

# Quantum Direct Product Theorems for Symmetric Functions and Time-Space Tradeoffs

Andris Ambainis\*  
University of Waterloo  
ambainis@math.uwaterloo.ca

Robert Špalek†  
CWI, Amsterdam  
sr@cw.nl

Ronald de Wolf‡  
CWI, Amsterdam  
rdewolf@cw.nl

## Abstract

A direct product theorem upper-bounds the overall success probability of algorithms for computing many independent instances of a computational problem. We prove a direct product theorem for 2-sided error algorithms for symmetric functions in the setting of quantum query complexity, and a stronger direct product theorem for 1-sided error algorithms for threshold functions. We also present a quantum algorithm for deciding systems of linear inequalities, and use our direct product theorems to show that the time-space tradeoff of this algorithm is close to optimal.

---

\*Institute for Quantum Computing and Department of Combinatorics and Optimization, University of Waterloo. Supported by NSERC, ARO, CIAR and IQC University Professorship.

†Supported in part by the EU fifth framework project RESQ, IST-2001-37559.

‡Supported by a Veni grant from the Netherlands Organization for Scientific Research (NWO) and by the EU fifth framework project RESQ, IST-2001-37559.

# 1 Introduction

## 1.1 Direct product theorems for symmetric functions

Consider an algorithm that simultaneously needs to compute  $k$  independent instances of a function  $f$  (denoted  $f^{(k)}$ ). Direct product theorems deal with the optimal tradeoff between the resources and success probability of such algorithms. Suppose we need  $t$  “resources” to compute a single instance  $f(x)$  with bounded error probability. These resources could for example be time, space, ink, queries, communication, etc. A typical direct product theorem (DPT) has the following form:

Every algorithm with  $T \leq \alpha kt$  resources for computing  $f^{(k)}$  has success probability  $\sigma \leq 2^{-\Omega(k)}$  (where  $\alpha > 0$  is some small constant).

This expresses our intuition that essentially the best way to compute  $f^{(k)}$  on  $k$  independent instances is to run separate  $t$ -resource algorithms for each of the instances. Since each of those will have success probability less than 1, we expect that the probability of simultaneously getting all  $k$  instances right goes down exponentially with  $k$ . DPT’s can be stated for classical algorithms or quantum algorithms, and  $\sigma$  could measure worst-case success probability or average-case success probability under some input distribution. DPT’s are generally hard to prove, and Shaltiel [Sha01] even gives general examples where they are just not true (with  $\sigma$  average success probability), the above intuition notwithstanding. Klauck, Špalek, and de Wolf [KŠW04] recently examined the case where the resource is query complexity and  $f = \text{OR}$ , and proved an optimal DPT both for classical algorithms and for quantum algorithms (with  $\sigma$  worst-case success probability).

In this paper we generalize their results to the case where  $f$  can be any symmetric function, i.e., a function depending only on the Hamming weight  $|x|$  of its input. In the case of classical algorithms the situation is quite simple. Every  $n$ -bit symmetric function  $f$  has classical bounded-error query complexity  $R_2(f) = \Theta(n)$  and block sensitivity  $bs(f) = \Theta(n)$ , hence an optimal classical DPT follows immediately from [KŠW04, Theorem 3]. Classically, all symmetric functions essentially “cost the same” in terms of query complexity. This is different in the quantum world. For instance, the OR function has bounded-error quantum query complexity  $Q_2(\text{OR}) = \Theta(\sqrt{n})$  [Gro96, BHMT02], while Parity needs  $n/2$  quantum queries [BBC<sup>+</sup>01, FGGS98]. If  $f$  is a  $t$ -threshold function ( $f(x) = 1$  iff  $|x| \geq t$ , with  $t \leq n/2$ ), then  $Q_2(f) = \Theta(\sqrt{tn})$  [BBC<sup>+</sup>01].

Our main result is an essentially optimal quantum DPT for all symmetric functions:

There is a constant  $\alpha > 0$  such that for every symmetric  $f$  and every positive integer  $k$ :  
Every 2-sided error quantum algorithm with  $T \leq \alpha k Q_2(f)$  queries for computing  $f^{(k)}$  has success probability  $\sigma \leq 2^{-\Omega(k)}$ .

