

Minimizing State Preparations for VQE

Pranav Gokhale¹, Olivia Angiuli², Yongshan Ding¹, Kaiwen Gui¹,
Teague Tomesh³, Margaret Martonosi³, Frederic T. Chong¹

¹University of Chicago

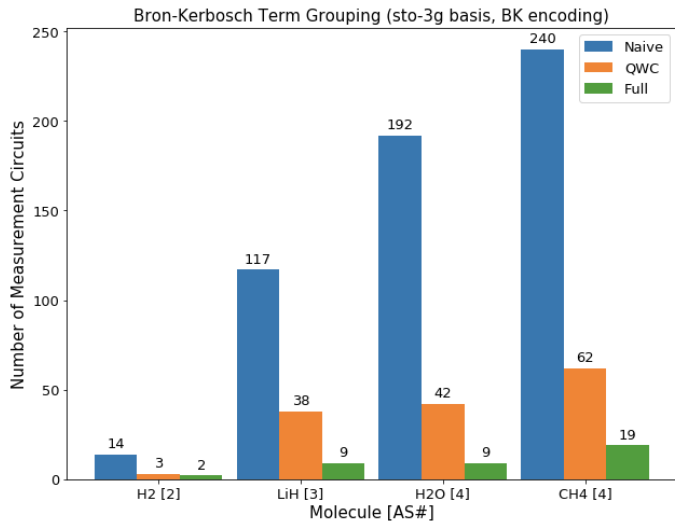
²UC Berkeley

³Princeton University

EPIQC: Enabling Practical-scale Quantum Computation

QRE 2019, June 22

Results



Background: Ground State Estimation

- Important problem in chemistry: what is ground state energy of a molecule?

Background: Ground State Estimation

- Important problem in chemistry: what is ground state energy of a molecule?
- Determine reaction rates, molecular geometry, etc.

Background: Ground State Estimation

- Important problem in chemistry: what is ground state energy of a molecule?
- Determine reaction rates, molecular geometry, etc.
- Solving is equivalent to finding the min eigenvalue (min energy) of the Hamiltonian matrix, H : $H|j_{\min}\rangle = E_{\min}|j_{\min}\rangle$.

Background: Ground State Estimation

- Important problem in chemistry: what is ground state energy of a molecule?
- Determine reaction rates, molecular geometry, etc.
- Solving is equivalent to finding the min eigenvalue (min energy) of the Hamiltonian matrix, H : $H|j_{\min}\rangle = E_{\min}|j_{\min}\rangle$.
- Classically, diagonalize exponentially-sized matrix.

Background: Ground State Estimation

- Important problem in chemistry: what is ground state energy of a molecule?
- Determine reaction rates, molecular geometry, etc.
- Solving is equivalent to finding the min eigenvalue (min energy) of the Hamiltonian matrix, H : $H|j_{\min}\rangle = E_{\min}|j_{\min}\rangle$.
- Classically, diagonalize exponentially-sized matrix.
- Quantum Phase Estimation algorithm showed how to solve in poly-time.

Variational Quantum Eigensolver (VQE)

Invented in 2014, suitable for near-term/NISQ quantum.

Variational Quantum Eigensolver (VQE)

Invented in 2014, suitable for near-term/NISQ quantum.

Variational Method

$E_j = \langle \psi_j | H | \psi_j \rangle$ is an overestimate of the lowest eigenvalue (energy).

Variational Quantum Eigensolver (VQE)

Invented in 2014, suitable for near-term/NISQ quantum.

Variational Method

$\langle H \rangle_{|\psi\rangle}$ is an overestimate of the lowest eigenvalue (energy).

Algorithm:

- 1 **Guess.** Prepare an ansatz state, $|\psi\rangle$.

Variational Quantum Eigensolver (VQE)

Invented in 2014, suitable for near-term/NISQ quantum.

Variational Method

$\langle j | H | j \rangle$ is an overestimate of the lowest eigenvalue (energy).

Algorithm:

- 1 **Guess.** Prepare an ansatz state, $|j\rangle$.
- 2 **Check.** Measure $\langle j | H | j \rangle$.

