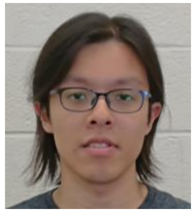
Efficient variational quantum eigensolvers for NISQ hardware

Sophia Economou

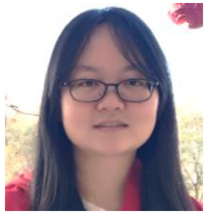
VIRGINIA TECH

Frontiers of Quantum Computing and Quantum Dynamics, KITP (virtual), October 19 2020

Collaborators in these projects



Ho Lun Tang



Linghua Zhu

George Barron

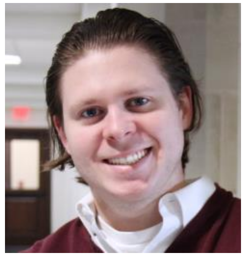
Vlad Shkolnikov

Fernando Calderon-Vargas

Bryan Gard (now at GTRI)

Edwin Barnes
Virginia Tech

Nick Mayhall
Virginia Tech



Harper Grimsley
Virginia Tech

VIRGINIA TECH

Analog vs digital simulation

- Feynman's original vision was for analog quantum simulation
  - Create Hamiltonian of system of interest on simulator
  - Simulator is controllable, has tunable parameters
  - Study system in various regimes

- Digital quantum simulation
  - In quantum computing, every evolution can be decomposed in set of elementary quantum gates
  - Algorithm for finding eigenenergies of many body fermion systems

Abrams and Lloyd, Phys. Rev. Lett. **79**, 2586 (1997)

Digital quantum simulation

- Fermionic Hamiltonian $H = -\sum_i \frac{\nabla^2_{r_i}}{2} - \sum_{i,j} \frac{Z_i}{|R_i - r_j|} + \sum_{i,j>i} \frac{Z_i Z_j}{|R_i - R_j|} + \sum_{i,j>i} \frac{1}{|r_i - r_j|}$
- Second quantization (basis chosen, Coulomb integrals computed)

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

- Number of qubits = number of orbitals
- Jordan-Wigner transformation
  - Each orbital is mapped onto a qubit: |0> → unoccupied orbital, |1> → occupied orbital
  - Need to preserve fermionic anticommutation relations

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

  - Qubits are distinguishable → Pauli Z strings

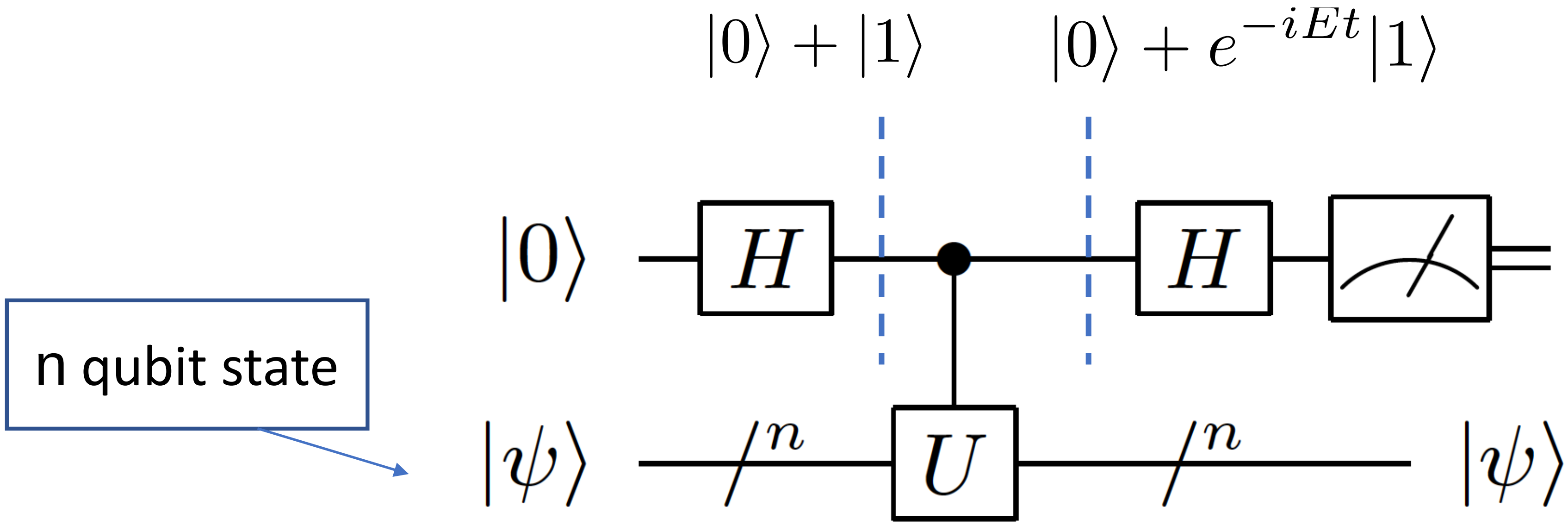
$$a_p^\dagger = \left(\prod_{m<p} \sigma_m^z\right) \sigma_p^+$$

  - The Hamiltonian is then expressed as $H = \sum_{\alpha=1}^{T} h_\alpha P_\alpha$

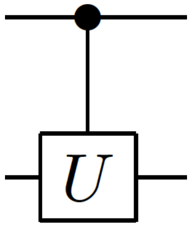
For recent reviews see
Rev. Mod. Phys. **92**, 015003 (2020)
Chem. Rev. 2019, 119, 19, 10856 (2019)

Phase estimation algorithm

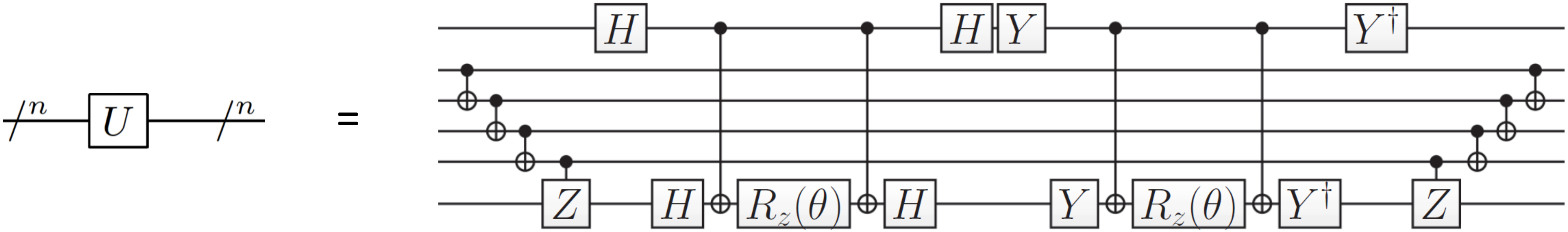
- Find eigenenergy of many-body state
- n qubits to encode eigenstate + ancilla qubits
- Conditionally apply time evolution $U = e^{-i\mathcal{H}t}$



- Conditional-U through Trotterization: $U=(e^{-iH_aT/N}\, e^{-iH_bT/N})^N$
- Exact for $N \rightarrow \infty$

Extremely long circuits required for the multi-qubit gate 

E.g., exponentiating the hopping term $c^{\dagger}_{p,\sigma}c_{q,\sigma} + c^{\dagger}_{q,\sigma}c_{p,\sigma}$:



As a result, PEA is beyond the scope of existing and near-future devices

NISQ era

- Building a universal quantum computer is a formidable task
- Can we do something technologically interesting before that, even with **n**oisy **i**ntermediate **s**cale **q**uantum (NISQ) devices?
- Simulation of many-body systems is probably the most interesting known application of quantum processors
- Use of hybrid classical-quantum algorithms

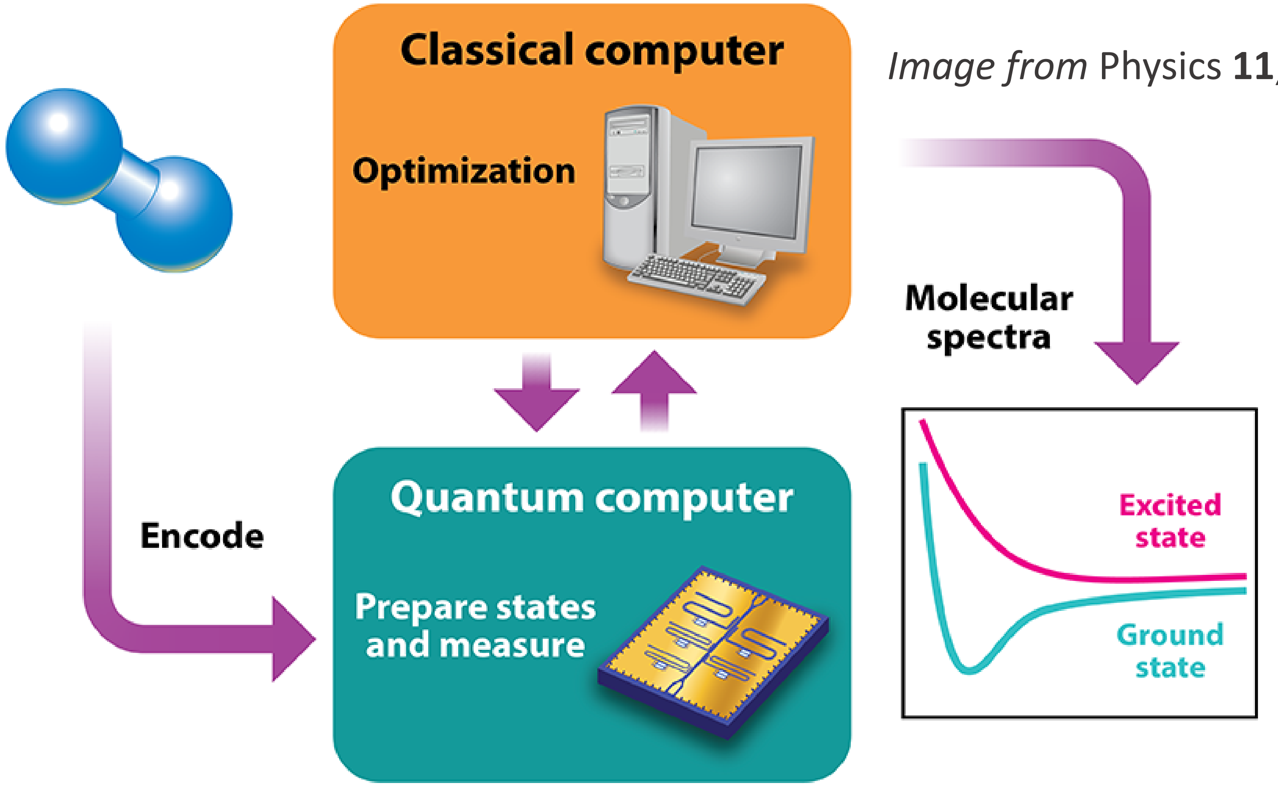
Preskill, arXiv:1801.00862

Variational quantum eigensolvers

Nat. Comm. **5**, 4213 (2014)
New J. Phys. **18**, 023023 (2016)

