


Introducing structure to expedite quantum searching

Marcin Briński^{✉,*}, Jan Gwinner,[†] Vladyslav Hlembotskyi,[‡] Witold Jarnicki^{✉,§}, Szymon Pliś^{✉,||} and Adam Szady[¶]
BEIT, Mogilska 43 31-545 Kraków, Poland

 (Received 15 December 2020; accepted 7 June 2021; published 28 June 2021)

We present a quantum algorithm for solving the unstructured search problem with one marked element. Our algorithm allows generating quantum circuits that use asymptotically fewer additional quantum gates than the famous Grover's algorithm and may be successfully executed on noisy intermediate-scale quantum devices. We prove that our algorithm is optimal in the total number of elementary gates up to a multiplicative constant. As many NP-hard problems are, in fact, not unstructured, we also describe the *partial uncompute* technique which exploits the oracle structure and allows a significant reduction in the number of elementary gates required to find the solution. Combining these results allows us to use an asymptotically smaller number of elementary gates than Grover's algorithm in various applications, keeping the number of queries to the oracle essentially the same. We show how the results can be applied to solve hard combinatorial problems, for example, Unique k -SAT. Additionally, we show how to asymptotically reduce the number of elementary gates required to solve the unstructured search problem with multiple marked elements.

DOI: [10.1103/PhysRevA.103.062425](https://doi.org/10.1103/PhysRevA.103.062425)

I. INTRODUCTION

In the quantum *unstructured search problem*, the task is to find one marked element out of N elements corresponding to the computational basis. We want to accomplish that by the least possible number of queries to a given phase oracle, the only action of which is changing the signs of the coordinates corresponding to the marked elements. For more details, see Sec. II.

The celebrated Grover's algorithm [1] is one of the main achievements of quantum computing. It locates a marked element using only $O(\sqrt{N})$ queries to the oracle and $O(\sqrt{N} \log N)$ additional (i.e., nonoracle) elementary gates. Grover's result has been used extensively as a subroutine in many quantum algorithms (for examples, see [2–4]). We show how to reduce the average number of additional gates per oracle query while keeping the number of oracle queries as close to the optimum as we wish. We also prove that our algorithm is optimal up to a multiplicative constant.

A. Prior work

Since the invention of Grover's algorithm, there were several attempts to improve it further. In [5] the author improves the number of nonoracle quantum gates. Using a simple pattern of small diffusion operators the following result is obtained.

Theorem 1 ([5]). For every $\alpha > 2$ and any sufficiently large N there exists a quantum algorithm that finds the unique marked element among N with probability tending to 1, using fewer than $\frac{\pi}{4}\sqrt{N}(\frac{1}{1-(\log_2 N)^{-\alpha}})$ oracle queries and no more than $\frac{9}{8}\pi\alpha\sqrt{N} \log_2 \log_2 N$ nonoracle gates.

Later, in [6] the authors reduce the number of nonoracle gates even further.

Theorem 2 ([6]). For any integer $r > 0$ and sufficiently large N of the form $N = 2^n$, there exists a quantum algorithm that finds the unique marked element among N with probability 1, using $[\frac{\pi}{4} + o(1)]\sqrt{N}$ queries and $O(\sqrt{N} \log^r N)$ gates. For every $\varepsilon > 0$ and sufficiently large N of the form $N = 2^n$, there exists a quantum algorithm that finds the unique marked element among N with probability 1, using $\frac{\pi}{4}\sqrt{N}(1 + \varepsilon)$ queries and $O(\sqrt{N} \log(\log^* N))$ gates.

In the same paper, the authors raise questions regarding removing the $\log(\log^* N)$ factor in gate complexity, which we answer in the affirmative in Theorem 3, and dealing with oracles that mark multiple elements. Note that both aforementioned results assume that the given oracle marks only a single element.

The concept of benefits arising from the use of local diffusion operators has been studied in other papers, e.g., Ref. [7].

B. Our results

We present an algorithm which uses only $O(\sqrt{N})$ nonoracle gates while making only $O(\sqrt{N})$ oracle queries. Additionally, to remedy the objections against optimizing the average number of additional elementary gates per oracle query mentioned in [6], we introduce the concept of *partial uncompute*, a technique that achieves asymptotical improvement in the total number of elementary gates in many combinatorial problems, such as Unique k -SAT (see, e.g., [8])

*marbri@beit.tech

†jan.gwinner@beit.tech

‡vlad@beit.tech

§witek@beit.tech

||szymon@beit.tech

¶adsz@beit.tech

for the definition of Unique k -SAT). The high-level idea of the technique is to utilize the structure of the given oracle and store some intermediate information on ancilla qubits when implementing the oracle. If between two consecutive oracle queries we applied elementary gates only on a small number of qubits, we expect that the most of intermediate information has not changed at all. Leveraging this phenomenon, we can reduce the asymptotic number of gates needed to implement the circuit.

In Grover's algorithm the diffusion operator is applied on $O(\log N)$ qubits, so we cannot benefit from partial uncompute. We need to have an algorithm that on average affects only a small subset of qubits between consecutive oracle queries. To handle this problem we introduce an algorithm for generating quantum circuits that drastically reduces the average number of additional gates. The algorithm can be used to generate circuits that work for any number of qubits and can be potentially implemented on noisy intermediate-scale quantum (NISQ) devices. Moreover, the algorithm improves on the results of [5,6] and can be summarized as follows.

Theorem 3. Fix any $\varepsilon \in (0, 1)$, and any $N \in \mathbb{N}$ of the form $N = 2^n$. Suppose we are given a quantum oracle O operating on n qubits that marks exactly one element. Then there exists a quantum circuit \mathcal{A} which uses the oracle O at most $(1 + \varepsilon)\frac{\pi}{4}\sqrt{N}$ times and uses at most $O(\log(1/\varepsilon)\sqrt{N})$ non-oracle basic gates, which finds the element marked by O with certainty.

It is important to note that the constant hidden by O notation in Theorem 3 is independent of both N and ε . Moreover, any quantum algorithm tackling this problem must perform at least $\frac{\pi}{4}\sqrt{N}$ oracle calls (see [9]).

The algorithm \mathcal{A} can be, in broad strokes, explained as follows. We build a quantum circuit recursively according to some simple rules. The resulting circuit concentrates enough amplitude in the marked element. After that, we apply amplitude amplification [10] to it. The main idea in \mathcal{A} is to explore small diffusion operators (diffusion operators applied on a small subset of qubits). They are obviously easier to implement than large ones and require fewer elementary gates. Moreover, if they are applied wisely, they can be extremely efficient in concentrating amplitude in the marked element.

If we combine the partial uncompute technique with Theorem 3 to solve a Unique k -SAT problem, we get the following corollary.

Corollary 1. Consider the Unique k -SAT problem with n variables and c clauses. There exists a quantum circuit that uses $O[c \log(c)2^{n/2}/n]$ total (oracle and non-oracle) gates and solves the problem with certainty.

It is worth mentioning that it is a slight improvement over the naïve application of Grover's algorithm to solve the Unique k -SAT problem because Grover's algorithm requires $O[(n + c)2^{n/2}]$ elementary gates to solve the problem with certainty.

By result of [9], the optimal number of queries to the oracle required for solving unstructured search problem with certainty is $\frac{\pi}{4}\sqrt{N}$. We show that the tradeoff between the number of oracle queries and non-oracle gates from Theorem 3 is optimal up to a constant factor.

Corollary 2. There exists a number $\delta > 0$ such that for any $\varepsilon \in (0, 1)$ and for any quantum circuit \mathcal{A} the following holds. If \mathcal{A} uses at most $\delta \log(1/\varepsilon)\sqrt{N}$ non-oracle gates and finds the

element marked by O with certainty, then \mathcal{A} uses the oracle O at least $(1 + \varepsilon)\frac{\pi}{4}\sqrt{N}$ times.

Last but not least, following the approach of [11], we asymptotically reduce the overhead incurred when reducing the unstructured search problem with multiple marked elements to the unstructured search problem with exactly one marked element. We modify the oracle in a classical randomized way so that the modified oracle marks exactly one element with constant probability. This is achieved by randomly choosing an affine hash function that excludes some elements from the search space. If the number of marked elements K is known in advance, we will sample a hash function from such a set so that the expected number of marked elements after combining the oracle with the function is equal to one. We formulate this result as the following theorem.

Theorem 4. Let $N \in \mathbb{N}$ be of the form $N = 2^n$. Assume that we are given a phase oracle O that marks K elements, and we know the number k given by $k = 1 + \lceil \log_2 K \rceil$. Then one can find an element marked by O with probability at least $\frac{1}{16}$, using at most $O(\sqrt{\frac{N}{K}})$ oracle queries and at most $O(\log K \sqrt{\frac{N}{K}})$ non-oracle basic gates.

What is more, we can extend this approach to the case when the number of marked elements is unknown by trying different values of K and applying the same algorithm. This can be done in such a way that the number of oracle queries and the average number of additional elementary gates per oracle query are asymptotically the same as in the case of known K .

C. Further remarks

While our results describe asymptotic behavior, the techniques used to achieve them are quite practical. As described in [12], they may be applicable for achieving the improvements in implementations of unstructured search on existing and near-future NISQ devices. The previous implementations of unstructured search beyond spaces spanned by 3-qubits were unsuccessful [13]; perhaps techniques described here can allow searching larger spaces on current hardware.

Organization

In Sec. II we briefly discuss the computational model and notation used throughout this paper. In Sec. III we describe our main algorithm for constructing quantum circuits. Next, in Sec. IV we prove that our algorithm is optimal (up to a constant factor) in the number of additional elementary gates. Later, in Sec. V we introduce the partial uncompute technique and show an example application to a hard combinatorial problem. Finally, in Sec. VI we proceed to reduce the unstructured search problem with multiple marked elements to the unstructured search problem with one marked element.

II. PRELIMINARIES

In the *unstructured search problem* we are given a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ for some $n \in \mathbb{N}$ and we wish to find $x \in \{0, 1\}^n$ such that $f(x) = 1$. We will call such x *marked*. The function can be evaluated at N points in total, where $N = 2^n$, and the goal is to find a marked element while minimizing

the total number of evaluations of f . In the quantum version of the problem the function f is given as a *phase oracle* O , i.e., a unitary transformation given by $O|x\rangle = (-1)^{f(x)}|x\rangle$ for every computational basis vector $|x\rangle$. We still want to query O the least possible number of times to find a marked element. Sometimes this problem is called the *database search* problem. We use the standard gate model of quantum computations. We assume that our elementary operations are the universal set of quantum gates consisting of CNOT and arbitrary one-qubit gates. We will refer to these gates as *basic gates*. We note that this gate set can simulate any other universal gate set with bounded gate size with at most constant overhead; the details can be found in [14].

In all following equations all operators are to be understood as applied right to left (i.e., as in standard operator composition), while in figures the application order is left to right, as is the standard when drawing quantum circuits.

Given a positive integer k , the *uniform superposition state on k qubits*, denoted $|u_k\rangle$, is defined as $|u_k\rangle = \frac{1}{\sqrt{2^k}} \sum_{b \in \{0,1\}^k} |b\rangle$. We extend this definition to the special case of $k = 0$ by setting $|u_0\rangle = 1$. A useful identity which we will use throughout the derivations to come is $|u_a\rangle|u_b\rangle = |u_{a+b}\rangle$ for $a, b \in \mathbb{N}$.

The *mixing operator of size k* (alternatively also called the diffusion operator, or simply the diffuser), denoted G_k , is defined as $G_k = 2|u_k\rangle\langle u_k| - \text{Id}_k$, where Id_k is the identity matrix of size 2^k . From [14] we know that we can implement G_k using $O(k)$ basic gates (and this is best possible).

To prove optimality of our results and to define the partial uncompute technique we consider what happens when operators do not act on some subset of qubits. Intuitively, it means that we do not need to use these qubits when implementing this operator using basic gates. We say that a unitary matrix A operating on n qubits (here denoted $\{q_1, \dots, q_n\}$) *does not act on the qubit q_i* if

$$A = \text{SWAP}(q_i, q_n)(A' \otimes \text{Id}_1)\text{SWAP}(q_i, q_n),$$

where A' is some unitary matrix operating on $n - 1$ qubits, and

$\text{SWAP}(a, b) = \text{CNOT}(a, b)\text{CNOT}(b, a)\text{CNOT}(a, b)$. Otherwise, we say that A *acts on* qubit q_i . We say that operator A *may act on* qubits q_{i_1}, \dots, q_{i_m} if it does not act on qubits $\{q_1, \dots, q_n\} \setminus \{q_{i_1}, q_{i_2}, \dots, q_{i_m}\}$.

In the proof of Theorem 3 we will need the following result from [10], which we will refer to as *amplitude amplification*.

Theorem 5 ([10], p. 7, Theorem 2). Let \mathcal{A} be any quantum algorithm operating on n qubits that uses no measurements, and let $f : \{0, 1\}^n \rightarrow \{0, 1\}$ be any Boolean function with a corresponding phase oracle O . Let a be the probability that measuring $\mathcal{A}|00 \dots 0\rangle$ yields $|t\rangle$ such that $f(t) = 1$, and assume that $a \in (0, 1)$. Let $\theta \in (0, \pi/2)$ be such that $(\sin \theta)^2 = a$, and let $s = \lfloor \frac{\pi}{4\theta} \rfloor$. Then measuring $(-\mathcal{A}F_0\mathcal{A}^\dagger O)^s \mathcal{A}|00 \dots 0\rangle$ yields $|t\rangle$ such that $f(t) = 1$ with probability at least $\max\{1 - a, a\}$, where

$$F_0|t\rangle = \begin{cases} |t\rangle, & \text{if } t \neq 0 \\ -|t\rangle, & \text{if } t = 0. \end{cases}$$

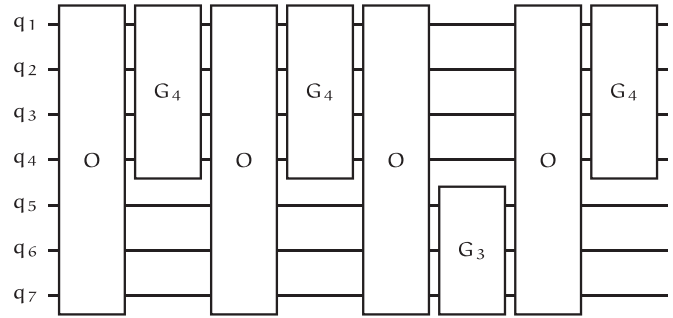


FIG. 1. W_2 for $\bar{k} = (4, 3)$.

