## Adaptive quantum simulation algorithms

Sophia Economou

VIRGINIA TECH...



CQCWS3, IPAM, Nov. 6, 2023

## Quantum simulation

- Predicting and designing properties of chemicals important for industry, medicine, agriculture, ...
- Understanding/designing physical systems with interesting properties: superconductivity, ...





The electrons in molecules and crystals behave quantum mechanically  $\rightarrow$  The size of the space scales exponentially with the number of electrons

## Quantum simulation

In general, representing n interacting two-level systems involves storing  $2 * 2^n$  real numbers. Scaling is exponential.

Even storing moderately-sized quantum states is infeasible with classical hardware

# of qubits	RAM required to store state
1	32 B
10	16 kiB
20	16 MiB
30	16 GiB
40	16 TiB
46	1 PiB

Use quantum systems that can be controlled to simulate other quantum systems

Digital simulation

#### System to be simulated



#### Quantum computer



#### Observable (Hamiltonian) in second quantized form

$$\circ \text{ Fermionic Hamiltonian } \quad H = -\sum_{i} \frac{\nabla_{r_{i}}^{2}}{2} - \sum_{i,j} \frac{Z_{i}}{|R_{i} - r_{j}|} + \sum_{i,j>i} \frac{Z_{i}Z_{j}}{|R_{i} - R_{j}|} + \sum_{i,j>i} \frac{1}{|r_{i} - r_{j}|}$$

• Second quantization (basis chosen, Coulomb integrals computed—*classical preprocessing step*)

$$\hat{H} = \sum_{i,j} h_{ij} a_i^{\dagger} a_j + \frac{1}{2} \sum_{i,j,k,l} h_{ijkl} a_i^{\dagger} a_j^{\dagger} a_k a_l$$

 ○ Wavefunction antisymmetric under exchange of two particles → fermionic operators satisfy anticommutation relations

$$\left\{ \hat{a}_{\alpha}, \hat{a}_{\beta} \right\} = 0, \quad \left\{ \hat{a}_{\alpha}^{\dagger}, \hat{a}_{\beta}^{\dagger} \right\} = 0, \quad \left\{ \hat{a}_{\alpha}, \hat{a}_{\beta}^{\dagger} \right\} = \delta_{\alpha,\beta}$$

In second quantization, the states correspond to occupation numbers of the spin-orbitals in the basis:

E.g. 1 electron in state  $\phi_2(\vec{x})$ , 1 electron in state  $\phi_4(\vec{x})$ 

$$\frac{1}{\sqrt{2}} \left( \phi_2(\vec{x}_1) \phi_4(\vec{x}_2) - \phi_4(\vec{x}_1) \phi_2(\vec{x}_2) \right) \to |0, 1, 0, 1, 0, 0, \dots \rangle$$

# A natural mapping to qubits: Jordan-Wigner

- We associate one qubit with each of our single-particle states Unoccupied  $\rightarrow |0\rangle$ , occupied  $\rightarrow |1\rangle$
- Map creation/annihilation operators onto qubit raising/lowering operators, being careful to preserve anticommutation relations  $a_i \rightarrow \prod_{j < i} Z_j \frac{1}{2} (X_i - i Y_i)$   $a_i^{\dagger} \rightarrow \prod_{j < i} Z_j \frac{1}{2} (X_i + i Y_i)$

Now, we can represent molecular electronic states/operators with states/operators on our quantum computer!

## What is the objective of quantum simulation?

- Create the state of interest (often ground or low excited state)
- Measure properties (total energy, correlation functions, response to fields, ...)
- Do this accurately and definitely better than classical approaches

## Analog vs digital simulation

#### • Analog quantum simulation

- $\circ~$  Create Hamiltonian of system on simulator
- Simulator has tunable parameters, study various regimes
- $\circ~$  Limited by native interactions of simulator
- Digital quantum simulation
  - $\circ~$  Evolution decomposed in elementary gates
  - $\circ$  Any problem can in principle be solved

# Phase estimation algorithm

- Find eigenenergy of many-body state with  $H_{tot} = H_a + H_b + \dots + H_k$
- n qubits to encode eigenstate  $|\Psi
  angle$  & ancilla qubits
- Conditionally apply time evolution  $U = \dots e^{-iH_{tot}t}$



