Grover Search for Inverting a Quantum Function

January 7, 2020

This lemma was initially part of a "quantum slowdown" lemma, but it has no use without a corresponding "quantum speedup" step which I've been unable to prove. I wrote it because it may(?) be of independent interest and because I found writing out the details to be a useful exercise. This writeup has *not* been edited. If you find errors, please let me know.

Lemma 1 (Inverting a quantum function with classical input)**.** *Given a $t(n)$-time* BQP *algorithm $\mathcal{A}(|x\rangle, y) : \mathbb{C}^{2^n} \times \{0,1\}^m \to \{0,1\}$ that takes as input a classical input $y \in \{0,1\}^m$ and a possibly quantum n-qubit input $|x\rangle \in \mathbb{C}^{2^n}$, there exists a quantum algorithm $\mathcal{B}$ running in time $O(2^{\frac{m}{2}} t)$ that, given $O(m)$ unentangled copies of $|x\rangle$,*

- *Outputs* 1 *with probability at least $\frac{2}{3}$ if there exists a y such that* $\Pr[\mathcal{A}(|x\rangle, y) = 1] \geq \frac{2}{3}$
- *Outputs* 0 *with probability at least $\frac{2}{3}$ if for all y* $\Pr[\mathcal{A}(|x\rangle, y) = 1] \leq \frac{1}{3}$.

We'd like to use Grover search to remove the quantifier. The immediate issue that arises when we make f a quantum algorithm is that of entanglement. Part of vanilla Grover's algorithm involves implementing the following sequence of transformations:

$$\begin{aligned} |y\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) |0\rangle &\mapsto |y\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) |f_x(y)\rangle && \text{Reversibly computing } f_x \\ &\mapsto |(-1)^{f_x(y)} y\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) |f_x(y)\rangle && \text{XORing with the first ancilla and moving the sign} \\ &\mapsto |(-1)^{f_x(y)} y\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) |0\rangle && \text{Uncomputing.} \end{aligned}$$

Let's refer to the qubits of $|y\rangle$ as the "target qubits", the qubit containing $\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)$ as the "first ancilla qubit", and the qubit containing $|f_x(y)\rangle$ as the "function qubit". The uncomputation is critical because it allows us to avoid entangling the function qubit with the target qubits. This means that any downstream computation on the target qubits - in particular, the operation $2|u\rangle\langle u| - I$ - may safely ignore all the ancilla qubits altogether as they are the same for every state of the target qubits. Without this, we run into issues with amplitude cancellation, as states of the target qubits that "want" to cancel cannot because they differ in the ancilla bits. Uncomputation also allows us to reuse these qubits later, but this is of secondary importance.

The uncomputation in the vanilla case works cleanly because $f_x(y)$ is either $|0\rangle$ or $|1\rangle$, because in this case we don't need to worry about entangling the function qubit with the first ancilla qubit. More explicitly (omitting the target qubits),

$$\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)|1\rangle \xmapsto{\text{XOR}} \left(\frac{|1\rangle - |0\rangle}{\sqrt{2}}\right)|1\rangle \tag{1}$$

$$\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)|0\rangle \xmapsto{\text{XOR}} \left(\frac{|1\rangle - |0\rangle}{\sqrt{2}}\right)|0\rangle . \tag{2}$$

The point is that the state at the end remains separable and we can apply uncomputation on just the function qubits.

When f is a quantum function, however, this is no longer true. We now have $|f_{|x\rangle}(y)\rangle = a\,|0\rangle + b\,|1\rangle$, and

$$\left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right)(a\,|0\rangle + b\,|1\rangle) \xmapsto{\text{XOR}} \frac{a\,|00\rangle + b\,|11\rangle - a\,|10\rangle - b\,|01\rangle}{\sqrt{2}}, \tag{3}$$

which in general is not separable. When we try to uncompute $f_{|x\rangle}(y)$, we fail because the XOR operation has entangled it with the first ancilla qubit. However, all hope is not lost! The following idea is implicit in [Aar06]. Let $\bar{f}$ denote the *rounded* form of f, defined as

$$\bar{f}(|x\rangle, y) = 1 \iff \Pr[f(|x\rangle, y) = 1] \geq \frac{2}{3}. \tag{4}$$

Proof. Let $r = O(m)$ for constant to be determined.

1. Apply Hadamard gates to m ancilla qubits initially set to $|0\rangle$ and measure as many bits as you need to get a uniformly sampled $k \in [0, \frac{1}{\sin \frac{2}{2^{\frac{m}{2}}}}]$.

2. Apply k modified Grover iterations (details below) using $|x_1\rangle \otimes \cdots \otimes |x_r\rangle$ and the function f to a state initialized to the uniform state on m qubits to obtain a final state $|\alpha\rangle$.

