CAFQA: Clifford Ansatz For Quantum Accuracy

Gokul Subramanian Ravi*[1], Pranav Gokhale[2], Yi Ding[3], William M. Kirby[4], Kaitlin N. Smith[1], Jonathan M. Baker[1], Peter J. Love[4], Henry Hoffmann[1], Kenneth R. Brown[5], and Frederic T. Chong[1,2]

[1]University of Chicago, [2]Super.tech, [3]MIT, [4]Tufts University, [5]Duke University

Abstract—**Variational Quantum Algorithms (VQAs) rely upon the iterative optimization of a parameterized unitary circuit with respect to an objective function. Since quantum machines are noisy and expensive resources, it is imperative to choose a VQA's ansatz appropriately and its initial parameters to be as close to optimal as possible, as this would improve and accelerate the accurate convergence of the algorithm's execution on the quantum device.**

This work tackles the problem of finding initial ansatz parameters by proposing CAFQA, a Clifford ansatz for quantum accuracy. The CAFQA ansatz is a hardware efficient circuit built with only Clifford gates. In this ansatz, the initial parameters for the tunable gates are chosen by searching efficiently through the Clifford parameter space via classical simulation, and thereby producing a suitable stabilizer state. The stabilizer states produced are shown to always equal or outperform the traditional classical initialization approach (e.g., Hartree-Fock) of finding a suitable computational basis state, and often produce high accuracy estimations even prior to execution and exploration on a quantum device. Furthermore, the technique is suited to classical compute since a) Clifford-only quantum circuits can be exactly simulated classically in polynomial time and b) the discrete Clifford space, while scaling exponentially in the number of qubits, is searched efficiently via Bayesian Optimization.

For the Variational Quantum Eigensolver (VQE) task of ground state energy estimation of molecular systems up to 20 qubits, CAFQA's Clifford Ansatz achieves a mean accuracy of near 99% and is able to recover as much as 99.99% of the molecular correlation energy over Hartree-Fock initialization. Notably, the scalability of the approach allows for preliminary ground state energy estimation of the challenging chromium dimer (Cr_2) with an accuracy greater than that achieved by Hartree-Fock. CAFQA is also evaluated on optimization tasks, specifically MAXCUT problems of up to 18 qubits. With CAFQA's high accuracy initialization, the convergence of VQAs is accelerated by a factor of 2.5x.

In all, this work shows that stabilizer states are a highly accurate ansatz initialization for variational algorithms. Furthermore, it highlights the potential for quantum-inspired classical techniques as both an alternative as well as a supporting methodology for VQAs in the NISQ era and beyond.

I. CAFQA Overview

Quantum computing is a revolutionary computational model that leverages quantum mechanical phenomena for solving certain classically intractable problems. Quantum computers (QCs) evaluate quantum circuits or programs in a manner similar to a classical computer, but quantum information's
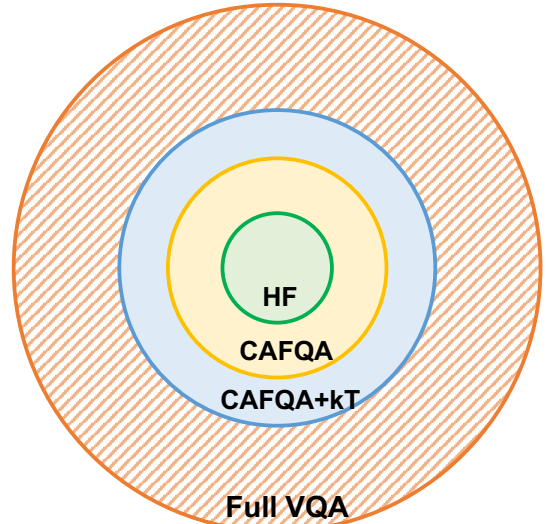
* Correspondence: gravi@uchicago.edu

Fig. 1: Figure illustrates the VQA state space explored by different techniques. Hartree-Fock (HF), in green, only explores the computational basis state space. CAFQA goes a step further by exploring the Clifford space (yellow) for higher accuracy. The advantage of exploring these spaces for VQA initialization is that they can be simulated classically in an efficient manner. Classical exploration can further be extended to incorporate a few T gates (blue) - in the future we look to extend CAFQA to CAFQA+T. Finally the rest of the state space is shown in orange - this space cannot be executed efficiently via classical compute and requires quantum exploration. VQA problems with potential quantum advantage are expected to span beyond the classically simulable state space - CAFQA provides a good ansatz initialization for quantum exploration.

ability to leverage superposition, interference, and entanglement is projected to give QCs a significant advantage in cryptography [73], chemistry [41], optimization [52], and machine learning [9].

In the ongoing Noisy Intermediate-Scale Quantum (NISQ) era, we expect to work with quantum machines which comprise hundreds to thousands of imperfect qubits [63]. NISQ era machines will be unable to execute large-scale quantum algorithms like Shor Factoring [73] and Grover Search [33], which would require error correction comprised of millions of qubits to create fault-tolerant quantum systems [58] wherein

1

many physical device qubits work in careful synchronization to create a single logical qubit that is used at the algorithm level.

Variational Quantum Algorithms: However, recently, variational quantum algorithms (VQAs) have been introduced that are expected to be a good match for NISQ machines where logical and physical qubits have a one-to-one mapping. This class of algorithms has a wide range of applications such as electronic energy estimations of molecules [61], MAXCUT approximation [52] and so on. The fundamental aspect of VQAs that make them suitable to today's quantum devices is that these algorithms adapt to the characteristics and noise profile of the quantum machine they are executing on. The quantum circuit for a VQA is parameterized by a list of angles which are optimized by a classical optimizer over the course of many iterations to maximize or minimize a specific target objective which is representative of the VQA problem, while navigating through the device's inherent noisy characteristics.

NISQ impediments to VQAs: Noise has prevented current quantum computers from surpassing the capabilities of classical computers in almost all applications, including VQAs. For instance, molecular chemistry often requires energy estimations which are extremely accurate, with estimation errors less than 1.6×10^{-3} Hartree needed to be useful towards analyses of chemical reactions. Unfortunately, the accuracy obtained on today's NISQ machines, even with the addition of sophisticated error mitigation techniques, is often considerably far from these accuracy requirements, especially as we scale to larger problem sizes [41], [65], [85]. Considering the promise of VQAs on NISQ machines, it is necessary to explore novel techniques to improve the quality of their execution on today's quantum devices.

Aiding VQAs in the NISQ Era: Since today's quantum machines are often too noisy for most real world VQA usecases, it is imperative to choose a VQA's parameterized circuit (ansatz) wisely and its initial parameters to be as close to optimal as possible, even prior to quantum exploration, as this would improve and accelerate the accurate convergence of the algorithm's execution on the noisy quantum device [48], [86]. While some applications benefit from physical domain knowledge that can inspire the choice of specific parameterized circuits and initial parameters (e.g., UCCSD [67]), these choices are less suited to execution on today's quantum devices due to their considerable quantum circuit depth. Suitable circuits for today's devices, referred to as "hardware efficient ansatz" [41], are often application agnostic and can benefit significantly from a wise choice of initial parameters, but these can be challenging to estimate - thus, novel approaches are required.

Classical support for VQAs with Cliffords: This work aids VQAs via classical computing support, specifically through simulation. In general, classical simulation of quantum tasks is not a scalable solution and is very limited in its capabilities, primarily only suited to today's trivial quantum problems, and is, in fact, the motivation for quantum machines to begin with.

An exception to the above is the classical simulation of the Clifford space. Circuits made up of only Clifford operations can be exactly simulated in polynomial time on a classical computer as put forth by the Gottesman-Knill theorem [30]. Of course, Clifford operations do not provide a universal set of quantum gates - hence, the stabilizer states produced by Clifford-only circuits are limited in how effectively they can explore the quantum space of a given problem such as those targeted by VQAs. However, exploring the Clifford space of the VQA problem through classical simulation can potentially find useful initial states by identifying appropriate VQA parameters in a suitable parameterizable circuit.

CAFQA: This work tackles the problem of finding initial ansatz parameters by proposing CAFQA, a Clifford ansatz for quantum accuracy. The CAFQA Clifford Ansatz is a hardware efficient parameterizable circuit that is parameterized with only Clifford gates. In this ansatz, the initial parameters for the tunable gates are chosen by searching efficiently through the Clifford parameter space via classical simulation, and thereby producing a suitable stabilizer state. The proposed approach is attractive for multiple reasons: ⓐ the Clifford-initialized ansatz produces stabilizer states that perform equal to or better than the traditional classical approach of finding a suitable computational basis state (e.g., Hartree-Fock [34]) because (i) it can explore a larger state space and (ii) the stabilizer can have explicit chemical relevance in some molecular systems; ⓑ the Clifford-only quantum circuits can be perfectly simulated in polynomial (quadratic or even linear) time on classical computers; ⓒ the discrete Clifford space, while scaling exponentially in number of qubits, is searched efficiently via Bayesian Optimization; ⓓ while the Clifford space is significantly smaller compared to the entire quantum space, the stabilizer state produced from the optimal Clifford parameters are able to achieve solutions of high accuracy even prior to execution/exploration on a quantum device; and finally ⓔ the selected ansatz can then be tuned over the entire parameter space on a quantum device, allowing for improved and accelerated accurate convergence to chemical accuracy on NISQ devices (and beyond).

Fig.1 provides a break down of the VQA parameter space and CAFQA's scope. An illustration of the CAFQA framework is shown in Fig.2.
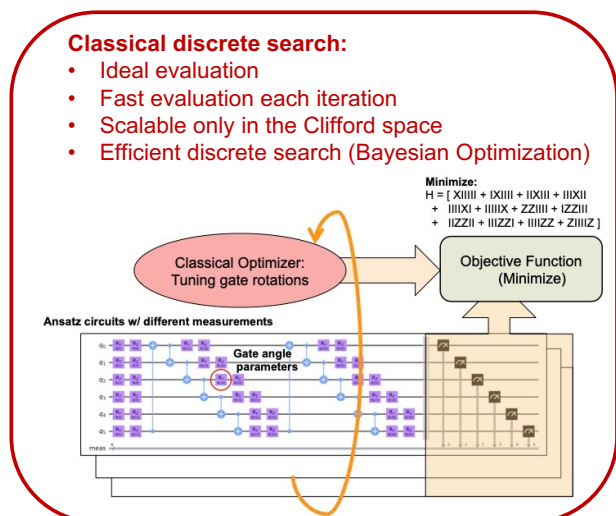
Contributions / Key results:

① CAFQA builds a Clifford ansatz by tuning a hardware-efficient ansatz in the Clifford space, thereby producing stabilizer states for high accuracy VQA initialization.
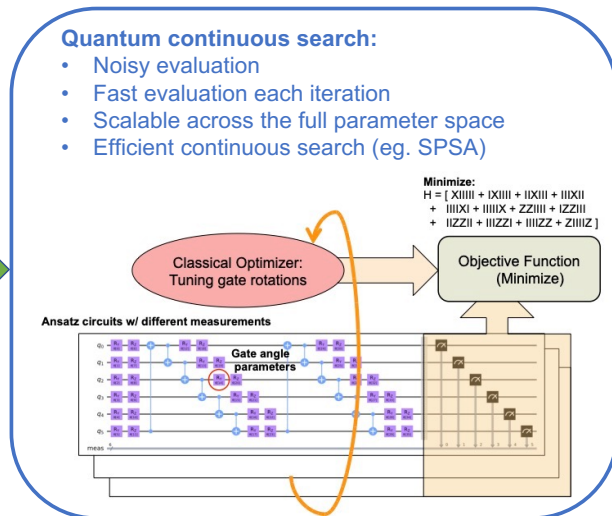
② CAFQA's benefits are significant because it is scalable (since Clifford circuits are simulable), it searches the search space efficiently (through Bayesian Optimization, using a random forest surrogate model and a greedy acquisition function) and its per-iteration measurements are ideal (since the simulation is noise free).

