① Optimality of Grover's algorithm

- **Recall**: We have a Boolean function $f:\{0,1\}^n \to \{0,1\}$ and we want to identify/search for a "marked" item — an item $\vec{x}^*$ such that $f(\vec{x}^*)=1$.

- Grover's algorithm makes use of $O(\sqrt{N})$ queries to the phase oracle

  $Z_f|\vec{x}\rangle = (-1)^{f(\vec{x})}|\vec{x}\rangle$, and finds a marked element with high probability.

- Is $O(\sqrt{N})$ the best we can do? YES!

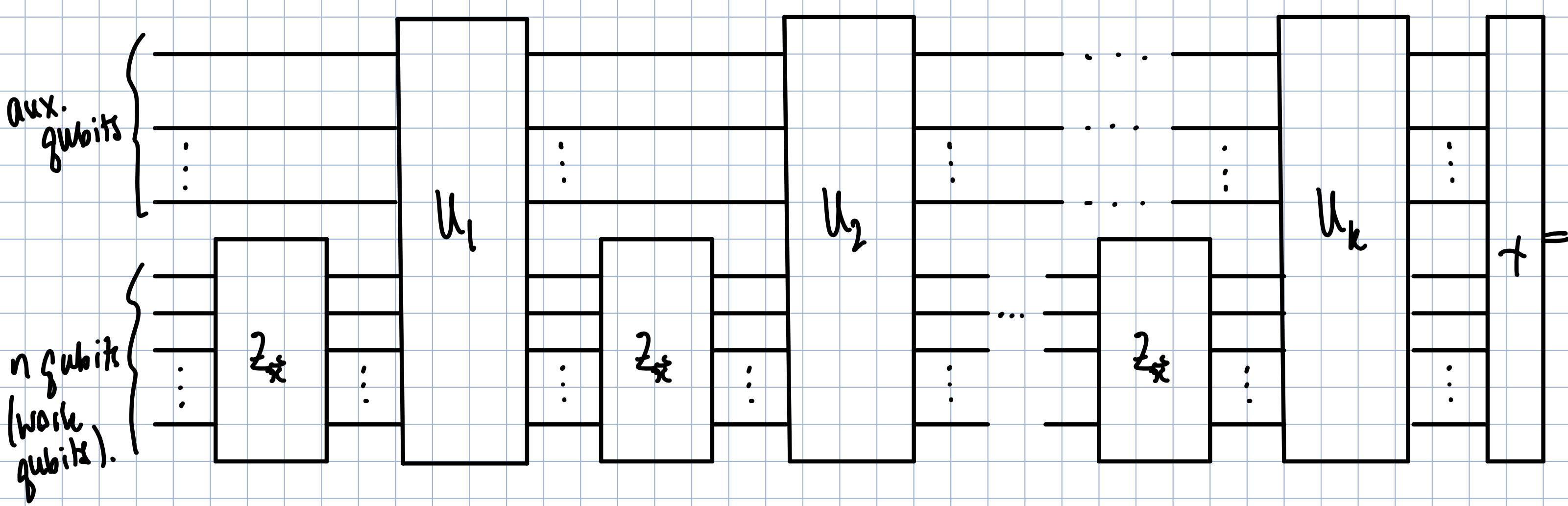
$$Z_f|\vec{x}\rangle = \begin{cases} |\vec{x}\rangle & \text{if } \vec{x} \text{ is not marked} \\ -|\vec{x}\rangle & \text{if } \vec{x} \text{ is marked.} \end{cases}$$

(a) Proof of Optimality: Setup

- Consider exactly one marked element $\vec{x}^* \in \{0,1\}^n$. The phase oracle for this is
  $$Z_f \equiv Z_{\vec{x}^*} = \mathbb{1} - 2|\vec{x}^*\rangle\langle\vec{x}^*|.$$
  ↳ for simplicity (it can be relaxed)

  Suppose the marked element is uniformly random. (also an assumption for simplicity, which can be relaxed)

- Consider the most general algorithm that uses the oracle k times (including algorithms with additional qubits):



Initial state is $|\psi_0\rangle$ (over all qubits).