3. Measure $|\alpha\rangle$ in the computational basis to obtain a candidate solution α.

4. Compute $f_{|x_{r+1}\rangle}(\alpha)$ and measure the output qubit. If the output is $|1\rangle$ then return 0 and vice-versa.

When describing the function, we will refer to the m qubits that hold our candidate solution state as the "target qubits". Recall that a Grover iteration for inverting an arbitrary function $g\colon \{0,1\}^m \to \{0,1\}$ consists of two steps:

1. Phase inversion: flipping the sign of basis vectors of the target qubits corresponding to a solution to g.

2. Mean inversion: applying the operation $2\,|u\rangle\langle u| - I$, where $|u\rangle$ is the uniform vector on m qubits.

Mean inversion will be the same as before, but phase inversion - the part that depends on the function being inverted - is different, instead performing an amplified computation of $f_{|x\rangle}(y)$ by making use of the extra copies.

Phase inversion proceeds as follows.

1. Use the r states $\{|x_i\rangle\}$ to compute $f_{|x_i\rangle}(y)$ r times in parallel, but don't measure the individual output qubits.
2. Apply a majority gate to the r output qubits. This will be our "actual" output of the computation of $f_{|x\rangle}(y)$.
3. XOR the output of the majority gate into an ancilla qubit in state $\left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right)$.
4. Uncompute everything involved in evaluating $f_{ketx}(y)$.

Now that we've described the algorithm, let's prove that it works.

For a fixed $|x\rangle$, let b_y denote the "result" of $f_{|x\rangle}(y)$. $b = 1$ when f accepts with probability at least $\frac{2}{3}$ and $b = 0$ in the other case.

Let's start by understanding the way that a single Grover iteration acts on a given basis state $|y\rangle$ of the target qubits. Suppose each of the r parallel runs uses K qubits, and WLOG let the output qubit be the first of these K for each parallel run. Then, $\{|b_{ij}\rangle \otimes |w_{ij}\rangle$ - where each $|w_{ij}\rangle$ is a $(K-1)$-qubit basis vector and $|b_{ij}\rangle$ captures the state of the i^{th} output qubit in this run - is a basis for all the qubits involved in the parallel runs. Note that $|w_{ij}\rangle$ includes $|x_i\rangle$. Because the parallel runs don't talk to each other, we can write the state that results after Step 1 of phase inversion as

$$\prod_{0\leq i\leq r}\left(\sum_{0\leq j\leq 2^K}\alpha_{ij}\,|b_{ij}\rangle\,|w_{ij}\rangle\right), \tag{5}$$

or equivalently as,

$$\sum_{j_1,\ldots,j_r\in[2^K]}\beta_{j_1\ldots,j_r}\,|b_{1j_1}b_{2j_2}\ldots b_{rj_r}\rangle\prod_{i\in[r]}|w_{ij_i}\rangle \tag{6}$$

where $\beta_{j_1,\ldots,j_r} := \alpha_{1j_1}\alpha_{2j_2}\ldots\alpha_{rj_r}$.

Constant completeness error: If $|x\rangle$ were the "good" state from the existential quantifier before, our new good state is $|x\rangle^{\otimes m}$.

Claim 2. *After Step 1 of phase inversion,*

$$\sum_{\substack{j_1,\ldots,j_r\in[2^K]\\ \textit{at least half of } j_i \textit{ are } b_y}}\beta^2_{j_1,\ldots,j_r} > \sqrt{1-\epsilon} \tag{7}$$

Proof. Observe that

$$\begin{aligned}\sum_{j_1,\ldots,j_r\in[2^K]}\beta^2_{j_1,\ldots,j_r}\,|b_{1j_1}b_{2j_2}\ldots b_{rj_r}\rangle\prod_{i\in[r]}|w_{ij_i}\rangle &= \prod_{i\in[r]}\sum_{j\in[2^K]}\alpha^2_{ij}\,|b_{ij}\rangle\,|w_{ij}\rangle\\ &= \prod_{i\in[r]}|0\rangle\left(\sum_{\substack{j\in[2^{K-1}]\\ b_{ij}=0}}\alpha^2_{ij}\,|w_{ij}\rangle\right)+|1\rangle\left(\sum_{\substack{j\in[2^{K-1}]\\ b_{ij}=1}}\alpha^2_{ij}\,|w_{ij}\rangle\right).\end{aligned}$$

