Lectures 3 & 4 – A linear algebra perspective and solving another hard problem with interference

Recap from last time

- Solving problems in interference model
- interf 1, interf 2
- Decision vs search problems
- Black box model - goal is to minimize number of queries

	classical	Interference
– Deutsch-Jozsa	3	1
Bernstein-Vazirani	n	1

These lectures

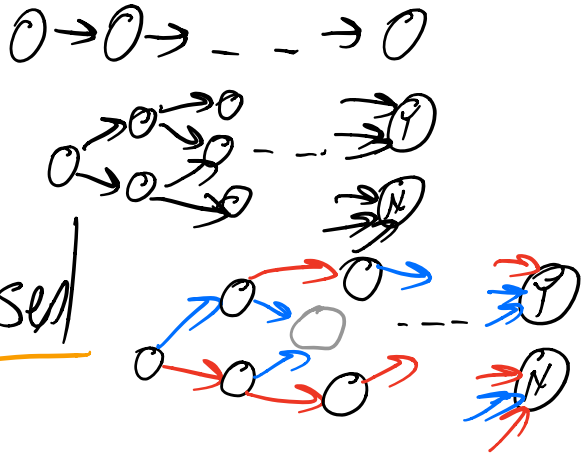
- A linear algebra perspective
- Simon's problem (exponential separation between classical and interference based computation)

A linear algebra perspective (Lecture 3)

Deterministic

Probabilistic

Interference-based



States

Suppose each node is labelled by an m-bit string $w \in \{0,1\}^m$

Let $|w\rangle$ be a 2^m component column vector that is 0 everywhere except in the w'th component where it's 1

$$|w\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} 000\ldots0 & 0 \\ 000\ldots1 & 1 \\ \vdots & \vdots \\ w & w \\ \vdots & \vdots \\ 111\ldots1 & 2^m-1 \end{matrix}$$

E.g. $m = 2$

$w = 01$

$$|w\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \begin{matrix} 00 \\ 01 \\ 10 \\ 11 \end{matrix}$$

Let $\langle w|$ be the transposed raw vector

$$\langle w| = (0\ 1\ 0\ 0)$$

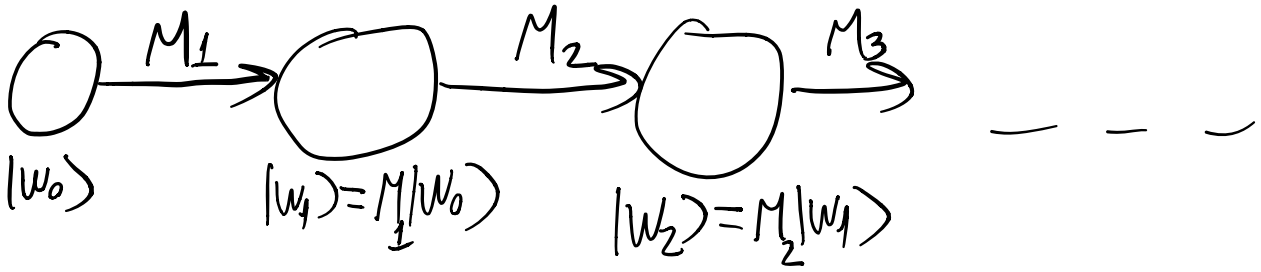
$|w\rangle$ - bra vector; $\langle w|$ - ket vector

$$\langle w|w\rangle = (\ldots \underset{w}{1} \ldots) \begin{matrix} \\ \\ w \\ \\ \\ \end{matrix} \begin{pmatrix} \vdots \\ \\ 1 \\ \\ \vdots \end{pmatrix} = 1$$

Operations

Let M be a $2^m \times 2^m$ matrix for which each raw and each column has exactly one 1 entry and all other entries are 0 (permutation matrix)

$$M|w\rangle = |w'\rangle$$



Suppose now that M has real entries in $[0,1]$ and that the values in each row and column sum to 1 (doubly-stochastic matrix)

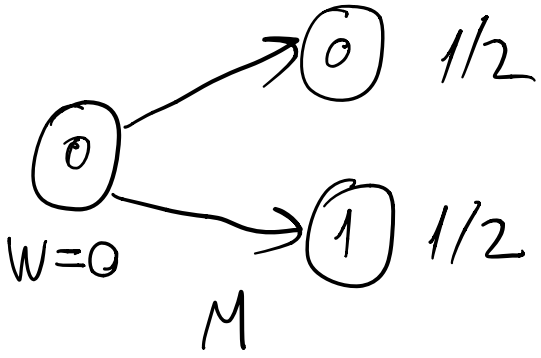
$$\Rightarrow M|w\rangle = \sum_{i \in \{0,1\}^m} M(i,w)|i\rangle$$

$$\sum_i M(i,w) = 1$$

E.g. $M = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}$ $\quad w = 0$

$$|w\rangle = \begin{matrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \begin{matrix} 0 \\ 1 \end{matrix} \end{matrix}$$

