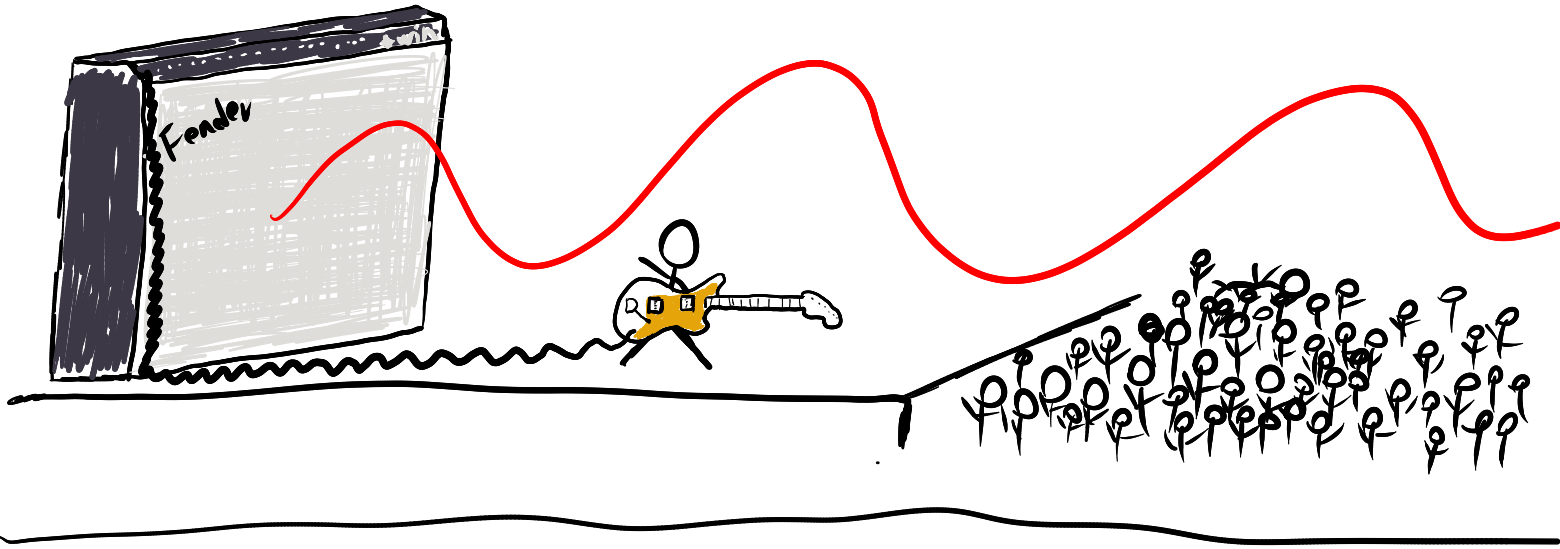
# CMPT Lecture 26

## Amplitude Amplification




Last class we learned about **Grover's algorithm** which solves the unstructured search problem over a search space of size $N$ with $O(\sqrt{N})$ queries, a **square-root speed-up** over the classical case of $O(N)$. Today we discuss the generalization of Grover's algorithm to a quantum algorithm that is used in many other algorithms as a sub-routine:

**Amplitude amplification**

## (Picking out one of many solutions)

To see how Grover's algorithm generalizes, let's first consider the problem of finding **some** solution to a Boolean oracle $f:\{0,1\}^n \to \{0,1\}$. For instance, a solution $f(x)=1$ may be a satisfying assignment to a propositional formula as in SAT solving. Intuitively, Grover's algorithm should work the same way — the key is in the analysis of how many iterations we need to get a reasonably high probability.

