Introduction to Quantum Computing

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 500

1. Grover's Algorithm and Generalizations
   - Grover's Algorithm
   - Amplitude Amplification
   - Optimality of Grover's Algorithm
   - Derandomization and Amplitude Amplification
   - Unknown Number of Solutions

Subsection 1

Grover's Algorithm

The Problem

- Grover's algorithm uses **amplitude amplification** to search an unstructured set of N elements.
- Suppose the property being searched for is given in terms of a Boolean function, or predicate,

$$P : \{0, \ldots, N-1\} \to \{0, 1\}.$$

- The goal of the problem is to find a solution.
- That is, identify an element x, such that

$$P(x) = 1.$$

The Classical Complexity

- As in Simon's problem and the Deutsch-Jozsa problem, the predicate P is viewed as an oracle or black box.
- So our focus is on the query complexity, the number of calls made to the oracle P.
- Given a black box that outputs $P(x)$ upon input of x, the best classical approaches must, in the single solution case, inspect an average of $\frac{N}{2}$ values.
- That is, the classical approach requires an average of $\frac{N}{2}$ evaluations of the predicate $P(x)$.

The Quantum Complexity

- Suppose, we are given a quantum black box U_P that sends

$$\sum_x c_x|x\rangle|0\rangle \to \sum_x c_x|x\rangle|P(x)\rangle.$$

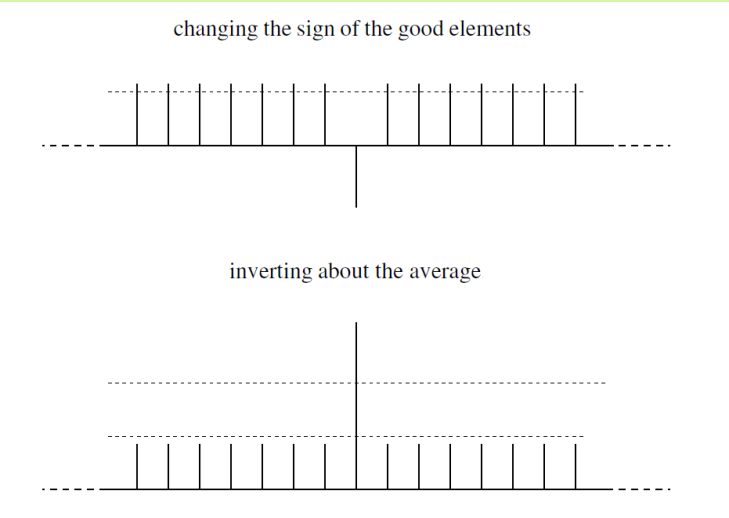
- Grover's algorithm finds a solution, in the single solution case, with only $O(\sqrt{N})$ calls to U_P.
- Grover's algorithm works by iteratively increasing the amplitudes c_x of those values x with $P(x) = 1$.
- As a result, a final measurement will return a value x of interest with high probability.
- For practical applications of Grover's algorithm, the predicate P:
  - Must be efficiently computable;
  - Should lack such structure as allows classical methods to gain advantage over the quantum algorithm.

Outline

- Grover's algorithm starts with an equal superposition of all N values of the search space,

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_x |x\rangle.$$

- It repeatedly performs the same sequence of transformations:
  1. Apply U_P to $|\psi\rangle$.
  2. Flip the sign of all basis vectors that represent a solution.
  3. Perform inversion about the average, a transformation that maps every amplitude $A - \delta$ to $A + \delta$, where A is the average of the amplitudes.



Setup

- Without loss of generality, let $N = 2^n$ for some integer n.
- Let X be the state space generated by $\{|0\rangle, \ldots, |N-1\rangle\}$.
- Let U_P be a quantum black box that acts as

$$U_P : |x, a\rangle \to |x, P(x) \oplus a\rangle,$$

  for all $x \in X$ and all single-qubit states $|a\rangle$.
- Denote the sets of good and bad values, respectively, by

$$G = \{x : P(x)\} \quad \text{and} \quad B = \{x : \neg P(x)\}.$$

- Let the number of good states be a small fraction of the total number of states, written

$$|G| \ll N.$$

Setup (Cont'd)

- Consider the even superpositions:
  - Of all good states,
$$|\psi_G\rangle = \frac{1}{\sqrt{|G|}} \sum_{x \in G} |x\rangle;$$
  - Of all bad states,
$$|\psi_B\rangle = \frac{1}{\sqrt{|B|}} \sum_{x \in B} |x\rangle.$$
- Then $|\psi\rangle = W|0\rangle$, an equal superposition of all N values, can be written as a superposition of $|\psi_G\rangle$ and $|\psi_B\rangle$

$$|\psi\rangle = \frac{1}{\sqrt{2^n}} \sum_{x=0}^{2^n-1} |x\rangle = g_0|\psi_G\rangle + b_0|\psi_B\rangle,$$

where $g_0 = \sqrt{\frac{|G|}{N}}$ and $b_0 = \sqrt{\frac{|B|}{N}}$.

Setup (Cont'd)

- The core of Grover's algorithm is the repeated application of a unitary transformation

$$Q : g_i|\psi_G\rangle + b_i|\psi_B\rangle \rightarrow g_{i+1}|\psi_G\rangle + b_{i+1}|\psi_B\rangle$$

  that increases the amplitude g_i of good states (and decreases b_i).
- This is done until a maximal value is reached.
- After applying Q an appropriate number of times j, almost all amplitude will have shifted to good states, so that $|b_j| \ll |g_j|$.
- At this point, measurement will return an $x \in G$ with high probability.
- The exact number of times Q needs to be applied is on the order of $\sqrt{N}$ and depends on both N and $|G|$.

Iteration Step: Changing the Sign of the Good Elements

- To change the sign in a superposition $\sum c_x|x\rangle$ of exactly those $|x\rangle$ such that $x \in G$, apply S_G^π.
- A sign change is simply a phase shift by $e^{\boldsymbol{i}\pi} = -1$.
- We showed that

$$U_P(|\psi\rangle \otimes H|1\rangle) = (S_G^\pi|\psi\rangle) \otimes H|1\rangle.$$

- Changing the sign of the good elements is accomplished by

$$U_P : (g_i|\psi_G\rangle + b_i|\psi_B\rangle) \otimes H|1\rangle \to (-g_i|\psi_G\rangle + b_i|\psi_B\rangle) \otimes H|1\rangle.$$

- The number of gates needed to change the sign on the good elements does not depend on N, but rather on how many gates it takes to compute U_P.

Iteration Step: Inversion About the Average

- Let A be the average of the amplitudes of all basis vectors in the superposition.
- Inversion about the average sends

$$a|x\rangle \to (2A - a)|x\rangle.$$

- The transformation

$$\sum_{i=0}^{N-1} a_i|x_i\rangle \to \sum_{i=0}^{N-1} (2A - a_i)|x_i\rangle$$

is performed by the unitary matrix

$$D = \begin{pmatrix} \frac{2}{N} - 1 & \frac{2}{N} & \cdots & \frac{2}{N} \\ \frac{2}{N} & \frac{2}{N} - 1 & \cdots & \frac{2}{N} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{2}{N} & \frac{2}{N} & \cdots & \frac{2}{N} \end{pmatrix}.$$

Iteration Step: Inversion About the Average (Cont'd)

- We implement D with $O(n) = O(\log_2(N))$ quantum gates.
- Following Grover, we define

$$D = -WS_0^{\pi}W,$$

where:

- W is the Walsh-Hadamard transform;
- S_0^{π} is the phase shift by π of the basis vector $|0\rangle$,

$$S_0^{\pi} = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots \\ 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}.$$

- We now show that this $D = -WS_0^{\pi}W$ is the one we need.

Iteration Step: Inversion About the Average (Cont'd)

- Let

$$R = \begin{pmatrix} 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots \\ 0 & \cdots & \cdots & 0 \\ 0 & \cdots & 0 & 0 \end{pmatrix}.$$

- We have $S_0^\pi = I - R$.
- So we get

$$-WS_0^\pi W = W(R - I)W = WRW - I.$$

- Bur $R_{ij} = 0$, for $i \neq 0$ or $j \neq 0$.
- Thus,

$$(WRW)_{ij} = W_{i0} R_{00} W_{0j} = \frac{2}{N}.$$

Iteration Step: Inversion About the Average (Cont'd)

- We got

$$-WS_0^{\pi}W \quad \text{and} \quad (WRW)_{ij} = \frac{2}{N}.$$

- So we obtain

$$-WS_0^{\pi}W = WRW - I = D.$$

- We finally put together:
  - Inversion about the average;
  - Changing the sign of the good elements.
- This yields the iteration transformation

$$Q = -WS_0^{\pi}WS_G^{\pi}.$$

Number of Iterations: Intuition

- We examine the result of multiple application of the iteration step Q.
- The goal is to determine the optimal number of times to apply Q.
- We show that:
  - Q is a fixed rotation;
  - The amplitude g_i of good states varies periodically with the number of iterations.
- To find a solution with high probability, the number of iterations i must be chosen carefully.
- To determine the correct number of iterations to use, we describe the result of applying Q in terms of recurrence relations on g_i and b_i.

