In this exercise sheet we will prove two things: Grover's search algorithm for the unstructured search is quadratically better than any classical algorithm; and no quantum algorithm can perform the search significantly faster than Grover's algorithm, i.e. Grover is optimal.

Exercise 1. *Warm-up: classical query complexity for unstructured search*

As for the 'quantum data base search' discussed in the lecture we consider the problem of the 'classical data base search' in the following form:

> *Given an oracle O_f for some function $f : \{0,1\}^n \to \{0,1\}$ and assuming we are promised that exactly one input $x = w$ is such that $f(w) = 1$, find w, i.e.*
>
> $$f(x) = \begin{cases} 1, & \text{if } x = w, \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

The name 'unstructured search' says that there is no inherent structure in the data one could make use of to accelerate the algorithm. In complexity theory it is often asked what is the minimal number of queries to an oracle necessary to achieve a certain task – this is called the query complexity.

(a) If we are interested in a deterministic classical algorithm determining w, how often does this algorithm have to query the oracle in the worst case? Be precise in the argumentation.

(b) Assume we consider (not necessarily deterministic) algorithms whose output is w only with success probability p. What is the worst case number of queries?

(c) Conclude that a classical algorithm solving the search problem has to consult the oracle at least $\Omega(N)$ times.[1]

Solution.

(a) We note a few crucial observations before combining them into a simple stringent argument. First and foremost, the only way a classical algorithm can have success is if it inputs w into the oracle and thereby observes the outcome $f(w) = 1$. Second, an algorithm that is said to solve the problem has to find the marked item w for all possible functions f. Third, the order according to which the possible inputs $x \in \{0,1\}^n$ are checked cannot be faster or slower than an other one in general. For any given order there is a function f whose marked item is the last one to be checked. Finally, there are no shortcuts because it is assumed that there is no inherent structure in the data (unstructured search). Otherwise one could imagine that knowing that x_1 and x_2 are not the marked item one can conclude that x_3 is not marked either, hence saving one query.

We conclude that any deterministic algorithm has to specify an order according to which the different inputs x are input into the oracle. It can stop if and only if it found w, i.e. an output of 1, or if it checked all but the last possible input, which can then be concluded to be the marked one. For any checking order, there is a function f that marks the last input of that order, hence requiring $2^n - 1 \equiv N - 1$ queries.

(b) We build on the previous argument. In a nondeterministic algorithm the order may not predefined. It is possible that from time to time, or always, the next input is chosen at random. However, the previous argument can be extended to that case: any order,

[1] For two functions h, g one says $h = \Omega(g)$, 'h is Big Omega of g', if $\exists c > 0, \exists N_0$ s.t. $\forall N > N_0$: $h(N) \geq c \cdot g(N)$.

whether predetermined or random, can encounter a function f that shows the output 1 only in the last possible input.

If the success probability is lowered from one to $0 \leq p \leq 1$ the game changes, but only slightly. An algorithm still has to check almost all possible inputs. Suppose it checked $N-k$ of the N possible inputs without finding w. The probability to guess correctly which one of the k remaining item is marked is $\frac{1}{k}$. Hence, if the success probability has to be at least p, it may happen that $N - \lfloor \frac{1}{p} \rfloor$ items have to be checked.[2]

(c) We conclude that any classical algorithm for the unstructured search with success probability p, whether deterministic or randomized, has to query the oracle $N - \lfloor \frac{1}{p} \rfloor$ times in the worst case. This function is Big Omega of N, $\Omega(N)$ (see definition in the above footnote).

Exercise 2. ***Optimality of Grover's search algorithm in quantum computation***

We have seen in the lecture that Grover's search algorithm consults the oracle only $O(\sqrt{N})$ times[3] ($N = 2^n$). In this exercise we show that no quantum algorithm can perform this task using fewer than $\Omega(\sqrt{N})$ queries, hence Grover's algorithm is optimal. For simplicity we assume that there is a unique solution. The oracle is then described by the unitary operation $\tilde{U}_w = \mathbb{1} - 2|w\rangle\langle w|$, as encountered in the lecture.[4]