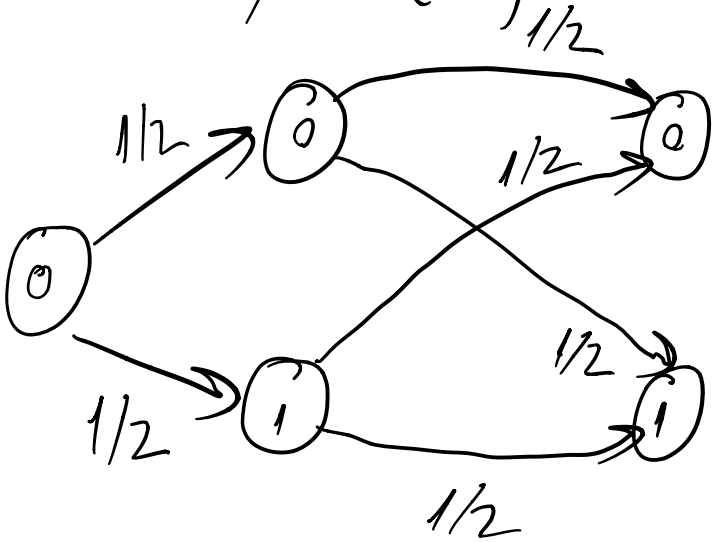
$$M|w\rangle = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$$



What about $M \cdot (M|w\rangle) = M^2 \cdot |w\rangle =$

$$= \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \cdot \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix} \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} \\ \frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}$$



Finally suppose M has entries in $[-1, 1]$ and the raws and columns have sum to 1 when taking the absolute values of the entries (L_1 norm)

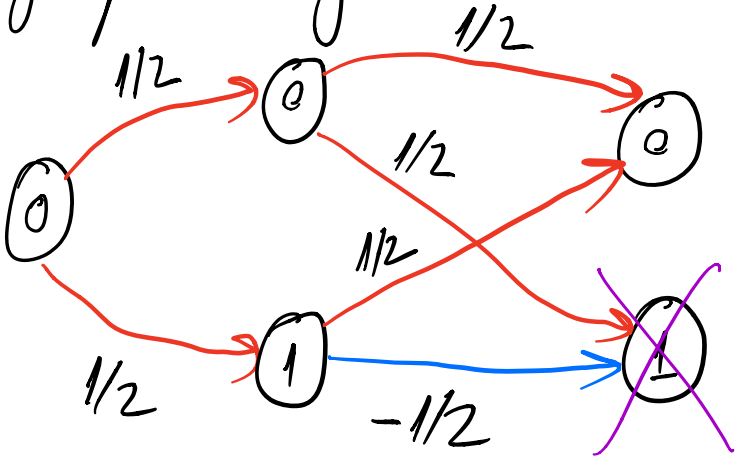
E.g. $M = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ $|w\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

What is $M^2 \cdot |w\rangle$?

$$M|w\rangle = \frac{1}{2}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$M \cdot M|w\rangle = \frac{1}{4}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 1+1 \\ 1-1 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 0 \end{pmatrix}$$

This is the amplitude vector. Have to renormalize to get probability vector.



When we are given oracle access to some function $f: \{0,1\}^n \to \{0,1\}$ we can imagine the oracle as a matrix operation that maps $|x,0\rangle$ to $|x, f(x)\rangle$