③ For the VQE task of ground state energy estimations of molecular systems up to 20 qubits, CAFQA's Clifford Ansatz is observed to achieve a mean accuracy of near 99% and mean absolute error of less than 0.05 Hartree (compared to

Fig. 2: The red box to the left shows CAFQA framework. CAFQA performs parameter tuning on the ansatz which is inspired by traditional quantum variational tuning. But CAFQA's tuning is suitable to classical compute since it restricts the search space to the Clifford parameter space alone (in a parameterized circuit which has also fixed components to also be Clifford, often the case with a hardware-efficient ansatz). The search space is discrete and is searched with Bayesian Optimization. While the search space is limited, its evaluation is classically efficient, scalable and ideal (i.e., noise-free). On the other hand, traditional quantum variational tuning, while being scalable and suited to the entire quantum space, is extremely noisy in the NISQ era. Once a suitable Clifford ansatz is found, traditional VQA tuning such as Simultaneous Perturbation Stochastic Approximation is performed on the ansatz (blue box to the right), leading to faster and more accurate convergence.

ideal evaluations). Further, the Clifford Ansatz is observed to recover as much as 99.99% of the molecular correlation energy over the standard Hartree-Fock initialization as bond lengths increase. CAFQA is also evaluated on optimization tasks, specifically MAXCUT problems of up to 18 qubits.

④ Quantum exploration post-CAFQA initialization can lead to faster and highly accurate convergence even on reasonably noisy quantum machines - we show 2.5x faster convergence compared to HF in this work.

⑤ Notably, the scalability of the approach allows for accurate ansatz initialization for ground state energy estimation of the challenging chromium dimer, often considered a benchmark for variational quantum advantage, with greater than Hartree-Fock accuracy.

Section II showcases a stabilizer initial state achieving high accuracy for LiH ground state energy estimation. Section III discusses the efficient discrete search over the Clifford Space. Sections IV and V provide evaluation results and Section VI concludes the paper.

More details are provided in the Appendices. Appendix A: more background and motivation content. Appendix B: CAFQA framework in a step-by-step manner. Appendix C: CAFQA benefits on a simple micro-benchmark. Appendix D: methodology for evaluation. Appendix E: discussion and future work. Appendix F: related work.

II. HIGH ACCURACY CAFQA STABILIZER STATES

In Fig.3 we break down the expectation value returned by the Clifford Ansatz in comparison to Hartree-Fock and ideal / exact minimum from noise free simulation. This is shown for the ground state energy estimation of the LiH molecule, represented by a 4 qubit Hamiltonian system, at a bond length of 4.8Å. The Y-axis shows the expectation value of each Pauli term and the X-axis lists the Pauli terms in the Hamiltonian.

① In the figure, the resultant expectation value of each Pauli term for the Hartree-Fock (HF) initialization is shown in blue. Since HF is also Clifford, as noted in Appendix B, all expectation values are +1 / -1 / 0. Note that for HF, all non-diagonal Pauli terms (any terms apart from the tensor products of I and Z) have an expectation value of zero i.e., the state being found is a computational basis state. Only calculating the expectation values for the diagonal terms leads to ignoring the correlation energy, which is known to lead to serious errors for some larger molecules.

② The expectation value of each Pauli term for the Clifford ansatz produced by CAFQA is shown in red. Again, note that all expectation values are +1 / -1 / 0. Moreover, note that for the Clifford ansatz, there are multiple non-diagonal Pauli terms which have an expectation of +1 / -1. The selection of non-diagonal terms is indicative of CAFQA producing a non-computational basis state, albeit a Clifford one. By doing so, it is able to capture some of the correlation energy that is
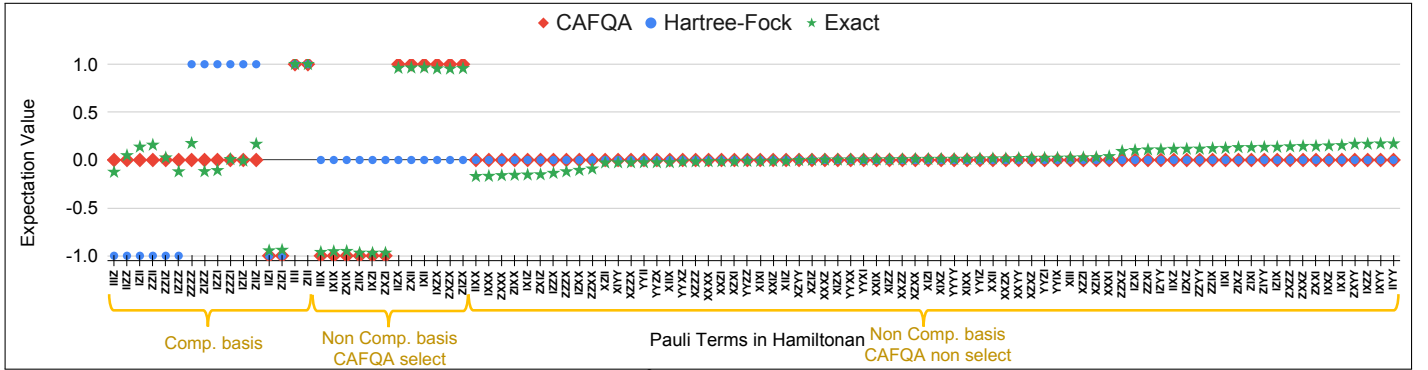
Fig. 3: LiH ground state energy at bond length of 4.8Å (3x the equilibrium length). Expectation value of each Pauli term, as obtained from different methods, is shown. While Hartree-Fock is only able to obtain non zero expectations (of +/- 1) for diagonal Pauli (computational basis) terms, the Clifford Ansatz enables expectations of +/- 1 for non-diagonal Pauli terms as well. Further, the similarity in expectation value between the Clifford Ansatz and the Ideal / Exact LiH tuning over the entire parameter space is evident. The Pauli terms along the X-axis are arranged as: i) Computational basis terms, ii) Non computational basis terms selected by CAFQA, and iii) remaining terms which are beyond the Clifford reach (sorted by Exact expectation value).
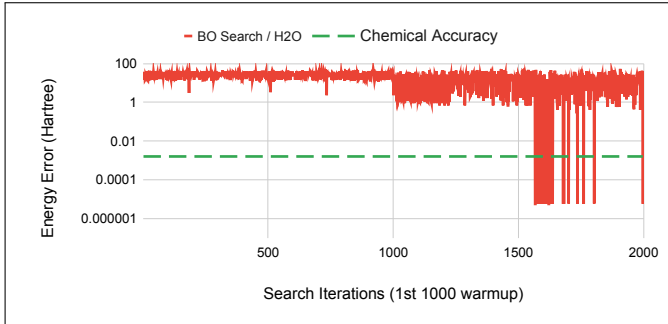

Fig. 4: H2O ground state energy estimation via CAFQA's Clifford ansatz discrete search at bond length of 4Å (4x the equilibrium length). The first 1000 iterations are a warm up period for the BO search algorithm. Beyond this, the search is able to achieve accuracy within chemical accuracy target in an additional 600 iterations. In this instance, post-CAFQA variational tuning on a quantum machine is not required.

contributed by the non-diagonal Pauli terms. This is important because it is indicative of the potential for high(er) estimation accuracy through CAFQA as problem complexity scales up.

③ The expectation value of each Pauli term for the exact minimum from the ideal noise-free simulation is shown in green. While expectation values range from -1 to 1, it is evident that the expectation values are close to those produced by the Clifford Ansatz, both for diagonal terms and the non-diagonal terms. This is a clear indicator of the effectiveness of CAFQA for ground state energy estimation for the LiH moecule (at the chosen bond length). This is confirmed later in Section IV-B which shows that CAFQA is able to achieve high accuracy in the range of 10^{-2} Hartree for LiH.

III. DISCRETE SEARCH OVER THE CLIFFORD SPACE

To efficiently search through the discrete Clifford parameter space, CAFQA requires a sample-efficient search technique to find the performant ansatz parameters as quickly as possible. Bayesian optimization is such a technique that actively and intelligently queries the most informative samples at each round to reduce the number of samples required [25]. Due to this merit, Bayesian optimization has been successfully applied to different domains in computer systems optimization such as compiler tuning [56], resource allocation [59], and configuration optimization [6], [19], [20], [69]. Bayesian optimization iteratively alternates between intelligent sampling and model updates. As such, it includes two components: a surrogate model and an acquisition function. The surrogate model tries to learn the unknown underlying function that maps the search parameters to the problem objective (e.g., ground state energy). The acquisition function is the search strategy that selects the next sample to query to update the surrogate model.

CAFQA searches through the Clifford space to identify optimal Clifford gates for the tunable circuit parameters. Each tunable parameter is able to take one of four rotational angles as was shown in Fig.10. This creates a discrete search space complexity of $O(4^{\#params})$ which scales exponentially in the number of parameters although it is considerably smaller than the entire quantum tuning space. To search efficiently through this discrete space, CAFQA utilizes Bayesian Optimization as introduced in Section III.

Since the Clifford parameter space is discrete, CAFQA chooses the random forest as the surrogate model as it is flexible enough to model the discrete space and scales well [56]. CAFQA uses a greedy acquisition function [19], [56] to select samples with the lowest predicted energy estimates from the surrogate model. Empirically, the combination of the random forest surrogate model and the greedy acquisition function gives highly accurate results, as illustrated in Section IV.

Details on the search algorithm implementation can be found in [56]. Here we limit ourselves to an illustrative example.

Fig.4 shows the discrete search employed by CAFQA to produce a Clifford Ansatz for H_2O ground state energy estimation at a bond length of 4Å. The first 1000 iterations are a warm-up period which involves randomly sampling and mapping the search space - a key component to BO. The search algorithm then uses these random samples to efficiently search through the parameter space. In the figure, note that as soon as the random sampling is complete, the search algorithm begins to find better expectation values compared to random. A potentially global minimum is found after an additional 600 search iterations. Notably, in this use case, the identified minimum is well within chemical accuracy requirements, obviating the need for traditional variational tuning on a quantum device. While 2000 iterations are shown here, the search can be constrained by a tuning budget or by the observed saturation of the obtained minimum.

IV. EVALUATION: CAFQA FOR MOLECULAR CHEMISTRY

In Fig.5 we showcase CAFQA for ground state energy estimation through VQE across 5 molecules' bond dissociation over different bond lengths. Evaluations are shown over multiple metrics. Each row represents different evaluation metrics for each molecule - the left column shows the ground state energy (Hartree), the center column shows the error in energy estimation, and the right column shows the correlation energy recovered. We compare CAFQA in green against exact evaluations in orange; and Hartree-Fock initialization (HF) in blue. More details on these molecules, metrics and comparisons are discussed in Appendix D.

A. H_2

First, we look at H_2 shown in Fig.5 (a)-(c). In (a) and (b), we see that the HF steadily deviates away from exact energies as bond lengths increase. This is unsurprising since HF is known to work best at / near equilibrium geometry, as discussed in Appendix A. CAFQA matches HF at low bond lengths but achieves lower energy estimates at higher bond lengths, thus being closer to exact estimates.

(a) also shows CAFQA energy estimates for the H_2^+ cation. The cation is at a higher energy state than its neutral counterpart - this is intuitive since H_2 does not naturally ionize. For a given molecular system, the Fock space combines the Hilbert spaces of molecular forms with all possible numbers of electrons and all electron-spin combinations [70]. When given the molecular Hamiltonian as above and solving VQE for the upper bound of the ground state energy of simple neutral molecules, the lack of specific electron preservation constraints does not cause inaccuracy because the energy of the neutral molecule is usually lower than that of the corresponding cations (otherwise the molecule would auto-ionize) and is also often lower than that of corresponding anions (unless the molecule has high positive electron affinity). On the other hand, when solving for energy estimates of cations and anions explicit enforcement of electron preservation constraints is

often required. For the H_2^+ cation, if the H_2 Hamiltonian was used as is with no electron constraints, CAFQA's produced energy estimates would be the same as the line in green since these are the lower bounds of the estimates for the given Hamiltonian, thus producing incorrect energy estimates for the cation. With CAFQA, we impose electron count constraints through the search objective function so that only electronic energies corresponding to a one electron system are considered thereby producing accurate estimates. Thus, CAFQA ensures that that any required constraints are maintained, along the lines of prior work [70].

