

VQE for the J1-J2 model

Niyousha Najmaei
Praveen Viswanathan



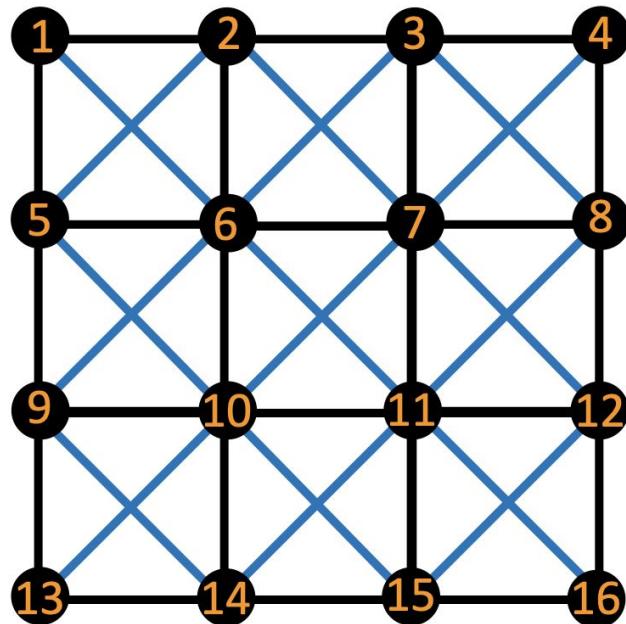
AP3421-PR
Group 3

Delft University of Technology, The Netherlands
January 24, 2023

Overview

- Introduction to J1-J2 model and VQE
- Issues encountered with “ordinary” VQE on J1-J2
- Ideas to tackle these issues
- Results
- Future and outlook

J1-J2 Model



Spins on a square lattice with

- nearest neighbour (J_1), and
- next-to-nearest neighbour (J_2) interactions

These can each be ferromagnetic or antiferromagnetic

The Hamiltonian,

$$\mathcal{H} = -J_1 \sum_{\langle i,j \rangle} \vec{S}_i \cdot \vec{S}_j - J_2 \sum_{\langle\langle i,j \rangle\rangle} \vec{S}_i \cdot \vec{S}_j$$

Open vs Periodic boundaries

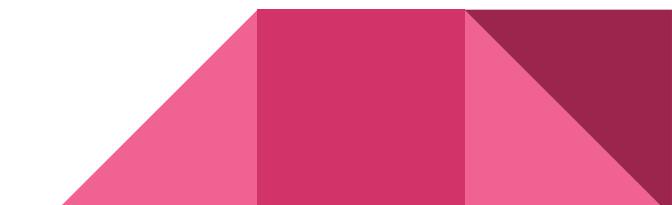


Figure: [1]

J1-J2 Model

- Why J1-J2? Because it can model certain physical systems of interest:
 - Magnetic properties of CuO₂ planes in high-Tc cuprate superconductors: spin ½, with antiferromagnetic signs of both exchange constants^[3]
 - Vanadium oxides: spin ½, exchange interactions not only of the antiferromagnetic type^[3]
- Anti-ferromagnetic, so $J_1, J_2 < 0$
- In our experiments, $\frac{J_2}{J_1} = 0.5$
 - Classical numerical methods are incompatible for $0.4 \gtrsim J_2/J_1 \lesssim 0.6$ ^[1, 2]

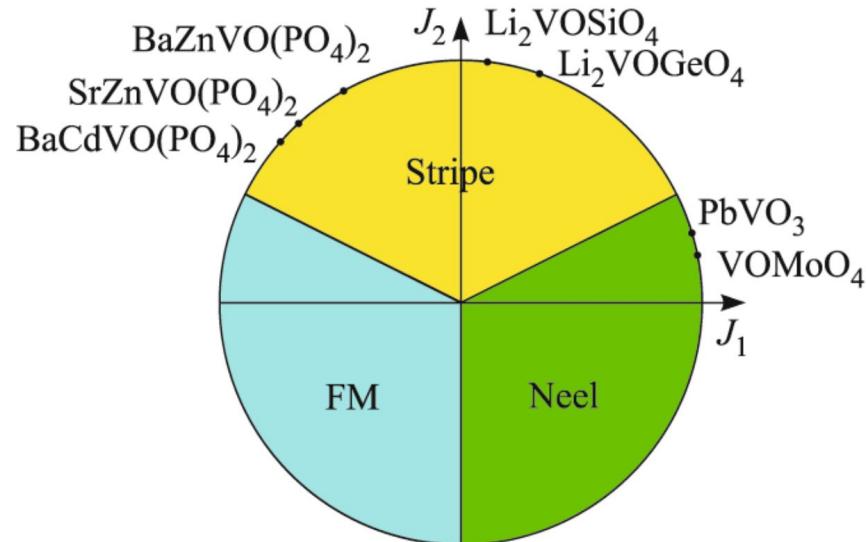
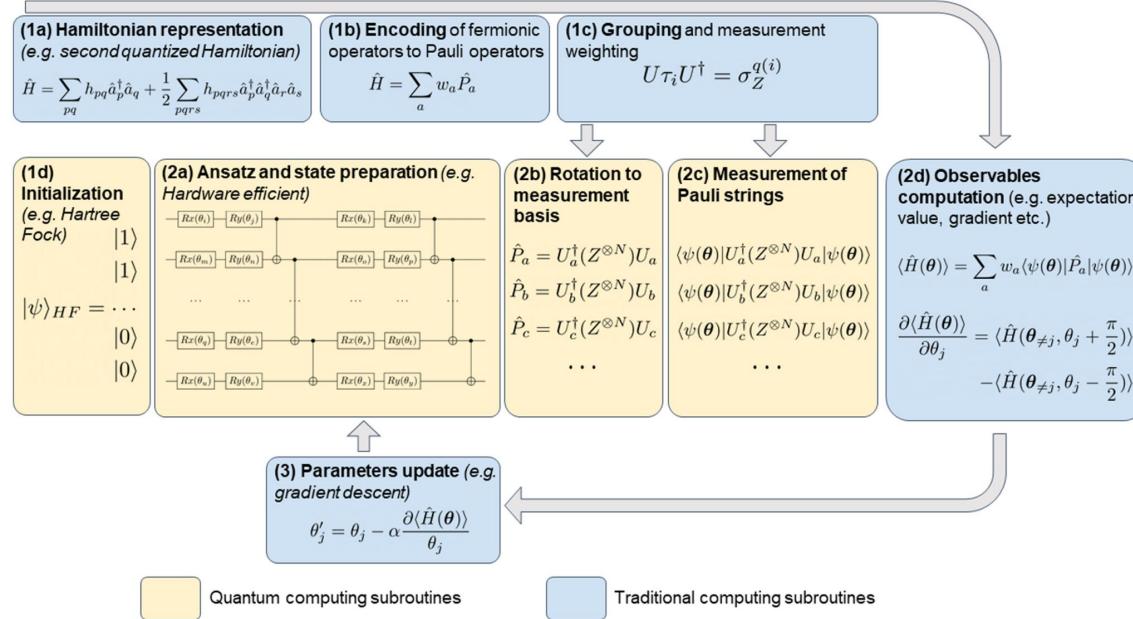


Figure: [3]

VQE: Variational Quantum Eigensolvers



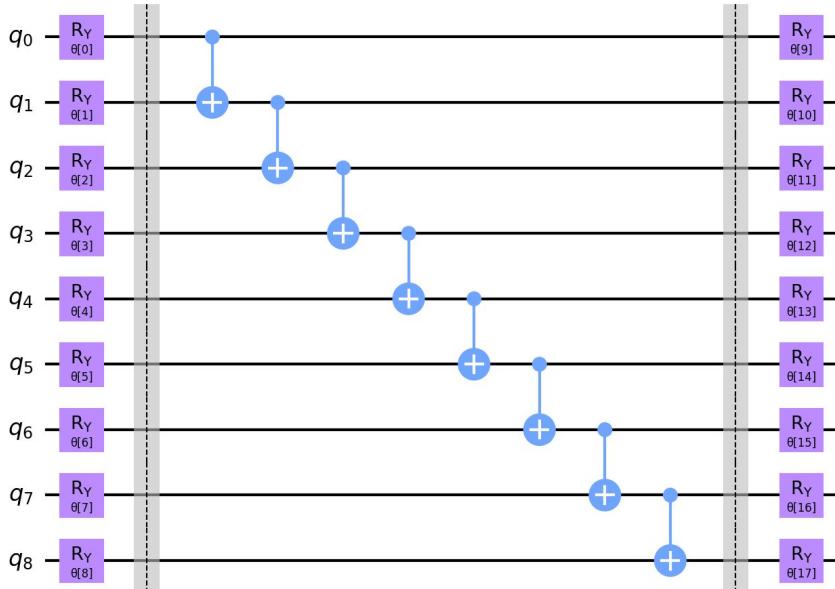
- Diagonalization - infeasible
- Ansatz
 - Only single and two qubit gates
- Global minima, and importance of the initial parameters
- Simulations!
 - Available quantum processors were not large enough :(

Figure: [4]

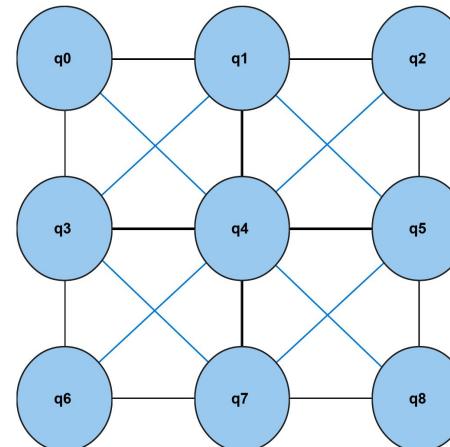
Overview

- Introduction to J1-J2 model and VQE
- Issues encountered with “ordinary” VQE on J1-J2
- Ideas to tackle these issues
- Results
- Future and outlook