We know that at worst for all $i \in [r]$, or equivalently, for each parallel run,

$$\sum_{\substack{j \in [2^{K-1}] \\ b_{ij} = b}} \alpha_{ij}^2 \geq \frac{2}{3} \tag{8}$$

and

$$\sum_{\substack{j \in [2^{K-1}] \\ b_{ij} = \neg b}} \alpha_{ij}^2 \leq \frac{1}{3}. \tag{9}$$

Therefore, the probability we want is actually just

$$\sum_{k \geq \frac{r}{2}} \binom{r}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{r-k}. \tag{10}$$

and by a Chernoff bound this exceeds $\sqrt{1-\epsilon}$. □

Let $|w_0\rangle$ denote the initial state of every workspace qubit (including the $|x_i\rangle$) except for the sign ancilla and the output qubit of the majority gate. Then, Step 1 of phase inversion sends

$$\sum_y \alpha_y |y\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) |0\rangle |w_0\rangle \mapsto \sum_y \alpha_y |y\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) |f_{|x\rangle}(y)\rangle |w_f\rangle . \tag{11}$$

Let $|\phi\rangle := \sum_y \alpha_y |y\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) |f_{|x\rangle}(y)\rangle |w_f\rangle$ and $|\bar{\phi}\rangle := \sum_y \alpha_y |y\rangle \left(\frac{|0\rangle - |1\rangle}{\sqrt{2}}\right) |y_b\rangle |w_f\rangle$. Claim 2 tells us that $|f_{|x\rangle}(y)\rangle$ is very close to $|b_y\rangle$ for every $|y\rangle$. In particular, $|\phi\rangle \langle\bar{\phi}| \geq \sqrt{1-\epsilon}$ and hence

$$\begin{aligned} || \, |\phi\rangle \langle\phi| - |\bar{\phi}\rangle \langle\bar{\phi}| \, || &\leq \sqrt{1 - \langle\phi|\bar{\phi}\rangle^2} \\ &\leq \sqrt{\epsilon} \end{aligned}$$

The next part of this argument is annoying to write up so instead please see this diagram. The point is that we can "almost" uncompute even when our oracle is quantum. In the diagram, $|w'_{y,f}\rangle$ is the state over all qubits except for the output of the majority gate.

$$\begin{tikzcd}
\sum_y \alpha_y |y\rangle \left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) |w_0\rangle \arrow[r, mapsto, "\text{Step 1}"] & \sum_y \alpha_y |y\rangle \left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) |w_{y,f}\rangle \arrow[r, mapsto, "\mathsf{XOR}"] \arrow[d, squiggly, "\text{Trace dist. } \leq\sqrt{\epsilon}"] & |\phi\rangle \arrow[r, mapsto, "\text{Uncompute}"] & |\psi\rangle \\
& \sum_y \alpha_y |y\rangle \left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) |b_y\rangle |w'_{y,f}\rangle \arrow[r, mapsto, "\mathsf{XOR}"] & \sum_y \alpha_y (-1)^{b_y} |y\rangle \left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) |b_y\rangle |w'_{y,f}\rangle \arrow[r, mapsto, "\text{Uncompute}"] \arrow[d, squiggly, "\text{Trace dist. } \leq\sqrt{\epsilon}"] & |\psi'\rangle \\
& & \sum_y \alpha_y (-1)^{b_y} |y\rangle \left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) |w_{y,f}\rangle \arrow[r, mapsto, "\text{Uncompute}"] & \sum_y \alpha_y (-1)^{b_y} |y\rangle \left(\frac{|0\rangle-|1\rangle}{\sqrt{2}}\right) |w_0\rangle
\end{tikzcd}$$

The point is that the result of these steeps, ψ, is within $2\sqrt{\epsilon}$ in trace distance from the state we would have if we used $\bar{f}$ (the "rounded" form of f) rather than f. Our error accrues each Grover iteration, so after at most $2^{\frac{m}{2}}$ Grover iterations we end up at most $2\sqrt{\epsilon}2^{\frac{m}{2}} < \frac{1}{8}$

from where we would be had we performed the entire algorithm with $\bar{f}$. $\bar{f}$ can be thought of as a *classical* function and so we can apply the ideas of normal Grover search. By repeating a random number of iterations above a certain threshold we can show that we accept using $\bar{f}$ with probability at least $\frac{1}{4}$, meaning that our acceptance probability with f is at least $\frac{1}{8}$.

Constant soundness error: Because the copies of $|x\rangle$ are not entangled, our Chernoff bound works and we can follow the same approach as in the proof of completeness.

Runtime: For each of the $O(2^{\frac{m}{2}})$ Grover iterations we make m parallel runs of t elementary operations, along with $O(m)$ work at the end of each iteration. We also use $O(m)$ operations at the beginning for our randomness. Overall, the runtime is thus $O(mt2^{\frac{m}{2}})$. $\square$