Learning Beyond Stabilizer States

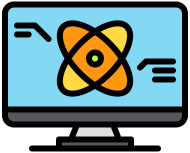
Based on: arXiv:2305.13409

Sabee Grewal, Vishnu Iyer, William Kretschmer,
Daniel Liang

Rice University Department of Computer Science

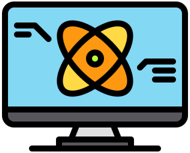
October 18th, 2023

Quantum State Tomography

 $\rightarrow |\psi\rangle$

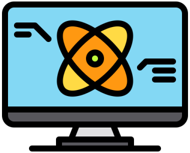
Input: Black-box access to copies of $|\psi\rangle$.

Quantum State Tomography

 $\rightarrow |\psi\rangle$

Input: Black-box access to copies of $|\psi\rangle$.
Output: Approximation $|\psi'\rangle \approx |\psi\rangle$ as a classical description.

Quantum State Tomography



$\rightarrow |\psi\rangle$

Input: Black-box access to copies of $|\psi\rangle$.
Output: Approximation $|\psi'\rangle \approx |\psi\rangle$ as a classical description.

Provably hard!

Efficient Quantum State Tomography

How do we get around this exponential barrier?

Efficient Quantum State Tomography

How do we get around this exponential barrier?
The same way we do in classical learning!

Solution

- Move the goal post:
  - PAC learning [Aar07]
  - Shadow tomography [Aar18, HKP20]
  - Distinguishing/property testing [GNW21, GIKL23d]

Efficient Quantum State Tomography

How do we get around this exponential barrier?
The same way we do in classical learning!

Solution

- Move the goal post:
  - PAC learning [Aar07]
  - Shadow tomography [Aar18, HKP20]
  - Distinguishing/property testing [GNW21, GIKL23d]
- Restrict the states:
  - Free-fermion states [AG23]
  - Low-Degree Phase states [ABDY23]
  - **Stabilizer States** [Mon17]

Stabilizer States & Clifford Unitaries

Definition

A *Clifford unitary* is any unitary generated by H, S, and CNOT.

$$H := \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad S := \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} \qquad \mathrm{CNOT} := \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

Stabilizer States & Clifford Unitaries

Definition

A *Clifford unitary* is any unitary generated by H, S, and CNOT.

$$H := \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} \qquad S := \begin{bmatrix}1 & 0\\ 0 & i\end{bmatrix} \qquad \mathrm{CNOT} := \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{bmatrix}$$

Definition

A *stabilizer state* is a state generated by a Clifford unitary on $|0^n\rangle$.

Stabilizer States & Clifford Unitaries

Definition

A *Clifford unitary* is any unitary generated by H, S, and CNOT.

$$H := \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} \qquad S := \begin{bmatrix}1 & 0\\ 0 & i\end{bmatrix} \qquad \mathrm{CNOT} := \begin{bmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{bmatrix}$$

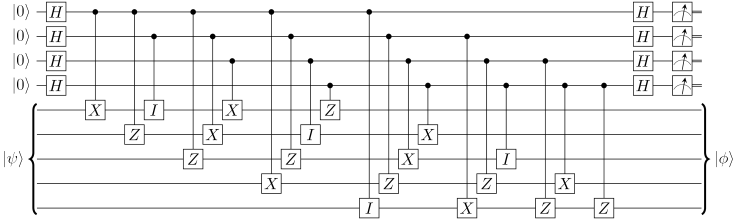
Definition

A *stabilizer state* is a state generated by a Clifford unitary on $|0^n\rangle$.

Not a universal gate set!