- Conditional-U through Trotterization:  $U = (e^{-iH_a t/N}e^{-iH_b t/N} \dots e^{-iH_k t/N})^N$
- Exact for  $N \rightarrow \infty$

# Phase estimation algorithm

- Find eigenenergy of many-body state with  $H_{tot} = H_a + H_b + \dots + H_k$
- n qubits to encode eigenstate  $|\Psi
  angle$  & ancilla qubits
- Conditionally apply time evolution  $U = \dots e^{-iH_{tot}t}$



- Conditional-U through Trotterization:  $U = (e^{-iH_at/N}e^{-iH_bt/N} \dots e^{-iH_kt/N})^N$
- Exact for  $N \rightarrow \infty$

Linear combination of unitaries



Variational Quantum Eigensolver (VQE)



Recent reviews:

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

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

#### How do the parameterized circuits look?

- Generally, there is a lot of freedom in selecting the gates (unitary operators) that enter the circuit
- Example:



# Outline

• Typical ansatze & issues

• Our approach: Adaptive quantum simulation

• Control-VQE: optimizing at the pulse level

### Most widely considered ansatze

#### I. Hardware-efficient ansatz

- Tailored to chosen platform
- Inefficient—too much of the Hilbert space sampled
- Difficult to optimize (barren plateaus) McClean et al., Nat. Commun. 9, 4812 (2018)





Kandala et al, Nature 549, 242 (2017)

### Most widely considered ansatze

#### II. Chemistry-inspired ansatz (UCC)

- Generalizes classical simulations
- Impractically long circuits
- 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)

Most widely considered ansatze

- Quantum-resource inefficient: long circuits
- Not problem-tailored
- Possible trainability issues

### 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,  $A_{\rm m}$
- Applies unitaries one by one :  $U_m = \exp(\theta_m A_m)$  to a reference state



### 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  $\theta_j$

$$\frac{\partial}{\partial \theta_j} \langle \Psi_0 | e^{-\theta_j A_j} H e^{\theta_j A_j} | \Psi_0 \rangle |_{\theta_j = 0}$$

### 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  $\theta_j$

$$\frac{\partial}{\partial \theta_{j}} \langle \Psi_{0} | e^{-\theta_{j}A_{j}} H e^{\theta_{j}A_{j}} | \Psi_{0} \rangle |_{\theta_{j}=0} = \langle \Psi_{0} | [H, A_{j}] | \Psi_{0} \rangle$$

## 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  $\theta_j$

$$\frac{\partial}{\partial \theta_{j}} \langle \Psi_{0} | e^{-\theta_{j}A_{j}} H e^{\theta_{j}A_{j}} | \Psi_{0} \rangle |_{\theta_{j}=0} = \langle \Psi_{0} | [H, A_{j}] | \Psi_{0} \rangle$$

New operator  $\rightarrow$  measure on hardware

#### ADAPT-VQE overview

Terminate: when

#### Inputs:

- Hamiltonian
- Initial state



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

#### Convergence in ADAPT-VQE

$$\frac{\partial}{\partial \theta_{j}} \langle \Psi_{0} | e^{-\theta_{j}A_{j}} H e^{\theta_{j}A_{j}} | \Psi_{0} \rangle |_{\theta_{j}=0} = \langle \Psi_{0} | [H, A_{j}] | \Psi_{0} \rangle$$

The algorithm terminates when this is zero (i.e., very small)

- This is satisfied by any energy eigenstate
- Strictly speaking, ADAPT finds energy eigenstates (not necessarily the g.s.)
- In practice, it always finds the g.s. given a proper pool

ADAPT-VQE with fermionic pool (classical simulations)



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

#### Comparing ADAPT to other orderings



 $BeH_2$ bond distance 2.39 Å

Grimsley, Economou, Barnes, Mayhall, Nature Commun. 10, 3007 (2019)

# Trainability of ADAPT-VQE

- ADAPT produces compact, problem tailored ansatze
- 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)

#### Intuitive picture



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

### How should the operator pool be chosen?



• We can choose it according to hardware constraints

### How should the operator pool be chosen?



- We can choose it according to hardware constraints
  - Hardware-efficient pool  $\{e^{i\theta_j P_j}\}$ , where  $P_i$  is a Pauli string with up to 4 Paulis
  - For simplicity, take doubles and drop Z's: XYYY, XXXY ("qubit-ADAPT-VQE")

#### Qubit ADAPT-VQE—results



Tang, Shkolnikov, Barron, Grimsley, Mayhall, Barnes, Economou, PRX Quantum 2, 020310 (2021)

## How should the operator pool be chosen?



- We can choose it according to hardware constraints
- Ideally, the pool size should be as small as possible, to reduce quantum resources
- How can we guarantee that it contains enough and the right type of operators?

