Introduction to Quantum Computing

Optimality of Grover's algorithm

Silvio Martinico

June 02, 2022

Abstract

In this paper we prove the optimality of the Grover's Quantum Search algorithm. First of all, we introduce the Grover's algorithm and the corresponding quantum circuit; then, we analyze its computational complexity and its probability of success. Finally, we proceed with the proof of the optimality and we make some considerations about its complexity.

1 Grover's algorithm

One of the most relevant quantum algorithms is the Grover's quantum search algorithm. This algorithm was devised to solve the problem described below.

1.1 The problem

Let $X := \{0, ..., N-1\}$; we assume for simplicity that $N = 2^n$ (otherwise we can just add ancillary elements until we get a power of two).
Let $O : X \longrightarrow \{0, 1\}$ be a function such that $\exists!\ \beta \in X \mid O(\beta) = 1$. We will refer to this function as **oracle**.

We want to know the identity of β. In classical computation this requires on average $\frac{N}{2}$ calls to the oracle and N in the worst case (we have to check every element of X since the problem has no structure we can exploit).

In 1996 Lov Grover found a way to perform this task in $O(\sqrt{N})$ calls to the oracle, which means a quadratic speedup. Subsequently, it was shown that no quantum algorithm can solve this problem in less than $\Omega(\sqrt{N})$ calls to the oracle and, therefore, that Grover's Algorithm is asymptotically optimal.

1.1.1 Grover's algorithm as search algorithm

Grover's algorithm is often addressed as a database search algorithm. Let's see how we can exploit it in order to perform this task.

Let $D = [x_1, ..., x_N]$ be an unstructured database and let x be an element we want to search in D. Assume we have an oracle O_x such that $O_x(k) = 1 \iff D[k] = x$. In this way, we can find the position of x inside D querying the oracle. Again, this requires on average $\frac{N}{2}$ calls to the oracle with classical computation.

So, if we knew how to build such an oracle and how to handle classical data on a quantum computer (or how to store quantum data), we could again exploit Grover's algorithm to solve this problem in $\Omega(\sqrt{n})$ calls to the oracle. Another issue is that we have to build a different oracle for each x we want to search for, so this operation must be very cheap in order to have a real advantage over classical computation.

1.2 The algorithm

The idea of Grover's algorithm can be summarized in these three steps:

i Create a superposition of all the computational basis states;

ii Amplify the amplitude of $|\beta\rangle$ so that we will measure it with very high probability;

iii Measure the state in order to find $|\beta\rangle$.

The first point is immediate: just take $|0\rangle^{\otimes n}$ as initial state and apply an Hadamard gate to each qubit; the resultant state will be

$$|\varphi\rangle := H^{\otimes n} |0\rangle^{\otimes n} .$$

From now on, we will indicate with $|x\rangle$, where $x \in \{0, ..., N-1\}$, a state of the computational basis of $\mathbb{C}^N$; so, when there is no ambiguity, we write just $|0\rangle$ instead of $|0\rangle^{\otimes n}$.

Let's see how to implement the second step.
Given a computational basis state $|\beta\rangle$, we define first a fundamental gate for this algorithm:

$$U_\beta \;:=\; \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & -1 & & & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix} \;=\; I - 2\,|\beta\rangle\,\langle\beta|\;\;,$$

where I is the N-dimensional identity matrix. We have $U_\beta\,|x\rangle = (-1)^{O(x)}\,|x\rangle\;\;\forall\; x \in \{0, ..., N-1\}$.
We can decompose the superposition of all computational basis states $|\varphi\rangle$ as:

$$|\varphi\rangle \;=\; \underbrace{\frac{1}{\sqrt{N}} \sum_{\substack{x=0 \\ x\neq\beta}}^{N-1} |x\rangle}_{|\widetilde{\psi}\rangle} \;+\; \frac{1}{\sqrt{N}}\,|\beta\rangle\;\;.$$

If we define $|\psi\rangle := \frac{|\widetilde{\psi}\rangle}{|||\widetilde{\psi}\rangle||}$, it is clear that $|\psi\rangle$ and $|\beta\rangle$ are orthogonal states and that we can write

$$|\varphi\rangle \;=\; \frac{\sqrt{N-1}}{\sqrt{N}}\,|\psi\rangle \;+\; \frac{1}{\sqrt{N}}\,|\beta\rangle\;\;.$$

