

# The VQE Energy Landscape

Harper Grimsley

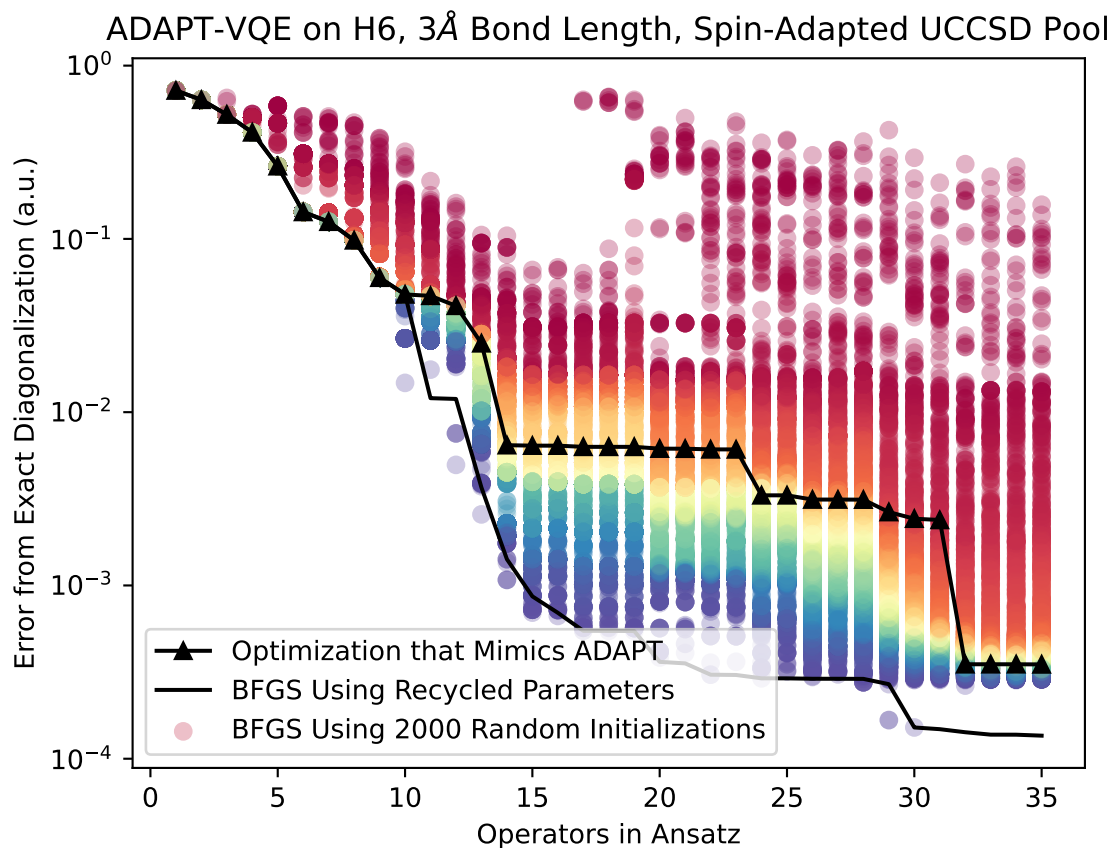
July 20, 2021

## 1 Motivation

Generalizability of the ADAPT-VQE ansatz to different Hamiltonians (e.g.,  $\hat{H}$  for the same molecule at different geometries) remains an open problem. Unfortunately, it has proven challenging to recover a continuous potential energy surface for a "fixed" ADAPT ansatz subject to parameter re-optimization. The underlying cause appears to be a failure of classical optimizers to identify the global energy minimum.

## 2 Spurious Solutions

Very large numbers of local minima apparently exist in the ADAPT-VQE ansatz, forming nearly continuous spectra in some cases:



Note in particular that random guesses (each parameter is initialized randomly between -1 and 1) are occasionally better than a recycled ADAPT guess. This occurrence appears to be unique to the Fermionic pool, or at least is so much more common that we only observe it in those pools. (It occurs regardless of the use spin-adaptation, spin-complementation, or neither.) By contrast, while many local minima are also seen in the Pauli pool, they are never better than the recycled guess.