Variational Quantum Eigensolver (VQE)

Invented in 2014, suitable for near-term/NISQ quantum.

Variational Method

$\langle j | H | j \rangle$ is an overestimate of the lowest eigenvalue (energy).

Algorithm:

- 1 **Guess.** Prepare an ansatz state, $|j\rangle$.
- 2 **Check.** Measure $\langle j | H | j \rangle$.
- 3 **Repeat.** Optimizer on classical co-processor guides next $|j\rangle$, with aim of minimizing $\langle j | H | j \rangle$.

Variational Quantum Eigensolver (VQE)

Invented in 2014, suitable for near-term/NISQ quantum

Variational Method

$\langle j | H | j \rangle$ is an overestimate of the lowest eigenvalue (energy).

Algorithm:

- 1 **Guess.** Prepare an ansatz state, $|j\rangle$.
- 2 **Check.** Measure $\langle j | H | j \rangle$. **Our resource estimation & minimization focus.**
- 3 **Repeat.** Optimizer on classical co-processor guides next $|j\rangle$, with aim of minimizing $\langle j | H | j \rangle$.

Checking $\langle H \rangle$

- $\langle H \rangle$ cannot be measured directly on a quantum computer.

Checking $\langle H \rangle$

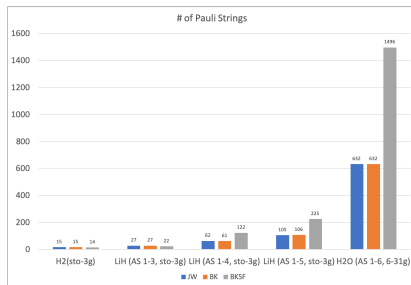
- $\langle H \rangle$ cannot be measured directly on a quantum computer.
- Instead, decompose H into lincomb of Pauli Strings $fI; X; Y; Zg^N$.

Checking $\langle H \rangle$

- $\langle H \rangle$ cannot be measured directly on a quantum computer.
- Instead, decompose H into lincomb of Pauli Strings $fI; X; Y; Zg^N$.
- E.g. $\langle H \rangle = \langle ZII \rangle + \langle IZI \rangle + \langle ZZI \rangle - \langle XXI \rangle - \langle YYI \rangle$.

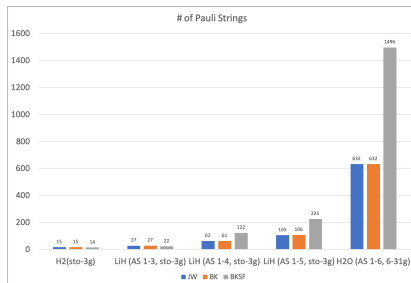
Checking hHi

- hHi cannot be measured directly on a quantum computer.
- Instead, decompose H into lincomb of Pauli Strings $fI; X; Y; Zg^N$.
- E.g. $hHi = hZII + hIZI + hZZI \quad hXXI \quad hYYI$.
- The number of Pauli Strings scales as N^4



Checking hHi

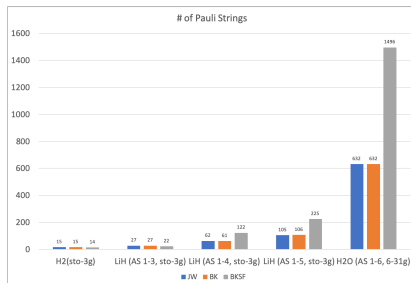
- hHi cannot be measured directly on a quantum computer.
- Instead, decompose H into lincomb of Pauli Strings $fI; X; Y; Zg^N$.
- E.g. $hHi = hZII + hIZI + hZZI \quad hXXI \quad hYYI$.
- The number of Pauli Strings scales as N^4



- Original VQE formulation, measure each term separately. Each measurement requires separate state preparation.