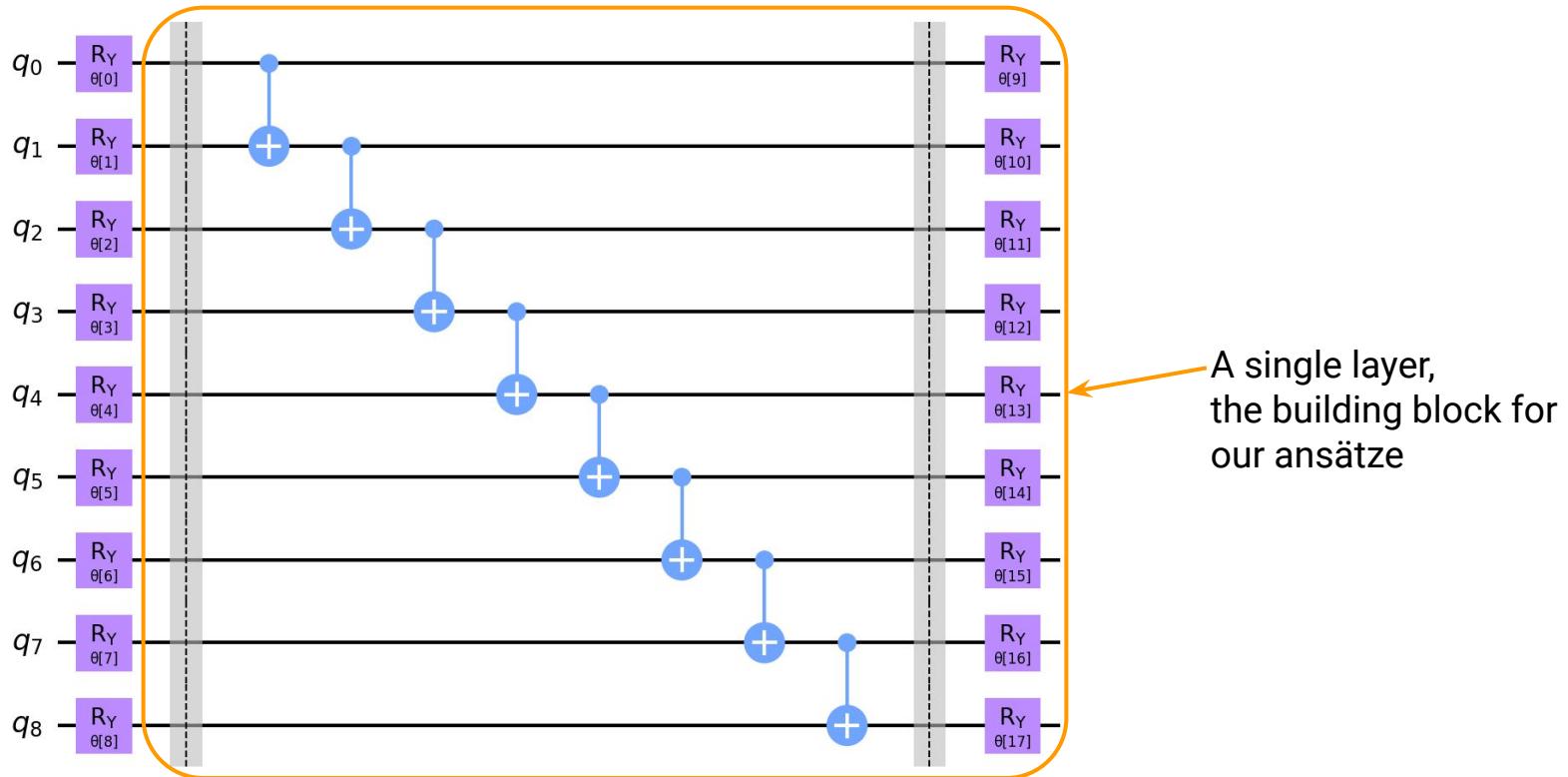
Ansatz 1: Two-Local - for 3x3 lattice



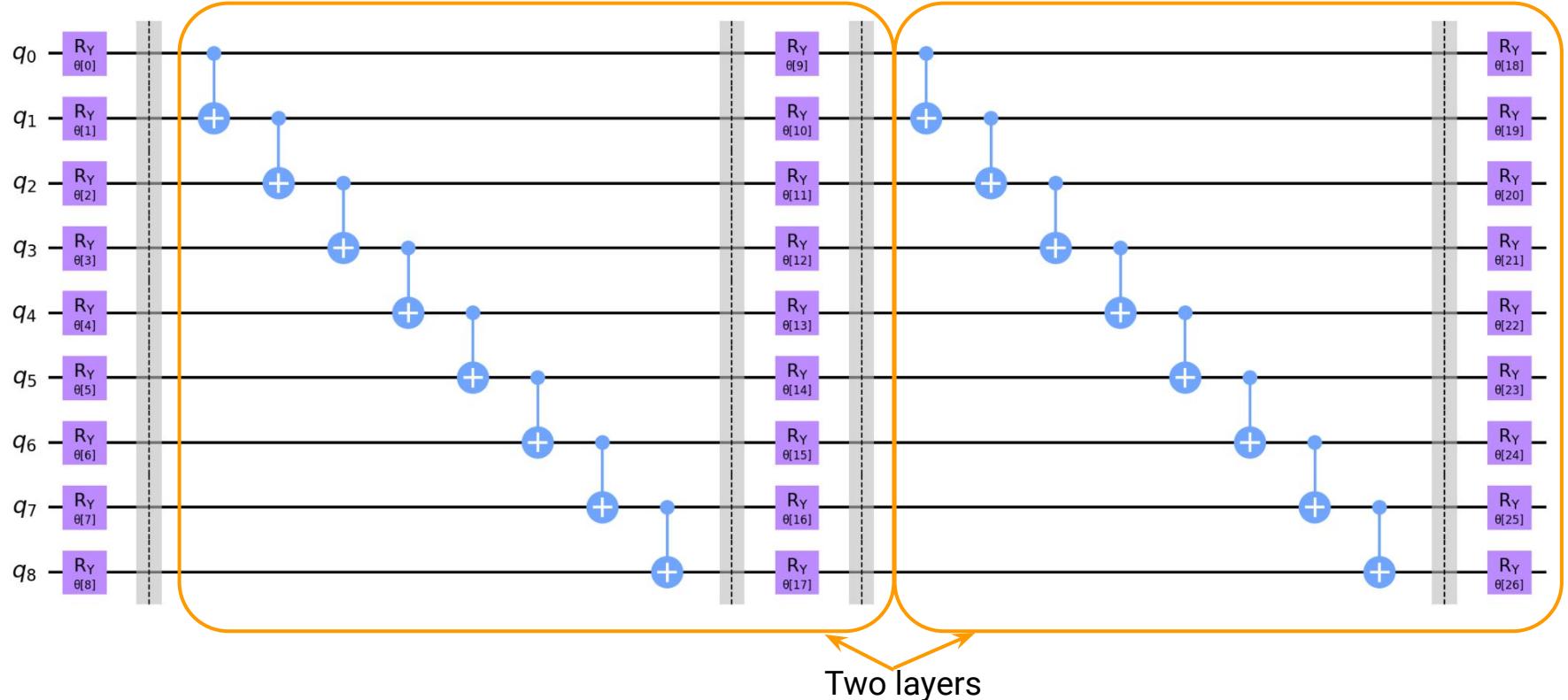
- Simple, hardware-efficient ansatz
- Requires only a 1D qubit array
 - Can be realized with quantum hardware sooner than ansaetze that require 2D qubit arrays
- We ignore the problem symmetry