Number of Iterations: Formalism

- The iteration step $Q = DS_G^\pi$ transforms

$$g_i|\psi_G\rangle + b_i|\psi_B\rangle \to g_{i+1}|\psi_G\rangle + b_{i+1}|\psi_B\rangle.$$

- First,

$$S_G^\pi : g_i|\psi_G\rangle + b_i|\psi_B\rangle \to -g_i|\psi_G\rangle + b_i|\psi_B\rangle.$$

- To compute the average amplitude, A_i, note that:
  - The term $-g_i|\psi_G\rangle$ contributes $|G|$ amplitudes $\frac{-g_i}{\sqrt{|G|}}$;
  - The term $b_i|\psi_B\rangle$ contributes $|B|$ amplitudes $\frac{b_i}{\sqrt{|B|}}$.
- Thus, altogether

$$A_i = \frac{\sqrt{|B|}b_i - \sqrt{|G|}g_i}{N}.$$

Number of Iterations: Inversion About the Average

- Next, we turn to inversion about the average transforms,

$$
\begin{aligned}
D : -g_i|\psi_G\rangle + b_i|\psi_B\rangle & \\
& \to \sum_{x\in G}\left(2A_i + \frac{g_i}{\sqrt{|G|}}\right)|x\rangle + \sum_{x\in B}\left(2A_i - \frac{b_i}{\sqrt{|B|}}\right)|x\rangle \\
& = (2A_i\sqrt{|G|} + g_i)|\psi_G\rangle + (2A_i\sqrt{|B|} - b_i)|\psi_B\rangle \\
& = g_{i+1}|\psi_G\rangle + b_{i+1}|\psi_B\rangle,
\end{aligned}
$$

where

$$
\begin{aligned}
g_{i+1} &= 2A_i\sqrt{|G|} + g_i, \\
b_{i+1} &= 2A_i\sqrt{|B|} - b_i.
\end{aligned}
$$

Number of Iterations: Solving the Recurrence Relations

- Let t denote the probability that a random value in $\{0, \ldots, N-1\}$ satisfies P.
- Then we have

$$t = \frac{|G|}{N} \quad \text{and} \quad 1 - t = \frac{|B|}{N}.$$

- Now we get

$$\begin{aligned} A_i\sqrt{|G|} &= \frac{\sqrt{|B||G|}b_i - |G|g_i}{N} = \sqrt{t(1-t)}b_i - tg_i; \\ A_i\sqrt{|B|} &= \frac{|B|b_i - \sqrt{|B||G|}g_i}{N} = (1-t)b_i - \sqrt{t(1-t)}g_i. \end{aligned}$$

- So, for the recurrence relations, we have:

$$\begin{aligned} g_{i+1} &= 2A_i\sqrt{|G|} + g_i \\ &= 2(\sqrt{t(1-t)}b_i - tg_i) + g_i \\ &= (1-2t)g_i + 2\sqrt{t(1-t)}b_i. \end{aligned}$$

Number of Iterations: Solving the Recurrences (Cont'd)

- Similarly,

$$\begin{aligned} b_{i+1} &= 2A_i\sqrt{|B|} - b_i \\ &= 2((1-t)b_i - \sqrt{t(1-t)}g_i) - b_i \\ &= (1-2t)b_i - 2\sqrt{t(1-t)}g_i. \end{aligned}$$

- We also have

$$g_0 = \sqrt{t} \quad \text{and} \quad b_0 = \sqrt{1-t}.$$

- We can verify that

$$g_i = \sin((2i+1)\theta), \quad b_i = \cos((2i+1)\theta)$$

is a solution to these equations with $\sin\theta = \sqrt{t} = \sqrt{\frac{|G|}{N}}$.

Number of Iterations: Computing the Optimum

- We are now ready to compute the optimum number of iterations of Q.
- We wish to find an element with the desired property P.
- This calls for maximizing the probability of measuring a good state.
- So we wish to choose i, such that

$$\sin\left((2i+1)\theta\right) \approx 1 \text{ or } (2i+1)\theta \approx \frac{\pi}{2}.$$

- For $|G| \ll N$, the angle θ becomes very small.
- So $\sqrt{\frac{|G|}{N}} = \sin\theta \approx \theta$.
- Thus, g_i will be maximal for

$$\begin{aligned} (2i+1)\sqrt{\tfrac{|G|}{N}} \approx \tfrac{\pi}{2} \quad &\Rightarrow \quad 2i+1 \approx \tfrac{\pi}{2}\sqrt{\tfrac{N}{|G|}} \\ &\Rightarrow \quad i \approx \tfrac{\pi}{4}\sqrt{\tfrac{N}{|G|}} - \tfrac{1}{2} \\ &\Rightarrow \quad i \approx \tfrac{\pi}{4}\sqrt{\tfrac{N}{|G|}}. \end{aligned}$$

Number of Iterations: Computing the Optimum (Cont'd)

- Additional iteration will reduce the success probability.
- This situation is in contrast to many classical algorithms in which the greater the number of iterations the better the results.
- Using the equations for g_i and b_i:
  - For $t = \frac{1}{4}$, the optimum number of iterations is 1.
    Indeed, we have

$$\begin{aligned} \sin\theta = \sqrt{t} = \tfrac{1}{2} \quad &\Rightarrow \quad \theta = \tfrac{\pi}{6}; \\ g_i = \sin\left((2i+1)\tfrac{\pi}{6}\right) \quad &\Rightarrow \quad i = 1. \end{aligned}$$

  - For $t = \frac{1}{2}$, no amount of iteration will improve the situation.
    Indeed, we have

$$\begin{aligned} \sin\theta = \sqrt{t} = \tfrac{\sqrt{2}}{2} \quad &\Rightarrow \quad \theta = \tfrac{\pi}{4}; \\ g_i = \sin\left((2i+1)\tfrac{\pi}{4}\right) \quad &\Rightarrow \quad i = 0. \end{aligned}$$

Revisiting the Geometric Interpretation

- Every step of the iteration process has been written as a linear combination of $|\psi_G\rangle$ and $|\psi_B\rangle$ with real coefficients.
- So Grover's algorithm can be viewed as acting in the real two-dimensional subspace spanned by $|\psi_G\rangle$ and $|\psi_B\rangle$.
- The algorithm simply shifts amplitude from $|\psi_B\rangle$ to $|\psi_G\rangle$.
- This picture leads to an elegant geometric interpretation of Grover's algorithm to be discussed shortly.
- First, we describe a generalization of Grover's algorithm, *amplitude amplification*, to which this geometric picture also applies.

Subsection 2

Amplitude Amplification

Generalizing Amplitude Amplification

- The first step of Grover's algorithm applies the iteration operator

$$Q = -WS_0^{\pi}WS_G^{\pi}$$

to the initial state $W|0\rangle$.

- W can be viewed as a trivial algorithm mapping $|0\rangle$ to all possible values.
- So it maps $|0\rangle$ to a solution with probability $\frac{|G|}{N}$.
- Suppose we have an algorithm U, such that $U|0\rangle$ gives an initial solution with a higher probability.
- We show that the previous analysis generalizes to any algorithm U, such that $U|0\rangle$ has some amplitude in the good states.
- *Amplitude amplification* generalizes Grover's algorithm by replacing the iteration operator $Q = -WS_0^{\pi}WS_G^{\pi}$ with

$$Q = -US_0^{\pi}U^{-1}S_G^{\pi}.$$

Normalized Projections

- Let $\mathcal{G}$ be the subspace spanned by $\{|x\rangle : x \in G\}$.
- Let $\mathcal{B}$ be the subspace spanned by $\{|x\rangle : x \notin G\}$.
- Let $P_{\mathcal{G}}$ and $P_{\mathcal{B}}$ be the associated projection operators.
- Let $|\psi_G\rangle$ be the normalized projection of $|\psi\rangle$ onto the good subspace,

$$|\psi_G\rangle = \frac{1}{g_0} P_{\mathcal{G}}|\psi\rangle, \quad g_0 = |P_{\mathcal{G}}|\psi\rangle|$$

- Let $|\psi_B\rangle$ be the normalized projection of $|\psi\rangle$ onto the bad subspace,

$$|\psi_B\rangle = \frac{1}{b_0} P_{\mathcal{B}}|\psi\rangle, \quad b_0 = |P_{\mathcal{B}}|\psi\rangle|.$$

- Let $|\psi\rangle = U|0\rangle$ be written as

$$|\psi\rangle = g_0|\psi_G\rangle + b_0|\psi_B\rangle.$$

Measurement and Probabilities

- For $U = W$, we take $|\psi_G\rangle$, $|\psi_B\rangle$, g_0 and b_0 are as before.
- Here g_0 and b_0 are not determined by the number of solutions, but rather by the properties of U relative to the good states.
- The states $|\psi_G\rangle$ and $|\psi_B\rangle$ need not be equal superpositions of the good and bad states respectively, but g_0 and b_0 are still real.
- Again, we let

$$t = g_0^2, \quad \text{with} \quad 1 - t = b_0^2,$$

  where t should be thought of as the probability that measurement of the superposition $U|0\rangle$ yields a state that satisfies predicate P.
- The operator U can be viewed as a reversible algorithm that maps $|0\rangle$ to a set of solutions in G with a probability $t = |g_0|^2$.