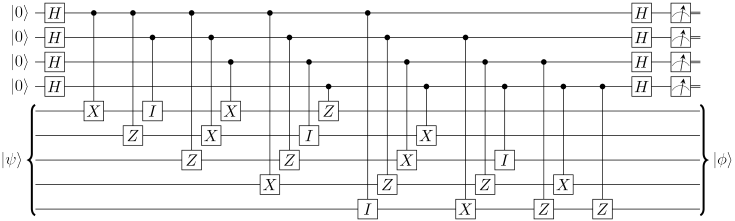
Applications

Error-correcting Codes

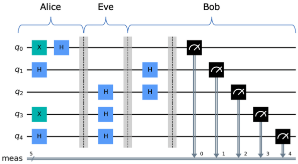


Applications

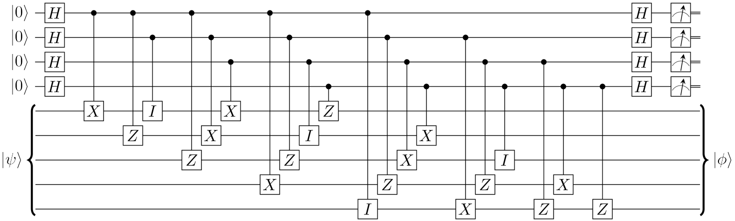
Error-correcting Codes



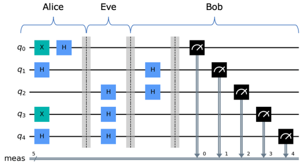
Quantum Key Distribution



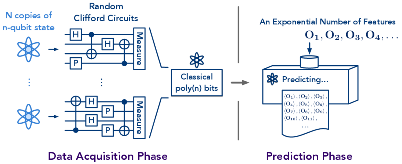
Error-correcting Codes



Quantum Key Distribution

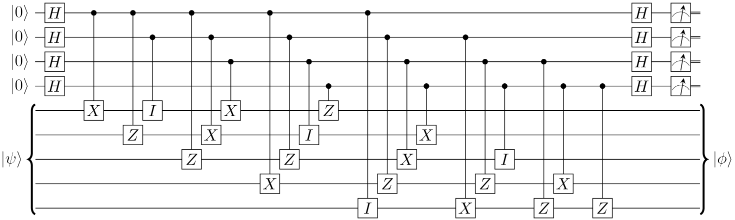


Learning Algorithms

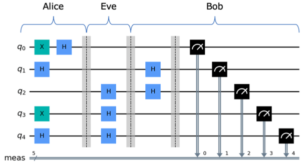


Applications

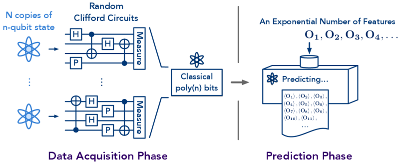
Error-correcting Codes



Quantum Key Distribution



Learning Algorithms



And more! Unitary Designs, Quantum Money, Classical Simulation, ...

Algebraic Structure of Stabilizer States

$$I := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad X := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad Z := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\mathcal{P}_n := \{I, X, Y, Z\}^{\otimes n}$$

Algebraic Structure of Stabilizer States

$$I := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad X := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad Z := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\mathcal{P}_n := \{I, X, Y, Z\}^{\otimes n}$$

Definition

$$\mathsf{Stab}(|\psi\rangle) := \left\{ W \in \mathcal{P}_n : |\langle\psi|W|\psi\rangle|^2 = 1 \right\}.$$

Algebraic Structure of Stabilizer States

$$I := \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad X := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad Y := \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad Z := \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\mathcal{P}_n := \{I, X, Y, Z\}^{\otimes n}$$

Definition

$$\text{Stab}(|\psi\rangle) := \left\{ W \in \mathcal{P}_n : |\langle\psi|W|\psi\rangle|^2 = 1 \right\}.$$

Fact: $|\text{Stab}(|\psi\rangle)| = 2^n$ if and only if $|\psi\rangle$ is a stabilizer state.

Learning Stabilizer States: A Warmup

Lemma ([AG04])

Given Stab$(|\varphi\rangle)$*, there exists a Clifford circuit* C *such that*

$$C|\varphi\rangle = |x\rangle$$

for some $x \in \{0,1\}^n$*.*

Moreover, C can be computed in time $O(n^2)$.

Learning Stabilizer States: A Warmup (cont.)

Lemma ([Mon17])

Given copies of a stabilizer state $|\varphi\rangle$, there exists a measurement to efficiently sample from the uniform distribution over Stab$(|\varphi\rangle)$.

Algorithm: Sample $O(n)$ times and output the group generated by the samples.

Learning Beyond Stabilizer States

Question

Various generalizations of stabilizer states:

- Low-stabilizer-rank states
- Low-degree phase states
- **Clifford + T states**

Can we learn any of them efficiently?

Clifford + T

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}$$

Clifford unitaries are not universal for computation, but Clifford + T is!

Clifford + T

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}$$

Clifford unitaries are not universal for computation, but Clifford + T is!

T gates take us further and further from the nice algebraic properties of stabilizer states:

- Classical simulation algorithms run in time $\mathrm{poly}(n)\exp(k)$.

Our Work

Theorem ([GIKL23a, LOH23, HG23])

Can learn any state produced by k T gates in time $\mathrm{poly}(n)\exp(k)$.

1

Our Work

Theorem ([GIKL23a, LOH23, HG23])

Can learn any state produced by k T gates in time $\mathrm{poly}(n)\exp(k)$.

Other results:

- Pseudorandom state distinguisher[1]: [GIKL23c].

[1] See Simons Talk for more information!

