \to s')$. Finally, we approximate the flow $F(s \to s')$ with a model $F_{\boldsymbol{\theta}}(s', s)$ with learnable parameter $\boldsymbol{\theta}$, which can be trained to minimize the loss $\mathcal{L}$ (Eq. 2) to satisfy the flow matching constraint.

$$\mathcal{L}(\boldsymbol{s}) = \sum_{s\in\boldsymbol{s}, s\neq s_0} \left(\log \frac{\sum_{(s''\to s)\in\mathcal{A}} F_{\boldsymbol{\theta}}(s'', s)}{\sum_{(s\to s')\in\mathcal{A}} F_{\boldsymbol{\theta}}(s, s')}\right)^2 \tag{2}$$

For the terminal state s_n, it is important to note that the denominator becomes the reward $R(s_n)$.

2.2 Hamiltonian terms grouping, measurements and GFlowNets implementation

Variational Quantum Eigensolver (VQE) techniques for the ESP are one of the most common applications of quantum algorithms in NISQ devices. These approaches rely on measuring the expectation value of the Hamiltonian. The molecular Hamiltonian is defined by the identity of the atoms, interatomic distances, and the basis set chosen for the calculation. For its measurement, a fermion-to-qubit mapping is required, e.g., Jordan-Wigner (JW) and Bravyi-Kitaev [26, 27]. We will focus on the JW mapping which takes the occupation number of orbitals and maps them directly to whether the qubit state is in $|0\rangle$ or $|1\rangle$ while maintaining the anticommutation relationships from fermionic operators by introducing a phase with $\hat{Z}$ qubit operators, leading to a qubit Hamiltonian in terms of Pauli products,

$$\hat{H}_q = \sum_k^{N_P} c_k \hat{P}_k, \quad \hat{P}_n = \bigotimes_{n=1}^{N_q} \sigma_n, \tag{3}$$

with every Pauli product $\hat{P}_n$ being a tensor product of Pauli operators and identities for the corresponding qubit $\sigma_n \in \{\hat{x}_n, \hat{y}_n, \hat{z}_n, \hat{1}_n\}$. N_P is the number of Pauli words in $\hat{H}_q$. Specific details on this mapping can be found in [27, 28].

To find molecular energies, VQE performs an iterative optimization of a parameterized wavefunction, $|\psi_\theta\rangle$, as $E_\theta = \min_\theta \langle\psi_\theta|\hat{H}_q|\psi_\theta\rangle$. Since $\hat{H}_q$ contains terms that do not commute with each other, a single shot measuring of the operator is not possible. For this reason, we are required to partition the Hamiltonian into compatible fragments [11] and measure each of them separately

$$\hat{H}_q = \sum_{\alpha=1}^{N_f} \hat{H}_\alpha; \quad E_\theta = \sum_{\alpha=1}^{N_f} \langle\psi_\theta|\hat{H}_\alpha|\psi_\theta\rangle. \tag{4}$$

Several grouping schemes exist, this work focuses on the fully commuting (FC) and qubit-wise commuting (QWC) groupings for the VQE problem. For FC, the requirement is that the Pauli products within a group commute with each other, $[\hat{P}_i, \hat{P}_j] = 0$. QWC is a more strict condition since, for a pair of Pauli products, every single-qubit operator needs to commute $[\hat{P}_i, \hat{P}_j]_{QWC} = 0$, e.g. the Pauli products $\hat{X}_1\hat{X}_2$ and $\hat{Y}_1\hat{Y}_2$ commute but not qubit-wise commute. Both of these partitionings can be achieved by constructing the corresponding commutativity graph and identifying the MCC. Each clique in the graph represents a distinct group. This is equivalent to coloring the complementary graph as is well-known in graph theory and as is often employed in quantum computing. Even though efficient algorithms are known for the coloring problem [29], these do not guarantee that the resulting partitioning would be the best performing when implemented in quantum devices, an issue that we aim to tackle with GFlowNets.