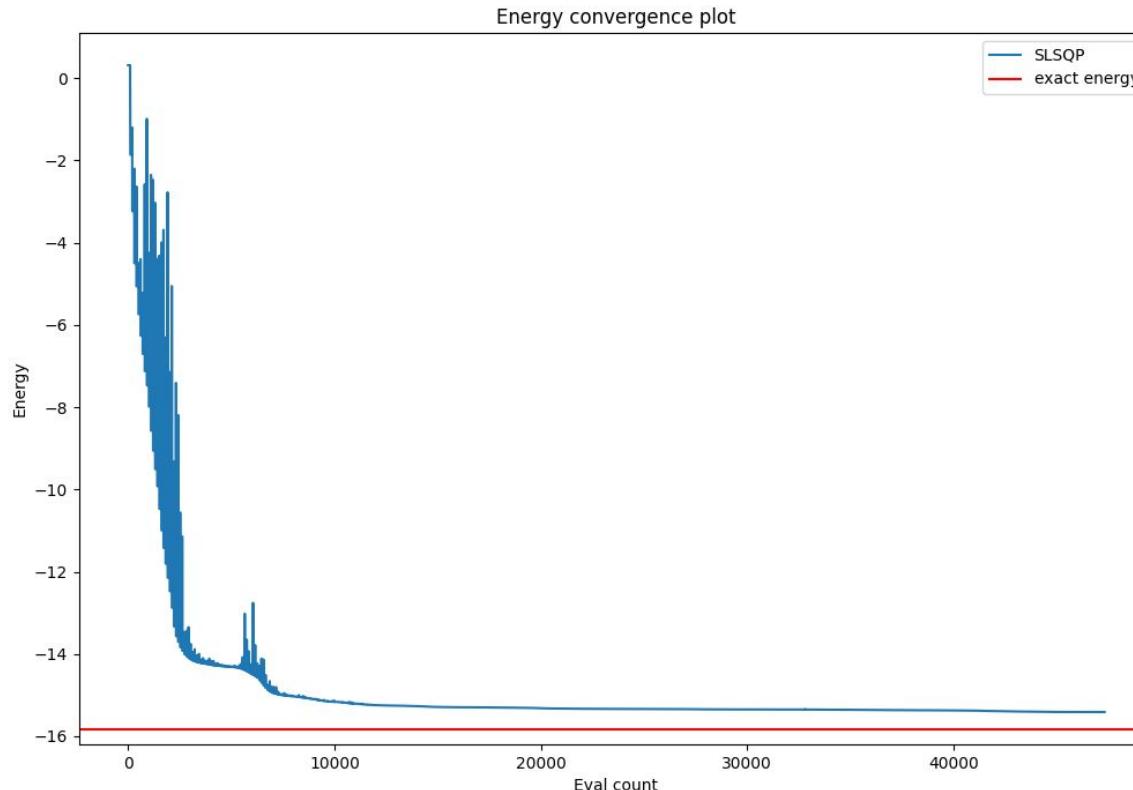
Ansatz 1: Two-Local - layers



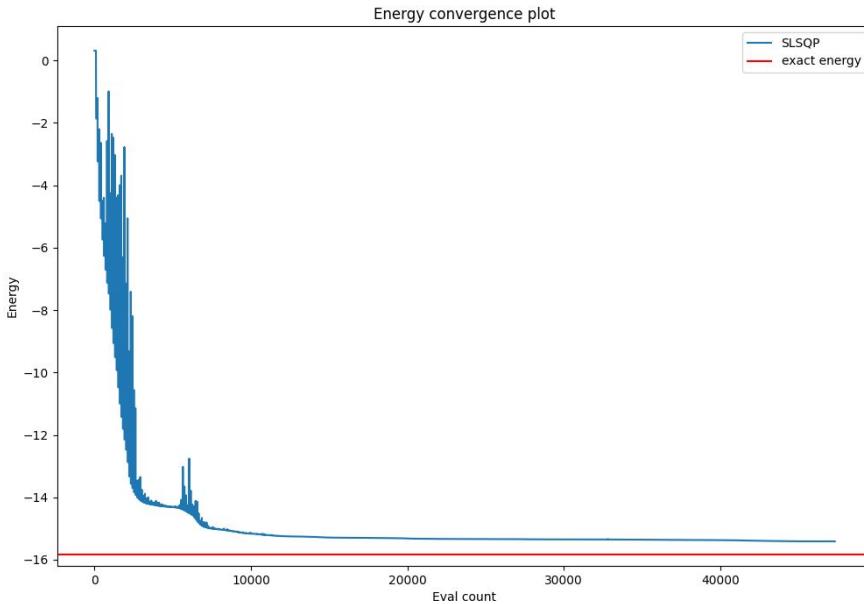
Ansatz 1: Two-Local - layers



Ansatz 1: Two-Local - VQE optimization



Ansatz 1: Two-Local - VQE optimization



For the 3x3 J1-J2 model with **10 layers** of the Two-Local ansatz:

Best VQE estimate obtained: -15.41

Through diagonalization:

Exact ground state energy: -15.837

First excited state energy: -13.086

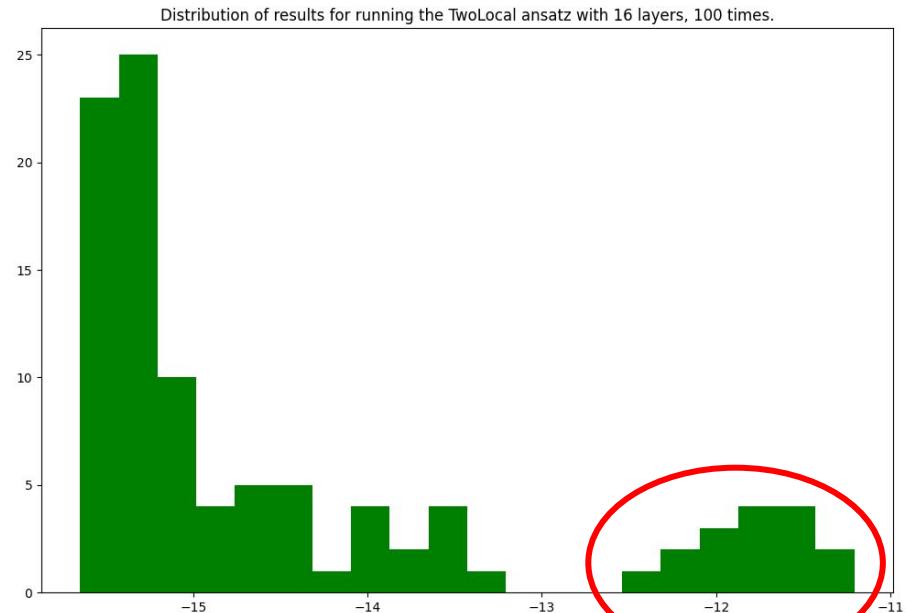
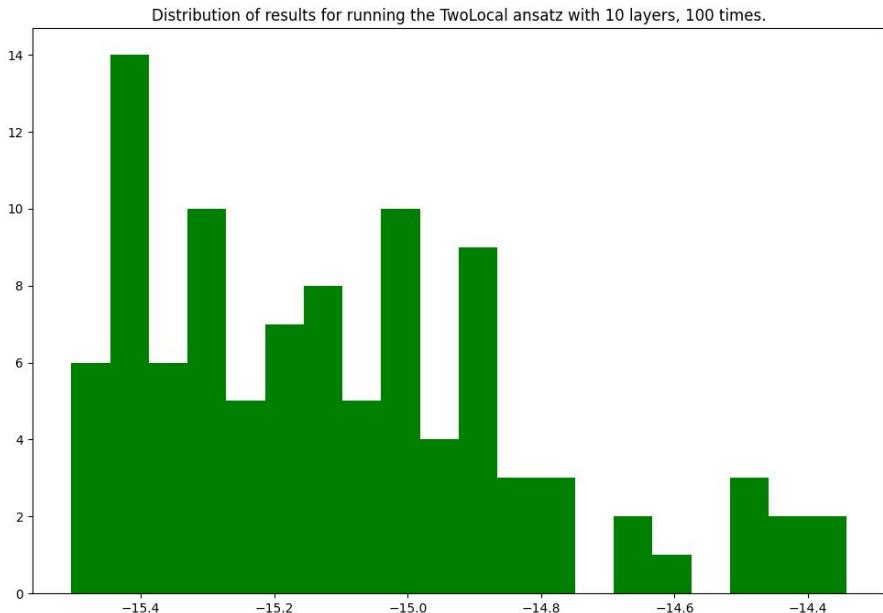
For a small system we would expect better accuracy

Ansatz 1: Two-Local

What if we just use more layers of the Two-Local ansatz?

Ansatz 1: Two-Local - Local Minima

What if we just use more layers of the Two-Local ansatz?
The optimizer gets stuck in local minima!



Ansatz 2: Feulner-Hartmann

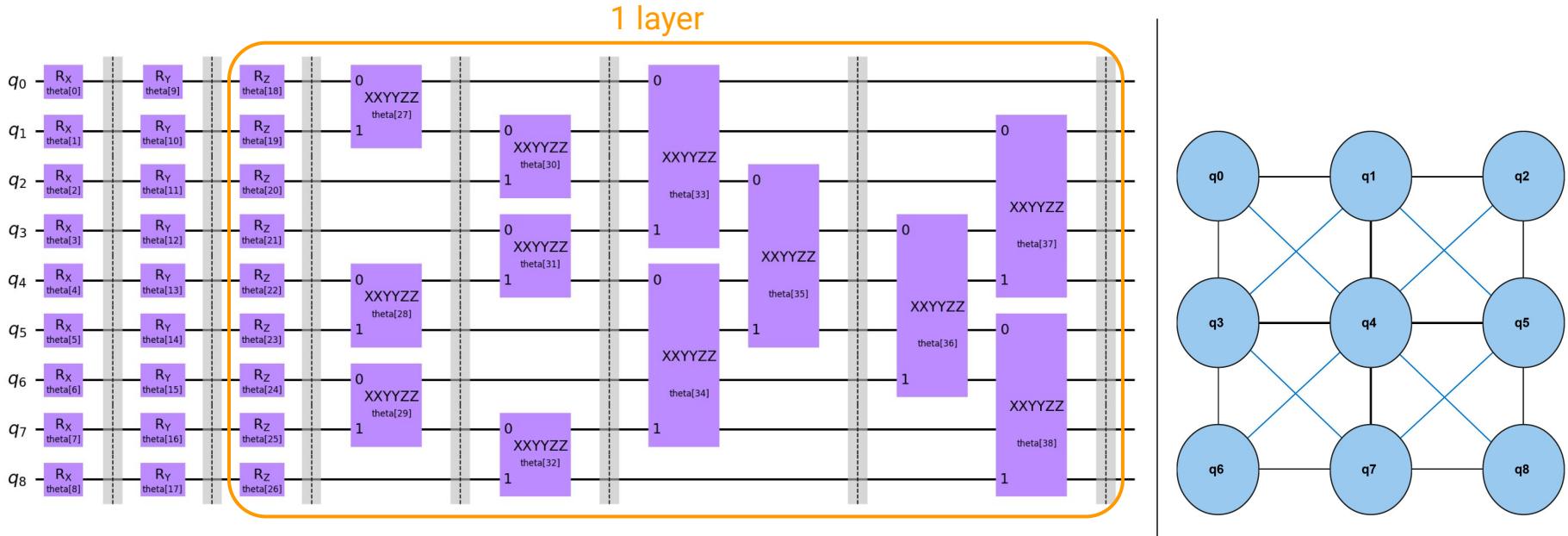
Choice of 2-qubit gate (inspired by Ref. [1])

The diagram illustrates the decomposition of a 4-qubit gate $XXYYZZ(\theta)$ into three 2-qubit gates: $XX(\theta)$, $YY(\theta)$, and $ZZ(\theta)$. On the left, a single box labeled $XXYYZZ(\theta)$ represents the full 4-qubit operation. An equals sign follows this, and then the circuit is shown as three separate boxes in sequence: $XX(\theta)$, $YY(\theta)$, and $ZZ(\theta)$.

$$XX(\theta) = (X \otimes X)^\theta = \begin{pmatrix} c & 0 & 0 & s \\ 0 & c & s & 0 \\ 0 & s & c & 0 \\ s & 0 & 0 & c \end{pmatrix}$$
$$YY(\theta) = (Y \otimes Y)^\theta = \begin{pmatrix} c & 0 & 0 & -s \\ 0 & c & s & 0 \\ 0 & s & c & 0 \\ -s & 0 & 0 & c \end{pmatrix}$$
$$ZZ(\theta) = (Z \otimes Z)^\theta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & w & 0 & 0 \\ 0 & 0 & w & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