Since $|\varphi\rangle$ is a unitary vector, there exists an angle θ such that $\frac{\sqrt{N-1}}{\sqrt{N}} = \cos\frac{\theta}{2}$ and $\frac{1}{\sqrt{N}} = \sin\frac{\theta}{2}$, which implies the following equality:

$$\frac{\theta}{2} = \arcsin\frac{1}{\sqrt{N}} \tag{1}$$

If we apply U_β to the superposition of all computational basis states, what happens is that the relative phase of $|\beta\rangle$ is flipped:

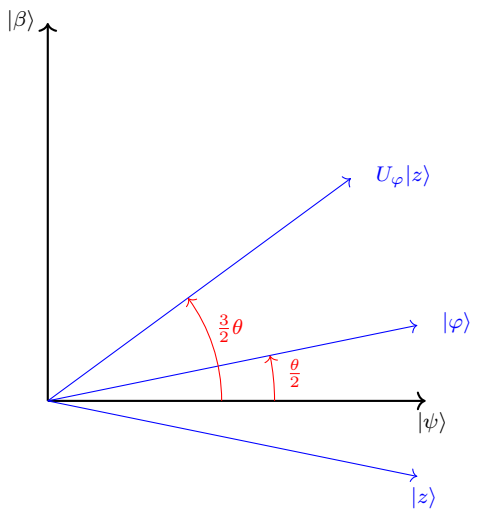
$$|z\rangle \;:=\; U_\beta\,|\varphi\rangle \;=\; \cos\frac{\theta}{2}\,|\psi\rangle \;-\; \sin\frac{\theta}{2}\,|\beta\rangle\;\;.$$

So, if we look at $|\beta\rangle$ and $|\psi\rangle$ as two orthogonal axis, $|\varphi\rangle$ is a vector between them (it is a positive linear combination of $|\beta\rangle$ and $|\psi\rangle$. In this setting, the gate U_β just reflects $|\varphi\rangle$ with respect to the axis $|\psi\rangle$. At this point we define the reflector around $|\varphi\rangle$ as follows:

$$U_\varphi \;:=\; 2\,|\varphi\rangle\,\langle\varphi| - I\;\;.$$

Applying this gate to $|z\rangle$ we make $|\varphi\rangle^{(1)} := U_\varphi\,|z\rangle$ closer to $|\beta\rangle$.

In the image below we can see what happens geometrically.



By iterating the process above a certain number of times k, we make the amplitude of $|\beta\rangle$ grow as much as possible. Passing from the iteration j to the iteration $j+1$, the angle of $|\varphi\rangle^{(j+1)}$ grows of θ. At the k-th step the angle of $|\varphi\rangle^{(k)}$ will be $\theta_k = (2k+1)\frac{\theta}{2}$.
Now the question is, how many Grover's iterations we want to do? We want to iterate until $\theta_k \leq \frac{\pi}{2}$.

$$\theta_k \leq \frac{\pi}{2} \iff k\theta + \frac{\theta}{2} \leq \frac{\pi}{2} \iff k \leq \frac{\pi}{2\theta} - \frac{1}{2} \,.$$

Now we recall that $\frac{\theta}{2} = \arcsin\frac{1}{\sqrt{N}}$ and, when $N \to \infty$, $\frac{1}{\sqrt{N}} \to 0$, so we can write the Taylor series of the arcsin function:

$$\frac{\theta}{2} \;=\; \arcsin\frac{1}{\sqrt{N}} \;=\; \frac{1}{\sqrt{N}} + \frac{1}{6\sqrt{N^3}} + \cdots \;=\; \frac{1}{\sqrt{N}} + o\left(\frac{1}{N}\right) \;\implies\; \theta = \frac{2}{\sqrt{N}} + o\left(\frac{1}{N}\right) \,.$$