To sample optimal groupings using GFlowNets, we need to estimate the number of measurements (M_{est}) required to achieve a certain accuracy (ε). This quantity can be found as, $M = \frac{1}{\varepsilon^2}\left(\sum_\alpha^{N_f} \sqrt{\text{Var}(\hat{H}_\alpha)}\right)^2$, [30, 10] where $\text{Var}(\hat{H}_\alpha)$ is the variance of the operator which requires an approximation of the wavefunction or, ideally, the full-CI wavefunction to get the variances of each fragment. We approximate the variances under the assumptions, $\text{Cov}(\hat{P}_j, \hat{P}_k) = 0$ and $\text{Var}(\hat{P}_j) \leq 1$, yielding the expression [31]

$$M_{est} \approx \frac{1}{\varepsilon^2}\left(\sum_i \sqrt{\sum_j c_{ij}^2}\right). \tag{5}$$

While it would be ideal to employ the exact wavefunction instead, this approximation is used in the reward function on the GFlowNets training stage due to its easiness of calculation.

The reward employed is based on 1) the number of groupings (colors) on the graph and 2) the number of measurements required to achieve chemical accuracy of 1 kcal/mol or 1.6 mHa. Explicitly, we defined the reward function as,

$$R(x) = (N_P - \text{max\_color}(x)) + \frac{\lambda_0}{M_{est}(x)} \tag{6}$$

where max_color is the maximum color for the generated graph which gives the number of groups, the lower this number is, the fewer circuits we need to run. λ_0 is a scaling factor, set to 10^6 to account for the order of magnitude of the number of measurements.

Now we proceed to describe our sampling algorithm which is schematized in Fig. 1. First, a set of Pauli products from the molecular qubit Hamiltonian is fed to GFlowNets with the commutativity graph. The terms are assigned to a vertex, and edges are given by the commutation/QW-commutation relations to generate the complementary graph. At every training iteration, colors are assigned sequentially using a categorical distribution with probabilities dictated by Eq. 1; a masking function is employed to limit the optimization space. The mask takes as the upper limit the number of groups from a classical heuristics-based algorithm, namely greedy coloring with a random sequential strategy, and it can be increased as the user requires. Each added color to the resulting graph represents a different state, x, through the action sequence for GFlowNets whose transitions come from adding a new colored term to the list. Once a terminal state is reached, the validity of the graph coloring is assessed, $R(x)$ is calculated, and the color probabilities per vertex given by GFlowNets are updated. Finally, the total number of colors in a terminal state is the number of groups to be measured and a new iteration begins.

3 Results

In this section, we present the results of applying GFlowNets to six molecular systems: H_2, H_4, LiH, BH, BeH_2, and N_2. For all systems, we used an inter-atomic distance of 1 Å, the STO-3G basis set [32], and the Jordan-Wigner mapping. Except for N_2, where only 100 training iterations were used, 1,000 iterations were found sufficient for the other systems. The number of Pauli terms (N_P) are reported in Table 1. For $F_{\boldsymbol{\theta}}(s', s)$, we employed a two-layer MLP with 512 neurons and the $\tanh$ activation function. For all simulations, we used an NVIDIA GeForce RTX 4080 SUPER GPU, Torch, the Adam optimizer [33] with a learning rate of 3×10^{-4} and gradient accumulation every ten steps.

Table 1: Comparison of M_{est} (in millions) required to achieve chemical accuracy for the different available methods. Number of colors reported in parentheses. Full is the M_{est} without grouping. We define the Reduction factor as the ratio between GFlowNets and the best-performing greedy NetworkX-method.

	GFlowNets (ours)		NX-lf		NX-dsat		Full	Reduction factor	
System [N_P]	FC	QWC	FC	QWC	FC	QWC	None	FC	QWC
H_2 [14]	**0.485 (2)**	**0.746 (5)**	0.757 (2)	**0.746 (5)**	0.757 (2)	**0.746 (5)**	2.56	0.640	1.000
H_4 [184]	**1.50 (19)**	**1.13 (71)**	2.87 (11)	1.29 (68)	3.79 (9)	1.27 (67)	76.7	0.524	0.877
LiH [275]	**2.10 (23)**	**4.10 (72)**	4.13 (21)	6.77 (64)	8.32 (12)	6.54 (63)	121	0.508	0.626
BH [275]	**3.23 (22)**	**4.92 (72)**	5.99 (21)	8.21 (64)	10.5 (12)	7.78 (63)	610	0.539	0.632
BeH_2 [326]	**5.71 (23)**	**9.78 (101)**	14.0 (16)	21.0 (100)	22.4 (9)	21.1 (99)	405	0.407	0.465
N_2^*[824]	**22.8 (49)**	**48.0 (340)**	51.2 (23)	111 (318)	125 (16)	110 (314)	12705	0.446	0.433

Table 1 compares the estimated required number of measurements for the groupings obtained by GFlowNets and the ones from greedy coloring in NetworkX following the Largest first (NX-lf), and saturation largest first (NX-dsat) strategies. We found that to achieve chemical accuracy, GFlowNets on average require 51% for FC and 67% for QWC, of measurements compared to heuristic methods. These promising results encourage further developments for incorporation of state-of-the-art techniques, such as k-commutativity groupings [34] and ghost Pauli products [35] to the GFlowNets workflow to further improve it.