Variationally change $\vec{\theta}$

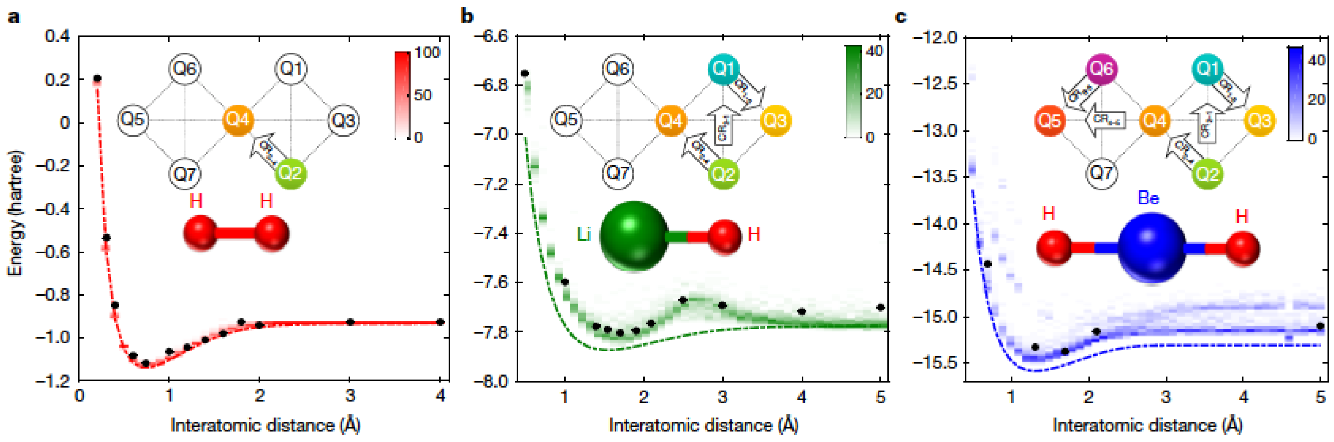
Image from Physics **11**, 14 (2018)



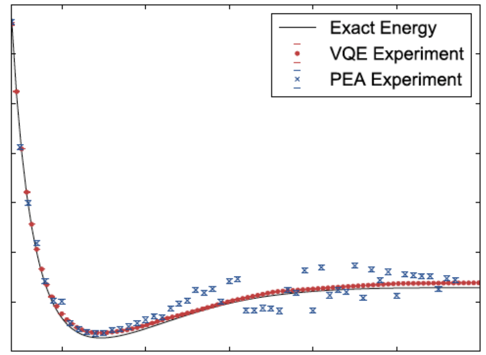
Ansatz $|\Psi(\vec{\theta})\rangle = T(\vec{\theta})|\Psi_{ref}\rangle$

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

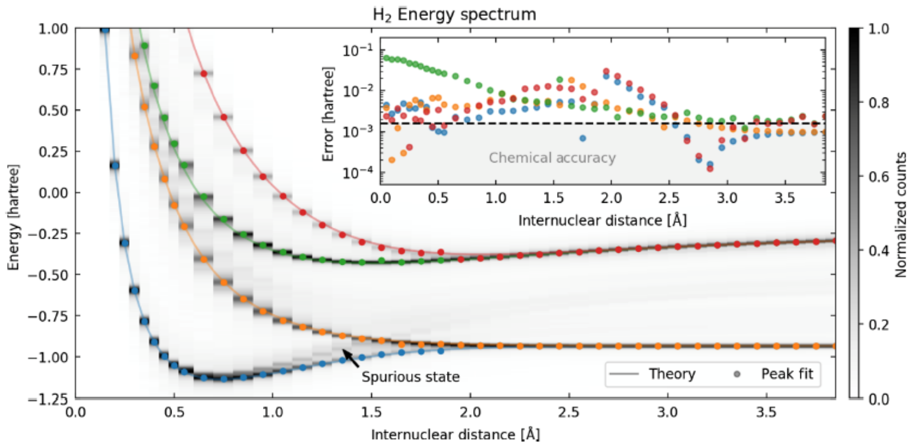
Recent highlights



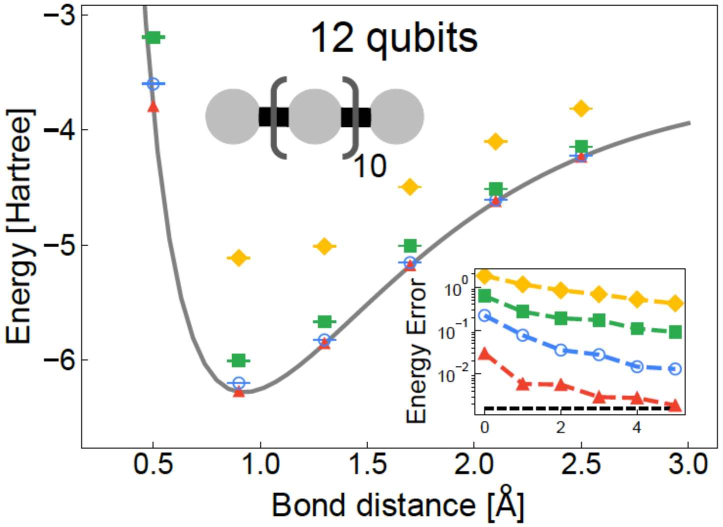
*IBM group, Nature **549**, 242 (2017)*



*Google group, PRX **6**, 031007 (2016)*



*Siddiqi group, Phys. Rev. X **8**, 011021 (2018)*



Google group, Science 369, 1084 (2020)

Outline

- Fermionic problems
  - Symmetry enforcing circuits
  - ADAPT-VQE algorithm

- Optimization (many body Ising) problems
  - ADAPT QAOA

Properties of a good ansatz

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

Choice of ansatz is crucial!

- Quantum coherence is very limited → shallow circuit
- Classical optimization is not infinitely powerful→ not too many optimization parameters
- Need to span the space where the solution lives (exactness)

Most widely considered ansatze

Hardware-efficient

Both are generic

Advantages:

- Designed to work with hardware
- Highly expressible

$$|\Phi(\boldsymbol{\theta})\rangle = \prod_{q=1}^{N} [U^{q,d}(\boldsymbol{\theta})] \times U_{\text{ENT}} \times \prod_{q=1}^{N} [U^{q,d-1}(\boldsymbol{\theta})] \times \cdots \times U_{\text{ENT}} \times \prod_{q=1}^{N} [U^{q,0}(\boldsymbol{\theta})] |00\ldots 0\rangle$$

Disadvantages:

- Ad hoc (generally not exact)
- Inefficient—too much of the Hilbert space sampled
- Barren plateaus[1] for generic circuits

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

Chemistry-inspired: UCCSD

Advantage:

- Performs well in classical simulation

$$|\psi\rangle = e^{\hat{T}}|\phi_0\rangle \qquad \hat{T} = \hat{T}_1 + \hat{T}_2$$

$$\hat{T}_1 = \sum_{ia} t_i^a \hat{a}_a^\dagger \hat{a}_i \qquad \hat{T}_2 = \frac{1}{4}\sum_{ijab} t_{ij}^{ab} \hat{a}_a^\dagger \hat{a}_b^\dagger \hat{a}_j \hat{a}_i$$

Disadvantages:

- Translating fermionic operators into efficient gate circuit challenging
- Trotterized form long, *not unique, do not always achieve chemical accuracy*[2]
- Not proven to be exact

[2] Grimsley, Claudino, et al., J. Chem. Theory Comput. 2020, 16, 1, 1-6

Our approach: problem-tailored ansatze

- Symmetry preserving circuits

- ADAPT-VQE

Features:

✓shallow circuits;

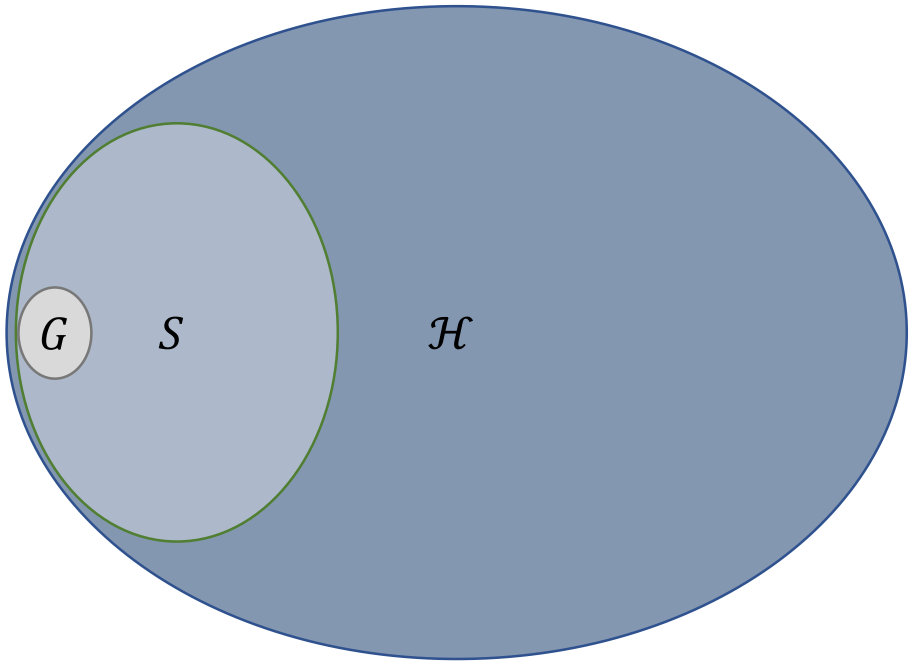
✓small/minimal number of optimization parameters;