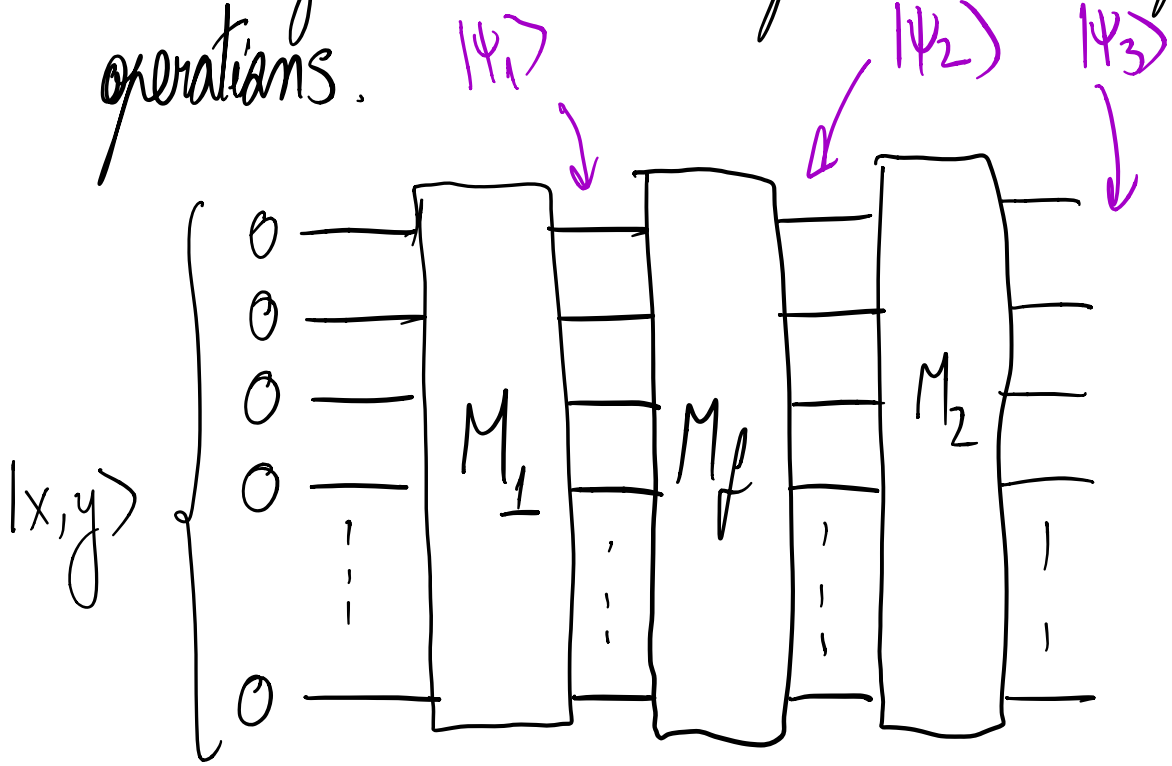
We'll denote it as M_f.

In general, M_f maps $|x,y\rangle$ to $|x, f(x) \oplus y\rangle$

(the reason for this choice will be made clear in future lectures)

Note that $|x,y\rangle$ (n, 1) is a 2^{n+1}-dimensional column vector.

We can give a nice circuit picture to all of these operations.



If our memory state starts out as 0^{n+1}, so $x = 0^n$, $y = 0$ our initial ket state is $|0^n, 0\rangle$ or just $|0^{n+1}\rangle$

We then perform some operation defined by the matrix M_1 (a $2^{n+1} \times 2^{n+1}$ matrix) and obtain the state denoted $|\psi_1\rangle = M_1 \cdot |0^{n+1}\rangle$

Note that $|\psi_1\rangle$ need not be an $|x,y\rangle$ type state, but it's some linear combination of such states.

We then get $|\psi_2\rangle = M_f \cdot M_1 |0^{n+1}\rangle$

$$|\psi_3\rangle = M_2 M_f M_1 |0^{n+1}\rangle$$

and so on.

Composing states and operations

The state $|x,y\rangle$ is a composition of the 2^n-dim vector $|x\rangle$ and the 2-dim vector $|y\rangle$. What is the composition operation?

I.e. What is $\otimes$ s.t. $|x\rangle \otimes |y\rangle = |x,y\rangle$?

$$|x\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} 00\ldots0 \\ 00\ldots1 \\ \vdots \\ x \\ \vdots \\ 11\ldots1 \end{matrix} \quad \underbrace{\qquad}_{n}$$

$$|y\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \text{ if } y=0$$

$$|y\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \text{ if } y=1$$

$$|x,0\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} 00\ldots0 \\ 00\ldots1 \\ \vdots \\ x,0 \\ x,1 \\ \vdots \\ 11\ldots1 \end{matrix} \quad \underbrace{\qquad}_{n+1}$$

$$|x,1\rangle = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \begin{matrix} 00\ldots0 \\ 00\ldots1 \\ \vdots \\ x,0 \\ x,1 \\ \vdots \\ 11\ldots1 \end{matrix} \quad \underbrace{\qquad}_{n+1}$$

Also note that $|x\rangle \otimes |y\rangle$ is not necessarily the same as $|y\rangle \otimes |x\rangle$ since $x,y \neq y,x$

Finally consider $(|x_1\rangle + |x_2\rangle) \otimes |y\rangle$

What should this correspond to?

$$|x_1,y\rangle + |x_2,y\rangle$$

Similarly $|x\rangle \otimes (|y_1\rangle + |y_2\rangle) = |x, y_1\rangle + |x, y_2\rangle$

This is all we need to determine $\otimes$

For 2 vectors $|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \vdots \\ \psi_{N_1} \end{pmatrix}$, $|\phi\rangle = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{N_2} \end{pmatrix}$

take $|\psi\rangle \otimes |\phi\rangle = \begin{pmatrix} \psi_1 \begin{pmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_{N_2} \end{pmatrix} \\ \psi_2 \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_{N_2} \end{pmatrix} \\ \vdots \\ \psi_{N_1} \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_{N_2} \end{pmatrix} \end{pmatrix}$ → $N_1 \cdot N_2$ - dim vector

This is called tensor product.