Checking $\langle H \rangle$

- $\langle H \rangle$ cannot be measured directly on a quantum computer.
- Instead, decompose H into lincomb of Pauli Strings $fI; X; Y; Zg^N$.
- E.g. $\langle H \rangle = \langle HZI \rangle + \langle HIZ \rangle + \langle HZZ \rangle + \langle HXX \rangle + \langle HYY \rangle$.
- The number of Pauli Strings scales as N^4



- Original VQE formulation, measure each term separately. Each measurement requires separate state preparation.
- But, **commuting terms can be measured simultaneously.**

Our Contributions

- Analysis of commutativity

Our Contributions

- Analysis of commutativity
 - | When do terms commute? What types of commutation relationships?

Our Contributions

- Analysis of commutativity
 - | When do terms commute? What types of commutation relationships?
 - | How to simultaneously measure?

Our Contributions

- Analysis of commutativity
 - | When do terms commute? What types of commutation relationships?
 - | How to simultaneously measure?
 - | How to group terms into large commuting families?

Our Contributions

- Analysis of commutativity
 - | When do terms commute? What types of commutation relationships?
 - | How to simultaneously measure?
 - | How to group terms into large commuting families?
- Impact on variance and study of covariances

Our Contributions

- Analysis of commutativity
 - | When do terms commute? What types of commutation relationships?
 - | How to simultaneously measure?
 - | How to group terms into large commuting families?
- Impact on variance and study of covariances
- Benchmarking & resource estimation for representative molecules

Pauli Commutativity Relations

Pauli Commutativity Relations

Commutator Notation

$$[A; B] = AB - BA \quad \begin{cases} = 0 & \text{if } A \text{ and } B \text{ commute} \\ \neq 0 & \text{if } A \text{ and } B \text{ do not commute} \end{cases}$$

Pauli Commutativity Relations

Commutator Notation

$$[A; B] = AB - BA \quad \begin{cases} = 0 & \text{if } A \text{ and } B \text{ commute} \\ \neq 0 & \text{if } A \text{ and } B \text{ do not commute} \end{cases}$$

For Pauli matrices, $P = \{I; X; Y; Z\}$:

- I commutes with everything. $[I;] = 0$

Pauli Commutativity Relations

Commutator Notation

$$[A; B] = AB - BA \quad \begin{cases} = 0 & \text{if } A \text{ and } B \text{ commute} \\ \neq 0 & \text{if } A \text{ and } B \text{ do not commute} \end{cases}$$

For Pauli matrices, $P = \{I; X; Y; Z\}$:

- I commutes with everything. $[I;] = 0$
- Everything commutes with itself. $[;] = 0$

Pauli Commutativity Relations

Commutator Notation

$$[A; B] = AB - BA \quad \begin{cases} = 0 & \text{if } A \text{ and } B \text{ commute} \\ \neq 0 & \text{if } A \text{ and } B \text{ do not commute} \end{cases}$$

For Pauli matrices, $P = \{I; X; Y; Z\}$:

- I commutes with everything. $[I;] = 0$
- Everything commutes with itself. $[;] = 0$
- Other Pauli commutators follow from cyclic multiplication property.:

Pauli Commutativity Relations

Commutator Notation

$$[A; B] = AB - BA \quad \begin{cases} = 0 & \text{if } A \text{ and } B \text{ commute} \\ \neq 0 & \text{if } A \text{ and } B \text{ do not commute} \end{cases}$$

For Pauli matrices, $P = \{I; X; Y; Z\}$:

- I commutes with everything. $[I;] = 0$
- Everything commutes with itself. $[;] = 0$
- Other Pauli commutators follow from cyclic multiplication property.:
 - | $XY = iZ$
 - | $YZ = iX$
 - | $ZX = iY$

Pauli Commutativity Relations

Commutator Notation

$$[A; B] = AB - BA \quad \begin{cases} = 0 & \text{if } A \text{ and } B \text{ commute} \\ \neq 0 & \text{if } A \text{ and } B \text{ do not commute} \end{cases}$$

For Pauli matrices, $P = \{I; X; Y; Z\}$:

- I commutes with everything. $[I;] = 0$
- Everything commutes with itself. $[;] = 0$
- Other Pauli commutators follow from cyclic multiplication property.:
 - | $XY = iZ = -YX$
 - | $YZ = iX = -ZY$
 - | $ZX = iY = -XZ$

Qubit-Wise Commutativity (QWC)

Consider two N -qubit Pauli Strings, A and B , $2 \text{ fl}; X; Y; Zg^N$.