Our Work

Theorem ([GIKL23a, LOH23, HG23])

Can learn any state produced by k T gates in time $\mathrm{poly}(n)\exp(k)$*.*

Other results:

- Pseudorandom state distinguisher[1]: [GIKL23c].
- First non-trivial estimator of stabilizer fidelity: [GIKL23c].

[1] See Simons Talk for more information!

Our Work

Theorem ([GIKL23a, LOH23, HG23])

Can learn any state produced by k T gates in time $\mathrm{poly}(n)\exp(k)$.

Other results:

- Pseudorandom state distinguisher[1]: [GIKL23c].
- First non-trivial estimator of stabilizer fidelity: [GIKL23c].
- Improved (tolerant) property tester for stabilizer states: [GIKL23c].

[1] See Simons Talk for more information!

Our Work

Theorem ([GIKL23a, LOH23, HG23])

Can learn any state produced by k T gates in time $\mathrm{poly}(n)\exp(k)$.

Other results:

- Pseudorandom state distinguisher[1]: [GIKL23c].
- First non-trivial estimator of stabilizer fidelity: [GIKL23c].
- Improved (tolerant) property tester for stabilizer states: [GIKL23c].
- **Single-copy learning algorithm:** [GIKL23b, CLL23].

[1] See Simons Talk for more information!

Learning States with Many Stabilizers

Lemma

Let $|\psi\rangle$ be produced by Clifford gates and at most k T gates. Then $|Stab(|\psi\rangle)| \geq 2^{n-k}$.

Learning States with Many Stabilizers

Lemma

Let $|\psi\rangle$ be produced by Clifford gates and at most k T gates. Then $|Stab(|\psi\rangle)| \geq 2^{n-k}$.

Suffices to consider the set of states such that Stab$(|\psi\rangle)$ is large!

Learning States with Many Stabilizers

Lemma

Let $|\psi\rangle$ *be produced by Clifford gates and at most* k T *gates. Then* $|\mathit{Stab}(|\psi\rangle)| \geq 2^{n-k}$.

Suffices to consider the set of states such that Stab$(|\psi\rangle)$ is large!

Theorem ([GIKL23a])

Can learn any state such that $|\mathit{Stab}(|\psi\rangle)| \geq 2^{n-k}$ *in time* $\mathrm{poly}(n)\exp(k)$.

A Compression Scheme

Critical Observation

Let $|\text{Stab}(|\psi\rangle)| \geq 2^{n-k}$. Then learning $\text{Stab}(|\psi\rangle)$ is enough to learn $|\psi\rangle$ in time $\text{poly}(n)\exp(k)$.

A Compression Scheme

Critical Observation

Let $|\text{Stab}(|\psi\rangle)| \geq 2^{n-k}$. Then learning $\text{Stab}(|\psi\rangle)$ is enough to learn $|\psi\rangle$ in time $\mathrm{poly}(n)\exp(k)$.

Given $\text{Stab}(|\psi\rangle)$, there exists a Clifford circuit C such that

$$C\,|\psi\rangle = |x\rangle \otimes \underbrace{|\varphi\rangle}_{k \text{ qubits}}$$

for some $x \in \{0,1\}^{n-k}$.

Moreover, C can be computed in time $O(n^2)$.

Initial Algorithm

Algorithm 1: First Approach

Input: Copies of $|\psi\rangle$ and description of Stab($|\psi\rangle$)

Promise: $|\text{Stab}(|\psi\rangle)| \geq 2^{n-k}$

Output: $|\widehat{\psi}\rangle \approx |\psi\rangle$

1. Find C such that $C\,|\psi\rangle = |x\rangle\,|\varphi\rangle$.
2. Measure first register of $C\,|\psi\rangle$ to learn x.
3. Perform pure state tomography on second register to get $|\widehat{\varphi}\rangle \approx |\varphi\rangle$.
4. Output $C^\dagger\,|x\rangle\,|\widehat{\varphi}\rangle$.

Initial Algorithm

Algorithm 1: First Approach

Input: Copies of $|\psi\rangle$ and description of Stab$(|\psi\rangle)$

Promise: $|\text{Stab}(|\psi\rangle)| \geq 2^{n-k}$

Output: $|\widehat{\psi}\rangle \approx |\psi\rangle$

1. Find C such that $C\,|\psi\rangle = |x\rangle\,|\varphi\rangle$.
2. Measure first register of $C\,|\psi\rangle$ to learn x.
3. Perform pure state tomography on second register to get $|\widehat{\varphi}\rangle \approx |\varphi\rangle$.
4. Output $C^\dagger\,|x\rangle\,|\widehat{\varphi}\rangle$.