In Fig. 2a, we show the 2D histogram of the maximum color found and the estimated number of measurements M_{est} for LiH and BeH_2 in the FC grouping. Let's reiterate that the maximum color

is the number of compatible groups. It is appreciable how the sampling algorithm veers towards configurations with a balance between the number of groups and M_{est}. This effect is of course dependent on λ_0 and optimization of this parameter is one of the perspectives of the present work. Decreasing this parameter, samples colorings with a lower number of groups while the opposite biases the algorithm towards lower M_{est} regardless of the group number. For future work, we plan to extend our algorithm to incorporate classically efficient wavefunctions, such as Coupled-Cluster, to improve variance estimation [10] thereby providing a more accurate reflection of our algorithm's practical performance.

Some of the sampled graphs generated by GFlowNets for H_2 and BeH_2 Hamiltonians are shown in Fig. 2b- 2c. As we observe, H_2 graphs I and II are equivalent colorings, yielding the same number of measurements. The set of equivalent graphs is responsible for the increased number of sampled graphs shown in the histograms for a particular M_{est}. On the other hand, the BeH_2 graphs show how inefficient groupings can increase dramatically the required number of measurements, producing a sixfold increase in M_{est}.

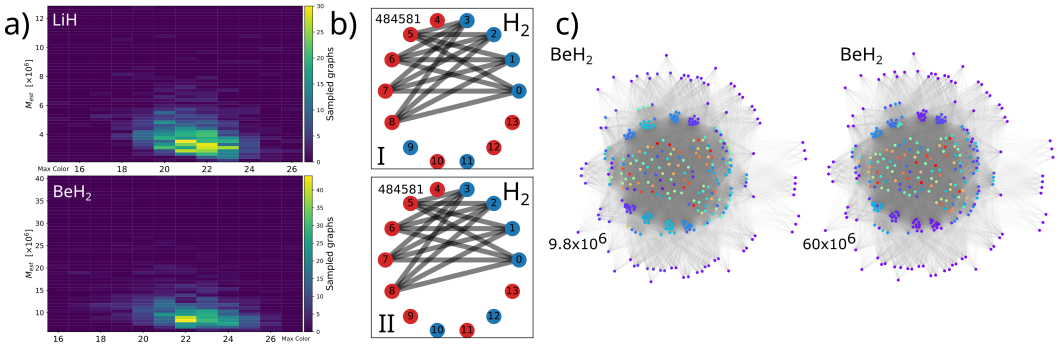

Figure 2: a) 2D histogram of number of groupings and estimated measurements (M_{est}) for LiH and BeH_2 with FC grouping. b) Equivalent graphs sampled for H_2 in FC grouping. c) Sampled graphs for BeH_2 in QWC grouping. M_{est} shown for each graph.

It is important to note that incorporating the number of measurements into the reward function, allows the algorithm to sample partitionings that may not minimize the number of cliques but result in lower measurement estimates. This flexibility in the choice of GFlowNets' reward function opens research opportunities for its implementation towards other resource optimization problems in quantum computing. Our results provide early evidence that GFlowNets, coupled with more competitive techniques, could become a promising alternative for grouping compatible operators.

4 Summary

In this work, we introduced GFlowNets for grouping compatible Hamiltonian terms in VQEs. Our results demonstrate that GFlowNets can achieve more efficient groupings compared to the deterministic methods available in NetworkX, as used in Pennylane [36]. This research serves as a promising foundation for leveraging GFlowNets to address the measurement problem. Given the flexibility of GFlowNets' reward function, potential extensions to our framework include exploring k-commutativity conditions [34], incorporating "ghost" Pauli products [35] or utilizing the Majorana tensor representation [37] for more general decompositions that further reduce measurement requirements. These are all directions we plan to investigate in the near future.

Acknowledgments

The authors thank Ignacio Loaiza for fruitful discussions. This research was partly enabled by support from the Digital Research Alliance of Canada and NSERC Discovery Grant No. RGPIN-2024-06594. GR acknowledges the support of CIFAR through the CCAI chair program.