(b) shows that CAFQA's error is always under 10^{-1} Hartree, and is able to achieve accuracy near chemical accuracy. On the other hand, as the bond lengths increase, HF error is 10^{-1} Hartree and greater. Thus, the benefits of CAFQA are clearly evident. Finally, (c) shows that CAFQA is able to recover up to 99.7% of the correlation energy as bond lengths increase.

Overall, CAFQA achieves better energy estimates than the best computational basis and is able to incorporate the expectation of non-diagonal Pauli terms. It thereby produces a non-computational basis state which is uncommon in other classical approaches - an example of this was illustrated earlier for LiH in Fig.3. Further, the produced stabilizer state appears intuitively suited to a molecule like H_2. The opposing attractive and repulsive forces of similar strength acting on its electrons (especially at higher bond lengths) can result in optimal configurations bearing resemblance to stabilizers.

B. LiH

Next, we look at LiH shown in Fig.5 (d)-(f). In (d) and (e), we see that HF deviates considerably away from exact at medium-high bond lengths but is nearer to exact at low bond lengths. As before, CAFQA is closest to exact, especially accurate at low and high bond lengths, but always achieves equal or more accurate energy estimates compared to HF. (e) shows that CAFQA's error is usually in the 10^{-1} to 10^{-2} Hartree range, with higher accuracy at high bond lengths. Finally, (f) shows that CAFQA is able to recover up to 93% of the correlation energy at medium-high bond lengths. As before, by improving beyond HF, CAFQA is able to produce a non-computational basis state as its ansatz initialization state.

C. H_2O

Next, we look at H_2O shown in Fig.5 (g)-(i). First to be noted is that HF Psi4 estimations do not converge at high bond lengths, so we extrapolate the expected trend as shown in (g). In (g) and (h) we see that HF steadily deviates away from exact at higher bond lengths. CAFQA matches HF at lower bond lengths but achieves considerably better energy estimates compared to HF at medium / higher bond lengths.

It is interesting to observe the kink in the energy estimation near a bond length of 1.5 Å. This appears to match the prior observation that this is caused by the energy crossing of the lowest singlet and triplet electronic states for the H_2O molecule [70]. Accordingly, the singlet and triplet states from CAFQA are plotted in pink and yellow respectively.

(a) H$_2$ Energy   **(b) H$_2$ Accuracy**   **(c) H$_2$ Correlation Energy**

(d) LiH Energy   **(e) LiH Accuracy**   **(f) LiH Correlation Energy**

(g) H$_2$O Energy   **(h) H$_2$O Accuracy**   **(i) H$_2$O Correlation Energy**

(j) H$_6$ Energy   **(k) H$_6$ Accuracy**   **(l) H$_6$ Correlation Energy**

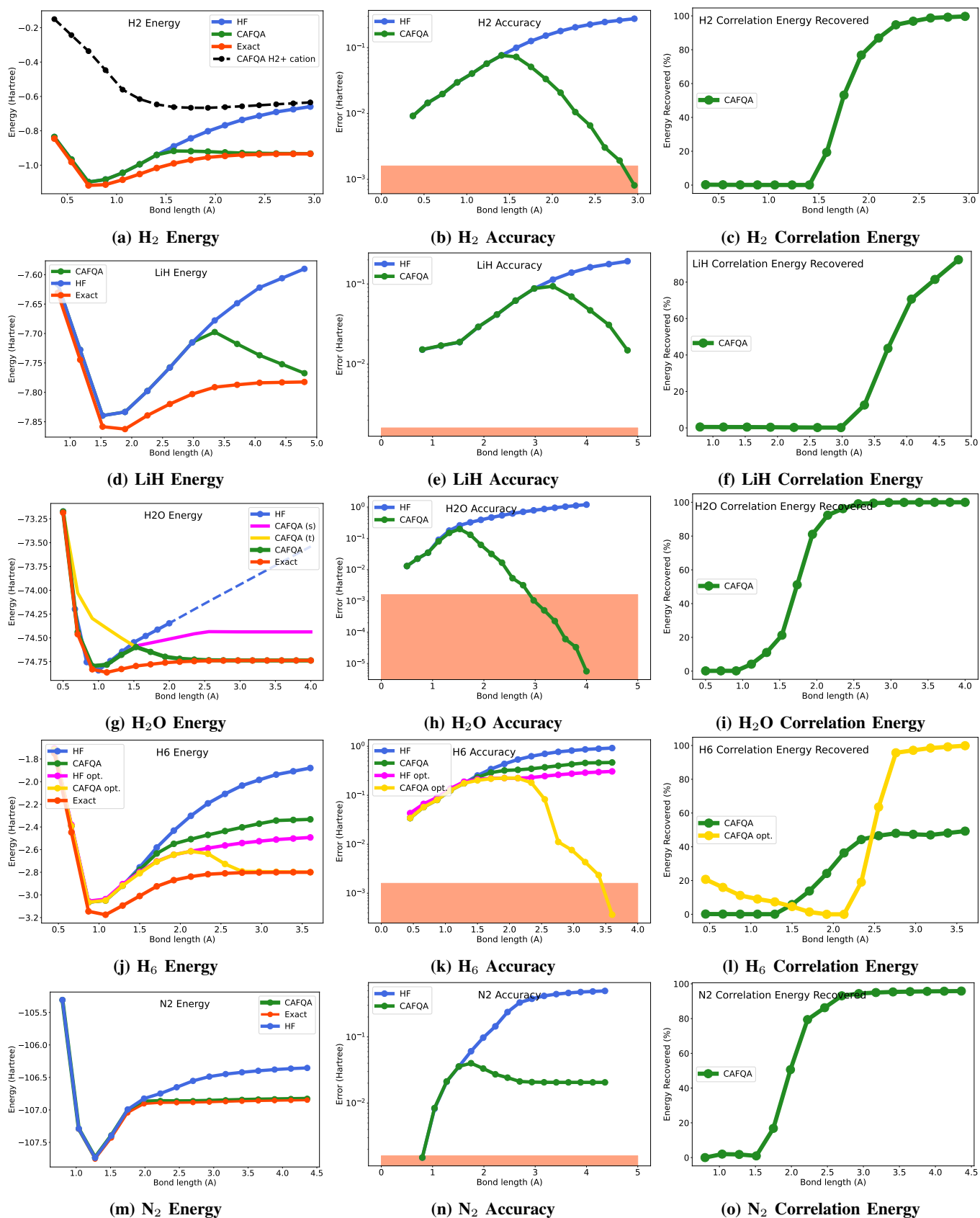**(m) N$_2$ Energy**   **(n) N$_2$ Accuracy**   **(o) N$_2$ Correlation Energy**

Fig. 5: Dissociation curves for different molecules. Evaluation of CAFQA in terms of ground state energy, energy estimation error and correlation energy recovered. Comparisons to Exact / Chemical Accuracy and Hartree-Fock are shown. CAFQA (s) and (t) refer to energies corresponding to singlet and triplet states. CAFQA opt. refers to best estimates over multiple spin-optimized Hamiltonians.

Next, (h) shows that CAFQA is able to impressively achieve chemical accuracy at higher bond lengths while HF has high error in the range of 10^{-1} Hartree. At lower bond lengths, CAFQA achieves error rates of around 10^{-2} Hartree. Finally, (i) shows that CAFQA is able to recover up to 99.998% of the correlation energy over HF as the bond lengths increase.

D. H_6

Next, we look at H_6 shown in Fig.5 (j)-(l). In (j) and (k) we show two versions of HF and CAFQA. The traditional HF and CAFQA results correspond to Hamiltonians generated for orbitals optimized for S=0 spin as mentioned in Appendix D. On the other hand, the optimized results (labeled 'opt.') are produced by generating unique Hamiltonians for different spins and with orbitals optimized accordingly. Then the HF and CAFQA results corresponding to the lowest estimates across all Hamiltonians are selected for every bond length. It is evident that 'opt.' results produce better energy estimates compared to those obtained from the S=0 Hamiltonian at higher bond lengths. This shows that optimizing the Hamiltonian to the best extent / as widely as possible can considerably improve VQE estimation, at the cost of increased compute. The HF techniques and CAFQA are far from exact while CAFQA opt. is near exact at high bond lengths. Deviations from exact are unsurprising since H_6 is known to have high correlation energy.

(k) shows that CAFQA errors are in the 10^{-1} Hartree range (except at high bond lengths when CAFQA opt. achieves chemical accuracy range), thus clearly requiring quantum exploration on a quantum device post ansatz selection. Finally, (l) shows that CAFQA is able to recover up to 50% of the correlation energy over HF as the bond lengths increase, whereas CAFQA opt. can achieve near 100% at high bond lengths. CAFQA is again able to produce a non-computational basis state as its ansatz initialization state, although it is evident that exploration of only the Clifford space limits accuracy.

E. N_2

Finally we look at N_2 shown in Fig.5 (m)-(o). In (m) and (n), we see that HF deviates from exact at higher bond lengths, whereas CAFQA has higher accuracy throughout. (n) shows that CAFQA error is usually in the 10^{-2} Hartree range whereas HF error is around 10^{-1} Hartree and worsens at high bond lengths. Finally, (o) shows that CAFQA is able to recover up to 96% of the correlation energy at medium-high bond lengths.

F. Chromium Dimer

In Fig.6 we evaluate CAFQA on the Cr_2 for ground state energy estimation. For Cr_2, we are unable to generate perfect evaluations since the size of the system is too large (discussed in Appendix D). We compare CAFQA to HF as well as experimental and approximate classical methods (CASPT2) [83]. CASPT or Complete active space perturbation theory is a multireference electron correlation method for computational investigation of molecular systems. It is evident from the figure
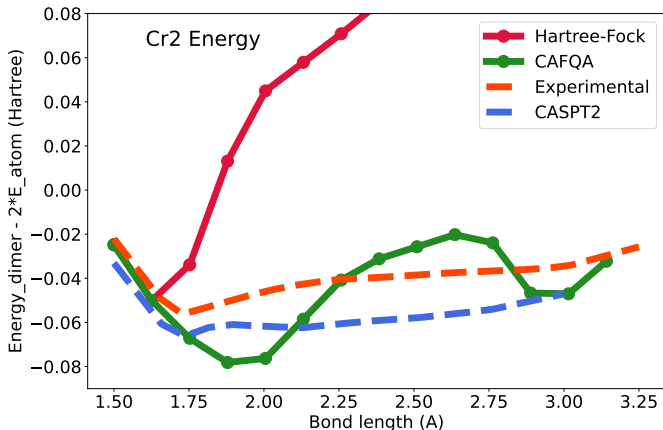

Fig. 6: Dissociation curves for Cr_2 with lower 18 out of 36 orbitals frozen. Evaluation of CAFQA in terms of ground state energy with comparisons to Hartree-Fock. CAFQA consistently outperforms Hartree-Fock. Note that CAFQA's estimates are limited by computational time / resources for the BO search, thus particular estimates can be uncharacteristically worse than others. Comparisons to experimental and classical methods (CASPT2) are shown [83].

that CAFQA consistently achieves better energy estimates compared to HF across all bond lengths. Further, CAFQA bears resemblance to prior estimates.

Due to the size of the problem, we are limited by resources in extensively running the BO search at each bond length. Our current estimates are obtained over a two-week period but these estimations can improve with more memory / compute, more execution time, better search strategies, efficient parallelization, limited additional exploration of non-Cliffords and so on - more in Appendix E. Limited search means that CAFQA at some bond lengths can produce sub-optimal estimates.

Note that CAFQA is still only a first step in VQA tasks, with the primary goal of producing an ansatz initial state which is well suited to further quantum exploration on a quantum device. While target accuracy can sometimes be obtained even without further quantum exploration (e.g., H_2O) this is likely to be uncommon, especially as complexity increases to systems such as Cr_2.