Recall that in our analysis of Grover, we wrote the superposition $\frac{1}{\sqrt{2^n}}\sum_x |x\rangle$ as

$$\frac{1}{\sqrt{2^n}}\sum_x |x\rangle = \frac{1}{\sqrt{2^n}}\sum_{x|f(x)=0} |x\rangle + \frac{1}{\sqrt{2^n}}\sum_{x|f(x)=1} |x\rangle$$

If $f$ has $K$ many solutions rather than 1, the probability of measuring one of those is $K/2^n$. Moreover, $\frac{1}{\sqrt{K}}\sum_{x|f(x)=1} |x\rangle$ is a unit vector, so we may write the uniform superposition as the **superposition** of two unit vectors,

$$|\psi_{good}\rangle = \frac{1}{\sqrt{K}}\sum_{x|f(x)=1} |x\rangle \qquad |\psi_{bad}\rangle = \frac{1}{\sqrt{2^n-K}}\sum_{x|f(x)=0} |x\rangle$$

and in particular,

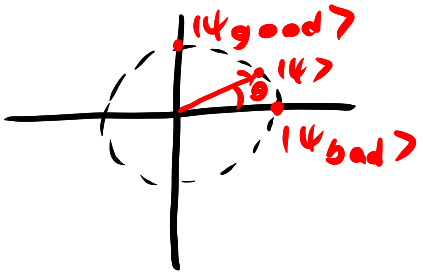
$$\frac{1}{\sqrt{2^n}}\sum_x |x\rangle = \frac{\sqrt{2^n-K}}{\sqrt{2^n}}|\psi_{bad}\rangle + \frac{\sqrt{K}}{\sqrt{2^n}}|\psi_{good}\rangle$$

At this point we may observe that:

1. $|\psi_{bad}\rangle$ & $|\psi_{good}\rangle$ are orthogonal
2. From 1, they span a 2-dimensional subspace of $\mathcal{H}$
3. $U_f$ is a reflection along $|\psi_{bad}\rangle$ in this subspace

Now, let $|\psi\rangle = \frac{1}{\sqrt{2^n}} \sum_x |x\rangle = \frac{\sqrt{2^n - K}}{\sqrt{2^n}} |\psi_{bad}\rangle + \frac{\sqrt{K}}{\sqrt{2^n}} |\psi_{good}\rangle$

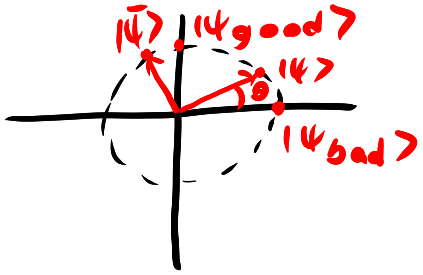
Setting $\frac{\sqrt{2^n - K}}{\sqrt{2^n}} = \cos\theta$ and $\frac{\sqrt{K}}{\sqrt{2^n}} = \sin\theta$, we can visualize $|\psi\rangle$ as a point at an angle of $\theta$ to $|\psi_{bad}\rangle$, i.e.




We can also visualize a state $|\bar{\psi}\rangle$ orthogonal to $|\psi\rangle$ in this view which would have an angle of $\theta$ to $|\psi_{good}\rangle$. Explicitly,

$$|\psi\rangle = \cos\theta |\psi_{bad}\rangle + \sin\theta |\psi_{good}\rangle$$

$$|\bar{\psi}\rangle = -\sin\theta |\psi_{bad}\rangle + \cos\theta |\psi_{good}\rangle$$