✓exactness

Symmetry preserving ansatze

In a nutshell

- Interested in creating states, not U
- Count and parameterize relevant states with given symmetry
- Impose the relevant symmetries at the circuit level



Enforcing particle number symmetry

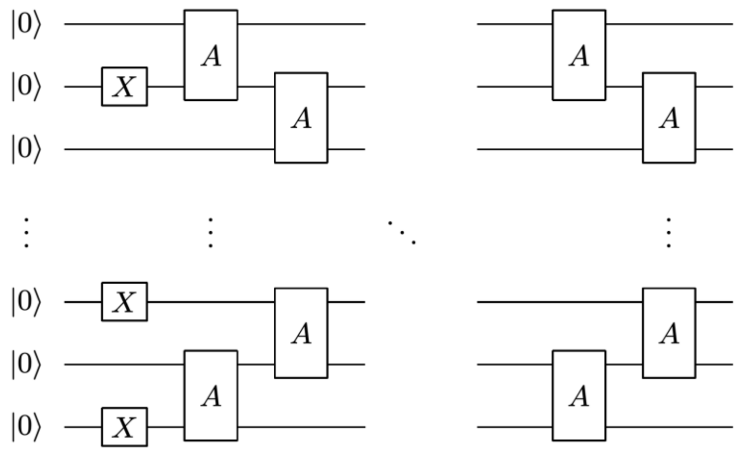
- System with *n* orbitals → *n* qubits; arbitrary state described by $2 * 2^n - 2$ real parameters
- For system of *m* fermions, min nr of variational parameters is $2 * \binom{n}{m} - 2$
- Key ingredient: particle preserving gate (Barkoutsos et al, PRA 98, 022322 (2018)):

$$A(\theta, \phi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & e^{i\phi}\sin\theta & 0 \\ 0 & e^{-i\phi}\sin\theta & -\cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

For n-orbital, m-fermion state:

- Put register into appropriate, separable basis state, e.g. |0101…0101>
- Apply layers of A gates until $\binom{n}{m}$ A gates are placed
- Fix any two of the ϕ parameters

✓ We can generate any state in the subspace with 100% fidelity
✓ Min number of optimization parameters
✓ Hardware-friendly: only requires nearest neighbor coupling



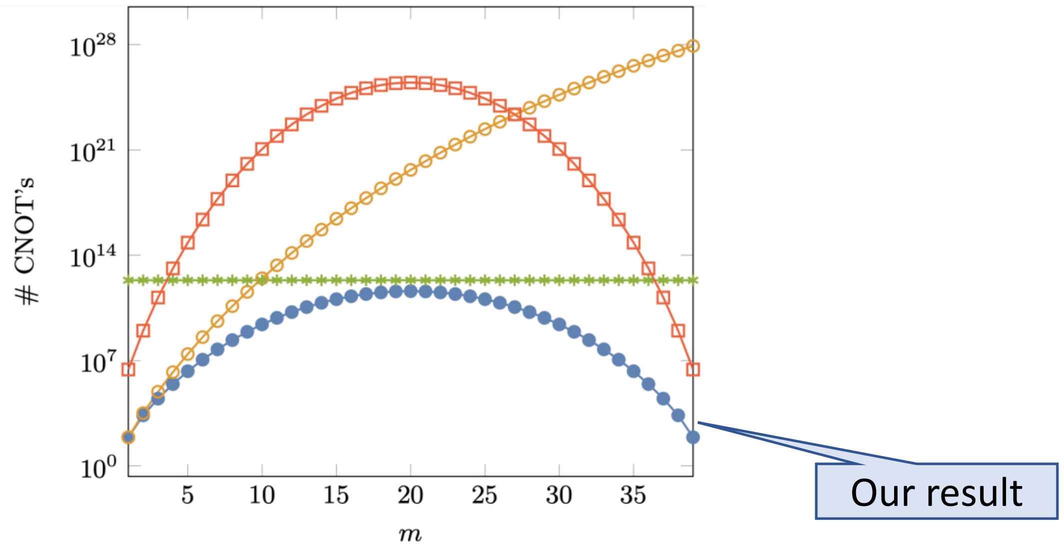
- Time-reversal symmetry: real states, number of parameters $\binom{n}{m} - 2$ → set $\varphi = 0$

Gard, Zhu, Barron, et al, *npj Quantum Inf* **6**, 10 (2020)

Enforcing particle number symmetry—results

CNOT count as function of nr of fermions (m)

nr of qubits = 40

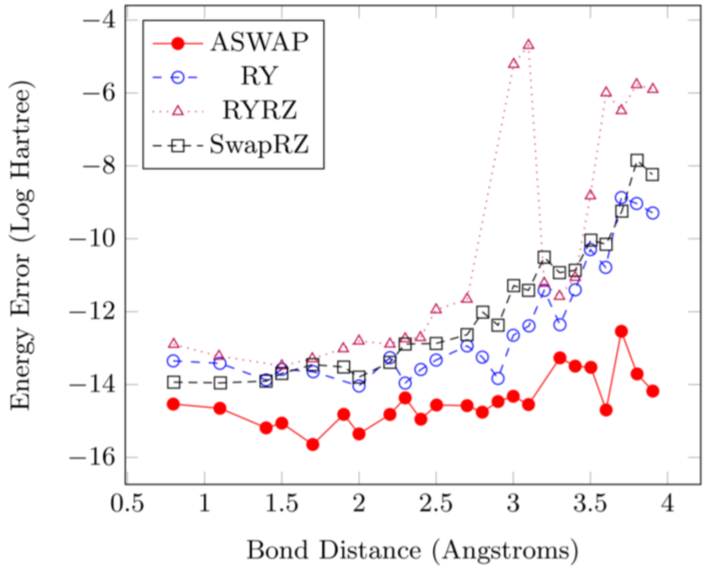


- Orders of magnitude lower CNOT count compared to prior works [1]

[1] PRA, 79, 042335 (2009); PRA 71, 052330 (2005); PRA 64, 022319 (2001)

H_2 ground state energy error

H_2 ground state energy error



- Lower nr of parameters (and often CNOT count) compared to other ansatze

Ansatz	CNOTs	Pars.
ASWAP	6	3
RY	18	16
RYRZ	18	32
SwapRZ	72	34

Gard, Zhu, Barron, et al, *npj Quantum Inf* **6**, 10 (2020)

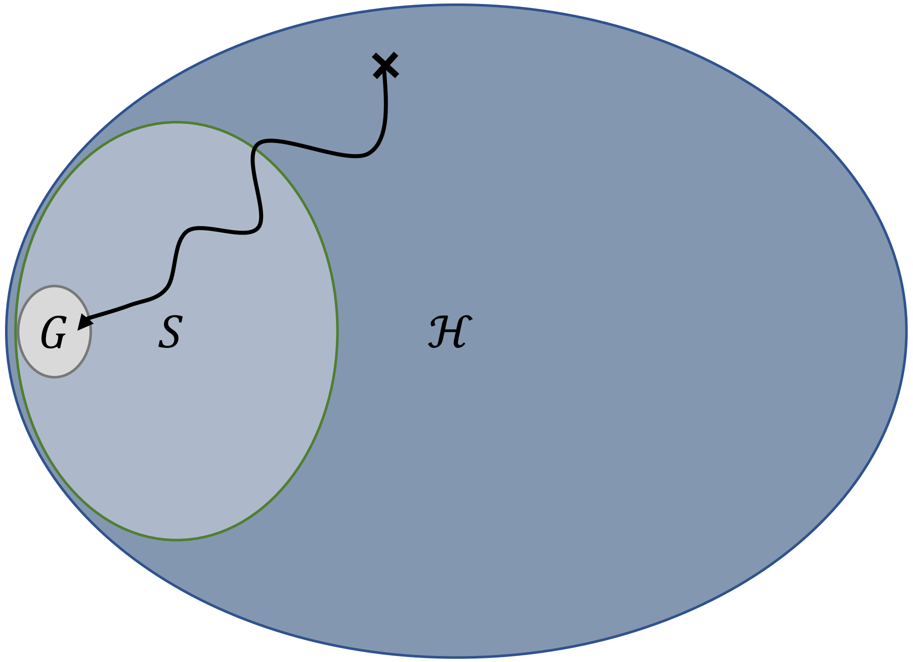
Barron, Gard, et al., arXiv: 2003.00171

Also performs well with noise included (taken from IBM processors)

Tailoring the ansatz to the Hamiltonian further

- The Symmetry Preserving Circuits enforce symmetries *if they are known*
- No other specific information about the Hamiltonian is input

- Can we find ansatze that are even more tailored to the Hamiltonian to be simulated while maintaining an economical structure of the ansatz?



