View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Publications of the IAS Fellows

arXiv:quant-ph/0007021v4 16 Oct 2001

The quantum complexity of set membership[*]

Jaikumar Radhakrishnan[†] Pranab Sen[‡] S. Venkatesh[§]

Abstract

We study the *quantum complexity* of the static set membership problem: given a subset S ($|S| \leq n$) of a universe of size m ($\gg n$), store it as a table, $T : \{0,1\}^r \rightarrow \{0,1\}$, of bits so that queries of the form 'Is x in S?' can be answered. The goal is to use a small table and yet answer queries using few bit probes. This problem was considered recently by Buhrman, Miltersen, Radhakrishnan and Venkatesh [BMRV00], who showed lower and upper bounds for this problem in the classical deterministic and randomised models. In this paper, we formulate this problem in the "quantum bit probe model". We assume that access to the table T is provided by means of a black box (oracle) unitary transform O_T that takes the basis state $|y, b\rangle$ to the basis state $|y, b \oplus T(y)\rangle$. The query algorithm is allowed to apply O_T on any superposition of basis states.

We show tradeoff results between space (defined as 2^r) and number of probes (oracle calls) in this model. Our results show that the lower bounds shown in [BMRV00] for the classical model also hold (with minor differences) in the quantum bit probe model. These bounds almost match the classical upper bounds. Our lower bounds are proved using linear algebraic arguments.

Keywords: Data structures, set membership, bit probe model, quantum black box model, linear algebraic methods, lower bounds, space-time tradeoffs.

1 Introduction

In this paper we study the *static membership* problem: Given a subset S ($|S| \leq n$) of a universe of size m ($\gg n$), store it efficiently and succinctly so that queries of the form

[*] A preliminary version of this paper appeared in the *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, pages 554–562, 2000.

[†] School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai 400005, India. Email: jaikumar@tcs.tifr.res.in.

[‡] Laboratoire de Recherche en Informatique, Université de Paris-Sud, 91405 Orsay, France. Email: pranab@lri.fr. Most of this work was done while the author was a graduate student at the Tata Institute of Fundamental Research.

[§] Center for Discrete Mathematics and Theoretical Computer Science (DIMACS), Rutgers University, Piscataway, NJ 08854, USA. Email: venkat@dimacs.rutgers.edu. Most of this work was done while the author was a graduate student at the Tata Institute of Fundamental Research.

"Is x in S?" can be answered quickly. This fundamental data structure problem has been studied earlier in various settings (e.g. by Minsky and Papert [MP69], Yao [Yao81], Fredman, Komlós and Szemerédi [FKS84], Pagh [Pag01]). Most of these results were in the classical deterministic *cell probe model*. Recently, this problem was considered by Buhrman, Miltersen, Radhakrishnan and Venkatesh [BMRV00] in the classical *bit probe model*, which was introduced in [MP69]; they studied tradeoffs between storage space and number of probes in the classical deterministic case, and also showed lower and upper bounds for the storage space when the query algorithm was randomised and made just *one* bit probe. In the classical bit probe model the storage scheme is deterministic and stores the given set as a string of bits. The query scheme is either deterministic or randomised and answers membership queries probing only one bit of the string at a time.

In this paper, we allow the query algorithm to perform quantum search on the table of bits representing the stored subset S. The storage scheme which encodes S a table of bits, $T : \{0,1\}^r \rightarrow \{0,1\}$, continues to be classical deterministic. To formalise this, we define the *quantum bit probe model*. The table of 2^r bits, T, is modelled using an oracle unitary transformation O_T that takes the basis state $|y, b\rangle$ to $|y, b \oplus T(y)\rangle$. Basically, the table of bits is accessed by the query algorithm as in the well-studied *quantum black box model* (see e.g. [BBC$^+$98]). If the inputs to the oracle O_T are restricted to basis states, O_T reduces to (the reversible version of) the classical table that stores one bit for each address in $\{0,1\}^r$. We, therefore, define the *space* used by the quantum bit probe scheme to be 2^r.

The main point of departure from the classical model, is in the query algorithm. We allow the algorithm to feed a superposition of basis states to the oracle. Each use of the oracle counts as one probe of the table, even if a superposition is supplied to the oracle, and the output depends on the value of several bits of the underlying table T. In the preparation of this superposition and in the processing of the output returned by the oracle, we allow arbitrary unitary transformations. It is known that this form of access often leads to significant improvements over classical algorithms for several problems (e.g. Grover's algorithm [Gro96] for searching an unordered database).

Previously, the number of probes to the black box as a complexity measure had been studied in the quantum setting (e.g. [BBBV97, BBC$^+$98, Amb00, Gro96]). Both lower bounds and upper bounds for various problems were proved. The main contribution of this paper is the study of tradeoffs between storage space and number of probes for a static data structure problem in the quantum setting. For the set membership problem, we show that several limitations of classical computation (shown in [BMRV00]) continue to persist even if quantum query algorithms are allowed. This is surprising, because for the (superficially) similar problem of searching an unordered database, quantum computation helps.

Our tradeoffs between storage space and the number of quantum probes are proved using linear algebraic arguments. Roughly speaking, we lower and upper bound the dimension of a set of unitary operators arising from the quantum query algorithm. The lower bound on the dimension arises from the 'correctness requirements' of the quantum algorithm. The upper bound on the dimension arises from limitations on the storage space and number of probes. By playing the lower and upper bounds against each other, we get the desired tradeoffs. To the best of our knowledge, this is the first time that linear algebraic arguments have been used to prove lower bounds for data structure problems, classical or quantum. Counting of dimensions has been previously used in quantum computing (see e.g. [AST$^+$98, BdW01]), but in quite different contexts and ways. Linear algebraic arguments similar to ours have been heavily used in combinatorics. For a delightful introduction, see the book by Babai and Frankl [BF92].

1.1 Our results

The exact quantum model: Buhrman *et al.* [BMRV00] have shown, for classical deterministic query algorithms, that any (s,t)-scheme (which uses space s and t bit probes) satisfies $\binom{m}{n} \leq \binom{s}{nt}2^{nt}$. We show a stronger (!) tradeoff result in the quantum bit probe model.