Note that this result requires us to know the value of a *precisely*. However, this is not a problem for us, as we shall later see.

There is a simple corollary one can obtain from the proof of Theorem 5 (it is noted as Theorem 4 in [10], however, the authors do not make the constants explicit in their formulation). The precise formula one gets for the probability of success when measuring $(-\mathcal{A}F_0\mathcal{A}^\dagger O)^m \mathcal{A}|00 \dots 0\rangle$ is in fact equal to $\sin^2[(2m + 1)\theta]$. If it were to happen that $r = \pi/(4\theta) - 1/2$ was an integer, then we could simply set the number of iterations to r and obtain a solution with certainty. Now it remains to note that we can easily modify \mathcal{A} to lower θ slightly so that the new value of r is indeed an integer. It is important for our results that the number of iterations is in fact bounded by $\lfloor \frac{\pi}{4\theta} \rfloor + 1$, which is formulated as the theorem below.

Theorem 6 ([10], Theorem 4 restated). Let \mathcal{A} be any quantum algorithm operating on n qubits that uses no measurements, and let $f : \{0, 1\}^n \rightarrow \{0, 1\}$ be any Boolean function. Let a be the probability that measuring $\mathcal{A}|00 \dots 0\rangle$ yields $|t\rangle$ such that $f(t) = 1$, and assume that $a \in (0, 1)$. Let $\theta \in (0, \pi/2)$ be such that $(\sin \theta)^2 = a$. Then, there exists a quantum algorithm that uses \mathcal{A} and \mathcal{A}^\dagger at most $\lfloor \frac{\pi}{4\theta} \rfloor + 2$ times each, which upon measurement yields $|t\rangle$ such that $f(t) = 1$ with certainty.

Note that the bound $\lfloor \frac{\pi}{4\theta} \rfloor + 2$ follows from the extra \mathcal{A} applied at the beginning of the amplitude amplification (as we are counting the applications of \mathcal{A} and \mathcal{A}^\dagger and not iterations).

III. STRUCTURE OF THE W_m CIRCUIT

Definition 1. Let $\bar{k} = (k_1, \dots, k_m)$ be a sequence of positive integers and let $n := \sum_{j=1}^m k_j$. Given a quantum oracle O , for $j \in \{0, \dots, m\}$ we define the circuit W_j recursively as follows:

$$W_0 := \text{Id}_n, \quad W_j := W_{j-1}(\text{Id}_{k_1+\dots+k_{j-1}} \otimes G_{k_j} \otimes \text{Id}_{k_{j+1}+\dots+k_m}) \times W_{j-1}^\dagger O W_{j-1}, \quad j \in \{1, 2, \dots, m\}.$$

For an example of how the circuits W_m look like, see figs. 1 and 2. Observe that the circuit W_m uses the oracle O exactly $(3^m - 1)/2$ times. Moreover, as the mixing operator G_k can be implemented using $O(k)$ basic quantum gates, for a given \bar{k} , one can implement W_m for these diffuser sizes using $O(\sum_{j=1}^m k_j 3^{m-j})$ basic quantum gates, not including the gates necessary for the implementation of the oracle.

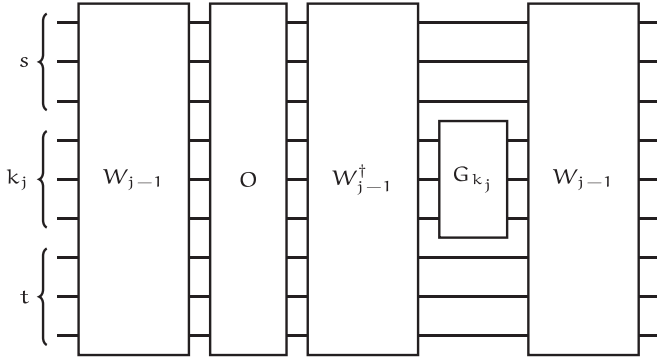


FIG. 2. Graphical representation of the W_j circuit. Note that the oracle in W_{j-1} manipulates all of the qubits, however, no other gate does so. In this picture $s = k_1 + \dots + k_{j-1}$ and $t = k_{j+1} + \dots + k_m$.

A. Obtaining the recurrence for amplitude in the target

In this section we aim to derive a recurrence formula that will allow us to compute the amplitude our circuit W_m concentrates in the unique marked state. We assume we are given a phase oracle O operating on n qubits, that marks a single state denoted target. We have also fixed a vector of positive integers $\bar{k} = (k_1, \dots, k_m)$, such that $k_1 + \dots + k_m = n$. For the duration of this section, we introduce the following notational conveniences. We split the marked state $|\text{target}\rangle$ according to \bar{k} as

$$|\text{target}\rangle = |\text{target}_1\rangle |\text{target}_2\rangle \dots |\text{target}_m\rangle,$$

where target_1 consists of bits of target numbered 1 to k_1 , target_2 of the bits numbered $k_1 + 1$ to $k_1 + k_2$, etc. Moreover, for given i, j we define the following product:

$$|\text{target}_i^j\rangle = |\text{target}_i\rangle |\text{target}_{i+1}\rangle \dots |\text{target}_j\rangle.$$

If the interval $[i, j]$ happens to be empty, we understand $|\text{target}_i^j\rangle$ to be the scalar 1. To shorten the derivations about to follow, we will also use these shorthands

$$|\overline{\text{target}}_j\rangle = \frac{1}{2^{k_j/2}} \sum_{\substack{b \in \{0,1\}^{k_j} \\ b \neq \text{target}_j}} |b\rangle,$$

$$|u_1^j\rangle = |u_s\rangle,$$

where $s = k_1 + \dots + k_j$ with the additional convention that $|u_1^0\rangle = 1$. Observe that we have the equations $|u_{k_j}\rangle = |\overline{\text{target}}_j\rangle + 2^{-k_j/2} |\text{target}_j\rangle$ and $\langle \text{target}_j | \overline{\text{target}}_j \rangle = 0$.

We begin by introducing two simple lemmas.

Lemma 1. Fix any $m \in \mathbb{N}_+$, and any $\bar{k} = (k_1, \dots, k_m) \in \mathbb{N}_+^m$, and let $n = \sum_{j=1}^m k_j$. Assume that we are given a phase oracle O that operates on n qubits and marks a single vector of the standard computational basis denoted target. Then for any $j \in \{0, \dots, m-1\}$, and any vector $|\phi\rangle \in (\mathbb{C}^2)^{\otimes t}$ (where $t = k_{j+1} + \dots + k_m$) such that $\langle \phi | \text{target}_{j+1}^m \rangle = 0$ we have

$$W_j(|u_1^j\rangle |\phi\rangle) = |u_1^j\rangle |\phi\rangle.$$

Proof. Observe that as $\langle \phi | \text{target}_{j+1}^m \rangle = 0$ the vector $|u_s\rangle |\phi\rangle$ is an eigenvector of the operator O with eigenvalue 1. Thus, the lemma's assertion will be proved, if we show that it is also

an eigenvector (with eigenvalue 1) of each diffusion operator that appears in W_j , that is $(\text{Id}_a \otimes G_b \otimes \text{Id}_{n-b-a})|u_1^j\rangle |\phi\rangle = |u_1^j\rangle |\phi\rangle$ whenever $a + b \leq k_1 + k_2 + \dots + k_j$, which we quickly verify by the direct calculation below:

$$\begin{aligned} & (\text{Id}_a \otimes G_b \otimes \text{Id}_{n-b-a})(|u_1^j\rangle |\phi\rangle) \\ &= (\text{Id}_a \otimes G_b \otimes \text{Id}_{n-b-a})(|u_a\rangle |u_b\rangle |u_{k_1+\dots+k_j-a-b}\rangle |\phi\rangle) \\ &= (\text{Id}_a |u_a\rangle)(G_b |u_b\rangle)(\text{Id}_{n-b-a} |u_{k_1+\dots+k_j-b-a}\rangle |\phi\rangle) \\ &= |u_a\rangle |u_b\rangle |u_{k_1+\dots+k_j-b-a}\rangle |\phi\rangle = |u_1^j\rangle |\phi\rangle. \end{aligned}$$

Lemma 2. Fix any $m \in \mathbb{N}_+$, and any $\bar{k} = (k_1, \dots, k_m) \in \mathbb{N}_+^m$, and let $n = \sum_{j=1}^m k_j$. Assume that we are given a phase oracle O that operates on n qubits and marks a single vector of the standard computational basis denoted target. Then for any $j \in \{1, \dots, m\}$ we have

$$\begin{aligned} & W_{j-1}(\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}) W_{j-1}^\dagger |\text{target}\rangle \\ &= \left(\frac{2}{2^{k_j}} - 1 \right) |\text{target}\rangle + |\vartheta\rangle, \end{aligned}$$

where $s = k_1 + \dots + k_j$, and $|\vartheta\rangle$ is some state orthogonal to $|\text{target}\rangle$.

Proof. Observe that each diffusion operator in W_{j-1} (and thus also in W_{j-1}^\dagger) operates on the qubits numbered $\{1, \dots, k_1 + \dots + k_{j-1}\}$, thus, there exists a vector $|\eta\rangle \in (\mathbb{C}^2)^{\otimes (k_1 + \dots + k_{j-1})}$ such that

$$W_{j-1}^\dagger |\text{target}\rangle = |\eta\rangle |\text{target}_j^m\rangle.$$

Equipped with this observation, we proceed to directly compute the desired result

$$\begin{aligned} & W_{j-1}(\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}) W_{j-1}^\dagger |\text{target}\rangle \\ &= W_{j-1}(\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}) |\eta\rangle |\text{target}_j^m\rangle \\ &= W_{j-1}(\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}) |\eta\rangle |\text{target}_j\rangle |\text{target}_{j+1}^m\rangle \\ &= W_{j-1} [|\eta\rangle (G_{k_j} |\text{target}_j\rangle) |\text{target}_{j+1}^m\rangle] \\ &= W_{j-1} \left[|\eta\rangle \left(\frac{2}{2^{k_j/2}} |u_{k_j}\rangle - |\text{target}_j\rangle \right) |\text{target}_{j+1}^m\rangle \right] \\ &= W_{j-1} \left(\frac{2}{2^{k_j}} - 1 \right) |\eta\rangle |\text{target}_j^m\rangle \\ &\quad + W_{j-1} \frac{2}{2^{k_j/2}} |\eta\rangle \langle \overline{\text{target}}_j | \text{target}_j^m \rangle |\text{target}_{j+1}^m\rangle \\ &= \left(\frac{2}{2^{k_j}} - 1 \right) |\text{target}\rangle + |\vartheta\rangle \end{aligned}$$

and observe that $|\vartheta\rangle$ is orthogonal to $|\text{target}\rangle$, as their respective preimages under W_{j-1} were orthogonal. \blacksquare

Lemma 3. Fix any $m \in \mathbb{N}_+$, and any $\bar{k} = (k_1, \dots, k_m) \in \mathbb{N}_+^m$, and let $n = \sum_{j=1}^m k_j$. Assume that we are given a phase oracle O that operates on n qubits and marks a single vector of the standard computational basis denoted target. Define the numbers

$$\alpha_j = \langle \text{target} | (W_j |u_1^j\rangle) | \text{target}_{j+1}^m \rangle$$

for $j \in \{0, 1, \dots, m\}$. Then, α_j satisfy the recurrence

$$\alpha_j = \begin{cases} 1, & \text{if } j = 0 \\ 2^{-k_j/2}(3 - 4 \times 2^{-k_j})\alpha_{j-1}, & \text{if } j > 0. \end{cases}$$

Proof. Clearly, $\alpha_0 = 1$ giving the base case. Now, let us assume that $j > 0$, and we will proceed to compute α_j by expanding the circuit W_j according to Definition 1. To maintain legibility we will split this computation into several steps. Let us define the intermediate states $|w_1\rangle, \dots, |w_5\rangle$ by the

following equations:

$$\begin{aligned} |w_1\rangle &= W_{j-1}(|u_1^j\rangle|\text{target}_{j+1}^m\rangle), \\ |w_2\rangle &= O|w_1\rangle, \\ |w_3\rangle &= W_{j-1}^\dagger|w_2\rangle, \\ |w_4\rangle &= (\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s})|w_3\rangle, \\ |w_5\rangle &= W_{j-1}|w_4\rangle, \end{aligned}$$

where $s = k_1 + \dots + k_j$;

$$\begin{aligned} |w_1\rangle &= W_{j-1}(|u_1^j\rangle|\text{target}_{j+1}^m\rangle) = W_{j-1}\left(\frac{1}{2^{k_j/2}}|u_1^{j-1}\rangle|\text{target}_j^m\rangle + |u_1^{j-1}\rangle\langle\overline{\text{target}}_j|\text{target}_{j+1}^m\rangle\right) \\ &= \frac{1}{2^{k_j/2}}W_{j-1}|u_1^{j-1}\rangle|\text{target}_j^m\rangle + |u_1^{j-1}\rangle\langle\overline{\text{target}}_j|\text{target}_{j+1}^m\rangle, \end{aligned} \quad (1)$$

where in Eq. (1) we relied on Lemma 1. Plugging this equation into the definition of $|w_2\rangle$ we obtain

$$\begin{aligned} |w_2\rangle &= O\left(\frac{1}{2^{k_j/2}}W_{j-1}|u_1^{j-1}\rangle|\text{target}_j^m\rangle + |u_1^{j-1}\rangle\langle\overline{\text{target}}_j|\text{target}_{j+1}^m\rangle\right) \\ &= \frac{1}{2^{k_j/2}}(W_{j-1}|u_1^{j-1}\rangle|\text{target}_j^m\rangle - 2\alpha_{j-1}|\text{target}\rangle) + |u_1^{j-1}\rangle\langle\overline{\text{target}}_j|\text{target}_{j+1}^m\rangle, \end{aligned} \quad (2)$$