How do we find Stab$(|\psi\rangle)$?

Characteristic Distribution p_ψ

For $W \in \{I, X, Y, Z\}^{\otimes n}$,

$$p_\psi(W) := \frac{1}{2^n} \langle \psi | W | \psi \rangle^2 .$$

Characteristic Distribution p_ψ

For $W \in \{I, X, Y, Z\}^{\otimes n}$,

$$p_\psi(W) := \frac{1}{2^n} \langle \psi | W | \psi \rangle^2 .$$

- p_ψ is a distribution [Mon17]
- Can sample from $q_\psi = p_\psi * p_\psi$ via Bell difference sampling [GNW21]

A Fourier Duality Theorem

Theorem ([GIKL23c])

Given a subgroup $G \subseteq \{I, X, Y, Z\}^{\otimes n}$*:*

$$\sum_{W \in G} p_\psi(W) = \frac{|G|}{2^n} \sum_{W \in G^\perp} p_\psi(W)$$

A Fourier Duality Theorem

Theorem ([GIKL23c])

Given a subgroup $G \subseteq \{I, X, Y, Z\}^{\otimes n}$*:*

$$\sum_{W \in G} p_\psi(W) = \frac{|G|}{2^n} \sum_{W \in G^\perp} p_\psi(W)$$

Definition

$G^\perp \subseteq \{I, X, Y, Z\}^{\otimes n}$ is the set of Pauli matrices that commutes with all of G.

Claim: $\left(G^\perp\right)^\perp = G$.

Corollary

The support of p_ψ lies in Stab$(|\psi\rangle)^\perp$.

Proof.

Corollary

The support of p_ψ *lies in Stab*$(|\psi\rangle)^\perp$.

Proof.

By duality theorem:

$$\sum_{W \in \text{Stab}(|\psi\rangle)} p_\psi(W) = \frac{|\text{Stab}(|\psi\rangle)|}{2^n} \sum_{W \in \text{Stab}(|\psi\rangle)^\perp} p_\psi(W).$$

Corollary

The support of p_ψ lies in Stab$(|\psi\rangle)^\perp$.

Proof.

By duality theorem:

$$\sum_{W \in \text{Stab}(|\psi\rangle)} p_\psi(W) = \frac{|\text{Stab}(|\psi\rangle)|}{2^n} \sum_{W \in \text{Stab}(|\psi\rangle)^\perp} p_\psi(W).$$

By definition of Stab$(|\psi\rangle)$:

$$\sum_{W \in \text{Stab}(|\psi\rangle)} p_\psi(W) = \sum_{W \in \text{Stab}(|\psi\rangle)} \frac{1}{2^n} \langle\psi|W|\psi\rangle^2 = \frac{|\text{Stab}(|\psi\rangle)|}{2^n}.$$

□

How to learn Stab($|\psi\rangle$)

Algorithm 2: Learning Algorithm v2

Input: Copies of $|\psi\rangle$
Promise: $|\text{Stab}(|\psi\rangle)| \geq 2^{n-k}$
Output: $|\widehat{\psi}\rangle \approx |\psi\rangle$

Draw $m = O(n)$ samples: $W_1, W_2, \cdots W_m \sim p_\psi$.
Compute $\widehat{\text{Stab}(|\psi\rangle)} := \langle W_1, W_2, \cdots, W_m \rangle^\perp$.

Claim: Given G, $G^\perp$ can be computed in time $O(n^3)$.

How to learn Stab($|\psi\rangle$)

Algorithm 2: Learning Algorithm v2

Input: Copies of $|\psi\rangle$

Promise: $|\text{Stab}(|\psi\rangle)| \geq 2^{n-k}$

Output: $|\widehat{\psi}\rangle \approx |\psi\rangle$

1 Draw $m = O(n)$ samples: $W_1, W_2, \cdots W_m \sim p_\psi$.

2 Compute $\widehat{\text{Stab}(|\psi\rangle)} := \langle W_1, W_2, \cdots, W_m \rangle^\perp$.

3 Run compression scheme from previous algorithm.

Claim: Given G, $G^\perp$ can be computed in time $O(n^3)$.

A More Robust Algorithm

Problem: Cannot always learn the support of p_ψ exactly.

A More Robust Algorithm

Problem: Cannot always learn the support of p_ψ exactly.

$$G \subsetneq \mathsf{Stab}(|\psi\rangle)^\perp \iff G^\perp \supsetneq \mathsf{Stab}(|\psi\rangle).$$

A More Robust Algorithm

Problem: Cannot always learn the support of p_ψ exactly.

$$G \subsetneq \text{Stab}(|\psi\rangle)^\perp \iff G^\perp \supsetneq \text{Stab}(|\psi\rangle).$$

Solution: Learning *almost* all of the support is sufficient!