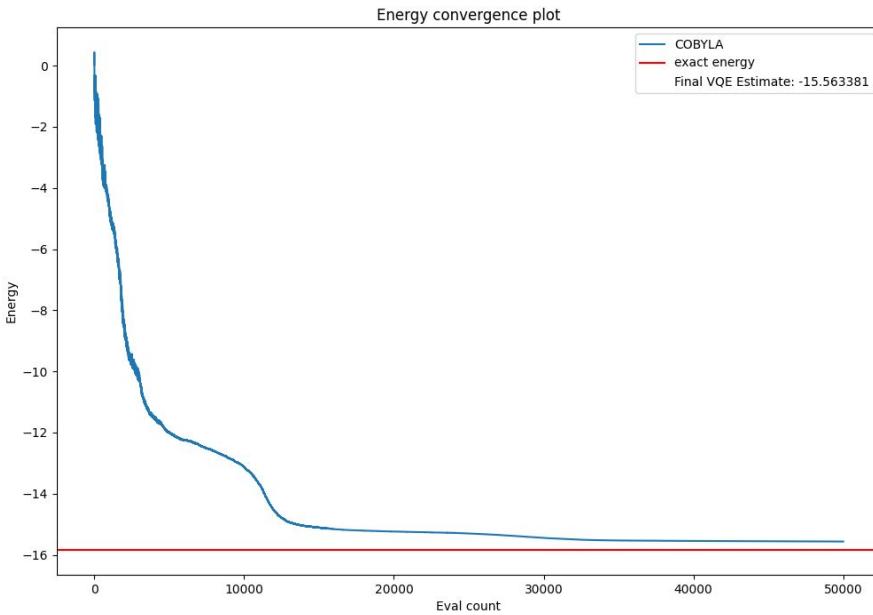
Figure: [1]

Ansatz 2: Feulner-Hartmann - for 3x3 lattice



- Incorporates problem symmetry (to some extent)
- 2-qubit gates only between adjacent qubits, not between diagonal qubits

Ansatz 2: Feulner-Hartmann - VQE optimization



For the 3x3 J1-J2 model with **7 layers** of the Feulner-Hartmann ansatz:

Best VQE estimate obtained: -15.563

Through diagonalization:

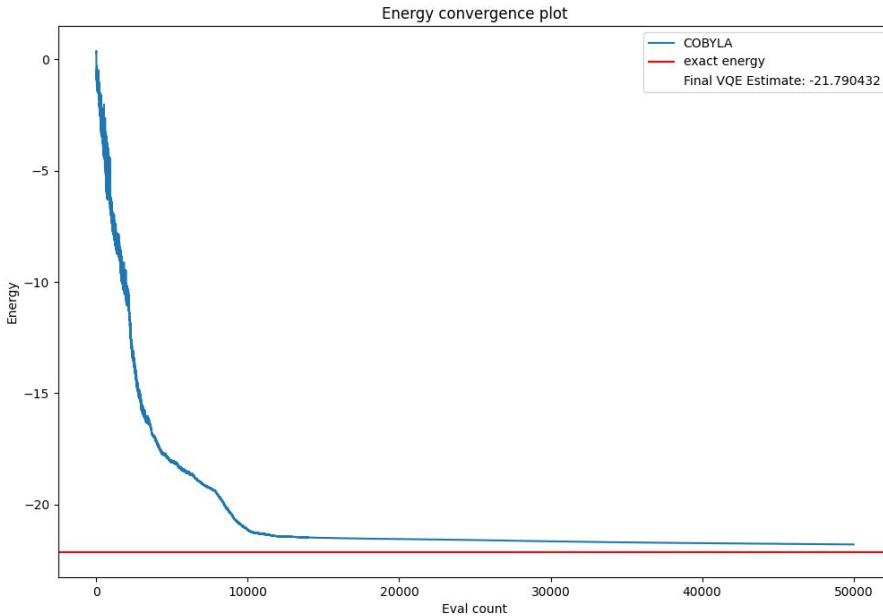
Exact ground state energy: -15.837

First excited state energy: -13.086

$$\frac{|E_0 - E_{VQE}|}{|E_0 - E_1|} \approx 10\%$$



Ansatz 2: Feulner-Hartmann - VQE optimization



For the 3x4 J1-J2 model with **7 layers** of the Feulner-Hartmann ansatz:

Best VQE estimate obtained: -21.790

Through diagonalization:

Exact ground state energy: -22.138

First excited state energy: -20.156

$$\frac{|E_0 - E_{VQE}|}{|E_0 - E_1|} \approx 19\%$$



Summary of Issues

- Getting stuck in local minima for large ansätze with high number of parameters
- Not good enough convergence even for large ansätze.
Can we perform better?
- Feulner-Hartmann ansatz has too many parameters. E.g. 265 parameters for 3x3 and 227 for 3x4 lattice
 - Takes a long time for optimization ~4 hours for the 3x4 lattice

	Ansatz	Result	#params
3x3 Exact: -15.837	Feulner-Hartmann with 7 layers	-15.801	165
	Two-Local with 17 layers	-11.913	162
	Two-Local with 10 layers	-15.412	99

3x4 Exact: -22.180	Feulner-Hartmann with 7 layers	-21.790	227
	Feulner-Hartmann with 7 layers in [1]	-21.980	227
	Two-Local with 18 layers	-20.015	228

Comparison between Feulner-Hartmann with 7 layers and Two Local ansatz with almost the same number of parameters for 3x3 and 3x4 lattices with $J_2=0.5$ and $J_1=1$. For the 3x3 case, the Two-Local ansatz with optimum number of parameters (10 layers) is given.

Overview

- Introduction to J1-J2 model and VQE
- Issues encountered with “ordinary” VQE on J1-J2
- Ideas to tackle these issues
- Results
- Future and outlook

Idea 1

Sequential Layer Addition:

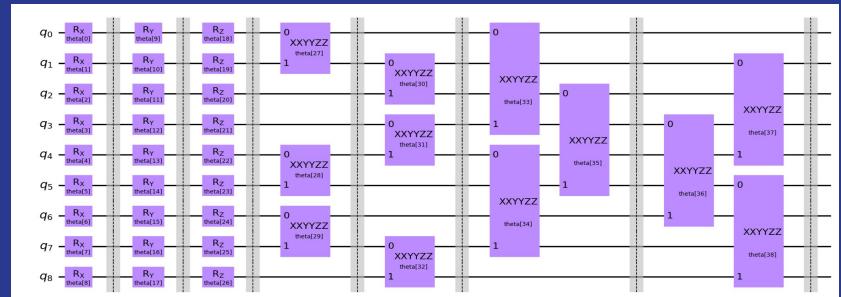
Start with $k < n$ layers and add one layers every t iterations until the number of layers is n .

Our hypothesis:

1. Lower probability of the optimizer getting stuck in local minima
2. Same (or Better?) results with lower number parameters (on average)

A similar approach for VQLS by Patil et al., 2022 [6]

Iteration = 0



Idea 1

Sequential Layer Addition:

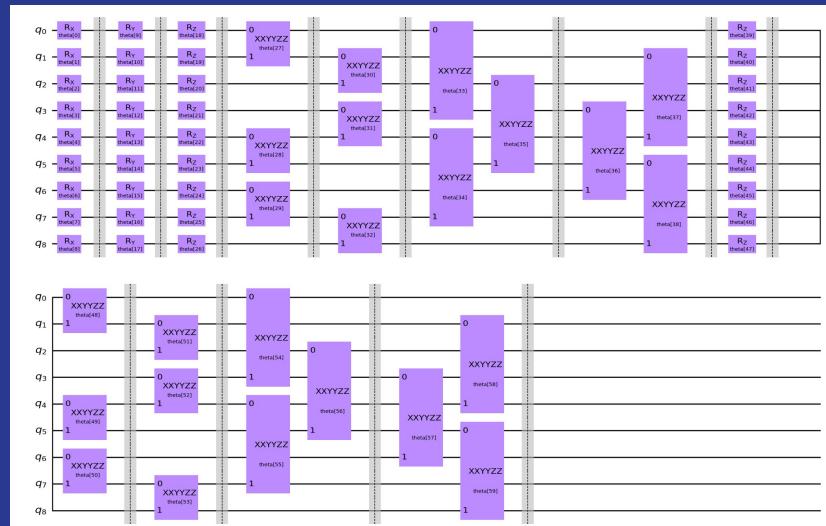
Start with $k < n$ layers and add one layers every t iterations until the number of layers is n .

Our hypothesis:

1. Lower probability of the optimizer getting stuck in local minima
2. Same (or Better?) results with lower number parameters (on average)

A similar approach for VQLS by Patil et al., 2022 [6]

Iteration = t



Idea 1

Sequential Layer Addition:

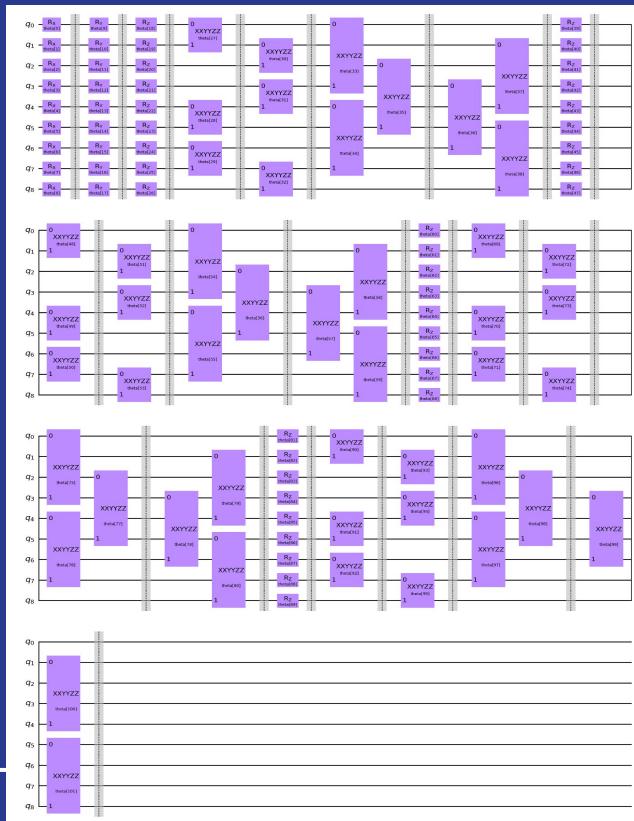
Start with $k < n$ layers and add one layers every t iterations until the number of layers is n .