Final state before measurement: $|\psi_k\rangle = U_k Z_{\vec{x}^*} U_{k-1} \cdots Z_{\vec{x}^*} U_1 Z_{\vec{x}^*} |\psi_0\rangle$ (with k oracle calls).

- Consider the simpler problem of figuring out whether there is a marked element at all or not. A lower bound on this implies a lower bound on actually finding the marked element.

⊛ Formally: there are two versions of the unstructured search problem:

(1) The search version: find a marked element — or conclude that a marked element does not exist. Q_{search} is the query complexity.

(2) The decision version: decide whether (a) $f(x)=0\ \forall x$; (b) There exists x^* such that $f(x^*)=1$. $Q_{decision}$ is the query complexity. (Note: Here we do not have to find a marked element.)

⊛ Observe: search ⇒ decision (any algorithm solving the search problem is a valid algorithm for solving the decision problem.)

Therefore: $Q_{decision} \leq Q_{search} \rightarrow Q_{decision}$ is a lower bound on Q_{search}!

⊛ Grover's algorithm gives an upper bound on Q_{search}: $Q_{search} \leq O(\sqrt{N})$.

$\Rightarrow \quad Q_{decision} \leq Q_{search} \leq O(\sqrt{N})$.

⊛ Our goal is to show that $Q_{decision} \geq \Omega(\sqrt{N})$. (to succeed w/ high probability).

- No marked element ⇒ $Z_f \equiv \mathbb{1}$. Final state is: $|\psi_k\rangle = U_k \cdots U_2 U_1 |\psi_0\rangle$

How well can we distinguish b/w $|\tilde{\psi}_k\rangle$ and $|\psi_k\rangle$, as a function of k and $N=2^n$? What k do we need to distinguish them with high probability?

one marked element

With no marked element

② Aside: Quantum State Distinguishability / Discrimination. (aka: quantum hypothesis testing).

- Given: two states ρ_0, ρ_1 with probabilities q_0, q_1.
- Problem: someone gives you one of these two states, but doesn't tell you which one.
- Goal: Do a measurement on the state and make the correct guess with the highest possible probability.

(a) Objective function: success probability.

- The measurement is described by a two-outcome POVM: $\{M_0, M_1\}$.

  ($M_0, M_1 \geq 0$ (PSD) and $M_0 + M_1 = \mathbb{1}$.)

- "Success" means that the correct guess is made.

Strategy:
- If measurement outcome is "0", guess ρ_0.
- If measurement outcome is "1", guess ρ_1.

$$\Pr[\text{Success}] = \underbrace{\Pr[\text{State is } \rho_0]}_{q_0} \cdot \underbrace{\Pr[\text{guess is "0"} \mid \text{State is } \rho_0]}_{\mathrm{Tr}[M_0\rho_0]}$$

$$+ \underbrace{\Pr[\text{State is } \rho_1]}_{q_1} \cdot \underbrace{\Pr[\text{guess is "1"} \mid \text{state is } \rho_1]}_{\mathrm{Tr}[M_1\rho_1]}.$$

$$= q_0 \mathrm{Tr}[M_0\rho_0] + q_1 \mathrm{Tr}[M_1\rho_1]$$

$M_0 + M_1 = \mathbb{1}$

$M_1 = \mathbb{1} - M_0$

$$= q_0 \mathrm{Tr}[M_0\rho_0] + \underbrace{q_1 \mathrm{Tr}[(\mathbb{1} - M_0)\rho_1]}$$

$\to q_1 \mathrm{Tr}[\rho_1] - q_1 \mathrm{Tr}[M_0\rho_1]$.

$$= \mathrm{Tr}[M_0(q_0\rho_0 - q_1\rho_1)] + q_1 \underbrace{\mathrm{Tr}[\rho_1]}_{=1}$$

$$= \mathrm{Tr}[M_0(q_0\rho_0 - q_1\rho_1)] + q_1$$

(Note: $0 \leq M_0 \leq \mathbb{1}$)

(b) Optimal Success Probability: Optimize over all measurements:

(Equivalently, optimize over all two-outcome POVMs.)

$$p_{succ}^{opt} := q_1 + \max_{0 \le M \le \mathbb{1}} \mathrm{Tr}\left[M\left(q_0 \rho_0^{\otimes n} - q_1 \rho_1^{\otimes n}\right)\right] = \frac{1}{2}\left(1 + \|q_0\rho_0^{\otimes n} - q_1 \rho_1^{\otimes n}\|_1\right)$$

We will prove this!