Robustness Lemma

Solution: Learning *almost* all of the support is sufficient!

Robustness Lemma

Solution: Learning *almost* all of the support is sufficient!

Lemma ([GIKL23a])

Let $G \subseteq$ *Stab*$(|\psi\rangle)^{\perp}$ *such that*

$$\sum_{W \in G} p_{\psi}(W) = 1 - \varepsilon^2.$$

Then $|\psi\rangle \approx_{\varepsilon} |\varphi\rangle$ *such that Stab*$(|\varphi\rangle) = G^{\perp}$*.*

Robustness Lemma

Solution: Learning *almost* all of the support is sufficient!

Lemma ([GIKL23a])

Let $G \subseteq \mathrm{Stab}(|\psi\rangle)^{\perp}$ *such that*

$$\sum_{W \in G} p_\psi(W) = 1 - \varepsilon^2.$$

Then $|\psi\rangle \approx_\varepsilon |\varphi\rangle$ *such that* $\mathrm{Stab}(|\varphi\rangle) = G^{\perp}$.

Can learn such a subgroup with $O(n/\varepsilon^2)$ samples.

The Learning Algorithm

Algorithm 3: Tomography of States with many Stabilizers

Input: Copies of $|\psi\rangle$
Promise: $|\text{Stab}(|\psi\rangle)| \geq 2^{n-k}$
Output: $|\widehat{\psi}\rangle \approx_\varepsilon |\psi\rangle$

1 Perform draw $m = O(n/\varepsilon^2)$ samples: $W_1, W_2, \cdots W_m$.
2 Compute $\langle W_1, W_2, \cdots, W_m \rangle^\perp$.
3 Apply C such that $C|\psi\rangle \approx_\varepsilon |x\rangle |\varphi\rangle$.

The Learning Algorithm

Algorithm 3: Tomography of States with many Stabilizers

Input: Copies of $|\psi\rangle$
Promise: $|\mathsf{Stab}(|\psi\rangle)| \geq 2^{n-k}$
Output: $|\widehat{\psi}\rangle \approx_\varepsilon |\psi\rangle$

1. Perform draw $m = O(n/\varepsilon^2)$ samples: $W_1, W_2, \cdots W_m$.
2. Compute $\langle W_1, W_2, \cdots, W_m \rangle^\perp$.
3. Apply C such that $C|\psi\rangle \approx_\varepsilon |x\rangle |\varphi\rangle$.
4. Measure first register of $C|\psi\rangle$ $O(1)$ times to learn x.

The Learning Algorithm

Algorithm 3: Tomography of States with many Stabilizers

Input: Copies of $|\psi\rangle$
Promise: $|\text{Stab}(|\psi\rangle)| \geq 2^{n-k}$
Output: $|\widehat{\psi}\rangle \approx_\varepsilon |\psi\rangle$

1. Perform draw $m = O(n/\varepsilon^2)$ samples: $W_1, W_2, \cdots W_m$.
2. Compute $\langle W_1, W_2, \cdots, W_m \rangle^\perp$.
3. Apply C such that $C|\psi\rangle \approx_\varepsilon |x\rangle |\varphi\rangle$.
4. Measure first register of $C|\psi\rangle$ $O(1)$ times to learn x.
5. Post-select on measuring $|x\rangle$ then run pure state tomography on $|\varphi\rangle$.

The Learning Algorithm

Algorithm 3: Tomography of States with many Stabilizers

Input: Copies of $|\psi\rangle$
Promise: $|\text{Stab}(|\psi\rangle)| \geq 2^{n-k}$
Output: $|\widehat{\psi}\rangle \approx_{\varepsilon} |\psi\rangle$

Perform draw $m = O(n/\varepsilon^2)$ samples: $W_1, W_2, \cdots W_m$.
Compute $\langle W_1, W_2, \cdots, W_m \rangle^{\perp}$.
Apply C such that $C|\psi\rangle \approx_{\varepsilon} |x\rangle |\varphi\rangle$.
Measure first register of $C|\psi\rangle$ $O(1)$ times to learn x.
Post-select on measuring $|x\rangle$ then run pure state tomography on $|\varphi\rangle$.
Return $C^{\dagger} |0^{n-k'}\rangle \otimes |\widehat{\varphi}\rangle$.