We can do the same thing with operations

A – $N_1 \times N_1$ matrix

B – $N_2 \times N_2$ matrix

$$A = \begin{pmatrix} a_{1,1} & \cdots & a_{1,N_1} \\ \vdots & & \vdots \\ a_{N_1,1} & \cdots & a_{N_1,N_1} \end{pmatrix}$$

$$B = \begin{pmatrix} b_{1,1} & \cdots & b_{1,N_2} \\ \vdots & & \vdots \\ b_{N_2,1} & \cdots & b_{N_2,N_2} \end{pmatrix}$$

$$A \otimes B = \begin{pmatrix} a_{1,1} \cdot B & \cdots & a_{1,N_1} \cdot B \\ \vdots & & \\ a_{N_1,1} \cdot B & \cdots & a_{N_1,N_1} \cdot B \end{pmatrix}$$

E.g. Denote $H = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

(note that $I \cdot |\psi\rangle = |\psi\rangle$)

$$H \otimes I = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{pmatrix}$$

$$I \otimes H = \frac{1}{2}\begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -1 \end{pmatrix}$$

Suppose we act with $I \otimes H$ on $|0\rangle \otimes |0\rangle$

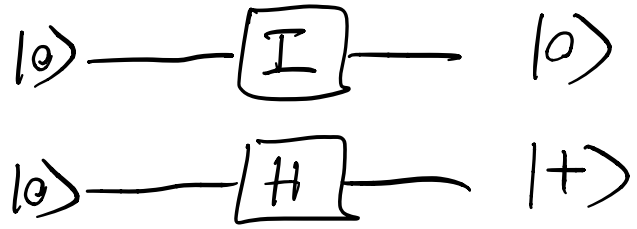
$$\Leftrightarrow I \otimes H \cdot \underbrace{(|0\rangle \otimes |0\rangle)}_{|00\rangle} = I \cdot |0\rangle \otimes H|0\rangle =$$

$$= |0\rangle \otimes |+\rangle$$

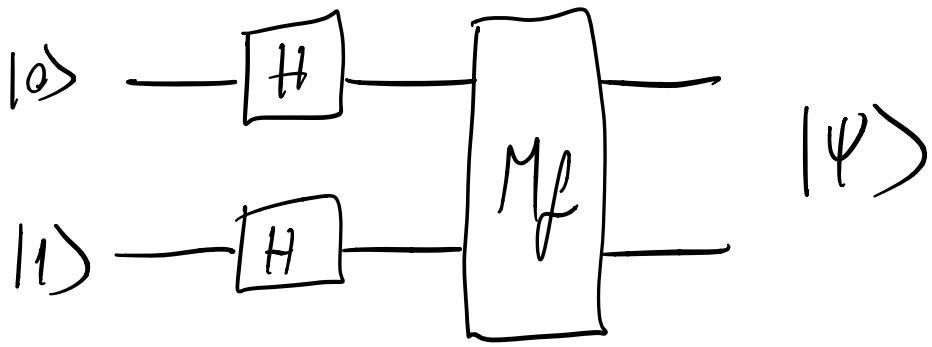
where $|+\rangle = \frac{1}{2}\begin{pmatrix}1\\1\end{pmatrix}$; also call $|-\rangle = \frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix}$

$= H|1\rangle$

In circuit form



Adding M_f into the mix, we'll consider the following circuit, for $f: \{0,1\} \to \{0,1\}$ (recall that $M_f|x,y\rangle = |x, f(x) \oplus y\rangle$; in this case $x, y \in \{0,1\}$)



What is $|\psi\rangle$?

$$|\psi\rangle = M_f \cdot (H \otimes H) \cdot (|0\rangle \otimes |1\rangle)$$

$$\Rightarrow |\psi\rangle = M_f \cdot \left(|+\rangle \otimes |-\rangle \right)$$

But $|+\rangle = \frac{1}{2}\left(|0\rangle + |1\rangle\right)$; $|-\rangle = \frac{1}{2}\left(|0\rangle - |1\rangle\right)$