Our hypothesis:

1. Lower probability of the optimizer getting stuck in local minima
2. Same (or Better?) results with lower number parameters (on average)

A similar approach for VQLS by Patil et al., 2022 [6]

Iteration = 2t



Idea 2

Gradient Based Gate Addition:

Start with k iterations ($k=1$ or $k=2$). Add gates which are among the α percent with the highest average gradient of the parameter, to the end of the circuit every t iterations.

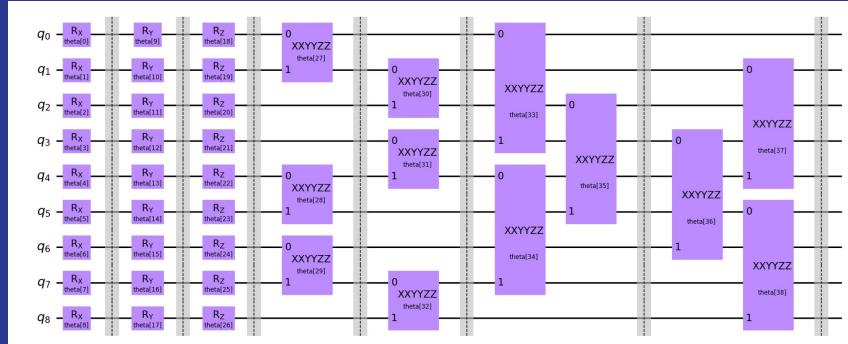
Our hypothesis:

1. The higher the gradient of the parameter, the more that gate is contributing to the convergence
2. Same (or Better?) results with lower number parameters
3. It is possible to start with a single layers and let it grow

Loosely based on Grimsley et al., 2019 [7]

Iteration = 0 , $\alpha = 10\%$

$10.1 \times 39 \rfloor = 7$, thus adding 7 gates



Idea 2

Gradient Based Gate Addition:

Start with k iterations ($k=1$ or $k=2$). Add gates which are among the α percent with the highest average gradient of the parameter, to the end of the circuit every t iterations.

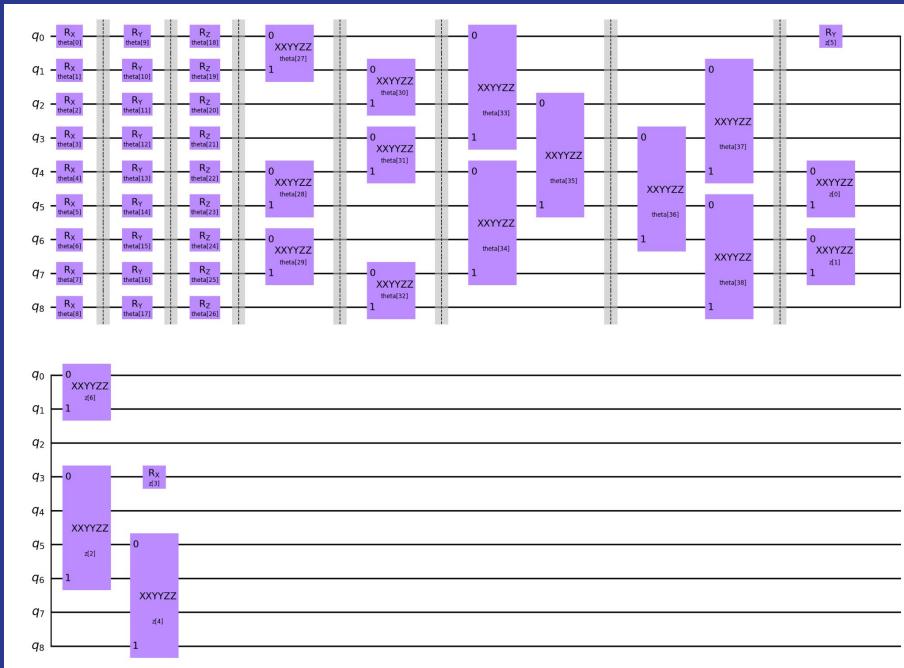
Our hypothesis:

1. The higher the gradient of the parameter, the more that gate is contributing to the convergence
2. Same (or Better?) results with lower number parameters
3. It is possible to start with a single layers and let it grow

Loosely based on Grimsley et al., 2019 [7]

$\text{Iteration} = t, \alpha = 10\%$

$\lfloor 0.1 * 39 \rfloor = 7$, thus adding 7 gates



Idea 2

Gradient Based Gate Addition:

Start with k iterations ($k=1$ or $k=2$). Add gates which are among the α percent with the highest average gradient of the parameter, to the end of the circuit every t iterations.

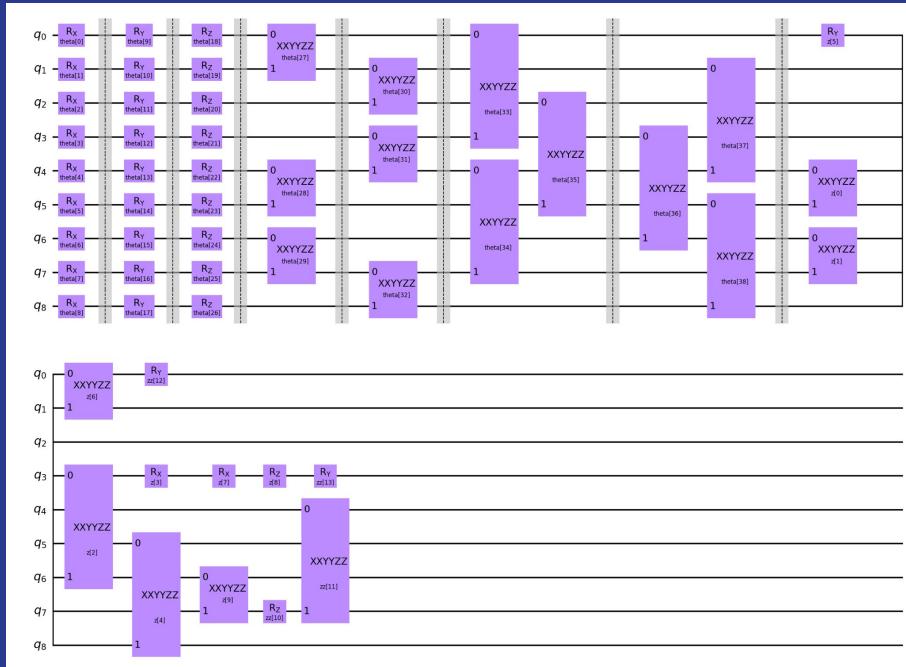
Our hypothesis:

1. The higher the gradient of the parameter, the more that gate is contributing to the convergence
2. Same (or Better?) results with lower number parameters
3. It is possible to start with a single layers and let it grow

Loosely based on Grimsley et al., 2019 [7]

$\text{Iteration} = 2t, \alpha = 10\%$

$\lfloor 0.1 * 39 \rfloor = 7$, thus adding 7 gates



Idea 2

Gradient Based Gate Addition:

Start with k iterations ($k=1$ or $k=2$). Add gates which are among the α percent with the highest average gradient of the parameter, to the end of the circuit every t iterations.

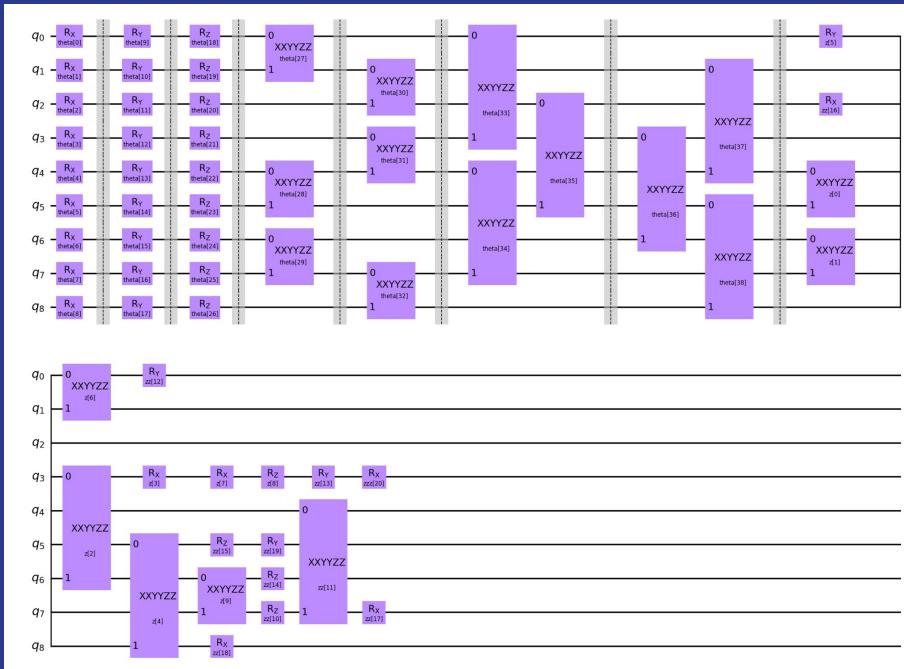
Our hypothesis:

1. The higher the gradient of the parameter, the more that gate is contributing to the convergence
2. Same (or Better?) results with lower number parameters
3. It is possible to start with a single layers and let it grow

Loosely based on Grimsley et al., 2019 [7]

$\text{Iteration} = 3t, \alpha = 10\%$

$\lfloor 0.1 * 39 \rfloor = 7$, thus adding 7 gates



Idea 2

Gradient Based Gate Addition:

Start with k iterations ($k=1$ or $k=2$). Add gates which are among the α percent with the highest average gradient of the parameter, to the end of the circuit every t iterations.