References

[1] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P. D. Sawaya, Sukin Sim, Libor Veis and Alán Aspuru-Guzik. Quantum chemistry in the age of quantum computing. *Chemical Reviews*, 119(19):10856–10915, 2019. doi: 10.1021/acs.chemrev.8b00803. URL https://doi.org/10.1021/acs.chemrev.8b00803. PMID: 31469277.

[2] Bela Bauer, Sergey Bravyi, Mario Motta and Garnet Kin-Lic Chan. Quantum algorithms for quantum chemistry and quantum materials science. *Chemical Reviews*, 120(22):12685–12717, 2020. doi: 10.1021/acs.chemrev.9b00829. URL https://doi.org/10.1021/acs.chemrev.9b00829. PMID: 33090772.

[3] Sam McArdle, Suguru Endo, Alán Aspuru-Guzik, Simon C. Benjamin and Xiao Yuan. Quantum computational chemistry. *Rev. Mod. Phys.*, 92:015003, Mar 2020. doi: 10.1103/RevModPhys.92.015003. URL https://link.aps.org/doi/10.1103/RevModPhys.92.015003.

[4] Kishor Bharti, Alba Cervera-Lierta, Thi Ha Kyaw, Tobias Haug, Sumner Alperin-Lea, Abhinav Anand, Matthias Degroote, Hermanni Heimonen, Jakob S. Kottmann, Tim Menke, Wai-Keong Mok, Sukin Sim, Leong-Chuan Kwek and Alán Aspuru-Guzik. Noisy intermediate-scale quantum algorithms. *Rev. Mod. Phys.*, 94:015004, Feb 2022. doi: 10.1103/RevModPhys.94.015004. URL https://link.aps.org/doi/10.1103/RevModPhys.94.015004.

[5] Jules Tilly, Hongxiang Chen, Shuxiang Cao, Dario Picozzi, Kanav Setia, Ying Li, Edward Grant, Leonard Wossnig, Ivan Rungger, George H. Booth and Jonathan Tennyson. The variational quantum eigensolver: A review of methods and best practices. *Physics Reports*, 986:1–128, 2022. ISSN 0370-1573. doi: https://doi.org/10.1016/j.physrep.2022.08.003. URL https://www.sciencedirect.com/science/article/pii/S0370157322003118. The Variational Quantum Eigensolver: a review of methods and best practices.

[6] Yuan Su, Dominic W Berry, Nathan Wiebe, Nicholas Rubin and Ryan Babbush. Fault-tolerant quantum simulations of chemistry in first quantization. *PRX Quantum*, 2(4):040332, 2021.

[7] William J Huggins, Jarrod R McClean, Nicholas C Rubin, Zhang Jiang, Nathan Wiebe, K Birgitta Whaley and Ryan Babbush. Efficient and noise resilient measurements for quantum chemistry on near-term quantum computers. *npj Quantum Information*, 7(1):23, 2021.

[8] Oumarou Oumarou, Maximilian Scheurer, Robert M. Parrish, Edward G. Hohenstein and Christian Gogolin. Accelerating Quantum Computations of Chemistry Through Regularized Compressed Double Factorization. *Quantum*, 8:1371, June 2024. ISSN 2521-327X. doi: 10.22331/q-2024-06-13-1371. URL https://doi.org/10.22331/q-2024-06-13-1371.

[9] Artur F Izmaylov, Tzu-Ching Yen, Robert A Lang and Vladyslav Verteletskyi. Unitary partitioning approach to the measurement problem in the variational quantum eigensolver method. *Journal of Chemical Theory and Computation*, 16(1):190–195, 2019.

[10] Tzu-Ching Yen, Aadithya Ganeshram and Artur F. Izmaylov. Deterministic improvements of quantum measurements with grouping of compatible operators, non-local transformations, and covariance estimates. *npj Quantum Information*, 9(1):14, Feb 2023. ISSN 2056-6387. doi: 10.1038/s41534-023-00683-y. URL https://doi.org/10.1038/s41534-023-00683-y.

[11] Vladyslav Verteletskyi, Tzu-Ching Yen and Artur F. Izmaylov. Measurement optimization in the variational quantum eigensolver using a minimum clique cover. *The Journal of Chemical Physics*, 152(12), 03 2020. ISSN 0021-9606. 124114.

[12] Tzu-Ching Yen and Artur F. Izmaylov. Cartan subalgebra approach to efficient measurements of quantum observables. *PRX Quantum*, 2:040320, Oct 2021.