Suppose the algorithm starts in a state $|\psi_0\rangle$ and applies the oracle $\tilde{U}_w$ exactly k times, interleaved with unitary operations[5] $U_1, \ldots, U_k$. Define

$$|\psi_k^w\rangle = U_k \tilde{U}_w U_{k-1} \tilde{U}_w \cdots U_1 \tilde{U}_w |\psi_0\rangle, \qquad \textit{(state with oracle operations)} \tag{2}$$

$$|\psi_k\rangle = U_k U_{k-1} \cdots U_1 |\psi_0\rangle, \qquad \textit{(state without oracle operations)} \tag{3}$$

and define the deviation after k steps caused by the oracle as

$$D_k = \sum_w \left\| |\psi_k^w\rangle - |\psi_k\rangle \right\|^2. \tag{4}$$

If D_k is small there is only a small difference between $|\psi_k^w\rangle$ and $|\psi_k\rangle$ and it is not possible to correctly identify w with high probability.

(a) Using Eqs. (5.1) and (6), show that $D_k \leq 4k^2$ by induction.

(b) Assume that for all possible functions f, i.e. all possible w, an observation yields a solution to the search with probability at least $1/2$. This is, $|\langle w|\psi_k^w\rangle|^2 \geq \frac{1}{2}$ for all w. Furthermore, assume[6] $\langle w|\psi_k^w\rangle = |\langle w|\psi_k^w\rangle|$. Using Eqs. (5.2) and (7), show that in this case $D_k \geq cN$ for some c and sufficiently large N.

Together these two points prove that $k = \Omega(\sqrt{N})$ if the algorithm is to succeed, hence any quantum algorithm solving the search problem has to query the oracle at least $\Omega(\sqrt{N})$ times.

Hints:

[2] If the algorithm checked only $N - \lfloor \frac{1}{p} \rfloor - 1$ or less the probability of guessing w correctly would be $p' = \frac{1}{\lfloor \frac{1}{p} \rfloor + 1} < p$.

[3] As a reminder: $h = O(g)$, 'h is Big O of g', if $\exists c > 0, \exists N_0$ s.t. $\forall\, N > N_0$: $h(N) \leq c \cdot g(N)$.

[4] The subscript of the oracle is chosen to be w instead of f (as it was done in the lecture) because this simplifies the notation later and because f is completely defined by w.

[5] In the Grover algorithm the U_k are all equal to $H^{\otimes n} \tilde{U}_0 H^{\otimes n}$, where $\tilde{U}_0$ is the new notation for $\tilde{U}_{f_0}$.

[6] Replacing $|w\rangle$ with $e^{i\theta}|w\rangle$ does not change the probability of success, so w.l.o.g. we may assume that $\langle w|\psi_k^w\rangle = |\langle w|\psi_k^w\rangle|$.

(i) The Cauchy-Schwarz inequality is helpful in various steps of this exercise.

(ii) For any two vectors a, b in a Hilbert space $\mathcal{H}$, show that

$$\|a+b\|^2 \leq \|a\|^2 + \|b\|^2 + 2\|a\|\|b\| \quad \textit{and} \quad \|a+b\|^2 \geq \|a\|^2 + \|b\|^2 - 2\|a\|\|b\|. \tag{5}$$

(iii) Let $\{a_i\}_{i=0}^{N-1}$ be an orthonormal basis of an N-dimensional Hilbert space $\mathcal{H}$ with inner product $(\cdot\,,\cdot)$ and $b \in \mathcal{H}$ normalized. Then

$$\sum_i |(a_i, b)|^2 = 1. \tag{6}$$

(iv) Same setting as in (iii) with $N = \dim(\mathcal{H})$, show that

$$\sum_i \|b - a_i\|^2 \geq 2N - 2\sqrt{N}. \tag{7}$$

Solution.

(a) Since $|\psi_0^w\rangle = |\psi_0\rangle$ the claim is clearly true for $k = 0$. For any k, notice that

$$D_{k+1} = \sum_w \left\| U_k \tilde{U}_w |\psi_k^w\rangle - U_k |\psi_k\rangle \right\|^2 = \sum_w \left\| \tilde{U}_w |\psi_k^w\rangle - |\psi_k\rangle \right\|^2 \tag{S.1}$$

$$= \sum_w \left\| \tilde{U}_w \left(|\psi_k^w\rangle - |\psi_k\rangle\right) + \left(\tilde{U}_w - \mathbb{1}\right)|\psi_k\rangle \right\|^2 \tag{S.2}$$