So: $|\psi\rangle = \frac{1}{4} \cdot M_f \cdot \left[\left(|0\rangle + |1\rangle\right) \otimes \left(|0\rangle - |1\rangle\right) \right]$

We know $\otimes$ distributes with respect to addition, so

$$|\psi\rangle = \frac{1}{4} \cdot M_f \cdot \left[|00\rangle - |01\rangle + |10\rangle - |11\rangle \right]$$

We know $M_f |x, y\rangle = |x, y \oplus f(x)\rangle$

$$\Rightarrow |\psi\rangle = \frac{1}{4} \left[|0, f(0)\rangle - |0, 1 \oplus f(0)\rangle + \right.$$
$$\left. + |1, f(1)\rangle - |1, 1 \oplus f(1)\rangle \right]$$

Let's now apply another H operation to the second register in $|\psi\rangle$ (so act with $I \otimes H$ on $|\psi\rangle$)

First let's see how H acts on $|f(x)\rangle$

If $f(x)=0$ we get $\frac{1}{2}(|0\rangle + |1\rangle)$

$f(x)=1$ we get $\frac{1}{2}(|0\rangle - |1\rangle)$

So $H\cdot|f(x)\rangle = \frac{1}{2}\left(|0\rangle + (-1)^{f(x)}|1\rangle\right)$

$$\Rightarrow I\otimes H|\psi\rangle = \frac{1}{8}\Big[|0\rangle\left(|0\rangle + (-1)^{f(0)}|1\rangle\right) -$$

$$- |0\rangle\left(|0\rangle - (-1)^{f(0)}|1\rangle\right) +$$

$$+ |1\rangle\left(|0\rangle + (-1)^{f(1)}|1\rangle\right) -$$

$$- |1\rangle\left(|0\rangle - (-1)^{f(1)}|1\rangle\right)\Big]$$

$$= \frac{1}{4}\Big[(-1)^{f(0)}|01\rangle + (-1)^{f(1)}|11\rangle\Big]$$

$$= \frac{1}{4}\Big[(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle\Big]|1\rangle$$

The top register is in the state

$(-1)^{f(0)}|0\rangle + (-1)^{f(1)}|1\rangle$, which in vector form is

$$\begin{pmatrix} (-1)^{f(0)} \\ (-1)^{f(1)} \end{pmatrix}$$

If we now also apply H to this state, we have

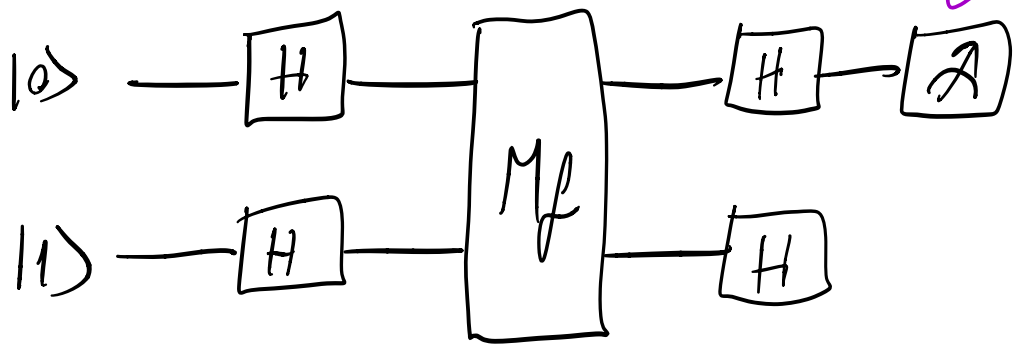
$$\frac{1}{2}\left((-1)^{f(0)}(|0\rangle+|1\rangle) + (-1)^{f(1)}(|0\rangle-|1\rangle)\right)$$

$$= \frac{1}{2}\Big(\left[(-1)^{f(0)} + (-1)^{f(1)}\right]|0\rangle + \left[(-1)^{f(0)} - (-1)^{f(1)}\right]|1\rangle \Big)$$

We'll call this state $|\phi\rangle$

The final missing piece is assigning probabilities to outcomes of a computation. For instance, in our circuit we'd like to know the probability of seeing a $|0\rangle$

Measurement



We said that the way we do this in our model is to take absolute values of amplitudes and renormalize.