This paper gives an interesting new twist to the rivalry between the polynomial and adversary methods. These are the two main quantum lower-bound methods. The polynomial method [NS94, BBC<sup>+</sup>01] works by lower-bounding the degree of a polynomial that in some way represents the desired success probability. The adversary method [Amb00] identifies a set of hard inputs and shows that we need many queries to distinguish all pairs that have different outputs (many formulations of this method exist, but they are essentially equivalent [ŠS05]). The methods are incomparable. For instance, the polynomial method gives optimal lower bounds for the collision problem and element distinctness [Aar02, Shi02], and also works well for analyzing zero-error or low-error quantum algorithms [BBC<sup>+</sup>01, BCWZ99]. In both cases it’s better than the adversary method. On the other hand, the adversary method proves stronger bounds than the polynomial method for certain iterated

functions [Amb03], and also gives tight lower bounds for constant-depth AND-OR trees [Amb00, HMW03], where we do not know how to analyze the polynomial degree.

Our new direct product theorem generalizes the polynomial-based results of [KŠW04] (which strengthened the polynomial-based [Aar04]), but our current proof is a version of the adversary method, extending the techniques recently introduced by Ambainis [Amb05]. We have not been able to prove it using the polynomial method. We can, however, use the polynomial method to prove an incomparable DPT. This result is worse than our main result in applying only to *1-sided error* quantum algorithms<sup>1</sup> for *threshold* functions; but it's better in not having any constraint on the threshold and giving a much stronger upper bound on the success probability:

There is a constant  $\alpha > 0$  such that for every  $t$ -threshold function  $f$  and every positive integer  $k$ : Every 1-sided error quantum algorithm with  $T \leq \alpha k Q_2(f)$  queries for computing  $f^{(k)}$  has success probability  $\sigma \leq 2^{-\Omega(kt)}$ .

A similar theorem can be proven for the  $k$ -fold  $t$ -search problem, where in each of  $k$  inputs of  $n$  bits, we want to find at least  $t$  ones. The different error bounds  $2^{-\Omega(kt)}$  and  $2^{-\Omega(k)}$  for 1-sided and 2-sided error algorithms intuitively say that imposing the 1-sided error constraint makes deciding the  $k$ -fold threshold problem as hard as actually *finding*  $t$  ones in each of the  $k$  inputs.

## 1.2 Application: Time-Space tradeoffs for systems of linear inequalities

As an application we obtain near-optimal time-space tradeoffs for deciding systems of linear equalities. Such tradeoffs between the two main computational resources are well known classically for problems like sorting, element distinctness, hashing, etc. In the quantum world, essentially optimal time-space tradeoffs were recently obtained for sorting and for Boolean matrix multiplication [KŠW04], but little else is known.

Let  $A$  be a fixed  $N \times N$  matrix of nonnegative integers. Our inputs are column vectors  $x = (x_1, \dots, x_N)$  and  $b = (b_1, \dots, b_N)$  of nonnegative integers. We are interested in the system

$$Ax \geq b$$

of  $N$  linear inequalities, and want to find out which of these inequalities hold (we could also mix  $\geq$ ,  $=$ , and  $\leq$ , but omit that for ease of notation).<sup>2</sup> We want to analyze the tradeoff between the time  $T$  and space  $S$  needed to solve this problem. Lower bounds on  $T$  will be in terms of query complexity. For simplicity we omit polylog factors in the following discussion.

In the classical world, the optimal tradeoff is  $TS = N^2$ , independent of the values in  $b$ . This follows from [KŠW04, Section 7]. The upper bounds are for deterministic algorithms and the lower bounds are for 2-sided error algorithms. In the quantum world the situation is more complex. Let us put an upper bound  $\max\{b_i\} \leq t$ . We have two regimes for 2-sided error quantum algorithms:

- Quantum regime. If  $S \leq N/t$  then the optimal tradeoff is  $T^2S = tN^3$  (better than classical).
- Classical regime. If  $S > N/t$  then the optimal tradeoff is  $TS = N^2$  (same as classical).

Our lower bounds hold even for the constrained situation where  $b$  is fixed to the all- $t$  vector,  $A$  and  $x$  are Boolean, and  $A$  is sparse in having only  $O(N/S)$  non-zero entries in each row.

Since our DPT for 1-sided error algorithms is stronger by an extra factor of  $t$  in the exponent, we obtain a stronger lower bound for 1-sided error algorithms:

<sup>1</sup>The error is 1-sided if 1-bits in the  $k$ -bit output vector are always correct.

