Variational quantum algorithms: adaptive and pulse-based approaches

Sophia Economou

QTML, Melbourne, Nov 2024

Outline

- Adaptive quantum algorithms (ADAPT-VQA)
  - Background & ADAPT-VQE
  - Adaptive Gibbs state preparation

- Control-VQE: optimizing at the pulse level

Variational Quantum Eigensolver (VQE)

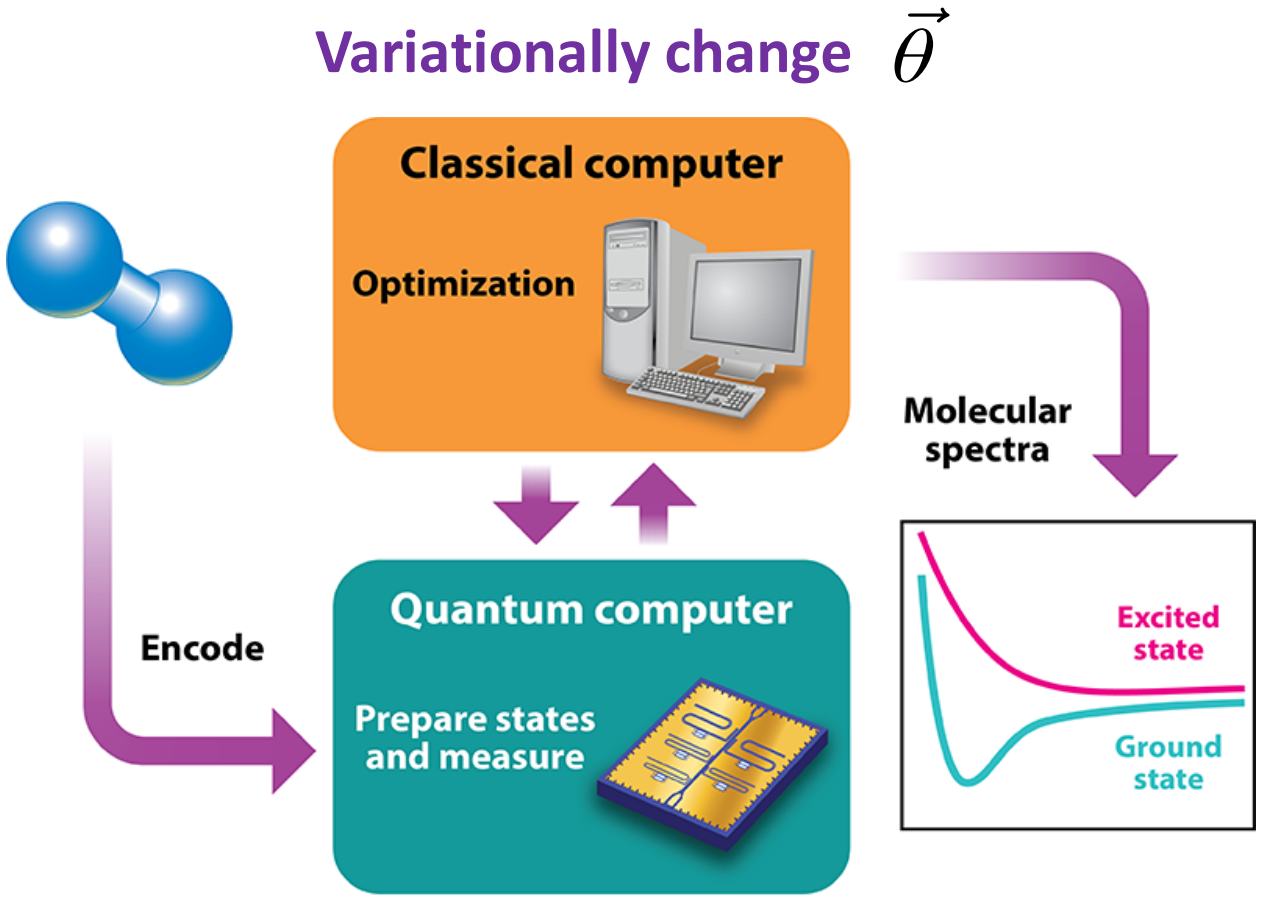
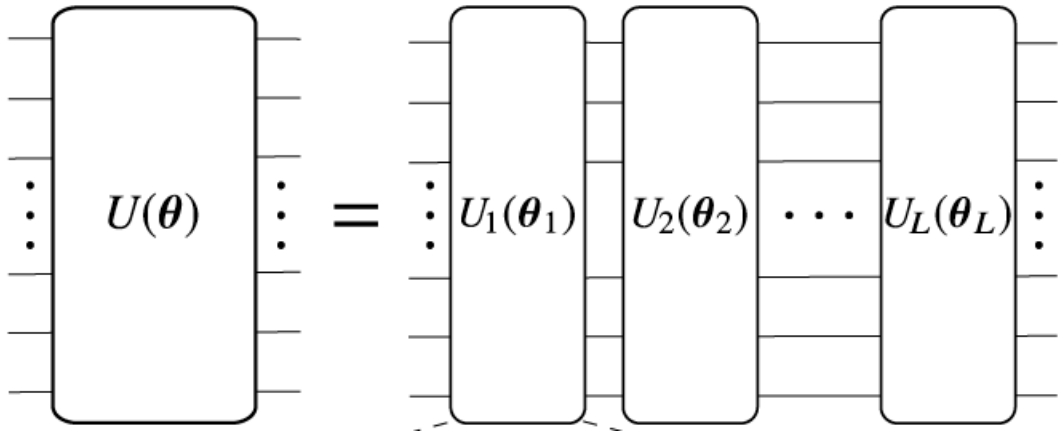


Image from Physics **11**, 14 (2018)

Ansatz $|\Psi(\vec{\theta})\rangle = U(\vec{\theta})|\Psi_{ref}\rangle$

Measured energy $E(\vec{\theta}) = \langle\Psi(\vec{\theta})|H|\Psi(\vec{\theta})\rangle$

Parametrized circuit



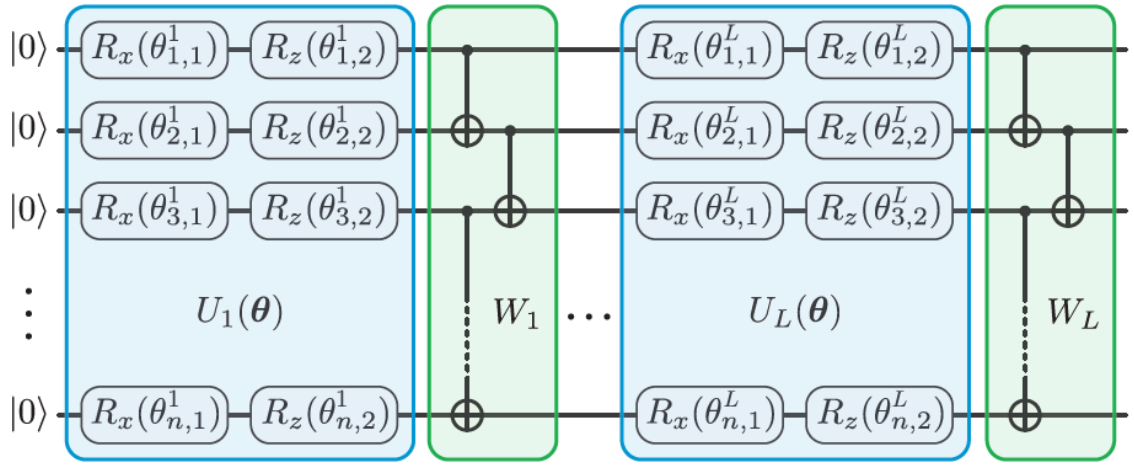
Reviews:
Cerezo et al, Nat. Rev. Phys. 3, 625 (2021)
Bharti et al, RMP 94, 015004 (2022)
Tilly et al, arXiv:2111.05176

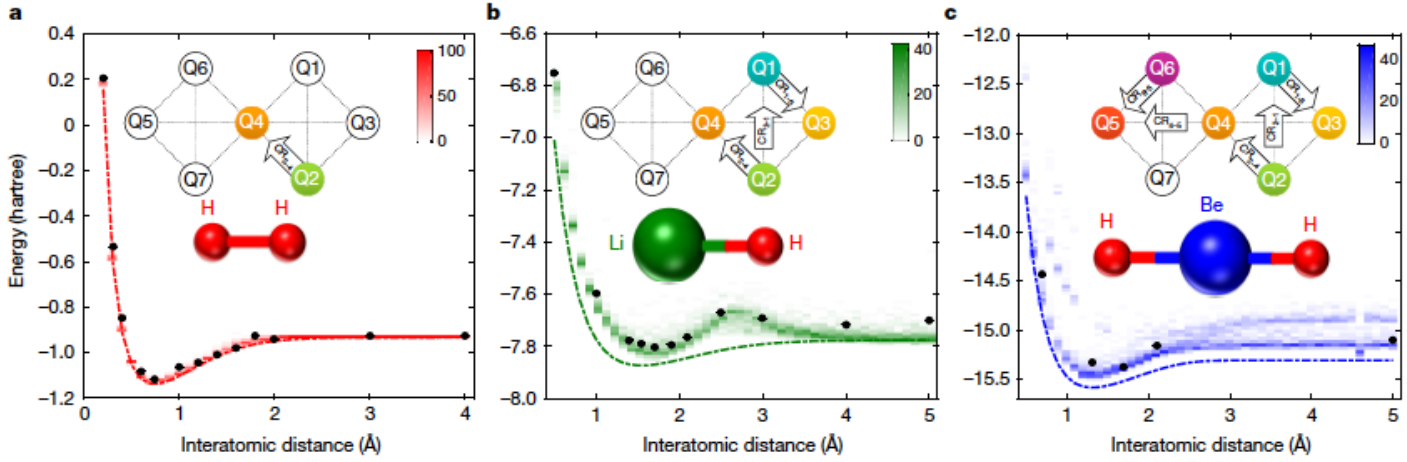
Hardware-efficient ansatz

- Tailored to chosen platform
- Inefficient—too much of the Hilbert space sampled
- *Not related to the problem* →
- Difficult to optimize (barren plateaus)

  McClean et al., Nat. Commun. 9, 4812 (2018)

  +lots of work from LANL





Kandala et al, Nature **549**, 242 (2017)

Problem-aware ansätze

- Instead of random circuits, use circuits that have knowledge of the problem
- Incorporate: symmetries (particle number, spin, spatial symmetry, etc), locality
- This can avoid issues with optimization, can lead to shorter circuits

Fermionic Hamiltonian and mapping

$$H = \sum_{p,q=1}^{N} h_{pq} a_p^\dagger a_q + \sum_{p,q,r,s=1}^{N} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s$$

Map (spin) orbitals to qubits: Jordan-Wigner mapping

$$a_i^\dagger \mapsto \frac{1}{2}(X_i - \mathrm{i}Y_i) \prod_{r=0}^{i-1} Z_r \qquad a_i \mapsto \frac{1}{2}(X_i + \mathrm{i}Y_i) \prod_{r=0}^{i-1} Z_r$$

Z strings O(*N*)

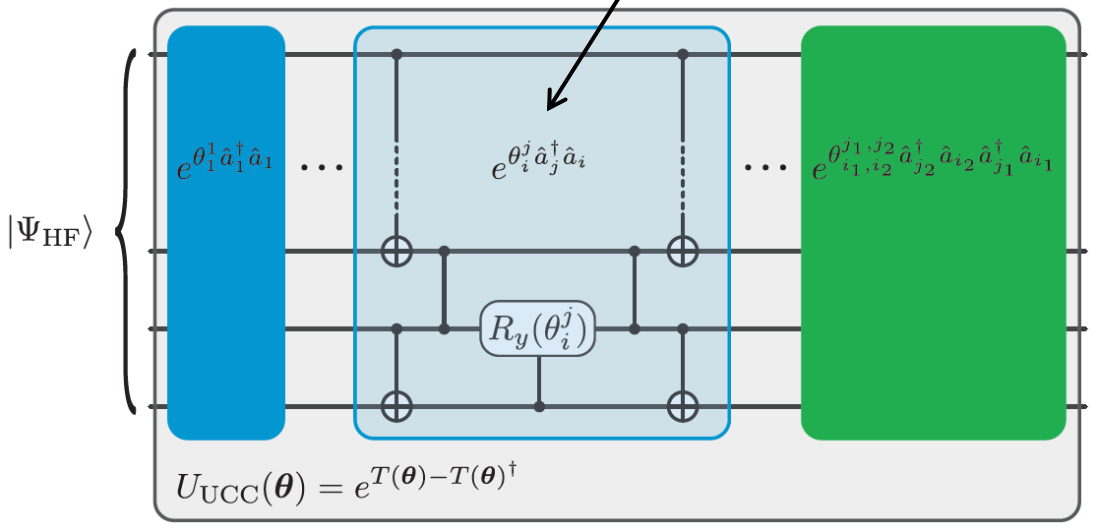
One of the first problem-aware ansätze: UCCSD