Recall trace norm: $\|\cdot\|_1$ is the sum of the singular values. For a Hermitian operator, singular values are sum of the absolute values of the eigenvalues:

$$H = \sum_{k=1}^{d} \underbrace{\lambda_k}_{\text{eigenvalues}} \underbrace{|\psi_k\rangle\langle\psi_k|}_{\text{eigenvectors.}} \Rightarrow \|H\|_1 = \mathrm{Tr}\left[\sqrt{H^\dagger H}\right]$$

$$= \mathrm{Tr}\left[\left(\sum_{k=1}^{d} \lambda_k^2 |\psi_k\rangle\langle\psi_k|\right)^{1/2}\right]$$

$$= \mathrm{Tr}\left[\sum_{k=1}^{d} \underbrace{\sqrt{\lambda_k^2}}_{=|\lambda_k| \text{ b/c } \lambda_k \text{ could be negative!}} |\psi_k\rangle\langle\psi_k|\right]$$

$$= \sum_{k=1}^{d} |\lambda_k|.$$

<u>Lemma</u>: For Hermitian H, $\|H\|_1 = \max\{\mathrm{Tr}[(M_0 - M_1)H] : M_0, M_1 \geq 0,\ M_0 + M_1 = \mathbb{1}\}$

<u>Proof</u>: $$H = \sum_{k=1}^{d} \lambda_k |\psi_k\rangle\langle\psi_k| = \underbrace{\sum_{k:\lambda_k \geq 0} \lambda_k |\psi_k\rangle\langle\psi_k|}_{\text{Positive part}} - \underbrace{\sum_{k:\lambda_k < 0} |\lambda_k| |\psi_k\rangle\langle\psi_k|}_{\text{Negative part.}}$$

$\Pi_+ \cdot \Pi_- = 0$

$\Pi_- \cdot \Pi_+ = 0$

P is a projector if $P^\dagger = P$, $P^2 = P$

Π_+ : projector onto positive part of H — $\Pi_+ = \sum_{k:\lambda_k \geq 0} |\psi_k\rangle\langle\psi_k|$.

Π_- : projector onto negative part of H — $\Pi_- = \sum_{k:\lambda_k < 0} |\psi_k\rangle\langle\psi_k|$.

These are orthogonal!

<u>Note</u>: $\Pi_+ + \Pi_- = \sum_{k=1}^{d} |\psi_k\rangle\langle\psi_k| = \mathbb{1}$ (b/c the eigenvectors form orthonormal basis.)

Then, $\mathrm{Tr}[H(\Pi_+ - \Pi_-)] = \mathrm{Tr}[H\Pi_+] - \mathrm{Tr}[H\Pi_-]$

$$H = \sum_{k=1}^{d} \lambda_k |\psi_k\rangle\langle\psi_k| = \underbrace{\sum_{k:\lambda_k\geq 0} \lambda_k |\psi_k\rangle\langle\psi_k|}_{\text{positive part}} \underbrace{- \sum_{k:\lambda_k<0} |\lambda_k| |\psi_k\rangle\langle\psi_k|}_{\text{negative part.}}$$

$$= \sum_{k:\lambda_k\geq 0} \lambda_k \underbrace{\mathrm{Tr}[|\psi_k\rangle\langle\psi_k|\Pi_+]}_{=1\ \forall k} - \sum_{k:\lambda_k<0} |\lambda_k| \underbrace{\mathrm{Tr}[|\psi_k\rangle\langle\psi_k|\Pi_+]}_{=0}$$