$$\begin{aligned} |w_3\rangle &= W_{j-1}^\dagger\left(\frac{1}{2^{k_j/2}}W_{j-1}|u_1^{j-1}\rangle|\text{target}_j^m\rangle - \frac{2}{2^{k_j/2}}\alpha_{j-1}|\text{target}\rangle + |u_1^{j-1}\rangle\langle\overline{\text{target}}_j|\text{target}_{j+1}^m\rangle\right) \\ &= \frac{1}{2^{k_j/2}}|u_1^{j-1}\rangle|\text{target}_j^m\rangle - \frac{2}{2^{k_j/2}}\alpha_{j-1}W_{j-1}^\dagger|\text{target}\rangle + |u_1^{j-1}\rangle\langle\overline{\text{target}}_j|\text{target}_{j+1}^m\rangle \end{aligned} \quad (3)$$

$$= |u_1^j\rangle|\text{target}_{j+1}^m\rangle - \frac{2}{2^{k_j/2}}\alpha_{j-1}W_{j-1}^\dagger|\text{target}\rangle, \quad (4)$$

$$\begin{aligned} |w_4\rangle &= \text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}\left(|u_1^j\rangle|\text{target}_{j+1}^m\rangle - \frac{2}{2^{k_j/2}}\alpha_{j-1}W_{j-1}^\dagger|\text{target}\rangle\right) \\ &= |u_1^j\rangle|\text{target}_{j+1}^m\rangle - \frac{2}{2^{k_j/2}}\alpha_{j-1}(\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s})W_{j-1}^\dagger|\text{target}\rangle, \end{aligned}$$

$$\begin{aligned} |w_5\rangle &= W_{j-1}\left(|u_1^j\rangle|\text{target}_{j+1}^m\rangle - \frac{2}{2^{k_j/2}}\alpha_{j-1}(\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s})W_{j-1}^\dagger|\text{target}\rangle\right) \\ &= \frac{1}{2^{k_j/2}}W_{j-1}|u_1^{j-1}\rangle|\text{target}_j^m\rangle + |u_1^{j-1}\rangle\langle\overline{\text{target}}_j|\text{target}_{j+1}^m\rangle - \frac{2}{2^{k_j/2}}\alpha_{j-1}W_{j-1}(\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s})W_{j-1}^\dagger|\text{target}\rangle \end{aligned} \quad (5)$$

$$= \frac{1}{2^{k_j/2}}W_{j-1}|u_1^{j-1}\rangle|\text{target}_j^m\rangle + |u_1^{j-1}\rangle\langle\overline{\text{target}}_j|\text{target}_{j+1}^m\rangle - \frac{2}{2^{k_j/2}}\alpha_{j-1}\left[\left(\frac{2}{2^{k_j}} - 1\right)|\text{target}\rangle + |\vartheta\rangle\right]. \quad (6)$$

Note that in Eq. (2) we used the definition of α_{j-1} , in Eq. (3) we applied Lemma 1, Eqs. (4) and (5) follow from the definition of $|\overline{\text{target}}_j\rangle$, while Eq. (6) we applied Lemma 2. Keeping in mind that $|\vartheta\rangle$ is orthogonal to $|\text{target}\rangle$ and equipped with Eq. (6) we may finally compute α_j as

$$\begin{aligned} \alpha_j &= \langle\text{target}|w_5\rangle = \frac{1}{2^{k_j/2}}\langle\text{target}|W_{j-1}|u_1^{j-1}\rangle|\text{target}_j^m\rangle + 2\alpha_{j-1}\left(1 - \frac{2}{2^{k_j}}\right)\langle\text{target}|\text{target}\rangle \\ &= \frac{1}{2^{k_j/2}}\left[\alpha_{j-1} + 2\alpha_{j-1}\left(1 - \frac{2}{2^{k_j}}\right)\right] = \frac{1}{2^{k_j/2}}(3 - 4 \times 2^{-k_j})\alpha_{j-1}. \end{aligned}$$

■

B. Proof of Theorem 3

Proof of Theorem 3. It clearly suffices to prove the theorem under assumption that ε is small enough; let us assume that is indeed the case.

Let $\bar{k} = (k_1, \dots, k_m)$ be some sequence of positive integers to be determined later, such that $\sum_{j=1}^m k_j = n$. We will use the circuit W_m with these diffuser sizes, and utilize Theorem 6 on top of this circuit. To estimate the number of

iterations made by amplitude amplification, we need a precise formula for amplitude in the marked state that the circuit $W_m H^{\otimes n}$ (the Walsh-Hadamard transform is only necessary because we assumed our circuit to be fed the state $|u_n\rangle$, while amplitude amplification assumes that the state $|00\dots 0\rangle$ is the one we work with) yields, denoted α_m . To this end we use the recurrence we have obtained in Lemma 3, to which we can provide a solution as a product¹

$$\begin{aligned} \alpha_m &= \prod_{j=1}^m [2^{-k_j/2} (3 - 4 \times 2^{-k_j})] = 2^{-n/2} \prod_{j=1}^m (3 - 4 \times 2^{-k_j}) \\ &= 2^{-n/2} 3^m \prod_{j=1}^m \left(1 - \frac{4}{3} \times 2^{-k_j}\right). \end{aligned} \quad (7)$$

Let us now consider the case of particular choice of \bar{k} , namely, $k_j = (x + 1)j$, where $x \in \mathbb{N}_+$ is some fixed constant. We will for now assume, for the sake of simplicity, that the number of qubits n is precisely equal to $(x + 1) + 2(x + 1) + \dots + m(x + 1) = (x + 1)m(m + 1)/2$. We will later argue that this assumption is not necessary. Observe that in particular we have

$$m \in \Theta(\sqrt{n/x}). \quad (8)$$

Thus, we can lower bound the product in α_m as follows:

$$\prod_{j=1}^m \left(1 - \frac{4}{3} \times 2^{-(x+1)j}\right) \geq \prod_{j=1}^m (1 - 2^{-xj}) \geq \prod_{j=1}^{\infty} (1 - 2^{-xj}).$$

We recall the identity due to Euler [15], which relates the infinite product on the right-hand side with pentagonal numbers

$$\prod_{j=1}^{\infty} (1 - z^j) = 1 + \sum_{j=1}^{\infty} (-1)^j (z^{(3j-1)j/2} + z^{(3j+1)j/2})$$

which we use to lower bound the product for $z \in [0, 1)$ as

$$\prod_{j=1}^{\infty} (1 - z^j) \geq 1 - z - z^2$$

by grouping latter terms in the series in consecutive pairs and observing that each such pair has a positive sum. This gives us the inequality

$$\alpha_m \geq 2^{-n/2} 3^m (1 - 2^{-x} - 2^{-2x}). \quad (9)$$

Using Theorem 6, we need at most

$$\left\lfloor \frac{\pi}{4\theta_m} \right\rfloor + 2$$

applications of our circuit W_m and its conjugate, where $\theta_m = \arcsin \alpha_m$. Using the standard inequality for $z \in (0, 1]$

$$\sin z \leq z$$

¹It is interesting to note that setting each $k_j = 2$ yields $\alpha_m = 1$ in which case amplitude amplification is not necessary, thus giving a simple algorithm solving the unstructured search problem with each diffuser size bounded by a constant. However, the number of oracle queries it makes is $O(3^{n/2})$.

which we can restate as

$$\frac{1}{\arcsin z} \leq \frac{1}{z}. \quad (10)$$

Inequalities (9) and (10) together imply that the number of applications of W_m and W_m^\dagger in amplitude amplification is bounded by

$$\frac{\pi}{4} \frac{1}{1 - 2^{-x} - 2^{-2x}} 2^{n/2} \times 3^{-m} + 2.$$

Observe that each W_m (and thus also W_m^\dagger) uses $(3^m - 1)/2$ oracle calls. Thus, the total number of oracle calls is bounded by

$$\frac{\pi}{4} \frac{1}{1 - 2^{-x} - 2^{-2x}} 2^{n/2} + 2 \times 3^m - 2,$$

thus, we are only a factor of $\frac{1}{1 - 2^{-x} - 2^{-2x}}$ away from optimal number of oracle calls, as by Eq. (8) the additive term is negligible.

Let us count the number of *nonoracle* gates used by our algorithm. Note that the overhead of operations used by amplitude amplification other than applications of W_m is negligible compared to the cost of the W_m circuit. Each W_m can be implemented using

$$O\left(\sum_{j=1}^m xj \times 3^{m-j}\right)$$

nonoracle gates, giving us at most

$$\begin{aligned} O\left[\left(\sum_{j=1}^m xj \times 3^{m-j}\right) 2^{n/2} \times 3^{-m}\right] \\ = O(x \times 2^{n/2} \sum_{j=1}^m j3^{-j}) = O(x \times 2^{n/2}) \end{aligned}$$

nonoracle gates used by the entire algorithm. Now, setting $x \in \Theta[\log(\varepsilon^{-1})]$ concludes the proof in this special case.

Now, we briefly explain how to deal with arbitrary number of qubits. We wish to get a suitable sequence \bar{k} for specific positive integers x and n . We do it as follows: let $m = \max\{k : \sum_{j \leq k} (x + 1)j \leq n\}$, and define for $j \in \{1, \dots, m\}$

$$k_j = \begin{cases} (x + 1)j, & \text{if } j < m \\ n - \sum_{k < m} (x + 1)k, & \text{if } j = m. \end{cases}$$

By the choice of m , we easily get that $k_m \in [(x + 1)m, 3(x + 1)m]$. Observe that the number of gates necessary to implement W_m goes up by a factor of at most 3, thus, that part of the calculation does not change. Next, observe that in Eq. (7), the final expression is monotonically increasing in k_j , thus our lower bound in inequality (9) still holds. Thus, further analysis also does not change, concluding the proof. ■

Remark 1. The above analysis could be generalized to the setting of underlying space being decomposable into a tensor product as

$$H_1 \otimes H_2 \otimes \dots \otimes H_m,$$

where of course the time complexity of the algorithm will depend on the relative dimensions of H_j . However, this would not improve the proof's clarity, and does not really provide a

significantly wider scope of applications, so we refrain from including it.

IV. OPTIMALITY

In this section we show the following lower bound for the number of oracle queries.

Theorem 7. Fix $p \in (0, 1)$, $n \in \mathbb{N}$, and $N = 2^n$. Let $T = T(N, p)$ be the number of oracle queries in the optimal (i.e., minimizing the number of oracle queries) search algorithm that is needed to find the marked element with probability at least p . There exists a constant $C > 0$, which possibly depends on p but does not depend on N , such that for any $\eta > 0$ and any algorithm \mathcal{A} the following holds. If \mathcal{A} uses at most ηT additional basic gates and finds the marked element with probability at least p , then \mathcal{A} needs to query oracle at least $T + \lfloor 2^{-C\eta} T \rfloor$ times.

As a by-product we reprove the Zalka’s estimation from [9] (Corollary 3) and at the end of the section we shortly explain how the above theorem implies Corollary 2.

Let $m \geq n$ be the number of qubits which we use. Assume that we have at our disposal a phase oracle O^y operating on n qubits with one marked element y . Any quantum algorithm that solves unstructured search problem has the following form: we start with some initial quantum state $|s\rangle$ and apply the alternating sequence of oracle queries O^y and unitary operators U_1, \dots, U_R (each of which acts on m qubits). Thus, as a result we get a state

$$|t\rangle = U_R O^y U_{R-1} O^y \dots U_1 O^y |s\rangle.$$

It is convenient to investigate the algorithm’s behavior for all possible $y \in \{0, 1\}^n$ simultaneously. For this purpose we consider the following sphere and its subset. Let

$$S = \{z \in ((\mathbb{C}^2)^{\otimes m})^N : |z| = \sqrt{N}\}$$

and

$$\hat{S} = \{(z_1, \dots, z_N) \in S : z_1 = \dots = z_N\}.$$

Let y_j for $j \in \{1, \dots, N\}$ be a sequence of all elements of $\{0, 1\}^n$. We use the following two actions of unitary group on the sphere S . For $U \in U(2^m)$ and z in S we put

$$Uz = (Uz_1, Uz_2, \dots, Uz_N)$$

and

$$O_U z = (UO^{y_1} z_1, UO^{y_2} z_2, \dots, UO^{y_N} z_N).$$

By straightforward calculations we get the following observation.

Observation 1. For z in \hat{S} we have

$$|O_U z - Uz| = 2.$$

We consider the following sequences of points on the sphere S :

$$\tilde{s} = (|s\rangle, \dots, |s\rangle) \text{ and } s_i = O_{U_R} \dots O_{U_{i+1}} U_i \dots U_1 \tilde{s}.$$

Let us recall the inequality proved in [16] (see also [9]) which is crucial for our considerations.

Lemma 4. If the algorithm finds marked element with probability at least p , then

$$|s_R - s_0|^2 \geq h(p), \tag{11}$$

where h is a function given by the formula

$$h(p) = 2N - 2\sqrt{N}\sqrt{p} - 2\sqrt{N}\sqrt{N-1}\sqrt{1-p}.$$

The advantage of working on the sphere is that the distance between points on the sphere S is connected with the angle between them. For $a, b \in S$ let $\varphi_{a,b}$ be the angle between them, i.e.,

$$\varphi_{a,b} = 2 \arcsin \left(\frac{|a-b|}{2\sqrt{N}} \right) \in [0, \pi]. \tag{12}$$

Such angle is proportional to the length of the shortest arc on S connecting a and b , so in particular it satisfies triangle inequality:

$$\varphi_{a,b} + \varphi_{b,c} \geq \varphi_{a,c}.$$

Put

$$\alpha = 2 \arcsin (1/\sqrt{N}). \tag{13}$$

Now let us consider distances between elements of sequence s_i .