V. OTHER EVALUATION RESULTS

A. Overall relative accuracy

Fig.7 shows the relative accuracy achieved by CAFQA across all applications (except Cr_2) - 8 VQE molecular chemistry ground state estimation tasks and 2 MaxCut optimizations. For molecular chemistry, the accuracy for each molecule is averaged over bond lengths. It is evident that the CAFQA is able to achieve high accuracy over all applications with a mean of 98.5%. The lowest accuracy is achieved in H_6, which, as explained in Appendix D, has a significant correlation energy component which cannot be entirely recovered by
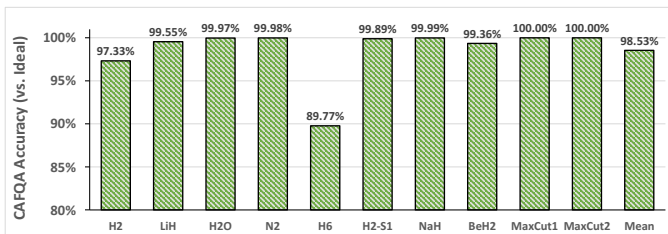
Fig. 7: Overall relative accuracy: CAFQA initialization achieves high accuracy estimates, potentially enabling faster convergence to target accuracy in post-CAFQA VQA tuning.
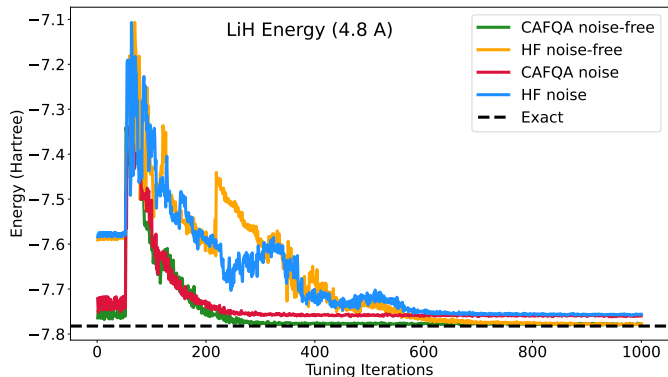

Fig. 8: Post-CAFQA VQA tuning for LiH. CAFQA initialization leads to 2.5x faster VQA convergence compared to Hartree-Fock, on both ideal as well as noisy quantum (simulated) machines.
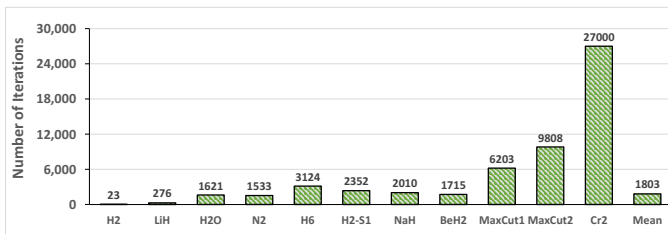

Fig. 9: BO search iterations for CAFQA to converge to the lowest expectation estimates in each targeted VQA problem. Search overheads are reasonable, especially considering the benefits of reduced variational tuning on noisy quantum devices.

only exploring the Clifford space. But it is evident that high VQA accuracy can be achieved by CAFQA even prior to traditional tuning on a quantum device. This allows for more focused tuning on actual quantum (NISQ) devices, meaning that there is lower potential for detrimental impact from noise and barren plateaus, potentially leading to faster and more accurate convergence.

B. Post-CAFQA VQA exploration

In Appendix A we discuss that real world chemistry use-cases require very high accuracy in ground state energy estimation. Similar accuracy requirements exist in other usecases and domains as well. Thus, it is almost always the case that the accuracy obtained from the best classical initializations are insufficient by themselves for real world accuracy. This is evident from Fig.5 - CAFQA can provide accurate initialization but most problems require post-CAFQA quantum exploration. Although today's NISQ machines are often too noisy to improve on CAFQA's estimates, we envision that NISQ machines in the near future will be able to do so. In this case, the CAFQA initialization usually leads to faster and more accurate VQA tuning convergence.

This is illustrated in Fig.8 which shows post-CAFQA VQE tuning for LiH ground state energy estimation. Evaluation is shown for tuning beginning from HF initialization and

CAFQA initialization, respectively. Furthermore, two sets of results are shown - one on ideal noise-free simulation and the other on noisy simulation modeled on real machine characteristics. In both sets of results it is evident that CAFQA-initialized exploration converges roughly 2.5x faster than HF-initialization, clearly indicative of the benefits of: a) CAFQA as a suitable classical initialization framework and b) stabilizers as a potentially accurate initialization state. Further, it can be observed that the ideal simulation produces near exact results, improving over the initialization. On the other hand, the energy estimation produced via the noisy simulation (error roughly $= 10^{-2}$ Hartree) is at par with the estimate obtained directly from CAFQA classical initialization itself. While we do not expect the latter trend to hold for more complex Hamiltonians and as machine noise reduces, CAFQA initialization will continue to be useful for fast and accurate convergence.

C. Discrete Search

Fig.9 shows the number of iterations consumed by CAFQA's discrete search to converge to a minimum energy estimate. It is evident that the number of iterations increases with the size of the problem because the number of tuning parameters increases. The number of iterations across all applications is very reasonable considering the benefits of reduced variational tuning on noisy quantum devices. Current runtime for CAFQA roughly varies from one hour (H_2) to two weeks (Cr_2). The execution time can be reduced via increased compute / memory resources, improved search algorithm, parallelizing the search etc. As discussed in Section IV-F, any such improvement would be especially useful for more complex systems such as Cr_2 and beyond.

VI. CONCLUSION

CAFQA tackles the problem of finding initial VQA parameters by proposing a "Clifford Ansatz" — an ansatz that is a hardware efficient circuit built with only Clifford gates. In this ansatz, the initial parameters for the tunable gates are chosen by searching efficiently through the Clifford parameter space via classical simulation, and thereby producing a suitable stabilizer state. The stabilizer states produced are shown to

equal or outperform the traditional classical (e.g., Hartree-Fock) approach of finding a suitable computational basis state, and often achieve high accuracy estimations even prior to execution/exploration on a quantum device. Therefore, they are attractive as an initialization for VQAs. Furthermore, the technique is classically efficient since Clifford-only quantum circuits can be perfectly simulated in polynomial time on classical computers and the discrete Clifford space, while scaling exponentially in the number of qubits, is searched efficiently via Bayesian Optimization using a Random Forest surrogate model and a greedy acquisition function.

In all, this work shows that stabilizer states produced from a Clifford ansatz are a highly accurate initialization for variational algorithms. Further, there is considerable potential to extend these findings beyond Cliffords and to other circuit structures, targeting a variety of problems in molecular chemistry and beyond (Appendix E). Finally, CAFQA is a promising example of the potential for quantum-inspired classical techniques as both an alternative as well as a supporting methodology for VQAs (and quantum computing in general) in the NISQ era and beyond.

REFERENCES

[1] "Ibm quantum su2 ansatz," https://qiskit.org/documentation/stubs/qiskit.circuit.library.EfficientSU2.html, accessed: 2021-12-01.

[2] "Microsoft docs hartree–fock theory," https://docs.microsoft.com/en-us/azure/quantum/user-guide/libraries/chemistry/concepts/hartree-fock.

[3] "Nist atomic spectra database ionization energies form," https://physics.nist.gov/PhysRefData/ASD/ionEnergy.html, accessed: 2022-01-01.

[4] "Techniques and applications of quantum monte carlo," https://web.ornl.gov/~kentpr/thesis/pkthnode13.html.

[5] H. Abraham, AduOffei, R. Agarwal, I. Y. Akhalwaya, G. Aleksandrowicz, T. Alexander, M. Amy, E. Arbel, Arijit02, A. Asfaw, A. Avkhadiev, C. Azaustre, AzizNgoueya, A. Banerjee, A. Bansal, P. Barkoutsos, G. Barron, G. S. Barron, L. Bello, Y. Ben-Haim, D. Bevenius, A. Bhobe, L. S. Bishop, C. Blank, S. Bolos, S. Bosch, Brandon, S. Bravyi, Bryce-Fuller, D. Bucher, A. Burov, F. Cabrera, P. Calpin, L. Capelluto, J. Carballo, G. Carrascal, A. Chen, C.-F. Chen, E. Chen, J. C. Chen, R. Chen, J. M. Chow, S. Churchill, C. Claus, C. Clauss, R. Cocking, F. Correa, A. J. Cross, A. W. Cross, S. Cross, J. Cruz-Benito, C. Culver, A. D. Córcoles-Gonzales, S. Dague, T. E. Dandachi, M. Daniels, M. Dartiailh, DavideFrr, A. R. Davila, A. Dekusar, D. Ding, J. Doi, E. Drechsler, Drew, E. Dumitrescu, K. Dumon, I. Duran, K. EL-Safty, E. Eastman, G. Eberle, P. Eendebak, D. Egger, M. Everitt, P. M. Fernández, A. H. Ferrera, R. Fouilland, FranckChevallier, A. Frisch, A. Fuhrer, B. Fuller, M. GEORGE, J. Gacon, B. G. Gago, C. Gambella, J. M. Gambetta, A. Gammanpila, L. Garcia, T. Garg, S. Garion, A. Gilliam, A. Giridharan, J. Gomez-Mosquera, S. de la Puente González, J. Gorzinski, I. Gould, D. Greenberg, D. Grinko, W. Guan, J. A. Gunnels, M. Haglund, I. Haide, I. Hamamura, O. C. Hamido, F. Harkins, V. Havlicek, J. Hellmers, Ł. Herok, S. Hillmich, H. Horii, C. Howington, S. Hu, W. Hu, J. Huang, R. Huisman, H. Imai, T. Imamichi, K. Ishizaki, R. Iten, T. Itoko, JamesSeaward, A. Javadi, A. Javadi-Abhari, Jessica, M. Jivrajani, K. Johns, S. Johnstun, Jonathan-Shoemaker, V. K, T. Kachmann, N. Kanazawa, Kang-Bae, A. Karazeev, P. Kassebaum, J. Kelso, S. King, Knabberjoe, Y. Kobayashi, A. Kovyrshin, R. Krishnakumar, V. Krishnan, K. Krsulich, P. Kumkar, G. Kus, R. LaRose, E. Lacal, R. Lambert, J. Lapeyre, J. Latone, S. Lawrence, C. Lee, G. Li, D. Liu, P. Liu, Y. Maeng, K. Majmudar, A. Malyshev, J. Manela, J. Marecek, M. Marques, D. Maslov, D. Mathews, A. Matsuo, D. T. McClure, C. McGarry, D. McKay, D. McPherson, S. Meesala, T. Metcalfe, M. Mevissen, A. Meyer, A. Mezzacapo, R. Midha, Z. Minev, A. Mitchell, N. Moll, J. Montanez, M. D. Mooring, R. Morales, N. Moran, M. Motta, MrF, P. Murali, J. Müggenburg, D. Nadlinger, K. Nakanishi, G. Nannicini, P. Nation, E. Navarro, Y. Naveh, S. W. Neagle, P. Neuweiler, J. Nicander, P. Niroula, H. Norlen, NuoWenLei, L. J. O'Riordan, O. Ogunbayo, P. Ollitrault, R. Otaolea, S. Oud, D. Padilha, H. Paik, S. Pal, Y. Pang, S. Perriello, A. Phan, F. Piro, M. Pistoia, C. Piveteau, P. Pocreau, A. Pozas-iKerstjens, V. Prutyanov, D. Puzzuoli, J. Pérez, Quintiii, R. I. Rahman, A. Raja, N. Ramagiri, A. Rao, R. Raymond, R. M.-C. Redondo, M. Reuter, J. Rice, M. L. Rocca, D. M. Rodríguez, RohithKarur, M. Rossmannek, M. Ryu, T. SAPV, SamFerracin, M. Sandberg, H. Sandesara, R. Sapra, H. Sargsyan, A. Sarkar, N. Sathaye, B. Schmitt, C. Schnabel, Z. Schoenfeld, T. L. Scholten, E. Schoute, J. Schwarm, I. F. Sertage, K. Setia, N. Shammah, Y. Shi, A. Silva, A. Simonetto, N. Singstock, Y. Siraichi, I. Sitdikov, S. Sivarajah, M. B. Sletfjerding, J. A. Smolin, M. Soeken, I. O. Sokolov, I. Sokolov, SooluThomas, Starfish, D. Steenken, M. Stypulkoski, S. Sun, K. J. Sung, H. Takahashi, T. Takawale, I. Tavernelli, C. Taylor, P. Taylour, S. Thomas, M. Tillet, M. Tod, M. Tomasik, E. de la Torre, K. Trabing, M. Treinish, TrishaPe, D. Tulsi, W. Turner, Y. Vaknin, C. R. Valcarce, F. Varchon, A. C. Vazquez, V. Villar, D. Vogt-Lee, C. Vuillot, J. Weaver, J. Weidenfeller, R. Wieczorek, J. A. Wildstrom, E. Winston, J. J. Woehr, S. Woerner, R. Woo, C. J. Wood, R. Wood, S. Wood, S. Wood, J. Wootton, D. Yeralin, D. Yonge-Mallo, R. Young, J. Yu, C. Zachow, L. Zdanski, H. Zhang, C. Zoufal, Zoufalc, a kapila, a matsuo, bcamorrison, brandhsn, nick bronn, chlorophyll zz, dekel.meirom, dekelmeirom, dekool, dime10, drholmie, dtrenev, ehchen, elfrocampeador, faisaldebouni, fanizzamarco, gabrieleagl, gadial, galeinston, georgios ts, gruu, hhorii, hykavitha, jagunther, jliu45, jscott2, kanejess, klinvill, krutik2966, kurarrr, lerongil, ma5x, merav aharoni, michelle4654, ordmoj, sagar pahwa, rmoyard, saswati qiskit, scottkelso, sethmerkel, strickroman, sumitpuri, tigerjack, toural, tsura crisaldo, vvilpas, welien, willhbang, yang.luh, yotam-vakninibm, and M. Čepulkovskis, "Qiskit: An open-source framework for quantum computing," 2019.