Proof of Robustness Lemma

Lemma ([GIKL23a])

Let $G \subseteq$ *Stab*$(|\psi\rangle)^\perp$ *such that*

$$\sum_{W \in G} p_\psi(W) = 1 - \varepsilon^2.$$

Then $|\psi\rangle \approx_\varepsilon |\varphi\rangle$ *such that Stab*$(|\varphi\rangle) = G^\perp$*.*

Proof of Robustness Lemma

Lemma ([GIKL23a])

Let $G \subseteq$ *Stab*$(|\psi\rangle)^{\perp}$ *such that*

$$\sum_{W \in G} p_{\psi}(W) = 1 - \varepsilon^2.$$

Then $|\psi\rangle \approx_{\varepsilon} |\varphi\rangle$ *such that Stab*$(|\varphi\rangle) = G^{\perp}$*.*

Goal:

$$|\psi\rangle \coloneqq C^{\dagger} \sum_{x \in \{0,1\}^{n-l}} \alpha_x |x\rangle \otimes |\varphi_x\rangle .$$

$$\max_{x \in \{0,1\}^{n-l}} |\alpha_x|^2 \geq 1 - \varepsilon^2$$

Collision Probability and p_ψ

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Collision Probability and p_ψ

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Then the collision probability of the first register is:

$$\sum_{x \in \{0,1\}^{n-l}} |\alpha_x|^4$$

Collision Probability and p_ψ

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Then the collision probability of the first register is:

$$\sum_{x \in \{0,1\}^{n-l}} |\alpha_x|^4 = \sum_{x \in \{0,1\}^{l}} \mathrm{Tr}\left[\left(|x\rangle\langle x| \otimes I^l \right)^{\otimes 2} |\psi\rangle\langle\psi|^{\otimes 2} \right]$$

Collision Probability and p_ψ

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Then the collision probability of the first register is:

$$\sum_{x \in \{0,1\}^{n-l}} |\alpha_x|^4 = \sum_{x \in \{0,1\}^l} \mathrm{Tr}\left[\left(|x\rangle\langle x| \otimes I^l\right)^{\otimes 2} |\psi\rangle\langle\psi|^{\otimes 2}\right]$$

$$= 2^l \cdot \sum_{W \in \mathcal{Z}^{n-l}} p_\psi(W)$$

Collision Probability and p_ψ

$$\mathcal{Z}^t \coloneqq \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Then the collision probability of the first register is:

$$\sum_{x\in\{0,1\}^{n-l}} |\alpha_x|^4 = \sum_{x\in\{0,1\}^l} \mathrm{Tr}\left[\left(|x\rangle\langle x| \otimes I^l\right)^{\otimes 2} |\psi\rangle\langle\psi|^{\otimes 2}\right]$$

$$= 2^l \cdot \sum_{W\in\mathcal{Z}^{n-l}} p_\psi(W)$$

$$\max_{x\in\{0,1\}^{n-l}} |\alpha_x|^2 = \max_{x\in\{0,1\}^{n-l}} |\alpha_x|^2 \cdot 1$$

Collision Probability and p_ψ

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Then the collision probability of the first register is:

$$\sum_{x \in \{0,1\}^{n-l}} |\alpha_x|^4 = \sum_{x \in \{0,1\}^l} \mathrm{Tr}\left[\left(|x\rangle\langle x| \otimes I^l\right)^{\otimes 2} |\psi\rangle\langle\psi|^{\otimes 2}\right]$$
$$= 2^l \cdot \sum_{W \in \mathcal{Z}^{n-l}} p_\psi(W)$$

$$\max_{x \in \{0,1\}^{n-l}} |\alpha_x|^2 = \max_{x \in \{0,1\}^{n-l}} |\alpha_x|^2 \cdot \sum_{x \in \{0,1\}^{n-1}} |\alpha_x|^2$$

Collision Probability and p_ψ

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Then the collision probability of the first register is:

$$\sum_{x \in \{0,1\}^{n-l}} |\alpha_x|^4 = \sum_{x \in \{0,1\}^l} \mathrm{Tr}\left[\left(|x\rangle\langle x| \otimes I^l\right)^{\otimes 2} |\psi\rangle\langle\psi|^{\otimes 2}\right]$$

$$= 2^l \cdot \sum_{W \in \mathcal{Z}^{n-l}} p_\psi(W)$$

$$\max_{x \in \{0,1\}^{n-l}} |\alpha_x|^2 = \max_{x \in \{0,1\}^{n-l}} |\alpha_x|^2 \cdot \sum_{x \in \{0,1\}^{n-1}} |\alpha_x|^2 \geq \sum_{x \in \{0,1\}^{n-l}} |\alpha_x|^4$$

Proof (cont.)

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Let $G \subset \{I, X, Y, Z\}^{\otimes n}$ such that $G^{\perp} = \mathcal{Z}^{n-l}$.