Observation 2. For $i \in \{0, \dots, R-1\}$ we have

$$|s_i - s_{i+1}| = 2 \text{ and } \varphi_{s_i, s_{i+1}} = \alpha.$$

Proof. By Observation 1 we have

$$|s_i - s_{i+1}| = |O_U z - Uz| = 2,$$

where $U = U_{i+1}$ and $z = U_i \dots U_1 s_0$. The second part follows trivially from Eq. (12) and the choice of α . ■

Observation 3. For any $i, c \in \mathbb{N}$ such that $i+c \leq R$, the following inequality holds:

$$\varphi_{s_i, s_{i+c}} \leq c\alpha.$$

Proof. The inequality holds by the Observation 2 and the triangle inequality for angles. ■

If we look at Grover’s algorithm we can notice the following facts.

Observation 4. In case of Grover’s algorithm we have equality in the inequality given by Observation 3 for $i+c \leq (\frac{\pi}{\alpha} - 1)/2$.

Observation 5. In case of Grover’s algorithm all points s_0, \dots, s_R lie on a great circle of the sphere S .

Lemma 5. For a given $R \leq (\frac{2\pi}{\alpha} - 1)/2$ the expression $|s_R - s_0|$ is maximized by Grover’s algorithm.

Proof. Since in the case of Grover’s algorithm points s_0, \dots, s_R lie on the great circle we get $\varphi_{s_0, s_R} = R\alpha$ and thus the distance between s_0 and s_R is maximized. ■

Let us here recall the result of Zalka. By the above lemma, Observation 4, and Lemma 4 we get the following.

Corollary 3 (Zalka’s lower bound for search algorithm). Let $R \leq (\frac{2\pi}{\alpha} - 1)/2$. The Grover’s algorithm that makes R oracle queries gives maximal probability of measuring marked element among all quantum circuits that solve unstructured quantum search problem using at most R queries.

Note that for large N the number $\frac{\pi}{2\alpha}$ is close to $\frac{\pi}{4}\sqrt{N}$. Now let us see what happens after two steps of the algorithm. Put $d_K = 16(K - 1)/K$ for $K \geq 1$.

Observation 6. If $z \in \hat{S}$, then $|O_{Id}O_U z - Uz|^2 \leq d_N$. In particular for $i \in \{0, \dots, R - 2\}$ we have

$$|s_i - s_{i+2}|^2 \leq d_N.$$

Proof. It is the direct consequence of Observation 3 for $c = 2$. ■

The key observation for our lower bound is better estimation for unitary operators that act on bounded number of qubits. From this point of view, we consider that each oracle query can be performed on arbitrary qubits in arbitrary order (or we can think that we just add SWAP gates). To stress this we use here the symbols O and O' for oracle operators.

Lemma 6. Let $z = (|z\rangle, \dots, |z\rangle) \in \hat{S}$. If U acts at most on k qubits, then $|O'_{Id}O_U z - Uz|^2 \leq d_{K^2}$ where $K = 2^k$.

Proof. Let A_U be the set of k qubits on which U acts. Oracle O acts on qubits Q_1, \dots, Q_n and O' on Q'_1, \dots, Q'_n (here of course the order of qubits is important). Let J_O be a set of all such indices i that $Q_i \in A_U$, and $J_{O'}$ be a set of all indices j that $Q'_j \in A_U$. Without loss of generality we can assume that $J = J_O \cup J_{O'} = \{n + 1 - a, \dots, n\}$. Note that $a \leq 2k$ since $|J_O|, |J_{O'}| \leq |A_U| \leq k$. Put

$$B = A_U \cup \{Q_{n+1-a}, \dots, Q_n\} \cup \{Q'_{n+1-a}, \dots, Q'_n\}.$$

Let B' be a set of all other qubits. By the assumption above, it is a prefix of the set of all qubits.

Let us fix for a moment $y = (y_1, \dots, y_n) \in \{0, 1\}^n$. Let $q = (y_1, \dots, y_{n-a}) \in \{0, 1\}^{n-a}$ and $r = (y_{n-a+1}, \dots, y_n) \in \{0, 1\}^a$. We will also write qr in place of y . With these notions introduced, we can write $|z\rangle$ as

$$|z\rangle = \sum_{|x\rangle \in B'} \alpha_x |x\rangle |S_x\rangle,$$

where B' is a computational basis in the space related to qubits from B' , α_x are complex numbers, and $|S_x\rangle$ are states in the space related to qubits in B . It is clear that

$$\sum_{|x\rangle \in B'} |\alpha_x|^2 = 1.$$

We group elements of B' into four disjoint sets. Let B_1^q be the set of all $|x\rangle$ that agree with y_j on qubits Q_j and Q'_j for all $j \leq n - a$. Let B_2^q (respectively B_3^q) be the set of all $|x\rangle$ that agree with p_j on qubits Q_j (respectively Q'_j) for all $j \leq n - a$ but differs on at least one qubit Q'_j (respectively Q_j) for some $j \leq n - a$. And, finally, we put $B_4^q = B' \setminus (B_1^q \cup B_2^q \cup B_3^q)$. Now we have

$$|z\rangle = \sum_{i=1}^4 |z_i^q\rangle,$$

where

$$|z_i^q\rangle = \sum_{|x\rangle \in B_i^q} \alpha_x |x\rangle |S_x\rangle.$$

We have

$$\begin{aligned} U|z_i^q\rangle &= \sum_{|x\rangle \in B_i^q} \alpha_x |x\rangle U|S_x\rangle, \\ O^y U O^y |z_i^q\rangle &= \sum_{|x\rangle \in B_i^q} \alpha_x |x\rangle O^r U O^r |S_x\rangle, \end{aligned}$$

$$O^y U O^y |z_2^q\rangle = \sum_{|x\rangle \in B_2^q} \alpha_x |x\rangle U O^r |S_x\rangle,$$

$$O^y U O^y |z_3^q\rangle = \sum_{|x\rangle \in B_3^q} \alpha_x |x\rangle O^r U |S_x\rangle,$$

$$O^y U O^y |z_4^q\rangle = U |z_4^q\rangle,$$

where O^r (and respectively O'^r) are oracles on a qubits that mark element of the computational basis if for $k > n - a$ on Q_k (respectively on Q'_k) this element is y_k . We get

$$\begin{aligned} |O^y U O^y |z\rangle - U|z\rangle|^2 &= \sum_{|x\rangle \in B_1^q} |\alpha_x|^2 |(U - O^r U O^r)|S_x\rangle|^2 \\ &\quad + \sum_{|x\rangle \in B_2^q} |\alpha_x|^2 |(1 - O^r)|S_x\rangle|^2 \\ &\quad + \sum_{|x\rangle \in B_3^q} |\alpha_x|^2 |(1 - O'^r)U|S_x\rangle|^2. \end{aligned}$$

Now we are ready to sum up above expression with respect to r . For a fixed q , by applying Observation 1, we get

$$\begin{aligned} &\sum_r |O^{qr} U O^{qr} |z\rangle - U|z\rangle|^2 \\ &= \sum_{|x\rangle \in B_1^q} |\alpha_x|^2 \sum_r |(O^{qr} U O^{qr} - U)|S_x\rangle|^2 \\ &\quad + 4 \sum_{|x\rangle \in B_2^q} |\alpha_x|^2 + 4 \sum_{|x\rangle \in B_3^q} |\alpha_x|^2 \\ &\leq d_{2^a} \sum_{|x\rangle \in B_1^q} |\alpha_x|^2 + 4 \sum_{|x\rangle \in B_2^q \cup B_3^q} |\alpha_x|^2 \\ &\leq d_{K^2}/2 \left(\sum_{|x\rangle \in B_1^q \cup B_2^q} |\alpha_x|^2 + \sum_{|x\rangle \in B_1^q \cup B_3^q} |\alpha_x|^2 \right). \end{aligned}$$

The inequality in the fourth line holds by applying Observation 6 in case of $N = 2^a$.

The next step is sum the bound above with respect to q . Notice that

$$\bigcup_q B_1^q \cup B_2^q = \bigcup_q B_1^q \cup B_3^q = B',$$

since if for a fixed q one oracle marks some state $|z\rangle \in B'$, then the other oracle either agrees with it (putting $|z\rangle$ in B_1^q) or not (putting it in B_2^q or B_3^q , respectively). Because of that and the fact that for different q 's oracle marks disjoint $|z\rangle$'s, the \cup symbol is to be understood as disjoint set union. Therefore, we have

$$|z\rangle = \sum_q \sum_{|x\rangle \in B_1^q \cup B_2^q} \alpha_x |x\rangle |S_x\rangle = \sum_q \sum_{|x\rangle \in B_1^q \cup B_3^q} \alpha_x |x\rangle |S_x\rangle$$

and thus

$$\sum_q \sum_{|x\rangle \in B_1^q \cup B_2^q} |\alpha_x|^2 = \sum_q \sum_{|x\rangle \in B_1^q \cup B_3^q} |\alpha_x|^2 = 1.$$

Finally, we can conclude

$$\begin{aligned} |O'_{Id}O_U z - Uz|^2 &= \sum_y |O'^y U O^y |z\rangle - U|z\rangle|^2 \\ &= \sum_q \sum_r |O'^{qr} U O^{qr} |z\rangle - U|z\rangle|^2 \leq d_{K^2}. \end{aligned}$$

Proof of Theorem 7. Let us choose $k = 8\eta$. Note that if $k > n/4$ then for any $C \geq 32$ we get $T2^{-C\eta} \leq (\pi\sqrt{N}/4 + 1)/N < 1$ and we are done, so we can assume that $k \leq n/4$.

By Observation 3 for all $i \in \{0, \dots, R - 2\}$ we have

$$\varphi_{s_i, s_{i+2}} \leq 2\alpha = 2 \arcsin(\sqrt{d_N}/2\sqrt{N}),$$

and by Lemma 6, if operator U_{i+1} acts on at most k qubits then

$$\varphi_{s_i, s_{i+2}} \leq 2 \arcsin\left(\frac{\sqrt{d_{K^2}}}{2\sqrt{N}}\right).$$

We will use either of these bounds depending on whether an operator acts on more than k qubits or not. Note that the second bound is always better than the first one, as we assumed that $k \leq n/4$.

Since arcsin has the derivative greater or equal to one, for U_{i+1} acting on at most k qubits we can bound

$$\begin{aligned} \varphi_{s_i, s_{i+2}} &\leq 2 \arcsin\left(\frac{\sqrt{d_N}}{2\sqrt{N}}\right) - 2 \frac{\sqrt{d_N} - \sqrt{d_{K^2}}}{2\sqrt{N}} \\ &\leq 2\alpha - \frac{D}{K^2\sqrt{N}}, \end{aligned}$$

where the constant D (as well as constants D' and C below) does not depend on N and K . The inequality $\arcsin x \leq 2x$ for $x = 1/\sqrt{N}$ combined with Eq. (13) yields

$$\varphi_{s_i, s_{i+2}} \leq 2\alpha(1 - D/8K^2).$$

From the triangle inequality for angles we can now establish the following bound:

$$\varphi_{s_R, s_0} \leq \hat{\alpha} + \sum_{t \in \{l \in \mathbb{N} : 2l+2 \leq R\}} \varphi_{s_{2t}, s_{2t+2}},$$

where

$$\hat{\alpha} = \begin{cases} \alpha, & \text{if } R \text{ is an odd number} \\ 0, & \text{if } R \text{ is an even number.} \end{cases}$$

Note that, since each basic gate acts on at most two qubits, at least half of operators U_{2t+1} act on at most k qubits. Therefore, we can bound half of the angles by $2\alpha(1 - D/8K^2)$ and the rest by 2α , which gives us

$$\varphi_{s_R, s_0} \leq R\alpha - \frac{D'R}{K^2}\alpha \leq \alpha R(1 - 2^{-C\eta}).$$

On the other hand, by Lemma 4 and Observation 4 we have

$$\varphi_{s_R, s_0} > (T - 1)\alpha,$$

and thus

$$R > (T - 1)/(1 - 2^{-C\eta}) \geq T - 1 + 2^{-C\eta}T,$$

and the theorem follows. ■

Proof of Corollary 2. One can see that any algorithm needs more than $\frac{\pi}{4}\sqrt{N} - 1$ steps to find the marked element with

certainty (compare Corollary 3). Using Theorem 7 we get the result for $\delta = 1/C$ and for large enough n . If necessary, we may decrease δ so that $\delta \log(1/\epsilon)\sqrt{N} < 1$ for smaller values of n . ■

Remark 2. Note that we do not allow measurements before the end of the algorithm. It is not clear to the authors how measurements performed inside a circuit can reduce the expected number of oracle queries made. In particular, how Zalka's results about optimality of Grover's algorithm applies to this more general class of quantum algorithms is far from obvious. Example of measurements speeding up quantum procedures can be found in Sec. 4 of [16] or in the last section of this paper.

V. PARTIAL UNCOMPUTE

A. Motivation and intuition

The motivation for this section comes from the fact that many natural implementations of phase oracles mimic parallel classical computation by the following pattern of operations.

(1) We perform a long series of operations that do not alter the original n qubits (or alter them temporarily), but modify some number of ancilla qubits that were initially zero, usually by *CCX* gates.

(2) We perform a *Z* gate operation to flip the phase on the interesting states.

(3) We undo all the operations from step 1, not to hinder the amplitude interference in the subsequent mixing operators.

Step 3 offers the benefit of being able to reuse the (now cleared) ancilla qubits, but it is not the main motivation of performing it.

If a mixing operator used only acts on a subset of qubits, maybe not all gates from step 1 interfere with proper amplitude interference? It turns out that very often it suffices to undo only a fraction of gates. We will establish the proper language to express that in Sec. VB.

It turns out that the state that allows safe application of the mixing operator is much closer (in the metric of number of gates) to the state from after step 1 than to the base state with all ancilla qubits zeroed. Our approach will initially compute all the ancilla qubits, perform all the mixing "close" to this state, and finally uncompute the ancillas.

Intuitively, we shall follow the following scheme:

- (1) Compute all the ancilla qubits.
- (2) For all mixing operators,
 - (a) perform the phase flip,
 - (b) undo the ancilla computation that would interfere with the upcoming mixing operator,
 - (c) perform the mixing,
 - (d) redo the computation from step 2(b).
- (3) Uncompute all the ancilla qubits.