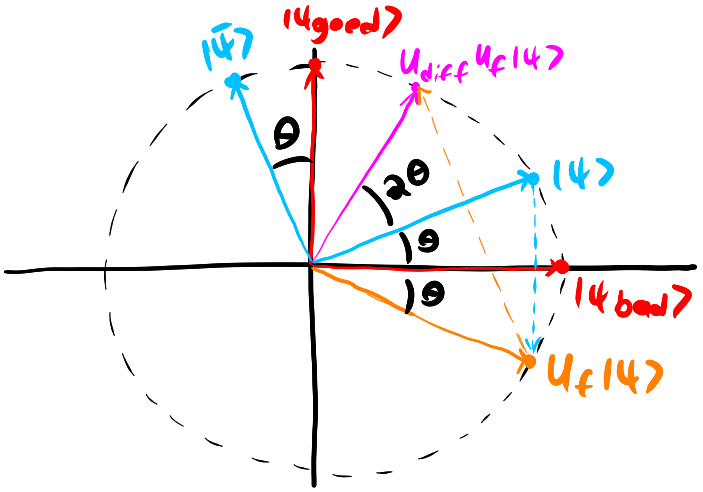



Now, recall that the Grover diffusion operator

$$U_{diff} = 2|\psi\rangle\langle\psi| - I$$

was a reflection along the line $|\psi\rangle$ — so Grover iteration amounts to reflecting along $|\psi_{bad}\rangle$ with $U_f$, then along $|\psi\rangle$ with $U_{diff}$. We can once again visualize this geometrically.

(Grover iteration)




So after one iteration, our new state makes an angle of $3\theta$ with $|\psi_{bad}\rangle$. If we were to repeat this again, we go from $3\theta$ to $-3\theta$ when we reflect along $|\psi_{bad}\rangle$, then we have an angle of $-4\theta$ with $|\psi\rangle$, which after reflecting along $|\psi\rangle$ becomes $4\theta$, or $5\theta$ from $|\psi_{bad}\rangle$. In general, we add $2\theta$ to our angle with every iteration, which we now prove.

(Grover iteration as a rotation)

Let $|\psi_{bad}\rangle$ and $|\psi_{good}\rangle$ be defined as above. Then the Grover iterant $Q = U_{diff}U_f$ is a rotation of $2\theta$ in the subspace spanned by $|\psi_{bad}\rangle$ and $|\psi_{good}\rangle$ — that is

$$Q = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}$$

Proof

We'll just write $U_f$ and $U_{diff}$ as matrices on this subspace and multiply...

As an operator on $\{|\psi_{bad}\rangle, |\psi_{good}\rangle\}$,

$$U_f = |\psi_{bad}\rangle\langle\psi_{bad}| - |\psi_{good}\rangle\langle\psi_{good}| = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

Likewise, $U_{diff}$ is diagonal in $\{|\psi\rangle, |\bar{\psi}\rangle\}$, which we can now express over $\{|\psi_{bad}\rangle, |\psi_{good}\rangle\}$

$$\begin{aligned} U_{diff} &= |\psi\rangle\langle\psi| - |\bar{\psi}\rangle\langle\bar{\psi}| \\ &= (\cos\theta|\psi_{bad}\rangle + \sin\theta|\psi_{good}\rangle)(\cos\theta\langle\psi_{bad}| + \sin\theta\langle\psi_{good}|) \\ &\quad - (-\sin\theta|\psi_{bad}\rangle + \cos\theta|\psi_{good}\rangle)(-\sin\theta\langle\psi_{bad}| + \cos\theta\langle\psi_{good}|) \\ &= \begin{bmatrix} \cos^2\theta - \sin^2\theta & 2\cos\theta\sin\theta \\ 2\cos\theta\sin\theta & \sin^2\theta - \cos^2\theta \end{bmatrix} \\ &= \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \end{aligned}$$

by double-angle + trig identities

Now,

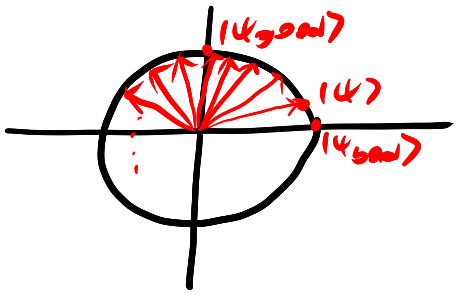
$$Q = U_{diff} U_f = \begin{bmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}$$

□

## (How many iterations are needed?)

Grover iteration hence has the effect of rotating our initial state $|\psi\rangle$ towards $|\psi_{good}\rangle$ at an angle of $2\theta$ every iteration. Since the initial state is at an angle of $\theta$, and by a fact we proved in an assignment, after $k$ iterations we're rotated by an angle of $(2k+1)\theta$. If $(2k+1)\theta = \frac{\pi}{2}$ we've rotated all the way to $|\psi_{good}\rangle$, but if we keep going we overshoot $|\psi_{good}\rangle$ and start moving towards $-|\psi_{bad}\rangle$.