Qubit-Wise Commutativity (QWC)

Consider two N -qubit Pauli Strings, A and B , $2 \text{ fl}; X; Y; Z g^N$.

Definition: QWC

A and B QWC iff $[A_i; B_i] = 0 \ \forall i$

Qubit-Wise Commutativity (QWC)

Consider two N -qubit Pauli Strings, A and B , $2 \text{ fl}; X; Y; Z g^N$.

Definition: QWC

A and B QWC iff $[A_i; B_i] = 0 \ \forall i$

Examples

QWC: $\begin{array}{cccccc} X & Y & Z & Y & I & Y \\ I & Y & Z & I & I & I \end{array}$

Qubit-Wise Commutativity (QWC)

Consider two N -qubit Pauli Strings, A and B , $2 \leq N$.

Definition: QWC

A and B QWC iff $[A_i; B_j] = 0 \ \forall i$

Examples

QWC: $\begin{matrix} X & Y & Z & Y & I & Y \\ I & Y & Z & I & I & I \end{matrix}$

Not QWC: $\begin{matrix} X & Y & Z & Y & I & Y \\ I & Y & Z & X & I & I \end{matrix}$

Simultaneous QWC Measurement

Consider terms matching $(I \text{ or } Z)(I \text{ or } X)(I \text{ or } Z)$:

Simultaneous QWC Measurement

Consider terms matching (I or Z)(I or X)(I or Z):

I I I
I I Z
I X I
I X Z
Z I I
Z I Z
Z X I
Z X Z

Simultaneous QWC Measurement

Consider terms matching (I or Z)(I or X)(I or Z):

I I I
I I Z
I X I
I X Z
Z I I
Z I Z
Z X I
Z X Z

These terms are a QWC family.

Simultaneous QWC Measurement

Consider terms matching (I or Z)(I or X)(I or Z):

I I I
I I Z
I X I
I X Z
Z I I
Z I Z
Z X I
Z X Z

These terms are a QWC family.

Measure in:

- Z basis for 1st qubit

Simultaneous QWC Measurement

Consider terms matching $(I \text{ or } Z)(I \text{ or } X)(I \text{ or } Z)$:

I I I
I I Z
I X I
I X Z
Z I I
Z I Z
Z X I
Z X Z

These terms are a QWC family.

Measure in:

- Z basis for 1st qubit
- X basis for 2nd qubit

Simultaneous QWC Measurement

Consider terms matching (I or Z)(I or X)(I or Z):

I I I
I I Z
I X I
I X Z
Z I I
Z I Z
Z X I
Z X Z

These terms are a QWC family.

Measure in:

- Z basis for 1st qubit
- X basis for 2nd qubit
- Z basis for 3rd qubit

Simultaneous QWC Measurement

Consider terms matching $(I \text{ or } Z)(I \text{ or } X)(I \text{ or } Z)$:

I I I
I I Z
I X I
I X Z
Z I I
Z I Z
Z X I
Z X Z

These terms are a QWC family.

Measure in:

- Z basis for 1st qubit
- X basis for 2nd qubit
- Z basis for 3rd qubit

In general: QWC simultaneous measurements requires $O(N)$ single qubit gates (depth = 1).

Full Commutativity

Full Commutativity

Is there life beyond QWC? **Yes!**

Full Commutativity

Is there life beyond QWC? **Yes!**

Example

X X
Y Y are a commuting family! But not QWC.
Z Z

Full Commutativity

Is there life beyond QWC? **Yes!**

Example

X X
 Y Y are a commuting family! But not QWC.
 Z Z

General Commutativity of Pauli Strings

If an **even** number of indices don't commute, then A and B commute.

Full Commutativity

Is there life beyond QWC? **Yes!**

Example

X X
 Y Y are a commuting family! But not QWC.
 Z Z

General Commutativity of Pauli Strings

If an **even** number of indices don't commute, then A and B commute.

Proof: All Pauli strings are either commuting or anti-commuting, so we know that either $AB = BA$ or $AB = -BA$. Each commuting index multiplies by $+1$, each non-commuting index multiplies by -1 . Need even number of non-commuting indices to have a total $+1$.