[13] Ignacio Loaiza and Artur F. Izmaylov. Block-invariant symmetry shift: Preprocessing technique for second-quantized hamiltonians to improve their decompositions to linear combination of unitaries. *Journal of Chemical Theory and Computation*, 19(22):8201–8209, Nov 2023. ISSN 1549-9618. doi: 10.1021/acs.jctc.3c00912. URL https://doi.org/10.1021/acs.jctc.3c00912.

[14] Ignacio Loaiza, Alireza Marefat Khah, Nathan Wiebe and Artur F Izmaylov. Reducing molecular electronic hamiltonian simulation cost for linear combination of unitaries approaches. *Quantum Science and Technology*, 8(3):035019, may 2023. doi: 10.1088/2058-9565/acd577. URL https://dx.doi.org/10.1088/2058-9565/acd577.

[15] Seonghoon Choi, Ignacio Loaiza and Artur F. Izmaylov. Fluid fermionic fragments for optimizing quantum measurements of electronic Hamiltonians in the variational quantum eigensolver. *Quantum*, 7:889, January 2023. ISSN 2521-327X. doi: 10.22331/q-2023-01-03-889. URL https://doi.org/10.22331/q-2023-01-03-889.

[16] Hsin-Yuan Huang, Richard Kueng and John Preskill. Predicting many properties of a quantum system from very few measurements. *Nature Physics*, 16(10):1050–1057, 2020.

[17] Charles Hadfield, Sergey Bravyi, Rudy Raymond and Antonio Mezzacapo. Measurements of quantum hamiltonians with locally-biased classical shadows. *Communications in Mathematical Physics*, 391(3):951–967, 2022.

[18] Guillermo García-Pérez, Matteo AC Rossi, Boris Sokolov, Francesco Tacchino, Panagiotis Kl Barkoutsos, Guglielmo Mazzola, Ivano Tavernelli and Sabrina Maniscalco. Learning to measure: Adaptive informationally complete generalized measurements for quantum algorithms. *PRX Quantum*, 2(4):040342, 2021.

[19] Zi-Jian Zhang, Kouhei Nakaji, Matthew Choi and Alán Aspuru-Guzik. A composite measurement scheme for efficient quantum observable estimation. *arXiv preprint arXiv:2305.02439*, 2023.

[20] Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup and Yoshua Bengio. Flow network based generative models for non-iterative diverse candidate generation. *Advances in Neural Information Processing Systems*, 34:27381–27394, 2021.

[21] Yoshua Bengio, Salem Lahlou, Tristan Deleu, Edward J Hu, Mo Tiwari and Emmanuel Bengio. Gflownet foundations. *The Journal of Machine Learning Research*, 24(1):10006–10060, 2023.

[22] Dinghuai Zhang, Hanjun Dai, Nikolay Malkin, Aaron C Courville, Yoshua Bengio and Ling Pan. Let the flows tell: Solving graph combinatorial problems with gflownets. *Advances in neural information processing systems*, 36:11952–11969, 2023.

[23] Moksh Jain, Tristan Deleu, Jason Hartford, Cheng-Hao Liu, Alex Hernandez-Garcia and Yoshua Bengio. Gflownets for ai-driven scientific discovery. *Digital Discovery*, 2(3):557–577, 2023.

[24] Dinghuai Zhang, Nikolay Malkin, Zhen Liu, Alexandra Volokhova, Aaron Courville and Yoshua Bengio. Generative flow networks for discrete probabilistic modeling. In *International Conference on Machine Learning*, pages 26412–26428. PMLR, 2022.

[25] Nikolay Malkin, Moksh Jain, Emmanuel Bengio, Chen Sun and Yoshua Bengio. Trajectory balance: Improved credit assignment in gflownets. *Advances in Neural Information Processing Systems*, 35:5955–5967, 2022.

[26] Sergey B. Bravyi and Alexei Yu. Kitaev. Fermionic quantum computation. *Annals of Physics*, 298(1):210–226, 2002. ISSN 0003-4916. doi: https://doi.org/10.1006/aphy.2002.6254. URL https://www.sciencedirect.com/science/article/pii/S0003491602962548.

[27] Pascual Jordan and Eugene P. Wigner. About the Pauli exclusion principle. *Z. Phys.*, 47: 631–651, 1928. doi: 10.1007/BF01331938.