[6] O. Alipourfard, H. H. Liu, J. Chen, S. Venkataraman, M. Yu, and M. Zhang, "Cherrypick: Adaptively unearthing the best cloud configurations for big data analytics," in *Proceedings of the 14th USENIX Conference on Networked Systems Design and Implementation*, ser. NSDI'17. USA: USENIX Association, 2017, p. 469–482.

[7] G. S. Barron and C. J. Wood, "Measurement error mitigation for variational quantum algorithms," 2020.

[8] K. Beer, D. Bondarenko, T. Farrelly, T. J. Osborne, R. Salzmann,

D. Scheiermann, and R. Wolf, "Training deep quantum neural networks," *Nature Communications*, vol. 11, no. 1, Feb 2020. [Online]. Available: http://dx.doi.org/10.1038/s41467-020-14454-2

[9] J. Biamonte, P. Wittek, N. Pancotti, P. Rebentrost, N. Wiebe, and S. Lloyd, "Quantum machine learning," *Nature*, vol. 549, no. 7671, pp. 195–202, 2017.

[10] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, "Characterizing quantum supremacy in near-term devices," *Nature Physics*, vol. 14, no. 6, p. 595–600, Apr 2018. [Online]. Available: http://dx.doi.org/10.1038/s41567-018-0124-x

[11] L. Botelho, A. Glos, A. Kundu, J. A. Miszczak, Özlem Salehi, and Z. Zimborás, "Error mitigation for variational quantum algorithms through mid-circuit measurements," 2021.

[12] S. Bravyi, D. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, "Simulation of quantum circuits by low-rank stabilizer decompositions," *Quantum*, vol. 3, p. 181, Sep 2019. [Online]. Available: http://dx.doi.org/10.22331/q-2019-09-02-181

[13] S. Bravyi and D. Gosset, "Improved classical simulation of quantum circuits dominated by clifford gates," *Physical Review Letters*, vol. 116, no. 25, Jun 2016. [Online]. Available: http://dx.doi.org/10.1103/PhysRevLett.116.250501

[14] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, "Cost function dependent barren plateaus in shallow parametrized quantum circuits," *Nature Communications*, vol. 12, no. 1, Mar 2021. [Online]. Available: http://dx.doi.org/10.1038/s41467-021-21728-w

[15] A. Cervera-Lierta, J. S. Kottmann, and A. Aspuru-Guzik, "Meta-variational quantum eigensolver: Learning energy profiles of parameterized hamiltonians for quantum simulation," *PRX Quantum*, vol. 2, no. 2, May 2021. [Online]. Available: http://dx.doi.org/10.1103/PRXQuantum.2.020329

[16] A. S. Christensen, S. K. Sirumalla, Z. Qiao, M. B. O'Connor, D. G. A. Smith, F. Ding, P. J. Bygrave, A. Anandkumar, M. Welborn, F. R. Manby, and et al., "Orbnet denali: A machine learning potential for biological and organic chemistry with semi-empirical cost and dft accuracy," *The Journal of Chemical Physics*, vol. 155, no. 20, p. 204103, Nov 2021. [Online]. Available: http://dx.doi.org/10.1063/5.0061990

[17] P. Czarnik, A. Arrasmith, P. J. Coles, and L. Cincio, "Error mitigation with clifford quantum-circuit data," 2020.

[18] H. De Raedt, F. Jin, D. Willsch, M. Willsch, N. Yoshioka, N. Ito, S. Yuan, and K. Michielsen, "Massively parallel quantum computer simulator, eleven years later," *Computer Physics Communications*, vol. 237, p. 47–61, Apr 2019. [Online]. Available: http://dx.doi.org/10.1016/j.cpc.2018.11.005

[19] Y. Ding, A. Pervaiz, M. Carbin, and H. Hoffmann, "Generalizable and interpretable learning for configuration extrapolation," in *Proceedings of the 29th ACM Joint Meeting on European Software Engineering Conference and Symposium on the Foundations of Software Engineering*, ser. ESEC/FSE 2021. New York, NY, USA: Association for Computing Machinery, 2021, p. 728–740. [Online]. Available: https://doi.org/10.1145/3468264.3468603

[20] Y. Ding, A. Pervaiz, S. Krishnan, and H. Hoffmann, "Bayesian learning for hardware and software configuration co-optimization."

[21] Y. Ding, P. Gokhale, S. F. Lin, R. Rines, T. Propson, and F. T. Chong, "Systematic crosstalk mitigation for superconducting qubits via frequency-aware compilation," *arXiv preprint arXiv:2008.09503*, 2020.

[22] A. Dommer, L. Casalino, F. Kearns, M. Rosenfeld, N. Wauer, S.-H. Ahn, J. Russo, S. Oliveira, C. Morris, A. Bogetti *et al.*, "# covidisairborne: Ai-enabled multiscale computational microscopy of delta sars-cov-2 in a respiratory aerosol," *bioRxiv*, 2021.

[23] V. E. Elfving, B. W. Broer, M. Webber, J. Gavartin, M. D. Halls, K. P. Lorton, and A. Bochevarov, "How will quantum computers provide an industrially relevant computational advantage in quantum chemistry?" 2020.

[24] E. Farhi, J. Goldstone, and S. Gutmann, "A quantum approximate optimization algorithm," 2014.

[25] P. I. Frazier, "A tutorial on bayesian optimization," *arXiv preprint arXiv:1807.02811*, 2018.

[26] X. Fu, M. A. Rol, C. C. Bultink, J. van Someren, N. Khammassi, I. Ashraf, R. F. L. Vermeulen, J. C. de Sterke, W. J. Vlothuizen, R. N. Schouten, C. G. Almudever, L. DiCarlo, and K. Bertels, "An experimental microarchitecture for a superconducting quantum processor," in *Proceedings of the 50th Annual IEEE/ACM International Symposium on Microarchitecture*, ser. MICRO-50 '17. New York, NY, USA: Association for Computing Machinery, 2017, p. 813–825. [Online]. Available: https://doi.org/10.1145/3123939.3123952

[27] T. Giurgica-Tiron, Y. Hindy, R. LaRose, A. Mari, and W. J. Zeng, "Digital zero noise extrapolation for quantum error mitigation," in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*. IEEE, 2020, pp. 306–316.

[28] P. Gokhale, O. Angiuli, Y. Ding, K. Gui, T. Tomesh, M. Suchara, M. Martonosi, and F. T. Chong, "Minimizing state preparations in variational quantum eigensolver by partitioning into commuting families," *arXiv preprint arXiv:1907.13623*, 2019.

[29] P. Gokhale, Y. Ding, T. Propson, C. Winkler, N. Leung, Y. Shi, D. I. Schuster, H. Hoffmann, and F. T. Chong, "Partial compilation of variational algorithms for noisy intermediate-scale quantum machines," *Proceedings of the 52nd Annual IEEE/ACM International Symposium on Microarchitecture*, Oct 2019. [Online]. Available: http://dx.doi.org/10.1145/3352460.3358313

[30] D. Gottesman, "The heisenberg representation of quantum computers," *arXiv preprint quant-ph/9807006*, 1998.

[31] D. Gottesman and I. L. Chuang, "Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations," *Nature*, vol. 402, no. 6760, pp. 390–393, 1999.

[32] H. R. Grimsley, S. E. Economou, E. Barnes, and N. J. Mayhall, "An adaptive variational algorithm for exact molecular simulations on a quantum computer," *Nature Communications*, vol. 10, no. 1, Jul 2019. [Online]. Available: http://dx.doi.org/10.1038/s41467-019-10988-2

[33] L. K. Grover, "A fast quantum mechanical algorithm for database search," in *ANNUAL ACM SYMPOSIUM ON THEORY OF COMPUTING*. ACM, 1996, pp. 212–219.

[34] D. R. Hartree and W. Hartree, "Self-consistent field, with exchange, for beryllium," *Proceedings of the Royal Society of London. Series A-Mathematical and Physical Sciences*, vol. 150, no. 869, pp. 9–33, 1935.

[35] M. Heinrich and D. Gross, "Robustness of magic and symmetries of the stabiliser polytope," *Quantum*, vol. 3, p. 132, Apr 2019. [Online]. Available: http://dx.doi.org/10.22331/q-2019-04-08-132

[36] O. Higgott, D. Wang, and S. Brierley, "Variational quantum computation of excited states," *Quantum*, vol. 3, p. 156, Jul 2019. [Online]. Available: http://dx.doi.org/10.22331/q-2019-07-01-156

[37] A. A. Holmes, H. J. Changlani, and C. J. Umrigar, "Efficient heat-bath sampling in fock space," *Journal of Chemical Theory and Computation*, vol. 12, no. 4, p. 1561–1571, Mar 2016. [Online]. Available: http://dx.doi.org/10.1021/acs.jctc.5b01170

[38] Z. Holmes, K. Sharma, M. Cerezo, and P. J. Coles, "Connecting ansatz expressibility to gradient magnitudes and barren plateaus," 2021.

[39] T. Häner and D. S. Steiger, "0.5 petabyte simulation of a 45-qubit quantum circuit," *Proceedings of the International Conference for High Performance Computing, Networking, Storage and Analysis*, Nov 2017. [Online]. Available: http://dx.doi.org/10.1145/3126908.3126947

[40] M. A. Jones, H. J. Vallury, C. D. Hill, and L. C. L. Hollenberg, "Chemistry beyond the hartree-fock limit via quantum computed moments," 2021.

[41] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, "Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets," *Nature*, vol. 549, no. 7671, pp. 242–246, 2017.

[42] W. M. Kirby, A. Tranter, and P. J. Love, "Contextual subspace variational quantum eigensolver," *Quantum*, vol. 5, p. 456, May 2021. [Online]. Available: http://dx.doi.org/10.22331/q-2021-05-14-456

[43] W. Kohn, A. D. Becke, and R. G. Parr, "Density functional theory of electronic structure," *The Journal of Physical Chemistry*, vol. 100, no. 31, pp. 12 974–12 980, 1996.