### 3 Gimbal Locks

Francesco Evangelista, Garnet Chan, and Gus Scuseria published a paper in 2019 on the expressibility of UCC-type ansätze. Notable points of the paper include:

- The UCCSD...N ansatz  $|\psi\rangle = e^{\hat{T}-\hat{T}^\dagger} |0\rangle$  is able to exactly describe a CISD...N wavefunction  $|\Psi\rangle$ , assuming there are no pathological critical surfaces where the Jacobian of  $|\psi\rangle$ ,  $\hat{J}$  is rank-deficient.

$$\hat{J}_{ij} = \frac{\partial |\psi\rangle_i}{\partial t_j} \quad (1)$$

- Some (i.e., a non-empty set for any given  $|\Psi\rangle$ ) 1st-order, Pseudo-Trotterized (No operators are repeated, but free control of the t-amplitudes is assumed) UCCSD...N ansätze can express a CISD...N wavefunction  $|\Psi\rangle$  unconditionally.
- An infinite number of 1- and 2-particle, hole-particle generators can be used to approximate a pseudo-Trotterized UCCSD...N ansatz to arbitrary precision. That is, ignoring local minima and failures to add operators, ADAPT will eventually be exact when using the simple Fermionic pool.

Jacobian analysis of  $|\Psi\rangle$  without looking at the energy has proven interesting. For example, the first time a recycled guess is worse than a random one is the 10th operator. The solution ansatz's Jacobian obtained by initializing BFGS at the previous solution parameters, a very popular local minimum, has a distinctly small singular value, with the right-hand singular vector suggesting that the 1st and 10th parameter can be moved in opposite directions, with some minor relaxation between them, and the wavefunction will change very little. This suggests the presence of a “Gimbal lock,” where the ansatz experiences an approximate loss of linear degrees of freedom.

### 4 Periodicity Issues with Symmetry-Adapted UCCSD

Among the issues with global optimization of the Trotterized UCCSD landscape is the fact that symmetry-adapted Fermionic operators do not have the simple periodic structures possessed by, say, Pauli strings. This makes it difficult to efficiently (or even fairly) sample the various amplitudes when doing things like choosing initial parameters for BFGS. As an example, we consider the first two operators in the 3 Å H<sub>6</sub> ansatz. Some notational conventions will make this easier to follow:

$$\hat{S}_{ij}^{ab} = \frac{1}{2} \left( \hat{a}_{ij}^{a\bar{b}} + \hat{a}_{ji}^{a\bar{b}} + \hat{a}_{ij}^{b\bar{a}} + \hat{a}_{ji}^{b\bar{a}} \right) \quad (2)$$

$$|S_{ij}^{ab}\rangle = \frac{1}{2} \left( |\phi_{ij}^{a\bar{b}}\rangle + |\phi_{ji}^{a\bar{b}}\rangle + |\phi_{ij}^{b\bar{a}}\rangle + |\phi_{ji}^{b\bar{a}}\rangle \right) \quad (3)$$

$$|S_{k\bar{k}ij}^{c\bar{c}ab}\rangle = \frac{1}{2} \left( |\phi_{k\bar{k}ij}^{c\bar{c}a\bar{b}}\rangle + |\phi_{k\bar{k}ji}^{c\bar{c}a\bar{b}}\rangle + |\phi_{k\bar{k}ij}^{c\bar{c}b\bar{a}}\rangle + |\phi_{k\bar{k}ji}^{c\bar{c}b\bar{a}}\rangle \right) \quad (4)$$