#### Complete pools

$$|\psi^{ADAPT}(\vec{\theta})\rangle = e^{\theta_n A_n} \dots e^{\theta_2 A_2} e^{\theta_1 A_1} |\psi^{ref}\rangle = e^{\sum_i \phi_i B_i} |\psi^{ref}\rangle$$

where 
$$\{B_i\} = \{A_1, A_2, ..., [A_1, A_2], ..., [A_1, [A_2, A_3]], ...\}$$

We have a complete pool when states  $B_i |\psi\rangle$  form a complete basis (where  $|\psi\rangle$  is an arbitrary state)

Minimal complete pools (MCPs)

*Minimal complete pool:* smallest sized complete Pauli pool The minimal size of complete pools is linear in the nr of qubits: 2n-2

Example of min complete pool
$$\begin{cases}
G_1 = ZYII \dots I, & G_2 = IZYII \dots I, \\
G_3 = IIZYII \dots I, & \dots, & G_{n-2} = II \dots IZYI, & G_{n-1} = II \dots IZY, \\
G_n = YII \dots I, & G_{n+1} = IYII \dots I, \\
G_{n+2} = IIYII \dots I, & \dots, & G_{2n-3} = II \dots IYII, & G_{2n-2} = II \dots IYI
\end{cases}$$

We also have a simple criterion to determine whether a minimal pool is complete that scales polynomially with system size

Shkolnikov, Mayhall, Economou, Barnes, Quantum 7, 1040 (2023)

## Minimal complete pools—cont.

Given a pool of size 2n-2, how do we know it's complete?

- Taking all the commutators to check the full algebra scales exponentially with system size
- We have a simple, polynomially scaling criterion to determine whether a minimal Pauli pool is complete:

If the pool operators cannot be split into two fully mutually commuting sets, then the MCP is complete



Shkolnikov, Mayhall, Economou, Barnes, Quantum 7, 1040 (2023)

## Incorporating symmetries

- Using random MCPs not aware of symmetries can cause convergence issues
- This can be addressed using symmetry-preserving MCPs
- For starting the algorithm, include "starters", operators that preserve symmetries
- These need to involve two-particle excitations (four nontrivial Paulis), since HF is the lowest-energy product state

- Symmetry-preserving MCPs have fewer than 2n 2 operators
- This is because we do not need to span the full Hilbert space



Shkolnikov, Mayhall, Economou, Barnes, Quantum 7, 1040 (2023)

### Is ADAPT-VQE the most efficient way to construct circuits?



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, Chen, et al, arXiv:2209.10562

#### **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



TETRIS-ADAPT-VQE: results



Blue: TETRIS-ADAPT Orange: standard ADAPT

Anastasiou, Chen, et al, arXiv:2209.10562

# Trainability of TETRIS-ADAPT-VQE



Anastasiou, Chen, et al, arXiv:2209.10562

#### ADAPT-VQE overview



# Measuring gradients in ADAPT

$$\frac{\partial E}{\partial \theta_i}\Big|_{\theta_i=0} = \left[\frac{\partial}{\partial \theta_i} \left\langle \Psi^{(k)} \middle| e^{-\theta_i P_i} H e^{\theta_i P_i} \middle| \Psi^{(k)} \right\rangle\right]\Big|_{\theta_i=0} = \left\langle \Psi^{(k)} \middle| [H, P_i] \middle| \Psi^{(k)} \right\rangle$$
Number of terms Size of Pauli pool goes as  $O(N^4)$  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? Not ideal approach, scaling  $\sim O(N^{16})$

#### 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!

### Measuring gradients in ADAPT—a better approach



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

- 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 ٠ shot-noise
- Measuring the pool gradient is only  $\sim N$  times as expensive as a naïve VQE iteration ٠

## Beyond molecules: lattice models

#### Challenge for extended Hamiltonians:

- Full Pauli pool scales as 4<sup>N</sup>
- Using ad hoc smaller pool does not respect periodicity

