Assessing VQE's Viability As a Ground State Energy Solver For Simple Molecules

J.M. Mason

School of Physics, University of Bristol.
(Dated: August 14, 2025)

This project evolved out of the need to find alternative methods in computational chemistry to tackle the fields scaling issues. VQE's performance as a ground state energy solver for small molecules was analysed by comparing 27 different combinations of mappers, ansatze, and optimisers, across both noiseless and noisy simulations. The scalability of the results was then assessed and compared to classical methods like HF and MP2. Under noiseless simulation, UCCSD was found to be the most accurate ansatz, providing a near exact value for H2 (-1.137 Ha), and under noise TwoLocal with the NFT optimiser was found to be the most accurate combination. Under noiseless conditions, VQE out performed HF for all molecules, but only H2 for MP2. Under noisy conditions, VQE was outperformed by both HF and MP2. VQE showed promise for small molecules, but for it to become a viable method in computational chemistry, advancements will have to be made in both the hardware and error mitigation techniques used.

INTRODUCTION

Computational quantum chemistry faces a serious challenge: the scaling of exact methods (like FCI) with system size [1]. To mitigate this problem, classical approximations like CCSD and MP2 can be used for smaller molecules, but they struggle with larger, highly correlated systems. The Variational Quantum Eigensolver (VQE) [2], which combines both quantum and classical computation and is one of the most widely used algorithms in the NISQ era, offers a promising alternative to this problem - however its effectiveness depends on the careful selection of mappers, ansatze, and optimisers. In this project, VQE's viability as a ground state energy finder for simple molecules (H2, LiH, BeH2) was assessed through three key goals: 1. Evaluate the effectiveness of different combinations of mappers (Jordan-Wigner, Parity, Bravyi-Kitaev), anstaze (UCCSD, TwoLocal, EfficientSU2), and optimisers (cobyla, spsa, nft), 2. Assess the scalability of the VQE algorithm with progressively larger molecules, and 3. Compare the VQE results to classical benchmarks like Hartree-Fock (HF), and Moller-Plesset (specifically the second order, MP2).

These molecules were chosen specifically for their simplicity and well understood electronic structure. This enabled clear interpretation of results and well-established classical benchmarks to compare against. The findings of this project discuss tradeoffs made in chemical structure accuracy, computational power, and dealings with noise, all of which are critical moving forward into this new era of computational chemistry.

THEORY

The ability to describe a molecule's structure as a Hamiltonian is one of the pillars of modern computational chemistry. The VQE algorithm builds upon this further by 'mapping' this Hamiltonian into a qubit Hamiltonian. In doing this, the original Hamiltonian is expressed in terms of Pauli operators and is now in a form that can be passed to the chosen ansatz where together with the trial wavefunction prepped by the quantum circuit, an expectation value can be calculated. This expectation value is then minimised by repeatedly optimising the ansatz and re-calculating the expectation value. Thanks to the variational principle, which states that the expectation value of the energy calculated by any trial wavefunction cannot be lower than the true ground state energy of the system, the convergence of the expectation value yields an upper bound to E_{True}. The accuracy depends on the ansatz's ability to represent the true ground state, hence the importance of ansatz selection.

The choice of mappers, ansatze, and optimisers was carefully thought out to give a fair and wide range of combinations to analyse. All three of the chosen mappers (Jordan-Wigner, Parity, and Bravyi-Kitaev) use the same amount of qubits under standard conditions (although this may increase for JW when symmetry

is taken into account), however the qubit connectivity varies. JW requires all-to-all connectivity, whereas Parity only requires linear connectivity, and BK offers a middle ground between the two by making use of tree-like connections. Whilst JW is the simplest to understand, its connectivity requirements make it impractical for hardware. The low connectivity requirements of Parity make it a good choice for NISQ-era devices. BK scales better than both other mappers, but its own connection requirements are more complex, and so is better suited to fault-tolerant devices.