We want that $k \leq \left(\frac{\pi}{2} - \frac{\theta}{2}\right)\frac{1}{\theta}$ and k must be an integer, so the best choice for it is to take:

$$k = \left\lfloor \left(\frac{\pi}{2} - \frac{\theta}{2}\right)\frac{1}{\theta} \right\rfloor \,.$$

Let's do some rewriting in order to give an estimation of k in terms of N:

$$\left(\frac{\pi}{2} - \frac{\theta}{2}\right)\frac{1}{\theta} \;=\; \frac{\pi}{2\theta} - \frac{1}{2} \;=\; \frac{\pi}{2\left(\frac{2}{\sqrt{N}} + o\left(\frac{1}{N}\right)\right)} - \frac{1}{2}$$

$$= \frac{\pi}{4}\left(\frac{1}{\frac{1}{\sqrt{N}}\left(1 + o\left(\frac{1}{N}\right)/\frac{1}{\sqrt{N}}\right)}\right) - \frac{1}{2} \;=\; \underbrace{\frac{\pi\sqrt{N}}{4}\left(\frac{1}{1 + o\left(\frac{1}{N}\right)\sqrt{N}}\right) - \frac{1}{2}}_{h(N)}$$

Now we can state two things:

- $h(N) \approx \frac{\pi}{4}\sqrt{N}$, i.e. $\frac{h(N)}{\frac{\pi}{4}\sqrt{N}} \overset{N\to\infty}{\longrightarrow} 1$, (since $o\left(\frac{1}{N}\right)\sqrt{N} \to 0$);
- $h(N) \leq \frac{\pi}{4}\sqrt{N}$ since the o-notation is hiding the Taylor expansion of $\arcsin\frac{1}{\sqrt{N}}$, which has only "+" signs and $\sqrt{N} > 0$; furthermore, it hides some positive small constant from the steps above.

At this point, we have all is needed to define the circuit for the Grover's algorithm and to state the following theorem:

Theorem 1.1. If we perform $k = \left\lfloor \left(\frac{\pi}{2} - \frac{\theta}{2}\right)\frac{1}{\theta} \right\rfloor$ Grover's iterations, the probability of measuring the target state $|\beta\rangle$ from $|\varphi\rangle^{(k)}$ is:

$$\mathbb{P}(|\beta\rangle) \;\geq\; 1 - \frac{2}{N} + o\left(\frac{1}{N}\right).$$

Proof:

$$\mathbb{P}(|\beta\rangle) \;=\; \left|\langle\beta|\varphi\rangle^{(k)}\right|^2 \;=\; \sin^2\left((2k+1)\frac{\theta}{2}\right) \;=\; \sin^2\underbrace{\left(\left(2\left\lfloor\frac{\pi}{2\theta} - \frac{1}{2}\right\rfloor + 1\right)\frac{\theta}{2}\right)}_{arg}.$$

Now, removing the "floor" operator, we can bound arg:

$$\frac{\pi}{2} - \frac{2}{\sqrt{N}} - \frac{1}{N} \;\leq\; \frac{\pi}{2} - \theta \;=\; \left(2\left(\frac{\pi}{2\theta} - \frac{1}{2} - 1\right) + 1\right)\frac{\theta}{2} \;<\; arg \;\leq\; \left(2\left(\frac{\pi}{2\theta} - \frac{1}{2}\right) + 1\right)\frac{\theta}{2} \;=\; \frac{\pi}{2}.$$

Observing that $\sin x$ in increasingly monotone in this interval, we can then derive:

$$\sin^2\left(\frac{\pi}{2} - \frac{2}{\sqrt{N}} - \frac{1}{N}\right) \;<\; \sin^2(arg) \;\leq\; \sin^2\frac{\pi}{2} \;=\; 1. \tag{2}$$

Finally, using the fact that $\sin\left(\frac{\pi}{2} - x\right) = \cos x$ and expanding $\cos x$ we get:

$$\sin^2\left(\frac{\pi}{2} - \frac{2}{\sqrt{N}} - \frac{1}{N}\right) \;=\; \cos^2\left(\frac{2}{\sqrt{N}} + \frac{1}{N}\right) \;=\; \left(1 - \frac{2}{N} + o\left(\frac{1}{N}\right)\right)^2,$$

which, together with the inequality in (2), gives us the thesis:

$$1 - \frac{4}{N} + o\left(\frac{1}{N}\right) \;<\; \mathbb{P}(|\beta\rangle) \;\leq\; 1$$