Simultaneous Full Commutativity Measurement

Steps

Simultaneous Full Commutativity Measurement

Steps

- 1 Find stabilizer generators of commuting term group:

$$fXX;YY;ZZg \quad hXX;ZZi$$

Simultaneous Full Commutativity Measurement

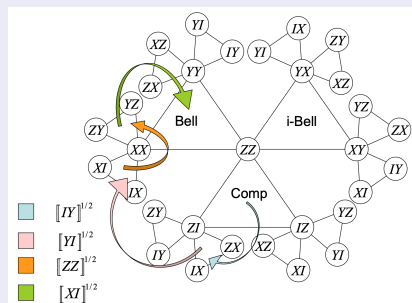
Steps

- 1 Find stabilizer generators of commuting term group:
 $fXX;YY;ZZg ! hXX;ZZi$
- 2 Transform from computational basis generators to new generators
 $hZI;IZi ! hXX;ZZi$ via rules of Clifford Algebra

Simultaneous Full Commutativity Measurement

Steps

- 1 Find stabilizer generators of commuting term group:
 $fXX; YY; ZZg \quad hXX; ZZi$
- 2 Transform from computational basis generators to new generators
 $hZI; IZi \quad hXX; ZZi$ via rules of Clifford Algebra



[Devoret Notes]

Simultaneous Full Commutativity Measurement

Steps

- 1 Find stabilizer generators of commuting term group:
 $fXX;YY;ZZg ! hXX;ZZi$
- 2 Transform from computational basis generators to new generators
 $hZI;IZi ! hXX;ZZi$ via rules of Clifford Algebra

Simultaneous Full Commutativity Measurement

Steps

- 1 Find stabilizer generators of commuting term group:
 $fXX;YY;ZZg ! hXX;ZZi$
- 2 Transform from computational basis generators to new generators
 $hZI;IZi ! hXX;ZZi$ via rules of Clifford Algebra
- 3 Apply $O(N^2)$ gates from resulting circuit (our software gives explicit circuit decomposition):

Simultaneous Full Commutativity Measurement

Steps

- 1 Find stabilizer generators of commuting term group:
 $fXX;YY;ZZg ! hXX;ZZi$
- 2 Transform from computational basis generators to new generators
 $hZI;IZi ! hXX;ZZi$ via rules of Clifford Algebra
- 3 Apply $O(N^2)$ gates from resulting circuit (our software gives explicit circuit decomposition):
- 4 Result: Bell Basis Measurement for this example

Simultaneous Full Commutativity Measurement

Steps

- 1 Find stabilizer generators of commuting term group:
 $fXX; YY; ZZg ! hXX; ZZi$
- 2 Transform from computational basis generators to new generators
 $hZI; IZi ! hXX; ZZi$ via rules of Clifford Algebra
- 3 Apply $O(N^2)$ gates from resulting circuit (our software gives explicit circuit decomposition):
- 4 Result: Bell Basis Measurement for this example

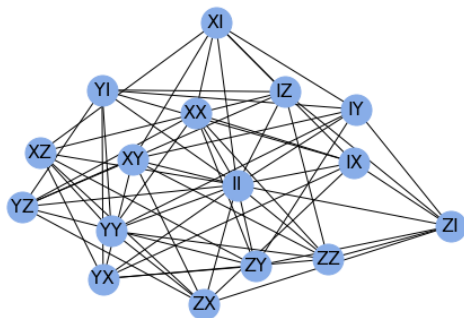
$O(N^2)$ gates is fine, because UCCSD ansatz prep is $O(N^3)$ or $O(N^4)$.

Grouping Commuting Terms

Commutativity is not transitive—complicates grouping.

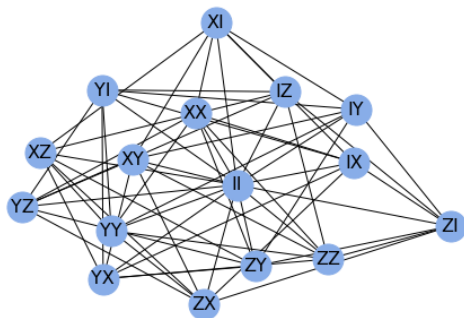
Grouping Commuting Terms

Commutativity is not transitive—complicates grouping.