The Effect of Applying Q

- To understand the effect of $Q = -US_0^\pi U^{-1} S_G^\pi$, recall that $S_0^\pi|\varphi\rangle$ can be written as
$$S_0^\pi|\varphi\rangle = |\varphi\rangle - 2\langle 0|\varphi\rangle|0\rangle.$$
- For an arbitrary state $|\psi\rangle$,
$$\begin{aligned} US_0^\pi U^{-1}|\psi\rangle &= U(U^{-1}|\psi\rangle - 2\langle 0|U^{-1}|\psi\rangle|0\rangle) \\ &= |\psi\rangle - 2\langle 0|U^{-1}|\psi\rangle U|0\rangle \\ &= |\psi\rangle - 2\overline{\langle\psi|U|0\rangle}\, U|0\rangle. \end{aligned}$$

The Effect of Applying Q (Cont'd)

- We got

$$US_0^\pi U^{-1}|\psi\rangle = |\psi\rangle - 2\overline{\langle\psi|U|0\rangle}\,U|0\rangle.$$

- Now recall that

$$S_G^\pi|\psi_G\rangle = -|\psi_G\rangle \quad \text{and} \quad S_G^\pi|\psi_B\rangle = |\psi_B\rangle.$$

- So we get

$$\begin{aligned} Q|\psi_G\rangle &= -US_0^\pi U^{-1}S_G^\pi|\psi_G\rangle \\ &= US_0^\pi U^{-1}|\psi_G\rangle \\ &= |\psi_G\rangle - 2\overline{g_0}\,U|0\rangle \\ &= |\psi_G\rangle - 2\overline{g_0}g_0|\psi_G\rangle - 2\overline{g_0}b_0|\psi_B\rangle \\ &= (1-2t)|\psi_G\rangle - 2\sqrt{t(1-t)}|\psi_B\rangle. \end{aligned}$$

The Effect of Applying Q (Cont'd)

- Similarly, we have

$$\begin{aligned}
Q|\psi_B\rangle &= -US_0^\pi U^{-1} S_G^\pi |\psi_B\rangle \\
&= -US_0^\pi U^{-1}|\psi_B\rangle \\
&= -|\psi_B\rangle + 2\overline{\langle\psi_B|U|0\rangle}\, U|0\rangle \\
&= -|\psi_B\rangle + 2\overline{b_0}\, U|0\rangle \\
&= -|\psi_B\rangle + 2\overline{b_0} g_0|\psi_G\rangle + 2\overline{b_0} b_0|\psi_B\rangle \\
&= -|\psi_B\rangle + 2(1-t)\tfrac{g_0}{b_0}|\psi_G\rangle + 2(1-t)|\psi_B\rangle \\
&= (1-2t)|\psi_B\rangle + 2\sqrt{t(1-t)}|\psi_G\rangle.
\end{aligned}$$

The Effect of Applying Q (Cont'd)

- We obtained

$$\begin{aligned} Q|\psi_G\rangle &= (1-2t)|\psi_G\rangle - 2\sqrt{t(1-t)}|\psi_B\rangle; \\ Q|\psi_B\rangle &= (1-2t)|\psi_B\rangle + 2\sqrt{t(1-t)}|\psi_G\rangle. \end{aligned}$$

- An arbitrary real superposition of $|\psi_G\rangle$ and $|\psi_B\rangle$ is transformed by Q as follows:

$$\begin{aligned} Q(g_i|\psi_G\rangle + b_i|\psi_B\rangle) &= (g_i(1-2t) + 2b_i\sqrt{t(1-t)})|\psi_G\rangle \\ &\quad +(b_i(1-2t) - 2g_i\sqrt{t(1-t)})|\psi_B\rangle. \end{aligned}$$

- This leads to the same recurrence relation as in the previous section,

$$\begin{aligned} g_{i+1} &= (1-2t)g_i + 2\sqrt{t(1-t)}b_i; \\ b_{i+1} &= (1-2t)b_i - 2\sqrt{t(1-t)}g_i. \end{aligned}$$

- It has the solution

$$g_i = \sin((2i+1)\theta), \quad b_i = \cos((2i+1)\theta), \quad \sin\theta = \sqrt{t} = g_0.$$

Number of Iterations

- Thus, for small g_0, the amplitude g_i will be maximal after

$$i \approx \frac{\pi}{4}\frac{1}{g_0}$$

  iterations.
- If the algorithm U succeeds with probability t, then simple classical repetition of U requires an average of $\frac{1}{t}$ iterations to find a solution.
- Amplitude amplification speeds up this process so that it takes only $O\left(\sqrt{\frac{1}{t}}\right)$ tries to find a solution.

Comments

- If U has no amplitude in the good states, g_0 will be zero and amplitude amplification will have no effect.
- Recall that no amount of iteration in Grover's algorithm improves the probability if $t = \frac{1}{2}$.
- Similarly, if g_0 is large, amplitude amplification cannot improve the situation.
- For this reason, amplitude amplification applied to an algorithm U that is the result of amplitude amplification does not improve the results.

Geometry of Amplitude Amplification

- Let $|\psi_G\rangle$, $|\psi_B\rangle$ and $Q = -US_0^{\pi}U^{-1}S_G^{\pi}$ be as defined before.
- We show that the entire discussion of amplitude amplification, and Grover's algorithm in particular, reduces to a simple geometric argument about rotations in the two-dimensional real subspace generated by $\{|\psi_G\rangle, |\psi_B\rangle\}$.

Geometry (Cont'd)

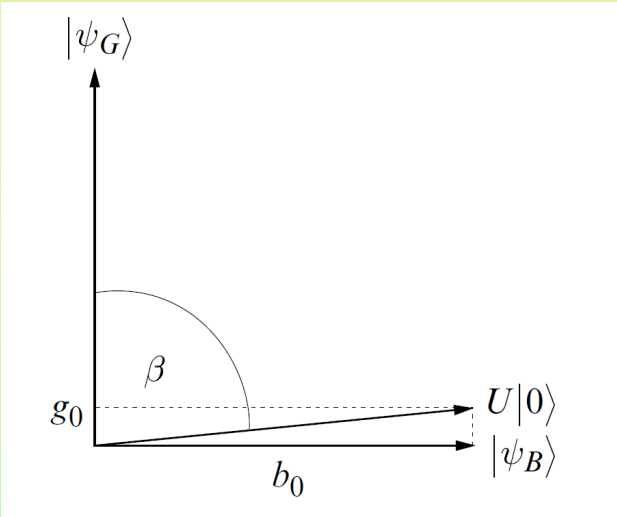
- By the definition of $|\psi_G\rangle$ and $|\psi_B\rangle$, the initial state

$$U|0\rangle = g_0|\psi_G\rangle + b_0|\psi_B\rangle$$

  has real amplitudes g_0 and b_0.
- So it is in the two-dimensional real plane spanned by $\{|\psi_G\rangle, |\psi_B\rangle\}$.
- The smaller the success probability t, the closer $U|0\rangle$ is to $|\psi_B\rangle$.

- Let β be the angle between $U|0\rangle$ and $|\psi_G\rangle$.
- The angle β depends only on the probability $t = g_0^2$ that the initial state $U|0\rangle$, if measured, gives a solution

$$\cos(\beta) = \langle\psi_G|U|0\rangle = g_0.$$



Geometry (The Goal)

- The rest of this section explains how each iteration of Grover's algorithm rotates the state by a fixed angle in the direction of the desired state.
- To maximize the amplitude in the good states, we iterate until the state is close to $|\psi_G\rangle$.
- From the simple geometry of the situation, we can determine:
  - The optimal number of iterations;
  - The probability that the run succeeds.