□

Recalling that $|\varphi\rangle$ is the uniform superposition of all the computational basis states, we do a last observation about the writing of the diffuser U_φ:

$$U_\varphi \;=\; 2\,|\varphi\rangle\langle\varphi| - I \;=\; H^{\otimes n}\Big(2\,|0\rangle\langle 0| - I\Big)H^{\otimes n}$$

We can finally draw the circuit for the Grover's algorithm:

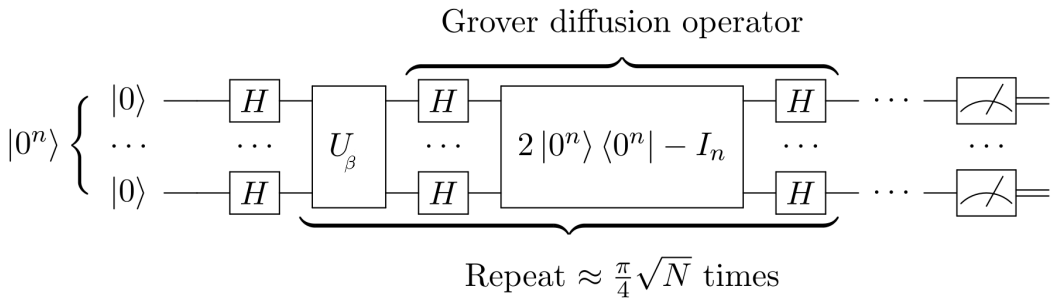


Figure 1: Circuit for the Grover's algorithm

2 Optimality of Grover's algorithm

We now prove the optimality of the algorithm. The idea of the following proof comes from [1].
Assume that the algorithm starts from a state $|\psi\rangle$ and applies the oracle (that we will call O_β in this section, in order to avoid confusing it with other unitary transformations) k times, possibly alternating it with unitary transformations $U_1, ..., U_k$. Therefore we define:

$$|\psi_k^\beta\rangle := U_k O_\beta U_{k-1} O_\beta \, ... \, U_1 O_\beta \, |\psi\rangle \ , \qquad |\psi_k\rangle := U_k U_{k-1} \, ... \, U_1 \, |\psi\rangle \ .$$

We want to find a relation between k (the number of calls to the oracle) and N. In order to do this, we will bound the following quantity:

$$D_k := \sum_{x=0}^{N-1} \Big|\Big| \, |\psi_k^x\rangle - |\psi_k\rangle \, \Big|\Big|^2 .$$

The idea is that D_k gives us a measure of the deviation caused by the oracle.
From now on we will omit the "ket" notation for simplicity.

Lemma 2.1. $D_k = O(k^2)$. More precisely, it holds: $D_k \leq 4k^2$.

Proof:
The proof proceeds by induction on k.

- **Basic step**: This is obvious since $D_0 = 0$;
- **Inductive step**: Let's see how to exploit the inductive hypothesis (i.h.).

$$D_{k+1} \ = \ \sum_{x=0}^{N-1} \Big|\Big|\psi_{k+1}^x - \psi_{k+1}\Big|\Big|^2 \ = \ \sum_{x=0}^{N-1} \Big|\Big|U_{k+1}O_x\psi_k^x - U_{k+1}\psi_k\Big|\Big|^2$$

$$\overset{(*)}{=} \ \sum_{x=0}^{N-1} \Big|\Big|O_x\psi_k^x - \psi_k\Big|\Big|^2 \ = \ \sum_{x=0}^{N-1} \Big|\Big|O_x(\psi_k^x - \psi_k) + (O_x - I)\psi_k\Big|\Big|^2$$

$$\overset{\text{t.i.}}{\leq} \ \sum_{x=0}^{N-1} \Big|\Big|\psi_k^x - \psi_k\Big|\Big|^2 \ + \ \sum_{x=0}^{N-1} \Big|\Big|(O_x - I)\psi_k\Big|\Big|^2 \ + \ 2\sum_{x=0}^{N-1} \Big|\Big|\psi_k^x - \psi_k\Big|\Big| \cdot \Big|\Big|(O_x - I)\psi_k\Big|\Big| ,$$

where the step $(*)$ is implied by the fact that U_{k+1} is a unitary matrix, while the inequality follows from the triangular inequality (t.i.).