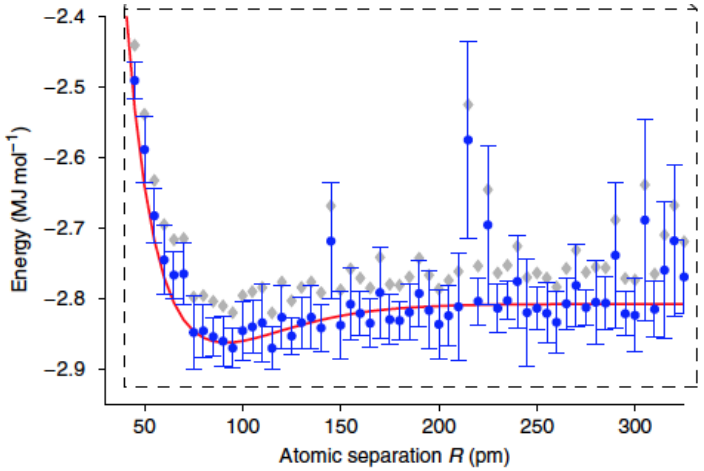
gates built from symmetry-adapted fermionic operators

Chemistry-inspired (UCCSD) ansatz

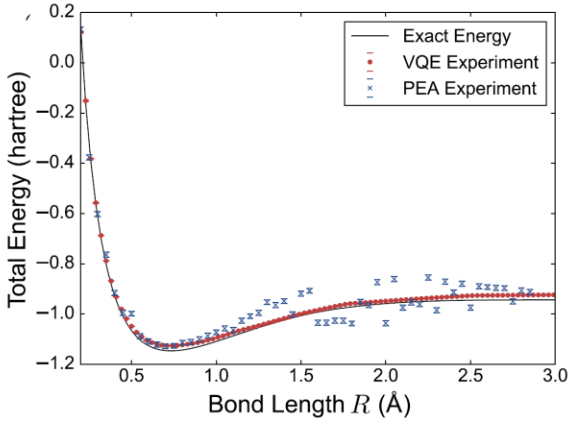
- Incorporates symmetries
- Generalizes classical simulations
- Impractically long circuits, many CNOTs
- Inconsistent under low-order Trotterization

Grimsley et al., JCTC 2020, 16, 1, 1-6

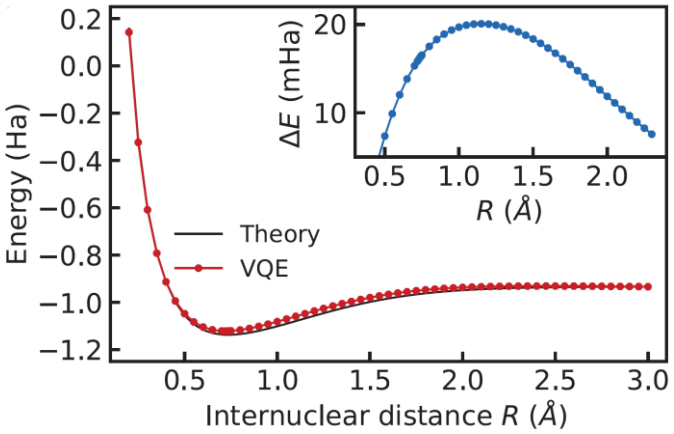




*Peruzzo et al, Nature Comm. **5**, 1 (2014)*



*O'Malley et al, PRX **6**, 031007 (2016)*



Xue et al, Nature 601, 343 (2022)

Constructing ansatz: desiderata

Quantum circuit that is:

- Short
- Expressive
- Trainable

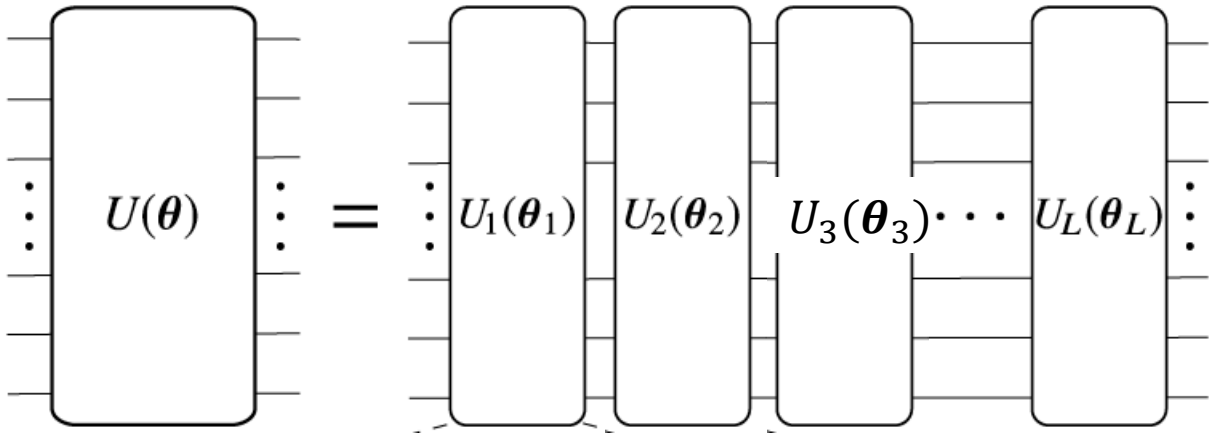
Outline

- Adaptive quantum simulation (ADAPT-VQE)
  - Molecules
  - Periodic systems

- Control-VQE: optimizing at the pulse level

Adaptive, problem-tailored VQE (ADAPT-VQE)

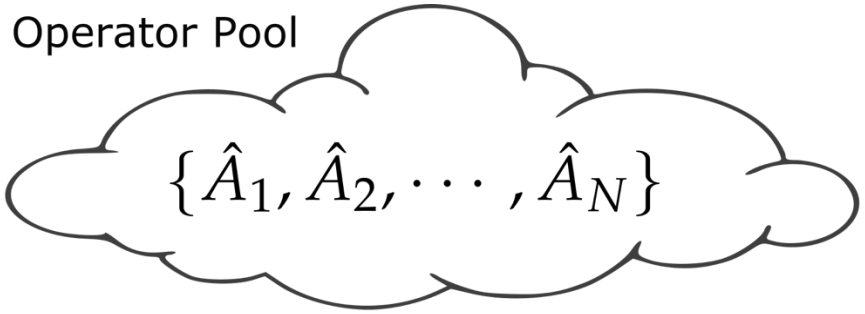
- Start from a simple reference state
- Quantum resources are precious: Only add as many operators as needed
- Problem-tailor the ansatz: Use the QC to determine how to grow the ansatz further



Grimsley, Economou, Barnes, Mayhall, Nature Communications **10**, 3007 (2019)

ADAPT-VQE ingredients: (i) operator pool

- ADAPT-VQE uses a pool of operators, A_m
- Applies unitaries one by one : $U_m = \exp(\theta_m A_m)$ to a reference state



Grimsley, Economou, Barnes, Mayhall, Nature Communications **10**, 3007 (2019)

ADAPT-VQE ingredients: (ii) update criterion

- Identify which $e^{\theta_j A_j}$ to apply to reference state $|\Psi_0\rangle$
- Take gradient of mean energy wrt θ_j

$$\frac{\partial}{\partial \theta_j} \langle \Psi_k | e^{-\theta_j A_j} H e^{\theta_j A_j} | \Psi_k \rangle |_{\theta_j = 0} = \langle \Psi_k | \underbrace{[H, A_j]} | \Psi_k \rangle$$

New operator → measure on hardware

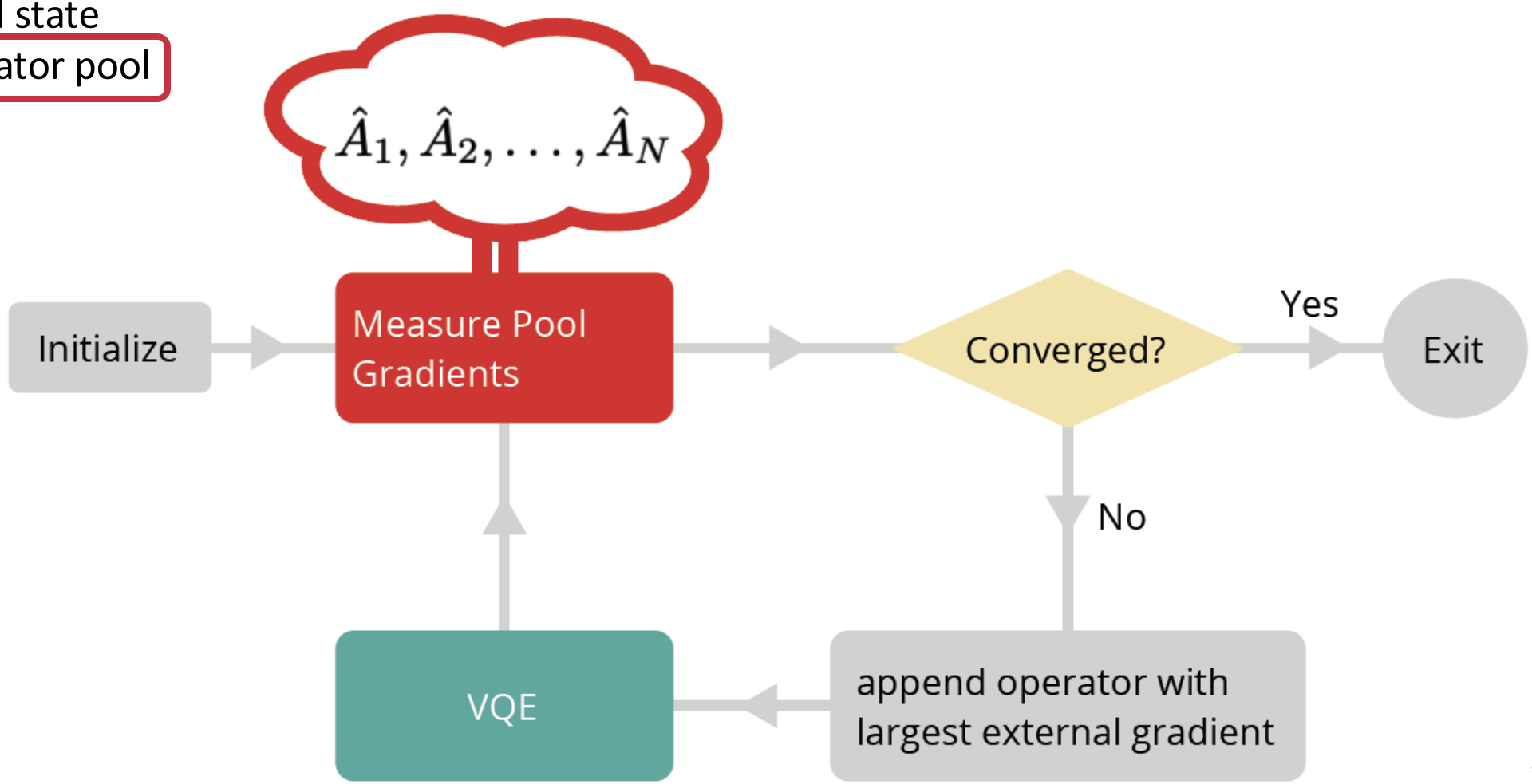
k: ADAPT iteration step

Grimsley, Economou, Barnes, Mayhall, Nature Communications **10**, 3007 (2019)

ADAPT-VQE overview

Inputs:

- Hamiltonian
- Initial state
- Operator pool



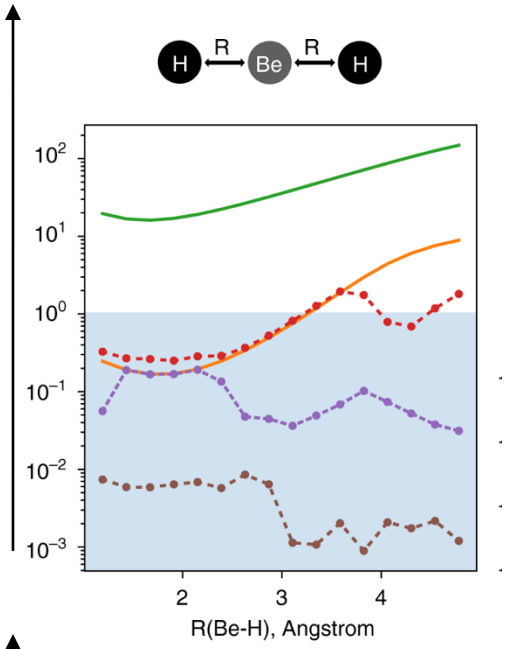
Grimsley, Economou, Barnes, Mayhall, Nature Communications **10**, 3007 (2019)