<sup>2</sup>Note that if  $A$  and  $x$  are Boolean and  $b = (t, \dots, t)$ , this gives  $N$  overlapping  $t$ -threshold functions.

- If  $t \leq S \leq N/t^2$  then the optimal tradeoff for 1-sided error algorithms is  $T^2S \geq t^2N^3$ .
- If  $S > N/t^2$  then the optimal tradeoff for 1-sided error algorithms is  $TS = N^2$ .

We do not know whether the lower bound in the first case is optimal (probably it is not), but note that it is stronger than the optimal bounds that we have for 2-sided error algorithms. This is the first separation of 2-sided and 1-sided error algorithms in the context of quantum time-space tradeoffs.<sup>3</sup>

#### Remarks:

1. Klauck et al. [KŠW04] gave direct product theorems not only for quantum query complexity, but also for 2-party quantum communication complexity, and derived some communication-space tradeoffs in analogy to the time-space tradeoffs. This was made possible by a translation of communication protocols to polynomials due to Razborov [Raz03], and the fact that the DPTs of [KŠW04] were polynomial-based. Some of the results in this paper can similarly be ported to a communication setting, though only the ones that use the polynomial method.

2. The time-space tradeoffs for 2-sided error algorithms for  $Ax \geq b$  similarly hold for a system of  $N$  equalities,  $Ax = b$ . The upper bound clearly carries over, while the lower holds for equalities as well, because our DPT holds even under the promise that the input has weight  $t$  or  $t - 1$ . In contrast, the stronger 1-sided error time-space tradeoff does not automatically carry over to systems of equalities, because we do not know how to prove the DPT with bound  $2^{-\Omega(kt)}$  under this promise.

## 2 Preliminaries

We assume familiarity with quantum computing [NC00] and sketch the model of quantum query complexity, referring to [BW02] for more details, also on the close relation between query complexity and degrees of multivariate polynomials. Suppose we want to compute some function  $f$ . For input  $x \in \{0, 1\}^N$ , a *query* gives us access to the input bits. It corresponds to the unitary transformation

$$O_x : |i, b, z\rangle \mapsto |i, b \oplus x_i, z\rangle.$$

Here  $i \in [N] = \{1, \dots, N\}$  and  $b \in \{0, 1\}$ ; the  $z$ -part corresponds to the workspace, which is not affected by the query. We assume the input can be accessed only via such queries. A  $T$ -query quantum algorithm has the form  $A = U_T O_x U_{T-1} \cdots O_x U_1 O_x U_0$ , where the  $U_k$  are fixed unitary transformations, independent of  $x$ . This  $A$  depends on  $x$  via the  $T$  applications of  $O_x$ . The algorithm starts in initial  $S$ -qubit state  $|0\rangle$  and its *output* is the result of measuring a dedicated part of the final state  $A|0\rangle$ . For a Boolean function  $f$ , the output of  $A$  is obtained by observing the leftmost qubit of the final superposition  $A|0\rangle$ , and its *acceptance probability* on input  $x$  is its probability of outputting 1. We mention some well known quantum algorithms that we use as subroutines.

- **Quantum search.** Grover's search algorithm [Gro96, BBHT98] can find an index of a 1-bit in an  $n$ -bit input in expected number of  $O(\sqrt{n/(|x| + 1)})$  queries, where  $|x|$  is the Hamming weight (number of ones) in the input. If  $|x|$  is known, the algorithm can be made exact rather than expected [BHMT02]. By repeated search, we can find  $t$  ones in an  $n$ -bit input with  $|x| \geq t$ , using  $\sum_{i=|x|-t+1}^{|x|} O(\sqrt{n/(i + 1)}) = O(\sqrt{tn})$  queries.

---

<sup>3</sup>Strictly speaking, there's a quadratic gap for OR, but space  $\log n$  suffices for the fastest 1-sided and 2-sided error algorithms so there's no real tradeoff in that case.

- **Quantum counting** [BHMT02, Theorem 13]. There is a quantum algorithm that uses  $M$  queries to  $n$ -bit  $x$  to compute an estimate  $w$  of  $|x|$  such that with probability at least  $8/\pi^2$

$$|w - |x|| \leq 2\pi \frac{\sqrt{|x|(n - |x|)}}{M} + \pi^2 \frac{n}{M^2}.$$

For investigating time-space tradeoffs we use the circuit model. A circuit accesses its input via an oracle like a query algorithm. Time corresponds to the number of gates in the circuit. We will, however, usually consider the number of queries to the input, which is obviously a lower bound on time. A circuit uses space  $S$  if it works with  $S$  bits/qubits only. We require that the outputs are made at predefined gates in the circuit, by writing their value to some extra bits/qubits that may not be used later on.