Degree of problem tailoring

Key ideas in our algorithm (ADAPT-VQE):

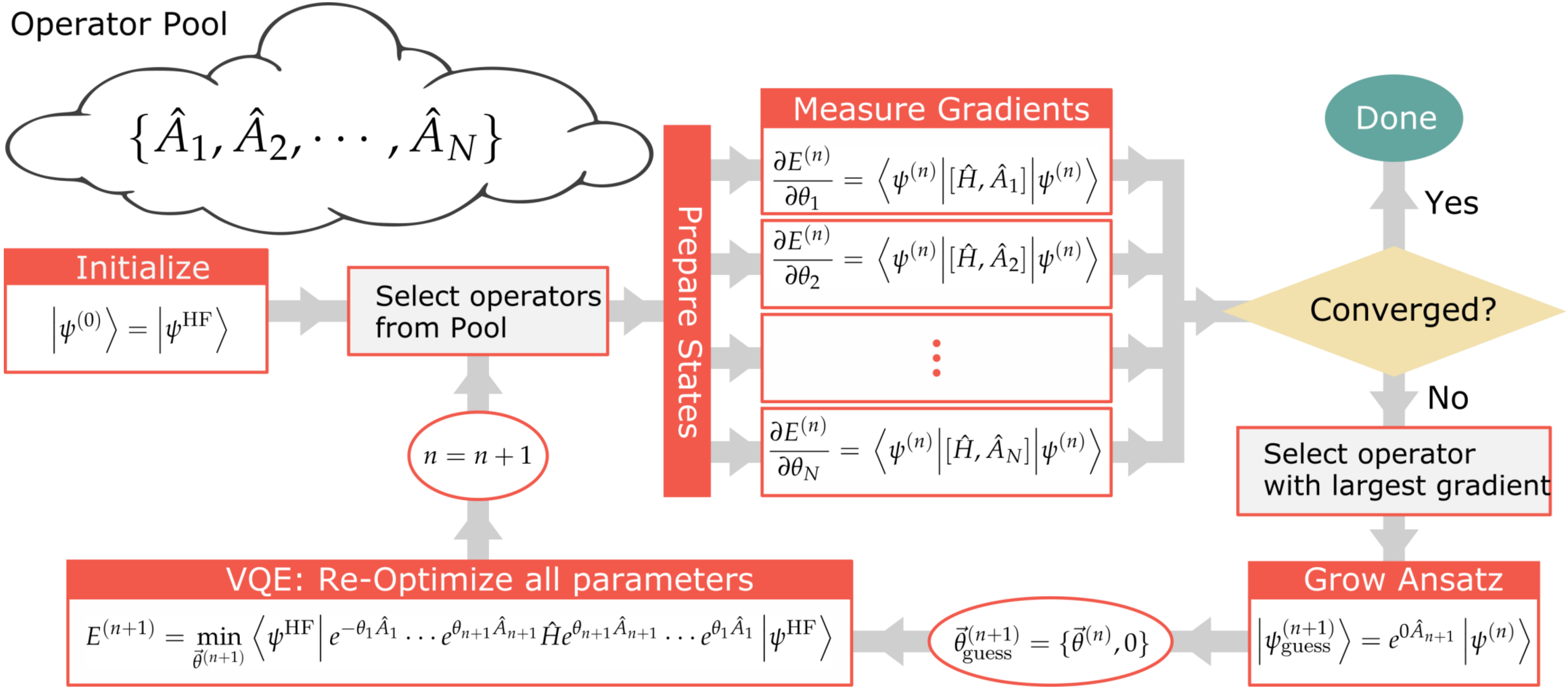
- ✓ Allow the simulated system to dictate its own ansatz
- ✓ Compact ansatz, grown one unitary operator at a time

ADAPT uses a pool of operators, A_m

Applies iteratively unitaries: $U_m = \exp(\theta_m A_m)$

ADAPT-VQE overview

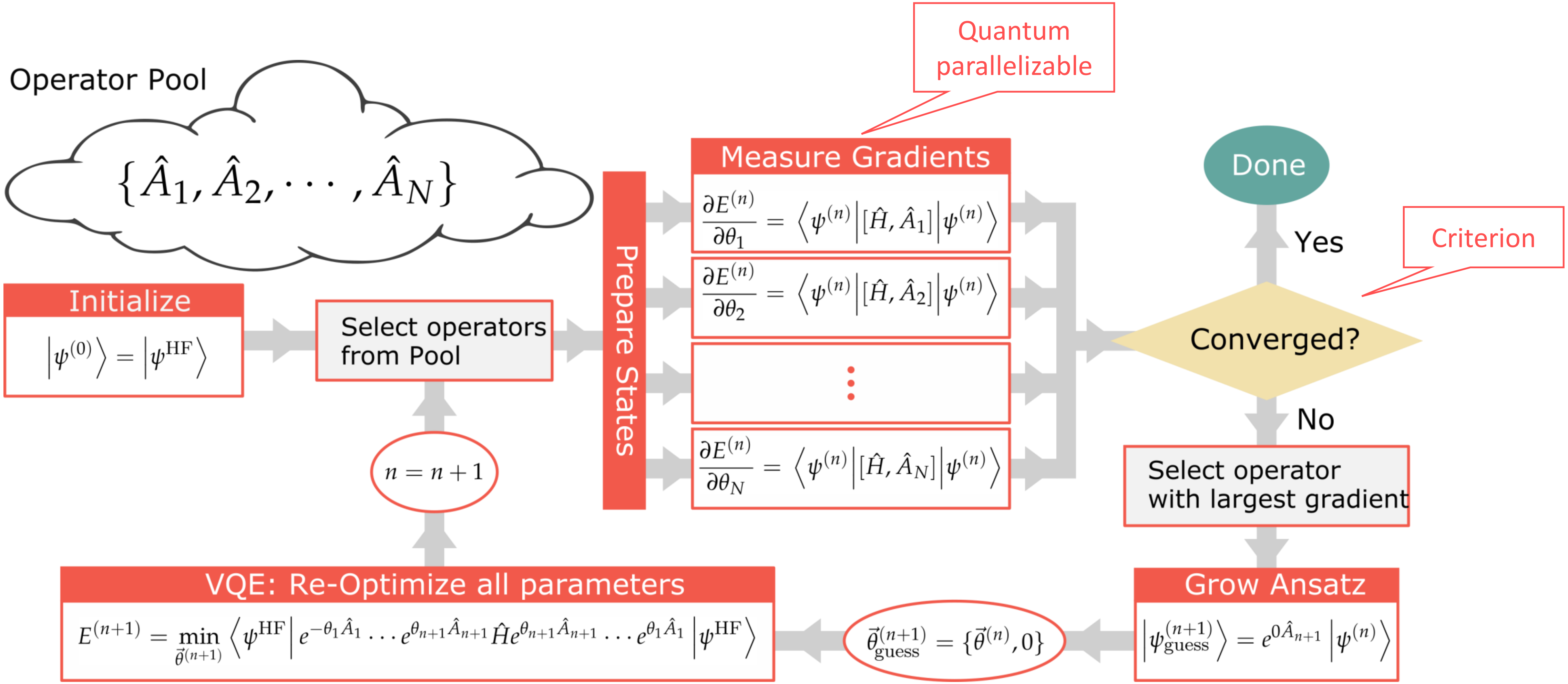
Adaptive **D**erivative **A**ssembled **P**roblem **T**ailored VQE



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

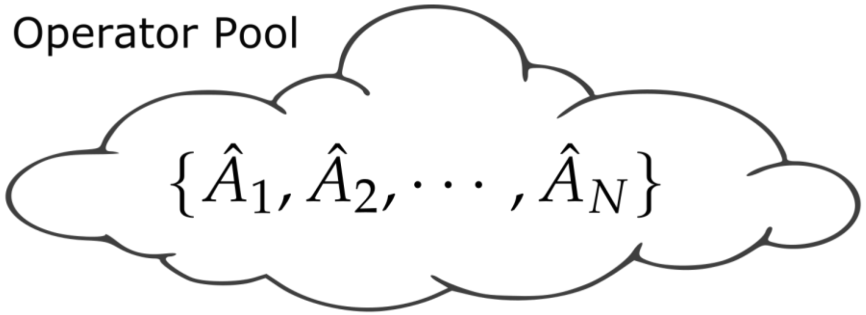
ADAPT-VQE overview

Adaptive **D**erivative **A**ssembled **P**roblem **T**ailored VQE



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

Operator pool a crucial component of ADAPT



How should it be chosen?
How do different pools perform?

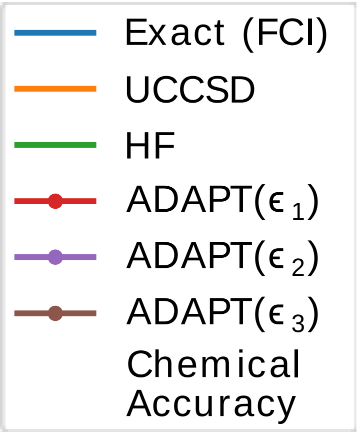
ADAPT with fermionic pool

$$\hat{A}_m = \left\{ \left(\hat{\tau}_p^q + \hat{\tau}_{\bar{p}}^{\bar{q}} \right), \left(\hat{\tau}_{pq}^{rs} + \hat{\tau}_{\bar{p}\bar{q}}^{\bar{r}\bar{s}} \right), \left(\hat{\tau}_{p\bar{q}}^{r\bar{s}} + \hat{\tau}_{\bar{p}q}^{\bar{r}s} \right) \right\}$$