Now, at the left-hand side, the first term is D_k. Furthermore, $(O_x - I)\psi_k = -2\langle x|\psi_k\rangle |x\rangle$, so the second term becomes the sum of the squares of the components of $|\psi_k\rangle$ multiplied by 4, i.e. 4 times the square of the norm of $|\psi_k\rangle$ (which is equal to 1).

Using the observations above and then using the Cauchy-Schwartz inequality (C-S), we get:

$$D_{k+1} \;\leq\; D_k \;+\; 4 \;+\; 4\sum_{x=0}^{N-1} \Big|\Big|\psi_k^x - \psi_k\Big|\Big| \cdot \Big|\Big| \langle x|\psi_k\rangle |x\rangle \Big|\Big|$$

$$\overset{\text{C-S}}{\leq} \; D_k \;+\; 4 \;+\; 4 \underbrace{\sqrt{\sum_{x=0}^{N-1} \Big|\Big|\psi_k^x - \psi_k\Big|\Big|^2}}_{\sqrt{D_k}} \cdot \underbrace{\sqrt{\sum_{x=0}^{N-1} \Big|\Big| \langle x|\psi_k\rangle |x\rangle \Big|\Big|^2}}_{||\psi_k||=1}$$

$$= D_k \;+\; 4 \;+\; 4\sqrt{D_k} \;\overset{\text{i.h.}}{\leq}\; 4(k^2+2k+1) \;=\; 4(k+1)^2$$

□

Lemma 2.2. Assume that we want to measure $|x\rangle$ from $|\psi_k^x\rangle$ with probability at least $\frac{1}{2}$, i.e. $|\langle x|\psi_k^x\rangle| \geq \frac{1}{2}\ \forall x$. Then D_k can't grow slower than N. More precisely $D_k \geq cN$ for a certain constant $c \in \mathbb{R}^+$.

Proof:

Let's estimate the distance between ψ_k^x and x assuming, without loss of generality, that $\langle x|\psi_k^x\rangle = \big|\langle x|\psi_k^x\rangle\big|$ (we can just multiply x by a total phase, which does not change the probability of measuring it):

$$\big|\big|\psi_k^x - x\big|\big|^2 \;=\; \big|\big|\psi_k^x\big|\big|^2 + \big|\big|x\big|\big|^2 - 2\underbrace{\big|\langle x|\psi_k^x\rangle\big|}_{\geq \frac{1}{\sqrt{2}}} \;\leq\; 2 - \sqrt{2}\ .$$

This implies the following inequality:

$$E_k := \sum_{x=0}^{N-1} \big|\big|\psi_k^x - x\big|\big|^2 \;\leq\; (2-\sqrt{2})N\ . \tag{3}$$

Now, defining $F_k := \sum_{x=0}^{N-1} \big|\big|x - \psi_k\big|\big|^2$, we get:

$$D_k \;=\; \sum_{x=0}^{N-1} \big|\big|(\psi_k^x - x) + (x - \psi_k)\big|\big|^2$$

$$\stackrel{\text{t.i.}}{\geq} \underbrace{\sum_{x=0}^{N-1} \big|\big|\psi_k^x - x\big|\big|^2}_{E_k} \;+\; \underbrace{\sum_{x=0}^{N-1} \big|\big|x - \psi_k\big|\big|^2}_{F_k} \;-\; 2\sum_{x=0}^{N-1} \big|\big|\psi_k^x - x\big|\big| \cdot \big|\big|x - \psi_k\big|\big| \qquad (4)$$

$$\stackrel{\text{C-S}}{\geq} \; E_k \;+\; F_k \;-\; 2\sqrt{E_k F_k} \;=\; \big(\sqrt{F_k} - \sqrt{E_k}\big)^2$$

In order to conclude the proof, we show that $F_k \geq 2N - 2\sqrt{N}$.