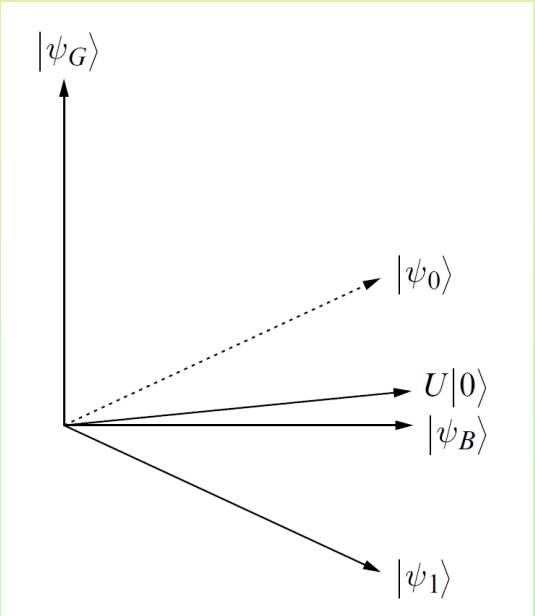
Geometry (Reflection)

- Amplitude amplification, and Grover's algorithm as the special case when $U = W$, consists of repeated applications of

$$Q = -US_0^{\pi}U^{-1}S_G^{\pi}.$$

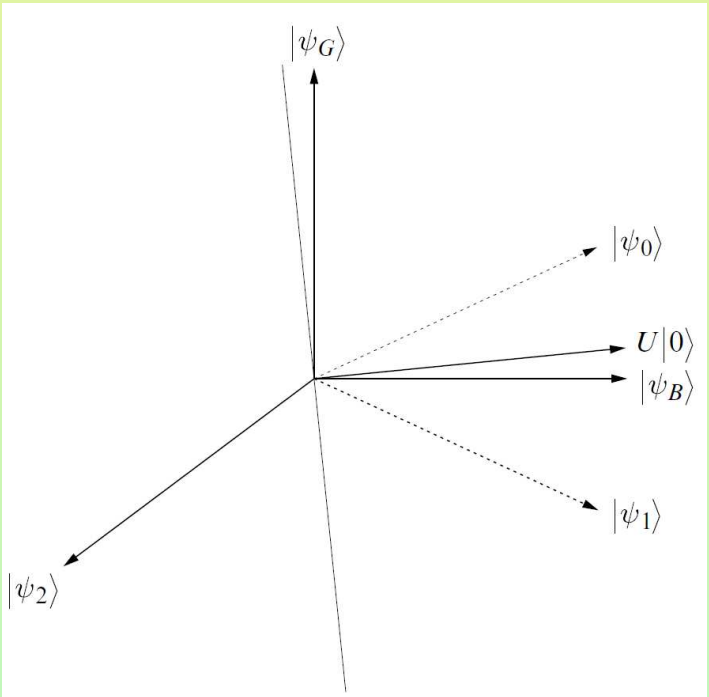
- Recall that the transformation S_G^{π} can be viewed as a reflection about the hyperplane perpendicular to $|\psi_G\rangle$.
- In the plane spanned by $\{|\psi_G\rangle, |\psi_B\rangle\}$, this hyperplane reduces to the one-dimensional space spanned by $|\psi_B\rangle$.
- In the figure S_G^{π} maps an arbitrary state $|\psi_0\rangle$ in the $\{|\psi_G\rangle, |\psi_B\rangle\}$ subspace to

$$|\psi_1\rangle = S_G^{\pi}|\psi_0\rangle.$$



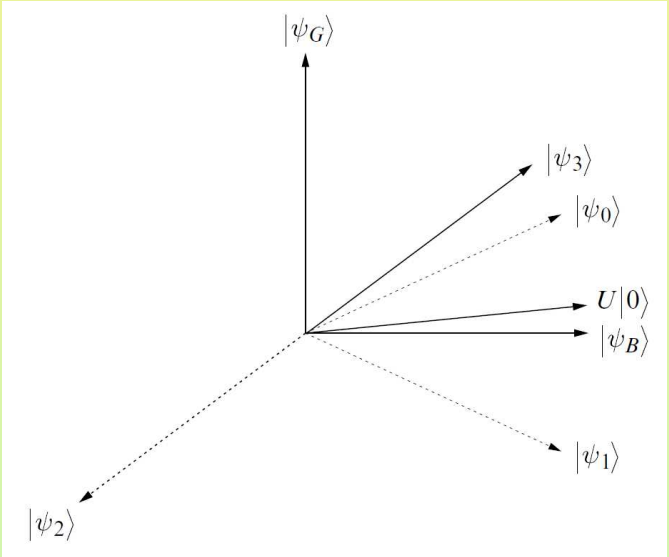
Geometry (Another Reflection)

- Similarly, the transformation S_0^π is a reflection about the hyperplane orthogonal to $|0\rangle$.
- Since $US_0^\pi U^{-1}$ differs from S_0^π by a change of basis, it is a reflection about the hyperplane orthogonal to $U|0\rangle$.
- The effect of this transformation on $|\psi_1\rangle$ is shown below:



Geometry (The Negative Sign)

- The final negative sign reverses the direction of the state vector.
- Strictly speaking, this negative sign is unnecessary, since it does nothing to the quantum state.
- It is a global phase change, so it is physically irrelevant.



- However, since we are drawing our pictures in the plane, not in projective space, the negative sign makes it easier to see what is going on.

Geometry (Rotation)

- The concatenation of two reflections is a rotation of twice the angle between the axes of the two reflections.
- The two axes of reflection in this case are perpendicular to $U|0\rangle$ and $|\psi_G\rangle$ respectively.
- So the angle between the axes of reflection is $-\beta$ where $\cos\beta = g_0$.
- The two reflections perform a rotation by -2β.
- The final negation amounts to a rotation by π.
- Thus, each step Q performs a rotation by $\pi - 2\beta$.

Geometry (Conclusion)

- Each step Q performs a rotation by $\pi - 2\beta$.
- Let $\theta = \frac{\pi}{2} - \beta$, the angle between $U|0\rangle$ and $|\psi_B\rangle$.
- So $\sin\theta = g_0$.
- Each iteration of Q rotates the state by 2θ.
- So the angle after i steps is $(2i+1)\theta$.
- As before, the amplitude in the good states after i steps is given by

$$g_i = \sin\left((2i+1)\theta\right).$$

- We solve for the optimal number of iterations just as we did before.

Subsection 3

Optimality of Grover's Algorithm

Optimality

- Even before Grover discovered his algorithm, researchers had proved a lower bound on the query complexity of any possible quantum algorithm for exhaustive search.
- It turns out that no quantum algorithm can use fewer than $\Omega(\sqrt{N})$ calls to the predicate U_P.
- Thus, Grover's algorithm is optimal.

Speedup

- The exponential size of the quantum state space gives naive hope that quantum computers could provide an exponential speedup for all computations.
- A less naive guess would be that quantum computers can provide exponential speedup for any computation that:
  - Can be parallelized;
  - Requires only a single answer output.
- The optimality of Grover's algorithm shows that even that hope is too optimistic.
- Exhaustive search is easily parallelized and requires a single answer.
- But quantum computers can provide only a relatively small speedup.

Role of S_x^π

- We showed how S_x^π can be computed from U_P.
- We use S_x^π as the interface to the oracle.
- We do not lose any generality in doing so.
  - The process of computing S_x^π from U_P is reversible;
  - So any algorithm using S_x^π could be rewritten in terms of U_P and vice versa.
- The oracle U_P provides us with the only way to access any information about the element x we are searching for.

Role of S_x^π (Cont'd)

- It follows that an arbitrary quantum search algorithm can be viewed as an algorithm that alternates between:
  - Unitary transformations independent of x;
  - Calls to S_x^π.
- That is, any quantum search algorithm can be written as

$$|\psi_k^x\rangle = U_k S_x^\pi U_{k-1} S_x^\pi \cdots U_1 S_x^\pi U_0 |0\rangle,$$

  where the U_i are unitary transformations that do not depend on x.
- The argument does not change if we allow the use of additional qubits.
- We simply use $I \otimes S_x^\pi$ instead of S_x^π.
- Moreover, as N is now larger, the algorithm will be less efficient.

Independence from x

- It is important to recognize that the algorithm must work no matter which x is the solution.
  - For any particular x, there are transformations that find x very quickly.
  - We want an algorithm that finds x quickly no matter what x is.
- Any search algorithm worth the name must return x with reasonable probability, for all possible values of x.
- We consider only quantum search algorithms that return x with at least probability $p = \frac{1}{2}$.
- It is easy for the reader to check that any value $0 < p < 1$ results in a $O(\sqrt{N})$ bound, just with a different constant.
- We will show that if the state $|\psi_k^x\rangle$, obtained after k steps of the form $U_i S_x^\pi$, satisfies $|\langle x|\psi_k^x\rangle|^2 \geq \frac{1}{2}$, for all x, then k must be $\Omega(\sqrt{N})$.