#### Solution:

- Exploit periodicity (translational symmetry) of lattice
- Algorithm:
  - Run ADAPT with full pool on small tile
  - If there are gradient degeneracies, repeat multiple times
  - Collect operators chosen by ADAPT
  - Tile these operators to larger instances (pad with identities)
- The resulting pool is *linear* in N
- We have identified the conditions for the pool to be *complete*



Van Dyke et al, arXiv:2206.14215 (to be updated soon)

## Tiling example: XXZ model

- Use resulting Pauli pool for small instance (1x3 tile for 1D, 2x2 tile for 2D): take chosen operators, tile them to create pool for larger instances
- Run 1000 trials, pick operators randomly in case of degenerate gradients

1D case:

- For a 1x3 tile, ADAPT always converges in three steps
- ADAPT selects operator set IXY, IYX, XYI, XIY, YXI, YIX, ZXY, YZX, YXZ, XZY, ZYX, XYZ across several runs

2D case:

- For a 2x2 tile, ADAPT always converges in five steps
- ADAPT selects operator set IIXY, IIYX, XYII, YXII, XYZI, IZXY, XZIY, ZYIX, YZXI, YIZX, XIYZ, IZYX, IXZY, XYIZ, ZIYX, ZXYI, YXZI, YXIZ, IYXZ, ZIXY, ZZYX, YXZZ, ZZXY, XYZZ

XXZ model with tiling: results for up to 16 qubits for 1D

$$H = \sum_{j=1}^{L-1} \left( \frac{J_{xy}}{2} (S_{j+1}^+ S_j^- + h.c.) + J_z S_{j+1}^z S_j^z \right)$$



Number of qubits

*Exponential reduction in pool size* 

Van Dyke et al, arXiv:2206.14215 (to be updated soon)

## XXZ model with tiling: results for up to 12 qubits for 1D

- Computational complexity greater due to increased connectivity of lattice
- Tiling strategy able to achieve good accuracy (error < 0.05%)





## Beyond pure states: 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

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

- This has useful applications in quantum simulation, quantum machine learning, quantum optimization, etc., but is a hard problem in general
- Need additional ancilla qubits to purify the state on the whole system

• Standard approach: prepare state that minimizes 
$$F\left(
ho(ec{ heta})
ight) = ext{Tr}ig(
ho(ec{ heta})H + T
ho(ec{ heta})\ln
ho(ec{ heta})ig)$$

- 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) Some remarks on prior work

- The thermofield double has been considered as a tool
- Create it based on QAOA-like ansatz [1]

$$H_{AD} = \sum_{k=1}^{N_D} (X_{D_k} X_{A_k} + Y_{D_k} Y_{A_k} + Z_{D_k} Z_{A_k})$$

- Algorithm starts from highly entangled state (fully thermal)
- But interesting/challenging regime is low T (modest/low entanglement)
- Algorithm largely undoes entanglement  $\rightarrow$  wasteful in q resources, deep circuits

Eigenstate

of  $H_{AD}$ 

## Gibbs-ADAPT-VQE

• Our approach:

✓ New objective function that is easier to measure:  $C\left(\rho(\vec{\theta})\right) = -\operatorname{Tr}\left(\rho_G(T)\rho(\vec{\theta})\right) + \frac{1}{2}\operatorname{Tr}\left(\rho(\vec{\theta})^2\right)$ ✓ ADAPT-VQE approach to grow the ansatz

- Two strategies, one good for low T, one for high T
- Low *T* approach:
  - Start with partially entangled data-ancilla state
  - Operator pool: all 1- and 2-qubit Paulis on combined data/ancilla system
  - Pool requires initial state be only partially entangled (use non-optimized random y-rotations and CNOTs)
  - Use ADAPT criterion to minimize Gibbs energy iteratively

#### Warren et al, arXiv: 2203.12757

#### Gibbs-ADAPT-VQE: numerical results for XY Hamiltonian

- Pool requires initial state be only partially entangled (use non-optimized random y-rotations and CNOTs)
- Ancilla system can be any size, allowing resource savings (though  $N_A = N_D$  needed at larger T)
- Improvement (reduction) in number of CNOTs, number of ancilla qubits