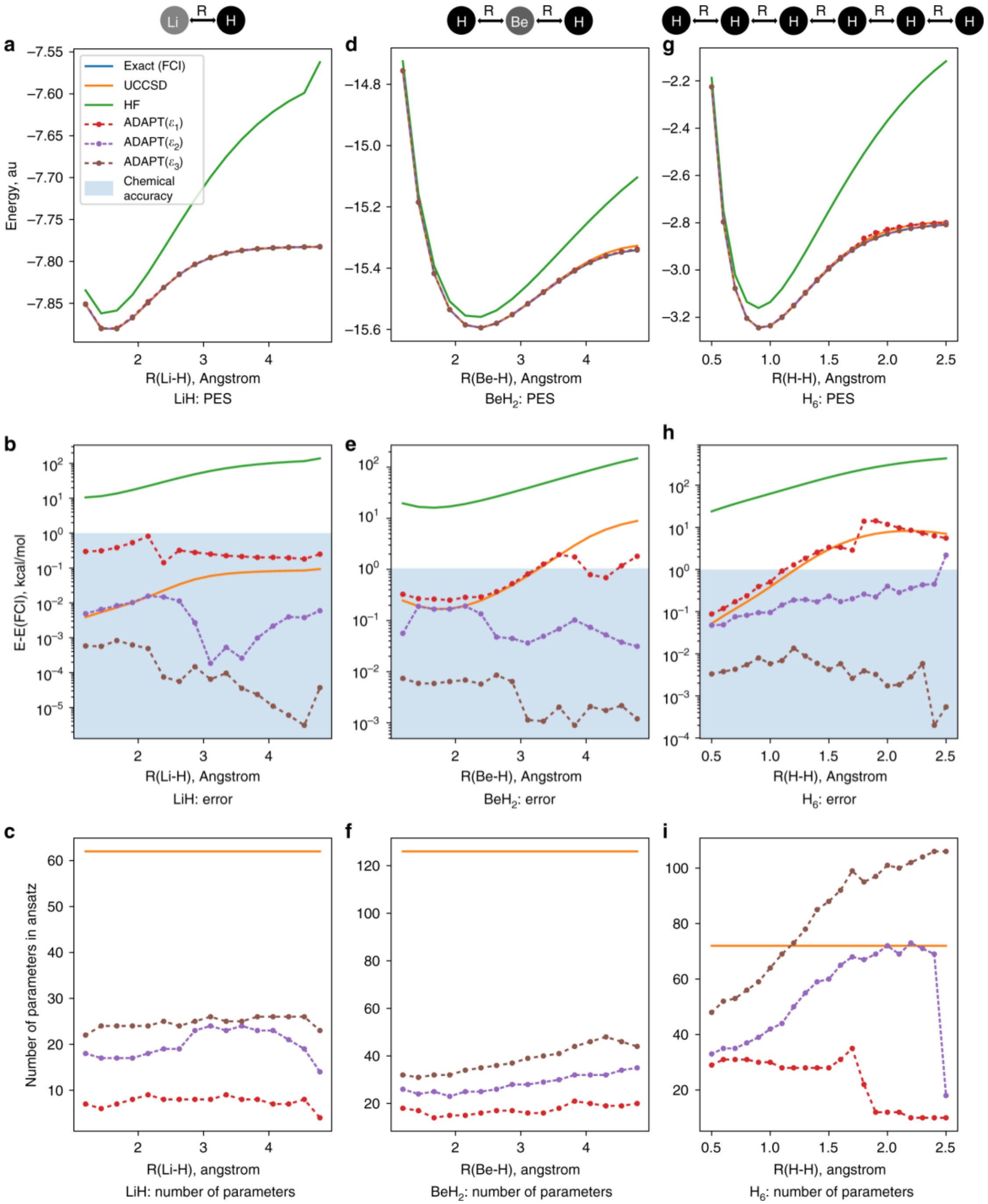
$$\hat{\tau}_p^q = a_q^\dagger a_p$$

$$\hat{\tau}_{pq}^{rs} = a_r^\dagger a_s^\dagger a_p a_q$$

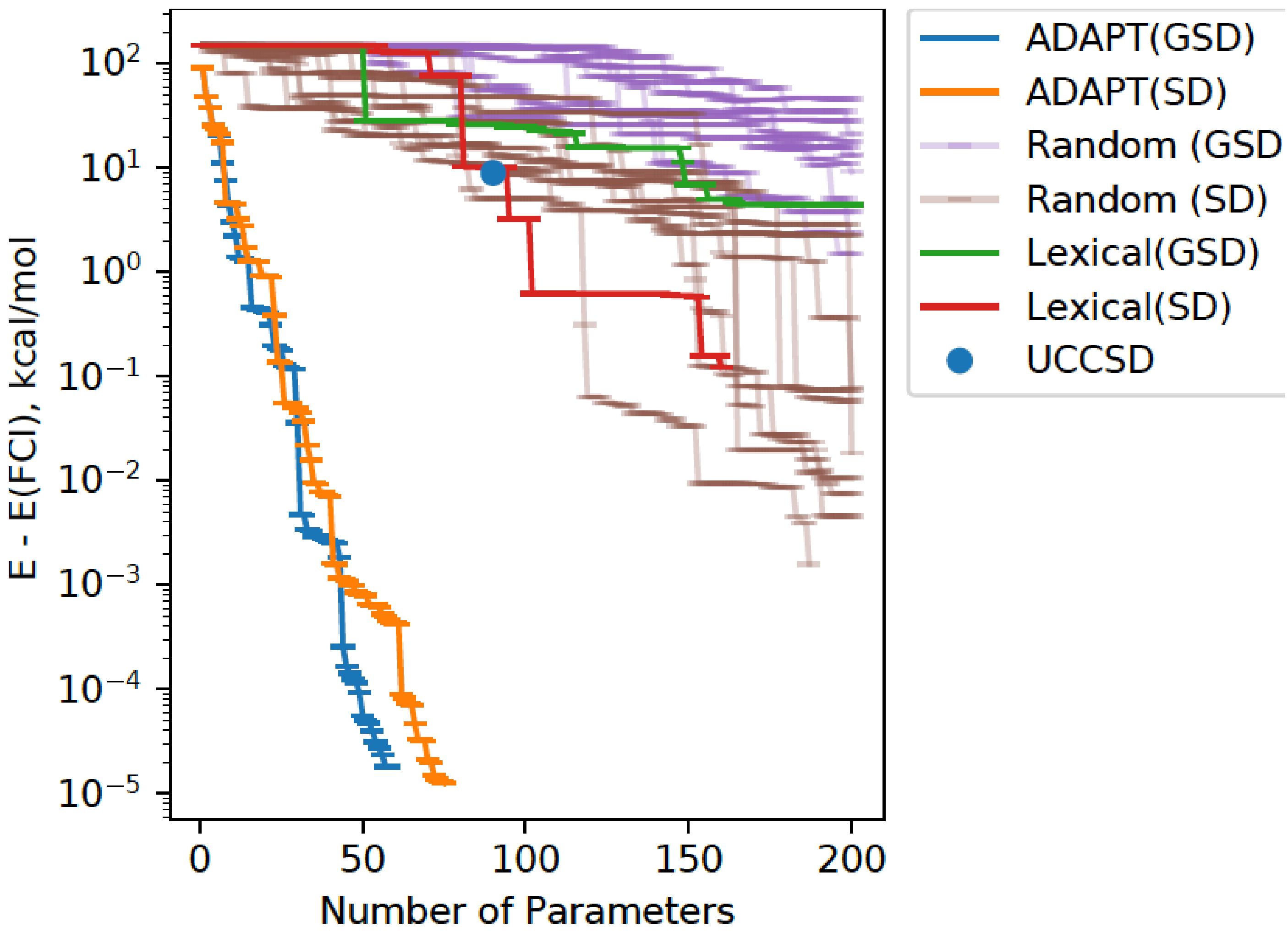
Results



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



Comparing ADAPT to other pseudo-Trotter orderings



BeH_2
bond distance 2.39 Å

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

ADAPT with hardware-efficient pool

- So far, we started with fermionic operators, then transformed them into qubit operators
- Each fermionic operator gives *O(n)* gates

- Alternative strategy for potentially shorter circuits: replace fermionic pool with *‘qubit’ pool* (‘Qubit-ADAPT-VQE’)
- Pool of operators can be dictated by hardware (e.g., nearest-neighbor coupling)

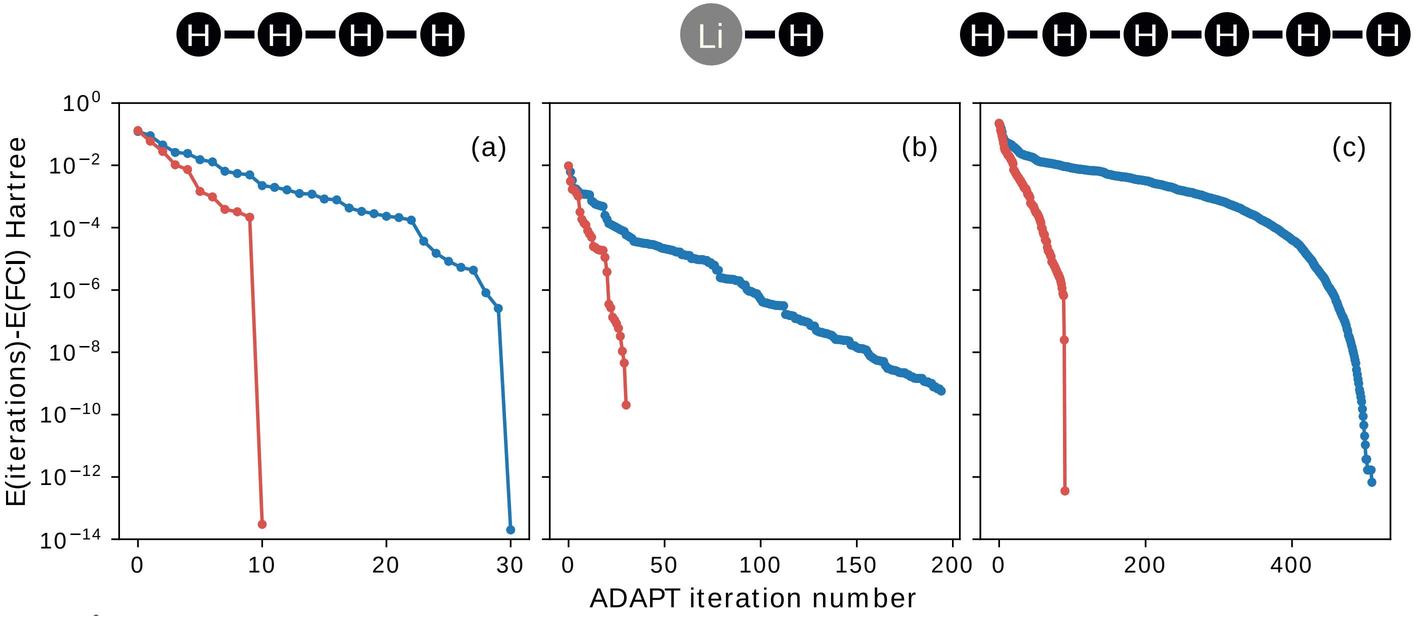
Qubit ADAPT-VQE: choice of pool

- Begin by taking operators of the form $e^{i\theta_j P_j}$ where P_j is a Pauli string

- Caveat: only imaginary operators in pool →antisymmetric pool—odd nr of Y operators (to respect time reversal)