[44] W. Lavrijsen, A. Tudor, J. Müller, C. Iancu, and W. de Jong, "Classical optimizers for noisy intermediate-scale quantum devices," in *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*, 2020, pp. 267–277.

[45] G. Li, Y. Shi, and A. Javadi-Abhari, "Software-hardware co-optimization for computational chemistry on superconducting quantum processors," 2021.

[46] Y. Li and S. C. Benjamin, "Efficient variational quantum simulator incorporating active error minimization," *Phys. Rev. X*, vol. 7, p. 021050, Jun 2017. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevX.7.021050

[47] C. O. Marrero, M. Kieferová, and N. Wiebe, "Entanglement induced barren plateaus," 2021.

[48] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," *Nature Communications*, vol. 9, no. 1, Nov 2018. [Online]. Available: http://dx.doi.org/10.1038/s41467-018-07090-4

[49] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, "The theory of variational hybrid quantum-classical algorithms," *New Journal of Physics*, vol. 18, no. 2, p. 023023, 2016.

[50] K. Mitarai, Y. Suzuki, W. Mizukami, Y. O. Nakagawa, and K. Fujii, "Quadratic clifford expansion for efficient benchmarking and initialization of variational quantum algorithms," 2020.

[51] M. Mohseni, P. Read, H. Neven, S. Boixo, V. Denchev, R. Babbush, A. Fowler, V. Smelyanskiy, and J. Martinis, "Commercialize quantum technologies in five years," *Nature*, vol. 543, p. 171–174, 2017. [Online]. Available: http://www.nature.com/news/commercialize-quantum-technologies-in-five-years-1.21583

[52] N. Moll, P. Barkoutsos, L. S. Bishop, J. M. Chow, A. Cross, D. J. Egger, S. Filipp, A. Fuhrer, J. M. Gambetta, M. Ganzhorn *et al.*, "Quantum optimization using variational algorithms on near-term quantum devices," *Quantum Science and Technology*, vol. 3, no. 3, p. 030503, 2018.

[53] P. Murali, J. M. Baker, A. Javadi-Abhari, F. T. Chong, and M. Martonosi, "Noise-adaptive compiler mappings for noisy intermediate-scale quantum computers," in *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*, 2019, pp. 1015–1029.

[54] P. Murali, D. C. McKay, M. Martonosi, and A. Javadi-Abhari, "Software mitigation of crosstalk on noisy intermediate-scale quantum computers," in *Proceedings of the Twenty-Fifth International Conference on Architectural Support for Programming Languages and Operating Systems*, 2020, pp. 1001–1016.

[55] K. M. Nakanishi, K. Mitarai, and K. Fujii, "Subspace-search variational quantum eigensolver for excited states," *Physical Review Research*, vol. 1, no. 3, Oct 2019. [Online]. Available: http://dx.doi.org/10.1103/PhysRevResearch.1.033062

[56] L. Nardi, D. Koeplinger, and K. Olukotun, "Practical design space exploration," 2019.

[57] M. Ostaszewski, E. Grant, and M. Benedetti, "Structure optimization for parameterized quantum circuits," *Quantum*, vol. 5, p. 391, Jan 2021. [Online]. Available: http://dx.doi.org/10.22331/q-2021-01-28-391

[58] J. O'Gorman and E. T. Campbell, "Quantum computation with realistic magic-state factories," *Physical Review A*, vol. 95, no. 3, Mar 2017. [Online]. Available: http://dx.doi.org/10.1103/PhysRevA.95.032338

[59] T. Patel and D. Tiwari, "Clite: Efficient and qos-aware co-location of multiple latency-critical jobs for warehouse scale computers," in *2020 IEEE International Symposium on High Performance Computer Architecture (HPCA)*, 2020, pp. 193–206.

[60] T. L. Patti, K. Najafi, X. Gao, and S. F. Yelin, "Entanglement devised barren plateau mitigation," *Phys. Rev. Research*, vol. 3, p. 033090, Jul 2021. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevResearch.3.033090

[61] 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*, vol. 5, p. 4213, 2014.

[62] K. A. Peterson, D. Feller, and D. A. Dixon, "Chemical accuracy in ab initio thermochemistry and spectroscopy: current strategies and future challenges," *Theoretical Chemistry Accounts*, vol. 131, no. 1, pp. 1–20, 2012.

[63] J. Preskill, "Quantum computing in the nisq era and beyond," *Quantum*, vol. 2, p. 79, 2018.

[64] L. Ratini, C. Capecci, F. Benfenati, and L. Guidoni, "Wave function adapted hamiltonians for quantum computing," *Journal of Chemical Theory and Computation*, 2022.

[65] G. S. Ravi, K. N. Smith, P. Gokhale, A. Mari, N. Earnest, A. Javadi-Abhari, and F. T. Chong, "Vaqem: A variational approach to quantum error mitigation," 2021.

[66] J. Roffe, "Quantum error correction: an introductory guide," *Contemporary Physics*, vol. 60, no. 3, p. 226–245, Jul 2019. [Online]. Available: http://dx.doi.org/10.1080/00107514.2019.1667078

[67] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. Love, and A. Aspuru-Guzik, "Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz," 2018.

[68] E. Rosenberg, P. Ginsparg, and P. L. McMahon, "Experimental error mitigation using linear rescaling for variational quantum eigensolving with up to 20 qubits," *Quantum Science and Technology*, Nov 2021. [Online]. Available: http://dx.doi.org/10.1088/2058-9565/ac3b37

[69] R. B. Roy, T. Patel, V. Gadepally, and D. Tiwari, *Bliss: Auto-Tuning Complex Applications Using a Pool of Diverse Lightweight Learning Models*. New York, NY, USA: Association for Computing Machinery, 2021, p. 1280–1295. [Online]. Available: https://doi.org/10.1145/3453483.3454109

[70] I. G. Ryabinkin, S. N. Genin, and A. F. Izmaylov, "Constrained variational quantum eigensolver: Quantum computer search engine in the fock space," 2018.

[71] I. G. Ryabinkin, T.-C. Yen, S. N. Genin, and A. F. Izmaylov, "Qubit coupled cluster method: a systematic approach to quantum chemistry on a quantum computer," *Journal of chemical theory and computation*, vol. 14, no. 12, pp. 6317–6326, 2018.

[72] F. Sauvage, S. Sim, A. A. Kunitsa, W. A. Simon, M. Mauri, and A. Perdomo-Ortiz, "Flip: A flexible initializer for arbitrarily-sized parametrized quantum circuits," 2021.

[73] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *SIAM Journal on Computing*, vol. 26, no. 5, p. 1484–1509, Oct 1997. [Online]. Available: http://dx.doi.org/10.1137/S0097539795293172

[74] K. N. Smith, G. S. Ravi, P. Murali, J. M. Baker, N. Earnest, A. Javadi-Abhari, and F. T. Chong, "Error mitigation in quantum computers through instruction scheduling," *arXiv preprint arXiv:2105.01760*, 2021.

[75] A. Strikis, D. Qin, Y. Chen, S. C. Benjamin, and Y. Li, "Learning-based quantum error mitigation," 2021.

[76] Q. Sun, T. C. Berkelbach, N. S. Blunt, G. H. Booth, S. Guo, Z. Li, J. Liu, J. McClain, E. R. Sayfutyarova, S. Sharma, S. Wouters, and G. K.-L. Chan, "The python-based simulations of chemistry framework (pyscf)," 2017.

[77] R. Takagi, S. Endo, S. Minagawa, and M. Gu, "Fundamental limits of quantum error mitigation," 2021.

[78] S. S. Tannu and M. K. Qureshi, "Not all qubits are created equal: a case for variability-aware policies for nisq-era quantum computers," in *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*, 2019, pp. 987–999.

[79] K. Temme, S. Bravyi, and J. M. Gambetta, "Error mitigation for short-depth quantum circuits," *Physical review letters*, vol. 119, no. 18, p. 180509, 2017.

[80] J. Tilly, H. Chen, S. Cao, D. Picozzi, K. Setia, Y. Li, E. Grant, L. Wossnig, I. Rungger, G. H. Booth, and J. Tennyson, "The variational quantum eigensolver: a review of methods and best practices," 2021.

[81] J. M. Turney, A. C. Simmonett, R. M. Parrish, E. G. Hohenstein, F. A. Evangelista, J. T. Fermann, B. J. Mintz, L. A. Burns, J. J. Wilke, M. L. Abrams *et al.*, "Psi4: an open-source ab initio electronic structure program," *Wiley Interdisciplinary Reviews: Computational Molecular Science*, vol. 2, no. 4, pp. 556–565, 2012.

[82] A. V. Uvarov and J. D. Biamonte, "On barren plateaus and cost function locality in variational quantum algorithms," *Journal of Physics A: Mathematical and Theoretical*, vol. 54, no. 24, p. 245301, May 2021. [Online]. Available: http://dx.doi.org/10.1088/1751-8121/abfac7

[83] S. Vancoillie, P. Å. Malmqvist, and V. Veryazov, "Potential energy surface of the chromium dimer re-re-revisited with multiconfigurational perturbation theory," *Journal of chemical theory and computation*, vol. 12, no. 4, pp. 1647–1655, 2016.

[84] V. Veitch, S. A. Hamed Mousavian, D. Gottesman, and J. Emerson, "The resource theory of stabilizer quantum computation," *New Journal of Physics*, vol. 16, no. 1, p. 013009, Jan 2014. [Online]. Available: http://dx.doi.org/10.1088/1367-2630/16/1/013009

[85] S. Wang, P. Czarnik, A. Arrasmith, M. Cerezo, L. Cincio, and P. J. Coles, "Can error mitigation improve trainability of noisy variational quantum algorithms?" 2021.

[86] S. Wang, E. Fontana, M. Cerezo, K. Sharma, A. Sone, L. Cincio, and P. J. Coles, "Noise-induced barren plateaus in variational quantum algorithms," *arXiv preprint arXiv:2007.14384*, 2020.

[87] J. Wurtz and P. Love, "Classically optimal variational quantum algorithms," 2021.

[88] F. Zhang, N. Gomes, Y. Yao, P. P. Orth, and T. Iadecola, "Adaptive variational quantum eigensolvers for highly excited states," *Physical Review B*, vol. 104, no. 7, Aug 2021. [Online]. Available: http://dx.doi.org/10.1103/PhysRevB.104.075159

[89] R. Zhang, G. Wang, and P. Johnson, "Computing ground state properties with early fault-tolerant quantum computers," 2021.

[90] Y. Zhou, E. M. Stoudenmire, and X. Waintal, "What limits the simulation of quantum computers?" *Phys. Rev. X*, vol. 10, p. 041038, Nov 2020. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevX.10.041038

APPENDICES

A. Background and Motivation

VQAs in the NISQ era: Today, Noisy Intermediate Scale Quantum (NISQ) devices are error prone and only around 100 qubits in size [63]. These devices are extremely sensitive to external influences and require precise control, and as a result, some of the biggest challenges that limit scalability include limited coherence, gate errors, readout errors, and connectivity.

Variational quantum algorithms are especially important in the near-term because they comply with the constraints of NISQ hardware. In particular, variational algorithms have innate error resilience, due to the hybrid alternation with a noise-robust classical optimizer [49], [61]. There are multiple applications which are part of the VQA domain such as Quantum Approximate Optimization Algorithm (QAOA) [24] and Variational Quantum Eigensolver (VQE) [61]. While CAFQA is suited widely across variationally algorithms, our discussions and evaluations in this paper primarily focus on VQE.

The Variational Quantum Eigensolver (VQE) [61] is used to compute an upper bound of a Hamiltonian's ground state energy. This is done by finding a circuit parameterization such that the expectation value of the target Hamiltonian is minimized. Ideally the expectation should be indistinguishable from the ground state energy to the level of accuracy desired [80]. At a high level, VQE can be conceptualized as a guess-check-repeat algorithm [28]. The check stage involves the preparation of a quantum state corresponding to the guess. This preparation stage is done in polynomial time on a quantum computer, but would incur exponential cost in general on a classical computer. This contrast gives rise to a potential quantum speedup for VQE [29].