The last step 2(d) and step 3 could be even skipped, if the ancilla computation does not modify the original qubits. However we are not doing this optimization in the formal approach, as the benefits are minimal. Naturally, this is a very imprecise description. Full details are presented in Sec. VC.

We are aware that many (if not all) of these operations are performed by modern quantum circuit optimizers and preprocessors. The aim is to give structure to the process and

understand how many gates are guaranteed to be removed from the circuit.

B. Definitions

Recall from Sec. II that a phase oracle O of a function $f : \{0, 1\}^n \rightarrow \{0, 1\}$ is a unitary transformation given by $O|x\rangle = (-1)^{f(x)}|x\rangle$ for all vectors $|x\rangle$ in the computational basis.

Definition 2. We will say that a phase oracle O admits an *uncomputable decomposition* (O_u, O_p) if $O = O_u^\dagger \circ O_p \circ O_u$. We call O_u and O_p the uncomputable part and the phase part, respectively.

Remark 3. Note that neither O_u nor O_p need to be phase oracles themselves. Naturally, every phase oracle O on n qubits admits a trivial decomposition (Id_n, O) . However, in practice, many real-life examples give more interesting decompositions. The intuitive goal is to make the phase part as simple as possible. The more gates needed to implement O we manage to move to the uncomputable part, the more gates we can hope to cancel out. A common pattern in many settings is to use an ancilla to mark the sought state by a bit flip, apply Z gate on said ancilla, followed by uncomputing the ancilla. In such case, there is a natural uncomputable decomposition of O .

Definition 3. Let A_1, \dots, A_ℓ be a chronologically ordered sequence of unitary matrices corresponding to the gates in a quantum circuit operating on n qubits (ties being broken

arbitrarily). We will define the fact that A_j depends on qubit q_i , $i \in \{1, \dots, n\}$, by induction on $j \in \{1, \dots, \ell\}$.

We say that A_j depends on q_i if A_j acts on q_i or if there exist $i_0 \in \{1, \dots, n\}$ and $j_0 \in \{1, \dots, j - 1\}$ such that the following hold:

- (i) A_{j_0} depends on q_i .
- (ii) A_{j_0} acts on q_{i_0} .
- (iii) A_j acts on q_{i_0} .

In this section, by $\mathcal{P}(X)$ we will denote the powerset of X , that is, the set of all subsets of X .

Definition 4. Let O be a phase oracle on n qubits, $\ell \in \mathbb{N}$, $d : \{1, \dots, \ell\} \rightarrow \mathcal{P}(\{q_1, \dots, q_n\})$, and $U : \{1, \dots, \ell\} \ni j \mapsto U_j \in U(n)$. Assume that U_j is an arbitrary unitary operator acting on the qubit set $d(j)$, $j \in \{1, \dots, \ell\}$. We define a *generic oracle circuit* $V(\ell, d, U, O)$ by the following formula (the product is to be understood as right-to-left operator composition):

$$V(\ell, d, U, O) := \prod_{j=1}^{\ell} (U_j \circ O).$$

Remark 4. Observe that W_m from Definition 1 is a generic oracle circuit.

Proof. For $j \in \{1, \dots, m\}$ put $\ell_j := (3^j - 1)/2$ and define $d_j : \{1, \dots, \ell_j\} \mapsto \mathcal{P}(\{1, \dots, k_1 + \dots + k_j\})$ recursively as follows:

$$d_j(i) := \begin{cases} d_{j-1}(i), & 1 \leq i \leq \ell_{j-1} \\ \{k_1 + \dots + k_{j-1} + 1, \dots, k_1 + \dots + k_j\}, & i = \ell_{j-1} + 1 \\ d_{j-1}(2\ell_{j-1} + 2 - i), & \ell_{j-1} + 2 \leq i \leq 2\ell_{j-1} + 1 \\ d_{j-1}(i - 2\ell_{j-1} - 1), & 2\ell_{j-1} + 2 \leq i \leq \ell_j. \end{cases}$$

We then set U_j to be the mixing operator $G_{|d_m(j)|}$ applied onto the qubits in the set $d_m(j)$, $j \in \ell_m$. Observe that $W_m = V(\ell_m, d_m, U, O)$. ■

C. Reducing the number of gates

Theorem 8. Let (O_u, O_p) be an uncomputable decomposition of O and let D_u , and D_p , be the total number of gates used in O_u , and O_p , respectively. Let \bar{D}_s denote the number of gates within O_u that depend on any of the qubits in s , for all $s \in \mathcal{P}(\{1, \dots, n\})$.

For a given generic oracle circuit $V(\ell, d, U, O)$ one can implement an equivalent circuit \tilde{V} that uses a total of $2D_u + \ell D_p + 2 \sum_{j=1}^{\ell} \bar{D}_{d(j)}$ gates for oracle queries. This results in no more than the following average number of gates per oracle query:

$$D_p + 2 \frac{D_u}{\ell} + 2 \frac{\sum_{j=1}^{\ell} \bar{D}_{d(j)}}{\ell}.$$

Proof. Let O_s (respectively \tilde{O}_s) be the in-order composition of all gates in O_u that depend (respectively do not depend) on any of the qubits in s , for all $s \in \mathcal{P}(\{1, \dots, n\})$.

First, let us observe that, for a fixed $s \in \mathcal{P}(\{1, \dots, n\})$, every gate in $A \in O_s$ commutes with every gate in \tilde{O}_s that originally appears later than A , as there are no common qubits that they act on. This implies that $O_u = O_s \circ \tilde{O}_s$, for all $s \in \mathcal{P}(\{1, \dots, n\})$. Similarly U_j and $\tilde{O}_{d(j)}$ commute, as there are no common qubits that they act on, $j = 1, \dots, \ell$.

We now proceed to apply these properties to $V(\ell, d, U, O)$ in order to obtain \tilde{V} . This will happen in the following five steps.

(1) Append an identity $O_u^\dagger \circ O_u$ operation to each factor of $V(\ell, d, U, O)$ (see Fig. 3).

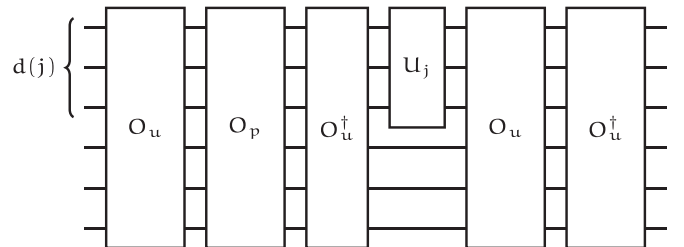


FIG. 3. Graphical representation of the j th factor after applying step 1. For simplicity we assumed that $d(j)$ consists of first $|d(j)|$ qubits.

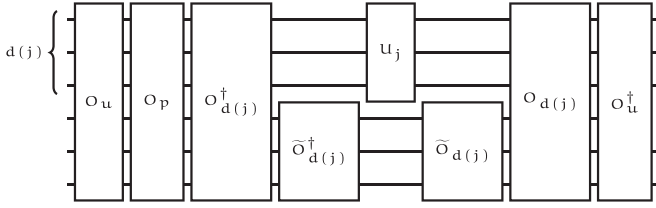


FIG. 4. Graphical representation of the j th factor after applying step 2. For simplicity we assumed that $d(j)$ consists of first $|d(j)|$ qubits.

(2) Express O_u (respectively O_u^\dagger) occurring next to U_j as $O_{d(j)} \circ \tilde{O}_{d(j)}$ (respectively $\tilde{O}_{d(j)}^\dagger \circ O_{d(j)}^\dagger$) in the j th factor, $j \in \{1, \dots, \ell\}$ (see Fig. 4).

(3) Swap $\tilde{O}_{d(j)}$ and U_j in the j th factor, $j \in \{1, \dots, \ell\}$ (see Fig. 5).

(4) Remove the identity operation $\tilde{O}_{d(j)}^\dagger \circ \tilde{O}_{d(j)}$ in the j th factor, $j \in \{1, \dots, \ell\}$ (see Fig. 6).

(5) Remove the identity operation $O_u \circ O_u^\dagger$ on the boundary between each two consecutive factors (see Fig. 7).

More concisely, we put

$$\tilde{V} := O_u^\dagger \circ \prod_{j=1}^{\ell} (O_{d(j)} \circ U_j \circ O_{d(j)}^\dagger \circ O_p) \circ O_u.$$

As discussed above $V(\ell, d, U, O)$ and \tilde{V} are equal as unitary operators. The desired gate count follows directly from the definition. ■

Corollary 4. Let (O_u, O_p) be an uncomputable decomposition of O and let D_u and D_p be the total number of gates used in O_u and O_p , respectively. Let D_i be the number of gates within O_u that depend on the i th qubit, $i \in \{1, \dots, n\}$, and let \bar{D}_j be the average of D_i taken over the qubits $i \in d(j)$, i.e., the average number of gates within O_u that depend on a single qubit from the $d(j)$, $j = 1, \dots, \ell$.

For a given generic oracle circuit $V(\ell, d, U, O)$ one can implement an equivalent circuit \tilde{V} that uses no more than the following average number of gates per oracle query:

$$D_p + 2 \frac{D_u + \sum_{j=1}^{\ell} |d(j)| \bar{D}_j}{\ell}.$$

Proof. This follows directly from Theorem 8 and the fact that the number of gates that depend on any of the qubits from a given set is no greater than the sum of the numbers of gates that depend on individual qubits from this set. ■

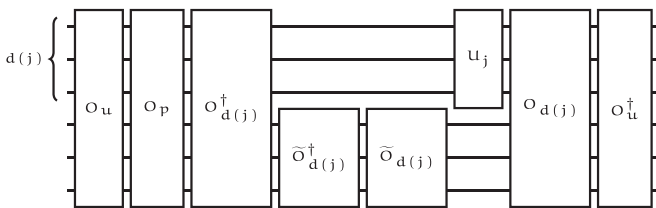


FIG. 5. Graphical representation of the j th factor after applying step 3. For simplicity we assumed that $d(j)$ consists of first $|d(j)|$ qubits.

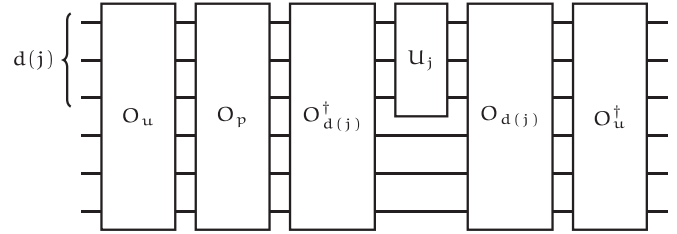


FIG. 6. Graphical representation of the j th factor after applying step 4. For simplicity we assumed that $d(j)$ consists of first $|d(j)|$ qubits.

Corollary 5. Let (O_u, O_p) be an uncomputable decomposition of O and let D_u , and D_p , be the total number of gates used in O_u , and O_p , respectively. Let D_i be the number of gates within O_u that depend on the i th qubit, $i \in \{1, \dots, n\}$, and let D be the average of D_i taken over all qubits $i \in \{1, \dots, n\}$.

For a given generic oracle circuit $V(\ell, d, U, O)$ one can implement an equivalent circuit \tilde{V} that uses no more than the following average number of gates per oracle query:

$$D_p + \frac{2D_u}{\ell} + 2D \frac{\sum_{j=1}^{\ell} |d(j)|}{\ell}.$$

In particular, one can implement a circuit equivalent to W_m that uses an average of no more than the following number of gates per oracle query:

$$D_p + \frac{4D_u}{3^m - 1} + 4D \frac{\sum_{j=1}^m k_j 3^{m-j}}{3^m - 1}.$$

Using the notation of Theorem 3, this asymptotic average number of gates is $O[D_p + \log(1/\varepsilon)D]$.

Proof. Without loss of generality, we may assume that $(D_i)_{i=1}^n$ is nondecreasing. Similarly, without loss of generality, we may assume that the weights $w_i = |\{j \in \{1, \dots, \ell\} : i \in d(j)\}|/\ell$, $i = 1, \dots, n$, of subsequent qubits form a nonincreasing sequence. Then a weighted average of D_i 's,

$$\frac{\sum_{i=1}^n w_i D_i}{\sum_{i=1}^n w_i} = \frac{\sum_{j=1}^{\ell} \sum_{i \in d(j)} D_i}{\ell \sum_{i=1}^n w_i} = \frac{\sum_{j=1}^{\ell} |d(j)| \bar{D}_j}{\sum_{j=1}^{\ell} |d(j)|},$$

will not be greater than their arithmetic mean D . This completes the first part of the proof. To obtain bound for circuit W_m , recall from Sec. III that $\ell = (3^m - 1)/2$ and that j th

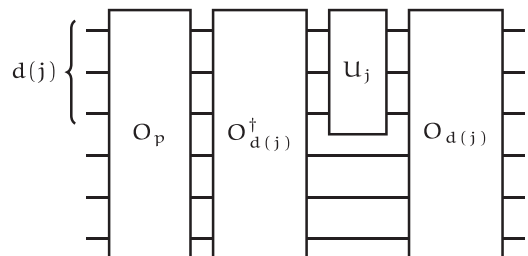


FIG. 7. Graphical representation of the j th factor after applying step 5. For simplicity we assumed that $d(j)$ consists of first $|d(j)|$ qubits. Notice that this picture is only valid for $1 < j < \ell$ as the first factor will retain O_u^\dagger , while the last one will retain O_u .

diffuser (of size k_j) appears in W_m exactly 3^{m-j} times, so $\sum_{j=1}^{\ell} |d(j)| = \sum_{j=1}^m k_j 3^{m-j}$.