ADAPT-VQE with fermionic pool

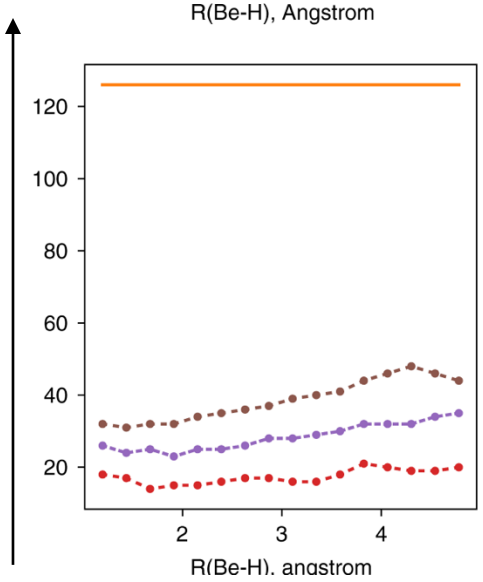
UCCSD
HF
ADAPT(ϵ_1)
ADAPT(ϵ_2)
ADAPT(ϵ_3)
Chemical Accuracy

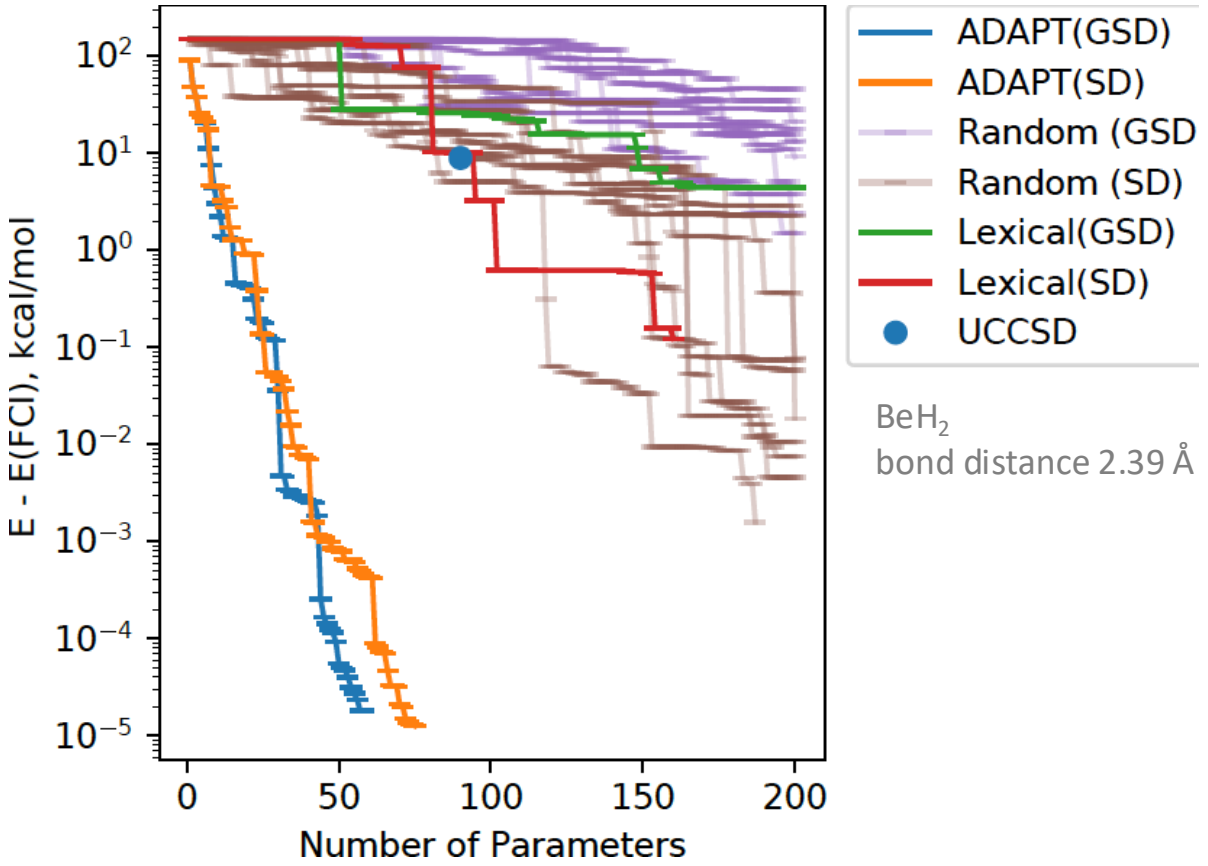
$\epsilon_1 = 0.1$
$\epsilon_2 = 0.01$
$\epsilon_3 = 0.001$

energy error



nr of variational parameters





BeH_2
bond distance 2.39 Å

Grimsley, Economou, Barnes, Mayhall,
Nature Communications **10**, 3007 (2019)

ADAPT-VQE vs other ansätze

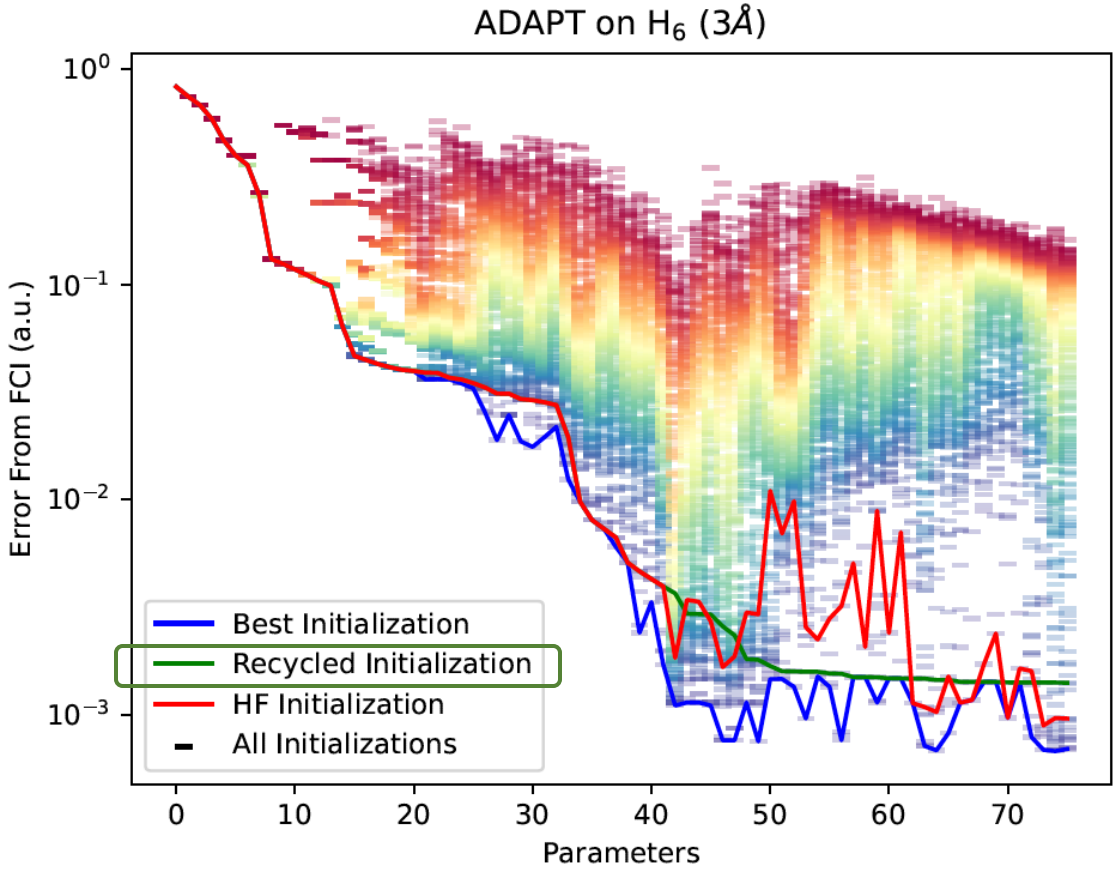
***Problem-tailored* vs problem-aware**

Example: ansatz at different points on dissociation curve of a given molecule

- UCCSD: exact same across curve
- ADAPT-VQE: changes across the curve, picks up operators as needed for strongly correlated regions

Trainability of ADAPT-VQE

- ADAPT produces compact tailored ansätze
- Shallow circuit → the landscape is generally too rugged
- ADAPT avoids the issues associated with trainability
- By construction resistant to barren plateaus



Grimsley et al, npj Quantum Information **9**, 19 (2023)

ADAPT-VQE vs other ansätze

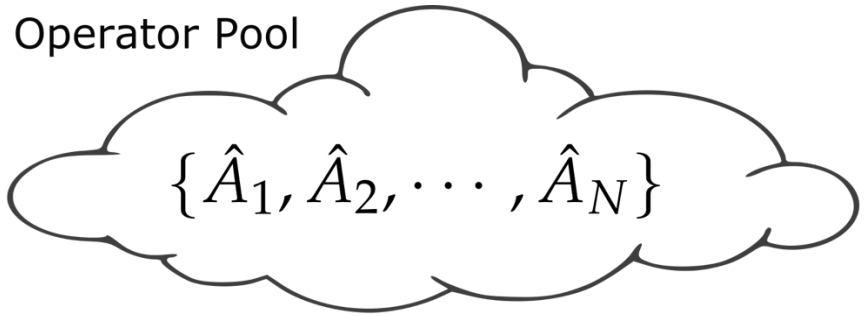
- ***Problem-tailored* vs problem-aware**

  Example: ansatz at different points on dissociation curve of a given molecule

  - UCCSD: exact same across curve
  - ADAPT-VQE: changes across the curve, picks up operators as needed for strongly correlated regions

- ***Expressivity potential* vs expressive circuits**
  - Random circuits: very expressive, hard to optimize
  - ADAPT-generated ansatz:
    - bespoke to problem, each instance not random
    - parameters warm-started

How should the operator pool be chosen?



Criteria & desiderata

- Problem constraints: **symmetries**, etc
- Hardware constraints: **connectivity, locality**, etc
- Small pool size
- …

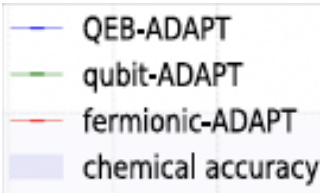
An obvious choice is the ***fermionic pool***

- Preserves particle number
- Can be spin-adapted (preserves Sz)

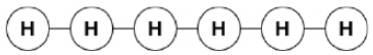
But:

- Hardware inefficient
- Grows with system size

ADAPT-VQE: Pool Choices

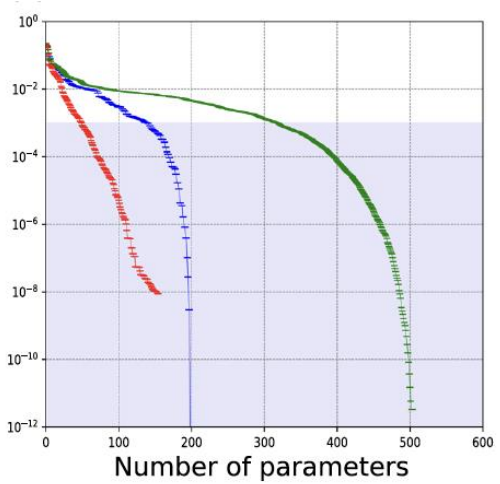


The choice of operator pool impacts the circuit- and parameter-efficiency of the final ansatz. Popular pool types include:



- Fermionic Pools: Operators directly correspond to fermionic excitations

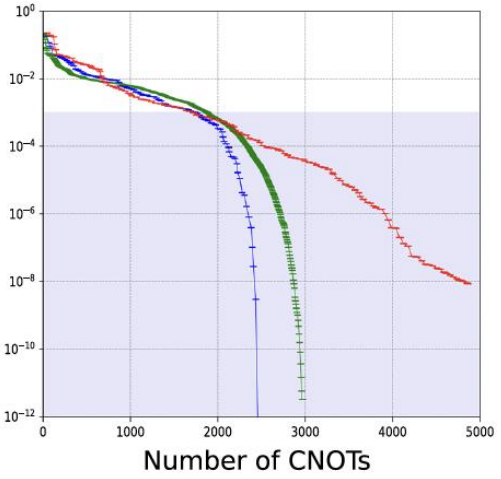
$$T_{ijkl} \equiv a_i^\dagger a_j^\dagger a_k a_l - a_k^\dagger a_l^\dagger a_i a_j$$



- Qubit Excitation (QE) Pools: Operators correspond to fermionic excitations up to anticommutation effects

$$\tilde{T}_{ijkl} \equiv Q_i^\dagger Q_j^\dagger Q_k Q_l - Q_k^\dagger Q_l^\dagger Q_i Q_j$$

(h)



- Qubit Pools: Operators are individual Pauli strings

$$i \cdot Y_q X_p X_s X_r, \quad i \cdot X_q Y_p X_s X_r, \quad i \cdot X_q X_p Y_s X_r, \quad i \cdot X_q X_p X_s Y_r,$$
$$i \cdot X_q Y_p Y_s Y_r, \quad i \cdot Y_q X_p Y_s Y_r, \quad i \cdot Y_q Y_p X_s Y_r, \quad i \cdot Y_q Y_p Y_s X_r.$$

Yordanov et al, *Commun Phys* **4**, 228 (2021)
Tang et al, *PRX Quantum* **2**, 020310 (2021)

We can do even better by further incorporating symmetries: Coupled-excitation operator pool