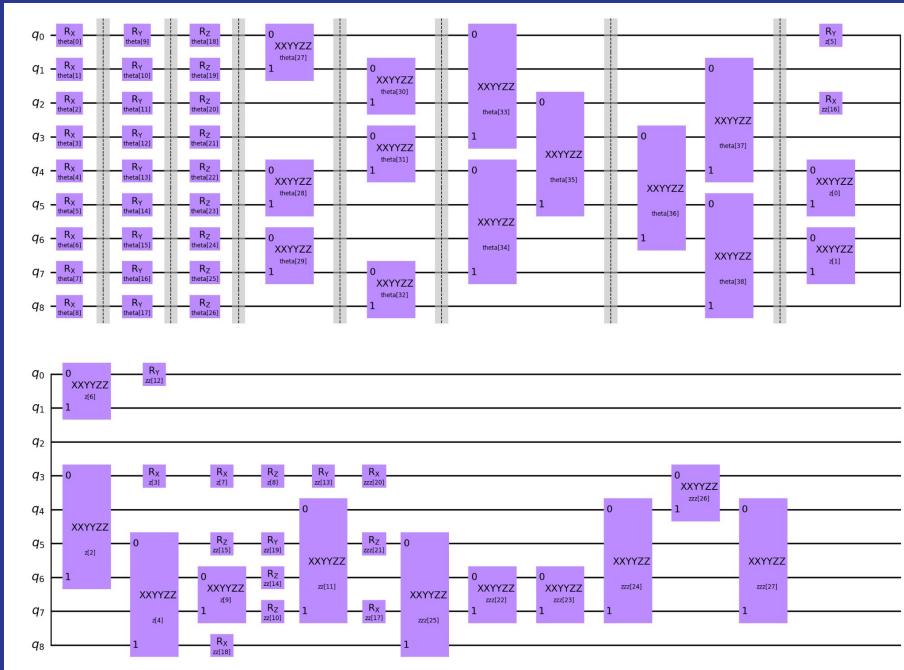
Our hypothesis:

1. The higher the gradient of the parameter, the more that gate is contributing to the convergence
2. Same (or Better?) results with lower number parameters
3. It is possible to start with a single layers and let it grow

Loosely based on Grimsley et al., 2019 [7]

$\text{Iteration} = 4t, \alpha = 10\%$

$\lfloor 0.1 * 39 \rfloor = 7$, thus adding 7 gates



Iteration = $5t$, $\alpha = 10\%$

Idea 2

Gradient Based Gate Addition:

Start with k iterations ($k=1$ or $k=2$). Add gates which are among the α percent with the highest average gradient of the parameter, to the end of the circuit every t iterations.

Our hypothesis:

1. The higher the gradient of the parameter, the more that gate is contributing to the convergence
2. Same (or Better?) results with lower number parameters
3. It is possible to start with a single layers and let it grow

Loosely based on Grimsley et al., 2019 [7]

$10.1 \times 39 \downarrow = 7$, thus adding 7 gates



Iteration = $6t$, $\alpha = 10\%$

Idea 2

Gradient Based Gate Addition:

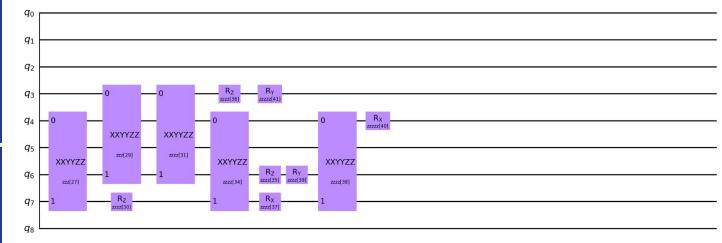
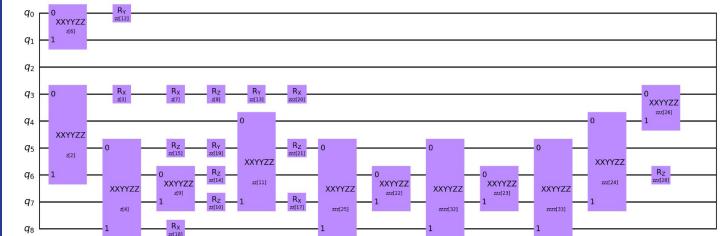
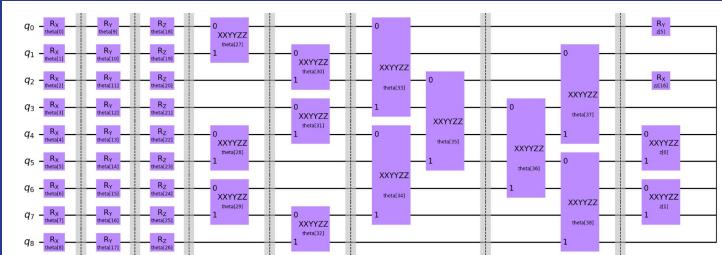
Start with k iterations ($k=1$ or $k=2$). Add gates which are among the α percent with the highest average gradient of the parameter, to the end of the circuit every t iterations.

Our hypothesis:

1. The higher the gradient of the parameter, the more that gate is contributing to the convergence
2. Same (or Better?) results with lower number parameters
3. It is possible to start with a single layers and let it grow

Loosely based on Grimsley et al., 2019 [7]

$10.1 \times 39 = 7$, thus adding 7 gates



Iteration = $7t$, $\alpha = 10\%$

Idea 2

Gradient Based Gate Addition:

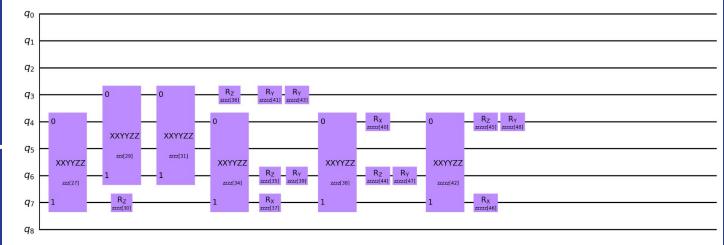
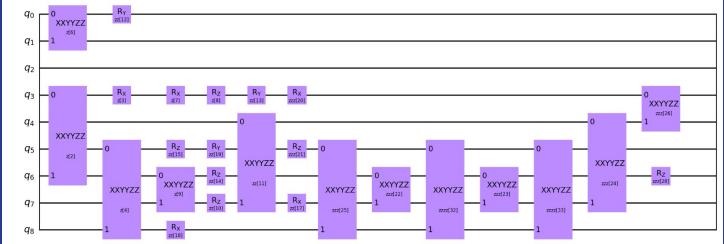
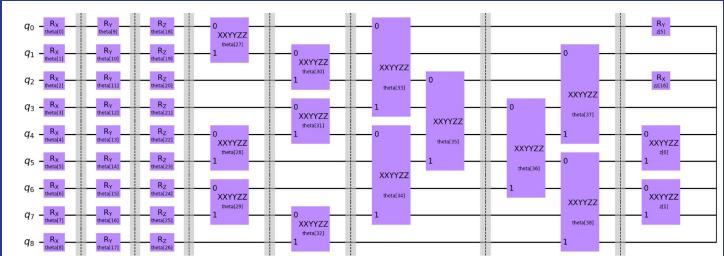
Start with k iterations ($k=1$ or $k=2$). Add gates which are among the α percent with the highest average gradient of the parameter, to the end of the circuit every t iterations.

Our hypothesis:

1. The higher the gradient of the parameter, the more that gate is contributing to the convergence
2. Same (or Better?) results with lower number parameters
3. It is possible to start with a single layers and let it grow

Loosely based on Grimsley et al., 2019 [7]

$10.1 \times 39 = 7$, thus adding 7 gates



Overview

- Introduction to J1-J2 model and VQE
- Issues encountered with “ordinary” VQE on J1-J2
- Ideas to tackle these issues
- Results
- Future and outlook

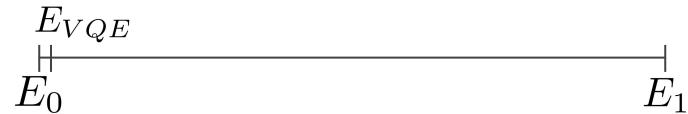
Dynamic Ansatz Results

	Static/Dynamic	Result	Final #params	Avg. #params
3x3 Exact: -15.837	Static Ansatz	-15.801	165	165
	Sequential Layer Addition (up to 7)	-15.836	165	102
	Gradient Based Gate Addition (Starting from two, $\alpha = 20\%$)	-15.645	144	102

	Static/Dynamic	Result	Final #params	Avg. #params
3x4 Exact: -22.180	Static Ansatz	-21.790	227	227
	Sequential Layer Addition (up to 7)	-22.130	227	140
	Gradient Based Gate Addition (Starting from two, $\alpha = 20\%$)	-21.984	146	114

Comparison between static and dynamic ansatz using Sequential Layer Addition and Gradient Based Gate Addition with Feulner-Hartmann with 7 layers for 3x3 and 3x4 lattices with $J_2=0.5$ and $J_1=1$

For Sequential Layer Adding up to 7 layers: $\frac{|E_0 - E_{VQE}|}{|E_0 - E_1|} \approx 2.5\%$



Limits of Sequential Layer Adding

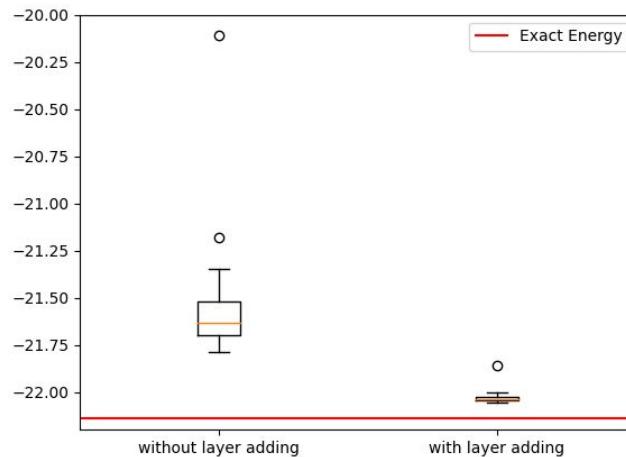
Using Sequential Layer Adding for 3x4 lattice and adding layers up to 4 layer, we still get better results than using the static ansatz with 227 parameters.