Warren et al, arXiv: 2203.12757

Other applications of ADAPT-VQE

• Optimization: ADAPT-QAOA, where we use ADAPT to determine mixers

Zhu et al, PRR 4, 033029 (2022) Chen et al, arxiv:2205.12283





- Nuclear physics: LMG ('Lipkin' model), where ADAPT finds g.s. across phase transition Romero et al, Phys. Rev. C 105, 064317
- ADAPT + randomness

Magann, Economou, Arenz, PRR 5, 033227 (2023)

• Time dynamics: ADAPT-VQE for time-evolved state (minimize 'McLachlan distance') Yao et al, PRX Quantum 2, 030307 (2021)

# Outline

• Typical ansatze & issues

• Our approach: Adaptive quantum simulation

• Control-VQE: optimizing at the pulse level

### Hardware currently noisy even for VQEs

Pulse-level optimization:

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

Meitei et al, npj Quantum Information 7, 155 (2021); arXiv:2008.04302

See also Yang et al, Phys. Rev. X 7, 021027 (2017) for pulse-based QAOA





Evolution time (ns)

Orders of magnitude improvement:

E.g., for LiH: 80,000ns (gate-based UCCSD) vs. 50ns (ctrl-V<sup>and-leakage (bottom)</sup> along the time evolution steps with op-

FIG. 5: Energy difference between ctrl-VQE and FCI (top) and leakage (bottom) along the time evolution steps with optimal pulses of different durations for H<sub>2</sub> with a bond distance of 0.75 Å. The red and purple lines with T = 29 ns are distinct solutions to the same optimization. Optimized pulse parameters are provided in the SI.

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

## Ctrl-VQE: minimal evolution time

#### In more recent work we showed:

- Using excited states outside qubit space *speeds up evolution*
- Shortening the total pulse time leads to 'bang-bang' control
- Below a certain evolution time there is no solution





Simulations for H2 at bond distance 1.5 A

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

## Summary

Adaptive variational quantum eigensolver (ADAPT-VQE)

• TETRIS-ADAPT-VQE

- Efficient measurement of gradients
- Algorithms for periodic (pure and thermal) systems

Control-VQE







2



## Acknowledgements

#### Student/postdocs:





Senior collaborators (VT):

Peter Anastasiou Ho Lun Tang

Ada Warren

Vlad Shkolnikov



Yanzhu Chen Karunya Shirali

Funding:



Edwin Barnes Virginia Tech



Nick Mayhall Virginia Tech



![](_page_61_Picture_16.jpeg)

![](_page_61_Picture_17.jpeg)

![](_page_61_Picture_18.jpeg)

*External collaborators:* Christian Arenz, ASU Alicia Magann, Sandia

### Postdoc positions open

![](_page_62_Picture_1.jpeg)

![](_page_62_Picture_2.jpeg)

![](_page_62_Picture_3.jpeg)

#### economou@vt.edu

![](_page_62_Picture_5.jpeg)

![](_page_62_Picture_6.jpeg)

![](_page_62_Picture_7.jpeg)

![](_page_62_Picture_8.jpeg)

![](_page_62_Picture_9.jpeg)

![](_page_62_Picture_10.jpeg)

Sophia Economou Physics **VTQ Director** 

![](_page_62_Picture_12.jpeg)

Nick

Mayhall

Chemistry

Gretchen Matthews Math

![](_page_62_Picture_14.jpeg)

Park

**Physics** 

Ed Barnes **Physics** 

![](_page_62_Picture_16.jpeg)

Charles Cao Physics

![](_page_62_Picture_18.jpeg)

![](_page_62_Picture_20.jpeg)

![](_page_62_Picture_22.jpeg)

Arpit Dua **Physics** 

![](_page_62_Picture_24.jpeg)

Seva lvanov **Physics** 

![](_page_62_Picture_26.jpeg)

![](_page_62_Picture_27.jpeg)

Sumeet Khatri CS

CS

Atul Mantri CS

![](_page_62_Picture_30.jpeg)

![](_page_62_Picture_31.jpeg)

Vito Scarola **Physics** 

![](_page_62_Picture_33.jpeg)

Shao

ECE

![](_page_62_Picture_35.jpeg)

![](_page_62_Picture_36.jpeg)

![](_page_62_Picture_37.jpeg)

Tianci Zhou Physics