$$\overset{(5.1)}{\leq} \sum_w \left(\left\| \tilde{U}_w \left(|\psi_k^w\rangle - |\psi_k\rangle\right) \right\|^2 + 2 \left\| \tilde{U}_w \left(|\psi_k^w\rangle - |\psi_k\rangle\right) \right\| \left\| \left(\tilde{U}_w - \mathbb{1}\right)|\psi_k\rangle \right\| + \left\| \left(\tilde{U}_w - \mathbb{1}\right)|\psi_k\rangle \right\|^2 \right)$$

$$= \sum_w \left(\left\| |\psi_k^w\rangle - |\psi_k\rangle \right\|^2 + 4 \left\| |\psi_k^w\rangle - |\psi_k\rangle \right\| \left| \langle w|\psi_k\rangle \right| + 4 \left| \langle w|\psi_k\rangle \right|^2, \right) \tag{S.3}$$

where we used $\left(\tilde{U}_w - \mathbb{1}\right)|\psi_k\rangle = -2\langle w|\psi_k\rangle|w\rangle$ in the last line as well as the fact that unitary operations leave the norm of a vector invariant. By Eq. (6) and Cauchy-Schwarz we obtain

$$\begin{aligned} D_{k+1} &\leq D_k + 4 \left(\sum_w \left\| |\psi_k^w\rangle - |\psi_k\rangle \right\|^2 \right)^{1/2} \left(\sum_w \left| \langle w|\psi_k\rangle \right|^2 \right)^{1/2} + 4 \\ &= D_k + 4\sqrt{D_k} + 4 \end{aligned} \tag{S.4}$$

The induction hypothesis tells us $D_k \leq 4k^2$, hence we obtain

$$D_{k+1} \leq 4k^2 + 8k + 4 = 4(k+1)^2, \tag{S.5}$$

which completes the first part of the proof.

(b) By $|\langle w|\psi_k^w\rangle|^2 \geq \frac{1}{2}$ for all w and $\langle w|\psi_k^w\rangle = |\langle w|\psi_k^w\rangle|$,

$$\left\| |\psi_k^w\rangle - |w\rangle \right\|^2 = 2 - 2\left|\langle w|\psi_k^w\rangle\right| \leq 2 - \sqrt{2}. \tag{S.6}$$

Define now $E_k = \sum_w \left\| |\psi_k^w\rangle - |w\rangle \right\|^2$ and $F_k = \sum_w \left\| |\psi_k\rangle - |w\rangle \right\|^2$. By the above: $E_k \leq (2 - \sqrt{2})N$; and by Eq. (7): $F_k \geq 2N - 2\sqrt{N}$. We are now in the position to prove an asymptotic lower bound on D_k:

$$D_k = \sum_w \left\| \left(|\psi_k^w\rangle - |w\rangle\right) + \left(|w\rangle - |\psi_k\rangle\right) \right\|^2 \tag{S.7}$$

$$\overset{(5.2)}{\geq} \sum_w \left\| |\psi_k^w\rangle - |w\rangle \right\|^2 - 2 \sum_w \left\| |\psi_k^w\rangle - |w\rangle \right\| \left\| |w\rangle - |\psi_k\rangle \right\| + \sum_w \left\| |w\rangle - |\psi_k\rangle \right\|^2 \tag{S.8}$$

$$= E_k + F_k - 2 \sum_w \left\| |\psi_k^w\rangle - |w\rangle \right\| \left\| |w\rangle - |\psi_k\rangle \right\|. \tag{S.9}$$

Using again Cauchy-Schwarz the last term can be bounded by $2\sqrt{E_k F_k}$, hence using the bounds for F_k and E_k

$$D_k \geq E_k + F_k - 2\sqrt{E_k F_k} = \left(\sqrt{F_k} - \sqrt{E_k}\right)^2 \geq cN \tag{S.10}$$

for some c and large enough N.

Comment 1: If there are M solutions to the search problem it can be shown analogously that any successful quantum algorithm needs at least $\Omega(\sqrt{N/M})$ queries. Also in this case Grover's algorithm is optimal (up to constant factors in the number of queries).

Comment 2: Notice that it was not crucial to request $|\langle w|\psi_k^w\rangle|^2 \geq \frac{1}{2}$ for all w with probability bound 1/2. It could have been any strictly positive bound p (smaller than 1, of course).