$$- \sum_{k:\lambda_k\geq 0} \lambda_k \underbrace{\mathrm{Tr}[|\psi_k\rangle\langle\psi_k|\Pi_-]}_{=0} + \sum_{k:\lambda_k<0} |\lambda_k| \underbrace{\mathrm{Tr}[|\psi_k\rangle\langle\psi_k|\Pi_-]}_{=1}$$

$$= \sum_{k:\lambda_k\geq 0} |\lambda_k| + \sum_{k:\lambda_k<0} |\lambda_k| \qquad x = |x|\ \forall\, x \geq 0$$

$$= \sum_{k=1}^{d} |\lambda_k| = \|H\|_1.$$

So $\max\{\mathrm{Tr}[H(M_0 - M_1)] : M_0, M_1 \geq 0,\ M_0 + M_1 = \mathbb{1}\} \geq \|H\|_1$

Pick $M_0 = \Pi_+$, $M_1 = \Pi_-$

Now show the opposite inequality.

Pick an arbitrary pair $M_0, M_1 \geq 0$, $M_0 + M_1 = \mathbb{1}$.

$\mathrm{Tr}[H(M_0 - M_1)] = \mathrm{Tr}[HM_0] - \mathrm{Tr}[HM_1]$

$$H = \sum_{k=1}^{d} \lambda_k |\psi_k\rangle\langle\psi_k| = \underbrace{\sum_{k:\lambda_k\geq 0} \lambda_k |\psi_k\rangle\langle\psi_k|}_{\text{positive part}} \underbrace{- \sum_{k:\lambda_k<0} |\lambda_k| |\psi_k\rangle\langle\psi_k|}_{\text{negative part.}}$$

$$= \sum_{k:\lambda_k\geq 0} \lambda_k \mathrm{Tr}[|\psi_k\rangle\langle\psi_k|M_0] - \sum_{k:\lambda_k<0} |\lambda_k| \underbrace{\mathrm{Tr}[|\psi_k\rangle\langle\psi_k|M_0]}_{\geq 0\ \forall k}$$

$$- \sum_{k:\lambda_k\geq 0} \lambda_k \underbrace{\mathrm{Tr}[|\psi_k\rangle\langle\psi_k|M_1]}_{\geq 0\ \forall k} + \sum_{k:\lambda_k<0} |\lambda_k| \mathrm{Tr}[|\psi_k\rangle\langle\psi_k|M_1]$$

$$\leq \sum_{k:\lambda_k\geq 0} \lambda_k \underbrace{\mathrm{Tr}[|\psi_k\rangle\langle\psi_k|M_0]}_{\leq 1\ \forall k} + \sum_{k:\lambda_k<0} |\lambda_k| \underbrace{\mathrm{Tr}[|\psi_k\rangle\langle\psi_k|M_1]}_{\leq 1\ \forall k}$$

$$\leq \sum_{k:\lambda_k\geq 0} \lambda_k + \sum_{k:\lambda_k<0} |\lambda_k|$$

$$= \sum_{k=1}^{d} |\lambda_k| = \|H\|_1 .$$

So $\mathrm{Tr}[H(M_0 - M_1)] \leq \|H\|_1$ for all pairs $M_0, M_1 \geq 0$, $M_0 + M_1 = \mathbb{1}$.

$\Rightarrow \max\{\mathrm{Tr}[H(M_0 - M_1)] : M_0, M_1 \geq 0,\ M_0 + M_1 = \mathbb{1}\} \leq \|H\|_1$

$\left.\begin{matrix} x \geq y \\ x \leq y \end{matrix}\right\} \Rightarrow x = y$

Combine with $\max\{\mathrm{Tr}[H(M_0 - M_1)] : M_0, M_1 \geq 0,\ M_0 + M_1 = \mathbb{1}\} \geq \|H\|_1$

to conclude equality. ■

Corollary: For Hermitian H, $\|H\|_1 = -\mathrm{Tr}[H] + 2 \cdot \max\{\mathrm{Tr}[HM] : 0 \leq M \leq \mathbb{1}\}$.