The 3 chosen ansatze consisted of one physically accurate ansatz and two hardware efficient ansatze (UCCSD, TwoLocal, EfficientSU2). UCCSD is a chemistry inspired ansatz and provides the best theoretical results, but due to its high complexity and deep circuits it is impractical on real hardware unless used for small molecules. Of the hardware efficient ansatze, TwoLocal is the simplest, offering shallow circuits and flexibility, making it a good choice for NISQ devices. EfficientSU2 offers a middle ground, remaining hardware efficient but not as simplistic as TwoLocal, and more scalable than UCCSD, making it ideal for medium sized VQE problems. The classical benchmarks used for comparison were chosen for their reproducability across all three of the molecules considered in this project, their suitability for the computing power available for this project, and their reputation as classical methods.

The optimisers were chosen with the requirements that they were strong and established performers, which took different approaches to the same problem, and offered good noise resilience in noisy simulations (particularly NFT for this reason). This was vital as the project addressed both noiseless and noisy simulations, with noiseless simulations being run first to show proof of concept and demonstrate theoretical limits, and noisy simulations being run to examine the effects of errors and to give an insight of what to expect from a quantum device.

RESULTS AND DISCUSSION

All 27 combinations were first run without noise and with H2 as the molecule of concern, as this was the simplest and quickest case to run (Appendix table I). The top 3 results (from both the noiseless and noisy simulations) were then carried forwards and used to achieve values for both LiH and BeH2 when considering scalability and comparison to classical methods. 7 of the 27 combinations achieved a top value of $-1.137306 \, Ha$. All 3 mappers cropped up in the top performers, which was to be expected in a simulation without noise. UCCSD was the top performing ansatz, appearing in 3 of the 7 combinations, again, expected from the most physically accurate ansatz in a noiseless simulation. NFT was the most consistent optimiser, appearing in 4 out of the 7 combinations. This is due to NFT's versatility, as it appeared in combinations with every ansatz and every mapper. An honourable mention goes to the COBYLA optimiser, which achieved the same results as NFT, making up for the remaining 3 combinations, but only with the UCCSD ansatz. This is due to the landscape that UCCSD provides being well suited to COBYLA's approach to optimisation. In response to 7 combinations achieving the same result, 3 were selected that offered variability and opportunity for interesting comparison going forward with the project:

1. Jordan-Wigner + UCCSD + COBYLA: -1.137306 Ha

2. Parity + TwoLocal + NFT: -1.137306 Ha

3. Bravyi-Kitaev + UCCSD + COBYLA: -1.137306 Ha

An energy convergence plot was generated to accompany each of these combinations:

Noise was then added in the form of a depolarising error and all the combinations were re-run (Appendix II), giving the top 3 noisy combinations to be:

1. Bravyi-Kitaev + TwoLocal + NFT: -1.077173 Ha

2. Parity + TwoLocal + NFT: -1.065129 Ha

3. Jordan-Wigner + TwoLocal + NFT: -1.058371 Ha

TwoLocal, being the shallowest ansatz, performed the best under noise [3], in agreement with contemporary research, and NFT, with its natural gradient
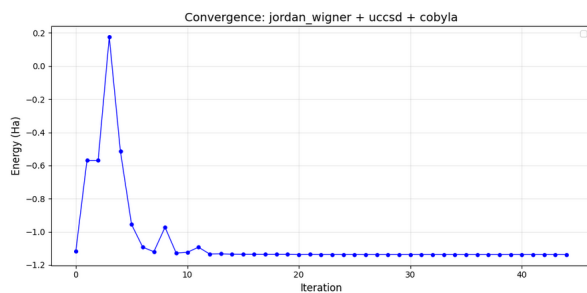
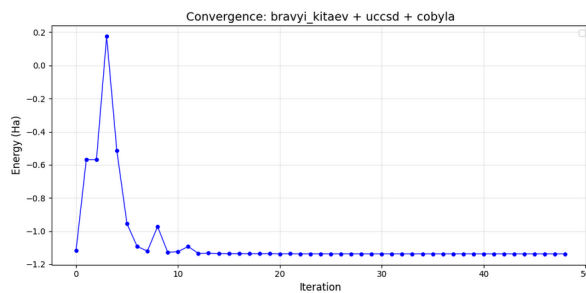FIG. 1: Noiseless energy convergence plot for JW + UCCSD + COBYLA