Grouping Commuting Terms

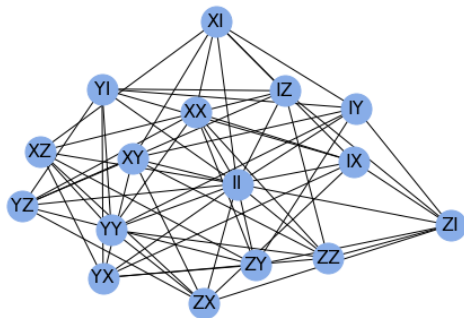
Commutativity is not transitive—complicates grouping.



We seek MIN-CLIQUE-COVER.

Grouping Commuting Terms

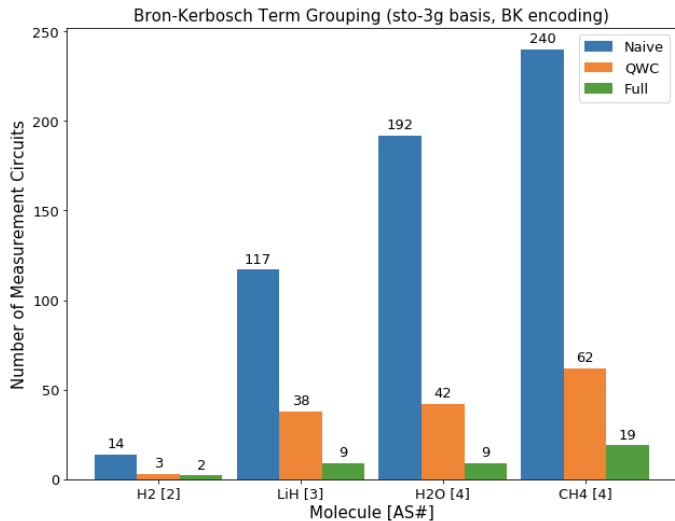
Commutativity is not transitive—complicates grouping.



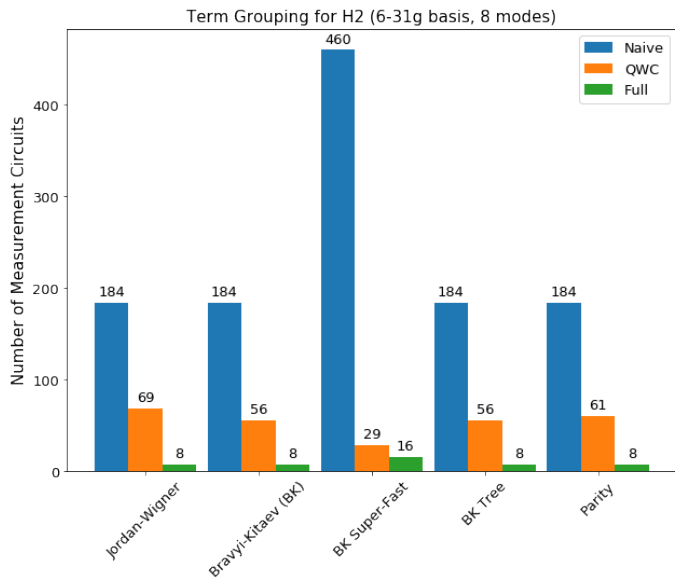
We seek MIN-CLIQUE-COVER.

Problem is NP-Hard, but use heuristics (Boppana-Halldórsson). Note that our problem is also NP-HARD by reduction from MIN-CLIQUE-COVER.