Proof: Start with $\|H\|_1 = \max\{\mathrm{Tr}[(M_0 - M_1)H] : M_0, M_1 \geq 0,\ M_0 + M_1 = \mathbb{1}\}$

Eliminate the variable M_1: $M_1 = \mathbb{1} - M_0$

$M_1 \geq 0 \Rightarrow \mathbb{1} - M_0 \geq 0 \Rightarrow M_0 \leq \mathbb{1}$.

$\mathrm{Tr}[(M_0 - M_1)H] = \mathrm{Tr}[(M_0 - (\mathbb{1} - M_0))H] = 2\mathrm{Tr}[M_0 H] - \mathrm{Tr}[H]$.

$M \leq \mathbb{1} \Rightarrow \underbrace{\mathbb{1} - M \geq 0}$

So $\|H\|_1 = \max\{\underbrace{-\mathrm{Tr}[H]}_{\text{does not depend on } M_0 \,\to\, \text{bring outside max.}} + 2\mathrm{Tr}(M_0 H) : 0 \leq M_0 \leq \mathbb{1}\}$

$= -\mathrm{Tr}[H] + 2 \cdot \max\{\mathrm{Tr}[M_0 H] : 0 \leq M_0 \leq \mathbb{1}\}$. ■

Theorem: $P_{succ}^{opt} := q_1 + \max_{0 \leq M \leq \mathbb{1}} \mathrm{Tr}[M(q_0\rho_0 - q_1\rho_1)] = \frac{1}{2}(1 + \|q_0\rho_0 - q_1\rho_1\|_1)$

Proof: From the corollary above,

$$\max_{0 \leq M \leq \mathbb{1}} \mathrm{Tr}[M(q_0\rho_0 - q_1\rho_1)] = \frac{1}{2}\Big(\|q_0\rho_0 - q_1\rho_1\|_1 + \underbrace{\mathrm{Tr}[q_0\rho_0 - q_1\rho_1]}_{= q_0\mathrm{Tr}[\rho_0] - q_1\mathrm{Tr}[\rho_1] = q_0 - q_1}\Big).$$

$$\Rightarrow p^{opt}_{succ} = q_1 + \tfrac{1}{2}\|q_0\rho_0 - q_1\rho_1\|_1 + \tfrac{1}{2}q_0 - \tfrac{1}{2}q_1$$

$$= \tfrac{1}{2}\big(\underbrace{q_0 + q_1}_{=1} + \|q_0\rho_0 - q_1\rho_1\|_1\big) = \tfrac{1}{2}\big(1 + \|q_0\rho_0 - q_1\rho_1\|_1\big). \blacksquare$$

(c) Remarks

- $P^{opt}_{succ} \geq \frac{1}{2}$ always — b/c we can always randomly guess one of the two states!

- If $\rho_0 = |\gamma\rangle\langle\gamma|$ and $\rho_1 = |\psi\rangle\langle\psi|$ (both are pure states), and also $q_0 = q_1 = \frac{1}{2}$, then

$$P^{opt}_{succ} = \frac{1}{2} + \frac{1}{4}\underbrace{\big\|\,|\gamma\rangle\langle\gamma| - |\psi\rangle\langle\psi|\,\big\|_1}_{=2\sqrt{1-|\langle\gamma|\psi\rangle|^2}} = \frac{1}{2} + \frac{1}{2}\sqrt{1-|\langle\gamma|\psi\rangle|^2}$$

- If $|\gamma\rangle$ and $|\psi\rangle$ are orthogonal, then $\langle\gamma|\psi\rangle = 0 \Rightarrow P^{opt}_{succ} = 1$

  ⊛ Orthogonal states can be perfectly distinguished!