Result 1 *Suppose there exists a scheme for storing subsets S of size at most n from a universe $\mathbf{U}$ of size m that uses s bits of storage and answers membership queries, with zero error probability, with t quantum probes. Then,*

$$\sum_{i=0}^{n} \binom{m}{i} \leq \sum_{i=0}^{nt} \binom{s}{i}$$

This has two immediate consequences. First, by setting $t = 1$, we see that if only one probe is allowed, then m bits of storage are necessary. (In [BMRV00], for the classical model, this was justified using an ad hoc argument.) Thus, the classical deterministic *bit vector* scheme that stores the characteristic vector of the set S and answers membership queries using one bit probe, is optimal even up to quantum. Second, it follows (see [BMRV00] for details) that the classical deterministic scheme of Fredman, Komlós and Szemerédi [FKS84], which uses $O(n \log m)$ bits of storage and answers membership queries using $O(\log m)$ bit probes, is optimal even up to quantum — quantum schemes that use $O(n \log m)$ bits of storage must make $\Omega(\log m)$ probes if $n \leq m^{1-\Omega(1)}$. Recently, Pagh [Pag01] has shown classical deterministic schemes using the information-theoretic minimum space $O(n \log(m/n))$ and making $O(\log(m/n))$ bit probes, which is optimal even up to quantum, by the above result. For t between 1 and $O(\log(m/n))$, Buhrman *et al.* [BMRV00] have given classical deterministic schemes making t bit probes, which $O(m^{3/t} n \log m)$ bits of storage. A lower bound of $\Omega(nt(m/n)^{1/t})$ for storage space, for suitable values of the various parameters, follows from Result 1. Thus, if we only care about space up to a polynomial, classical deterministic schemes that make t bit probes for t between 1 and $O(\log(m/n))$, and which use storage space almost matching the exact quantum lower bounds, exist.

Interestingly, the above theorem holds even in the presence of errors, provided the error is restricted to positive instances, that is the query algorithm sometimes (with probability < 1) returns the answer 'No' for a query x that is actually in the set S. This was not observed earlier even in the classical model, although one can easily modify the proof of the tradeoff result in [BMRV00] to give this.

The ϵ-error model: In the classical model, there exists a scheme for storing subsets of size at most n from a universe of size m that answers membership queries, with two-sided error at most $\epsilon < 1/16$, using just *one* bit probe, and using storage space $O(\frac{n \log m}{\epsilon^2})$. Also, any such one probe scheme making two-sided error at most ϵ must use space $\Omega(\frac{n \log m}{\epsilon \log(1/\epsilon)})$. Both the upper bound and the lower bound have been proved in [BMRV00]. By two-sided error, we mean that the query algorithm can make an error for both positive instances (the query element is a member of the stored set), as well as negative instances (the query element is not a member of the stored set). Since different sets must be represented by different tables, every scheme, no matter how many probes the query algorithm is allowed, must use $\Omega(n \log(m/n))$ bits of storage, even in the bounded two-sided error quantum model. However, one might ask if the dependence of space on ϵ is significantly better in the quantum probe model. We show the following lower bound which implies that a quantum scheme needs significantly more than the information-theoretic optimal space if sub-constant error probabilities are desired.

> **Result 2** *Let $n/m < \epsilon < 1/8$. Suppose there is a scheme with two-sided error ϵ which stores subsets of size at most n from a universe of size m and answers membership queries, with two-sided error at most ϵ, using one quantum probe. It must use space*
>
> $$s = \Omega\left(\frac{n \log(m/n)}{\epsilon^{1/6} \log(1/\epsilon)}\right)$$

The method used to prove this result can be generalised to algorithms that use more probes than one.

> **Result 2'** *For any $p \geq 1$ and $n/m < \epsilon < 2^{-3p}$, suppose there is a scheme which stores subsets of size at most n from a universe of size m and answers membership queries, with two-sided error at most ϵ, using p quantum probes. Define $\delta \stackrel{\Delta}{=} \epsilon^{1/p}$. It must use space*
>
> $$s = \Omega\left(\frac{n \log(m/n)}{\delta^{1/6} \log(1/\delta)}\right)$$

Such a tradeoff between space and error probability for multiple probes was not known earlier, even in the classical randomised model. We note that for p bit probes, an upper bound of $O(\frac{n \log m}{\epsilon^{4/p}})$ on the storage space, for $\epsilon < 2^{-p}$, follows by taking the storage scheme of [BMRV00] for error probability $\frac{\epsilon^{2/p}}{4}$, and repeating the (classical randomised) single probe query scheme p times. This diminishes the probability of error to ϵ. Thus, our lower bounds for two-sided error quantum schemes roughly match the two-sided error classical randomised upper bounds.

The results described above are inspired by similar results proved earlier in [BMRV00] in the classical model. However, the methods used for classical models, which were based on combinatorial arguments involving set systems, seem to be powerless in giving the results in the quantum model. Our results are based on linear algebraic arguments, involving counting the dimensions of spaces of various operators that arise in the quantum query algorithm.

Bounds for classical models: As stated above, Result 1 is stronger than what was known earlier, even in the classical deterministic model. One might wonder if this stronger result is somehow easier to prove in the classical deterministic model. We show that the linear algebraic techniques used in the proof of Result 1 can be considerably simplified when we assume the classical deterministic model, and give the same inequality as stated in Result 1. Also, one can easily modify the proof to yield the same tradeoff for randomised query schemes where the error is restricted to positive instances (i.e. when the query element is a member of the stored set).

The proof in [BMRV00] of the space lower bound for a classical two-sided ϵ-error randomised query scheme, namely $s = \Omega(\frac{n \log m}{\epsilon \log(1/\epsilon)})$, involved some tricky use of both upper and lower bounds for r-cover-free families shown by Nisan and Wigderson [NW94], Erdős, Frankl and Füredi [EFF85], Dyachkov and Rykov [DR82]. The proof of the analogous bound (Result 2) in the quantum model completely avoids these. In fact, we give a proof of a (slightly weaker) lower bound $s = \Omega(\frac{n \log m}{\epsilon^{2/5} \log(1/\epsilon)})$ in the classical randomised model by adapting the ideas used in the proof of Result 2 to the classical setting. We first diminish the error probability of the one-probe query algorithm by repetition and then we can, by fixing the random coin tosses, make it a deterministic query algorithm which however, uses more than one probe. We then apply our (classical) deterministic space-time tradeoff equation to complete the proof. This approach is inspired by our proof of Result 2. Besides being simpler, this proof has the advantage that it generalises readily to more than one probe. No such result was known earlier in the classical setting.

Result 3 *Let $p \geq 1$, $18^{-p} > \epsilon > 1/m^{1/3}$ and $m^{1/3} > 18n$. Define $\delta \stackrel{\Delta}{=} \epsilon^{1/p}$. Any classical scheme which stores subsets of size n from a universe of size m and answers membership queries, with two-sided error at most ϵ, using at most p bit probes must use space*

$$\Omega\left(\frac{n \log m}{\delta^{2/5} \log(1/\delta)}\right)$$

1.2 Organisation of the paper

In the next section, we describe our quantum bit probe model formally and give the framework of our proofs. Detailed proofs of all our results (for both quantum and classical models) appear in Section 3. We conclude in Section 4 with some open problems.

2 Definitions and notations

In this section we first describe our *quantum bit probe model* and then give some formal definitions and notations which will be used in the proofs of the theorems.

2.1 The model

Our model is a quantum analogue of the classical bit probe model which has been extensively used in the past to study data structure problems (see e.g. [MP69, Mil93, BMRV00, Pag01]).