Intuition Behind the Proof

- We require hat the algorithm work for any x.
- So, if the oracle interface is S_x^π, then the result of applying

$$U_k S_x^\pi U_{k-1} S_x^\pi \cdots U_1 S_x^\pi U_0 |0\rangle$$

must be a state $|\psi_k^x\rangle$ sufficiently close to $|x\rangle$ so that x will be obtained upon measurement with high probability.

- Note that two elements of the standard basis $|x\rangle$ and $|y\rangle$ cannot be closer than a certain constant.
- So the final states of the algorithm for different S_x^π and S_y^π must be sufficiently far apart.

Intuition Behind the Proof (Cont'd)

- Now the U_i are all the same.
- It follows that any difference in the result arises from calls to S_x^π.
- The algorithms all start with the same state $U_0|0\rangle$.
- We want to obtain a bound on k, the number of calls to the oracle interface S_x^π.
- For this we need to bound from above the amount each step increases the distance between $|\psi_i^x\rangle$ and $|\psi_i^y\rangle$.

Intuition Behind the Proof (Cont'd)

- In other words, we want to bound from above the amount this distance can increase by:
  - Applying $U_i S_x^\pi$ to $|\psi_{i-1}^x\rangle$;
  - Applying $U_i S_y^\pi$ to $|\psi_{i-1}^y\rangle$.
- Let $|\psi_i\rangle$ be the state obtained by applying U_0 up through U_i without any intervening calls to S_x^π.
- To obtain the bound, we compare both $|\psi_i^x\rangle$ and $|\psi_i^y\rangle$ with $|\psi_i\rangle$.
- We first give the details of how to use inequalities based on these ideas to prove that $\Omega(\sqrt{N})$ calls to the oracle are required.
- Then we give detailed proofs of each of the inequalities.

Phase-Adjustment

- The proof considers the relation between three classes of quantum states:
  - The desired result $|x\rangle$;
  - The state of the computation $|\psi_k^x\rangle$ after k steps;
  - The state $|\psi_k\rangle = U_k U_{k-1} \cdots U_1 U_0 |0\rangle$ obtained by performing the sequence of transformations U_i without consulting the oracle.
- The analysis simplifies if we sometimes consider, instead of $|x\rangle$, a phase-adjusted version of $|x\rangle$,

$$|x_k'\rangle = e^{\boldsymbol{i}\theta_k^x}|x\rangle, \quad e^{\boldsymbol{i}\theta_k^x} = \frac{\langle x|\psi_k^x\rangle}{|\langle x|\psi_k^x\rangle|}.$$

Phase-Adjustment (Cont'd)

- $e^{\boldsymbol{i}\theta_k^x}$ is chosen so that $\langle x_k'|\psi_k^x\rangle$ is positive real for all k.
- Indeed, we have

$$\begin{aligned}\langle x_k'|\psi_k^x\rangle &= e^{-\boldsymbol{i}\theta_k^x}\langle x|\psi_k^x\rangle \\ &= \frac{\overline{\langle x|\psi_k^x\rangle}}{|\langle x|\psi_k^x\rangle|}\langle x|\psi_k^x\rangle \\ &= \frac{|\langle x|\psi_k^x\rangle|^2}{|\langle x|\psi_k^x\rangle|} \\ &= |\langle x|\psi_k^x\rangle| \geq 0.\end{aligned}$$

- $|x_k'\rangle$ differs from $|x\rangle$ only in a phase.
- So we have

$$\begin{aligned}|\langle x|\psi_k^x\rangle|^2 \geq \tfrac{1}{2} \quad &\Rightarrow \quad |\langle x_k'|\psi_k^x\rangle|^2 \geq \tfrac{1}{2} \\ &\Rightarrow \quad \langle x_k'|\psi_k^x\rangle \geq \frac{1}{\sqrt{2}}.\end{aligned}$$

Distances Between States

- We consider the distances between certain pairs of these states:

$$d_{kx} = \||\psi_k^x\rangle - |\psi_k\rangle\|, \quad a_{kx} = \||\psi_k^x\rangle - |x_k'\rangle\|, \quad c_{kx} = \||x_k'\rangle - |\psi_k\rangle\|.$$

- We establish bounds involving the average of these distances squared,

$$D_k = \frac{1}{N}\sum_x d_{kx}^2, \quad A_k = \frac{1}{N}\sum_x a_{kx}^2, \quad C_k = \frac{1}{N}\sum_x c_{kx}^2.$$

- The reason for considering the sum, or equivalently the average, is that the algorithm must efficiently find x for all possible x.
- The proof relies on three inequalities involving D_k, A_k, and C_k.
- Before proving the inequalities, we describe them and show how they imply a lower bound on the number of calls to the oracle.

The Three Inequalities

- The first inequality bounds from above A_k, the average squared distance between the state $|\psi_k^x\rangle$ obtained after k steps, and $|x_k'\rangle$.

  We will show that in order to obtain a success probability of $|\langle x|\psi_k^x\rangle|^2 \geq \frac{1}{2}$, we must have $A_k \leq 2 - \sqrt{2}$.
- The second inequality bounds from below C_k, the sum of the squared distances between the vector $|\psi_k\rangle$ and all basis vectors $|j\rangle$.

  We see that, as long as $N \geq 4$, $C_k \geq 1$.
- The third inequality bounds the growth of D_k, the average squared distance between $|\psi_k^x\rangle$ and $|\psi_k\rangle$ as k increases, $D_k \leq \frac{4k^2}{N}$.
- The three quantities d_{kx}, a_{kx} and c_{kx} are related as follows:

$$d_{kx} = \||\psi_k^x\rangle - |\psi_k\rangle| = \||\psi_k^x\rangle - e^{\boldsymbol{i}\theta_x^k}|x\rangle + e^{\boldsymbol{i}\theta_x^k}|x\rangle - |\psi_k\rangle| \geq a_{kx} - c_{kx}.$$

Averages and Number of Iterations

- Using the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} D_k &= \frac{1}{N}\sum_x d_{kx}^2 \\ &\geq \frac{1}{N}\left(\sum_x a_{kx}^2 - 2\sum_x a_{kx}c_{kx} + \sum_x c_{kx}^2\right) \\ &\geq \frac{1}{N}\sum_x a_{kx}^2 - \frac{2}{N}\sqrt{\left(\sum_x a_{kx}^2\right)\left(\sum_x c_{kx}^2\right)} + \frac{1}{N}\sum_x c_{kx}^2 \\ &\geq A_k - 2\sqrt{A_k C_k} + C_k. \end{aligned}$$

- We make use of this inequality, together with

$$A_k \leq 2 - \sqrt{2}, \quad C_k \geq 1, \quad D_k \leq \frac{4k^2}{N}.$$

- We bound $\frac{4k^2}{N}$ from below by a constant.

Averages and Number of Iterations (Cont'd)

- We use
$$D_k \geq A_k - 2\sqrt{A_k C_k} + C_k$$
together with $A_k \leq 2 - \sqrt{2}$, $C_k \geq 1$ (for $N \geq 4$) and $D_k \leq \frac{4k^2}{N}$.
- We obtain
$$\begin{aligned} \frac{4k^2}{N} &\geq D_k \\ &\geq A_k - 2\sqrt{A_k C_k} + C_k \\ &= (\sqrt{C_k} - \sqrt{A_k})^2 \\ &\geq (1 - \sqrt{2 - \sqrt{2}})^2. \quad (1 \geq 2 - \sqrt{2} \geq A_k) \end{aligned}$$
- Thus, for $N \geq 4$, and taking $q = 1 - \sqrt{2 - \sqrt{2}}$, at least
$$k \geq \frac{q}{2}\sqrt{N}$$
iterations are required for success probability $|\langle x|\psi_k^x\rangle|^2 \geq \frac{1}{2}$, for all x.