FIG. 3: Noiseless energy convergence plot for BK + UCCSD + COBYLA
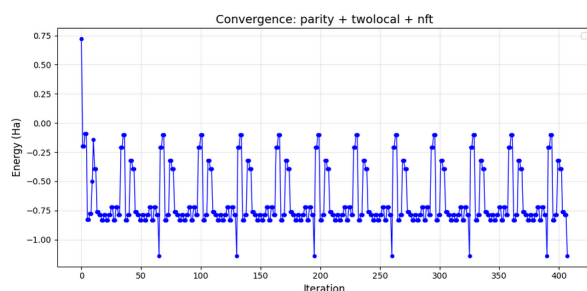

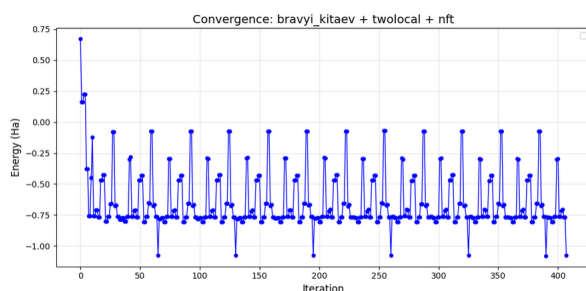FIG. 2: Noiseless energy convergence plot for Parity + TwoLocal + NFT

FIG. 4: Noisy energy convergence plot for BK + TwoLocal + NFT

method, was the most resilient to noise out of the optimisers. The mapper choice didn't seem to have too large of an effect on the results, but it makes sense when dealing with H2 that BK came up as the top mapper and JW as the worst. Accompanying plots were also generated for these noisy results.

In FIG. 1 and FIG. 3, a clear convergence can be seen, occurring well before the the max iteration limit. In the rest of the plots where TwoLocal + NFT are used, an interesting oscillation pattern can be seen in the plot. This is likely due to two factors: 1. The ansatz is over-paramaterised and not well matched to the Hamiltonian - this is a common problem when using TwoLocal as it is a very general ansatz, and the excess parameters create a flat problem landscape which means large jumps in the energy are note very costly and occur more often, and, when combined with the fact that overparamaterisation often creates symmetries in the circuit, the energy can be seen to 'bounce' between states, creat-

ing this spiking effect. 2. For estimators with a finite shot count, like the ones used in this project, each energy evaluation is slightly noisy, and when this is combined with a sensitive optimiser, this can cause oscillations. These effects should be borne in mind when reproducing any of the results from this project or scaling up these same methods to larger molecules, especially for the latter as necessary mitigations will have to be made. For the rest of this project, however, no adjustments were made as reasonable results were achieved given time limitations, leaving these energy oscillations as simply an interesting occurrence to comment upon.

After having achieved 3 top combinations for both the noiseless and noisy simulations, the next consideration of the project was to see how scalable VQE was for larger molecules through these combinations. This was done through comparison of how long each combination took to run for each case (Appendix Tables IV and V). As expected, a drastic increase can be seen in
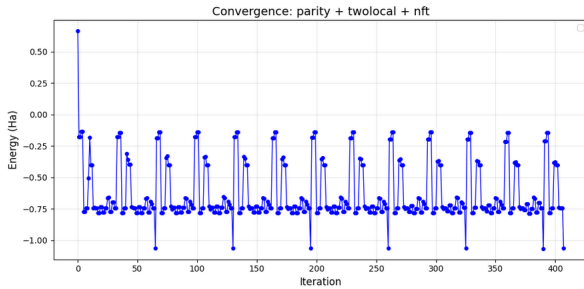
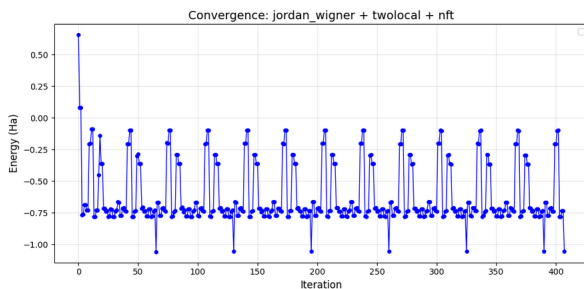FIG. 5: Noisy energy convergence plot for Parity +
TwoLocal + NFT