VQE can be a critical step in computing the energy properties of molecules and materials. While conventional computational chemistry provides efficient methods to approximate such properties, they can lack sufficient accuracy in molecular systems due to an inadequate treatment of the correlations between constituent electrons. These interactions require computation that scales exponentially in the size of the system, rendering exact quantum chemistry methods generally intractable with conventional computing [80], [90].

Estimating the VQE global optimal with high accuracy has proven challenging in the NISQ era even with sophisticated optimizers, a well chosen ansatz, and error mitigation [7], [11], [17], [65], [68], [77], [80], [85]. As an example, ground state energy estimation of molecules (a key use case for VQE) requires energy estimates with an estimation error of less than 1.6×10^{-3} Hartree, or what is known as "chemical accuracy" [62]. Such energy estimates are then used in molecular chemistry for determining properties of chemical reactions and so forth. Unfortunately, energy estimates on NISQ machines are far from accurate, especially as problem size increases. For instance, prior work on the ground state energy estimation of BeH_2 on a superconducting transmon machine results in an error greater than 10^{-1} Hartree, which is roughly 100x lower than the required accuracy [41].

Considering the significant disparity in NISQ-device VQA accuracy and real world requirements, it is imperative to aid VQA methods to the best extent possible. CAFQA looks to classical computing support (as has prior work [21], [26], [27], [46], [53], [54], [74], [78], [79]) and puts it to use efficiently in an accurate and scalable manner.

VQA Ansatz and its Initialization: The VQA problem is translated into a task that can be executed on a quantum computer by defining an ansatz wave function that can be implemented on a quantum device as a series of quantum gates - a circuit with parameterized unitary operations is used [80]. An ansatz with its parameterized gate rotation angles is shown in Fig. 10.

Many ansatz structures are suitable for VQAs. In the context of VQE for molecular chemistry, the Unitary Coupled Cluster Single-Double (UCCSD) ansatz is considered the gold standard [29], [67]. Unfortunately the UCCSD ansatz is generally of considerable circuit depth, making it less suitable for today's NISQ machines except for very small molecules like H_2. More suitable to the NISQ-era are hardware efficient ansatz like the SU2 [1] which are low depth parameterized circuits (but have Hilbert space coverage limitations [38], [80]). Fundamentally this ansatz is constructed by repeating blocks of parameterized single qubit rotation gates, and ladders of entangling gates [80]. Hardware efficient ansatz has been used on real NISQ machines targeting molecular chemistry tasks for molecules such as BeH_2 with promising ground state energy estimates but not within chemical accuracy [41]. CAFQA builds atop a traditional hardware efficient ansatz [41]. Suitability to other ansatz structures is discussed in Section E.

While a good choice of classical optimizer improves VQA convergence [44], VQAs can a) have complex optimization surfaces and b) suffer from a barren plateau problem. For simple VQA problems and in the presence of minimal noise, the optimization surface is convex around the global minimum and smooth [44]. As noise increases, in line with NISQ machines, the optimization surface can potentially become non-convex and non-smooth. The surface contour worsens as problem complexity increases because of increase in circuit depth, number of parameters (and thus the likelihood of noise-induced local minima), and entanglement spread out over many qubits (causing noise effects to become non-local) [44]. The barren plateau is the phenomenon that the gradients of the VQE parameters vanish exponentially. This could occur in increasing fashion with the number of qubits [48], in the expressibility of the ansatz [37], the degree of entanglement of the wavefunction [47], [60], or non-locality of the cost function [14], [82]. Important to this work, barren plateaus can become an especially significant problem in the presence of noise [86] and with poor (random) ansatz initialization [48]. Thus, well-chosen initialization of the ansatz can help avoid
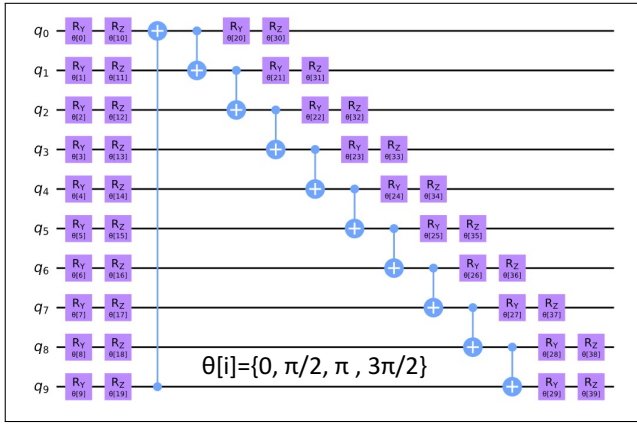
Fig. 10: A Clifford Ansatz - a circuit with only Clifford gates. In this example all fixed components are CXs and the tunable rotational gate angles are multiples of $\pi/2$.

barren plateaus apart from other effects of noise, and therefore enable faster and more accurate convergence on the VQA problem.

A popular and simple approach to construct the initial state for quantum systems is from Hartree-Fock (HF) theory [34]. HF yields a single Slater-determinant approximation to the ground state of a quantum system. While such approximation / optimization problems are generically hard, HF usually rapidly converges to a near-optimal solution, especially for closed-shell molecules at equilibrium geometries [2]. HF yields a quantum state that has no entanglement between the electrons - HF assumes that each electron's motion can be described by a single-particle function (orbital), independent of the instantaneous motion of other electrons. By assuming a single-determinant form for the wavefunction, HF neglects correlation between electrons, which is where classical computing is limited in solving such molecular chemistry problems. Thus, although HF is qualitatively correct in many materials and compounds, it is generally insufficient to make accurate quantitative predictions [4]. Therefore, obtaining accurate results, especially for strongly correlated systems and/or away from equilibrium geometry, necessitate quantum states that go beyond HF. CAFQA is one novel approach to do so, achieving both qualitative and quantitative accuracy.

Clifford Space: Classical simulation of quantum problems usually requires exponential resources (otherwise the need for a quantum computer is obviated) and thus are only useful for simulating small quantum systems. Even using high-performance supercomputers, simulation is restricted to under 100 qubits [10], [18], [39], [51], [80].

However, not all simulation is non-scalable. The Gottesman-Knill's theorem states that, *"Any quantum computer performing only: a) Clifford group gates, b) measurements of Pauli group operators, and c) Clifford group operations conditioned on classical bits, which may be the results of earlier measurements, can be perfectly simulated in polynomial time on a probabilistic classical computer"* [30].

While the Clifford group operations and Pauli group mea-

surements do not provide a universal set of quantum gates there are quantum domains that have applications focused on the Clifford-space including quantum communications / networks [84], error correction codes [66], teleportation [31] and error mitigation [17], [75].

CAFQA explores the benefits of Clifford-only circuits as an ansatz for variational algorithms. An example is shown in Fig.10. While it is a limited discrete tuning space, we observe such an ansatz to be tremendously effective in approaching required accuracy over a variety of applications. While we have restricted the ansatz to strictly the Clifford group in this paper, there is potential to introduce a small number of non-Clifford gates as well [13]. Such an ansatz could provide even greater accuracy before actual execution on the (noisy) quantum device. More details in Section E.

B. CAFQA Step-by-Step Breakdown

Fig.2 provides an illustrative overview of how CAFQA complements traditional variational tuning. The role of CAFQA is illustrated in the red box on the left of the figure and discussed below.

① CAFQA begins with a parameterized circuit in which all fixed gates are Clifford gates. This is usually the case with hardware-efficient ansatz, as described in Section A. In our view, focusing on a hardware-efficient ansatz is justified considering that other ansatz options are generally less suitable to noisy execution on today's NISQ devices. However, extensions beyond this are discussed in Section E.

② Given this parameterized circuit, CAFQA performs a discrete search over the tunable parameters of the circuit. The tunable search space is limited to angles which make the tunable gates strictly Clifford gates. Extensions are discussed in Section E.

③ Since both the fixed gates as well as the tunable gates are Clifford, the resulting circuit in each iteration of the tuning process can be simulated classically, even as the size of the circuits grow (as discussed in Section A).

④ Simulating the circuits corresponding to all the Pauli terms in the Hamiltonian and measuring the Hamiltonian expectation produces the objective function value for the iterative tuning process.

⑤ Explicit electron and spin preservation constrains, if required, can be added to the Hamiltonian [70] or directly to the objective function - CAFQA uses the latter.

⑥ Since simulations are performed classically, they are free of noise - thereby having the potential to eliminate a considerable portion of the noisy variational tuning on the real quantum device which could suffer from noise-induced barren plateaus [86] and other detrimental impacts of error.

⑦ Also worth noting is that only 1-shot simulation is required for each Pauli term / circuit since the expectation value produced by each is strictly +1, -1, or 0 for stabilizer states.

⑧ The discrete search is continued either until convergence of the minimum value obtained or until any specific tuning budget is consumed.
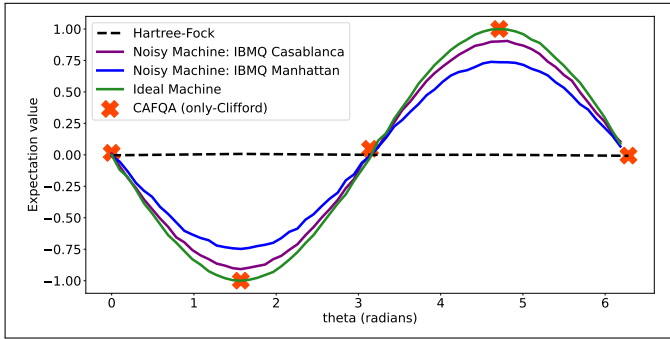
Fig. 11: Comparison of different methods of ansatz tuning for a 2qubit XX Hamiltonian employing an ansatz with only one tunable rotation angle parameter. While the search space for CAFQA is limited, it is able to achieve the global minimum of the ideal machine (which can tune over the entire continuous space). Further, by doing so it is able to outperform the noisy machines, which are limited by noise although they can explore the entire tuning space. On the other hand, the Hartree-Fock method of initialization is unable to recover the expectation value at all, since the XX Hamiltonian has no diagonal component.

⑨ The resulting circuit, with Clifford parameters corresponding to the minimum expectation value observed, is the Clifford ansatz and is ready for traditional VQA optimization. Note that for small problems it is possible that traditional VQA tuning is not required if CAFQA already produces an expectation value which is within the problem's tolerable bounds.

⑩ Traditional VQA tuning on the quantum device is noisy but is able to explore the entire quantum space allowed by the ansatz. The initialization state produced by CAFQA can enable fast and more accurate convergence of traditional VQA tuning.

C. CAFQA benefits on a micro-benchmark

In Fig.11 we use a 2-qubit 'XX' Hamiltonian system and a 2-qubit hardware-efficient ansatz with only one tuning parameter to show the benefits of CAFQA. The Y-axis shows the estimated expectation values of the Hamiltonian while the X-axis sweeps the tuning parameter. Results are discussed below:

① The green line represents tuning the ansatz (i.e., its one tuning parameter) on an ideal noise-free quantum device. Sweeping through all rotations produces an expectation value mimima = -1.0.

② Next, the same tuning is performed on two noisy quantum devices - IBMQ Casablanca and Manhattan (simulated with noise models). These are shown with the purple and blue lines. Clearly, the noisy devices are able to sweep through the entire parameter space, but the effect of noise limits the minima obtained, achieving only -0.7 / -0.85. Note that this microbenchmark is too simplistic to suffer from barren plateaus, etc, but it is expected that the deviation from the ideal / exact minimum will increase with more complex problems and increasing noise.

③ Next, the dashed line shows the expectation value produced by Hartree-Fock (HF) initialization. In this example, HF is unable to produce any useful result since the chosen Hamiltonian does not have any diagonal Pauli terms suited to HF.

④ Finally, CAFQA is shown with orange 'X's. Note that there are only 4 unique points in the Clifford space for the one tunable parameter. Even so, CAFQA is able to produce the expectation value minimum = -1.0 via one of those 4 Clifford points. Not only does the Clifford minimum match the minimum of the entire tuning space, but also the ability to simulate the Clifford space ideally without noise produces more accurate results compared to the noisy devices.