A static data structure problem is a function $f : D \times Q \to \{0,1\}$, where D is a finite set called the set of data to be stored and Q is a finite set called the set of queries. A classical (s,t)-bit probe scheme for a static data structure problem consists of a deterministic storage scheme which stores the given data d as a bit string of length s, and a query scheme which given a query q makes at most t bit probes to the stored string and computes $f(d,q)$. The query scheme can be either deterministic or randomised. For more details about the classical model, see [BMRV00].

A quantum (s,t)-bit probe scheme for a static data structure problem has two components: a classical deterministic storage scheme that stores the data d using s bits, and a quantum query scheme that answers queries by 'quantumly probing a bit at a time' at most t times.

The storage scheme: For the set membership problem, the data to be stored is a subset S of a universe $\mathbf{U}$ ($|S| \leq n$, $|\mathbf{U}| = m$). Let $x(S) \in \{0,1\}^s$ be the bit string that is stored by the storage scheme for recording S. The storage scheme is classical deterministic. The difference now, is that this bit string is made available to the query algorithm in the form of an oracle unitary transform O_S. To define O_S formally, we represent the basis states of the quantum query circuit as $|j,b,z\rangle$, where $j \in [s]$ is a binary string of length $\log s$ ('address qubits'), b is a single bit ('data qubit'), and z is a binary string of some fixed length ('work qubits'). Let $x(S)_j$ be the bit stored at the jth location in the string $x(S)$. The action of O_S on a basis state is described below.

$$O_S : |j,b,z\rangle \mapsto (-1)^{b \cdot x(S)_j} |j,b,z\rangle$$

Remark: The oracle described in the introduction mapped $|j,b,z\rangle$ to $|j, b \oplus x(S)_j, z\rangle$. It is known that the oracle O_S defined above is equivalent in power to this oracle.

Thus, information about the string $x(S)$ appears in the phase of the basis states in the output, and O_S is represented by a diagonal matrix (in the standard computational basis): the ith diagonal entry, where $i \equiv |j,b,z\rangle$, is

$$(O_S)_{i,i} = (-1)^{b \cdot x(S)_j}$$

For $T \subseteq [s]$ and $x \in \{0,1\}^s$, define $[x]_T \triangleq \sum_{i \in T} x_i \pmod 2$. In particular, $[x]_\emptyset = 0$. Thus, $(O_S)_{i,i} = (-1)^{[x(S)]_{l_i}}$, where l_i is some subset of $[s]$ of size 1 (when $b=1$, $l_i = \{j\}$) or 0 (when $b=0$, $l_i = \emptyset$).

Remark: Our model for storage does not permit O_S to be any arbitrary unitary transformation. However, this restricted form of the oracle is closer to the way bits are accessed in the classical case. Moreover, in most previous works, storage has been modelled using such an oracle (see e.g. [Gro96, BBBV97, BBC+98, Amb00]).

The query scheme: Suppose a subset $S \subseteq \mathbf{U}, |S| \leq n$, has been stored and $x(S) \in \{0,1\}^s$ is the corresponding bit string. A quantum query scheme with t probes is just a sequence of unitary transformations

$$U_0 \rightarrow O_S \rightarrow U_1 \rightarrow O_S \rightarrow \dots U_{t-1} \rightarrow O_S \rightarrow U_t$$

where U_j's are arbitrary unitary transformations that do not depend on the set S stored. For a query $q \in \mathbf{U}$, the computation starts in an observational basis state $|q\rangle|0\rangle$, where we assume that the ancilla qubits are initially in the basis state $|0\rangle$. Then we apply the operators $U_0, O_S, \dots, O_S, U_t$ and measure the final state. The result of the query is the rightmost bit of the state obtained by the measurement. The query scheme can be exact or have error; the error can be one-sided or two-sided. When the query scheme is exact, the measurement of the final state gives the correct answer with probability 1. If one-sided error ϵ is allowed, the measurement produces a 0 with probability 1 when the answer is 0, but when the answer is 1, is required to produce a 1 with probability only at least $1-\epsilon$. If two-sided error ϵ is allowed, the answer can be wrong, with probability at most ϵ, for both positive and negative instances.

The framework for the proofs: We now describe the general framework in which the various proofs are presented and also give some definitions and notations which will be used throughout the paper.

For a query $q \in \mathbf{U}$, define $|\phi_q\rangle \stackrel{\Delta}{=} |q\rangle|0\rangle$. The set of vectors $|\phi_q\rangle, q \in \mathbf{U}$ form an orthogonal system of vectors. They are independent of the set S stored.

Define two Hilbert spaces, A_0 and A_1, where A_i is the space of all state vectors that can be spanned by basis states having an i at the rightmost bit (i.e. if the state vector lies in A_i, then on measuring the rightmost bit at the output, one gets i with probability 1). Then the entire state space V decomposes as an orthogonal direct sum of the spaces A_0, A_1.

Define the unitary transformations $\{W_S\}_{S \subseteq \mathbf{U}, |S| \leq n}$ as follows.

$$W_S \stackrel{\Delta}{=} U_t O_S U_{t-1} O_S U_{t-2} O_S \cdots U_2 O_S U_1 O_S U_0$$

Thus when a set S is stored, in the exact quantum case, $W_S|\phi_i\rangle, i \in S$ lie in A_1, and $W_S|\phi_i\rangle, i \notin S$ lie in A_0. In the one-sided ϵ-error case, $W_S|\phi_i\rangle, i \notin S$ lie in A_0, but $W_S|\phi_i\rangle, i \in S$ may not lie entirely in A_1, but in fact may have a projection on A_0 of length at most $\sqrt{\epsilon}$. In the two-sided ϵ-error case, W_S has to send the vector $|\phi_i\rangle$ "approximately" to the correct space, i.e. the projection of $W_S|\phi_i\rangle$ on the correct space is of length more than $\sqrt{1-\epsilon}$.

Notation: In the proofs we have to take tensor products of vectors and matrices. For any vector v or matrix M, we have the following notation,

$$v^{\otimes t} \stackrel{\Delta}{=} \underbrace{v \otimes \cdots \otimes v}_{t \text{ times}}$$

$$M^{\otimes t} \stackrel{\Delta}{=} \underbrace{M \otimes \cdots \otimes M}_{t \text{ times}}$$

We note that since the entire state space V is the orthogonal direct sum of A_0 and A_1,

$$V^{\otimes t} = A_0^{\otimes t} \oplus (A_0^{\otimes t-1} \otimes A_1) \oplus \cdots \oplus A_1^{\otimes t}$$

and the 2^t vector spaces in the above direct sum are pairwise orthogonal.

Below, $A \triangle B$ stands for the symmetric difference of sets A and B; $M^\dagger$ stands for the conjugate transpose of the matrix M.

3 Proofs of theorems

3.1 Quantum schemes

We first prove our space v/s probes tradeoff result for exact quantum schemes.