FIG. 6: Noisy energy convergence plot for JW +
TwoLocal + NFT

the time taken between combinations that fits with the scaling times recognised for VQE. Understandably, in the noiseless comparison, UCCSD took the longest and scaled poorly with it being the most complex ansatz. Additionally, in the noisy results, as molecule size increased, Parity overtook BK as the top performing mapper, due to its linear connectivity making it less susceptible to noise. This fits with the expected results, and, when viewed in conjunction with the rapidly increasing run times for simulations carried out under noise, shows why the Parity mapper is often preferred for large scale VQE projects. Finally, it must be understood by the reader that whilst these simulations offer an insight into VQE's scalability, all these run times would only be exaggerated when run on real quantum hardware.

To complete this project, a final comparison between classical results and the data gathered so far was made, to see if VQE really can offer an alternative to established classical methods. HF was chosen because it

has been a historically relevant and pioneering method. MP2 was chosen as it offered another layer of abstraction on top of HF and more accurate results to compare against.

TABLE I: Reference Energies for Benchmarking (Ha)

Molecule	HF Energy	MP2 Energy
H_2	$-1.116\,759$	$-1.129\,897\,38$
LiH	$-7.861\,865$	$-7.874\,768\,87$
BeH_2	$-15.561\,278$	$-15.583\,854\,50$

Under noiseless conditions, VQE clearly out performs HF, as it should as HF is often used as the reference energy. VQE only out performs MP2 up to LiH, where it is then narrowly beaten, with the gap then widening at BeH2. When noise is added in, VQE fails to out perform either of HF and MP2. This is concurrent with modern findings: VQE is currently unable to out perform classical methods, although it does come close for small molecules when error mitigation techniques are used.

CONCLUSION

Whilst VQE showed some promise for smaller molecules, analysis had to be stopped there due to computational power and time limitations. Given more of both, larger molecules could be explored, and the active space trade-offs that were made such as freezing the core electrons could be foregone. Scalability was analysed in a simple manner, using simply the time taken to run each combination. By instead measuring directly other quantities like the number of Pauli terms or the number of CNOT gates for a problem, a much better analysis could be carried out, giving a greater insight into how VQE scales with molecule size. Classically, the methods available to use were again limited by time and computing power, but neither of these were a concern, a higher standard of comparison could be given in the form of CCSD(T) or even FCI. Regarding the noisy simulations, no error mitigation techniques were used, which although allowed the full effect of the errors to be observed, is not realistic, and would be one of the first points of improvement (like the methods demonstrated in [4]) were this project to be taken any further. This

project has shown that VQE has promise for smaller molecules but still has a long way to go before it is a viable alternative to the established classical methods so far. This is due to the key trade-offs between accuracy, active space size, and noise resilience that currently have to be made. If hardware and error mitigation techniques continues to improve at the current rate, these trade-offs will no longer become a necessity and VQE will be able to be implemented to its full potential.

APPENDIX

REFERENCES

[1] 2011, 'Focus on Computational Chemistry' Future Medicinal Chemistry, 3(8), pp. 909–921. doi: 10.4155/fmc.11.57.

[2] Peruzzo, A., McClean, J., Shadbolt, P. et al., 2014. 'A variational eigenvalue solver on a photonic quantum processor'. Nat Commun 5, 4213 (2014). https://doi.org/10.1038/ncomms5213

[3] Dao, H.L., 2025. Exploring new variational quantum circuit ansatzes for solving SU(2) matrix models. Eur. Phys. J. C 85, 705 (2025). https://doi.org/10.1140/epjc/s10052-025-14322-7

[4] Vovrosh, J. et al., 2021. Simple mitigation of global depolarizing errors in quantum simulations, Physical Review 104, 3, (2021). http://dx.doi.org/10.1103/PhysRevE.104.035309