So, we should stop when $k \approx \frac{\pi}{4\theta} - \frac{1}{2}$. However, this requires knowing $\theta$! In the context of SAT, this would correspond to knowing the number of satisfying assignments, a problem called #SAT which is in a complexity class likely harder than NP called #P. Still, in many cases it's reasonable to assume that we do know $\theta$, as we detail now.

## (Optimal number of iterations for Grover's algorithm)

In Grover's algorithm, we had exactly one solution, so

$$\sin\theta = \frac{1}{\sqrt{2^n}}$$

Since $\sin\theta \approx \theta$ for small $\theta$, we can say $\theta \approx \frac{1}{\sqrt{2^n}}$. So the optimal number of iterations is

$$k \approx \frac{\pi}{4\theta} - \frac{1}{2} \approx \frac{\pi}{4}\sqrt{2^n} - \frac{1}{2}$$

## (Optimal number for a known number of solutions)

Let $M$ be the number of solutions. Then

$$\theta \approx \sin\theta = \frac{\sqrt{M}}{\sqrt{2^n}}$$

So we would need $k \approx \frac{\pi}{4} \cdot \frac{\sqrt{2^n}}{\sqrt{M}} - \frac{1}{2}$ iterations.

## (Amplitude amplification)

Imagine we have a classical, probabilistic algorithm with success probability $p$ where the output can be efficiently checked or verified by a function $f$ which returns yes or no (i.e. 1 or 0). This algorithm, call it $Alg$, returns output $x \in \{0,1\}^n$ with probability $p_x$ such that

$$\sum_{x|f(x)=1} p_x = p$$

Now, implement $Alg$ as a unitary operator

$$A|00\cdots0\rangle = \sum_{x\in\{0,1\}^n} \sqrt{p_x}\,|x\rangle$$

We can see that this state, which just performs the classical algorithm on a quantum computer in superposition, can be written as

$$|\psi\rangle = A|00\cdots0\rangle = \sqrt{p}\,|\psi_{good}\rangle + \sqrt{1-p}\,|\psi_{bad}\rangle$$

Setting $\sin^2\theta = p$ and the diffusion operator as

$$\widetilde{U_{diff}} = 2|\psi\rangle\langle\psi| - I = A(2|0\rangle\langle 0| - I)A^\dagger$$

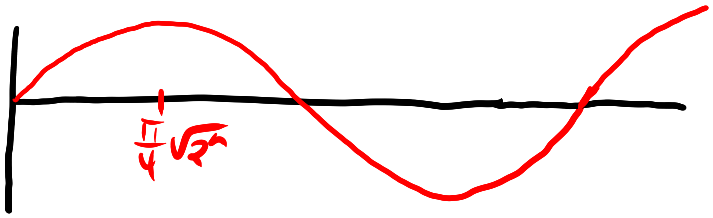
we can perform Grover's search with $Q = \widetilde{U_{diff}}\,U_f$ to amplify the success probability of $Alg$. Moreover, since we know $p$, we can calculate the optimal number of iterations as

$$\frac{\pi}{4\theta} - \frac{1}{2} = \frac{\pi}{4\sqrt{\arcsin p}} - \frac{1}{2} \in O\left(\frac{1}{\sqrt{p}}\right)$$

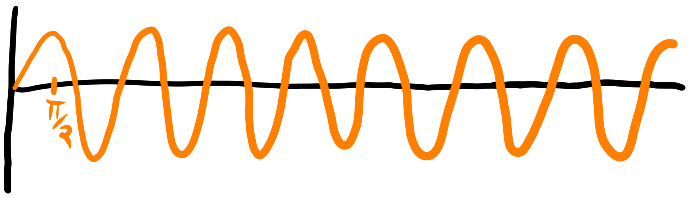
# (Working with an unknown success probability)

What if we truly don't know the approximate probability of success? One option is to just iterate a bunch of times and hope for the best. This isn't likely to be successful, as the distance from $|\psi_{good}\rangle$ — hence the probability of measuring a solution state — oscillates at wildly varying frequencies