- In the following, we choose P_j to be weight-4 Pauli strings

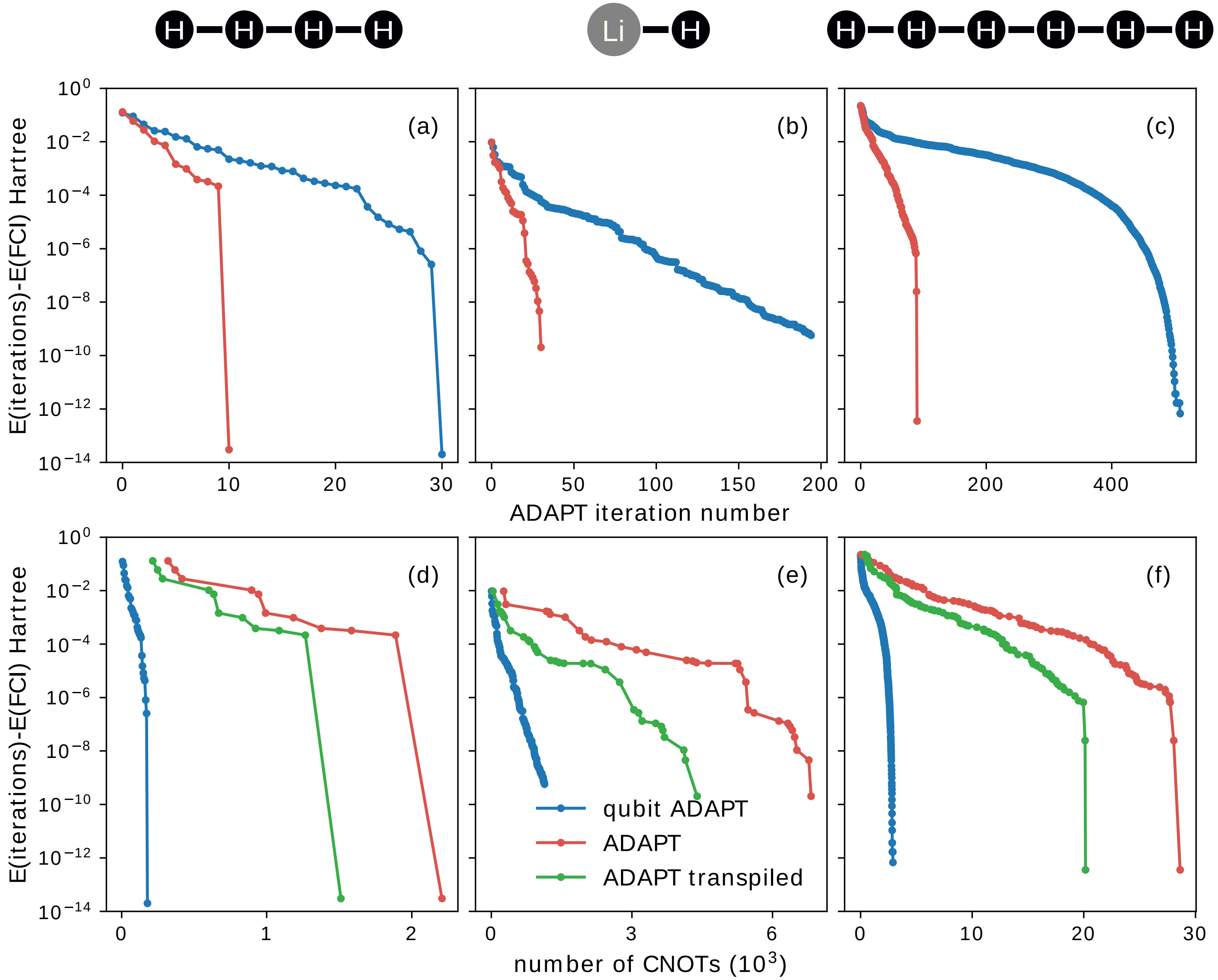
Qubit ADAPT-VQE—results



- 8, 12, 12 qubits respectively
- bond distances 1.5, 2, 1.5 Å

Tang, Shkolnikov, Barron, Grimsley, Mayhall, Barnes, Economou, arXiv:1911.10205

Qubit ADAPT-VQE—results

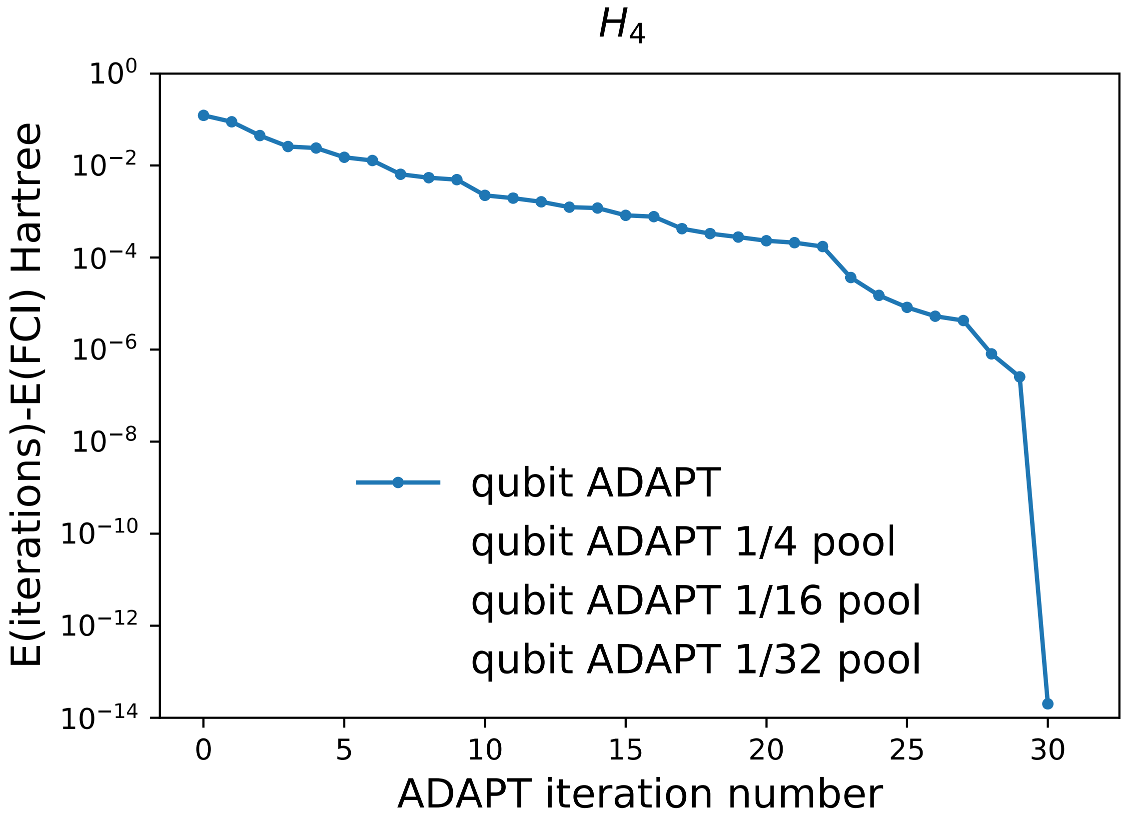


- 8, 12, 12 qubits respectively
- bond distances 1.5, 2, 1.5 Å

Tang, Shkolnikov, Barron, Grimsley, Mayhall, Barnes, Economou, arXiv:1911.10205

How big should the operator pool be?

- Our chosen pool is very large (for 8 qubits, >450 operators)
- Randomly reduce it and check convergence



Complete pools

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

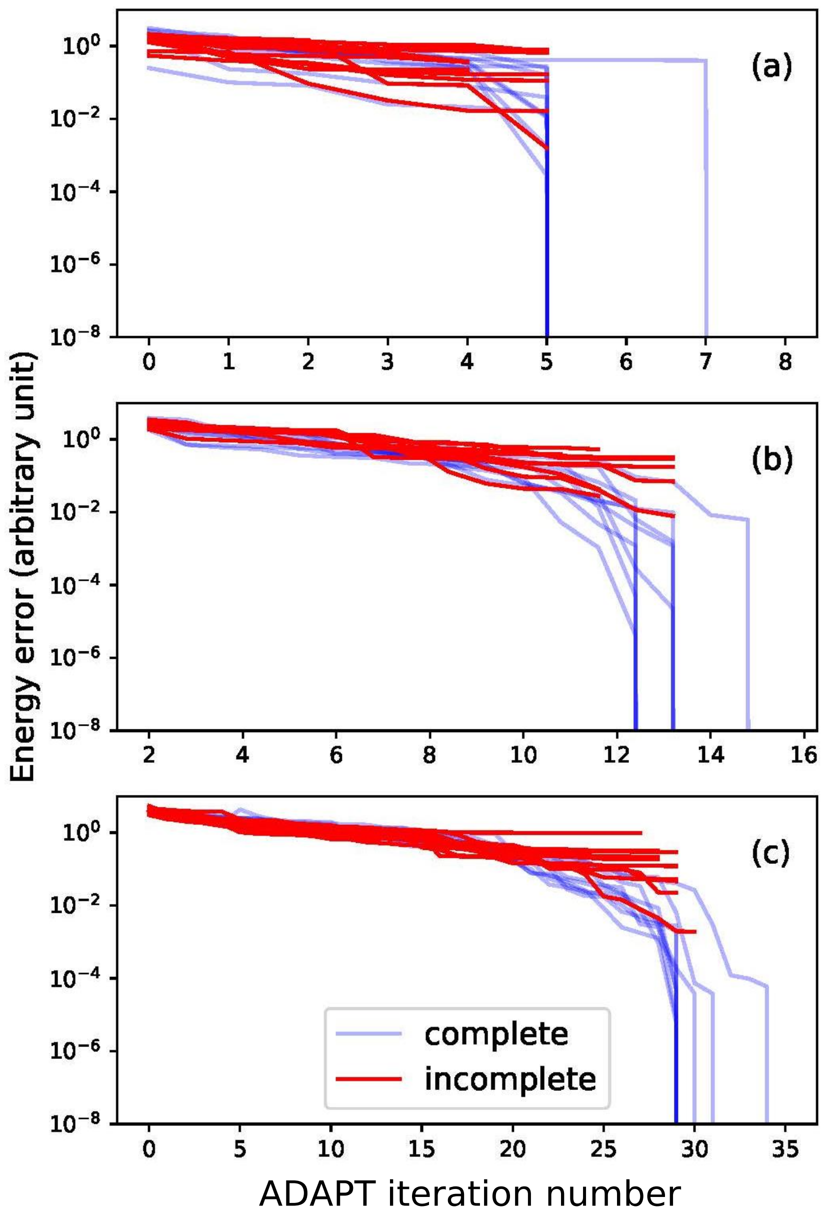
where $\{B_i\} = \{A_1, A_2, ..., [A_1, A_2], ..., [A_1, [A_2, A_3]], ...\}$