	Sequential Layer Adding	Result	Final #params	Avg. #params
3x4 Exact: -22.138	Up to 7 layers	-22.130	227	140
	Up to 6 layers	-22.120	198	125.5
	Up to 5 layers	-22.085	169	111
	Up to 4 layers	-22.024	140	96.5
	Up to 3 layers	-20.105	111	82
	Up to 2 layers	-11.621	82	67.5
	Static Ansatz (7 layers)	-21.790	227	227
	Static Ansatz (4 layers)	-21.667	140	140

Analysis of results

Static vs dynamic ansatz with Sequential Layer Adding

- Running each one 30 times with different initial values for the optimizer



	Static/Dynamic	mean	SD
3x4 Exact: -22.180	Static Ansatz	-21.560	0.301
	Sequential Layer Addition (up to 7)	-22.030	0.034

Comparison between the mean and standard deviation of the results obtained by running the static and dynamic ansatz with Sequential Layer Adding 30 times each for a 3x4 lattice

Analysis of results

Static vs dynamic ansatz with Sequential Layer Adding

- Running each one 30 times with different initial values for the optimizer

Comparing the two distributions with Welch's t-test (distributions with different variances):

- The 30 runs with the dynamic ansatz using Sequential Layer Adding ($M = -21.560$, $SD = 0.301$) compared to the 30 runs with the static ansatz ($M = -22.030$, $SD = 0.034$) demonstrated significantly better estimations of the ground state energy, $t(58) = -8.355$, $p = 2.689e-9$
- **The results are “highly significant” since p-value < 0.01**

Overview

- Introduction to J1-J2 model and VQE
- Issues encountered with “ordinary” VQE on J1-J2
- Ideas to tackle these issues
- Results
- Future and outlook

Summary and Outlook

- Sequential Layer Adding:
 - Significantly improves VQE estimates, with fewer number of parameters
 - Not so dependant on the initial parameter values
 - Potentially a tool for avoiding local minima
 - At the end, still requires optimization of a large number of parameters.
- Gradient Based Gate Addition:
 - Great estimates with fewer parameters
 - Potentially a tool for building ansätze given just the building blocks
 - May not be useful when the building blocks are non-entangling e.g.: TwoLocal ansatz.
 - Does it become less effective after a certain number of gate additions?
 - It does not give results as good as sequential layer adding, even if we allow it similar number of parameters. Can it be improved?

Future Work

- Idea 3: Gradient Based Gate Deletion
 - Similar to gradient based gate addition but removes β percent of gates with the smallest average gradient of parameters every t iterations
- Other values of J_2/J_1
- Periodic hamiltonian
- Incorporate noise!
- Experiment with Gradient Based Gate Addition for the Two-Local ansatz as well

References

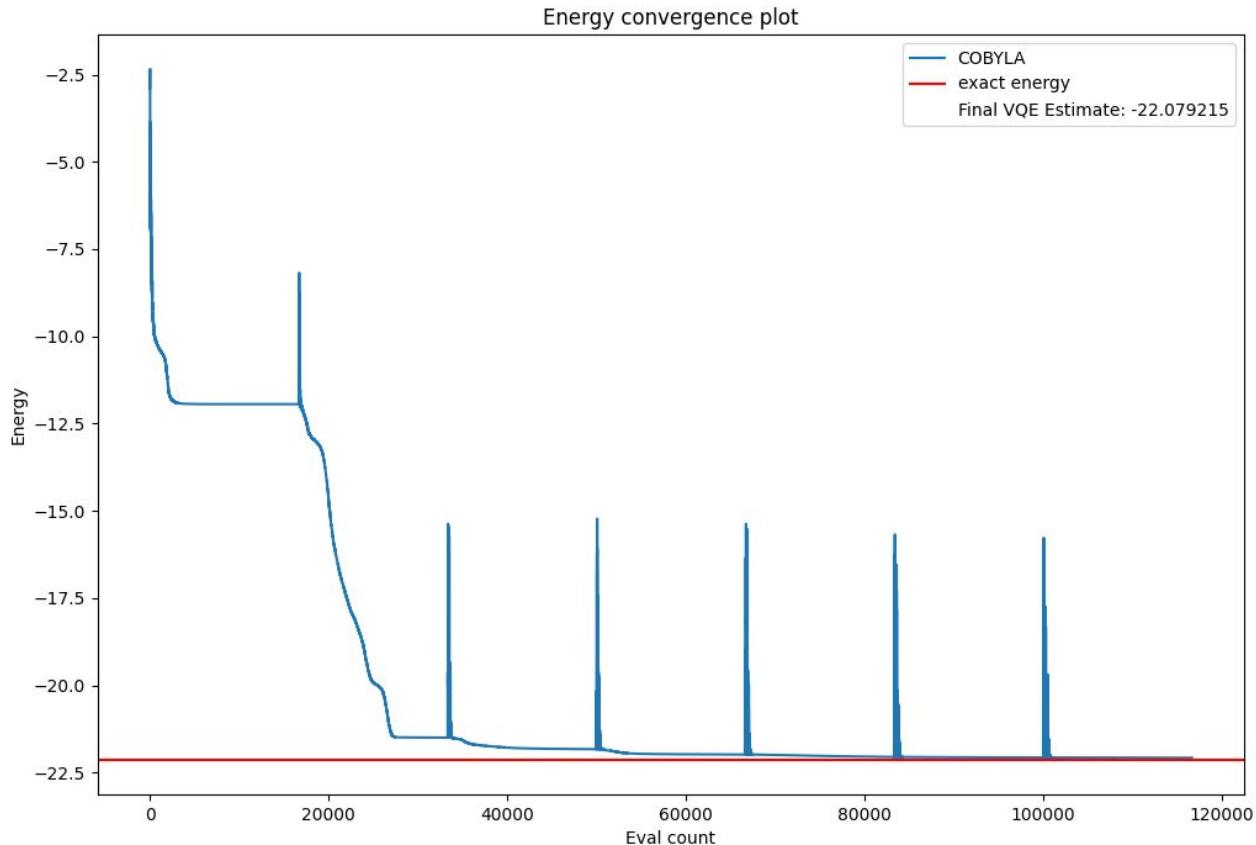
1. V. Feulner and M. J. Hartmann, "Variational quantum eigensolver ansatz for the J1-J2 model," *Phys. Rev. B*, vol. 106, p. 144 426, 14 Oct. 2022
2. Roscher, D., Gneist, N., Scherer, M.M., Trebst, S. and Diehl, S., 2019. Cluster functional renormalization group and absence of a bilinear spin liquid in the J 1– J 2 Heisenberg model. *Physical Review B*, 100(12), p.125130.
3. Mikheyenkov, A.V.E., Shvartsberg, A.V. and Barabanov, A.F., 2013. Phase transitions in the 2D J 1-J 2 Heisenberg model with arbitrary signs of exchange interactions. *JETP letters*, 98(3), pp.156-160.
4. Tilly, J., Chen, H., Cao, S., Picozzi, D., Setia, K., Li, Y., Grant, E., Wossnig, L., Rungger, I., Booth, G.H. and Tennyson, J., 2022. The variational quantum eigensolver: a review of methods and best practices. *Physics Reports*, 986, pp.1-128.
5. Boter, J.M., Dehollain, J.P., Van Dijk, J.P., Xu, Y., Hengsgens, T., Versluis, R., Naus, H.W., Clarke, J.S., Veldhorst, M., Sebastian, F. and Vandersypen, L.M., 2022. Spiderweb array: a sparse spin-qubit array. *Physical Review Applied*, 18(2), p.024053.
6. H. Patil, Y. Wang, and P. S. Krstic, "Variational quantum linear solver with a dynamic ansatz," *Physical Review A*, vol. 105, no. 1, Jan. 2022. Doi: 10.1103/physreva.105.012423.
7. 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. Doi: 10.1038/s41467-019-10988-2

Appendix A: Vanadates

Pharmacological properties [edit]