The first operator acts on the reference as:

$$|\psi_1\rangle = e^{t_{02}^{35}(\hat{S}_{02}^{35} - \hat{S}_{35}^{02})} |0\rangle = \cos^2 \left( \frac{t_{02}^{35}}{\sqrt{2}} \right) |0\rangle + \frac{1}{\sqrt{2}} \sin \left( \sqrt{2} t_{02}^{35} \right) |S_{02}^{35}\rangle + \sin^2 \left( \frac{t_{02}^{35}}{\sqrt{2}} \right) |\phi_{0022}^{3355}\rangle \quad (5)$$

Every coefficient in  $|\psi_1\rangle$  is periodic in  $t_{02}^{35}$  with period  $\sqrt{2}\pi$ . The second operator is where problems are introduced, however. It can be shown that:

$$|\psi_2\rangle = e^{t_{12}^{34}(\hat{S}_{12}^{34} - \hat{S}_{34}^{12})} |\psi_1\rangle = \cos^2\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \left[ \cos^2\left(\frac{t_{12}^{34}}{\sqrt{2}}\right) |0\rangle + \frac{1}{\sqrt{2}} \sin\left(\sqrt{2}t_{12}^{34}\right) |S_{12}^{34}\rangle + \sin^2\left(\frac{t_{12}^{34}}{\sqrt{2}}\right) |\phi_{1122}^{3344}\rangle \right] \quad (6)$$

$$+ \sin^2\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) |\phi_{0022}^{3355}\rangle + \frac{1}{\sqrt{2}} \sin\left(\sqrt{2}t_{02}^{35}\right) \left[ \cos\left(\frac{t_{12}^{34}}{2}\right) |S_{02}^{35}\rangle + \sin\left(\frac{t_{12}^{34}}{2}\right) |S_{2201}^{3345}\rangle \right] \quad (7)$$

Each coefficient in the above expansion is periodic (or constant) in  $t_{12}^{34}$ . However, the coefficients in the first bracket have period  $\sqrt{2}\pi$ , while those in the second bracket have period  $4\pi$ . These periods are *incommensurable*- they have no least common multiple, so that there is no overall  $t_{12}^{34}$  period of  $|\psi_2\rangle$ . This lack of periodicity extends to the energy functional:

$$E_c[t_{12}^{34}, t_{02}^{35}] = \frac{1}{2} \cos^4\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \sin^2\left(\sqrt{2}t_{12}^{34}\right) \langle S_{12}^{34} | \hat{H}_N | S_{12}^{34} \rangle \quad (8)$$

$$+ \cos^4\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \sin^4\left(\frac{t_{12}^{34}}{\sqrt{2}}\right) \langle \phi_{1122}^{3344} | \hat{H}_N | \phi_{1122}^{3344} \rangle \quad (9)$$

$$+ \sin^4\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \langle \phi_{0022}^{3355} | \hat{H}_N | \phi_{0022}^{3355} \rangle \quad (10)$$

$$+ \frac{1}{2} \sin^2\left(\sqrt{2}t_{02}^{35}\right) \cos^2\left(\frac{t_{12}^{34}}{2}\right) \langle S_{02}^{35} | \hat{H}_N | S_{02}^{35} \rangle \quad (11)$$

$$+ \frac{1}{2} \sin^2\left(\sqrt{2}t_{02}^{35}\right) \sin^2\left(\frac{t_{12}^{34}}{2}\right) \langle S_{2201}^{3345} | \hat{H}_N | S_{2201}^{3345} \rangle \quad (12)$$

$$+ \cos^2\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \sin\left(\sqrt{2}t_{12}^{34}\right) \langle 0 | \hat{H}_N | S_{12}^{34} \rangle \quad (13)$$

$$+ \sqrt{2} \cos^2\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \sin\left(\sqrt{2}t_{02}^{35}\right) \cos\left(\sqrt{2}t_{12}^{34}\right) \langle 0 | \hat{H}_N | S_{02}^{35} \rangle \quad (14)$$