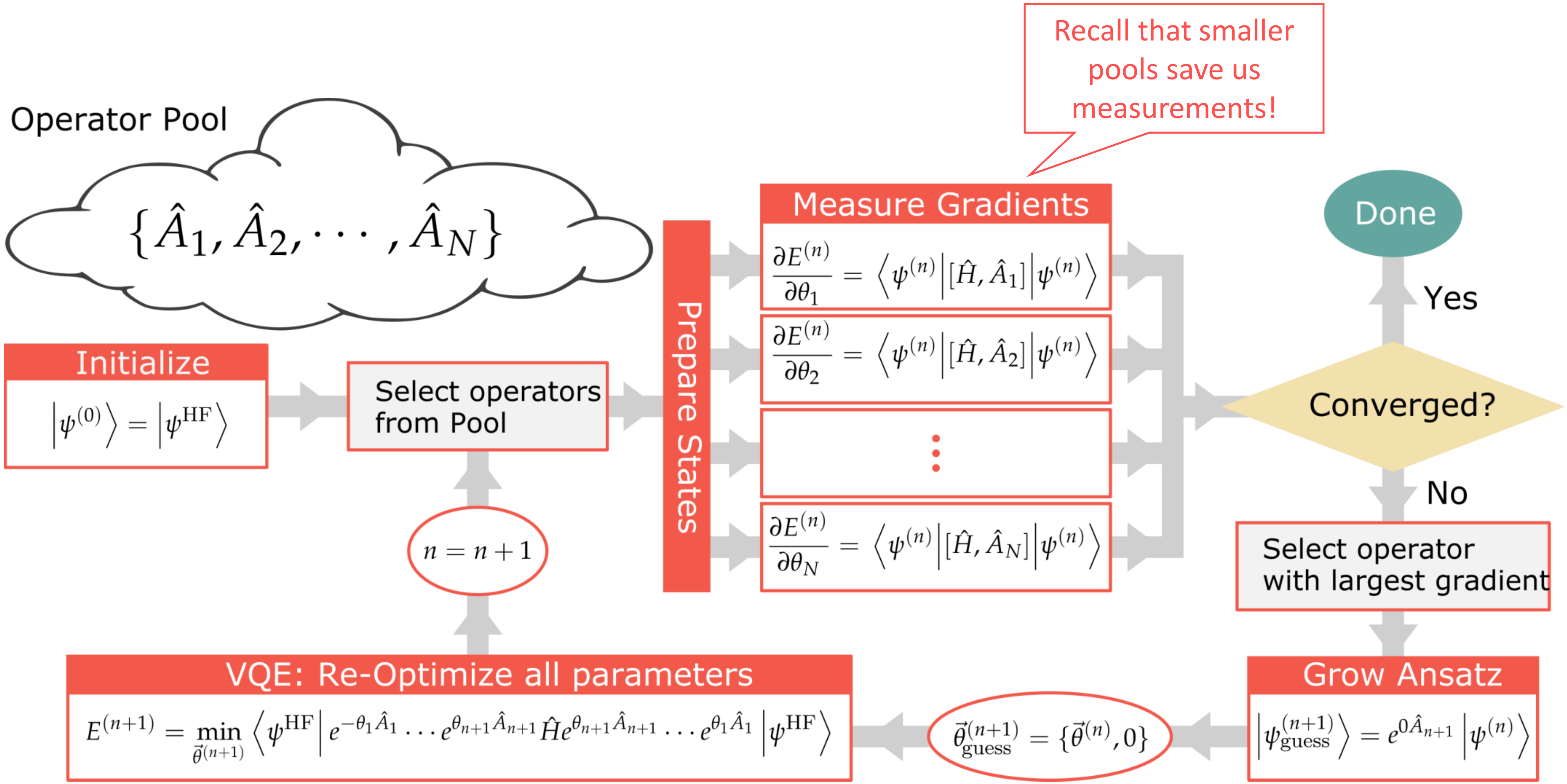
We have a complete pool and qubit-ADAPT is capable of converging to the exact ground state when states $B_i|\psi\rangle$ form a complete basis (where $|\psi\rangle$ is an arbitrary state)

Complete vs incomplete pool convergence

Test complete vs incomplete pools for random Hamiltonians for (a) 3 qubits, (b) 4 qubits, (c) 5 qubits

For pools that satisfy completeness criterion ADAPT always converges





Recall that smaller pools save us measurements!
Operator Pool
$\{\hat{A}_1, \hat{A}_2, \cdots, \hat{A}_N\}$
Initialize
$|\psi^{(0)}\rangle = |\psi^{\mathrm{HF}}\rangle$
Select operators from Pool
Prepare States
Measure Gradients
$\frac{\partial E^{(n)}}{\partial \theta_1} = \langle \psi^{(n)} | [\hat{H}, \hat{A}_1] | \psi^{(n)} \rangle$
$\frac{\partial E^{(n)}}{\partial \theta_2} = \langle \psi^{(n)} | [\hat{H}, \hat{A}_2] | \psi^{(n)} \rangle$
$\frac{\partial E^{(n)}}{\partial \theta_N} = \langle \psi^{(n)} | [\hat{H}, \hat{A}_N] | \psi^{(n)} \rangle$
Done
Yes
Converged?
No
Select operator with largest gradient
Grow Ansatz
$|\psi^{(n+1)}_{\mathrm{guess}}\rangle = e^{0\hat{A}_{n+1}} |\psi^{(n)}\rangle$
$\vec{\theta}^{(n+1)}_{\mathrm{guess}} = \{\vec{\theta}^{(n)}, 0\}$
VQE: Re-Optimize all parameters
$E^{(n+1)} = \min_{\vec{\theta}^{(n+1)}} \langle \psi^{\mathrm{HF}} | e^{-\theta_1 \hat{A}_1} \ldots e^{\theta_{n+1}\hat{A}_{n+1}} \hat{H} e^{\theta_{n+1}\hat{A}_{n+1}} \ldots e^{\theta_1 \hat{A}_1} | \psi^{\mathrm{HF}} \rangle$
$n = n+1$

Minimal complete pools

Minimal complete pool: smallest sized complete pool

The minimal size of complete pools is linear in the nr of qubits: *2n-2*

Examples of min complete pools

V pool:

$$V_1 = ZZ\ldots ZY, \quad V_2 = ZZ\ldots ZYI, \quad V_3 = ZZ\ldots ZYII, \quad \ldots, \quad V_{n-1} = ZYII\ldots I, \quad V_n = YII\ldots I,$$
$$V_{n+1} = ZZ\ldots ZIYI, \quad V_{n+2} = ZZ\ldots ZIYII, \quad \ldots, \quad V_{2n-3} = ZIYII\ldots I, \quad V_{2n-2} = IYII\ldots I$$

G pool:

$$G_1 = ZYII\ldots I, \quad G_2 = IZYII\ldots I, \quad G_3 = IIZYII\ldots I, \quad \ldots, \quad G_{n-2} = II\ldots IZYI, \quad G_{n-1} = II\ldots IZ$$
$$G_n = YII\ldots I, \quad G_{n+1} = IYII\ldots I, \quad G_{n+2} = IIYII\ldots I, \quad \ldots, \quad G_{2n-3} = II\ldots IYII, \quad G_{2n-2} = II\ldots I$$

E.g., for 3 qubits $V_1 = ZZY, \quad V_2 = ZYI, \quad V_3 = YII, \quad V_4 = IYI$

Tang, Shkolnikov, Barron, Grimsley, Mayhall, Barnes, Economou, arXiv:1911.10205

Outline

- Fermionic problems
  - Symmetry enforcing circuits
  - ADAPT-VQE algorithm
- Optimization (many body Ising) problems
  - ADAPT QAOA

Quantum Approximate Optimization Algorithm (QAOA)

- Optimization problems can be encoded in Ising Hamiltonians C
- Solution encoded in ground state
- E.g., for weighted Max-Cut problem $C = -\frac{1}{2}\sum_{i,j} w_{i,j}(I - Z_i Z_j)$

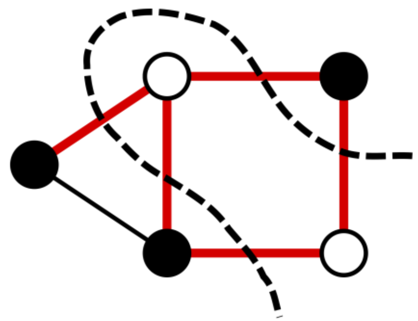


image from Wikipedia

QAOA algorithm (inspired by adiabatic theorem)

- Start from initial state $|+\rangle^{\otimes n}$, eigenstate of $B = \sum_{i=1}^{n} X_i$ — "mixer"
- QAOA ansatz:

$$|\psi_p(\vec{\gamma}, \vec{\beta})\rangle = e^{-i\beta_p B} e^{-i\gamma_p C} \ldots e^{-i\beta_1 B} e^{-i\gamma_1 C} |+\rangle^{\otimes n}$$

- Perform VQE to minimize $\langle\psi_p(\vec{\gamma}, \vec{\beta})|C|\psi_p(\vec{\gamma}, \vec{\beta})\rangle$

Farhi, Goldstone, Gutmann, MIT-CTP/4610, arXiv:1411.4028

ADAPT-QAOA

- Our approach:
  - Preserve QAOA ansatz structure $|\psi_p(\vec{\beta},\vec{\gamma})\rangle = e^{-i\beta_p A_p} e^{-i\gamma_p C} \ldots e^{-i\beta_1 A_1} e^{-i\gamma_1 C} |+\rangle^{\otimes n}$
  - Define mixer pools
  - Use ADAPT strategy to determine the mixers