So, for $|\phi\rangle$, the probability of $|0\rangle$ will be

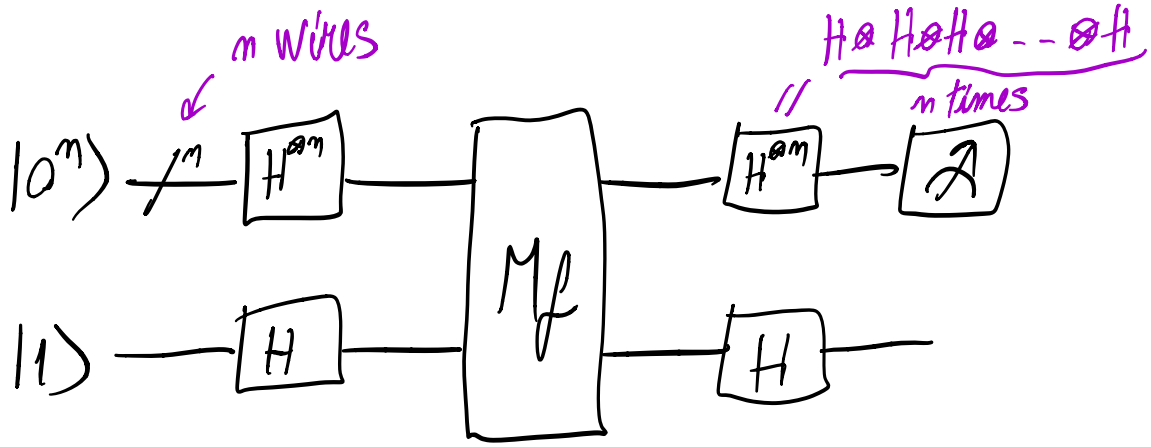
$$\frac{1}{2}\left|(-1)^{f(0)}+(-1)^{f(1)}\right|$$

If $f(0)=f(1)$ this is $\underline{1}$

$f(0)\neq f(1)$ this is 0

Look familiar?

If we generalize the above picture for $f:\{0,1\}^n\to\{0,1\}$ our circuit will look like this



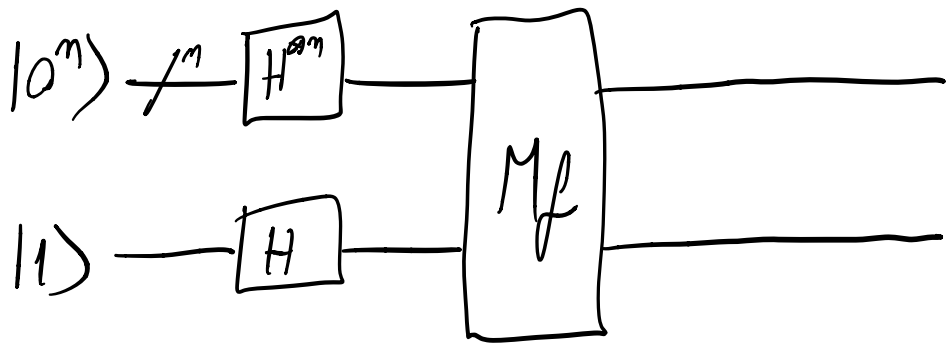
Where now we we want the probability of seeing $|0^n\rangle$ in the top register.

By generalising calculation from above, can be shown that this is

$$\frac{1}{2^n} \left| \sum_{x \in \{0,1\}^n} f(x) \right|$$

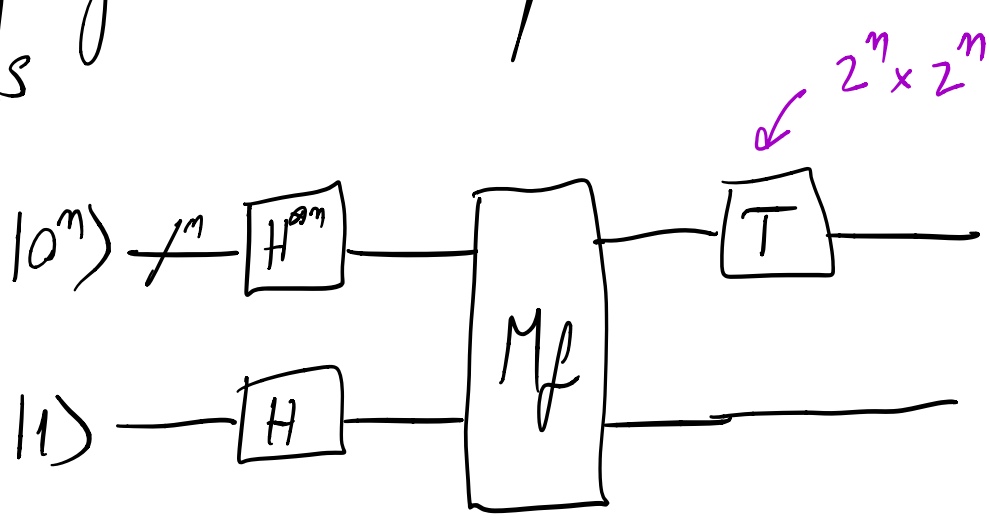
This is exactly what we were computing with interf1!