$$\big|\big|x - \psi_k\big|\big|^2 \;=\; \big|\big|x\big|\big|^2 + \big|\big|\psi_k\big|\big|^2 - 2\,\langle x|\psi_k\rangle \;=\; 2 - 2\,\langle x|\psi_k\rangle \;,$$

so $F_k \;=\; 2N - 2\sum_{x=0}^{N-1} \langle x|\psi_k\rangle$ and this quantity is $\geq 2N - 2\sqrt{N} \iff \sum_{x=0}^{N-1} \langle x|\psi_k\rangle \leq \sqrt{N}$, which follows from:

$$\sum_{x=0}^{N-1} \langle x|\psi_k\rangle \;\leq\; \sum_{x=0}^{N-1} \big|\,\langle x|\psi_k\rangle\,\big| \;=\; \big|\big|\psi_k\big|\big|_1 \;\leq\; \sqrt{N}\big|\big|\psi_k\big|\big|_2 \;=\; \sqrt{N}\;.$$

Finally, putting together this last inequality with (3) and (4) and observing that $\sqrt{F_k} - \sqrt{E_k} \geq 0$, we have:

$$D_k \;\geq\; \left(\sqrt{F_k} - \sqrt{E_k}\right)^2 \;\geq\; \left(\sqrt{2N - 2\sqrt{N}} - \sqrt{(2-\sqrt{2})N}\right)^2$$

$$=\; 2N - 2\sqrt{N} + 2N - \sqrt{2}N - 2\sqrt{(2N - 2\sqrt{N})(2-\sqrt{2})N}$$

$$=\; (4-\sqrt{2})N - 2\sqrt{N} - 2\sqrt{2(2-\sqrt{2})}N\left(\sqrt{1 - \frac{1}{\sqrt{N}}}\right)$$

$$\geq\; (4-\sqrt{2})N - 2\sqrt{N} - 2\sqrt{2(2-\sqrt{2})}N$$

$$=\; \left(4 - \sqrt{2} - 2\sqrt{4 - 2\sqrt{2}}\right)N - 2\sqrt{N} \;\geq\; c \cdot N\;,$$

where c is a constant ($0 < c < 4 - \sqrt{2} - 2\sqrt{4-2\sqrt{2}} \approx 0,42$) and, the larger N is, the closer c can be to the upper bound).

□

Finally, we have all that is needed to prove the optimality of Grover's algorithm.

Theorem 2.1. The Grover's algorithm is asymptotically optimal.

Proof:
From Lemma 2.1 and Lemma 2.2 we have the following inequalities:

$$cN \ \leq \ D_k \ \leq \ 4k^2 \ ,$$

which imply:

$$k \ \geq \ \frac{\sqrt{c}}{2}\sqrt{N} \ .$$

This means that $k = \Omega(\sqrt{N})$, i.e. we need to do at least $\frac{\sqrt{c}}{2}\sqrt{N}$ calls to the oracle in order to find the target state and therefore, the Grover's algorithm is asymptotically optimal.

□

2.1 Considerations about the result

If on one hand the optimality of the Grover's algorithm is a great result, on the other it is almost disappointing. The hope, in fact, was to have an exponential speed-up instead of just a quadratic one. Having a quantum algorithm that search an unstructured space of size N in $O(\log N)$ oracle calls means that we could solve NP-complete problems efficiently by just searching trough all the possible solutions. So, the optimality of Grover's algorithm breaks all hopes about a fast naive approach for attacking NP-complete problems with quantum computers. However, it can be used to quadratically speed-up the solution of problems in NP.

Due to what we said above, it is believed that $NPC \not\subset \mathbf{BQP}$. BQP is defined as the class of problems that can be solved in polynomial time with probability of error $\leq \frac{1}{3}$ on every input by a quantum computer. This class is taken as the formalization of the class of problems efficiently solvable by quantum computers. The error probability can be reduced effectively to a much smaller constant $\varepsilon > 0$ by just running the computation $O\big(\log(\frac{1}{\varepsilon})\big)$ times and taking the answer given by these runs more frequently.
For more insights on these considerations, see [2] and [1].

References

[1] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[2] Ronald de Wolf. *Quantum Computing: Lecture Notes*. 2019.