Theorem 1 *Suppose there exists a scheme for storing subsets S of size at most n from a universe $\mathbf{U}$ of size m that uses s bits of storage and answers membership queries, with zero error probability, with t quantum probes. Then,*

$$\sum_{i=0}^{n} \binom{m}{i} \leq \sum_{i=0}^{nt} \binom{s}{i}$$

Proof: We use the notation of Section 2. For any subset $S \subseteq \mathbf{U}$, $|S| \leq n$, let us define

$$W_S \stackrel{\Delta}{=} U_t O_S U_{t-1} O_S U_{t-2} O_S \cdots U_2 O_S U_1 O_S U_0$$

Claim 1 $\{W_S^{\otimes n}\}_{S \in \binom{\mathbf{U}}{\leq n}}$ *are linearly independent.*

Proof: Suppose there is a nontrivial linear combination

$$\sum_{S \in \binom{\mathbf{U}}{\leq n}} \alpha_S W_S^{\otimes n} = 0$$

Let T be a set of largest cardinality such that $\alpha_T \neq 0$ and let $T = \{i_1, \ldots, i_k\}$, $k \leq n$. We define a vector

$$|\phi_T\rangle \stackrel{\Delta}{=} |\phi_{i_1}\rangle^{\otimes(n-k+1)} \otimes |\phi_{i_2}\rangle \otimes \cdots \otimes |\phi_{i_k}\rangle$$

Applying $|\phi_T\rangle$ to the linear combination above, we have

$$\sum_{S \in \binom{\mathbf{U}}{\leq k}, S \neq T} \alpha_S W_S^{\otimes n} |\phi_T\rangle + \alpha_T W_T^{\otimes n} |\phi_T\rangle = 0 \qquad (1)$$

For any set S,

$$W_S^{\otimes n} |\phi_T\rangle = (W_S |\phi_{i_1}\rangle)^{\otimes(n-k+1)} \otimes W_S |\phi_{i_2}\rangle \otimes \cdots \otimes W_S |\phi_{i_k}\rangle$$

- If $S = T$, $W_T|\phi_{i_l}\rangle \in A_1$ for all l, $1 \le l \le k$ Hence $W_T^{\otimes n}|\phi_T\rangle \in A_1^{\otimes n}$.
- If $S \neq T$, there exists an element i_j in $T - S$ (by choice of T). $W_S|\phi_{i_j}\rangle \in A_0$. Hence $W_S^{\otimes n}|\phi_T\rangle \notin A_1^{\otimes n}$. In fact, $W_S^{\otimes n}|\phi_T\rangle$ is orthogonal to $A_1^{\otimes n}$.

Hence, in the above linear combination (equation 1), the only vector which has a nontrivial projection along $A_1^{\otimes n}$ is $W_T^{\otimes n}|\phi_T\rangle$. Hence, $\alpha_T = 0$ leading to a contradiction. ∎

Claim 2 $\{W_S^{\otimes n}\}_{S \in \binom{\mathrm{U}}{\le n}}$ *lie in a vector space of dimension at most* $\sum_{i=0}^{nt} \binom{s}{i}$.

Proof: By definition, for any set S, $|S| \le n$,

$$W_S \triangleq U_t O_S U_{t-1} O_S U_{t-2} O_S \cdots U_2 O_S U_1 O_S U_0$$

where $U_0, \ldots, U_t$ are unitary transformations (matrices) independent of the set stored.

For any pair of indices (i, j),

$$\begin{aligned}(W_S)_{i,j} &= \sum_{k_{t-1},\ldots,k_0} (U_t)_{i,k_{t-1}} (O_S)_{k_{t-1},k_{t-1}} (U_{t-1})_{k_{t-1},k_{t-2}} (O_S)_{k_{t-2},k_{t-2}} \\ &\qquad \cdots (U_1)_{k_1,k_0} (O_S)_{k_0,k_0} (U_0)_{k_0,j} \\ &= \sum_{k_{t-1},\ldots,k_0} (U_t)_{i,k_{t-1}} (-1)^{[x(S)]_{l_{k_{t-1}}}} (U_{t-1})_{k_{t-1},k_{t-2}} (-1)^{[x(S)]_{l_{k_{t-2}}}} \\ &\qquad \cdots (U_1)_{k_1,k_0} (-1)^{[x(S)]_{l_{k_0}}} (U_0)_{k_0,j}\end{aligned}$$

where, recalling the notation of Section 2, $x(S)$ is the string stored by the storage scheme for set S and l_i is either the single location in the string corresponding to index i or the empty set.

Therefore, if we define $T_{k_{t-1},\ldots,k_0} \triangleq l_{k_{t-1}} \triangle l_{k_{t-2}} \triangle \cdots \triangle l_{k_0}$ and $[x(S)]_T$ to be the parity of the bits stored in $x(S)$ at the locations of T, we have

$$\begin{aligned}(W_S)_{i,j} &= \sum_{k_{t-1},\ldots,k_0} (-1)^{[x(S)]_{T_{k_{t-1},\ldots,k_0}}} (U_t)_{i,k_{t-1}} (U_{t-1})_{k_{t-1},k_{t-2}} \cdots (U_1)_{k_1,k_0} (U_0)_{k_0,j} \\ &= \sum_{T \in \binom{[s]}{\le t}} (-1)^{[x(S)]_T} \sum_{\substack{k_{t-1},\ldots,k_0 \\ T_{k_{t-1},\ldots,k_0} = T}} (U_t)_{i,k_{t-1}} (U_{t-1})_{k_{t-1},k_{t-2}} \cdots (U_1)_{k_1,k_0} (U_0)_{k_0,j}\end{aligned}$$

Let us define for every set $T \subseteq [s]$, $|T| \le t$, a matrix A_T as follows:

$$(A_T)_{i,j} \triangleq \sum_{\substack{k_{t-1},\ldots,k_0 \\ T_{k_{t-1},\ldots,k_0} = T}} (U_t)_{i,k_{t-1}} (U_{t-1})_{k_{t-1},k_{t-2}} \cdots (U_1)_{k_1,k_0} (U_0)_{k_0,j}$$

Then we have,

$$W_S = \sum_{T \in \binom{[s]}{\le t}} (-1)^{[x(S)]_T} A_T \tag{2}$$

Hence,

$$\begin{aligned} (W_S)^{\otimes n} &= \left(\sum_{T_1\in\binom{[s]}{\le t}} (-1)^{[x(S)]_{T_1}} A_{T_1}\right)\otimes\cdots\otimes\left(\sum_{T_n\in\binom{[s]}{\le t}} (-1)^{[x(S)]_{T_n}} A_{T_n}\right)\\ &= \sum_{\substack{T_i\in\binom{[s]}{\le t}\\ 1\le i\le n}} (-1)^{[x(S)]_{T_1}}\cdots(-1)^{[x(S)]_{T_n}}(A_{T_1}\otimes\cdots\otimes A_{T_n})\\ &= \sum_{\tilde{T}\in\binom{[s]}{\le nt}} (-1)^{[x(S)]_{\tilde{T}}} B_{\tilde{T}} \end{aligned}$$

where for $\tilde{T}\in\binom{[s]}{\le nt}$,

$$B_{\tilde{T}} \triangleq \sum_{\substack{T_1\triangle\cdots\triangle T_n=\tilde{T}\\ T_i\in\binom{[s]}{\le t},1\le i\le n}} A_{T_1}\otimes\cdots\otimes A_{T_n}$$

Hence, we see that $\{B_{\tilde{T}}\}_{\tilde{T}\in\binom{[s]}{\le nt}}$ span $\{W_S^{\otimes n}\}_{S\in\binom{\mathbf{U}}{\le n}}$. So, $\{W_S^{\otimes n}\}_{S\in\binom{\mathbf{U}}{\le n}}$ lie in a vector space of dimension at most $\sum_{i=0}^{nt}\binom{s}{i}$. $\blacksquare$

Now the theorem is an easy consequence of the above two claims. $\blacksquare$

Remark: Equation 2 in the proof of Claim 2 above is similar to the statement of a lemma of Shi.