- Note for later: $\big\|\,|\gamma\rangle - |\psi\rangle\,\big\|_2^2 = (\langle\gamma| - \langle\psi|)(|\gamma\rangle - |\psi\rangle)$

$\langle\gamma|\gamma\rangle = 1 = \langle\psi|\psi\rangle$

$$= 1 - \langle\gamma|\psi\rangle - \langle\psi|\gamma\rangle + 1$$

$$= 2 - 2\,\mathrm{Re}\big(\langle\gamma|\psi\rangle\big).$$

$z + \bar{z} = 2\,\mathrm{Re}(z),\ z \in \mathbb{C}.$

$z \in \mathbb{C}$, $|z|^2 = \mathrm{Re}(z)^2 + \mathrm{Im}(z)^2 \geq \mathrm{Re}(z)^2$

$|z| \geq |\mathrm{Re}(z)| \geq \mathrm{Re}(z).$

$$\mathrm{Re}\big(\langle\gamma|\psi\rangle\big) \leq |\langle\gamma|\psi\rangle| \Rightarrow \geq 2 - 2|\langle\gamma|\psi\rangle|$$

$$= 2\big(1 - |\langle\gamma|\psi\rangle|\big).$$

Also, $1 - |\langle\gamma|\psi\rangle|^2 = \big(1 - |\langle\gamma|\psi\rangle|\big)\big(1 + \underbrace{|\langle\gamma|\psi\rangle|}_{\leq 1}\big) \leq 2\big(1 - |\langle\gamma|\psi\rangle|\big) \leq \big\|\,|\gamma\rangle - |\psi\rangle\,\big\|_2^2$

$$\Rightarrow \frac{1}{2}\big\|\,|\gamma\rangle\langle\gamma| - |\psi\rangle\langle\psi|\,\big\|_1 = \sqrt{1 - |\langle\gamma|\psi\rangle|^2} \leq \big\|\,|\gamma\rangle - |\psi\rangle\,\big\|_2$$

③ Back to the proof of optimality of Grover's algorithm.

(a) Proof of Optimality: Step 1

- State after an arbitrary algorithm with one marked element, with k calls to the phase oracle:

$$|\psi_k^{\vec{x}}\rangle = U_k Z_{\vec{x}} U_{k-1} \cdots Z_{\vec{x}} U_1 Z_{\vec{x}} |\psi_0\rangle$$

- Analogous state with no marked element:

$$|\psi_k\rangle = U_k \cdots U_2 U_1 |\psi_0\rangle$$

- How well can we distinguish b/w $|\psi_k^{\vec{x}}\rangle$ and $|\psi_k\rangle$, as a function of k and $N=2^n$? What k do we need to distinguish them with high probability?

- Optimal success probability of distinguishing these is

$$p^{opt}_{succ} = \frac{1}{2} + \frac{1}{4} \left\| |\psi_k^{\vec{x}}\rangle\langle\psi_k^{\vec{x}}| - |\psi_k\rangle\langle\psi_k| \right\|_1 \leq \frac{1}{2} + \frac{1}{2} \left\| |\psi_k^{\vec{x}}\rangle - |\psi_k\rangle \right\|_2$$

- Uniform average over $\vec{x}$ (b/c we don't know what $\vec{x}$ is):

$$\overline{p^{opt}_{succ}} \equiv \mathbb{E}_{\vec{x}}\Big[\underbrace{\frac{1}{2} + \frac{1}{4} \left\| |\psi_k^{\vec{x}}\rangle\langle\psi_k^{\vec{x}}| - |\psi_k\rangle\langle\psi_k| \right\|_1}_{p^{opt,x}_{succ}} \Big] = \frac{1}{N} \sum_{\vec{x}\in\{0,1\}^n} \left(\frac{1}{2} + \frac{1}{4} \left\| |\psi_k^{\vec{x}}\rangle\langle\psi_k^{\vec{x}}| - |\psi_k\rangle\langle\psi_k| \right\|_1 \right)$$

$$\leq \frac{1}{2} + \frac{1}{2} \sum_{\vec{x}\in\{0,1\}^n} \frac{1}{N} \left\| |\psi_k^{\vec{x}}\rangle - |\psi_k\rangle \right\|_2 .$$

- We want a constant advantage over random guessing:

$$\mathbb{E}_{\vec{x}}\left[\frac{1}{2}+\frac{1}{4}\left\| |\gamma_k^{\vec{x}}\rangle\langle\gamma_k^{\vec{x}}| - |\psi_k\rangle\langle\psi_k| \right\|_1\right] \geq \frac{1}{2}+\delta, \quad \delta > 0.$$

$$\Rightarrow \quad \delta \leq \frac{1}{2}\sum_{\vec{x}\in\{0,1\}^n} \frac{1}{N} \left\| |\gamma_k^{\vec{x}}\rangle - |\psi_k\rangle \right\|_2$$