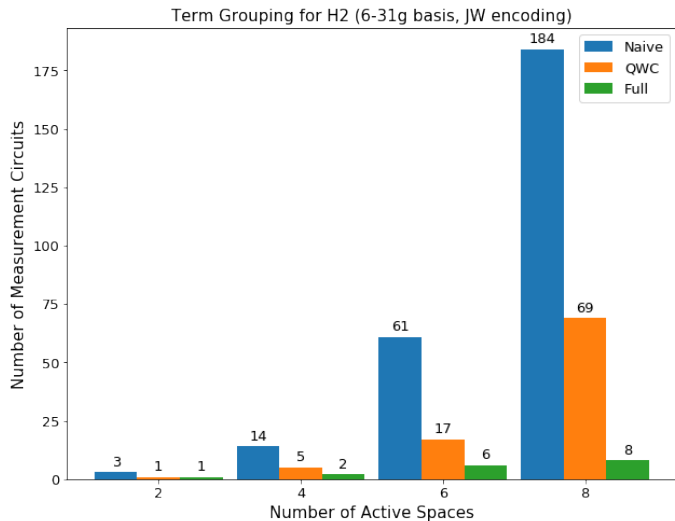
Results: Across Molecules



Results: Across Encodings



Results: Across Active Spaces



Qubit Tapering

- Create QWCommutativity by transforming Hamiltonian [Bravyi 2017].


Qubit Tapering

- Create QWCommutativity by transforming Hamiltonian [Bravyi 2017].

σ_1^z	σ_2^z	σ_3^z	σ_4^z
$\sigma_1^x \sigma_2^z$	$\sigma_1^z \sigma_3^z$	$\sigma_1^z \sigma_4^z$	
$\sigma_2^z \sigma_3^z$	$\sigma_2^z \sigma_4^z$	$\sigma_3^z \sigma_4^z$	
$\sigma_1^y \sigma_2^y \sigma_3^x \sigma_4^x$	$\sigma_1^x \sigma_2^y \sigma_3^y \sigma_4^x$	$\sigma_1^y \sigma_2^x \sigma_3^x \sigma_4^y$	$\sigma_1^x \sigma_2^x \sigma_3^y \sigma_4^y$

$$U_1 = \frac{1}{\sqrt{2}} (\sigma_2^x + \sigma_1^z \sigma_2^z), \quad U_2 = \frac{1}{\sqrt{2}} (\sigma_3^x + \sigma_1^z \sigma_3^z)$$

and $U_3 = \frac{1}{\sqrt{2}} (\sigma_4^x + \sigma_1^z \sigma_4^z).$



σ_1^z	$\sigma_1^z \sigma_2^x$	$\sigma_1^z \sigma_3^x$	$\sigma_1^z \sigma_4^x$
σ_2^x	σ_3^x	σ_4^x	
$\sigma_2^x \sigma_3^x$	$\sigma_2^x \sigma_4^x$	$\sigma_3^x \sigma_4^x$	
$\sigma_1^x \sigma_3^x \sigma_4^x$	$\sigma_1^x \sigma_4^x$	$\sigma_1^x \sigma_2^x \sigma_3^x$	$\sigma_1^x \sigma_2^x$

- In example, three qubits are tapered out of H2 Hamiltonian

Qubit Tapering Results

Hamiltonian	# of Qubits	# Tapered
H2 (small # active spaces)	4	3
H2 (large # active spaces)	8	2
H2O	8	2

Measurement Statistics¹

The optimal group depends on the ansatz state.

¹Example from [McClean et al 2015]

Measurement Statistics¹

The optimal group depends on the ansatz state.

Example

Consider $H = XX + YY + ZZ + IZ + ZI$, and state $|j\rangle = |j01\rangle$.

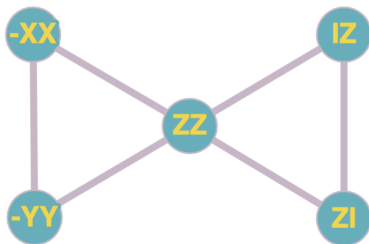
¹Example from [McClean et al 2015]

Measurement Statistics¹

The optimal group depends on the ansatz state.

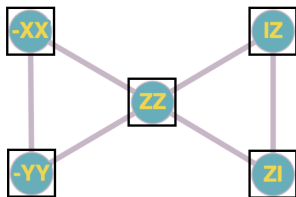
Example

Consider $H = -XX - YY + ZZ + IZ + ZI$, and state $|j\rangle = |j01i\rangle$.