The circuit



prepares the state $\sum_{x\in\{0,1\}^n} (-1)^{f(x)}|x\rangle$ in the top register.

Acting now with an operation T on that register yields



$$T \sum_{x\in\{0,1\}^n} (-1)^{f(x)}|x\rangle$$

If we denote $T(x,y)$, $x,y\in\{0,1\}^n$, as the (x,y) component of T, then the probability of seeing $|y\rangle$ at the output is just

$$\left| \sum_{x\in\{0,1\}^n} (-1)^{f(x)} \cdot T(x,y) \right|$$

suitably normalized

This is interf 2!

We now understand these operations from a linear algebraic perspective.

Simon's problem (Lecture 4)

In: oracle access to $f: \{0,1\}^n \to \{0,1\}^n$

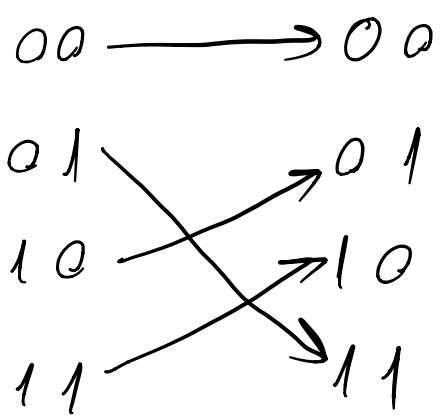
f is either 1-to-1 function (bijection)

Simon function: 2-to-1 with the property that $f(x) = f(y) \Leftrightarrow y = x \oplus s$, for some $s \in \{0,1\}^n$, $s \neq 0^n$

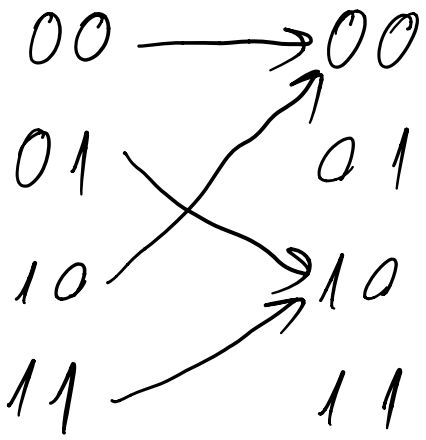
Out: Yes if Simon function
No if 1-to-1 function

E.g. $n = 2$

1-to-1



2-to-1 $s = 10$



How hard is Simon's problem classically?

Deterministic case

Query f on half of domain + 1 to see if there's a collision (a pair $x, y \in \{0,1\}^n$ for which $f(x) = f(y)$)

Takes $2^{n-1} + 1$ queries in the worst case.

As with Deutsch-Jozsa, we can't do any better

here. We can always be unlucky in our queries and think the function is 1-to-1 when it's 2-to-1, unless we query $2^{n-1}+1$ distinct elements

Probabilistic case (hand-wavy)

Suppose we query f on N random points. If f is 2-to-1, how many queries do we need to have a collision with high probability?

I.e. if the values are $x_1, x_2 \dots x_N$, we want to know if there is an i and a j, $i \neq j$, $1 \leq i, j \leq N$ such that $f(x_i) = f(x_j)$.

How many pairs can potentially form a collision?

$\binom{N}{2} \sim N^2$. How many collisions are there in total for f? 2^{n-1}

So if we take N s.t. that $N^2 \sim 2^{n-1}$ we should find a collision with high probability.

Thus we can solve the problem with around $\sqrt{2^n}$ queries.

This is related to the Birthday Paradox (how many people do you need until there's a high chance 2 of them share the same birthday? Around $\sqrt{365} \sim 20$; the number is 23, so very close).

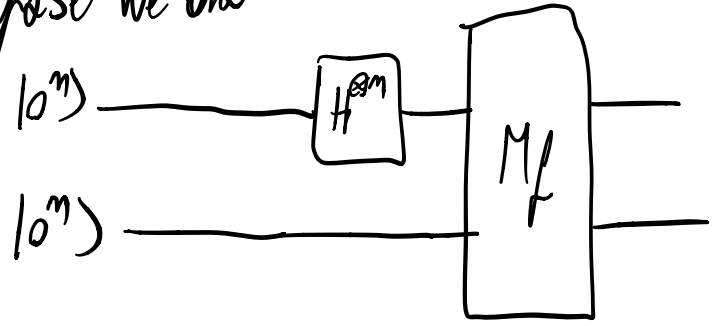
It turns out that this is optimal! (though we won't prove it)

Interference case

Let's use the circuit picture.