Lemma 1 ([Shi00, Lemma 2.4] rephrased) *Consider a quantum query algorithm with initial state vector $|0\rangle$, with the black box unitary transformation representing a bit string $x = x_1, \ldots, x_s$. Let $|\phi\rangle$ be the state vector of the circuit after t queries to the black box. Then*

$$|\phi\rangle = \sum_{T\in\binom{[s]}{\le t}} \hat{\phi}_T(-1)^{[x]_T}$$

where the $\hat{\phi}_T$ are independent of x.

Shi proved his lemma using the observation by Beals *et al.* (see [BBC$^+$98, Lemma 4.1]) that the amplitudes of the basis states in the state vector $|\phi\rangle$ are multilinear polynomials of degree at most t in $x_1, \ldots, x_s$.

The space-time tradeoff equation for the exact quantum case holds for the one-sided error case too, as shown below.

Theorem 2 *The tradeoff result of Theorem 1 also holds for a quantum scheme where the query scheme may err with probability less than 1 on the positive instances (i.e. if an element is present it may be erroneously reported absent), but not on the negative instances (i.e. if an element is absent it has to be reported absent).*

Proof: **(Sketch)** Essentially, the same proof of Theorem 1 goes through. Since the query scheme can make an error only if the element is present, we observe that the only vector in the linear combination (equation 1) that has a non-zero projection on the space $A_1^{\otimes n}$, is the vector $W_T^{\otimes n}|\phi_T\rangle$. Hence $\alpha_T = 0$, and the operators $\{W_S\}_{S \subseteq \mathbf{U}, |S| \leq n}$ continue to be linearly independent. Hence, the same tradeoff equation holds in this case too. ■

We now prove the lower bound on the space used by a two-sided ϵ-error 1-probe quantum scheme.

Theorem 3 *Let $n/m < \epsilon < 1/8$. Suppose there is a scheme which stores subsets S of size at most n from a universe $\mathbf{U}$ of size m that answers membership queries, with two-sided error at most ϵ, using one quantum probe. It must use space*

$$s = \Omega\left(\frac{n \log(m/n)}{\epsilon^{1/6} \log(1/\epsilon)}\right)$$

Proof: Since we are looking at a one probe quantum scheme, $W_S = U_1 O_S U_0$. We start by picking a family F of sets, $F = \{S_1, \ldots, S_k\}$, $S_i \subseteq \mathbf{U}$, $|S_i| = n$ and $|S_i \cap S_j| \leq n/2$ for all $i \neq j$. By picking the sets greedily [EFF85, NW94], one obtains a family F with

$$|F| \geq \frac{\binom{m}{n}}{\binom{n}{\frac{n}{2}}\binom{m-\frac{n}{2}}{\frac{n}{2}}} \geq \frac{\frac{m}{n}\frac{m-1}{n-1}\cdots\frac{m-n/2+1}{n-n/2+1}}{2^n} \geq \frac{\left(\frac{m}{n}\right)^{n/2}}{2^n} = \left(\frac{m}{4n}\right)^{n/2} \qquad (3)$$

Let $t \triangleq \left\lceil \frac{4 \log |F|}{n \log(1/(4\epsilon))} \right\rceil$. Since, $m/n \geq 1/\epsilon$,

$$\frac{4 \log |F|}{n \log(1/(4\epsilon))} \geq \frac{4n \log(m/(4n))}{2n \log(1/(4\epsilon))} \geq 2$$

Hence,

$$\frac{4 \log |F|}{n \log(1/(4\epsilon))} \leq t \leq \frac{4 \log |F|}{n \log(1/(4\epsilon))} + 1 \leq \frac{6 \log |F|}{n \log(1/(4\epsilon))} \qquad (4)$$

Claim 3 *$\{W_S^{\otimes nt}\}_{S \in F}$ are linearly independent.*

Proof: Suppose there is a non-trivial linear combination

$$\sum_{S \in F} \alpha_S W_S^{\otimes nt} = 0$$

Fix a $T \in F$. Let $T = \{i_1, \ldots, i_n\}$. Define

$$|\phi_T\rangle \triangleq (|\phi_{i_1}\rangle \otimes |\phi_{i_2}\rangle \otimes \cdots \otimes |\phi_{i_n}\rangle)^{\otimes t}$$

Applying ϕ_T to the above linear combination, we get

$$\sum_{S \in F} \alpha_S W_S^{\otimes nt} |\phi_T\rangle = 0$$

$$\Rightarrow \sum_{S\in F} \alpha_S (W_S|\phi_{i_1}\rangle \otimes \cdots \otimes W_S|\phi_{i_n}\rangle)^{\otimes t} = 0$$

Taking inner product of the above combination with the vector

$$W_T^{\otimes nt}|\phi_T\rangle = (W_T|\phi_{i_1}\rangle \otimes \cdots \otimes W_T|\phi_{i_n}\rangle)^{\otimes t}$$

we get

$$\sum_{S\in F} \alpha_S \langle (W_S|\phi_{i_1}\rangle \otimes \cdots \otimes W_S|\phi_{i_n}\rangle)^{\otimes t} | (W_T|\phi_{i_1}\rangle \otimes \cdots \otimes W_T|\phi_{i_n}\rangle)^{\otimes t}\rangle = 0$$

$$\Rightarrow \sum_{S\in F} \alpha_S (\langle \phi_{i_1}|W_S^\dagger W_T|\phi_{i_1}\rangle \cdots \langle \phi_{i_n}|W_S^\dagger W_T|\phi_{i_n}\rangle)^t = 0 \qquad (5)$$