$|\gamma_k^{\vec{x}}\rangle = U_k Z_{\vec{x}} U_{k-1} \cdots Z_{\vec{x}} U_1 Z_{\vec{x}} |\gamma_0\rangle$

$|\psi_k\rangle = U_k \cdots U_2 U_1 |\gamma_0\rangle$

(b) Proof of Optimality: Step 2

- Observe that $|\gamma_k^{\vec{x}}\rangle = U_k Z_{\vec{x}} |\gamma_{k-1}^{\vec{x}}\rangle$ for all $k \in \{1,2,\ldots\}$

$|\psi_k\rangle = U_k |\psi_{k-1}\rangle$ for all $k \in \{1,2,\ldots\}$, $|\psi_0\rangle = |\gamma_0\rangle$.

Then: $\left\| |\gamma_k^{\vec{x}}\rangle - |\psi_k\rangle \right\|_2 = \left\| U_k Z_f |\gamma_{k-1}^{\vec{x}}\rangle - U_k |\psi_{k-1}\rangle \right\|_2$

$= \left\| Z_{\vec{x}} |\gamma_{k-1}^{\vec{x}}\rangle - |\psi_{k-1}\rangle \right\|_2$ (b/c U_k is unitary).

$\left\| U|v\rangle \right\|_2 = \left\| |v\rangle \right\|_2$ $\forall$ unitaries U

$= \left\| Z_{\vec{x}} |\gamma_{k-1}^{\vec{x}}\rangle - Z_{\vec{x}} |\psi_{k-1}\rangle + Z_{\vec{x}} |\psi_{k-1}\rangle - |\psi_{k-1}\rangle \right\|_2$

$\leq \left\| Z_{\vec{x}} |\gamma_{k-1}^{\vec{x}}\rangle - Z_{\vec{x}} |\psi_{k-1}\rangle \right\|_2 + \left\| Z_{\vec{x}} |\psi_{k-1}\rangle - |\psi_{k-1}\rangle \right\|_2$

(triangle inequality).

$Z_{\vec{x}} = \mathbb{1} - 2|\vec{x}\rangle\langle\vec{x}|$

$= \left\| |\gamma_{k-1}^{\vec{x}}\rangle - |\psi_{k-1}\rangle \right\|_2 + \left\| \underbrace{(Z_{\vec{x}} - \mathbb{1})}_{= \mathbb{1} - 2|\vec{x}\rangle\langle\vec{x}| - \mathbb{1} = -2|\vec{x}\rangle\langle\vec{x}|} |\psi_{k-1}\rangle \right\|_2$

$= \left\| |\gamma_{k-1}^{\vec{x}}\rangle - |\psi_{k-1}\rangle \right\|_2 + \underbrace{\left\| (-2|\vec{x}\rangle\langle\vec{x}|) |\psi_{k-1}\rangle \right\|_2}_{= 2|\langle\vec{x}|\psi_{k-1}\rangle| \cdot \left\| |\vec{x}\rangle \right\|_2 = 2|\langle\vec{x}|\psi_{k-1}\rangle|.}$

$$\Rightarrow \| |\psi_k^{\vec{x}}\rangle - |\varphi_k\rangle \|_2 \le \| |\psi_{k-1}^{\vec{x}}\rangle - |\varphi_{k-1}\rangle \|_2 + 2\,|\langle \vec{x}|\varphi_{k-1}\rangle|.$$