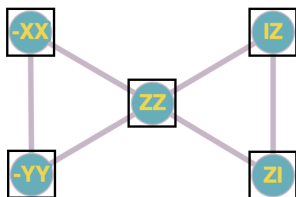


¹Example from [McClean et al 2015]

$k = 5$ Groups



$k = 5$ Groups



$$\begin{aligned} E(\# \text{ state preps}) &= \\ k \text{ Var}(XX) + \text{Var}(YY) + \text{Var}(ZZ) + \text{Var}(ZI) + \text{Var}(IZ) &= 2 \\ &= 5 \cdot 1 + 1 + 0 + 0 + 0 = 2 \\ &= \boxed{10} = 2^2 \end{aligned}$$

$k = 3$ Groups

k = 3 Groups

$$\begin{aligned}
 E(\# \text{ state preps}) &= \sum_i^2 = \\
 &= \sum_h^k \text{Var}(XX) + \text{Var}(f_{YY}; ZZg) + \text{Var}(f_{ZI}; IZg) = \\
 &= \sum_h^k \text{Var}(XX) + \text{Var}(YY) + \text{Var}(ZZ) + 2 \text{Cov}(YY; ZZ) \\
 &\quad + \text{Var}(ZI) + \text{Var}(IZ) + 2 \text{Cov}(IZ; ZI) = \\
 &= 3 \cdot 1 + 1 + 0 + 0 + 0 + 0 + 0 = 6 = 2^2
 \end{aligned}$$

$k = 2$ Groups

k = 2 Groups

$$\begin{aligned}
 & E(\# \text{ state preps}) = \\
 & \sum_h k \text{Var}(f_{XX}; YY; ZZ) + \sum_i \text{Var}(f_{ZI}; IZ) = 2 \\
 & = \sum_h k \text{Var}(XX) + \text{Var}(YY) + \text{Var}(ZZ) + \\
 & 2\text{Cov}(XX; YY) + 2\text{Cov}(XX; ZZ) + 2\text{Cov}(YY; ZZ) \\
 & \quad \text{Var}(ZI) + \text{Var}(IZ) + 2\text{Cov}(IZ; ZI) = 2 \\
 & = 2 \cdot 1 + 1 + 0 + 2 \cdot 1 + 0 + 0 + 0 + 0 + 0 = 2 = \boxed{8 = 2}
 \end{aligned}$$

Optimal Grouping

Bigger k means more state preparations per iteration, but more independent looks at data.

Optimal Grouping

Bigger k means more state preparations per iteration, but more independent looks at data.

However:

Optimal Initial Grouping

With no prior on the ansatz state ψ_i , optimal strategy is to pick the largest groups (smallest k).

Optimal Grouping

Bigger k means more state preparations per iteration, but more independent looks at data.

However:

Optimal Initial Grouping

With no prior on the ansatz state ψ_i , optimal strategy is to pick the largest groups (smallest k).

$E[\text{Cov}(A; B)] = 0$ for two commuting Pauli terms {proof via stabilizer formalism}. Thus, on average, all covariances are 0, and state preparations are reduced by lowering k .

Optimal Grouping

Bigger k means more state preparations per iteration, but more independent looks at data.

However:

Optimal Initial Grouping

With no prior on the ansatz state ψ_i , optimal strategy is to pick the largest groups (smallest k).

$E[\text{Cov}(A; B)] = 0$ for two commuting Pauli terms {proof via stabilizer formalism. Thus, on average, all covariances are 0, and state preparations are reduced by lowering k .

Work in progress: adaptively adjust groups after initial grouping.

Next Steps

Explicitly test on real hardware

Next Steps

Explicitly test on real hardware
Benchmark qubit tapering

Next Steps

Explicitly test on real hardware

Benchmark qubit tapering

Develop adaptive regrouping strategy in light of covariances

Next Steps

Explicitly test on real hardware

Benchmark qubit tapering

Develop adaptive regrouping strategy in light of covariances

More connections to stabilizer formalism?

Next Steps

- Explicitly test on real hardware
- Benchmark qubit tapering
- Develop adaptive regrouping strategy in light of covariances
- More connections to stabilizer formalism?

Thanks!