Couple certain excitations based on symmetry considerations

E.g., QE pool does not take $\alpha_1, \beta_1 \rightarrow \alpha_2, \beta_2$ on the same footing as $\alpha_1, \beta_2 \rightarrow \alpha_2, \beta_1$ (α: spin up; β: spin down)

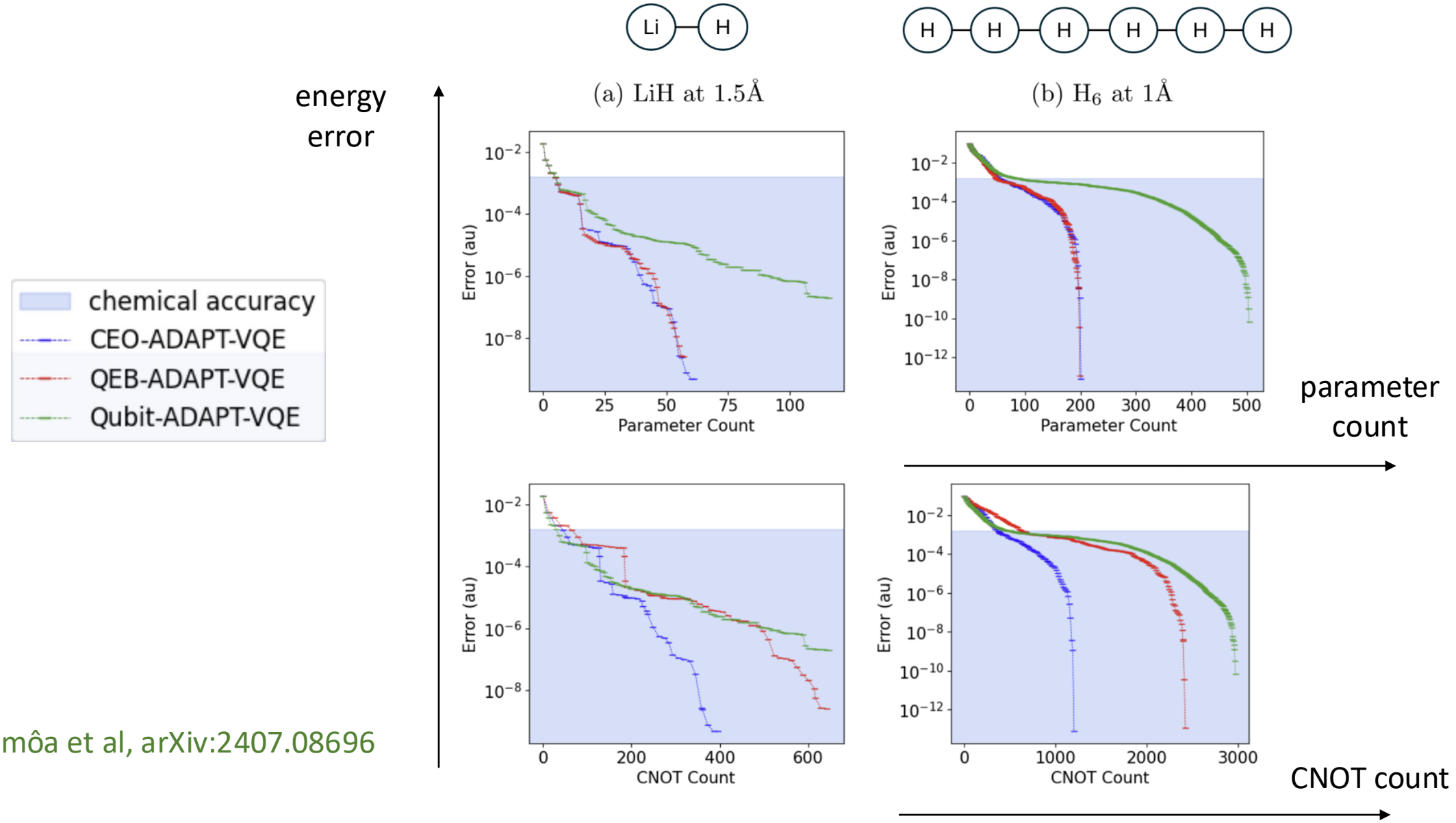
If we couple such operators, we obtain a pool with operators of the form:

$$T^{(MVP-CEO)}_{\alpha_1\beta_1\alpha_2\beta_2}(\theta_1,\theta_2) =$$
$$\frac{i}{8}[+(\theta_1+\theta_2)XXXY-(\theta_1+\theta_2)XXYX$$
$$+(\theta_1-\theta_2)XYXX+(\theta_1-\theta_2)XYYY$$
$$-(\theta_1-\theta_2)YXXX-(\theta_1-\theta_2)YXYY$$
$$+(\theta_1+\theta_2)YYXY-(\theta_1+\theta_2)YYYX].$$

Ramôa et al, arXiv:2407.08696

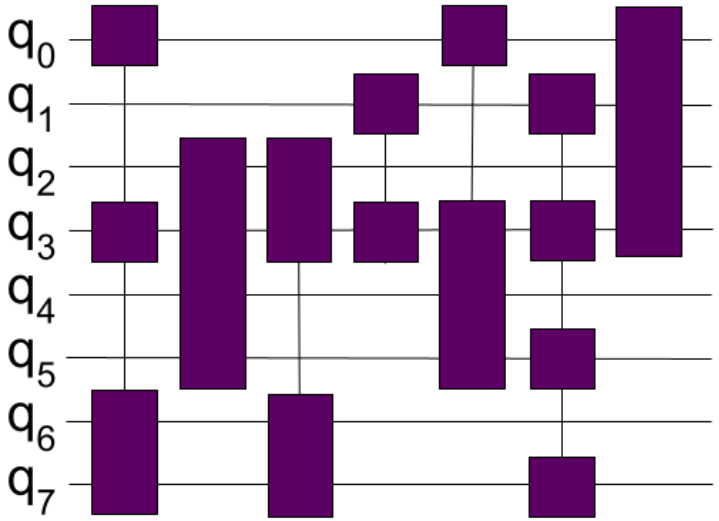
See Mafalda's talk at 4.15 pm

CEO pool, results & comparison to qubit and QE pools



Ramôa et al, arXiv:2407.08696

- Is ADAPT-VQE the most efficient way to construct circuits?
- Can we optimize further for circuit depth?

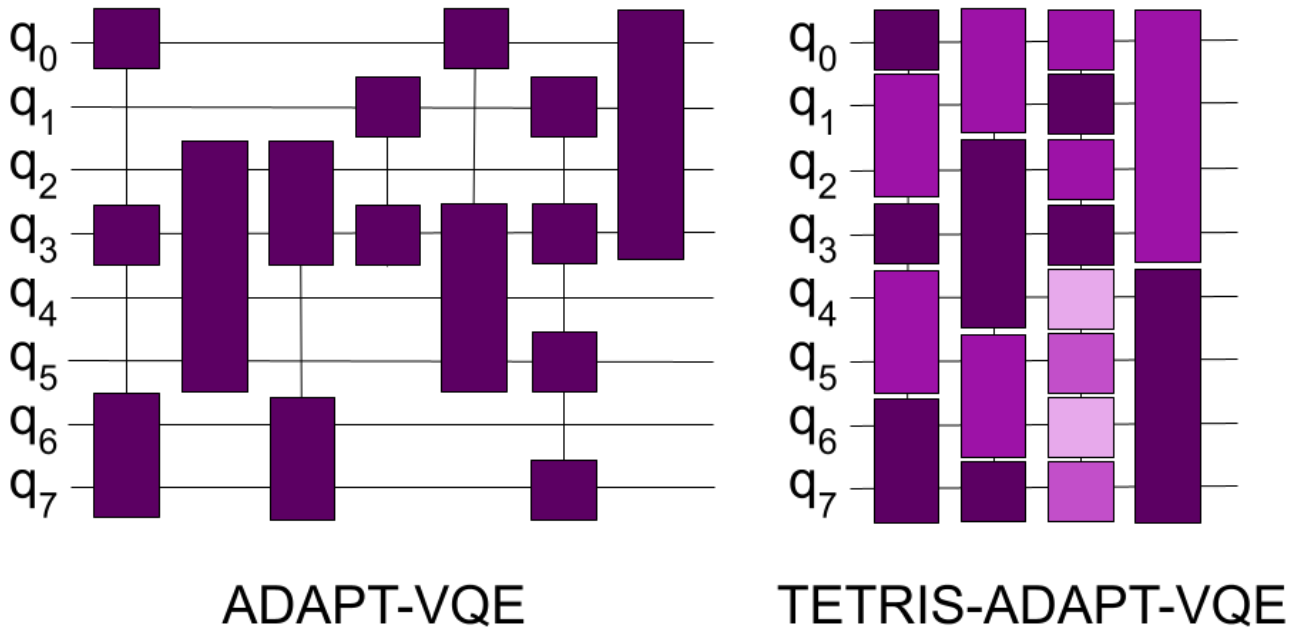


ADAPT-VQE

TETRIS-ADAPT-VQE: concept

Tiling Efficient Trial circuits with Rotations Implemented Simultaneously

Instead of one-at-a-time, add multiple operators at each step



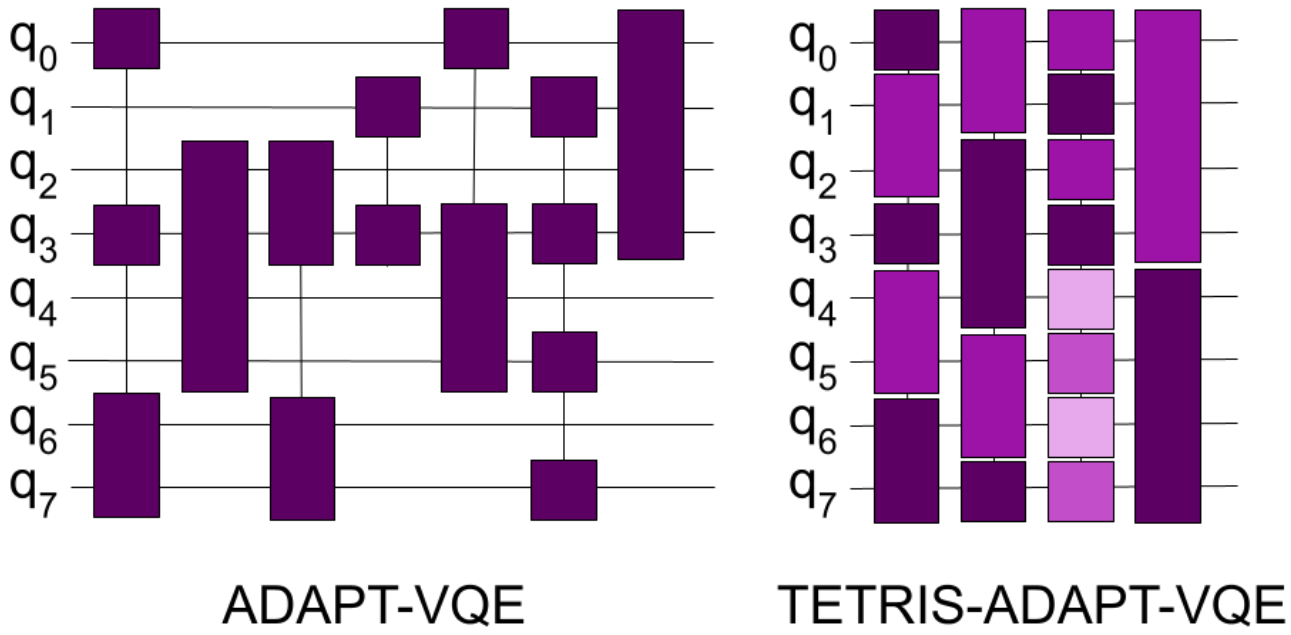
Anastasiou, et al Phys. Rev. Research 6, 013254 (2024)

TETRIS-ADAPT-VQE: concept

Tiling Efficient Trial circuits with Rotations Implemented Simultaneously

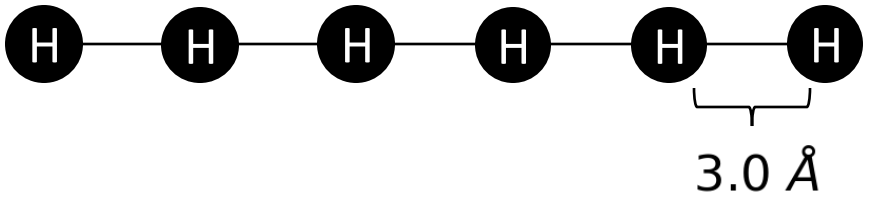
Instead of one-at-a-time, add multiple operators at each step according to:

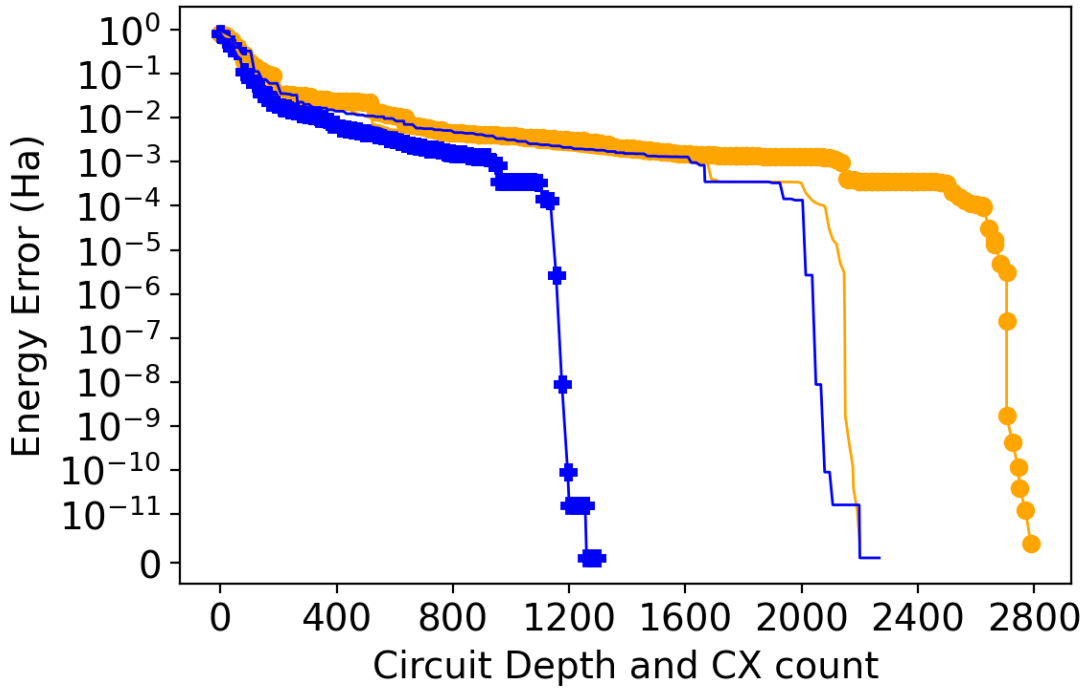
- Gradient magnitude
- $\mathcal{N}$th operator acting on different set of qubits from $(\mathcal{N}-j)$th



Anastasiou, et al Phys. Rev. Research 6, 013254 (2024)

TETRIS-ADAPT-VQE: results



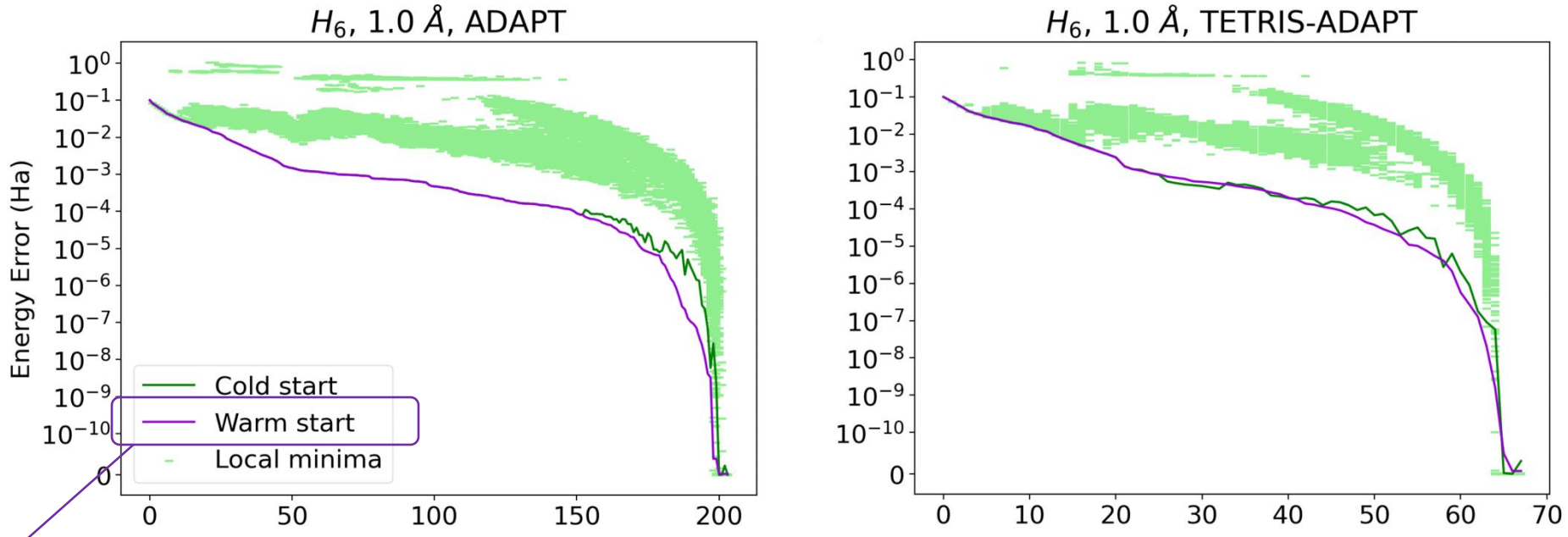


Blue: TETRIS-ADAPT

Orange: standard ADAPT

Anastasiou, et al Phys. Rev. Research 6, 013254 (2024)

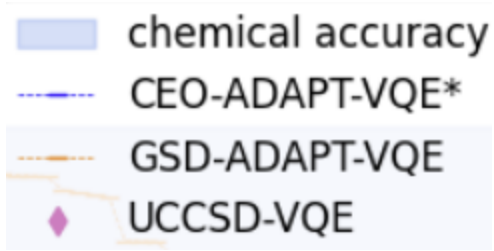
Trainability of TETRIS-ADAPT-VQE



ADAPT strategy

Anastasiou, et al Phys. Rev. Research 6, 013254 (2024)

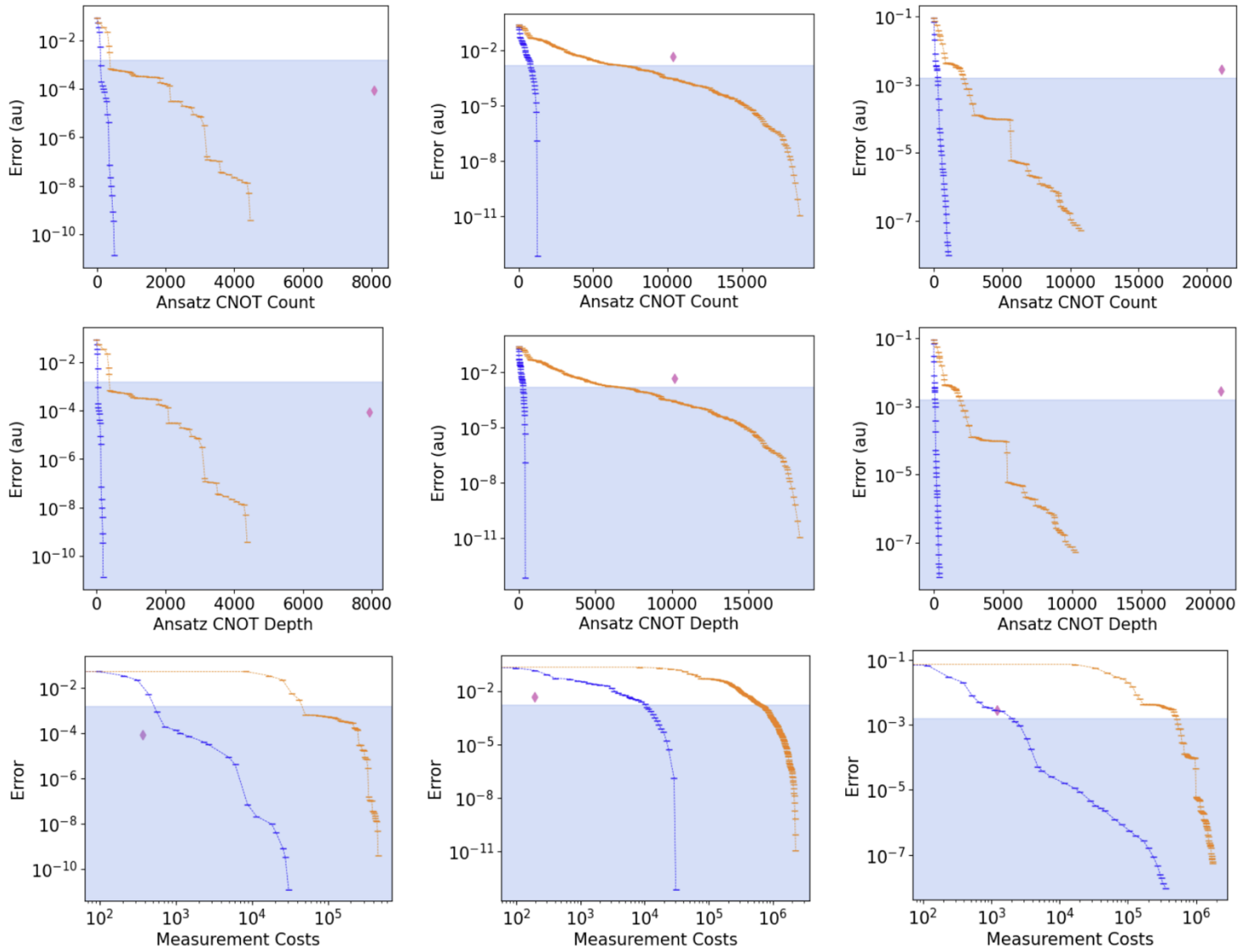
Putting it all together: CEO-ADAPT-VQE*
2019 vs 2024 ADAPT-VQE



- New (CEO) pool
- TETRIS strategy
- Recycling Hessian
- Grouping commuting operators

Compared to UCCSD:

- Order(s) of magnitude improvement in CNOT count/depth
- Comparable nr of measurements



(a) LiH at 3Å (b) H_6 at 1.5Å (c) BeH_2 at 2Å

Outline

- Adaptive quantum algorithms (ADAPT-VQA)
  - Background & ADAPT-VQE
  - Adaptive Gibbs state preparation
- Control-VQE: optimizing at the pulse level

Gibbs state preparation

- Given Hamiltonian H acting on data system D with N_D qubits, we wish to prepare mixed thermal states at arbitrary temperature T

$$\rho_G(T) = \frac{e^{-H/T}}{\mathrm{Tr}(e^{-H/T})}$$

- This has useful applications in quantum simulation, quantum machine learning, quantum optimization, etc., but is a hard problem in general

- Standard approach: prepare state that minimizes $F\left(\rho(\vec{\theta})\right) = \mathrm{Tr}\left(\rho(\vec{\theta})H + T\rho(\vec{\theta})\ln\rho(\vec{\theta})\right)$

- Challenges:
  - It's hard to know *a priori* what will make an efficient, effective ansatz
  - Measuring the entropy (and gradients) is difficult on hardware

Wu and Hsieh. PRL **123**, 220502 (2019)
Chowdhury, Low, and Wiebe. arXiv:2002:00055 (2020)
Wang, Li, and Wang. PRA **16**, 054035 (2021)

Gibbs-ADAPT-VQE

- Our approach:
  - ✓New objective function that is easier to measure:

$$C\left(\rho\left(\vec{\theta}\right)\right) = -\operatorname{Tr}\left(\rho_G(T)\rho\left(\vec{\theta}\right)\right) + \frac{1}{2}\operatorname{Tr}\left(\rho\left(\vec{\theta}\right)^2\right)$$

  - ✓ADAPT-VQE approach to grow the ansatz

Warren et al, arXiv: 2203.12757

Gibbs-ADAPT-VQE

- Our approach:
  - ✓New objective function that is easier to measure:

$$C\left(\rho\left(\vec{\theta}\right)\right) = -\mathrm{Tr}\left(\rho_G(T)\rho\left(\vec{\theta}\right)\right) + \frac{1}{2}\mathrm{Tr}\left(\rho\left(\vec{\theta}\right)^2\right)$$

  - ✓ADAPT-VQE approach to grow the ansatz

- Ancilla system can be any size, allowing resource savings (though $N_A = N_D$ needed at larger T)

- Operator pool: all 1- and 2-qubit Pauli's on combined data/ancilla system
- Pool requires initial state be only **partially entangled** (use non-optimized random y-rotations and CNOTs)

Warren et al, arXiv: 2203.12757

Gibbs-ADAPT-VQE

- Our approach:
  - ✓New objective function that is easier to measure:

$$C\left(\rho(\vec{\theta})\right) = -\mathrm{Tr}\left(\rho_G(T)\rho(\vec{\theta})\right) + \frac{1}{2}\mathrm{Tr}\left(\rho(\vec{\theta})^2\right)$$