$$+ \sqrt{2} \cos^4\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \sin\left(\sqrt{2}t_{12}^{34}\right) \sin^2\left(\frac{t_{12}^{34}}{\sqrt{2}}\right) \langle S_{12}^{34} | \hat{H}_N | \phi_{1122}^{3344} \rangle \quad (15)$$

$$+ \cos^2\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \sin\left(\sqrt{2}t_{02}^{35}\right) \sin\left(\sqrt{2}t_{12}^{34}\right) \cos\left(\sqrt{2}t_{12}^{34}\right) \langle S_{12}^{34} | \hat{H}_N | S_{02}^{35} \rangle \quad (16)$$

$$+ \cos^2\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \sin\left(\sqrt{2}t_{02}^{35}\right) \sin\left(\sqrt{2}t_{12}^{34}\right) \sin\left(\frac{t_{12}^{34}}{2}\right) \langle S_{12}^{34} | \hat{H}_N | S_{2201}^{3345} \rangle \quad (17)$$

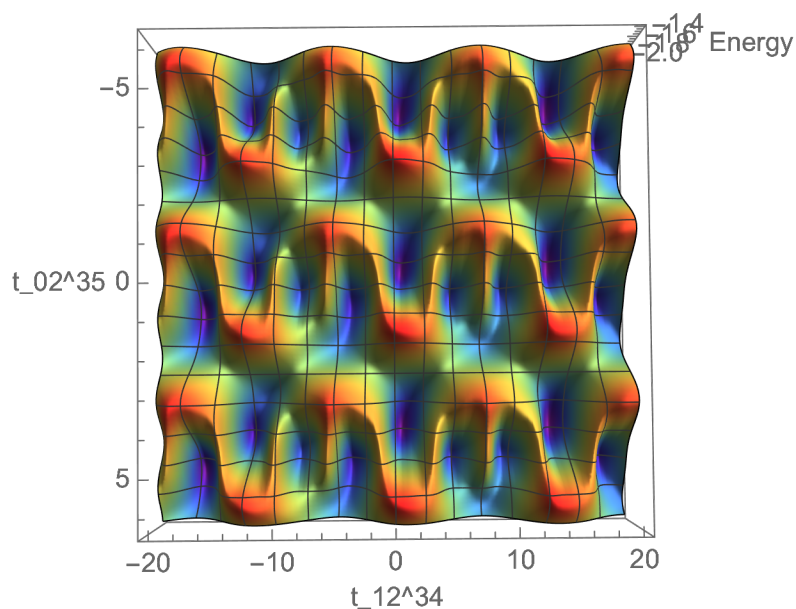
$$+ \sqrt{2} \cos^2\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \sin\left(\sqrt{2}t_{02}^{35}\right) \sin^2\left(\frac{t_{12}^{34}}{\sqrt{2}}\right) \sin\left(\frac{t_{12}^{34}}{2}\right) \langle \phi_{1122}^{3344} | \hat{H}_N | S_{2201}^{3345} \rangle \quad (18)$$

$$+ \sqrt{2} \sin^2\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \sin\left(\sqrt{2}t_{02}^{35}\right) \cos\left(\frac{t_{12}^{34}}{\sqrt{2}}\right) \langle \phi_{0022}^{3355} | \hat{H}_N | S_{02}^{35} \rangle \quad (19)$$