D. Methodology

Ground state energy estimation of molecules: We use VQE to estimate the ground state energy of the following molecules: H_2, LiH, H_2O, H_6, N_2, Cr_2, NaH, H_2-S1 and BeH_2. Hamiltonians are constructed in the STO-3G basis with parity mapping and Z2 symmetry / two qubit reduction. Hamiltonians are constructed for spin corresponding to S=0 which usually have the lowest energy near equilibrium geometries (more on this in Section IV-D). We provide detailed evaluations for the first six and only mean accuracy results for the other three. H_2-S1_STO-3G_singlet (H_2-S1) is obtained from Contextual Subspace VQE [42]. Details about these molecules and their representative Hamiltonians are provided in Table I.

Among these, experiments with quantum computing hardware have been attempted on at least H_2, LiH, H_2O, H_6, NaH and BeH_2 [23]. Note that many of the above were discretized into near-minimal basis sets i.e., STO-3G or similar, or utilized severe approximations to reduce the number of qubits [23].

H_6 is known to be a prototypical strongly correlated molecule, thereby widening the gap between ideal results and classical methods. Also notable is the Chromium Dimer (Cr_2) which has long been a benchmark molecule to evaluate the performance of different computational methods due to its unusual bonding properties in its ground and excited states [83]. This molecule, with its very short, formally sextuple bond, and a peculiarly shaped dissociation curve, has been viewed as a critical test for electronic structure methods [23]. Prior work has discussed that quantum estimations for Cr_2 could take years of computation [80]. Cr_2 is especially challenging to simulate, requiring a system of as many as 72 qubits - we are therefore unable to compare against its ideal energy estimations. Furthermore, we freeze the lower 18 (out of Cr_2's 36 orbitals) to reduce the system to 34 qubits, to ease the burden of iterative tuning given our reasonable yet limited computational resources - but this is not a strict limitation. Freezing lower orbitals is least detrimental to bond dissociation energy estimations - electrons closer to the nucleus are tightly attached and have considerably high ionization energies [3].

For all Hamiltonians above, we use a hardware-efficient SU2 parameterized circuit [1] with one layer of linear entanglement as the ansatz. Different initialization techniques are

Application	Qubits	Bond Length (Eq. geometry)	Bond Length (Range)	Mol Orbitals Total / Used	Exact?	HF?	CAFQA?
H_2	2	0.74 Å	0.37 - 2.96 Å	2 / 2	Yes	Yes	Yes
LiH	4	1.6 Å	0.8 - 4.8 Å	4 / 3	Yes	Yes	Yes
H_2O	12	1 Å	0.5 - 4.0 Å	7 / 7	Yes	Yes	Yes
H_6	10	0.9 Å	0.45 - 3.6 Å	6 / 6	Yes	Yes	Yes
N_2	12	1.09 Å	0.55 - 4.36 Å	10 / 7	Yes	Yes	Yes
Cr_2	34	1.68 Å	1.25 - 3.5 Å	36 / 18	No	Yes	Yes
NaH	12	1.9 Å	0.95 - 7.6 Å	10 / 7	Yes	Yes	Yes
H_2-S1	18	-	-	-	Yes	No	Yes
BeH_2	12	1.32 Å	0.66 - 5.28 Å	7 / 7	Yes	Yes	Yes
MaxCut1	15	-	-	-	Yes	No	Yes
MaxCut2	18	-	-	-	Yes	No	Yes

TABLE I: VQA applications and their characteristics. Columns 4-6 are specific to molecular chemistry. Columns 7-9 indicate if a particular technique was pursued for the application or not.

used in this circuit. An example of this for 10 qubits was shown in Fig.10.

Classical Optimization - MaxCut: While this paper has predominantly focused on VQE for ground state energy estimation targeting molecular chemistry, we also show the benefits of CAFQA for optimization applications like MaxCut. We evaluate CAFQA on two examples of MaxCut with 15 and 18 nodes, respectively. Here we use VQE to solve MaxCut with the same Clifford based SU2 ansatz. CAFQA can also be utilized in QAOA which uses an ansatz with fixed Clifford gates and tunable gates which can be constrained to be Clifford - thus inherently suitable as a Clifford ansatz.

Evaluation Comparisons: We compare the following -

① *CAFQA:* Our proposed approach that uses a Clifford-only ansatz wherein the tunable gates are searched efficiently to find the optimal Clifford parameters, potentially producing the best possible stabilizer state for the target Hamiltonian.

② *Exact:* The exact energy estimations computed classically (but possible only for small problem sizes).

③ *Hartree-Fock (HF):* HF is the best computational basis state for the target Hamiltonian under specified electron and spin preservation constraints. HF has high accuracy for systems which have near equilibrium geometry or have low correlation, but can lose accuracy otherwise since it ignores the correlations between electrons. HF rapidly converges to near-optimal solutions.

Evaluation Metrics: We evaluate CAFQA across four metrics detailed below:

① *Ground State Energy:* Potential energy as a function of nuclear coordinates, as estimated by different techniques, expressed in Hartree units

② *Energy estimation accuracy:* Absolute energy difference between energy estimations from different techniques and exact estimations, expressed in Hartree units. Chemical accuracy region is shown in orange.

③ *Recovered correlation energy:* Percentage of the difference between the Exact energy and the Hartree-Fock limit that is recovered by CAFQA.

④ *Relative accuracy:* Relative energy estimation accuracy between CAFQA and Exact.

Infrastructure: Hartree-Fock estimations are performed via Psi4 [81] while CAFQA evaluations are performed via Qiskit [5]. Qiskit interfaces with the PySCF library [76] in the process of constructing the Hamiltonians from the molecular specifications. The discrete search to find the optimal Clifford gates is performed through Bayesian Optimization via the HyperMapper [56] framework. Classical computations are predominantly carried out on the Google Compute Cloud.

E. Discussion and Future Work

Simulation beyond Cliffords: In CAFQA, we have limited the classical ansatz search to strictly the Clifford space since Clifford circuit simulation is classically efficient. Prior work has shown that efficient classical simulation can be extended beyond strictly Clifford circuits to constrained Clifford+T circuits wherein T refers to the single-qubit 45-degree phase shift [13]. While Clifford+T spans the entire quantum space, the number of T gates that can be added to still suit classical simulation, is constrained. From [13], the runtime is polynomial in the number of qubits and Clifford gates but exponential in the number of T gates. The exponential scaling is sufficiently mild, so classical simulation of Clifford+T circuits with a few hundred qubits and fewer than 50 T gates should be feasible with reasonable compute resources [13]. Other work extends the reach of the Clifford+T gate set to Clifford gates and arbitrary diagonal gates simulating 40-50 qubits and over 60 non-Clifford gates, without high-performance computers [12]. Along these lines, designing an optimal CAFQA ansatz with a mix of Clifford gates, minimal T gates, and appropriate initialization is worth exploration. It should be noted that these prescribed circuit sizes might have to scale down keeping in mind CAFQA's search algorithm runtime. Next, some simulation overheads (as well as VQA tuning overheads in general) can be reduced by employing Hamiltonian term truncation techniques [45], [49] if the truncation does not sacrifice the estimation accuracy or if the

accuracy loss can be bounded within the tolerable limits. Moreover, prior work has explored perturbative expansion of Hartree-Fock computational basis states by exploiting the classical simulability of Clifford circuits [50] (more on this in Section F). Similar expansion to CAFQA-produced stabilizer states can provide further improvements, again benefiting from Clifford simulation.

Beyond a hardware-efficient ansatz: In CAFQA, we have focused on initial parameter search for a hardware-efficient ansatz [41] because: a) the fixed gates in the circuit are Clifford and therefore amenable to classical simulation for appropriately chosen tunable parameters, and b) such an ansatz is suited in width and depth to NISQ devices. But a hardware-efficient ansatz can be limited in its capabilities due to being application agnostic, inefficient in the target Hilbert space coverage, requiring many tuning parameters and prone to barren plateaus [38], [80]. Thus expanding beyond this ansatz can be beneficial if efficiently suited to the CAFQA approach. Expanding the ansatz search to Clifford plus limited non-Clifford gates (discussed above) is a first step in this direction - potentially allowing for a dynamically evolving ansatz structure similar to ADAPT-VQE [32]. Increasing stabilizer rank (i.e., the number of stabilizer states in superposition) to tackle non-Clifford circuits is also worth exploration [12].

Other applications: In this work we have primarily focused on VQE for ground state energy estimation of molecules as well some examples of MaxCut optimization. CAFQA is suited to other applications both in molecular chemistry and outside of it. In molecular chemistry, other properties can be estimated such as Green's functions used to compute electron transport in materials and one-particle reduced density matrices used to compute the electric dipoles of molecules [89]. Excited state energy estimation is one such interesting direction [36], [55], [88]. Outside VQE, CAFQA is also suited to QAOA and potentially for QNNs [8]. Further, to predict the energy and characteristics of complex molecules, machine learning models are trained on very large known molecular data sets and are then used to predict on new molecules [16], [22]. In such proposals the energy estimates for the molecules in the training data set are usually obtained via scalable but less accurate classical methods such as density functional theory (DFT) [43]. The ML model then predicts the energy for new molecules but the error rates are similar to the training data. If the training data can have higher accuracy, the accuracy of the model can correspondingly increase. There is potential to use a CAFQA-style approach to generate training data of high accuracy (even if more limited due to computational cost) to advance existing training sets [16].

Optimization: The discrete search via Bayesian Optimization employed by CAFQA to search through the Clifford space is able to produce high accuracy results (Section V-A) and in a reasonable number of iterations and runtime (Section V-A). While this is clearly efficient for the target problem space, the search is not evaluated on significantly larger problems (our largest system is the partially frozen Cr_2, a 34-qubit system). This is especially important since the Clifford search space

scales exponentially in the number of tunable parameters (alternatively, the number of stabilizer states scales exponentially in the number of qubits [35]). Thus, optimizing the search strategy at the algorithmic as well as implementation levels can reap benefits. Beyond this, exploring the directions discussed above, such as simulation beyond Clifford or targeting an adaptive ansatz, will require more efficient search strategies. Intersections with other machine learning directions are also worth pursuing.

F. Related Work

Ansatz: One of the major hurdles to tuning VQE on quantum devices is the barren plateau problem which is often noise-induced. CAFQA has the potential to avoid such barren plateaus by navigating noise-free through a large part of the Hilbert space through the ideal classical search. Other techniques to address barren plateaus also often restrict the span of the ansatz to a section of the Hilbert space of interest [80]. Adaptive ansatz proposals are such a family of techniques including growing the quantum circuit iteratively [32], incorporating parts of the ansatz into the Hamiltonian by 'dressing' it [71] and variationally updating the gate structure of the entire ansatz [57]. Other work has proposed adapting the molecular Hamiltonians to a chosen ansatz by using a unitary transformation of molecular orbitals [64]. CAFQA can be used atop such proposals for potentially even greater gains.

Initialization: Some prior works have targeted improved initialization for VQAs. [62] proposes initializing ansatz parameters such that subsections of the ansatz do not form a '2-design'. This minimally avoids the optimization process from starting in a barren plateau. [50] proposes a perturbative expansion of the cost function from Hartree-Fock initialization. Being Clifford, the HF initial circuit is classically simulable and thus the first and second derivatives of the cost function can be efficiently calculated. This can produce an initialization state which is very close to the Hartree-Fock state in the Hilbert space but potentially performs better. FLIP [72] proposes initialization with the help of machine learning. It is tailored to learn the structure of successful parameters from a family of related problems which are used as a training set. MetaVQE [15] encodes the Hamiltonian parameters into the first layers of the quantum circuit and attempts to learn the Hamiltonian to obtain a good estimation of the ground state or useful initial states. [87] uses information from approximate classical solutions to construct problem instance-specific circuits, which is an alternative to optimizing the initial parameters. [40] implements the quantum computed moments (QCM) approach - Hamiltonian moments are computed with respect to the Hartree-Fock state and are employed in Lanczos expansion theory to determine an estimate for the ground-state energy which incorporates electronic correlations with high accuracy