### 3 Direct Product Theorem for Symmetric Functions (2-sided)

Consider some symmetric function  $f : \{0,1\}^n \rightarrow \{0,1\}$ . Let  $t$  denote the smallest nonnegative integer such that  $f$  is constant on the interval  $|x| \in [t, n - t]$ . We call this value  $t$  the “implicit threshold” of  $f$ . For instance, functions like OR and AND have  $t = 1$ , while Parity and Majority have  $t = \lfloor n/2 \rfloor$ . If  $f$  is the  $t$ -threshold function, then the implicit threshold is just the threshold. The implicit threshold is related to the parameter  $\Gamma(f)$  introduced by Paturi [Pat92] via  $t = n/2 - \Gamma(f)/2 \pm 1$ . It characterizes the bounded-error quantum query complexity of  $f$ :  $Q_2(f) = \Theta(\sqrt{tn})$  [BBC<sup>+</sup>01].

The main result of this paper is the following theorem.

**Theorem 1** *There is a constant  $\alpha > 0$  such that for every symmetric  $f$  and every positive integer  $k$ : Every 2-sided error quantum algorithm with  $T \leq \alpha k Q_2(f)$  queries for computing  $f^{(k)}$  has success probability  $\sigma \leq 2^{-\Omega(k)}$ .*

We actually prove a stronger statement, applying to any Boolean function  $f$  (total or partial) for which  $f(x) = 0$  if  $|x| = t - 1$  and  $f(x) = 1$  if  $|x| = t$ . In this section we give an outline of the proof. Most of the proofs of technical claims have been deferred to Appendix A.

Let  $\mathcal{A}$  be an algorithm that computes  $k$  instances of this weight- $(t-1)$  versus weight- $t$  problem. We recast  $\mathcal{A}$  into a different form, using a register that stores the input  $x^1, \dots, x^k$ . Let  $\mathcal{H}_A$  be the Hilbert space on which  $\mathcal{A}$  operates. Let  $\mathcal{H}_I$  be an  $((\binom{n}{t-1} + \binom{n}{t}))^k$ -dimensional Hilbert space whose basis states correspond to inputs  $(x^1, \dots, x^k)$  with  $|x^1| \in \{t-1, t\}, \dots, |x^k| \in \{t-1, t\}$ . We transform  $\mathcal{A}$  into a sequence of transformations on a Hilbert space  $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_I$ . A non-query transformation  $U$  on  $\mathcal{H}_A$  is replaced with  $U \otimes I$  on  $\mathcal{H}$ . A query is replaced by a transformation  $O$  that is equal to  $O_{x^1, \dots, x^k} \otimes I$  on the subspace consisting of states of the form  $|s\rangle_A \otimes |x^1, \dots, x^k\rangle_I$ . The starting state of the algorithm on Hilbert space  $\mathcal{H}$  is  $|\varphi_0\rangle = |\psi_{start}\rangle_A \otimes |\psi_0\rangle_I$  where  $|\psi_{start}\rangle$  is the starting state of  $\mathcal{A}$  as an algorithm acting on  $\mathcal{H}_A$  and  $|\psi_0\rangle$  is the uniform superposition of all basis states of  $\mathcal{H}_I$ :

$$|\psi_0\rangle = \frac{1}{(\binom{n}{t-1} + \binom{n}{t})^{k/2}} \sum_{\substack{x^1, \dots, x^k: \\ |x^1|, \dots, |x^k| \in \{t-1, t\}}} |x^1 \dots x^k\rangle.$$

Let  $|\varphi_d\rangle$  be the state of the algorithm  $\mathcal{A}$ , as a sequence of transformations on  $\mathcal{H}$ , after the  $d^{\text{th}}$  query. Let  $\rho_d$  be the mixed state in  $\mathcal{H}_I$  obtained from  $|\varphi_d\rangle$  by tracing out the  $\mathcal{H}_A$  register.