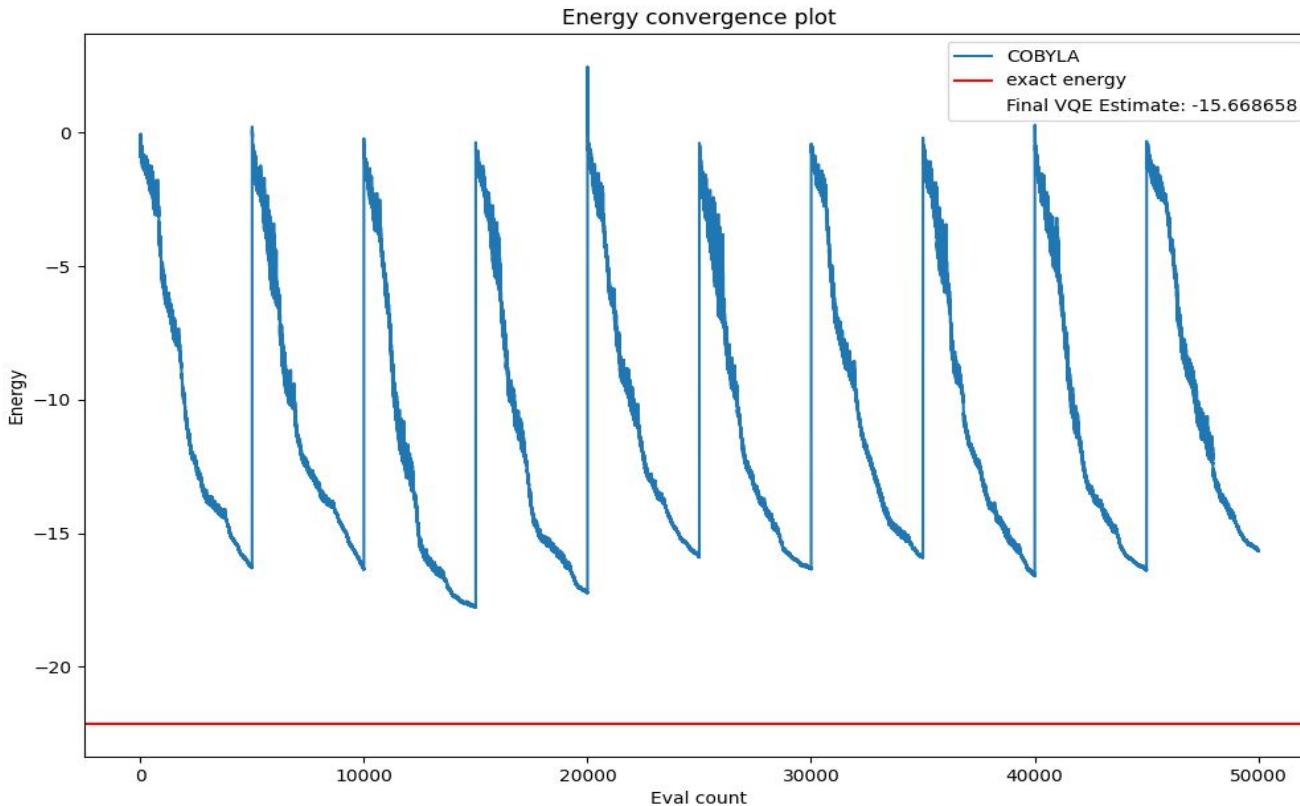
Vanadate is a potent inhibitor of certain plasma membrane [ATPases](#), such as Na^+/K^+ -ATPase and Ca^{2+} -ATPase ([PMCA](#)). Acting as a transition-state analog of phosphate, vanadate undergoes nucleophilic attack by water during phosphoryl transfer, essentially "trapping" P-type ATPases in their phosphorylated E2 state. [\[11\]](#)[\[12\]](#) However, it does not inhibit other ATPases, such as [SERCA](#) (sarco/endoplasmic reticulum Ca^{2+} -ATPase), actomyosin ATPase and mitochondrial ATPase. [\[13\]](#)[\[14\]](#)[\[15\]](#)

Appendix B: COBYLA as optimizer, static vs dynamic



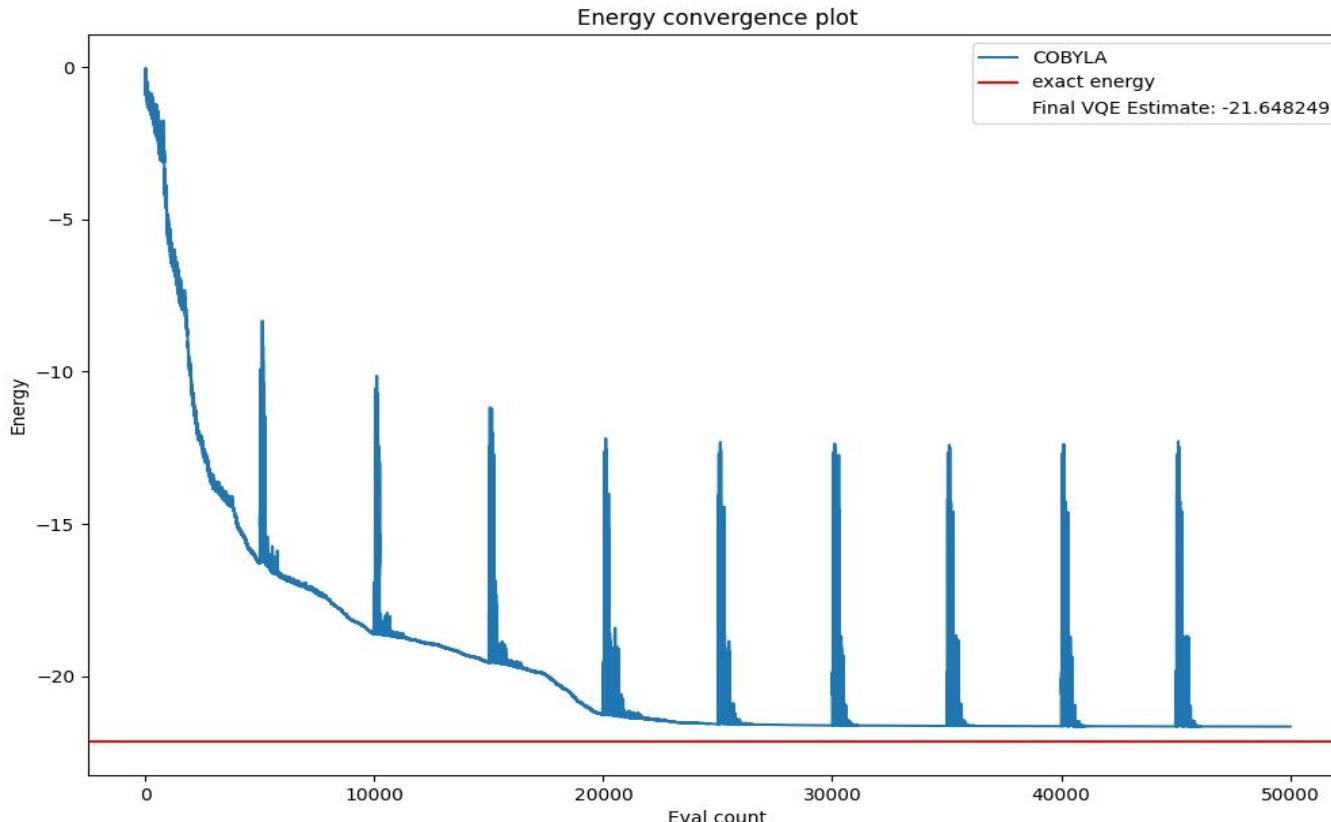
With interruptions,
NO parameter randomization,
and ansatz modifications
(sequential layer addition).

Appendix B: COBYLA as optimizer, static vs dynamic



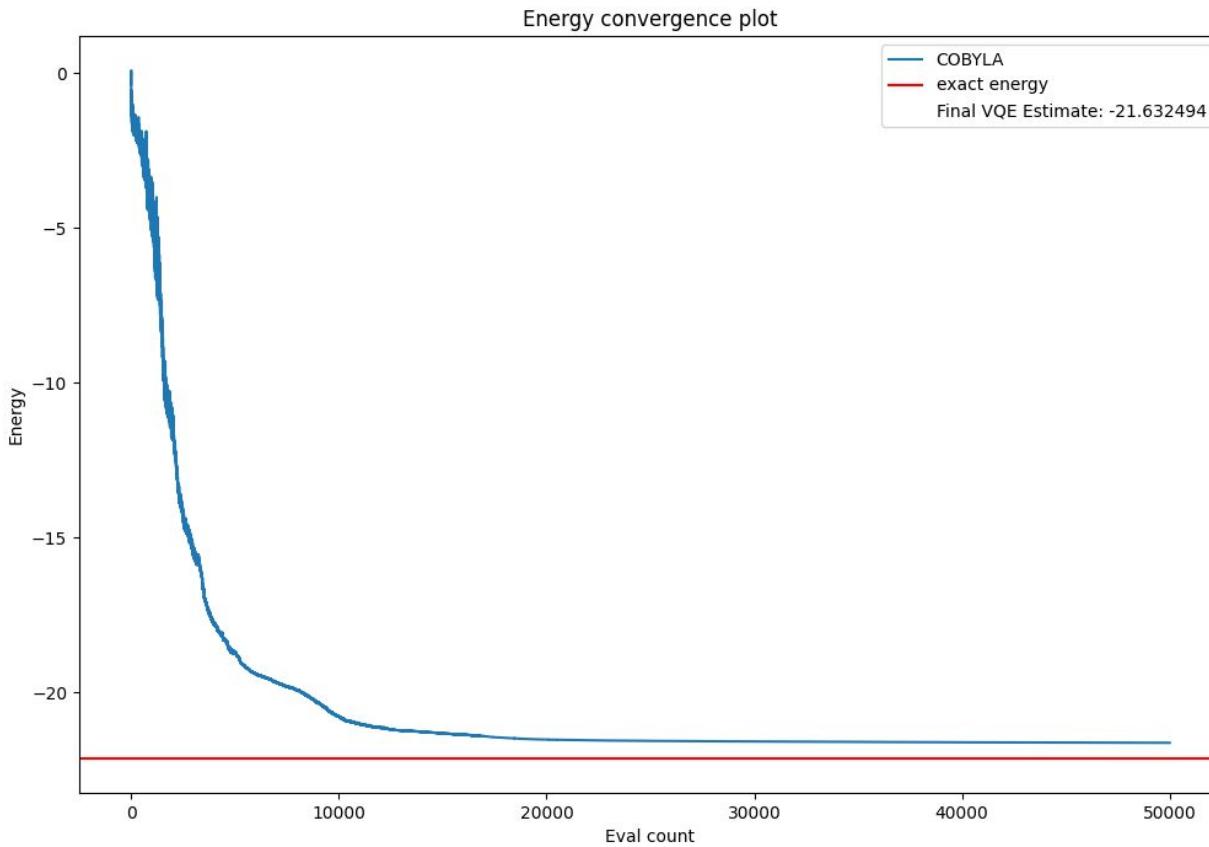
With interruptions,
and parameter randomization,
NO ansatz modifications.

Appendix B: COBYLA as optimizer, static vs dynamic



With interruptions,
NO parameter randomization,
NO ansatz modifications.

Appendix B: COBYLA as optimizer, static vs dynamic



NO interruptions.

Appendix C: Number of Parameters

- For Feulner-Hartmann ansatz:

$$2mn + \#layers * (mn + 2mn - (m + n))$$

- For Two-Local ansatz:

$$N * (\#layers + 1)$$

Appendix D: Outlier for Sequential Layer Adding