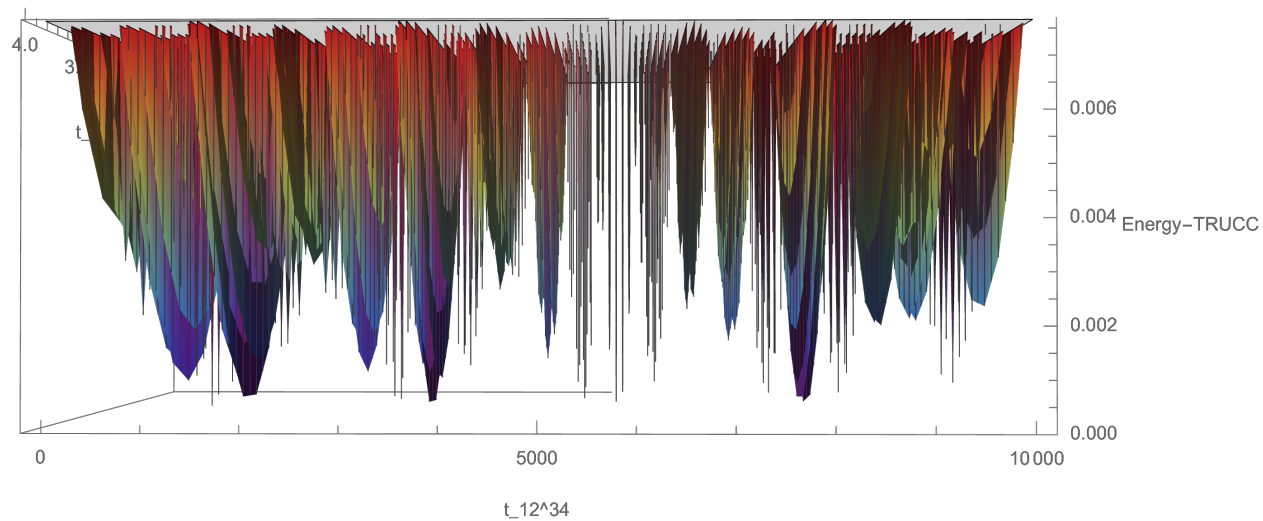
$$+ \sqrt{2} \sin^2\left(\frac{t_{02}^{35}}{\sqrt{2}}\right) \sin\left(\sqrt{2}t_{02}^{35}\right) \sin\left(\frac{t_{12}^{34}}{2}\right) \langle \phi_{0022}^{3355} | \hat{H}_N | S_{2201}^{3345} \rangle \quad (20)$$

$$+ \sin^2\left(\sqrt{2}t_{02}^{35}\right) \cos\left(\frac{t_{12}^{34}}{2}\right) \sin\left(\frac{t_{12}^{34}}{2}\right) \langle S_{02}^{35} | \hat{H}_N | S_{2201}^{3345} \rangle \quad (21)$$

For example, the first term has  $t_{12}^{34}$  periodic in  $\frac{\pi}{\sqrt{2}}$ , while the fourth term has it periodic in  $\frac{\pi}{2}$ . Thus, the energy cannot have an overall periodicity in  $t_{12}^{34}$ . The energy landscape for this problem is pictured below for a relatively small region.



Below is a plot over a much larger range of  $t_{12}^{34}$  values and a smaller range of  $t_{02}^{35}$  values. The grey plane at the top is the nearest minimum to the HF reference in the previous figure. The energies are given relative to TRUCC, described in the next section. The main takeaway is that if one assumes approximate periodicity in  $t_{12}^{34}$ , and doesn't sample it over the entire set of real numbers, they will not even get within chemical accuracy of the ansatz's full potential.



## Conjecture About Periodicity Prediction

Let  $\hat{A}$  be an anti-Hermitian operator. The action of  $e^{\hat{A}t}|\psi\rangle$  (not necessarily normalized), has a fundamental, simple period in  $t$  of  $\frac{2\pi}{\sqrt{-\lambda}}$  if and only if:

1.  $\hat{A}|\psi\rangle \neq 0$
2.  $\hat{A}|\psi\rangle$  is an eigenvector of  $\hat{A}^2$  with eigenvalue  $\lambda$ .

By a simple period, we mean that there are not multiple terms with unique commensurable periods. That is, two different periods with a least common multiple still aren't allowed. E.g.,  $\sin(2t) + \sin(4t)$  is not simple by our criterion. Note that if  $\hat{A}^2 = -\hat{\mathbf{1}}$ , as with the Pauli strings,  $e^{\hat{A}t}$  is immediately made periodic in  $2\pi$ , regardless of what it is acting on. I believe this conjecture can be used to show that the spin-complemented UCCSD pool should be periodic, at least for an  $S_z$ -pure reference.