$M=1$: $\langle\psi_{good}|Q^k|\psi\rangle$




$M=2^{n-1}$: $\langle\psi_{good}|Q^k|\psi\rangle$




If we could instead **estimate** the probability of success, we could then run Grover's search with an optimal number of iterations. In 1998, Brassard (McGill), Høyer (Calgary), and Tapp (Montreal) did exactly that by using phase estimation **on the Grover iterate** to estimate $\theta$. Their algorithm — **Quantum counting** — which was applied to the **unstructured search problem** with an unknown number of solutions $M$, like Grover's algorithm, can be generalized to the case of

$$A|00\cdots0\rangle = \sin\theta|\psi_{good}\rangle + \cos\theta|\psi_{bad}\rangle$$

where it is called **Amplitude estimation**.

| Problem | Algorithm $H^{\otimes n}\|00\cdots0\rangle=\frac{1}{\sqrt{2^n}}\sum_{f(x)=1}\|x\rangle+\frac{1}{\sqrt{2^n}}\sum_{f(x)=0}\|x\rangle$ | Generalization $A\|00\cdots0\rangle=\sqrt{p}\|\psi_{good}\rangle+\sqrt{1-p}\|\psi_{bad}\rangle$ |
|---|---|---|
| Searching | Grover | Amplitude amplification |
| Counting | Quantum Counting | Amplitude estimation |

# (Quantum Counting)

The quantum counting problem BHT 98 solved can be phrased as

input: A function $f: \{0,1\}^n \longrightarrow \{0,1\}$

goal: Compute $M = \#$ of solutions $f(x)=1$

We already know that:

1. $|\psi\rangle = H^{\otimes n}|00\cdots0\rangle = \frac{1}{\sqrt{2^n}} \sum_{f(x)=1} |x\rangle + \frac{1}{\sqrt{2^n}} \sum_{f(x)=0} |x\rangle$

$$= \sqrt{\frac{M}{2^n}}\,|\psi_{good}\rangle + \sqrt{\frac{2^n-M}{2^n}}\,|\psi_{bad}\rangle$$

2. $U_{diff} U_f = Q = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}$

3. $\sin^2\theta = \frac{M}{2^n}$

Now, what are the eigenvalues of $Q$?

$$\begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \begin{bmatrix} \cos 2\theta + i\sin 2\theta \\ \sin 2\theta - i\cos 2\theta \end{bmatrix}$$

$$= \begin{bmatrix} e^{i2\theta} \\ -i e^{i2\theta} \end{bmatrix}$$

$$= e^{i2\theta} \begin{bmatrix} 1 \\ -i \end{bmatrix}$$

$$\begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix} \begin{bmatrix} 1 \\ i \end{bmatrix} = \begin{bmatrix} \cos 2\theta - i\sin 2\theta \\ \sin 2\theta + i\cos 2\theta \end{bmatrix}$$

$$= e^{-i2\theta} \begin{bmatrix} 1 \\ i \end{bmatrix}$$

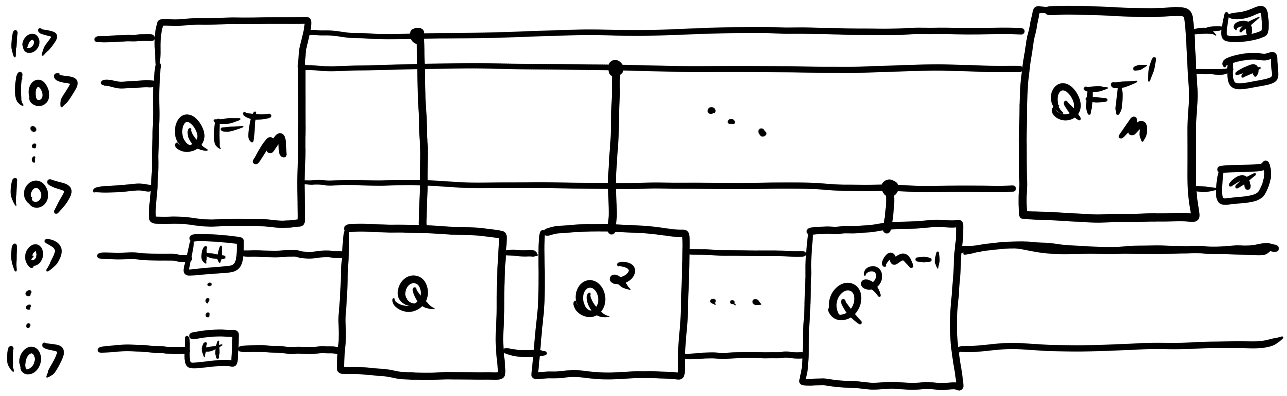
So knowing either eigenvalue of $Q$ gives us $\pm 2\theta$, which is enough to calculate $M$!