We define two decompositions of  $\mathcal{H}_I$  into a direct sum of subspaces. We have  $\mathcal{H}_I = (\mathcal{H}_{one})^{\otimes k}$  where  $\mathcal{H}_{one}$  is the Hilbert space with basis states  $|x\rangle$ ,  $x \in \{0, 1\}^n$ ,  $|x| \in \{t-1, t\}$ . Let

$$|\psi_{i_1, \dots, i_j}^0\rangle = \frac{1}{\sqrt{\binom{n-j}{t-1-j}}} \sum_{\substack{x_1, \dots, x_n: \\ x_1 + \dots + x_n = t-1, \\ x_{i_1} = \dots = x_{i_j} = 1}} |x_1 \dots x_n\rangle$$

and let  $|\psi_{i_1, \dots, i_j}^1\rangle$  be a similar state with  $x_1 + \dots + x_n = t$  instead of  $x_1 + \dots + x_n = t-1$ . Let  $T_{j,0}$  (resp.  $T_{j,1}$ ) be the space spanned by all states  $|\psi_{i_1, \dots, i_j}^0\rangle$  (resp.  $|\psi_{i_1, \dots, i_j}^1\rangle$ ) and let  $S_{j,l} = T_{j,l} \cap T_{j-1,l}^\perp$ . For a subspace  $S$ , we use  $\Pi_S$  to denote the projector onto  $S$ . Let  $|\tilde{\psi}_{i_1, \dots, i_j}^l\rangle = \Pi_{T_{j-1,l}^\perp} |\psi_{i_1, \dots, i_j}^l\rangle$ . For  $j < t$ , let  $S_{j,+}$  be the subspace spanned by the states

$$\frac{|\tilde{\psi}_{i_1, \dots, i_j}^0\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^0\|} + \frac{|\tilde{\psi}_{i_1, \dots, i_j}^1\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^1\|}$$

and  $S_{j,-}$  be the subspace spanned by

$$\frac{|\tilde{\psi}_{i_1, \dots, i_j}^0\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^0\|} - \frac{|\tilde{\psi}_{i_1, \dots, i_j}^1\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^1\|}$$

For  $j = t$ , we define  $S_{t,-} = S_{t,1}$  and there is no subspace  $S_{t,+}$ . Thus,  $\mathcal{H}_{one} = \bigoplus_{j=0}^{t-1} (S_{j,+} \oplus S_{j,-}) \oplus S_{t,-}$ . For the space  $\mathcal{H}_I$  (representing  $k$  inputs for the  $t$ -threshold function), we define

$$S_{j_1, \dots, j_k, l_1, \dots, l_k} = S_{j_1, l_1} \otimes S_{j_2, l_2} \otimes \dots \otimes S_{j_k, l_k}.$$

Let  $S_{m-}$  be the direct sum of all  $S_{j_1, \dots, j_k, l_1, \dots, l_k}$  such that exactly  $m$  of  $l_1, \dots, l_k$  are equal to  $-$ . Then,  $\mathcal{H}_I = \bigoplus_m S_{m-}$ . This is the first decomposition. In Appendix A.1 we prove:

**Lemma 2** *Let  $\rho$  be the reduced density matrix of  $\mathcal{H}_I$ . If the support of  $\rho$  is contained in  $S_{0-} \oplus S_{1-} \oplus \dots \oplus S_{m-}$ , then the probability that measuring  $\mathcal{H}_A$  gives the correct answer is at most  $\frac{\sum_{m'=0}^m \binom{k}{m'}}{2^k}$ .*

The following consequence of this lemma is proved in Appendix A.2:

**Corollary 3** *Let  $\rho$  be the reduced density matrix of  $\mathcal{H}_I$ . The probability that measuring  $\mathcal{H}_A$  gives the correct answer is at most*

$$\frac{\sum_{m'=0}^m \binom{k}{m'}}{2^k} + 4\sqrt{\text{Tr} \Pi_{(S_{0-} \oplus S_{1-} \oplus \dots \oplus S_{m-})^\perp} \rho}.$$

To define the second decomposition, we express  $\mathcal{H}_{one} = \bigoplus_{j=0}^{t/2} S'_j$  with  $S'_j = S_{j,+}$  for  $j < t/2$  and

$$S'_{t/2} = \bigoplus_{j \geq t/2} S_{j,+} \oplus \bigoplus_{j \geq 0} S_{j,-}.$$

Let  $V_m$  be the direct sum of all  $S'_{j_1} \otimes S'_{j_2} \otimes \dots \otimes S'_{j_k}$  satisfying  $j_1 + \dots + j_k = m$ . Then,  $\mathcal{H}_A = \bigoplus_{m=0}^{tk/2} V_m$ . This is the second decomposition.