For the asymptotic result, recall from the proof of Theorem 3 [Eq. (8)] that (after setting $x \in \Theta[\log(1/\varepsilon)]$ as in the proof) we have $m = \Theta(\sqrt{n/\log(1/\varepsilon)})$ and we invoke the W_m circuit $O(2^{n/2}/3^m)$ times with parameters $k_j \leq 3\lceil \log_2(1/\varepsilon) \rceil j$, $j \in \{1, \dots, m\}$. Additionally, observe that

$$\begin{aligned} \sum_{j=1}^m j 3^{m-j} &= \sum_{j=1}^m \frac{3^{m-j+1} - 1}{2} = \frac{3}{4}(3^m - 1) - \frac{m}{2} \\ &= O(3^m - 1). \end{aligned}$$

As each amplitude amplification step requires additional $O(n + D_p + D_u)$ gates, we get that the asymptotic average number of gates is

$$\begin{aligned} D_p + O(D_u/3^m) + \log(1/\varepsilon)O(D) + O[(n + D_p + D_u)/3^m] \\ = O[D_p + \log(1/\varepsilon)D]. \end{aligned}$$

■

Corollary 6. Let $K \in \mathbb{N}$ and consider all Unique k -SAT instances, each with n variables and c clauses. For each such an instance there exists a circuit equivalent to W_m that uses an average of no more than the following number of gates per query:

$$\begin{aligned} 1 + \frac{12Kc + 8c - 4}{3^m - 1} \\ + \frac{4Kc(4 + \lceil \log_2 K \rceil + \lceil \log_2 c \rceil) \sum_{j=1}^m k_j 3^{m-j}}{n(3^m - 1)}. \end{aligned}$$

Using the notation of Theorem 3, the asymptotic average per-oracle-query number of gates of a quantum circuit solving Unique k -SAT with certainty is $O[\log(1/\varepsilon)c \log(c)/n]$.

Proof. A straightforward implementation of the phase oracle O consists of $D_u \leq 3Kc + 2c - 1$ gates (each being an X , a CX , or a CCX) in the uncomputable part and one Z gate ($D_p = 1$) in the phase part. More precisely, we introduce the following ancilla qubits and the gates to compute them:

(1) c -qubit groups of K qubits corresponding to negation of all clause literals, each computed by one CX and at most one X .

(2) c -qubit groups of $K - 1$ qubits corresponding to conjunctions of qubits from 1, each computed by one CCX . They are arranged into a binary tree, so that only $\lceil \log_2 K \rceil$ gates depend on every qubit of 1.

(3) c qubits corresponding to all clauses, each computed by one CX and one X from a top-level qubit of 2.

(4) $c - 1$ qubits corresponding to conjunctions of qubits from (3), each computed by one CCX . Again, they are arranged into a binary tree, so that only $\lceil \log_2 c \rceil$ gates depend on every qubit from 3.

For a variable v appearing in c_v clauses, the number of gates depending on v is at most $c_v(2 + \lceil \log_2 K \rceil + 2 + \lceil \log_2 c \rceil)$, so we get the average $D \leq Kc(4 + \lceil \log_2 K \rceil + \lceil \log_2 c \rceil)/n$. By Corollary 5 we get both claims. ■

Corollary 1. Consider the Unique k -SAT problem with n variables and c clauses. There exists a quantum circuit that uses $O[c \log(c)2^{n/2}/n]$ total (oracle and nonoracle) gates and solves the problem with certainty.

Proof. This follows directly from Theorem 3 and Corollary 6. ■

VI. MULTIPOINT ORACLE

We now proceed to the unstructured search problem with multiple marked elements. As in previous sections we assume that number of qubits in the input of the oracle is n . Let S be the set of elements marked by oracle O and let $K = |S| > 0$. For convenience we mostly refer to the number $k = 1 + \lceil \log_2 K \rceil$. We begin our investigation with the assumption that k is known in advance and later proceed to consider the harder case of unknown k .

A. Known number of marked elements

In this section we assume that value k is known. It is weaker assumption than knowing K but it is sufficient for our purposes. We use the algorithm from Theorem 3 as a subroutine in algorithms in this section. By $\text{SinglePoint}(O, n)$ we denote the algorithm from Theorem 3 that solves unstructured search problem for oracle O which marks exactly one element and acts on n qubits. We want to reduce the problem of unstructured search with possibly many elements marked to the unstructured search with one marked element. To do this we construct a family of hash functions that allows us to effectively parametrize a subset of $\{0, 1\}^n$ which with high probability contains only one element from S . This technique is nearly identical to reduction from SAT to Unique SAT presented in [11]. Next, we improve some aspects of this reduction, so that methods of partial uncompute may be used to reduce the number of additional nonoracle basic gates.

Let us recall that family U of hash functions from X to Y , both being finite is called *pairwise independent* if for every $x \in X$ and every $y \in Y$ we have

$$\mathbb{P}_{h \in R^U}[h(x) = y] = \frac{1}{|Y|}$$

and for every $x_1, x_2 \in X, x_1 \neq x_2$ and every $y_1, y_2 \in Y$ we have

$$\mathbb{P}[h(x_1) = y_1 \wedge h(x_2) = y_2] = \frac{1}{|Y|^2}.$$

We use the following formulation of results from [11] that can be found in [17, p. 9.10 (180)].

Lemma 7 (Valiant-Vazirani). For any family of pairwise-independent hash functions H from $\{0, 1\}^n$ to $\{0, 1\}^k$ and $S \subset \{0, 1\}^n$ such that $2^{k-2} \leq |S| \leq 2^{k-1}$ we have that

$$\mathbb{P}_{h \in R^H}(|\{x \in S : h(x) = 0\}| = 1) \geq \frac{1}{8}.$$

A function $h : \{0, 1\}^m \rightarrow \{0, 1\}^l$ is called *affine* if we may represent it as $h(x) = Ax + b$ for some $A \in \{0, 1\}^{l \times m}$ and $b \in \{0, 1\}^l$. All arithmetical operations are performed modulo 2. The *kernel* of the affine function $h : \{0, 1\}^m \rightarrow \{0, 1\}^l$ of the form $h(x) = Ax + b$ is defined as $\ker h = h^{-1}(0)$. Whenever the kernel of the affine function h is not an empty set, we define the dimension of this kernel as $\dim \ker h = \dim \ker A$, where $\ker A$ is the null space of the matrix A . Whenever function h is clear from the context we will use $d = \dim \ker h$ for brevity. Our choice of the family of hash functions is as follows.

Definition 5. Define a family of hash functions $H_{n,k}$ as the set of all affine maps from $\{0, 1\}^n$ to $\{0, 1\}^k$:

$$H_{n,k} := \{h_{A,b} : A \in \{0, 1\}^{k \times n}, b \in \{0, 1\}^k, h_{A,b}(x) = Ax + b\}.$$

The first mention of this family is in [18], more detailed considerations can be found in [19]. The following standard result will be of use to us.

Observation 7 (Folklore). The family $H_{n,k}$ is pairwise independent.

We would like to run algorithm `SinglePoint` on the set $\ker h$. To do so we parametrize $\ker h$ by some injection $g : \{0, 1\}^{\dim \ker h} \rightarrow \ker h$ and build a quantum oracle O_g defined as follows.

Definition 6. Given a quantum oracle $O : (\mathbb{C}^2)^{\otimes n} \rightarrow (\mathbb{C}^2)^{\otimes n}$, and any function $g : \{0, 1\}^d \rightarrow \{0, 1\}^n$ we define the g -restricted oracle O_g as

$$O_g = D_g^{-1}(Id_d \otimes O)D_g,$$

where D_g is a unitary operator on $(\mathbb{C}^2)^{\otimes(d+n)}$ whose action on states $|i\rangle|0 \dots 0\rangle$ for $i \in \{0, 1\}^d$ is defined as

$$D_g|i\rangle|0 \dots 0\rangle = |i\rangle|g(i)\rangle.$$

Observation 8. If oracle O admits uncomputable decomposition (O_u, O_p) , then for any function $g : \{0, 1\}^d \rightarrow \{0, 1\}^n$ the g -restricted oracle O_g admits an uncomputable decomposition $((Id_d \otimes O_u)D_g, Id_d \otimes O_p)$.

Proof. It follows directly from definition

$$\begin{aligned} O_g &= D_g^{-1}(Id_d \otimes O)D_g \\ &= D_g^{-1}(Id_d \otimes O_u^{-1})(Id_d \otimes O_p)(Id_d \otimes O_u)D_g \\ &= ((Id_d \otimes O_u)D_g)^{-1}(Id_d \otimes O_p)((Id_d \otimes O_u)D_g). \end{aligned}$$

Lemma 8. For an affine injective function $g : \{0, 1\}^d \rightarrow \ker h$ of the form $g(x) = Cx + p$, where $C = (c_{ij})$, we can construct D_g using basic quantum gates so that the number of gates that depend on j th qubit of the first register is exactly equal to $|i : c_{ij} = 1|$.

Proof. It is easy to see that D_g can be implemented using gates $CX(f_j, s_i)$ where f_j is the j th qubit of the first register and s_i is the i th qubit of the second register for each i, j such that $c_{ij} = 1$ and using gates $X(s_i)$ for all i such that $p_i > 0$. ■

Now we are ready to construct g which effectively parametrizes $\ker h$.

Lemma 9. Given an affine function $h : \{0, 1\}^n \rightarrow \{0, 1\}^k$ of the form $h(x) = Ax + b$ with $\ker h \neq \emptyset$ we may construct in polynomial time an injective function $g : \{0, 1\}^d \rightarrow \{0, 1\}^n$ of the form $g(x) = Cx + p$ for some $C \in \{0, 1\}^{n \times d}$, $p \in \{0, 1\}^n$, and d , where $d = \dim \ker h$, such that $\text{Im} g = \ker h$. Moreover, we may choose C so that each of its columns has at most $n - d + 1$ ones.

Proof. We begin by obtaining an arbitrary affine parametrization of $\ker h$. To this end, fix some basis of $\ker A$, arrange it as columns into the matrix C' , and any solution p to the equation $Ax = -b$. All of this can be accomplished in polynomial time using Gaussian elimination [20]. Setting $f(x) = C'x + p$ gives us the desired parametrization.

To reduce the number of ones in the matrix C' , we can change basis of domain of f by an invertible matrix $Q \in$

$\{0, 1\}^{d \times d}$. As function f is an injection, matrix C' has d rows which are linearly independent and they form an invertible submatrix C'' . By picking $Q = (C'')^{-1}$ we ensure that for each column of matrix $C'Q$ at most one row among those d picked previously has one which is contained in this column. After these steps we end up with a function g of the form $g(x) = C'Qx + p$, where each column of $C = C'Q$ has at most $n - d + 1$ nonzero entries. ■

As we parametrize the kernel of random affine map we want to make sure that dimension of this kernel is not too big, as the number of oracle queries and the number of nonoracle basic gates used by `SinglePoint` depends exponentially on the dimension of the searched space.

Lemma 10. If $k < n - 2$, then $\mathbb{P}_{h \in_R H_{n,k}}(\dim \ker h \geq n - k + 2) \leq \frac{1}{16}$.

Proof. We prove the more general inequality $\mathbb{P}_{h \in_R H_{n,k}}(\dim \ker h \geq n - k + c) \leq \frac{1}{2^{c^2}}$ for $n > 4$, $k < n - 2$, and any natural $c \geq 2$. Set $\delta = k - c$. If $\delta < 0$, then the conclusion follows trivially. Otherwise, the event $\dim \ker h \geq n - k + c$ is equivalent to $n - \delta = n - k + c \leq \dim \ker h = \dim \ker A = n - \text{rank} A$ so we conclude $\text{rank} A \leq \delta$, meaning that vector subspace spanned by rows of matrix A must have dimension at most δ .

As vectors that span subspace of dimension at most δ are contained in some δ -dimensional subspace of $\{0, 1\}^n$ we can consider the probability of all k vectors being contained in a particular δ -dimensional subspace. Then we apply union bound by multiplying this probability by the number of δ -dimensional subspaces. The probability of choosing all k vectors from a single δ -dimensional space is $(\frac{1}{2^{n-\delta}})^k$, as δ -dimensional space contains 2^δ elements and we choose those vectors independently from each other. Let us recall that a number of δ -dimensional subspaces of n -dimensional space over \mathbb{F}_2 equals $\prod_{i=0}^{\delta-1} \frac{2^n - 2^i}{2^\delta - 2^i}$, this can be found in [21].

So, using the union bound the probability that k vectors span at most δ -dimensional subspace is bounded from above by

$$\begin{aligned} &\mathbb{P}_{h_{A,b} \in H_{n,k}}(\dim \ker h \geq n - \delta) \\ &\leq \left(\frac{1}{2^{n-\delta}}\right)^k \prod_{i=0}^{(\delta-1)} \frac{2^n - 2^i}{2^\delta - 2^i} \\ &= \left(\frac{1}{2^{n-\delta}}\right)^{k-\delta} \left(\frac{1}{2^{n-\delta}}\right)^\delta \prod_{i=0}^{(\delta-1)} \frac{2^n - 2^i}{2^\delta - 2^i} \\ &= \left(\frac{1}{2^{n-\delta}}\right)^{k-\delta} \prod_{i=0}^{(\delta-1)} \frac{2^\delta - 2^{i-n+\delta}}{2^\delta - 2^i} \\ &\leq \left(\frac{1}{2^{n-\delta}}\right)^{k-\delta} \prod_{i=0}^{(\delta-1)} \frac{2^\delta}{2^\delta - 2^i} \\ &= \left(\frac{1}{2^{n-\delta}}\right)^{k-\delta} \prod_{i=0}^{(\delta-1)} \frac{2^{\delta-i}}{2^{\delta-i} - 1} \\ &\leq \left(\frac{1}{2^{n-\delta}}\right)^{k-\delta} \prod_{j=1}^{\infty} \frac{2^j}{2^j - 1} \end{aligned}$$

Algorithm 1 Probabilistic algorithm for solving unstructured search with known k

```

1: procedure MULTIPPOINT ( $O, n, k$ )
2:   if  $k \geq n - 2$  then
3:      $x \leftarrow$  element from  $\{0, 1\}^n$  selected uniformly at random
4:     if  $x$  is marked then
5:       return  $x$ 
6:     end if
7:     return null
8:   end if
9:    $h \leftarrow$  random affine transformation from  $H_{n,k}$ 
10:   $d \leftarrow \dim \ker h$ 
11:  if  $d \geq n - k + 2$  then
12:    return null
13:  end if
14:   $O_g$  is built as described in Definition 6 using  $g$  from Lemma 9
15:  return SinglePoint( $O_g, d$ )
16: end procedure

```

$$\begin{aligned}
 &= \left(\frac{1}{2^{n-\delta}}\right)^{k-\delta} \frac{1}{\prod_{j=1}^{\infty} (1-2^{-j})} \\
 &\leq \left(\frac{1}{2^{n-\delta}}\right)^{k-\delta} \frac{1}{1-\frac{1}{2}-\frac{1}{4}} \\
 &= 4 \left(\frac{1}{2^{n-\delta}}\right)^{k-\delta} = 4 \left(\frac{1}{2^{n-k+c}}\right)^c = 4 \left(\frac{1}{2^{n-k}}\right)^c \left(\frac{1}{2^c}\right)^c,
 \end{aligned}$$

where the last inequality follows from Euler’s pentagonal numbers theorem [15]. The final expression is less than $\frac{1}{2^{c^2}}$ for $c \geq 2$ and $n - k > 2$. ■

Now, we may describe the algorithm for solving unstructured search problem with known value $k = 1 + \lceil \log_2 K \rceil$ where K is the number of marked elements.