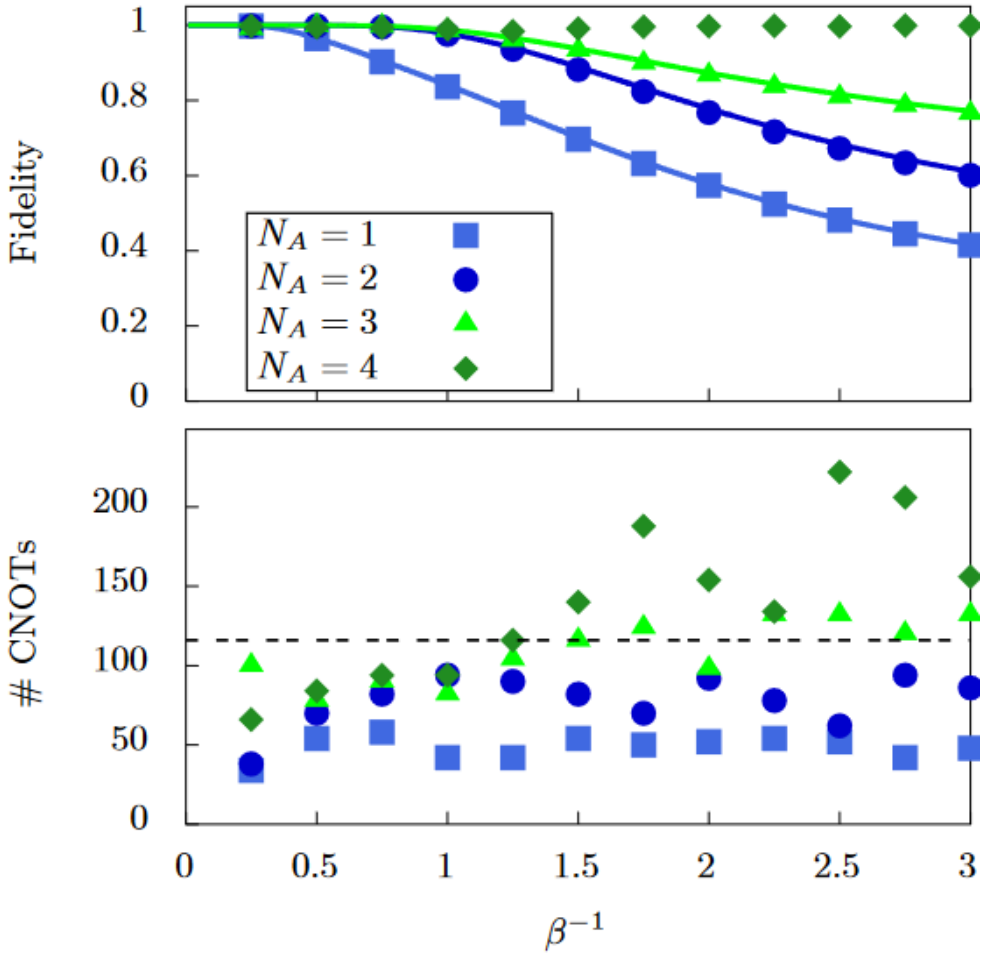
  - ✓ADAPT-VQE approach to grow the ansatz

- Ancilla system can be any size, allowing resource savings (though $N_A = N_D$ needed at larger T)

- Operator pool: all 1- and 2-qubit Pauli's on combined data/ancilla system
- Pool requires initial state be only **partially entangled** (use non-optimized random y-rotations and CNOTs)

Periodic XY Hamiltonian

$$H = -\sum_{i=1}^{N_D} X_i X_{i+1} + Y_i Y_{i+1} \qquad N_D = 4$$

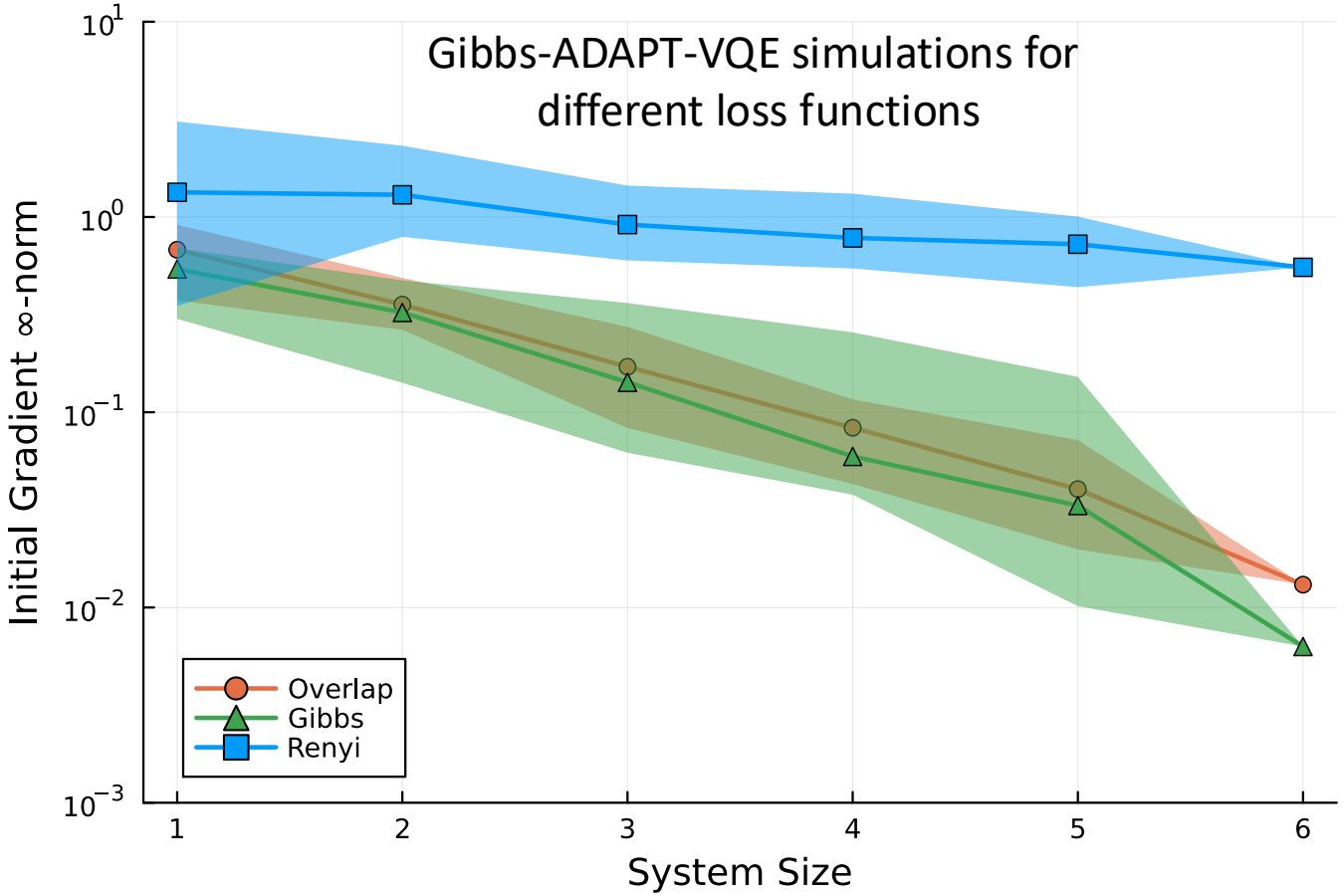


Warren et al, arXiv: 2203.12757

Avoiding barren plateaus in Gibbs-ADAPT-VQE

- Use Renyi divergence-based loss function: $$L(\sigma, \rho) = \log\left(\mathrm{Tr}\left(\sigma^2 \rho^{-1}\right)\right)$$
  - Unbounded → gradients remain large far from solution

Kieferová et al, arXiv:2106.09567



Sherbert et al, arXiv:2408.00218

Still no quantum advantage

What could work out in the near term?

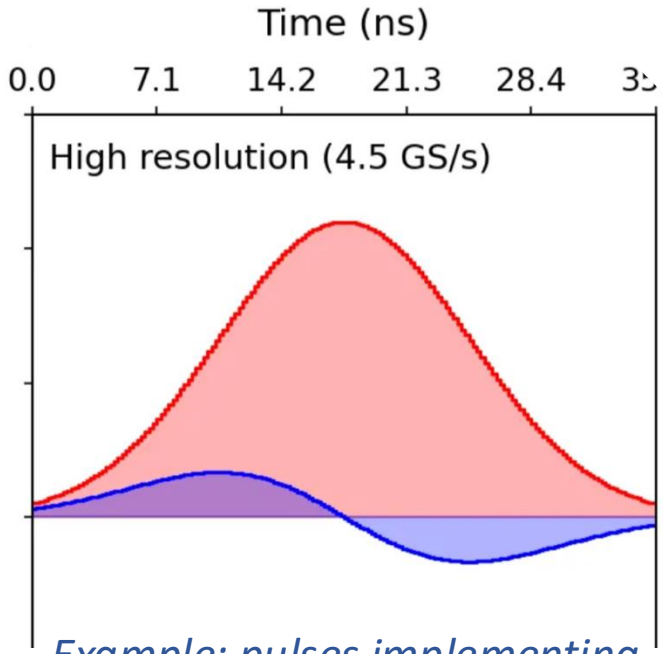
Outline

- Adaptive quantum algorithms (ADAPT-VQA)
  - Background & ADAPT-VQE
  - Adaptive Gibbs state preparation

- Control-VQE: optimizing at the pulse level

Optimizing at the pulse level

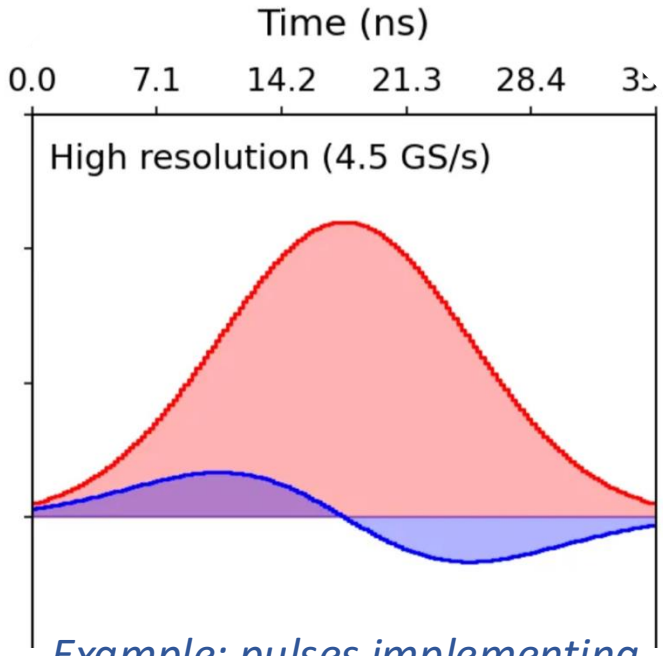
- All gates are made of electromagnetic pulses
- Using gates is a *digitized approach*



Example: pulses implementing IBM single-qubit gates

Optimizing at the pulse level

- All gates are made of electromagnetic pulses
- Using gates is a *digitized approach*



Example: pulses implementing IBM single-qubit gates

Pulse-level optimization:

- Throw out the gates, parameterize pulse directly
- Measure <H>
- Classical optimization → update pulse parameters
- Repeat until convergence

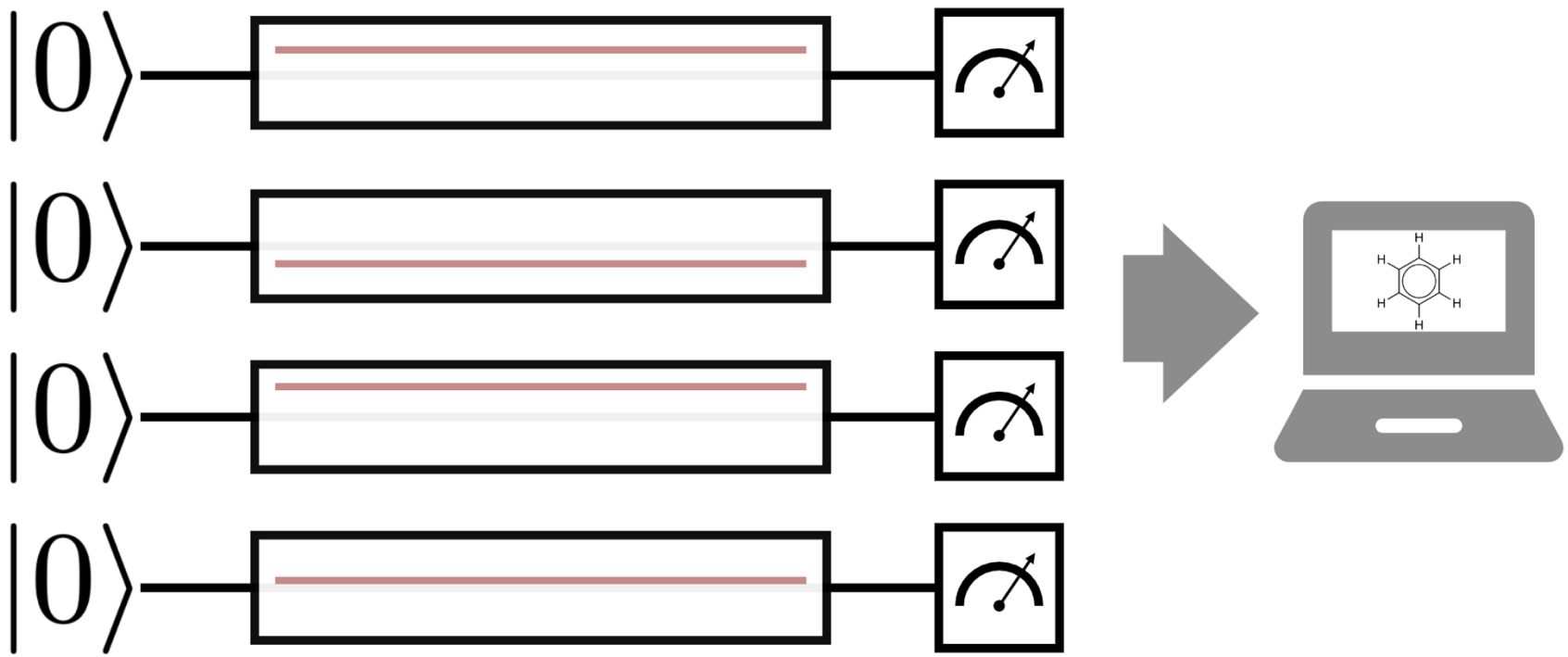
Meitei et al, npj Quantum Information 7, 155 (2021)