- For any $i_j \in S \cap T$, $|\langle \phi_{i_j}|W_S^\dagger W_T|\phi_{i_j}\rangle| \leq 1$.
- For any $i_j \in T$, $W_T|\phi_{i_j}\rangle = v_0 + v_1$ where $1 \geq \|v_1\| \geq \sqrt{1-\epsilon}$ and $\|v_0\| \leq \sqrt{\epsilon}$, $v_0 \in A_0$ and $v_1 \in A_1$. For any $i_j \in T - S$, $W_S|\phi_{i_j}\rangle = u_0 + u_1$ where $1 \geq \|u_0\| \geq \sqrt{1-\epsilon}$ and $\|u_1\| \leq \sqrt{\epsilon}$ and $u_0 \in A_0$ and $u_1 \in A_1$. Hence

$$\begin{aligned} |\langle \phi_{i_j}|W_S^\dagger W_T|\phi_{i_j}\rangle| &= |\langle u_0|v_0\rangle + \langle u_1|v_1\rangle| \\ &\leq \|u_0\|\|v_0\| + \|u_1\|\|v_1\| \\ &\leq 2\sqrt{\epsilon} \overset{\Delta}{=} \delta \end{aligned}$$

We now note that for every $T \in F$, we have a linear combination as in equation 5 above. We can write the linear combinations in the matrix form as $\alpha M = 0$, where $\alpha = (\alpha_S)_{S\in F}$ and M is a $|F| \times |F|$ matrix whose rows and columns are indexed by members of F. For $S, T \in F$,

$$M(S,T) = (\langle \phi_{i_1}|W_S^\dagger W_T|\phi_{i_1}\rangle \cdots \langle \phi_{i_n}|W_S^\dagger W_T|\phi_{i_n}\rangle)^t$$

where $T = \{i_1, \ldots, i_n\}$. The diagonal entries of M, $M(T,T)$, are 1. The non-diagonal entries satisfy $|M(S,T)| \leq (\delta)^{(n-|S\cap T|)t} \leq \delta^{nt/2}$.

Using the lower bound on t from (4), we get

$$|F|\delta^{tn/2} = |F|(4\epsilon)^{tn/4} \leq 1$$

Hence $(|F|-1)(\delta)^{tn/2} < 1$. This implies that M is non-singular. [Suppose not. Let y be a vector such that $My = 0$. Let i be the location of the largest coordinate of y. We can assume without loss of generality that $y_i = 1$. Now, the ith coordinate of the vector My is at least $1 - (|F|-1)(\delta)^{tn/2} > 0$ in absolute value, which is a contradiction.] So, $\alpha_S = 0$ for all $S \in F$. Hence $\{W_S^{\otimes nt}\}_{S\in F}$ are linearly independent. ■

Claim 4 *$\left\{W_S^{\otimes nt}\right\}_{S\in F}$ lie in a vector space of dimension at most $\sum_{j=0}^{nt}\binom{s}{j}$.*

Proof: Similar to proof of Claim 2 in Theorem 1. ■

Using the two claims above,

$$|F| \leq \sum_{j=0}^{nt} \binom{s}{j} \leq \binom{s+nt}{nt} \leq \left(\frac{(s+nt)e}{nt}\right)^{nt}$$

Using the upper bound on t from (4), we get

$$\left(\frac{1}{4\epsilon}\right)^{nt/6} \leq |F| \leq \left(\frac{(s+nt)e}{nt}\right)^{nt}$$

$$\Rightarrow s \geq \frac{nt}{e}\left(\left(\frac{1}{4\epsilon}\right)^{1/6} - e\right)$$

For values of ϵ such that $(1/4\epsilon)^{1/6} > 2e$, that is $\epsilon < 4^{-1}(2e)^{-6}$, using (3) and the lower bound on t from (4), we get

$$s \geq \frac{nt}{2e}\left(\frac{1}{4\epsilon}\right)^{1/6} \geq \frac{2\log|F|}{e(4\epsilon)^{1/6}\log(1/4\epsilon)} \geq \frac{n\log(m/4n)}{e(4\epsilon)^{1/6}\log(1/4\epsilon)}$$

$$\Rightarrow s = \Omega\left(\frac{n\log(m/n)}{\epsilon^{1/6}\log(1/\epsilon)}\right)$$

For $4^{-1}(2e)^{-6} \leq \epsilon < 1/8$, we recall the fact that $\Omega(n\log(m/n))$ is always a lower bound (the information-theoretic lower bound) for the storage space. Thus, for these values of ϵ too

$$s = \Omega\left(\frac{n\log(m/n)}{\epsilon^{1/6}\log(1/\epsilon)}\right)$$

Hence, the theorem is proved. ■

We now show how to extend the above argument for 2-sided ϵ-error quantum schemes which make p probes.

Theorem 4 *For any $p \geq 1$ and $n/m < \epsilon < 2^{-3p}$, suppose there is a scheme which stores subsets S of size at most n from a universe $\mathbf{U}$ of size m that answers membership queries, with two-sided error at most ϵ, using p quantum probes. Define $\delta \stackrel{\Delta}{=} \epsilon^{1/p}$. The scheme must use space*

$$s = \Omega\left(\frac{n\log(m/n)}{\delta^{1/6}\log(1/\delta)}\right)$$

Proof: **(Sketch)** The proof of this theorem is similar to the proof of Theorem 3. Pick a family F of sets, $F = \{S_1, \ldots, S_k\}$, $S_i \subseteq \mathbf{U}$, $|S_i| = n$, $|S_i \cap S_j| \leq n/2$ for all $i \neq j$, such that $|F| \geq (m/4n)^{n/2}$. One can prove that $\left\{W_S^{\otimes nt}\right\}_{S\in F}$, $t \stackrel{\Delta}{=} \left\lceil \frac{4\log|F|}{n\log(1/(4\epsilon))} \right\rceil$, are linearly independent in exactly the same fashion as Claim 3 in Theorem 3 was proved. The difference is that $\left\{W_S^{\otimes nt}\right\}_{S\in F}$ lie in a vector space of dimension at most $\sum_{j=0}^{pnt} \binom{s}{j}$

instead of $\sum_{j=0}^{nt} \binom{s}{j}$. This statement can be proved just as Claim 2 in Theorem 1 was proved. Therefore, by a argument similar to that at the end of the proof of Theorem 3, we get a lower bound

$$\Omega\left(\frac{n \log(m/n)}{\delta^{1/6} \log(1/\delta)}\right)$$

■

3.2 Classical schemes

We now give the proof for the space-time tradeoff equation in the classical deterministic case.

Theorem 5 *Suppose there exists a classical deterministic scheme for storing subsets S of size at most n from a universe $\mathbf{U}$ of size m which uses s bits of storage and answers membership queries with t classical bit probes. Then,*

$$\sum_{i=0}^{n} \binom{m}{i} \leq \sum_{i=0}^{nt} \binom{s}{i}$$

Proof: For $1 \leq i \leq m$, let $f_i : \{0,1\}^s \to \mathbb{R}$ denote the function for query i, which maps bit strings of length s to $\{0,1\} \subset \mathbb{R}$ i.e. f_i maps $x \in \{0,1\}^s$ to 1 iff the query scheme given query i and bit string x evaluates to 1. Consider a mapping $\Phi : \binom{\mathbf{U}}{\leq n} \to (\{0,1\}^s \to \mathbb{R})$ i.e. Φ takes a subset of the universe of size at most n to a function from bit strings of length s to the reals. Φ is defined as follows

$$\Phi(\{\}) \triangleq \text{constant 1 function}$$

$$\Phi(S) \triangleq f_{i_1} f_{i_2} \cdots f_{i_k}, \quad S = \{i_1, \cdots i_k\}, \ S \neq \{\}$$

Claim 5 $\{\Phi(S)\}_{S \in \binom{\mathbf{U}}{\leq n}}$ *are linearly independent over* $\mathbb{R}$.