## Conjecture Regarding A Fix to The Problem

Let  $f(x)$  and  $g(x)$  be incommensurable, periodic functions of  $x$  with periods  $T_f$  and  $T_g$ . For any  $\epsilon > 0$ ,  $y \in \mathbf{R}$ , there exists an integer  $N$  such that  $|g(x + NT_f) - g(y)| < \epsilon$ . If this is the case, it follows that in floating point arithmetic, the range of the objective functional  $\mathcal{L}[f(x), g(x)]$  is equivalent to that of  $\mathcal{L}[f(x), g(y)]$ , and the arguments of  $f$  and  $g$  can be treated as two independent parameters, each of which is periodic, giving overall bounds on the optimization and avoiding artificial dependencies between  $f$  and  $g$ . (This approximation can be made whether or not the conjecture is true, and the variational principle still holds.) I am tentatively calling it ‘‘Topologically Relaxed UCC,’’ or TRUCC. I can rigorously show that the range of the TRUCC wavefunction in a bounded region is equivalent to the global range of UCC for a specific two-operator problem, and I don't think a general proof would be all that hard.

My prescription would be to factorize each generator  $\hat{A}$  by sorting all product states into minimal groups which are closed under the action of powers of  $\hat{A}$ , ignoring non-zero scalars. For example, if  $\hat{A} = \hat{a}_i^a + \hat{a}_i^{\bar{a}} - \text{H.C.}$ , the the group containing  $|0\rangle$  would be

$$\mathcal{G}_{i,|0\rangle} = \{|0\rangle, |\phi_i^a\rangle + |\phi_i^{\bar{a}}\rangle, |\phi_i^{a\bar{a}}\rangle\} \quad (22)$$

The period of  $\hat{A}_i$  should be constant for any starting guess in  $\mathcal{G}_{i,|0\rangle}$ , so that if  $\mathcal{P}$  is the projector onto  $\mathcal{G}_{i,|0\rangle}$ , then  $e^{\mathcal{P}\hat{A}_i\mathcal{P}\theta_i}$  is periodic in  $\theta_i$  regardless of its target, with a well-defined period that can be classically determined. Many of the minimal groups for a fermionic generator will have cardinality 1, since only  $\hat{A}^0 = \mathbf{I}$  does not destroy them. (For Pauli strings, the groups would all have cardinality 2 due to the involutory nature of the operators.) If we construct a whole set of projectors  $\mathcal{P}_j$  this way:

$$e^{\theta_i \hat{A}_i} = e^{\theta_i \sum_j \mathcal{P}_j \hat{A}_i \mathcal{P}_j} \quad (23)$$

$$= e^{\theta_i \sum_j \mathcal{P}_j \hat{A}_i \mathcal{P}_j} e^{\theta_i \sum_k \mathcal{P}_k \hat{A}_i \mathcal{P}_k} \dots \quad (24)$$

If we do this type of partitioning so that each of the new  $e^{\theta_i \sum_j \mathcal{P}_j \hat{A}_i \mathcal{P}_j}$  rotations are mutually incommensurable in  $\theta_i$ , then, if I'm right, each rotation can have an independent parameter. By adding the auxiliary parameters, we now have a product of periodic rotations, rather than one non-periodic rotation. I believe these are the same. If not, this is still a unitary rotation, so it won't break the variational principle. We could simply replace the operator  $\hat{A}_i$  with the generators it would get broken into when we construct the pool. There are a limited number of group structures to compute for a given operator pool. Alternatively, we only actually need the groups generated by non-zero powers of  $\hat{A}$  on target states at any given time. I am personally partial to the former choice, since I can theoretically just work out the topologically relaxed operators for six types of excitations, and then ADAPT's operator addition criterion won't unphysically think of separately parameterized rotations as being dependent when it's deciding whether to add them.