Molecule	Mapping	Configuration	Energy
		COBYLA	$-1.137\,306$
	UCCSD	SPSA	$-1.137\,209$
		NFT	$-1.116\,999$
		COBYLA	$-1.133\,540$
Jordan-Wigner	TwoLocal	SPSA	$-0.554\,575$
		NFT	$-1.137\,306$
		COBYLA	$-1.131\,098$
	SU2	SPSA	$-0.450\,043$
		NFT	$-0.536\,370$
		COBYLA	$-1.137\,306$
	UCCSD	SPSA	$-1.137\,196$
		NFT	$-1.116\,999$
		COBYLA	$-1.127\,078$
Parity	TwoLocal	SPSA	$-0.540\,611$
		NFT	$-1.137\,306$
		COBYLA	$-1.133\,598$
	SU2	SPSA	$-0.350\,961$
		NFT	$-1.137\,306$
		COBYLA	$-1.137\,306$
	UCCSD	SPSA	$-1.137\,217$
		NFT	$-1.116\,999$
		COBYLA	$-1.131\,029$
Bravyi-Kitaev	TwoLocal	SPSA	$-0.666\,259$
		NFT	$-1.137\,306$
		COBYLA	$-1.126\,672$
	SU2	SPSA	$-0.282\,420$
		NFT	$-0.536\,370$

TABLE II: Noiseless results of all 27 combinations fun on H2

Molecule	Mapping	Configuration	Energy
		COBYLA	−0.927 049
	UCCSD	SPSA	−0.839 582
		NFT	−0.912 454
		COBYLA	−1.056 897
Jordan-Wigner	TwoLocal	SPSA	−0.807 764
		NFT	−1.058 371
		COBYLA	−1.005 339
	SU2	SPSA	−0.912 011
		NFT	−1.006 222
		COBYLA	−1.012 941
	UCCSD	SPSA	−1.005 640
		NFT	−0.991 511
		COBYLA	−1.050 651
Parity	TwoLocal	SPSA	−0.526 145
		NFT	−1.065 129
		COBYLA	−1.006 150
	SU2	SPSA	−0.341 766
		NFT	−1.006 090
		COBYLA	−0.995 846
	UCCSD	SPSA	−0.969 307
		NFT	−0.986 678
		COBYLA	−0.996 259
Bravyi-Kitaev	TwoLocal	SPSA	−0.795 478
		NFT	−1.077 173
		COBYLA	−0.919 020
	SU2	SPSA	−0.387 426
		NFT	−1.039 832

TABLE III: Noisy results of all 27 combinations run on H2

Molecule	Mapping	Configuration	Energy	Time
	Bravyi-Kitaev	TwoLocal + NFT	−1.077 173	3.290 306
H_2	Parity	TwoLocal + NFT	−1.065 129	3.289 715
	Jordan-Wigner	TwoLocal + NFT	−1.058 371	6.289 933
	Bravyi-Kitaev	TwoLocal + NFT	−7.811 283	39.238 196
LiH	Parity	TwoLocal + NFT	−7.818 005	24.435 142
	Jordan-Wigner	TwoLocal + NFT	−7.817 819	38.617 240
	Bravyi-Kitaev	TwoLocal + NFT	−15.436 360	151.314 859
BeH_2	Parity	TwoLocal + NFT	−15.460 218	154.732 811
	Jordan-Wigner	TwoLocal + NFT	−15.455 308	187.289 233

TABLE V: Noisy results of top combinations on all 3 molecules accompanied by run times for each combination

Molecule	Mapping	Configuration	Energy	Time
	Jordan-Wigner	UCCSD + COBYLA	−1.137 306	1.104 636
H_2	Parity	TwoLocal + NFT	−1.137 306	1.893 020
	Bravyi-Kitaev	UCCSD + COBYLA	−1.137 306	1.018 193
	Jordan-Wigner	UCCSD + COBYLA	−7.862 919	11.426 604
LiH	Parity	TwoLocal + NFT	−7.861 865	3.570 738
	Bravyi-Kitaev	UCCSD + COBYLA	−7.862 919	11.028 862
	Jordan-Wigner	UCCSD + COBYLA	−15.566 963	71.177 279
BeH_2	Parity	TwoLocal + NFT	−15.564 611	4.024 350
	Bravyi-Kitaev	UCCSD + COBYLA	−15.567 139	73.891 866

TABLE IV: Noiseless results of top combinations on all 3 molecules accompanied by run times for each combination