Proof (cont.)

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Let $G \subset \{I, X, Y, Z\}^{\otimes n}$ such that $G^\perp = \mathcal{Z}^{n-l}$.

$$1 - \varepsilon^2 \leq \sum_{W \in G} p_\psi(W)$$

Proof (cont.)

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Let $G \subset \{I, X, Y, Z\}^{\otimes n}$ such that $G^\perp = \mathcal{Z}^{n-l}$.

$$\begin{aligned} 1 - \varepsilon^2 &\leq \sum_{W \in G} p_\psi(W) \\ &= \frac{2^n}{|G|} \sum_{W \in G^\perp} p_\psi(W) \end{aligned}$$

Proof (cont.)

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Let $G \subset \{I, X, Y, Z\}^{\otimes n}$ such that $G^\perp = \mathcal{Z}^{n-l}$.

$$\begin{aligned} 1 - \varepsilon^2 &\leq \sum_{W \in G} p_\psi(W) \\ &= \frac{2^n}{|G|} \sum_{W \in G^\perp} p_\psi(W) \\ &= 2^l \cdot \sum_{W \in \mathcal{Z}^{n-l}} p_\psi(W) \end{aligned}$$

Proof (cont.)

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Let $G \subset \{I, X, Y, Z\}^{\otimes n}$ such that $G^\perp = \mathcal{Z}^{n-l}$.

$$\begin{aligned}
1 - \varepsilon^2 &\leq \sum_{W \in G} p_\psi(W) \\
&= \frac{2^n}{|G|} \sum_{W \in G^\perp} p_\psi(W) \\
&= 2^l \cdot \sum_{W \in \mathcal{Z}^{n-l}} p_\psi(W) \\
&\leq \max_{x \in \{0,1\}^{n-l}} |\alpha_x|^2.
\end{aligned}$$

Proof (cont.)

$$\mathcal{Z}^t := \{I, Z\}^{\otimes t} \otimes I^{\otimes n-t}$$

Let $G \subset \{I, X, Y, Z\}^{\otimes n}$ such that $G^\perp = \mathcal{Z}^{n-l}$.

$$\begin{aligned} 1 - \varepsilon^2 &\leq \sum_{W \in G} p_\psi(W) \\ &= \frac{2^n}{|G|} \sum_{W \in G^\perp} p_\psi(W) \\ &= 2^l \cdot \sum_{W \in \mathcal{Z}^{n-l}} p_\psi(W) \\ &\leq \max_{x \in \{0,1\}^{n-l}} |\alpha_x|^2. \end{aligned}$$

$$|\psi\rangle \approx_\varepsilon |x_{\text{max}}\rangle \, |\varphi_{x_{\text{max}}}\rangle$$

Finding C

Lemma ([GIKL23a])

Let $G \subseteq \{I, X, Y, Z\}^{\otimes n}$ *such that*

$$\sum_{W \in G} p_\psi(W) > \frac{3}{4}.$$

Then there exists a Clifford circuit C *such that* $C(G^\perp) = \mathcal{Z}^{n-l}$.

Finding C

Lemma ([GIKL23a])

Let $G \subseteq \{I, X, Y, Z\}^{\otimes n}$ *such that*

$$\sum_{W \in G} p_\psi(W) > \frac{3}{4}.$$

Then there exists a Clifford circuit C *such that* $C(G^\perp) = \mathcal{Z}^{n-l}$.

Let $|\psi'\rangle := C|\psi\rangle$:

Finding C

Lemma ([GIKL23a])

Let $G \subseteq \{I, X, Y, Z\}^{\otimes n}$ *such that*

$$\sum_{W \in G} p_\psi(W) > \frac{3}{4}.$$

Then there exists a Clifford circuit C *such that* $C(G^\perp) = \mathcal{Z}^{n-l}$.

Let $|\psi'\rangle := C|\psi\rangle$:

$$\sum_{x \in G} p_\psi(x) = \sum_{x \in C(G)} p_{\psi'}(x)$$

Algorithm Overview

Let $\varepsilon \in (0, 1)$:

Algorithm Overview

Let $\varepsilon \in (0, 1)$:

1. Draw $m = O\left(n/\varepsilon^2\right)$ samples from p_ψ: $W_1, W_2, \cdots W_m$.
   - Let $G := \langle W_1, W_2, \cdots W_m \rangle$. Then w.h.p.

$$\sum_{W \in G} p_\psi(W) \geq 1 - \varepsilon^2/4$$

Algorithm Overview

Let $\varepsilon \in (0,1)$:

1. Draw $m = O\left(n/\varepsilon^2\right)$ samples from p_ψ: $W_1, W_2, \cdots W_m$.
   - Let $G := \langle W_1, W_2, \cdots W_m \rangle$. Then w.h.p.

$$\sum_{W \in G} p_\psi(W) \geq 1 - \varepsilon^2/4$$

2. Compute $G^\perp$.
   - $G^\perp \supseteq \text{Stab}(|\psi\rangle)$.

Algorithm Overview

Let $\varepsilon \in (0,1)$:

1. Draw $m = O\left(n/\varepsilon^2\right)$ samples from p_ψ: $W_1, W_2, \cdots W_m$.
   - Let $G := \langle W_1, W_2, \cdots W_m \rangle$. Then w.h.p.

$$\sum_{W \in G} p_\psi(W) \geq 1 - \varepsilon^2/4$$

2. Compute $G^\perp$.
   - $G^\perp \supseteq \text{Stab}(|\psi\rangle)$.
3. Apply C such that $C(G^\perp) = \{I, Z\}^{\otimes n-l} \otimes I^{\otimes l}$.
   - $C\,|\psi\rangle = \sum_{x \in \{0,1\}^{n-l}} \alpha_x \,|x\rangle\,|\varphi_x\rangle$ such that

$$\max_{x \in \{0,1\}^{n-l}} |\alpha_x|^2 \geq 1 - \varepsilon^2/4.$$

Algorithm Overview (cont.)

4. Measure first register of $C|\psi\rangle$ $O(1)$ times, to learn x_{max} w.h.p.
   - Probability of measuring x_{max} is $|\alpha_{x_{\text{max}}}|^2 > 3/4$.

Algorithm Overview (cont.)

4. Measure first register of $C \left|\psi\right\rangle$ $O(1)$ times, to learn x_{max} w.h.p.
   - Probability of measuring x_{max} is $|\alpha_{x_{\text{max}}}|^2 > 3/4$.
5. Post-select on measuring $\left|x_{\text{max}}\right\rangle$.
   - Left with $\left|\eta\right\rangle := \left|x_{\text{max}}\right\rangle \left|\varphi\right\rangle$.
   - $d_{\text{Tr}}\left(C\left|\psi\right\rangle, \left|\eta\right\rangle\right) \leq \varepsilon/2$.

Algorithm Overview (cont.)

4. Measure first register of $C\left|\psi\right\rangle$ $O(1)$ times, to learn x_{max} w.h.p.
   - Probability of measuring x_{max} is $|\alpha_{x_{\text{max}}}|^2 > 3/4$.
5. Post-select on measuring $\left|x_{\text{max}}\right\rangle$.
   - Left with $\left|\eta\right\rangle := \left|x_{\text{max}}\right\rangle\left|\varphi\right\rangle$.
   - $d_{\text{Tr}}\left(C\left|\psi\right\rangle, \left|\eta\right\rangle\right) \leq \varepsilon/2$.
6. Run pure state tomography on second register such that $d_{\text{Tr}}\left(\left|\varphi\right\rangle, \left|\widehat{\varphi}\right\rangle\right) \leq \varepsilon/2$.
   - Total trace distance is at most $\varepsilon/2 + \varepsilon/2$ via triangle inequality.

Algorithm Overview (cont.)

4. Measure first register of $C|\psi\rangle$ $O(1)$ times, to learn x_{max} w.h.p.
   - Probability of measuring x_{max} is $|\alpha_{x_{\text{max}}}|^2 > 3/4$.
5. Post-select on measuring $|x_{\text{max}}\rangle$.
   - Left with $|\eta\rangle := |x_{\text{max}}\rangle|\varphi\rangle$.
   - $d_{\text{Tr}}(C|\psi\rangle, |\eta\rangle) \leq \varepsilon/2$.
6. Run pure state tomography on second register such that $d_{\text{Tr}}(|\varphi\rangle, |\widehat{\varphi}\rangle) \leq \varepsilon/2$.
   - Total trace distance is at most $\varepsilon/2 + \varepsilon/2$ via triangle inequality.
7. Output $C^\dagger|x_{\text{max}}\rangle|\widehat{\varphi}\rangle$.
   - Trace distance preserved by unitaries.

Open Questions

Lower Bounds

Current best-known lower bounds are $\approx \Omega(\sqrt[4]{k})$, due to unitary t-designs [HMMH$^+$23].

Open Questions

Lower Bounds

Current best-known lower bounds are $\approx \Omega(\sqrt[4]{k})$, due to unitary t-designs [HMMH$^+$23].

Proper Learning

Output state is not necessarily produced by $O(\log n)$ T-gates, can be as many as $\mathrm{poly}(n)$.

Thank You!