Theorem 4. Let $N \in \mathbb{N}$ be of the form $N = 2^n$. Assume that we are given a phase oracle O that marks K elements, and we know the number k given by $k = 1 + \lceil \log_2 K \rceil$. Then one can find an element marked by O with probability at least $\frac{1}{16}$, using at most $O(\sqrt{\frac{N}{K}})$ oracle queries and at most $O(\log K \sqrt{\frac{N}{K}})$ nonoracle basic gates.

Proof. To prove that Algorithm 1 finds a marked element with constant probability, let us see that if $k \geq n - 2$, then selecting a random element succeeds with probability at least $\frac{1}{16}$. Otherwise, from Lemma 7 with probability at least $\frac{1}{8}$ we have that $|K \cap \ker h| = 1$. From Lemma 10 with probability at least $\frac{15}{16}$ we have that $d < n - k + 2$. Combining those facts we obtain that with probability at least $\frac{1}{16}$ oracle O_g marks exactly one element and the number of qubits of its input does not exceed $n - k + 1$. So, Algorithm 1 succeeds with probability at least $\frac{1}{16}$, as from Theorem 3 we know that SinglePoint solves the unstructured problem with one marked element with certainty.

From Lemmas 9 and 8 we deduce that at most $O(k)$ additional basic gates from circuit D_g depend on each qubit. So, from Corollary 5 we deduce that on average we use $O(k)$ additional nonoracle basic gates per oracle query. There are $O(2^{\frac{d}{2}}) = O(2^{\frac{n-k}{2}}) = O(\sqrt{\frac{N}{K}})$ oracle queries in SinglePoint

Algorithmic 2 Probabilistic algorithm for solving unstructured search problem without an estimate of the number of marked elements

```

1: procedure MULTIPPOINTUNKNOWN ( $O, n, p$ )
2:   for  $i \leftarrow n + 2$  to 2 do
3:     for  $j \leftarrow n + 2$  to  $i$  do
4:        $x \leftarrow$  MultiPointAmplified( $O, n, j, p$ )
5:       if  $x$  is marked then
6:         return  $x$ 
7:       end if
8:     end for
9:   end for
10: end procedure

```

procedure so we need $O(k2^{\frac{n-k}{2}}) = O(\log K \sqrt{\frac{N}{K}})$ nonoracle gates to implement it. ■

Proposition 1. For any $p \in (0, 1)$ by repeating Algorithm 1 $O(\log \frac{1}{1-p})$ number of times we ensure that we find a marked element with probability at least p .

Proof. We may deduce that from the fact that all runs of this algorithm are independent and each finishes successfully with constant, nonzero probability. ■

Let us for any probability $p < 1$ define an algorithm MultiPointAmplified(O, n, k, p) which runs algorithm MultiPoint(O, n, k) minimal number of times to ensure probability of success higher than p .

B. Unknown number of marked elements

The technique presented next is similar to one used in [16], which finds element marked by an oracle O using on average $O(2^{\frac{n-k}{2}})$ calls to oracle O and on average $O(n2^{\frac{n-k}{2}})$ additional basic gates. We improve those results and propose the following algorithm that finds marked element using $O(k2^{\frac{n-k}{2}})$ nonoracle gates in expectation and makes $O(2^{\frac{n-k}{2}})$ queries to oracle O also in expectation.

Before we analyze Algorithm II, we note an observation:

Observation 9. For natural numbers x and m , and any real number r , such that $x \leq m$ and $r > 1$, we have

$$\begin{aligned}
 &\sum_{l=0}^x (m-l)r^{l/2} \\
 &= (m-x) \sum_{l=0}^x r^{l/2} + \sum_{l=0}^x (x-l)r^{l/2} \\
 &= (m-x)r^{x/2} \sum_{l=0}^x r^{(l-x)/2} + r^{x/2} \sum_{l=0}^x (x-l)r^{(l-x)/2} \\
 &\leq (m-x)r^{x/2} \sum_{i=0}^{\infty} r^{-i/2} + r^{x/2} \sum_{i=0}^{\infty} ir^{-i/2} \\
 &= C_1(m-x)r^{x/2} + C_2r^{x/2},
 \end{aligned}$$

where C_1, C_2 are positive constants which depend only on r .

Observation 10. For natural numbers x and m , and any real number r , such that $x \leq m$ and $r < 1$, we have

$$\sum_{l=0}^x (m-l)r^{l/2} \leq Cm,$$

where C is a positive constant which depends only on r .

Theorem 9. For p satisfying $2(1-p)^2 < 1$ the Algorithm 2 finds marked element with probability at least $1 - (1-p)^k$. Its expected number of oracle queries is $O(\sqrt{\frac{N}{K}})$ and its expected number of nonoracle basic gates is $O(\log K \sqrt{\frac{N}{K}})$, where $N = 2^n$ is the size of the search space and K is the number of elements marked by the oracle and $k = 1 + \lceil \log_2 K \rceil$.

Proof. In the complexity analysis we consider two phases of the Algorithm II. The first phase is when $i > k$. During this phase we never run algorithm `MultiPointAmplified`(O, n, j, p) with $j = k$, so let us assume that this algorithm never finds marked element in this phase.

In the second phase, i.e., for $i < k$ in each inner loop we run the procedure

`MultiPointAmplified`(O, n, j, p) with $j = k$ once, so during this loop we find marked element with probability at least p . So the probability that outer loop proceeds to the next iteration is at most $1 - p$. So, the overall bound on the expected number of oracle queries of this algorithm is given below, we also note that all constants hidden under big O notation depend either only on p or are universal:

$$\begin{aligned} & O\left(\sum_{i=k+1}^{n+2} \sum_{j=i}^{n+2} 2^{(n-j)/2} + \sum_{i=2}^k (1-p)^{k-i} \sum_{j=i}^{n+2} 2^{(n-j)/2}\right) \\ &= O\left(\sum_{i=k}^n \sum_{j=i}^n 2^{(n-j)/2} + \sum_{i=0}^k (1-p)^{k-i} \sum_{j=i}^n 2^{(n-j)/2}\right) \\ &= O\left(\sum_{i=k}^n \sum_{l=0}^{n-i} 2^{l/2} + \sum_{i=0}^k (1-p)^{k-i} \sum_{l=0}^{n-i} 2^{l/2}\right) \\ &= O\left(\sum_{i=k}^n 2^{(n-i)/2} + \sum_{i=0}^k (1-p)^{k-i} 2^{(n-i)/2}\right) \\ &= O\left(\sum_{s=0}^{n-k} 2^{s/2} + 2^{(n-k)/2} \sum_{s=0}^k (2(1-p)^2)^{s/2}\right) \\ &= O(2^{(n-k)/2}). \end{aligned}$$

To estimate the second summand we use the fact that $2(1-p)^2 < 1$. Using Observations 9 and 10 we calculate the similar

bound for the number of additional nonoracle basic gates, also take a note that hidden constants below depend only on p or are universal:

$$\begin{aligned} & O\left(\sum_{i=k+1}^{n+2} \sum_{j=i}^{n+2} j2^{(n-j)/2} + \sum_{i=2}^k (1-p)^{k-i} \sum_{j=i}^{n+2} j2^{(n-j)/2}\right) \\ &= O\left(\sum_{i=k}^n \sum_{j=i}^n j2^{(n-j)/2} + \sum_{i=0}^k (1-p)^{k-i} \sum_{j=i}^n j2^{(n-j)/2}\right) \\ &= O\left(\sum_{i=k}^n \sum_{l=0}^{n-i} (n-l)2^{l/2} + \sum_{i=0}^k (1-p)^{k-i} \sum_{l=0}^{n-i} (n-l)2^{l/2}\right) \\ &= O\left(\sum_{i=k}^n i2^{(n-i)/2} + \sum_{i=0}^k i(1-p)^{k-i} 2^{(n-i)/2}\right) \\ &= O\left(\sum_{s=0}^{n-k} (n-s)2^{s/2} + 2^{(n-k)/2} \sum_{s=0}^k (k-s)(2(1-p)^2)^{s/2}\right) \\ &= O(k2^{(n-k)/2}). \end{aligned}$$

So the complexity of the algorithm does not change even if the number of elements is not known beforehand. To calculate the probability of successfully finding the marked element, let us see that the outer loop runs less than $n + 1$ times only when the marked element was found. From the above considerations we know that this probability is bounded from below by $1 - (1-p)^k$. ■

ACKNOWLEDGMENTS

The following work has been partially supported by NCBR grant number POIR.01.01.01-00-0568/16-00.

APPENDIX: ANALYSIS OF THE TREE CIRCUIT

Due to limited nature of existing hardware, for small search spaces the circuits W_m are outperformed by a similar family of circuits, which we denote by D_m . The circuits were experimentally evaluated on current generation of superconducting quantum computers. The results are presented in [22]. For the sake of completeness we prove an analog of Theorem 3 that utilizes the D_m family of circuits.

Definition 7. Let $\vec{k} = (k_1, \dots, k_m)$ be a sequence of positive integers and let $n := \sum_{j=1}^m k_j$. Given a quantum oracle O , for $j \in \{0, \dots, m\}$, we define the circuit D_j recursively as follows:

$$D_j = \begin{cases} \text{Id}_n, & \text{if } j = 0 \\ D_{j-1}(\text{Id}_{k_1+\dots+k_{j-1}} \otimes G_{k_j} \otimes \text{Id}_{k_{j+1}+\dots+k_m})OD_{j-1}, & \text{if } j \neq 0. \end{cases}$$

Lemma 11. Let $m \in \mathbb{N}_+$ and $\vec{k} \in \mathbb{N}_+^m$ be fixed, and let $n = \sum_{j=1}^m k_j$. Assume we are given a phase oracle O that operates on n qubits and marks a single vector of the standard computational basis, which we then use in the circuits D_j . Then for

any $j \in \{0, \dots, m\}$ we have

$$D_jOD_j = O.$$

Proof. We proceed by induction on j . For $j = 0$ the claim is trivial. For $j > 0$ we expand D_j according to Definition 7 as follows

$$\begin{aligned} D_j O D_j &= D_{j-1} (\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}) O D_{j-1} O D_{j-1} \\ &\quad \times (\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}) O D_{j-1} \\ &= D_{j-1} (\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}) O O \end{aligned}$$

$$\times (\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}) O D_{j-1} \quad (\text{A1})$$

$$\begin{aligned} &= D_{j-1} (\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}) \\ &\quad \times (\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}) O D_{j-1} \\ &= D_{j-1} O D_{j-1} = O, \end{aligned} \quad (\text{A2})$$

where in Eqs. (A1) and (A2) we used the inductive hypothesis. ■

Lemma 12. Let $m \in \mathbb{N}_+$ and $\bar{k} \in \mathbb{N}_+^m$ be fixed, and let $n = \sum_{j=1}^m k_j$. Assume we are given a phase oracle O that operates on n qubits and marks a single vector of the standard computational basis denoted target. Define the numbers

$$\beta_j = \langle \text{target} | (D_j |u_1^j\rangle | \text{target}_{j+1}^m \rangle) \rangle$$

for $j \in \{0, \dots, m\}$. Then, β_j satisfy the recurrence

$$\beta_j = \begin{cases} 1, & \text{if } j = 0 \\ \frac{1}{2^{(k_1+\dots+k_j)/2}} \left(1 - \frac{2}{2^{k_j}}\right) + \frac{1}{2^{k_j/2}} \left(2 - \frac{2}{2^{k_j/2}}\right) \beta_{j-1}, & \text{if } j > 0. \end{cases}$$