- Operator pool $\{A_i\}$:
  - Single-qubit gate operators:

$$\left\{X_i, Y_i, \sum_{i=1}^{n} X_i, \sum_{i=1}^{n} Y_i\right\}$$

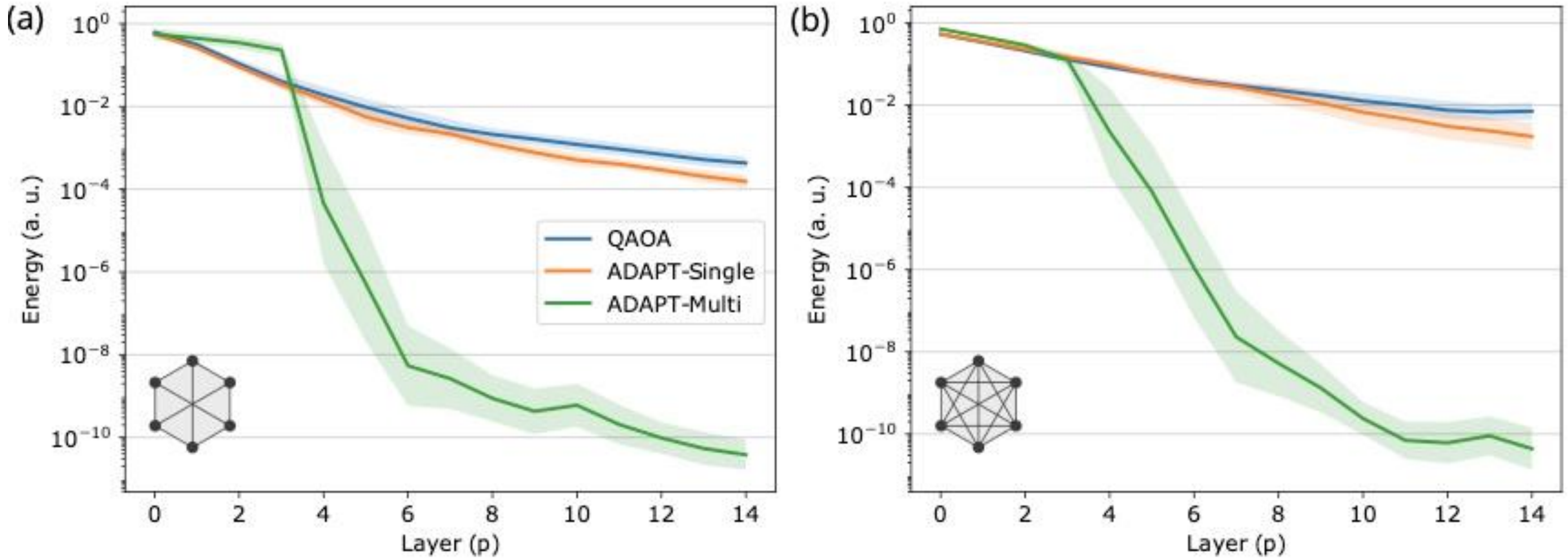
  - Single-qubit & entangling operators:

$$\left\{X_i, Y_i, \sum_{i=1}^{n} X_i, \sum_{i=1}^{n} Y_i, Z_iY_j, X_iY_j, Z_iZ_j, X_iX_j, X_iZ_j, Y_iY_j\right\}$$

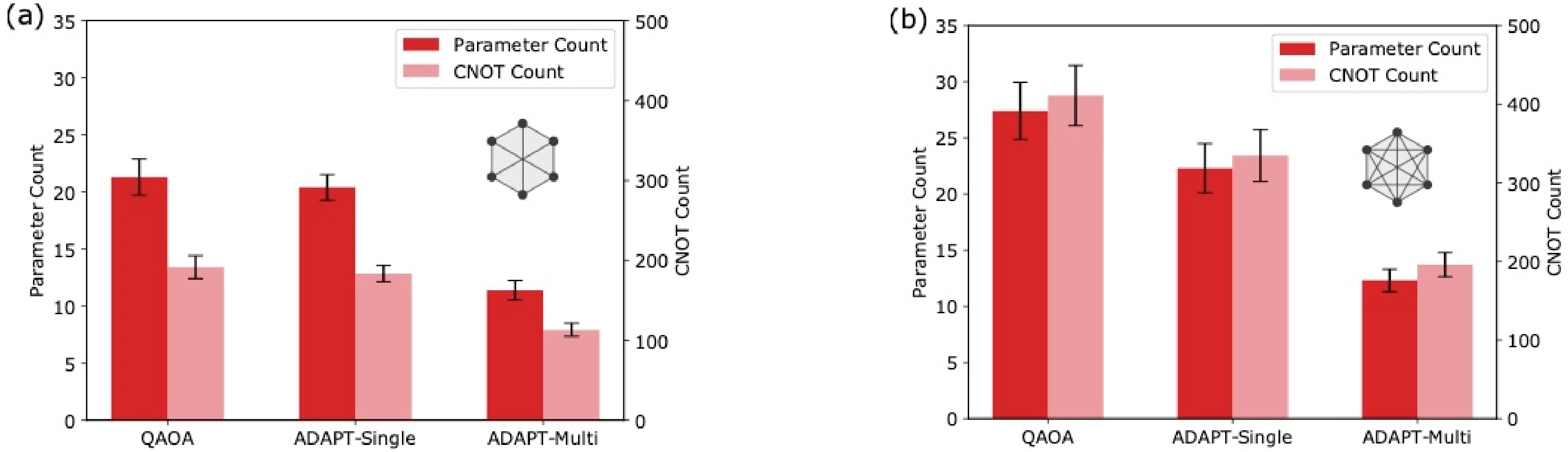
ADAPT-QAOA-results

Max-Cut problem for graphs with random edge weights

$$C = -\frac{1}{2}\sum_{i,j} w_{i,j}(I - Z_i Z_j)$$



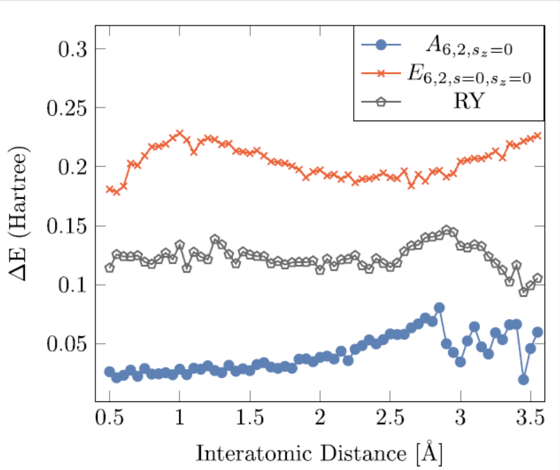
For fixed accuracy (10^{-3}):



Zhu, Tang, Barron et al., arxiv:2005.10258

Summary

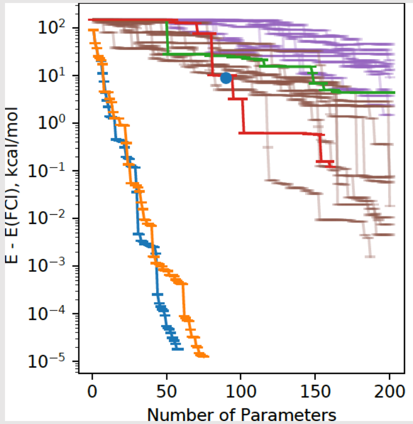
Symmetry preserving circuits



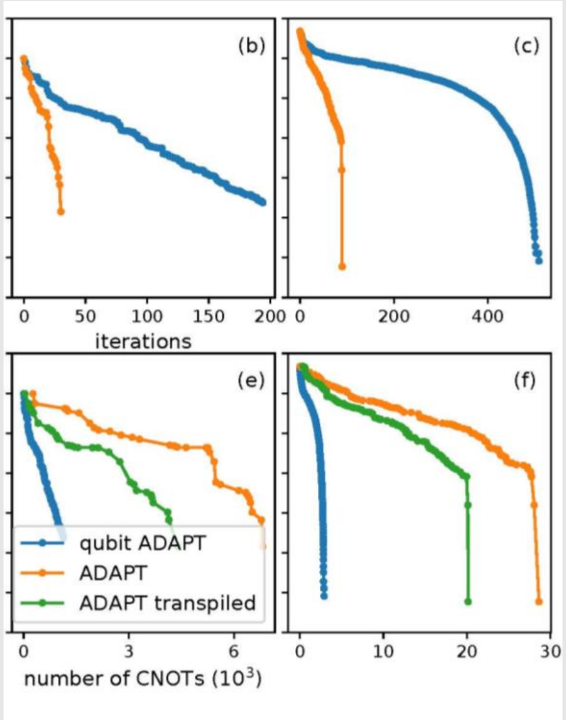
Gard, Zhu, Barron, et al, *npj Quantum Inf* **6**, 10 (2020)

Barron, Gard, Altman, et al., arXiv: 2003.00171

ADAPT-VQE

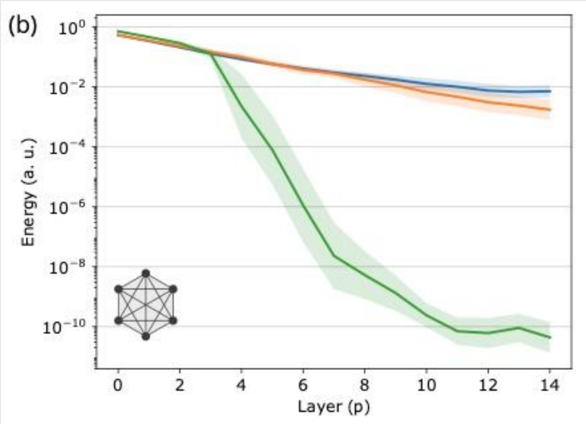


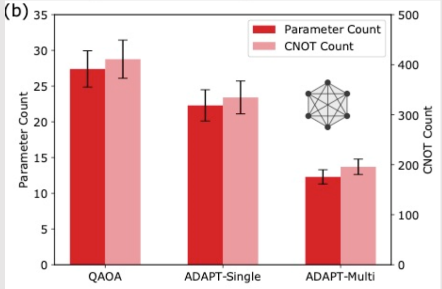
Grimsley, Economou, Barnes, Mayhall, Nature Commun. **10**, 3007 (2019)



Tang, Shkolnikov, Barron, et al, arXiv:1911.10205

ADAPT-QAOA





Zhu, Tang, Barron, et al., arXiv: 2005.10258