The Inequality for A_k

- By assumption, $|\langle \psi_k^x | x \rangle|^2 \geq \frac{1}{2}$.
- By the choice of phase $e^{\boldsymbol{i}\theta_k^x}$ relating $|x\rangle$ and $|x_k'\rangle$, $\langle \psi_k^x | x_k' \rangle \geq \frac{1}{\sqrt{2}}$.
- So

$$\begin{aligned} a_{kx}^2 &= \| |\psi_k^x\rangle - |x_k'\rangle \|^2 \\ &= \| |\psi_k^x\rangle \|^2 - 2\langle x_k' | \psi_k^x \rangle + \| |x_k'\rangle \|^2 \\ &\leq 2 - \sqrt{2}. \end{aligned}$$

- From this it follows that

$$A_k = \frac{1}{N} \sum_x a_{kx}^2 \leq 2 - \sqrt{2}.$$

The Inequality for C_k

- The terms c_{kx}^2 can be bounded as follows:

$$\begin{aligned} c_{kx}^2 &= \||x_k'\rangle - |\psi_k\rangle\|^2 \\ &= \|e^{\boldsymbol{i}\theta_k^x}|x\rangle - |\psi_k\rangle\|^2 \\ &= \||\psi_k\rangle\|^2 - \overline{e^{\boldsymbol{i}\theta_k^x}\langle\psi_k|x\rangle} - e^{\boldsymbol{i}\theta_k^x}\langle\psi_k|x\rangle + \||x\rangle\|^2 \\ &= 2 - 2\mathrm{Re}(e^{\boldsymbol{i}\theta_k^x}\langle\psi_k|x\rangle) \\ &\geq 2 - 2|\langle x|\psi_k\rangle|. \end{aligned}$$

The Inequality for C_k (Cont'd)

- We can now bound the average of these terms:

$$\begin{aligned} C_k &= \tfrac{1}{N}\textstyle\sum_x c_{kx}^2 \\ &\geq 2 - \tfrac{2}{N}\textstyle\sum_x |\langle x|\psi_k\rangle| \\ &\geq 2 - \tfrac{2}{\sqrt{N}}\sqrt{\textstyle\sum_x |x|\psi_k\rangle|^2} \quad \text{(Cauchy-Schwarz)} \\ &= 2 - \tfrac{2}{\sqrt{N}}. \\ &\quad (|\psi_k\rangle \text{ a unit vector and } \{|x\rangle\} \text{ a basis}) \end{aligned}$$

- Thus, $C_k \geq 1$, as long as $N \geq 4$.
- Note that this argument made no assumption about $|\psi_k\rangle$.
- So this bound holds for any quantum state $|\psi\rangle$,

$$\frac{1}{N}\sum_x \||x\rangle - |\psi\rangle|^2 \geq 2 - \frac{2}{\sqrt{N}}.$$

The Inequality for D_k

- First, we bound how much the distance between $|\psi_k^x\rangle$ and $|\psi_k\rangle$ can increase each step.
- Consider the following relation between d_{kx} and $d_{k+1,x}$,

$$\begin{aligned} d_{k+1,x} &= \||\psi_{k+1}^x\rangle - |\psi_{k+1}\rangle| \\ &= |U_{k+1}S_x^\pi|\psi_k^x\rangle - U_{k+1}|\psi_k\rangle| \\ &= |S_x^\pi|\psi_k^x\rangle - |\psi_k\rangle| \\ &= |S_x^\pi(|\psi_k^x\rangle - |\psi_k\rangle) + (S_x^\pi - I)|\psi_k\rangle| \\ &\leq |S_x^\pi(|\psi_k^x\rangle - |\psi_k\rangle)| + |(S_x^\pi - I)|\psi_k\rangle| \\ &= d_{kx} + 2|\langle x|\psi_k\rangle|. \end{aligned}$$

- This inequality shows that with each step the distance between $|\psi_k^x\rangle$ and $|\psi_k\rangle$ can increase by at most $2|\langle x|\psi_k\rangle|$.
- Using this bound, we prove by induction that $D_k = \frac{1}{N}\sum_x d_{kx}^2 \leq \frac{4k^2}{N}$.

The Inequality for D_k (Cont'd)

- **Base Case**: Let $k = 0$.

  Then, for all x,

  $$|\psi_0^x\rangle = U_0|0\rangle = |\psi_0\rangle.$$

  So $d_{0x} = 0$. Therefore, $D_0 = 0$.

- **Induction Step**:

$$\begin{aligned} D_{k+1} &= \frac{1}{N}\sum_x d_{k+1,x}^2 \\ &\leq \frac{1}{N}\sum_x (d_{kx} + 2|\langle x|\psi_k\rangle|)^2 \\ &= \frac{1}{N}\sum_x d_{kx}^2 + \frac{4}{N}\sum_x |\langle x|\psi_k\rangle|^2 + \frac{4}{N}\sum_x d_{kx}|\langle x|\psi_k\rangle| \\ &= D_k + \frac{4}{N} + \frac{4}{N}\sum_x d_{kx}|\langle x|\psi_k\rangle|. \end{aligned}$$

The Inequality for D_k (Cont'd)

- We obtained
$$D_{k+1} \leq D_k + \frac{4}{N} + \frac{4}{N} \sum_x d_{kx} |\langle x|\psi_k\rangle|.$$
- The Cauchy-Schwarz inequality gives
$$\begin{aligned} \frac{1}{N} \sum_x d_{kx} |\langle x|\psi_k\rangle| &\leq \frac{1}{N} \sqrt{(\sum_x d_{kx}^2)(\sum_x |\langle x|\psi_k\rangle|^2)} \\ &= \sqrt{\frac{D_k}{N}}. \end{aligned}$$
- By the induction assumption $D_k \leq \frac{4k^2}{N}$.
- So we have
$$D_{k+1} \leq D_k + \frac{4}{N} + 4\sqrt{\frac{D_k}{N}} \leq \frac{4(k+1)^2}{N}.$$

Subsection 4

Derandomization and Amplitude Amplification

Suggested Approached for Derandomization

- Unlike Shor's algorithm, Grover's algorithm is not inherently probabilistic.
- With a little cleverness, Grover's algorithm can be modified in such a way that:
  - It is guaranteed to find a solution;
  - It still preserves the quadratic speedup.
- More generally, amplitude amplification can be derandomized.
- Brassard, Høyer and Tapp suggest two approaches:
  - In the first, each iteration rotates by an angle that is slightly smaller than the one used previously;
  - The second changes only the last step to a smaller rotation.

Approach 1: Modifying Each Step (Idea)

- Suppose the angle θ in Grover's algorithm (or amplitude amplification) happened to be such that $\frac{\pi}{4\theta} - \frac{1}{2}$ is an integer.
- In this case, after $i = \frac{\pi}{4\theta} - \frac{1}{2}$ iterations, the amplitude g_i would be 1.
- Accordingly, the algorithm would output a solution with certainty.
- Recall that θ satisfies $\sin\theta = \sqrt{t} = g_0$.
- We hope to derandomize amplitude amplification for algorithm U with success probability g_0.
- We modify U to obtain an algorithm U' with success probability $g_0' < g_0$ such that, for θ' satisfying $\sin\theta' = g_0'$, the quantity

$$\frac{\pi}{4\theta'} - \frac{1}{2}$$

is an integer.

Approach 1: Modifying Each Step

- Intuitively, it seems as though it should not be hard to modify an algorithm U so that it is less successful.
- We must make sure that we can compute U' efficiently from U.
- The trick is to allow the use of an additional qubit b.
- We assume given an algorithm U with success probability g_0 acting on an n-qubit register $|s\rangle$.
- Let B be the single-qubit transformation

$$B = \sqrt{1 - \frac{g_0'}{g_0}}|0\rangle + \sqrt{\frac{g_0'}{g_0}}|1\rangle.$$

- We define U' to be the transformation

$$U \otimes B$$

on an $(n+1)$-qubit register $|s\rangle|b\rangle$.

Approach 1: Modifying Each Step (Cont'd)

- Let G' be the set of basis states $|x\rangle \otimes |b\rangle$, with $|x\rangle \in G$, $|b\rangle = |1\rangle$.
- It may be checked that the initial success probability

$$|P_{G'} U' |0\rangle| = g'_0.$$

- Perform amplitude amplification on an $(n+1)$-qubit state, with:
  - U' for U;
  - $S^{\pi}_{G'}$ for S^{π}_{G};
  - Iteration operator $Q' = -U' S^{\pi}_0 (U')^{-1} S^{\pi}_{G'}$.
- It succeeds with certainty after $i = \frac{\pi}{4\theta'} - \frac{1}{2}$ steps.
- This modified algorithm obtains a solution with certainty, using $O(\sqrt{\frac{1}{t}})$ calls to the oracle, at the cost of a single additional qubit.

Approach 2: Modifying Only the Last Step

- This approach results in a solution in $O(\sqrt{\frac{1}{t}})$ time with certainty without the need for an additional qubit.
- The idea is to modify S_G^π and S_0^π in the last step so that exactly the desired final state is obtained.
- We begin by analyzing general properties of transformations of the form

$$Q(\phi,\tau) = -US_0^\phi U^{-1} S_G^\tau,$$

where ϕ and τ are arbitrary angles and

$$S_X^\phi|x\rangle = \begin{cases} e^{\boldsymbol{i}\phi}|x\rangle, & \text{if } |x\rangle \in X, \\ |x\rangle, & \text{if } |x\rangle \notin X. \end{cases}$$

- We have showed how to implement S_X^ϕ efficiently.