Proof: Suppose there exists a non-trivial linear combination

$$\sum_{S \in \binom{\mathbf{U}}{\leq n}} \alpha_S \Phi(S) = 0$$

Pick a set T of smallest cardinality such that $\alpha_T \neq 0$. Let $x(T) \in \{0,1\}^s$ be the string stored by the storage scheme. Applying $x(T)$ to the above linear combination, we get

$$\sum_{S \in \binom{\mathbf{U}}{\leq n}} \alpha_S \Phi(S) x(T) = 0$$

If $S \neq T$, there exists an element $i \in \mathbf{U}$ such that $i \in S - T$. Then, $f_i(x(T)) = 0$, and hence, $\Phi(S)(x(T)) = 0$. If $S = T$, then $\Phi(S)(x(T)) = \Phi(T)(x(T)) = 1$. Hence, $\alpha_T = 0$ which is a contradiction. Hence the claim is proved. ■

Claim 6 $\{\Phi(S)\}_{S\in\binom{\mathbf{U}}{\leq n}}$ *lie in a vector space of dimension at most* $\sum_{i=0}^{nt}\binom{s}{i}$.

Proof: Since the query scheme is deterministic and makes at most t (classical) bit probes, given a query i, $1 \leq i \leq m$, the function f_i is modelled by a decision tree of depth at most t. Hence f_i can be represented over $\mathbb{R}$ as a sum of products of at most t linear functions, where the linear functions are either y_j (representing the value stored at location j in the bit string) or $1-y_j$ (representing the negation of the value stored at location j). Note that for any $y \in \{0,1\}^s$, at most one of these products evaluates to 1. Such a function can be represented as a multilinear polynomial in $y_1, y_2, \ldots, y_s$ of degree at most t. A product of at most n such functions can be represented as a multilinear polynomial of degree at most nt. Hence, $\{\Phi(S)\}_{S\in\binom{\mathbf{U}}{\leq n}}$ lie in the span of at most $\sum_{i=0}^{nt}\binom{s}{i}$ functions from $\{0,1\}^s$ to $\mathbb{R}$. From this, the claim follows. ■

From the above two claims, the theorem follows. ■

In fact, the tradeoff result can be extended to the one-sided error classical case too.

Theorem 6 *The tradeoff result of Theorem 5 also holds for classical schemes where the query scheme may err with probability less than* 1 *on the positive instances (i.e. if an element is present it might report it to be absent), but not on the negative instances (i.e, if an element is absent it has to be reported as absent). In fact, the tradeoff result holds for nondeterministic query schemes too.*

Proof: **(Sketch)** A proof very similar to that of Theorem 5 goes through. We just observe that now the query scheme is a logical disjunction over a family of deterministic query schemes. If the query element is present in the set stored, there is a decision tree in this family that outputs 1. If the query element is not present in the set stored, then all the decision trees output 0. Let us denote by F_i the family of decision trees corresponding to query element i, $1 \leq i \leq m$. For any decision tree D in F_i, let $g_D : \{0,1\}^s \to \{0,1\}$ be the function it evaluates.

Let us now define $f_i \triangleq \sum_{D\in F_i} g_D$. Then

$$f_i(x[T]) \begin{cases} \geq 1 & \text{if } i \in T \\ = 0 & \text{otherwise} \end{cases}$$

With this choice of f_i, the rest of the proof is the same as in the deterministic case. ■

Now we give a simple proof of the lower bound for the space used by a classical randomised scheme which answers membership queries with two-sided error at most ϵ and uses only one bit probe.

Theorem 7 *Let* $1/18 > \epsilon > 1/m^{1/3}$ *and* $m^{1/3} > 18n$. *Any classical scheme which stores subsets S of size at most n from a universe $\mathbf{U}$ of size m and answers membership queries, with two-sided error at most ϵ, using one bit probe must use space*

$$\Omega\left(\frac{n\log m}{\epsilon^{2/5}\log(1/\epsilon)}\right)$$

Proof: Suppose there is a classical scheme which stores subsets of size n from a universe of size m using s bits of storage, and answers membership queries using one bit probe with two-sided error at most ϵ. Define $k \triangleq \left\lceil \frac{4\log(27m)}{3\log(1/4e\epsilon)} \right\rceil$. Since $m^{1/3} > 1/\epsilon$, $\frac{4\log(27m)}{3\log(1/4e\epsilon)} \geq 4$. Therefore,

$$\frac{4\log(27m)}{3\log(1/4e\epsilon)} \leq k \leq \frac{4\log(27m)}{3\log(1/4e\epsilon)} + 1 \leq \frac{5\log(27m)}{3\log(1/4e\epsilon)} \tag{6}$$

We repeat the query scheme k times and accept only if more than $3k/4$ trials accept. Then the probability of making an error on a positive instance (i.e. the query element is present in the set stored) is bounded by

$$\binom{k}{k/4}\epsilon^{k/4} \leq (4e)^{k/4}\epsilon^{k/4} = (4e\epsilon)^{k/4}$$

The probability of making an error on a negative instance (i.e. the query element is not present in the set stored) is bounded by

$$\binom{k}{3k/4}\epsilon^{3k/4} \leq \left(\frac{4e\epsilon}{3}\right)^{3k/4} \leq (4e\epsilon)^{3k/4}$$

From lower bound on k from (6), we get

$$\begin{aligned} \Pr[\text{Error on a positive instance}] &\leq (4e\epsilon)^{k/4} \leq \tfrac{1}{(27m)^{1/3}} \leq \tfrac{1}{3n} \\ \Pr[\text{Error on a negative instance}] &\leq (4e\epsilon)^{3k/4} \leq \tfrac{1}{27m} \end{aligned}$$

Hence, the probability that a random sequence of coin tosses gives the wrong answer on some query $q \in \mathbf{U}$ and a particular set S stored, is at most

$$\frac{1}{3n} \times n + \frac{1}{27m} \times (m - n) < \frac{1}{2}$$

Call a sequence of coin tosses bad for a set S, if when S is stored, there is one query $q \in \mathbf{U}$ for which the query scheme with these coin tosses gives the wrong answer. Thus, at most half of the coin toss sequences are bad for a fixed set S. By an averaging argument, there exists a sequence of coin tosses which is bad for at most half of the sets $S \in \binom{\mathbf{U}}{n}$. By setting the coin tosses to that sequence, we now get a deterministic scheme which answers membership queries correctly for at least half the sets $S \in \binom{\mathbf{U}}{n}$, and uses k bit probes. From the proof of Theorem 5, we have that

$$\frac{1}{2}\binom{m}{n} \leq \sum_{i=0}^{nk}\binom{s}{i} \leq \binom{s+nk}{nk} \leq \left(\frac{e(s+nk)}{nk}\right)^{nk}$$

$$\Rightarrow \frac{1}{2}\left(\frac{m}{n}\right)^{n} \leq \left(\frac{e(s+nk)}{nk}\right)^{nk}$$