Pulse-based parameterization (“ctrl-VQE”)

Gate-based parameterization is a **special case** of pulse parameterization

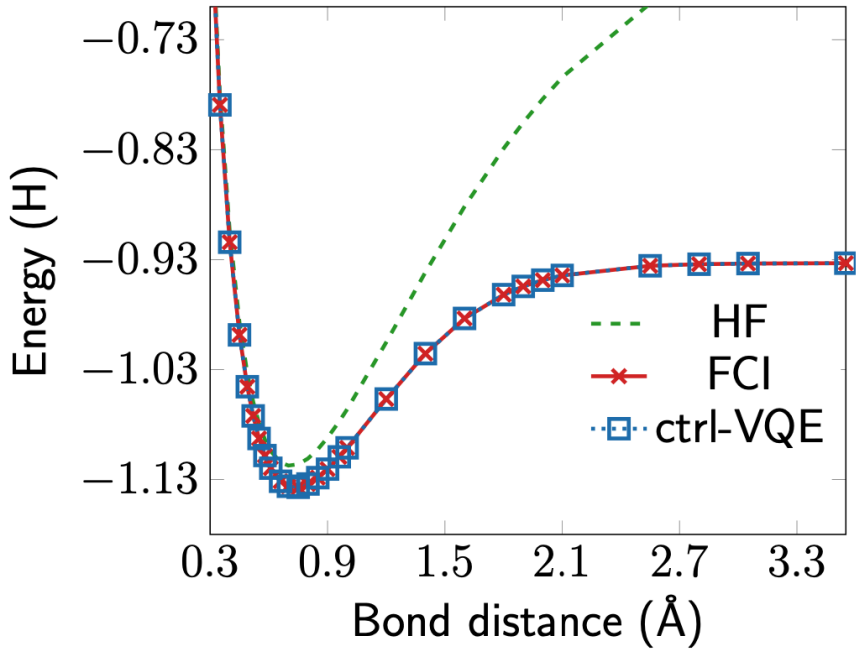
Lots of freedom in which terms to parameterize, and how to parameterize them

Meitei et al, npj Quantum Information 7, 155 (2021)

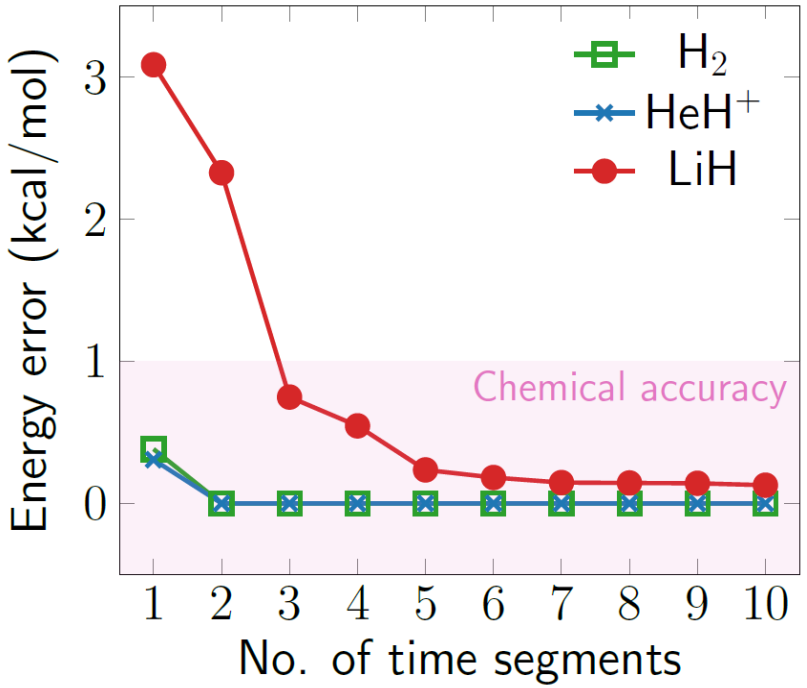

|0⟩
|0⟩
|0⟩
|0⟩
H
H
H
H
H
H

Numerical results

Dissociation curve for H2
(Indistinguishable from FCI)



Number of segments required for convergence



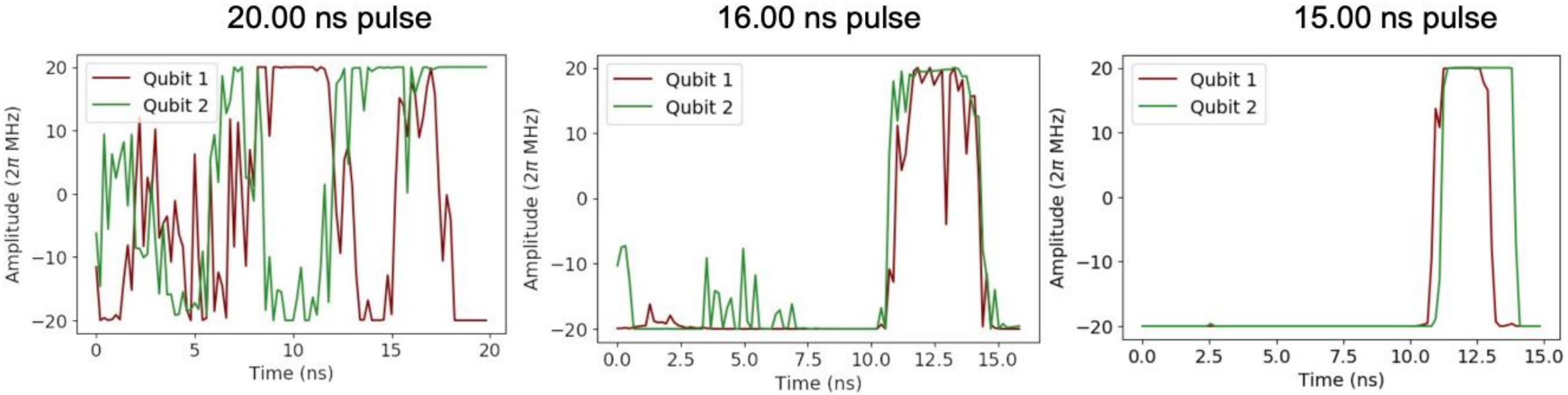
Orders of magnitude improvement:
E.g., for LiH: **80,000ns (gate-based UCCSD) vs. 50ns (ctrl-VQE)**

Meitei et al, npj Quantum Information 7, 155 (2021)

Ctrl-VQE: minimal evolution time

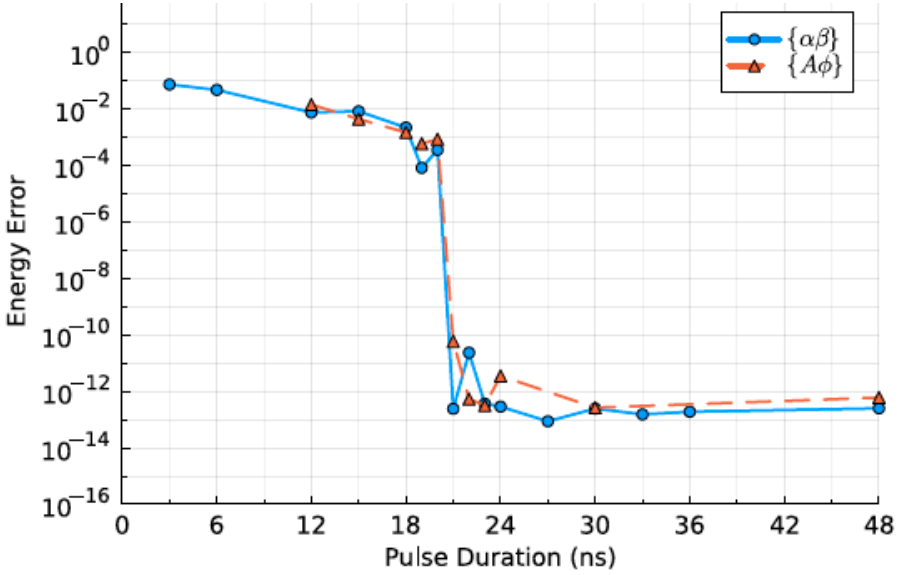
Minimal evolution time (MET): the min time required to transition between two particular quantum states (also referred to as quantum speed limit (QSL))

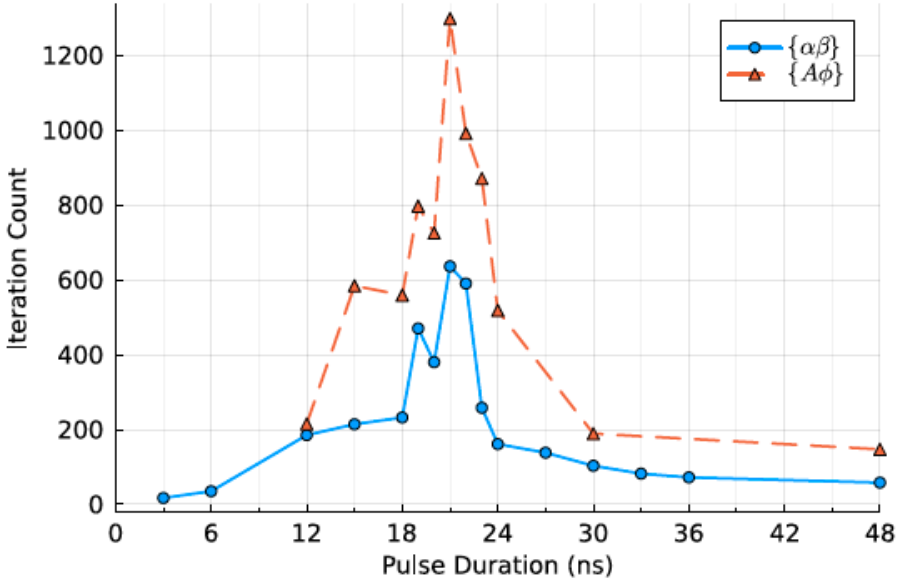
- Below a certain evolution time there is no solution
- Shortening the total pulse time leads to ‘bang-bang’ control (Pontryagin principle)



Asthana, Liu, et al,
Phys. Rev. Applied **19**, 064071 (2023)

Ctrl-VQE: parameterization and optimizability

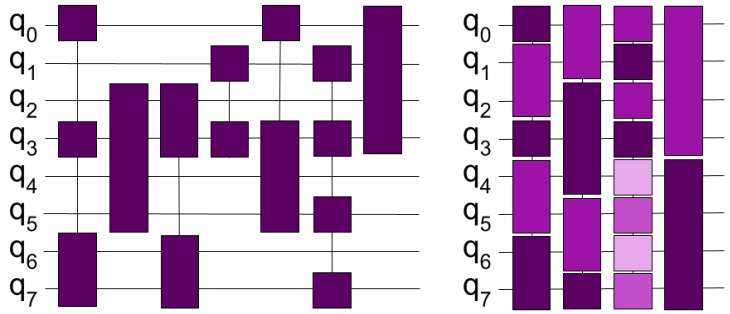


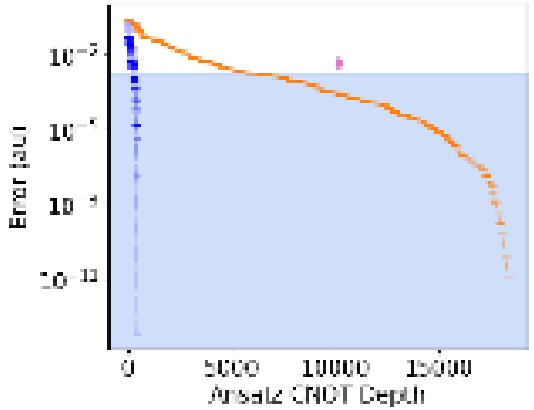


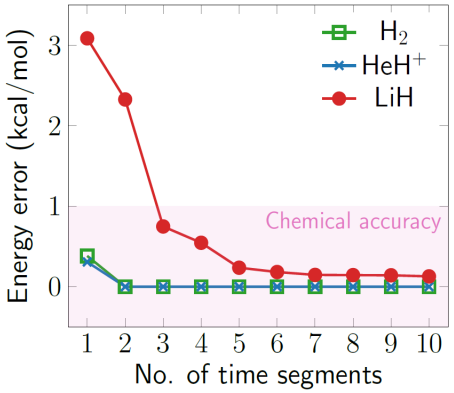
Sherbert, Amer, et al, arXiv:2405.15166

Summary

- ADAPT-VQE
- Algorithmic improvements since 2019
  - TETRIS-ADAPT-VQE (2022)
  - Efficient measurement of gradients (2023)
  - CEO pool (2024)
  - Hessian recycling (2024)
- Adaptive Gibbs state preparation
- Ctrl-VQE