Let  $V'_j = \bigoplus_{m=j}^{tk/2} V_m$ . We have  $S_{m-} \subseteq V'_{tm/2}$ . If we prove an upper bound on  $\text{Tr } \Pi_{V'_{tm/2}} \rho_d$ , where  $d$  is the total number of queries, this bound together with Corollary 3 implies an upper bound on the success probability of  $\mathcal{A}$ . To prove this, we consider the following potential function

$$P(\rho) = \sum_{m=0}^{tk/2} q^m \text{Tr } \Pi_{V_m} \rho,$$

where  $q = 1 + \frac{1}{t}$ . Then,

$$\text{Tr } \Pi_{V'_{tm/2}} \rho_d \leq P(\rho_d) q^{-tm/2} = P(\rho_d) e^{-(1+o(1))m/2}. \quad (1)$$

$P(\rho_0) = 1$ , because the initial state  $|\psi_0\rangle$  is a tensor product of the uniform superpositions

$$\frac{1}{\sqrt{\binom{n}{t-1} + \binom{n}{t}}} \left( \sum_{x:|x|=t-1} |x\rangle + \sum_{x:|x|=t} |x\rangle \right)$$

on each copy of  $\mathcal{H}_{one}$  and these uniform superpositions belong to  $S_{0,+}$ . In Appendix A.3 we prove

**Lemma 4**  $P(\rho_{j+1}) \leq \left(1 + \frac{C}{\sqrt{tn}}(q^{t/2} - 1) + \frac{C\sqrt{t}}{\sqrt{n}}(q - 1)\right) P(\rho_j)$ , for some constant  $C$ .

Since  $q = 1 + \frac{1}{t}$ , Lemma 4 means that  $P(\rho_{j+1}) \leq (1 + \frac{C\sqrt{e}}{\sqrt{tn}})P(\rho_j)$  and  $P(\rho_j) \leq (1 + \frac{C\sqrt{e}}{\sqrt{tn}})^j \leq e^{2Cj/\sqrt{tn}}$ . By equation (1),

$$\text{Tr } \Pi_{V'_{tm/2}} \rho_j \leq e^{2Cj/\sqrt{tn} - (1+o(1))m/2}.$$

We take  $m = k/3$ . Then, if  $j \leq m\sqrt{tn}/8C$ , this expression is exponentially small in  $k$ . Together with Corollary 3, this implies the theorem.

## 4 Direct Product Theorem for Threshold Functions (1-sided)

The previous section used the adversary method to prove a direct product theorem for 2-sided error algorithms computing  $k$  instances of some symmetric function. In this section we use the polynomial method to obtain stronger direct product theorems for 1-sided error algorithms for threshold functions. An algorithm for  $f^{(k)}$  has 1-sided error if the 1's in its  $k$ -bit output vector are always correct. Since our use of polynomials is a relatively small extension of the argument in [KŠW04], we have deferred it to Appendix B.

**Theorem 5** *There exists  $\alpha > 0$  such that for every threshold function  $T_t$  and positive integer  $k$ : Every 1-sided error quantum algorithm with  $T \leq \alpha k Q_2(T_t)$  queries for computing  $T_t^{(k)}$  has success probability  $\sigma \leq 2^{-\Omega(kt)}$ .*

**Proof.** We assume without loss of generality that  $t \leq n/20$ , the other cases can easily be reduced to this. We know that  $Q_2(T_t) = \Theta(\sqrt{tn})$  [BBC<sup>+</sup>01]. Consider a quantum algorithm  $A$  with  $T \leq \alpha k \sqrt{tn}$  queries that computes  $f^{(k)}$  with success probability  $\sigma$ . Roughly speaking, we use  $A$

to solve one big threshold problem on the total input, and then invoke the polynomial lemma to upper bound the success probability.

Define a new quantum algorithm  $B$  on an input  $x$  of  $N = kn$  bits, as follows:  $B$  runs  $A$  on a random permutation  $\pi(x)$ , and then outputs 1 iff the  $k$ -bit output vector has at least  $k/2$  ones.