Proof. By definition of D_j we have $\beta_0 = 1$ giving the base case. For $j > 0$, we proceed to compute β_j by expanding the circuit D_j according to the recursive definition. We split the computation into stages as follows:

$$\begin{aligned} |w_1\rangle &= D_{j-1} (|u_1^j\rangle | \text{target}_{j+1}^m \rangle), \quad |w_2\rangle = O |w_1\rangle, \\ |w_3\rangle &= (\text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s}) |w_2\rangle, \quad |w_4\rangle = D_{j-1} |w_3\rangle, \end{aligned}$$

where $s = k_1 + \dots + k_j$,

$$\begin{aligned} |w_1\rangle &= D_{j-1} \left(\frac{1}{2^{k_j/2}} |u_1^{j-1}\rangle | \text{target}_j^m \rangle + |u_1^{j-1}\rangle | \overline{\text{target}}_j \rangle | \text{target}_{j+1}^m \rangle \right) = \frac{1}{2^{k_j/2}} D_{j-1} |u_1^{j-1}\rangle | \text{target}_j^m \rangle + |u_1^{j-1}\rangle | \overline{\text{target}}_j \rangle | \text{target}_{j+1}^m \rangle, \\ |w_2\rangle &= O \left(\frac{1}{2^{k_j/2}} D_{j-1} |u_1^{j-1}\rangle | \text{target}_j^m \rangle + |u_1^{j-1}\rangle | \overline{\text{target}}_j \rangle | \text{target}_{j+1}^m \rangle \right) \\ &= \frac{1}{2^{k_j/2}} O D_{j-1} |u_1^{j-1}\rangle | \text{target}_j^m \rangle + |u_1^{j-1}\rangle | \overline{\text{target}}_j \rangle | \text{target}_{j+1}^m \rangle = \frac{1}{2^{k_j/2}} |\eta\rangle | \text{target}_j^m \rangle + |u_1^{j-1}\rangle | \overline{\text{target}}_j \rangle | \text{target}_{j+1}^m \rangle, \end{aligned}$$

where $|\eta\rangle$ is some state in $(\mathbb{C}^2)^{\otimes(k_1+\dots+k_{j-1})}$. We can write so, as all diffusers within D_{j-1} operate only on the prefix consisting of first $k_1 + \dots + k_{j-1}$ qubits, while O only changes relative phases:

$$\begin{aligned} |w_3\rangle &= \text{Id}_{s-k_j} \otimes G_{k_j} \otimes \text{Id}_{n-s} \left(\frac{1}{2^{k_j/2}} |\eta\rangle | \text{target}_j^m \rangle + |u_1^{j-1}\rangle | \overline{\text{target}}_j \rangle | \text{target}_{j+1}^m \rangle \right) \\ &= \frac{1}{2^{k_j/2}} |\eta\rangle (G_{k_j} | \text{target}_j \rangle) | \text{target}_{j+1}^m \rangle + |u_1^{j-1}\rangle (G_{k_j} | \overline{\text{target}}_j \rangle) | \text{target}_{j+1}^m \rangle \\ &= \frac{1}{2^{k_j/2}} |\eta\rangle \left(\frac{2}{2^{k_j/2}} |u_j\rangle - | \text{target}_j \rangle \right) | \text{target}_{j+1}^m \rangle + |u_1^{j-1}\rangle \left[\left(1 - \frac{2}{2^{k_j}}\right) |u_j\rangle + \frac{1}{2^{k_j/2}} | \text{target}_j \rangle \right] | \text{target}_{j+1}^m \rangle \\ &= \frac{1}{2^{k_j/2}} \left[\left(\frac{2}{2^{k_j}} - 1\right) |\eta\rangle + \left(2 - \frac{2}{2^{k_j}}\right) |u_1^{j-1}\rangle \right] | \text{target}_j^m \rangle + \left[\frac{2}{2^{k_j}} |\eta\rangle + \left(1 - \frac{2}{2^{k_j}}\right) |u_1^{j-1}\rangle \right] | \overline{\text{target}}_j \rangle | \text{target}_{j+1}^m \rangle, \\ |w_4\rangle &= D_{j-1} |w_3\rangle = \frac{1}{2^{k_j/2}} \left[\left(\frac{2}{2^{k_j}} - 1\right) O |u_1^{j-1}\rangle | \text{target}_j^m \rangle + \left(2 - \frac{2}{2^{k_j}}\right) D_{j-1} |u_1^{j-1}\rangle | \text{target}_j^m \rangle \right] \\ &\quad + \left[\frac{2}{2^{k_j}} |\eta\rangle + \left(1 - \frac{2}{2^{k_j}}\right) |u_1^{j-1}\rangle \right] | \overline{\text{target}}_j \rangle | \text{target}_{j+1}^m \rangle. \end{aligned}$$

Note that we used Lemma 11 when applying D_{j-1} in the first summand. Now we can plug $|w_4\rangle$ into the expression defining β_j . Observe that the second summand in the final expression is orthogonal to $|\text{target}\rangle$, thus can be safely discarded. We obtain

$$\begin{aligned} \beta_j &= \langle \text{target} | \left\{ \frac{1}{2^{k_j/2}} \left[\left(\frac{2}{2^{k_j}} - 1 \right) O |u_1^{j-1}\rangle | \text{target}_j^m \rangle + \left(2 - \frac{2}{2^{k_j}} \right) D_{j-1} |u_1^{j-1}\rangle | \text{target}_j^m \rangle \right] \right\} \\ &= \frac{1}{2^{k_j/2}} \left[\left(1 - \frac{2}{2^{k_j}} \right) \frac{1}{2^{(s-k_j)/2}} + \left(2 - \frac{2}{2^{k_j}} \right) \beta_{j-1} \right] = \frac{1}{2^{(k_1+\dots+k_j)/2}} \left(1 - \frac{2}{2^{k_j}} \right) + \frac{1}{2^{k_j/2}} \left(2 - \frac{2}{2^{k_j/2}} \right) \beta_{j-1} \end{aligned}$$

concluding the proof. ■

Theorem 10. Fix any $\varepsilon > 0$, and any $N \in \mathbb{N}$ of the form $N = 2^n$. Suppose we are given a quantum oracle O operating on n qubits that marks exactly one element. Then there exists a quantum circuit \mathcal{A} which uses the oracle O at most $(1 + \varepsilon) \frac{\pi}{4} \sqrt{N}$ times and uses at most $O(\log({}^1/\varepsilon) \sqrt{N})$ nonoracle basic gates, which finds the element marked by O with certainty.

Proof. We first begin by choosing a particular sequence of sizes for diffusers in the circuit D_m , namely, $k_j = (x + 1)j$ where $x \in \mathbb{N}_+$ is some parameter, and let us assume that the number of qubits we work with is precisely $(x + 1) + (x + 1) \times 2 + \dots + (x + 1)m = (x + 1)m(m + 1)/2$. From Lemma 12, we get that the amplitude the circuit D_m in the marked state is given by the following recurrence

$$\beta_j = \begin{cases} 1, & \text{if } j = 0 \\ \frac{1}{2^{(x+1)j(j+1)/4}} \left(1 - \frac{2}{2^{(x+1)j}} \right) + \frac{1}{2^{(x+1)j/2}} \left(2 - \frac{2}{2^{(x+1)j/2}} \right) \beta_{j-1}, & \text{if } j > 0. \end{cases}$$

To simplify the analysis of this recurrence, let us begin by substituting $\gamma_j = \beta_j 2^{(x+1)j(j+1)/4} \times 2^{-j}$, which yields

$$\gamma_j = \begin{cases} 1, & \text{if } j = 0 \\ (1 - 2 \times 2^{-(x+1)j}) 2^{-j} + (1 - 2^{-(x+1)j}) \gamma_{j-1}, & \text{if } j > 0. \end{cases}$$

We easily obtain the following inequality for $j > 0$:

$$\gamma_j \geq (1 - 2^{-xj})(2^{-j} + \gamma_{j-1}).$$

We may thus set

$$\delta_j = \begin{cases} 1, & \text{if } j = 0 \\ (1 - 2^{-xj})(2^{-j} + \delta_{j-1}), & \text{if } j > 0 \end{cases}$$

and we easily obtain the inequality $\gamma_j \geq \delta_j$. We can express the solution to this recurrence as a sum

$$\delta_m = \sum_{j=0}^{m-1} 2^{j-m} \prod_{k=m-j}^m (1 - q^k) + \prod_{k=1}^m (1 - q^k),$$

where $q = 2^{-x}$.

For $a \in \mathbb{N} \cup \{\infty\}$, let

$$\mathcal{P}(a) = \prod_{k=1}^a (1 - q^k).$$

In terms of $\mathcal{P}(a)$, we can lower bound δ_m , as each term in our product is strictly less than 1, thus,

$$\begin{aligned} \delta_m &\geq \sum_{j=0}^{m-1} 2^{-m+j} \mathcal{P}(m) + \mathcal{P}(m) \\ &= \sum_{j=0}^m 2^{-j} \mathcal{P}(m) = (2 - 2^{-m}) \mathcal{P}(m) \geq (2 - 2^{-m}) \mathcal{P}(\infty). \end{aligned}$$

We now need a lower bound on $\mathcal{P}(\infty)$, which we can obtain via Euler's Pentagonal Number Theorem [15], which states

that

$$\begin{aligned} \mathcal{P}(\infty) &= \prod_{k=1}^{\infty} (1 - q^k) \\ &= 1 + \sum_{k=1}^{\infty} (-1)^k (q^{(3k-1)k/2} + q^{(3k+1)k/2}) \end{aligned}$$

from which one can easily derive the inequality

$$\mathcal{P}(\infty) \geq 1 - q - q^2.$$

Combining these inequalities we get

$$\beta_m \geq (2 - 2^{-m})(1 - 2^{-x} - 2^{-2x}) \times 2^m \times 2^{-n/2}. \quad (\text{A3})$$

Using the same reasoning as in the proof of Theorem 3, the inequality (A3) allows us to bound the number of iterations of amplitude amplification by

$$\frac{\pi}{4} \frac{1}{2 - 2^{-m}} \frac{1}{1 - q - q^2} \times 2^{-m} \times 2^{n/2}.$$

Each D_m has exactly $2^m - 1$ oracle calls, so one iteration has $2^{m+1} - 1$ oracle calls (tree, its conjugate, and 1 extra call). Thus, the number of oracle calls is at most

$$\begin{aligned} &\frac{\pi}{4} \frac{1}{2 - 2^{-m}} \frac{1}{1 - q - q^2} 2^{-m} \times 2^{n/2} (2^{m+1} - 1) \\ &= \frac{\pi}{4} \frac{1}{1 - q - q^2} 2^{n/2} \end{aligned}$$

so we are only a factor of $\frac{1}{1 - 2^{-x} - 2^{-2x}}$ away from optimal number of oracle calls.

D_m can be implemented with $O(\sum_{k=1}^m ky \times 2^{m-k})$ gates. So we get at most

$$O\left[\left(\sum_{k=1}^m kx \times 2^{m-k}\right)2^{-m} \times 2^{n/2}\right] = O\left[\left(\sum_{k=1}^m kx2^{-k}\right)2^{n/2}\right]$$

nonoracle gates used by our algorithm. We use the following simple observation

$$\sum_{k=1}^m kx2^{-k} \leq \sum_{k=1}^{\infty} kx2^{-k} = 2x$$

to get that the total number of nonoracle gates used by our algorithm is bounded by $O(x2^{n/2})$. Thus, for any $\varepsilon > 0$ that is sufficiently small, we obtain an algorithm that makes at most

$$(1 + \varepsilon)\frac{\pi}{4}2^{n/2}$$

oracle calls, and uses at most

$$O[\log(\varepsilon^{-1})2^{n/2}]$$

nonoracle gates by setting $x \in \Theta[\log(\varepsilon^{-1})]$. ■

[1] L. K. Grover, A fast quantum mechanical algorithm for database search, in *Proceedings of the Twenty-eighth Annual ACM symposium on Theory of Computing* (Association for Computing Machinery, New York, 1996), pp. 212–219.

[2] G. Brassard, P. Høyer, and A. Tapp, Quantum cryptanalysis of hash and claw-free functions, *Lect. Notes Comput. Sci.* **1380**, 163 (1998).

[3] C. Durr and P. Høyer, A quantum algorithm for finding the minimum, [arXiv:quant-ph/9607014](https://arxiv.org/abs/quant-ph/9607014).

[4] C. Dürr, M. Heiligman, P. Høyer, and M. Mhalla, Quantum query complexity of some graph problems, *SIAM J. Comput.* **35**, 1310 (2006).

[5] L. K. Grover, Trade-offs in the quantum search algorithm, *Phys. Rev. A* **66**, 052314 (2002).

[6] S. Arunachalam and R. de Wolf, Optimizing the number of gates in quantum search, *Quantum Inf. Comput.* **17** (2015).

[7] K. Zhang and V. E. Korepin, Depth optimization of quantum search algorithms beyond Grover’s algorithm, *Phys. Rev. A* **101**, 032346 (2020).

[8] C. Calabro, R. Impagliazzo, V. Kabanets, and R. Paturi, The complexity of unique k-sat: An isolation lemma for k-cnfs, *J. Comput. Syst. Sci.* **74**, 386 (2008).

[9] C. Zalka, Grover’s quantum searching algorithm is optimal, *Phys. Rev. A* **60**, 2746 (1999).

[10] G. Brassard, P. Høyer, M. Mosca, and A. Tapp, Quantum amplitude amplification and estimation, *Contemp. Math.* **305**, 53 (2002).

[11] L. G. Valiant and V. V. Vazirani, Np is as easy as detecting unique solutions, in *Proceedings of the Seventeenth Annual ACM Symposium on Theory of Computing*, STOC ’85 (Association for Computing Machinery, New York, 1985), pp. 458–463.

[12] W. Burkot, J. Tułowiecki, V. Hlembotskyi, and W. Jarnicki, Quantum circuit and methods for use therewith, March 2020, US patent application No. 62990122.

[13] A. Mandviwalla, K. Ohshiro, and B. Ji, Implementing grover’s algorithm on the ibm quantum computers, in *2018 IEEE International Conference on Big Data (Big Data)* (IEEE, Piscataway, NJ, 2018), pp. 2531–2537.

[14] M. A. Nielsen and I. Chuang, Quantum computation and quantum information, *Am. J. Phys.* **70**, 558 (2002).

[15] L. Euler, Evolutio producti infiniti $(1-x)(1-xx)(1-x^3)(1-x^4)(1-x^5)(1-x^6)$ etc. in seriem simplicem, *Acta Academiae Scientiarum Imperialis Petropolitanae* (1783), pp. 47–44.

[16] M. Boyer, G. Brassard, P. Høyer, and A. Tapp, Tight bounds on quantum searching, *Fortschr. Phys.* **46**, 493 (1998).

[17] S. Arora and B. Barak, *Computational Complexity: A Modern Approach*, 1st ed. (Cambridge University Press, New York, 2009).

[18] J. L. Carter and M. N. Wegman, Universal classes of hash functions, *J. Comput. Syst. Sci.* **18**, 143 (1979).

[19] Y. Mansour, N. Nisan, and P. Tiwari, The computational complexity of universal hashing, *Theor. Comput. Sci.* **107**, 121 (1993).

[20] M. T. Nair and A. Singh, *Linear Algebra* (Springer, Berlin, 2018), pp. 107–161.

[21] J. Goldman and G.-C. Rota, On the foundations of combinatorial theory iv finite vector spaces and eulerian generating functions, *Stud. Appl. Math.* **49**, 239 (1970).

[22] J. Gwinner, M. Briański, W. Burkot, Ł. Czerwiński, and V. Hlembotskyi, Benchmarking 16-element quantum search algorithms on ibm quantum processors, [arXiv:2007.06539](https://arxiv.org/abs/2007.06539).