Acknowledgements

Student/postdocs:


Mafalda Ramoa


Peter Anastasiou


Karunya Shirali


Kyle Sherbert


Hisham Amer


Jim Furches

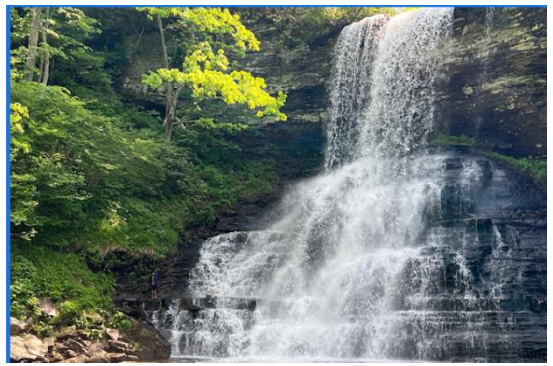
Senior collaborators (VT):


Edwin Barnes
Virginia Tech


Nick Mayhall
Virginia Tech

VIRGINIA TECH

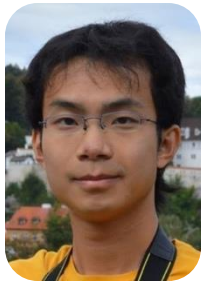
Postdoc positions open



VTQ

Sophia Economou
Physics
VTQ Director

Ed
Barnes
Physics



Charles
Cao
Physics

Arpit
Dua
Physics

Seva
Ivanov
Physics

Sumeet
Khatri
CS

Atul
Mantri
CS

Nick
Mayhall
Chemistry

Gretchen
Matthews
Math

Kyungwha
Park
Physics

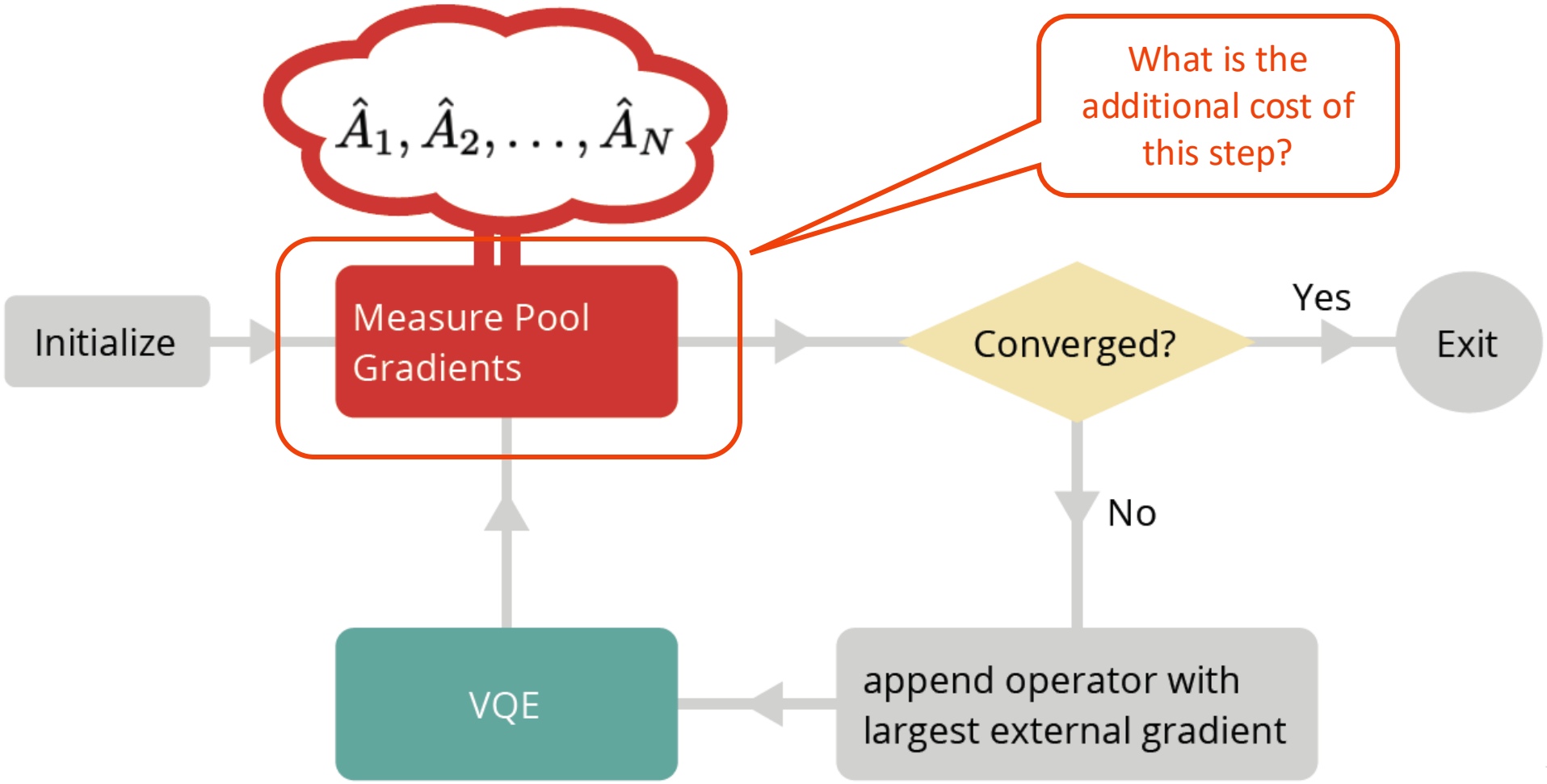
Vito
Scarola
Physics

Linbo
Shao
ECE

Jamie
Sikora
CS

Tianci
Zhou
Physics

ADAPT-VQE: measurement ‘overhead’



Let’s focus first on the qubit pool (XXXY, etc)

Measuring gradients in ADAPT

$$\left.\frac{\partial E}{\partial \theta_i}\right|_{\theta_i=0} = \left[\frac{\partial}{\partial \theta_i}\left\langle \Psi^{(k)}\right| e^{-\theta_i P_i} H e^{\theta_i P_i} \left|\Psi^{(k)}\right\rangle\right]\Bigg|_{\theta_i=0} = \left\langle \Psi^{(k)}\right| [H, P_i] \left|\Psi^{(k)}\right\rangle$$

Number of terms goes as $O(N^4)$

Size of Pauli pool also goes as $O(N^4)$

- How can we decrease the number of state preparations?

  Commuting observables can be measured simultaneously

- Which grouping heuristic should we use?

  Grouping Hamiltonian terms not ideal

Anastasiou et al, arXiv:2306.03227 (2023)

Measuring gradients in ADAPT—a better approach

For A, B, C Pauli strings:

If $[B, C] = 0$, then $[\,[A, B]\,, [A, C]\,] = 0$

By extension, the commutators of the elements of any mutually commuting set of Pauli strings with any one Pauli string commute

Anastasiou et al, arXiv:2306.03227 (2023)

Measuring gradients in ADAPT—a better approach

What does this mean for operators of the form $iY_iX_jX_kX_l$?

H term	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
					Y	Z	Y	Y	Z	Y	X	Z	X	X	Z	X
S_1	Y	X	X		X											
S_2	Y	X	X			X										
	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
S_3	Y													X	X	X

$O(N^3)$ commutators can be measured simultaneously!

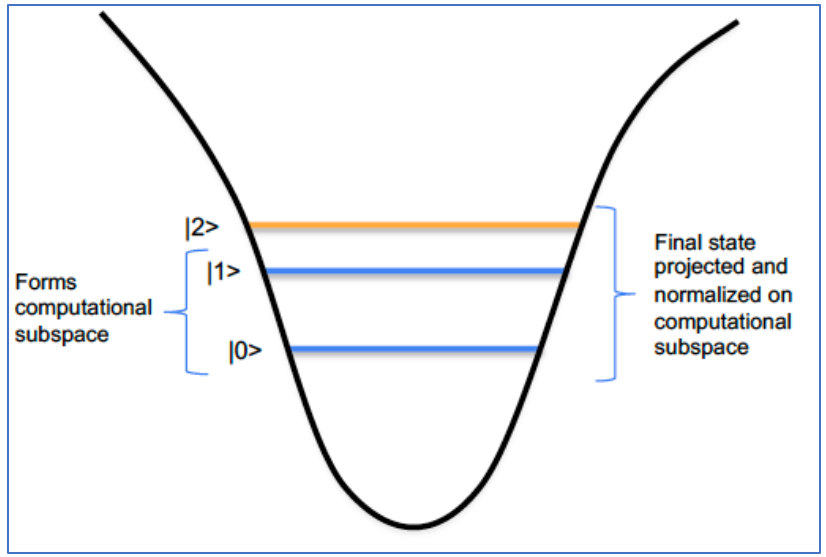
- We can group the commutators of any Hamiltonian term with all qubit operators into $2N$ groups
- Measuring term-by-term of the Hamiltonian means all observables in any given group have the same weight – lower shot-noise
- Measuring the pool gradient is only $\sim N$ times as expensive as a naïve VQE iteration
- QE and CEO pools contain same commuting terms → approach carries over

Anastasiou et al, arXiv:2306.03227 (2023)

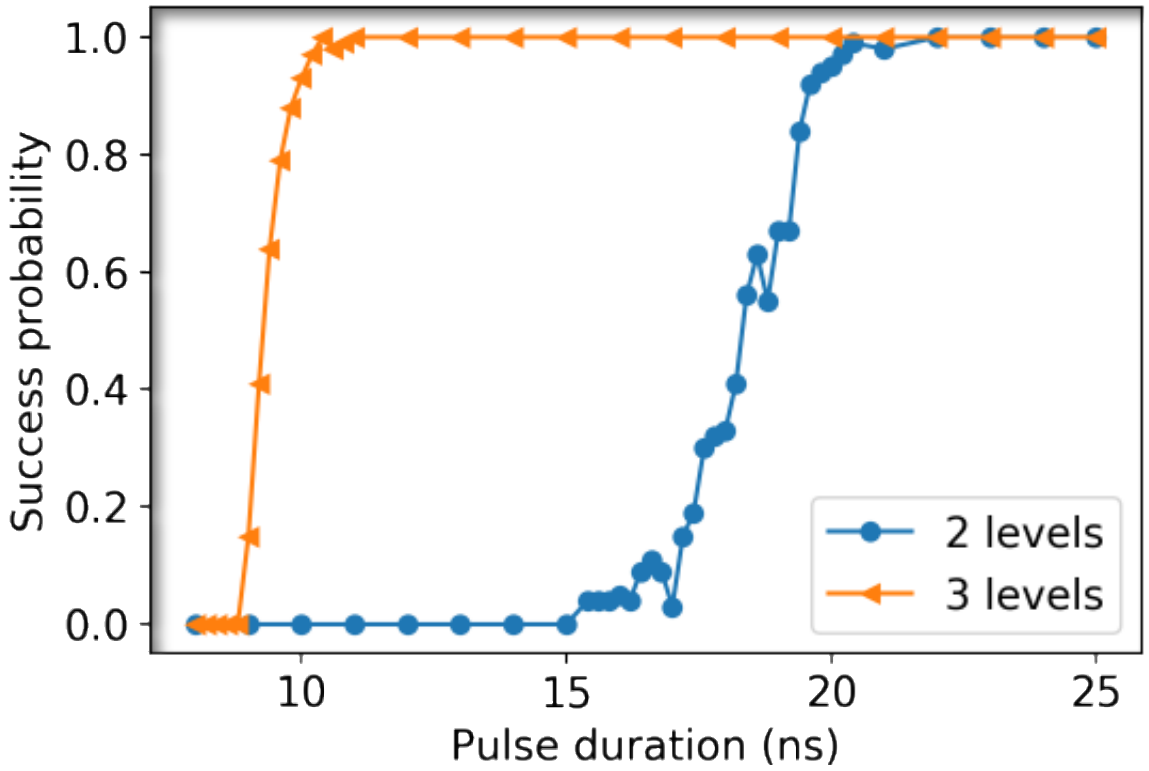
Ctrl-VQE: minimal evolution time

A (possibly surprising) finding:

Leakage to excited states outside qubit space ***speeds up evolution***



Probability to reach exact solution



Simulations for H2 at bond distance 1.5 A

Asthana, Liu, et al,
Phys. Rev. Applied **19**, 064071 (2023)