Using the upper bound on k in (6) and the fact that $m^{1/3} > 18n$, we get

$$\left(\left(\frac{1}{4e\epsilon}\right)^{3k/5}\right)^{2n/3} \leq (27m)^{2n/3} = \left(9m^{2/3}\right)^n \leq \frac{1}{2}\left(18m^{2/3}\right)^n \leq \frac{1}{2}\left(\frac{m}{n}\right)^n$$

$$\Rightarrow \left(\frac{1}{4e\epsilon}\right)^{2nk/5} \leq \left(\frac{e(s+nk)}{nk}\right)^{nk}$$

$$\Rightarrow s \geq \frac{nk}{e}\left(\left(\frac{1}{4e\epsilon}\right)^{2/5} - e\right)$$

Arguing as in the last part of the proof of Theorem 3, and recalling that since $m^{1/3} > 18n$, $\Omega(n \log m)$ is always a lower bound (the information-theoretic lower bound) for the storage space, we get

$$s = \Omega\left(\frac{n \log m}{\epsilon^{2/5}\log(1/\epsilon)}\right)$$

■

We can extend the classical randomised two-sided error space lower bound above to the case of multiple bit probes.

Theorem 8 *Let $p \geq 1$, $18^{-p} > \epsilon > 1/m^{1/3}$ and $m^{1/3} > 18n$. Define $\delta \triangleq \epsilon^{1/p}$. Any classical scheme which stores subsets S of size at most n from a universe $\mathbf{U}$ of size m and answers membership queries, with two-sided error at most ϵ, using at most p bit probes must use space*

$$\Omega\left(\frac{n \log m}{\delta^{2/5}\log(1/\delta)}\right)$$

Proof: **(Sketch)** The proof of this theorem is similar to the proof of Theorem 7 above. We repeat the query scheme $k \triangleq \left\lceil \frac{4\log(27m)}{3\log(1/4e\epsilon)} \right\rceil$ times and accept only if more than $3k/4$ trials accept. We "derandomise" the new query scheme in a manner similar to what was done in the proof of Theorem 7. We thus get a deterministic query scheme making kp bit probes and answering membership queries correctly for at least half the sets $S \in \binom{\mathbf{U}}{n}$. The rest of the proof now follows in the same fashion as the proof of Theorem 7. ■

4 Conclusion and open problems

In this paper, we introduce the quantum bit probe model and study the complexity of the static membership problem in this model. We study the problem in the exact and bounded (one-sided and two-sided) error versions of the model and give lower bounds which almost match the corresponding classical upper bounds. We also give stronger/simplified proofs of lower bounds for the problem in the classical setting.

The paper of Buhrman *et al.* [BMRV00] also considers classical schemes for the static membership problem where the error is bounded and restricted only to negative instances (i.e. when the query element is not a member of the stored set). For such schemes, they give almost matching upper and lower bounds. But for negative one-sided error quantum schemes, we can only prove similar lower bounds as for two-sided error quantum schemes. Also, we do not know if there are negative one-sided error quantum schemes better than the classical ones in [BMRV00]. Thus there is a gap between the upper and lower bounds here, and resolving it is an open problem.

The classical bit probe model has been used in the past to study other static data structure problems like, for example, perfect hashing and element containment (see e.g. [Pag01, Mil93]). The complexity of these problems in the quantum bit probe model is an important open problem.

Acknowledgements

We thank the anonymous referees for their comments, which have helped us to improve the presentation of the paper and fix some bugs in the parameters appearing in some theorems. We also thank them for pointing out the similarities between Claim 2 in Theorem 1, and the observations of Beals *et al.* [BBC+98] and Shi [Shi00].

References

[Amb00] A. Ambainis. Quantum lower bounds by quantum arguments. In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, pages 636–643, 2000. Also quant-ph/0002066.

[AST+98] A. Ambainis, L. J. Schulman, A. Ta-Shma, U. Vazirani, and A. Wigderson. The quantum communication complexity of sampling. In *Proceedings of the 39th Annual IEEE Symposium on Foundations of Computer Science*, pages 342–351, 1998.

[BBBV97] C. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computation. *SIAM Journal of Computing*, 26(3):1510–1523, 1997. Also quant-ph/9701001.

[BBC+98] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. In *Proceedings of the 39th Annual IEEE Symposium on Foundations of Computer Science*, pages 352–361, 1998. Full version to appear in the Journal of the ACM. Also quant-ph/9802049.

[BdW01] H. Buhrman and R. de Wolf. Communication complexity lower bounds by polynomials. In *Proceedings of the 16th Annual Conference on Computational Complexity*, pages 120–130, 2001. Also cs.CC/9910010.

[BF92] L. Babai and P. Frankl. *Linear Algebra Methods in Combinatorics (with applications to Geometry and Computer Science)*. Preliminary Version 2, Department of Computer Science, The University of Chicago, September 1992.

[BMRV00] H. Buhrman, P. B. Miltersen, J. Radhakrishnan, and S. Venkatesh. Are bitvectors optimal? In *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, pages 449–458, 2000.

[DR82] A. Dyachkov and V. Rykov. Bounds on the length of disjunctive codes. *Problemy Peredachi Informatsii*, 18(3):7–13, 1982. (In Russian).

[EFF85] P. Erdős, P. Frankl, and Z. Füredi. Families of finite sets in which no set is covered by the union of r others. *Israel Journal of Mathematics*, 51:79–89, 1985.

[FKS84] M. Fredman, J. Komlós, and E. Szemerédi. Storing a sparse table with $O(1)$ worst case access time. *Journal of the Association for Computing Machinery*, 31(3):538–544, 1984.

[Gro96] L. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, pages 212–219, 1996. Also quant-ph/9605043.

[Mil93] P. B. Miltersen. The bitprobe complexity measure revisited. In *Proceedings of the 10th Annual Symposium on Theoretical Aspects of Computer Science*, pages 662–671, 1993.

[MP69] M. Minsky and S. Papert. *Perceptrons*. MIT Press, Cambridge, Mass., USA, 1969.

[NW94] N. Nisan and A. Wigderson. Hardness vs randomness. *Journal of Computer and System Sciences*, 49:149–167, 1994.

[Pag01] R. Pagh. On the cell probe complexity of membership and perfect hashing. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 425–432, 2001.

[Shi00] Y. Shi. Lower bounds of quantum black-box complexity and degree of approximating polynomials by influence of boolean variables. *Information Processing Letters*, 75(1-2):79–83, 2000. Also quant-ph/9904107.

[Yao81] A. C-C. Yao. Should tables be sorted? *Journal of the Association for Computing Machinery*, 28(3):615–628, 1981.