Approach 2: An Equation

- First, we show that, for any quantum state $|v\rangle$,

$$US_0^{\phi}U^{-1}|v\rangle = |v\rangle - (1 - e^{\boldsymbol{i}\phi})\overline{\langle v|U|0\rangle}U|0\rangle.$$

- Write

$$|v\rangle = \sum_{i=1}^{N-1} \overline{\langle v|U|i\rangle}U|i\rangle + \overline{\langle v|U|0\rangle}U|0\rangle.$$

- Then

$$\begin{aligned} US_0^{\phi}U^{-1}|v\rangle &= US_0^{\phi}\left(\sum_{i=1}^{N-1}\overline{\langle v|U|i\rangle}|i\rangle + \overline{\langle v|U|0\rangle}|0\rangle\right) \\ &= U\left(\sum_{i=1}^{N-1}\overline{\langle v|U|i\rangle}|i\rangle + \overline{\langle v|Ue^{\boldsymbol{i}\phi}|0\rangle}|0\rangle\right) \\ &= \sum_{i=1}^{N-1}\overline{\langle v|U|i\rangle}U|i\rangle + e^{\boldsymbol{i}\phi}\overline{\langle v|U|0\rangle}U|0\rangle \\ &= |v\rangle - (1 - e^{\boldsymbol{i}\phi})\overline{\langle v|U|0\rangle}U|0\rangle. \end{aligned}$$

Approach 2: An Equation (Cont'd)

- Using this result, we now can see the effect of

$$Q(\phi,\tau) = US_0^{\phi}U^{-1}S_G^{\tau}$$

on any superposition $|v\rangle = g|v_G\rangle + b|v_B\rangle$ in the subspace spanned by $|v_G\rangle$ and $|v_B\rangle$.

- We have

$$\begin{aligned} Q(\phi,\tau)|v\rangle &= g(-e^{\boldsymbol{i}\tau}|v_G\rangle + e^{\boldsymbol{i}\phi}(1-e^{\boldsymbol{i}\phi})\overline{\langle v_G|U|0\rangle}U|0\rangle) \\ &\quad + b(-|v_B\rangle + (1-e^{\boldsymbol{i}\phi})\overline{\langle v_B|U|0\rangle}U|0\rangle). \end{aligned}$$

- After $s = \lfloor \frac{\pi}{4\theta} - \frac{1}{2} \rfloor$ iterations of amplitude amplification, we have the state

$$|\psi_s\rangle = \sin((2s+1)\theta)|\psi_G\rangle + \cos((2s+1)\theta)|\psi_B\rangle,$$

where $\sin\theta = \sqrt{t} = g_0$.

Approach 2: An Equation (Cont'd)

- Applying $Q(\phi,\tau)$ to the state $|\psi_G\rangle$, we obtain

$$\begin{aligned}
Q(\phi,\tau)|\psi_G\rangle &= -US_0^{\phi}U^{-1}S_G^{\tau}|\psi_G\rangle \\
&= -US_0^{\phi}U^{-1}(e^{\boldsymbol{i}\tau}|\psi_G\rangle) \\
&= -e^{\boldsymbol{i}\tau}|\psi_G\rangle + (1-e^{\boldsymbol{i}\phi})\overline{\langle e^{\boldsymbol{i}\tau}\psi_G|U|0\rangle}U|0\rangle \\
&= -e^{\boldsymbol{i}\tau}|\psi_G\rangle + e^{\boldsymbol{i}\tau}(1-e^{\boldsymbol{i}\phi})\overline{\langle\psi_G|g_0\psi_G + b_0\psi_B\rangle} \\
&\qquad\qquad (g_0\psi_G + b_0\psi_B) \\
&= e^{\boldsymbol{i}\tau}((1-e^{\boldsymbol{i}\phi})g_0^2 - 1)|\psi_G\rangle \\
&\qquad + e^{\boldsymbol{i}\tau}(1-e^{\boldsymbol{i}\phi})g_0 b_0|\psi_B\rangle).
\end{aligned}$$

- Similarly, applying $Q(\phi,\tau)$ to the state $|\psi_B\rangle$, we obtain

$$Q(\phi,\tau)|\psi_B\rangle = (1-e^{\boldsymbol{i}\phi})b_0 g_0|\psi_G\rangle + ((1-e^{\boldsymbol{i}\phi})b_0^2 - 1)|\psi_B\rangle).$$

Approach 2: An Equation (Cont'd)

- So $Q(\phi,\tau)|\psi\rangle = g(\phi,\tau)|\psi_G\rangle + b(\phi,\tau)|\psi_B\rangle$, where

$$\begin{aligned} g(\phi,\tau) &= \sin((2s+1)\theta)e^{\boldsymbol{i}\tau}((1-e^{\boldsymbol{i}\phi})g_0^2-1) \\ &\quad + \cos((2s+1)\theta)(1-e^{\boldsymbol{i}\phi})b_0 g_0 \\ b(\phi,\tau) &= \sin((2s+1)\theta)e^{\boldsymbol{i}\tau}(1-e^{\boldsymbol{i}\phi})g_0 b_0 \\ &\quad + \cos((2s+1)\theta)((1-e^{\boldsymbol{i}\phi})b_0^2-1). \end{aligned}$$

- Our aim now is to show that there exist ϕ and τ such that if

$$Q(\phi,\tau) = US_0^{\phi}U^{-1}S_G^{\tau}$$

is applied as a final step, a solution is obtained with certainty.

Approach 2: An Equation (Cont'd)

- To show that ϕ and τ can be chosen so that $Q(\phi,\tau)|\psi\rangle$ has all of its amplitude in the good states, we want $b(\phi,\tau) = 0$.
- That is, we need

$$\begin{aligned}(\sin((2s+1)\theta)e^{\boldsymbol{i}\tau}(1-e^{\boldsymbol{i}\phi})g_0 b_0)\\ +\cos((2s+1)\theta)((1-e^{\boldsymbol{i}\phi})b_0^2-1)=0.\end{aligned}$$

- Equivalently, since $b_0 = \sqrt{1-g_0^2}$,

$$\begin{aligned}e^{\boldsymbol{i}\tau}(1-e^{\boldsymbol{i}\phi})g_0\sqrt{1-g_0^2}\sin((2s+1)\theta)\\ =(1-(1-e^{\boldsymbol{i}\phi})(1-g_0^2)\cos((2s+1)\theta).\end{aligned}$$

- The right-hand side equals $(g_0^2(1-e^{\boldsymbol{i}\phi})+e^{\boldsymbol{i}\phi})\cos((2s+1)\theta)$.
- So we want ϕ and τ to satisfy

$$\cot((2s+1)\theta) = \frac{e^{\boldsymbol{i}\tau}(1-e^{\boldsymbol{i}\phi})g_0\sqrt{1-g_0^2}}{g_0^2(1-e^{\boldsymbol{i}\phi})+e^{\boldsymbol{i}\phi}}.$$

Approach 2: An Equation (Cont'd)

- Once ϕ is chosen, we choose τ to make the right-hand side real.
- To find ϕ, compute the magnitude squared of the right-hand side of the preceding equation

$$\frac{g_0^2 b_0^2 (2 - 2\cos\phi)}{g_0^4 (2 - 2\cos\phi) - g_0^2 (2 - 2\cos\phi) + 1}.$$

- The maximum value of the magnitude squared, obtained when $\cos\phi = -1$, is

$$\frac{4g_0^2 b_0^2}{4g_0^4 - 4g_0^2 + 1} = \frac{4g_0^2 b_0^2}{(2g_0^2 - 1)^2}.$$

- So the maximum magnitude is

$$\frac{2g_0 b_0}{2g_0^2 - 1} = \frac{2g_0 b_0}{g_0^2 - b_0^2} = \tan(2\theta),$$

where $\sin\theta = \sqrt{t} = g_0$ as before.

Approach 2: An Equation (Cont'd)

- We conclude that ϕ and τ can be chosen to make the right-hand side of the any real number between $[0, \tan(2\theta)]$.
- By the geometric interpretation, after $s = \lfloor \frac{\pi}{4\theta} - \frac{1}{2} \rfloor$ iterations, the state has been rotated to within 2θ of the desired state.
- We have shown that ϕ and τ can be chosen so that applying s iterations of Q, followed by one application of $Q(\phi, \tau)$, yields a solution with certainty.

Subsection 5

Unknown Number of Solutions

The Case of Unknown t

- Grover's algorithm requires that we know the relative number of solutions $t = \frac{|G|}{N}$ in order to determine how many times we should apply the transformation Q.
- More generally, amplitude amplification requires as input the success probability $t = |g_0|^2$ of $U|0\rangle$.
- We now sketch two approaches to handling cases in which we do not know t.
  - The first approach repeats Grover's algorithm multiple times, choosing a random number of iterations of Q in each run.
    It succeeds in finding a solution with high probability.
  - The second approach, called **quantum counting**, uses the quantum Fourier transform to estimate t.
- Both approaches require $O(\sqrt{N})$ calls to U_P.