[28] Andrew Tranter, Peter J. Love, Florian Mintert and Peter V. Coveney. A comparison of the bravyi–kitaev and jordan–wigner transformations for the quantum simulation of quantum chemistry. *Journal of Chemical Theory and Computation*, 14(11):5617–5630, Nov 2018. ISSN 1549-9618. doi: 10.1021/acs.jctc.8b00450. URL https://doi.org/10.1021/acs.jctc.8b00450.

[29] Thore Husfeldt. *Graph colouring algorithms*, page 277–303. Encyclopedia of Mathematics and its Applications. Cambridge University Press, 2015.

[30] Tzu-Ching Yen and Artur F. Izmaylov. Cartan subalgebra approach to efficient measurements of quantum observables. *PRX Quantum*, 2:040320, Oct 2021. doi: 10.1103/PRXQuantum.2.040320. URL https://link.aps.org/doi/10.1103/PRXQuantum.2.040320.

[31] Jérôme F. Gonthier, Maxwell D. Radin, Corneliu Buda, Eric J. Doskocil, Clena M. Abuan and Jhonathan Romero. Measurements as a roadblock to near-term practical quantum advantage in chemistry: Resource analysis. *Phys. Rev. Res.*, 4:033154, Aug 2022. doi: 10.1103/PhysRevResearch.4.033154. URL https://link.aps.org/doi/10.1103/PhysRevResearch.4.033154.

[32] W. J. Hehre, R. F. Stewart and J. A. Pople. Self-Consistent Molecular-Orbital Methods. I. Use of Gaussian Expansions of Slater-Type Atomic Orbitals. *The Journal of Chemical Physics*, 51(6):2657–2664, 09 1969. ISSN 0021-9606. doi: 10.1063/1.1672392. URL https://doi.org/10.1063/1.1672392.

[33] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[34] Ben DalFavero, Rahul Sarkar, Daan Camps, Nicolas Sawaya and Ryan LaRose. k-commutativity and measurement reduction for expectation values, 2024. URL https://arxiv.org/abs/2312.11840.

[35] Seonghoon Choi, Tzu-Ching Yen and Artur F. Izmaylov. Improving quantum measurements by introducing "ghost" pauli products. *Journal of Chemical Theory and Computation*, 18 (12):7394–7402, Dec 2022. ISSN 1549-9618. doi: 10.1021/acs.jctc.2c00837. URL https://doi.org/10.1021/acs.jctc.2c00837.

[36] Ville Bergholm, Josh Izaac, Maria Schuld, Christian Gogolin, Shahnawaz Ahmed, Vishnu Ajith, M. Sohaib Alam, Guillermo Alonso-Linaje, B. AkashNarayanan, Ali Asadi, Juan Miguel Arrazola, Utkarsh Azad, Sam Banning, Carsten Blank, Thomas R Bromley, Benjamin A. Cordier, Jack Ceroni, Alain Delgado, Olivia Di Matteo, Amintor Dusko, Tanya Garg, Diego Guala, Anthony Hayes, Ryan Hill, Aroosa Ijaz, Theodor Isacsson, David Ittah, Soran Jahangiri, Prateek Jain, Edward Jiang, Ankit Khandelwal, Korbinian Kottmann, Robert A. Lang, Christina Lee, Thomas Loke, Angus Lowe, Keri McKiernan, Johannes Jakob Meyer, J. A. Montañez-Barrera, Romain Moyard, Zeyue Niu, Lee James O'Riordan, Steven Oud, Ashish Panigrahi, Chae-Yeun Park, Daniel Polatajko, Nicolás Quesada, Chase Roberts, Nahum Sá, Isidor Schoch, Borun Shi, Shuli Shu, Sukin Sim, Arshpreet Singh, Ingrid Strandberg, Jay Soni, Antal Száva, Slimane Thabet, Rodrigo A. Vargas-Hernández, Trevor Vincent, Nicola Vitucci, Maurice Weber, David Wierichs, Roeland Wiersema, Moritz Willmann, Vincent Wong, Shaoming Zhang and Nathan Killoran. Pennylane: Automatic differentiation of hybrid quantum-classical computations, 2022. URL https://arxiv.org/abs/1811.04968.

[37] Ignacio Loaiza, Aritra Sankar Brahmachari and Artur F. Izmaylov. Majorana tensor decomposition: A unifying framework for decompositions of fermionic hamiltonians to linear combination of unitaries, 2024. URL https://arxiv.org/abs/2407.06571.