This holds for all $k \in \{1,2,\dots\}$

- Apply the analogous steps for $\| |\psi_{k-1}\rangle - |\varphi_{k-1}\rangle\|_2$, and do it recursively:

$$\begin{aligned}
\| |\psi_k^{\vec{x}}\rangle - |\varphi_k\rangle \|_2 &\le \| |\psi_{k-1}^{\vec{x}}\rangle - |\varphi_{k-1}\rangle \|_2 + 2\,|\langle \vec{x}|\varphi_{k-1}\rangle| \\
&\le \| |\psi_{k-2}^{\vec{x}}\rangle - |\varphi_{k-2}\rangle \|_2 + 2\,|\langle \vec{x}|\varphi_{k-2}\rangle| + 2\,|\langle \vec{x}|\varphi_{k-1}\rangle| \\
&\le \dots \\
&\le \| |\psi_1^{\vec{x}}\rangle - |\varphi_1\rangle \|_2 + 2\sum_{\ell=1}^{k-1} |\langle \vec{x}|\varphi_\ell\rangle| \\
&\le \underbrace{\| |\psi_0\rangle - |\varphi_0\rangle \|_2}_{=0.} + 2\sum_{\ell=0}^{k-1} |\langle \vec{x}|\varphi_\ell\rangle|
\end{aligned}$$

$$\Rightarrow \| |\psi_k^{\vec{x}}\rangle - |\varphi_k\rangle \|_2 \le 2\sum_{\ell=0}^{k-1} |\langle \vec{x}|\varphi_\ell\rangle|$$

- Now do the uniform average over $\vec{x}$:

$$\mathbb{E}_{\vec{x}}\left[\| |\psi_k^{\vec{x}}\rangle - |\varphi_k\rangle \|_2\right] \le 2\sum_{\ell=0}^{k-1} \mathbb{E}_{\vec{x}}\left[|\langle \vec{x}|\varphi_\ell\rangle|\right]$$

This depends on $\vec{x}$.

$$= \frac{1}{N}\sum_{\vec{x}\in\{0,1\}^n} |\langle \vec{x}|\varphi_\ell\rangle|$$

Apply the Cauchy-Schwarz inequality!

$$|\langle u|v\rangle| \le \sqrt{\langle u|u\rangle\langle v|v\rangle}$$

$$|u\rangle = \sum_{\vec{x}\in\{0,1\}^n} |\vec{x}\rangle \quad \text{(all-ones vector)}$$

$$|v\rangle = \sum_{\vec{x}\in\{0,1\}^n} |\langle \vec{x}|\varphi_\ell\rangle|\,|\vec{x}\rangle.$$

$$\le \frac{1}{N}\sqrt{\underbrace{\langle u|u\rangle}_{N}\,\langle v|v\rangle}$$

$$\langle v|v\rangle = \sum_{\vec{x}\in\{0,1\}^n} |\langle \vec{x}|\varphi_\ell\rangle|^2 = 1$$

Note: $\langle u|v\rangle = \sum_{\vec{x}\in\{0,1\}^n} |\langle \vec{x}|\varphi_\ell\rangle| = |\langle u|v\rangle|$

$$= \frac{1}{N}\sqrt{N} = \frac{1}{\sqrt{N}}.$$

$$\Rightarrow \mathbb{E}_{\vec{x}}\left[\left\| |\psi_k^{\vec{x}}\rangle - |\psi_k\rangle \right\|_2\right] \leq 2\sum_{\ell=0}^{k-1} \underbrace{\mathbb{E}_{\vec{x}}\left[|\langle \vec{x}|\psi_\ell\rangle|\right]}_{\leq \frac{1}{\sqrt{N}} \ \forall \ell} \leq \frac{2k}{\sqrt{N}}.$$

(c) Put everything together: $\overline{P_{succ}^{opt}} \geq \frac{1}{2} + \delta, \ \delta > 0$

$$\Rightarrow \quad \delta \leq \frac{1}{2}\mathbb{E}_{\vec{x}}\left[\left\| |\psi_k^{\vec{x}}\rangle - |\psi_k\rangle \right\|_2\right] \leq \frac{k}{\sqrt{N}}$$

$$\Rightarrow \underline{\underline{k \geq \delta \cdot \sqrt{N}}}$$

(*) B/c we want a constant δ (i.e., independent of N), we have the desired lower bound.

(*) Proving lower bounds for quantum algorithms often involves tools from quantum information, as we demonstrated here by using state discrimination/hypothesis testing.