Assume $M_f |x,y\rangle = |x, y \oplus f(x)\rangle$, where now $x, y \in \{0,1\}^n$.

Suppose we do



$$H^{\otimes n}|0^n\rangle = \frac{1}{2^n}\left(|0\rangle+|1\rangle\right)\left(|0\rangle+|1\rangle\right)\cdots\left(|0\rangle+|1\rangle\right)$$

$$= \frac{1}{2^n}\sum_{x\in\{0,1\}^n}|x\rangle$$

So after $H^{\otimes n}$ we have

$$\frac{1}{2^n}\sum_{x\in\{0,1\}^n}|x\rangle|0^n\rangle$$

Applying M_f we get

$$\frac{1}{2^n}\sum_{x\in\{0,1\}^n}|x\rangle|f(x)\rangle$$

When f is a Simon function, we can rewrite that as

$$\frac{1}{2^n}\sum_{x}\left(|x\rangle+|x\oplus s\rangle\right)|f(x)\rangle$$

↑ sum here is no longer over all of $\{0,1\}^n$ since we would be double-counting x and $x\oplus s$

So if we observe the second register and see $|y\rangle$

(with prob $\frac{1}{2^{n-1}}$ over $\mathrm{Im}(f)$), the first register will be $\frac{|x_1\rangle + |x_2\rangle}{2}$, with $f(x_1) = f(x_2) = y$

and $x_1 \oplus x_2 = s$

If we were to measure it, we'd see x_1 or x_2 with equal probability. But we won't measure it yet. Let's apply $H^{\otimes n}$ on it before we do.

$$|\psi\rangle = H^{\otimes n} \cdot \frac{1}{2}\left(|x_1\rangle + |x_2\rangle\right)$$

A relation we'll show next time

$$H^{\otimes n}|x\rangle = \frac{1}{2^n} \sum_{y \in \{0,1\}^n} (-1)^{x \cdot y} |y\rangle$$

$$\Rightarrow |\psi\rangle = \frac{1}{2} \cdot \frac{1}{2^n} \left[\sum_{y \in \{0,1\}^n} (-1)^{x_1 \cdot y} |y\rangle + \sum_{y \in \{0,1\}^n} (-1)^{x_2 \cdot y} |y\rangle \right]$$

$$= \frac{1}{2} \cdot \frac{1}{2^n} \cdot \sum_y \left((-1)^{x_1 \cdot y} + (-1)^{x_2 \cdot y} \right) |y\rangle$$

but $x_2 = x_1 \oplus s$

$$\Rightarrow |\psi\rangle = \frac{1}{2^{n+1}} \sum_y \left((-1)^{x_1 \cdot y} + (-1)^{x_1 \cdot y \oplus s \cdot y} \right) |y\rangle$$

$$= \frac{1}{2^{n+1}} \sum_y (-1)^{x_1 \cdot y} \left(1 + (-1)^{s \cdot y} \right) |y\rangle$$

Probability of seeing a particular $|y\rangle$ is going to be proportional to $\left| 1 + (-1)^{s \cdot y} \right|$

This is maximal when $s \cdot y = 0$

$\Rightarrow$ we get a y s.t $s \cdot y = 0$

Suppose we repeat this $O(n)$ times and get a bunch of y's values $y_1, y_2 \cdots y_{n-1}$ that are linearly independent.

We know that all are orthogonal to s, so...

$$y_1^1 s^1 \oplus y_1^2 s^2 \oplus \cdots \oplus y_1^n s^n = 0$$

$$y_2^1 s^1 \oplus y_2^2 s^2 \oplus \cdots \oplus y_2^n s^n = 0$$

$$\vdots$$

$$y_{n-1}^1 s^1 \oplus y_{n-1}^2 s^2 \oplus \cdots \oplus y_{n-1}^n s^n = 0$$

Can solve this with Gaussian elimination to find s!

Only required $O(n)$ queries to f (each M_f call is one query).

Can get the y values with interf 2

$$dom = \{(x, f(x)),\ x \in \{0,1\}^n\}$$

$$T = H^{\otimes n} \otimes I$$

$$T(u, v) = (-1)^{u[0] \cdot v[0]} \cdot (u[1] == v[1])$$

(u: $(x, f(x))$, v: $(y, f(y))$)

$fun = 1$

$y \leftarrow interf2(fun, T, dom)$

$s.t \quad y \cdot s = 0$

Repeat this many times and do Gaussian elimination to find s.

Note that we can check whether $f(x) = f(x \oplus s)$. If this is not the case, we conclude that f is 1-to-1.