Varying the Number of Iterations

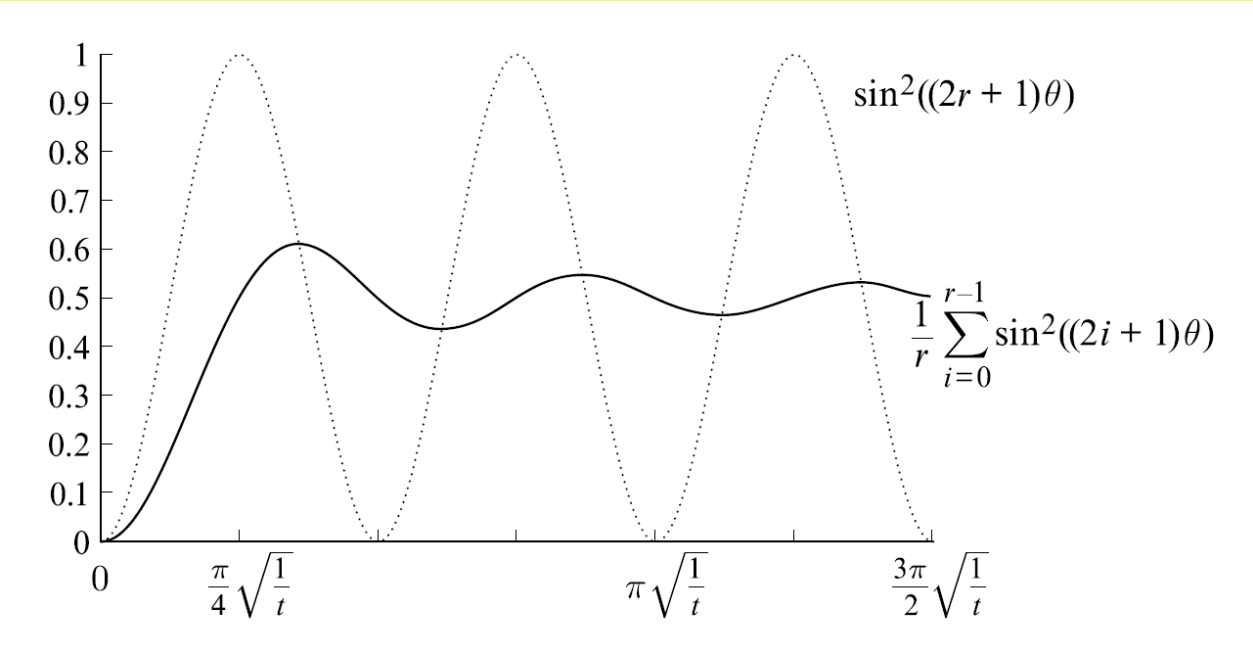
- Consider Grover's algorithm applied to a problem with tN solutions in a space of cardinality N.
- When t is unknown, a simple strategy is to repeatedly execute Grover's algorithm with a number of iteration steps picked randomly between 0 and $\frac{\pi}{4}\sqrt{N}$.
- For large values of t, this simple approach is clearly not optimal.
- Nevertheless, as we show, this simple strategy succeeds with at most $O(\sqrt{N})$ calls to U_P regardless of the value of t.
- Previous results imply that the average probability of success for a run with i iterations of Q, where i is randomly chosen between 0 and r, is given by

$$\Pr(i < r) = \frac{1}{r}\sum_{i=0}^{r-1}\sin^2\left((2i+1)\theta\right),$$

where $\sin\theta = \sqrt{t}$ as before.

Varying the Number of Iterations (Cont'd)

- A plot of the average success probability for different values of r is shown below.



- The graph will be identical for all values of t as long as $t \ll 1$.
- For comparison, the graph of the success probability after exactly r iteration steps of Grover's algorithm is also given.

Varying the Number of Iterations (Cont'd)

- It is easy to see from the graph that there is a constant c, such that

$$\Pr(i < r) > c, \quad \text{for all } r \geq \frac{\pi}{4}\sqrt{\frac{1}{t}}.$$

- Suppose $\frac{1}{t} \leq N$, guaranteeing at least one solution.
- Then, if we choose $r = \frac{\pi}{4}\sqrt{N}$, then

$$\Pr\left(i < \frac{\pi}{4}\sqrt{N}\right) \geq c.$$

- Thus, a single run of the algorithm, where the number of iterations of Q is chosen randomly between 0 and $\frac{\pi}{4}\sqrt{N}$, finds the solution with probability at least c.
- The expected number of calls to the oracle during such a run is therefore $O(\sqrt{N})$.

Varying the Number of Iterations (Cont'd)

- Take any probability $c' > c$.
- Then, there is a constant K, such that if Grover's algorithm is run K times, with the number of iterations for each run chosen as above, then a solution will be found with probability c'.
- Thus, for any c', the total number of times Q is applied is $O(\sqrt{N})$.
- Consequently, for any c', the total number of calls to the oracle is $O(\sqrt{N})$.

Quantum Counting

- Quantum counting takes a more quantum approach:
  - Create a superposition of results for different numbers of applications of Q;
  - Then use the quantum Fourier transform on that superposition to obtain a good estimate for the relative number of solutions t.
- The same strategy can be used for the amplitude amplification algorithm to estimate the success probability t of $U|0\rangle$.
- This approach also has query complexity $O(\sqrt{N})$.

Quantum Counting (Cont'd)

- Let U and Q be as defined in the amplitude amplification algorithm.
- Define a transformation **RepeatQ**, with input $|k\rangle$ and $|\psi\rangle$, that performs k iterations of Q on $|\psi\rangle$:

$$\textbf{RepeatQ}: \ |k\rangle \otimes |\psi\rangle \rightarrow |k\rangle \otimes Q^k|\psi\rangle.$$

- This transformation is more powerful than the classical ability to repeat Q because **RepeatQ** can be applied to a superposition.
- We apply **RepeatQ** to a superposition of all $k < M = 2^m$ tensored with the state $U|0\rangle$ to obtain

$$\frac{1}{\sqrt{M}} \sum_{k=0}^{M-1} |k\rangle \otimes U|0\rangle \rightarrow \frac{1}{\sqrt{M}} \sum_{k=0}^{M-1} |k\rangle \otimes (g_k|\psi_G\rangle + b_k|\psi_B\rangle),$$

where we ignore for the moment how M was chosen.

Quantum Counting (Cont'd)

- A measurement of the right register in the standard basis produces a state $|x\rangle$ that is one of the following:
  - A good state (orthogonal to $|\psi_B\rangle$);
  - A bad state (orthogonal to $|\psi_G\rangle$).
- Thus, the state of the left register collapses to either of:

$$|\psi\rangle = C \sum_{k=0}^{M-1} b_k |k\rangle \quad \text{or} \quad |\psi\rangle' = C' \sum_{k=0}^{M-1} g_k |k\rangle.$$

- Let us suppose the former state $|\psi\rangle$ is obtained.
- A similar reasoning applies for the latter case.
- Since, by a previous section, $b_k = \cos((2k+1)\theta)$, we get

$$|\psi\rangle = C \sum_{k=0}^{M-1} \cos((2k+1)\theta)|k\rangle.$$

Quantum Counting (Cont'd)

- Apply the quantum Fourier transform to this state to obtain

$$\mathcal{F} : C \sum_{k=0}^{M-1} b_k |k\rangle \rightarrow \sum_{j=0}^{M-1} B_j |j\rangle.$$

- We explained that, for a cosine function of period $\frac{\pi}{\theta}$, most of the amplitude is in those B_j that are close the single value $\frac{M\theta}{\pi}$.
- If we measure the state now, from the measured value $|j\rangle$ we obtain, with high probability, a good approximation of θ by taking $\theta = \frac{\pi j}{M}$.
- Thus, with high probability, the value $t = \sqrt{\sin\theta}$ is a good approximation for:
  - The ratio of solutions in the case of Grover's algorithm;
  - The success probability of $U|0\rangle$ in the case of amplitude amplification.

Quantum Counting (Cont'd)

- There is, of course, one issue remaining.
- We do not know a priori a proper value for M.
- This problem can be addressed by repeating the algorithm for increasing M until a meaningful value for j is read.
- We know that $\theta = \frac{j}{M}\pi$.
- So, for a given θ:
  - We will likely read an integer value $j \sim \frac{\theta M}{\pi}$;
  - j will be measured as 0 with high probability when M is chosen too small for the given problem.