Moreover, we already know how to estimate an eigenvalue of a unitary operator:

## Phase estimation

## (The quantum counting algorithm)

The quantum counting (and its generalization, amplitude estimation) really just amounts to phase estimation on $Q$, which we can draw as follows




After measurement we obtain some $y \in \mathbb{Z}_n$ which is an estimate of $\pm \frac{2M\theta}{\pi}$, so we return the estimate

$$M = 2^n \sin^2\theta = 2^n \sin^2\left(\pi \frac{y}{2^n}\right)$$

The number $N = 2^n$ is the bits of precision of the estimation. Note that the complexity is exponential in $n$, even in the query model! In particular, $Q^{2^k} = \underbrace{QQ\cdots Q}_{2^k \text{ times}}$, so $Q^{2^{n-1}}$ makes $O(2^n)$ queries alone. As a result, quantum counting or amplitude estimation is only really practical to low orders of precision.

(So why does phase estimation solve eigenvalue estimation and period finding efficiently, but not #SAT?)

The key lies in the ability to efficiently implement the $2^k$-th powers of $U$ in phase estimation. In period finding,

$$U_a|x\rangle = |a \cdot x \bmod M\rangle$$

so $U_a^k|x\rangle = |a^k \cdot x \bmod M\rangle = U_{a^k}|x\rangle$ — that is, $U_a^k = U_{a^k}$ which can be efficiently implemented.

Likewise, in Hamiltonian eigenvalue estimation,

$$U(t)|x\rangle = e^{-i\hat{H}t}|x\rangle$$

so $U(t)^k|x\rangle = (e^{-i\hat{H}t})^k|x\rangle = e^{-i\hat{H}tk}|x\rangle = U(tk)|x\rangle$, which is efficiently implementable (assuming $U(t)$ is).

By contrast, in amplitude estimation,

$$Q^k = (U_{diff}U_f)^k = ???$$

In particular, the only real way to implement $Q^k$ in general is to just do it $k$ times...

Woh Waaaaaah...

What if we were satisfied with a low-precision answer? Surely that would be enough to solve SAT since we only need to check whether $M \neq 0$. Well, it turns out that to do this with high certainty, we still need exponentially many queries. Oh well.

## (Other applications of quantum searching)

Grover's search paradigm provides speed-ups to a large number of general problems, like

1. Minimum finding
2. Element distinctness
3. Collision finding
4. Hash inversion
5. Combinatorial optimization
6. etc...

However, these speed-ups are only polynomial speed-ups – i.e. $O(N)$ to $O(\sqrt{N})$, so they don't magically make an intractable problem tractable. Moreover, the overheads associated with running them are massive. Case in point: inverting SHA256 should take $\approx 2^{256}$ queries classically. On a quantum computer, this gets down to $\approx \sqrt{2^{256}} = 2^{128}$ queries, but when you leave the black box model you get something like $2^{166}$ SEQUENTIAL operations, using $2402 \approx 2^{12}$ classical processors for error correction, so the speed-up is more like a factor of $2^{78}$, not $2^{128}$.


Amplitude amplification finds more practical use in conjunction with phase estimation (e.g. in HHL) or in Hamiltonian simulation via Linear Combinations of unitaries.