Let  $m = kt/2$ . Note that if  $|x| < m$ , then  $B$  always outputs 0 because the 1-sided error output vector must have fewer than  $k/2$  ones. Now suppose  $|x| = 8m = 4kt$ . Call an  $n$ -bit input block “full” if  $\pi(x)$  contains at least  $t$  ones in that block. Let  $F$  be the random variable counting how many of the  $k$  blocks are full. We claim that  $\Pr[F \geq k/2] \geq 1/9$ . To prove this, observe that the number  $B$  of ones in one fixed block is a random variable distributed according to a hypergeometric distribution ( $4kt$  balls into  $N$  boxes,  $n$  of which count as success) with expectation  $\mu = 4t$  and variance  $V \leq 4t$ . Using Chebyshev’s inequality we bound the probability that this block is not full:

$$\Pr[B < t] \leq \Pr[|B - \mu| > 3t] \leq \Pr[|B - \mu| > (3\sqrt{t}/2)\sqrt{V}] < \frac{1}{(3\sqrt{t}/2)^2} \leq \frac{4}{9}.$$

Hence the probability that the block is full ( $B \geq t$ ) is at least  $5/9$ . This is true for each of the  $k$  blocks, so using linearity of expectation we have

$$\frac{5k}{9} \leq \text{Exp}[F] \leq \Pr[F \geq k/2] \cdot k + (1 - \Pr[F \geq k/2]) \cdot \frac{k}{2}.$$

This implies  $\Pr[F \geq k/2] \geq 1/9$ , as claimed. But then on all inputs with  $|x| = 8m$ ,  $B$  outputs 1 with probability at least  $\sigma/9$ .

Algorithm  $B$  uses  $\alpha k \sqrt{tn}$  queries. By [BBC<sup>+</sup>01] and symmetrization,  $B$ ’s acceptance probability is a single-variate polynomial  $p$  of degree  $D \leq 2\alpha k \sqrt{tn}$  such that

$$\begin{aligned} p(i) &= 0 \text{ for all } i \in \{0, \dots, m-1\}, \\ p(8m) &\geq \sigma/9, \\ p(i) &\in [0, 1] \text{ for all } i \in \{0, \dots, N\}. \end{aligned}$$

The result now follows by applying Lemma 18 (Appendix B) with  $N = kn$ ,  $m = kt/2$ ,  $E = 10$ , and  $\alpha$  a sufficiently small positive constant.  $\square$

## 5 Time-Space Tradeoff for Systems of Linear Inequalities

Let  $A$  be a fixed  $N \times N$  matrix of nonnegative integers and let  $x, b$  be two input vectors of  $N$  nonnegative integers smaller or equal to  $t$ . A *matrix-vector product with upper bound*, denoted by  $y = (Ax)_{\leq b}$ , is a vector  $y$  such that  $y[i] = \min((Ax)[i], b[i])$ . An *evaluation of a system of linear inequalities*  $Ax \geq b$  is the  $N$ -bit vector of the truth values of the individual inequalities. Here we present a quantum algorithm for matrix-vector product with upper bound that satisfies time-space tradeoff  $T^2S = O(tN^3(\log N)^5)$ . We then use our direct product theorems to show this is close to optimal.

### 5.1 Upper bound

It is simple to prove that matrix-vector products with upper bound  $t$  can be computed classically with  $TS = O(N^2 \log t)$ , as follows. Let  $S' = S/\log t$  and divide the matrix  $A$  into  $(N/S')^2$  blocks

of size  $S' \times S'$  each. The output vector is evaluated row-wise as follows: (1) Clear  $S'$  counters, one for each row, and read  $b[i]$ . (2) For each block, read  $S'$  input variables, multiply them by the corresponding submatrix of  $A$ , and update the counters, but do not let them grow larger than  $b[i]$ . (3) Output the counters. The space used is  $O(S' \log t) = O(S)$  and the total query complexity is  $T = O(\frac{N}{S'} \cdot \frac{N}{S'} \cdot S') = O(N^2 \log t / S)$ .

The quantum algorithm works in a similar way. We compute the matrix product in groups of  $S' = S / \log N$  rows, read input variables, and update the counters accordingly. The advantage over the classical algorithm is that we use the faster quantum search and quantum counting for finding non-zero entries.

The  $u$ -th row is called *open* if its counter has not yet reached  $b[u]$ . The subroutine *SmallMatrixProduct* maintains a set of open rows  $U \subseteq \{1, \dots, S'\}$  and counters  $0 \leq y[u] \leq b[u]$  for all  $u \in U$ . We process the input  $x$  in blocks, each containing between  $S' - O(\sqrt{S'})$  and  $2S' + O(\sqrt{S'})$  non-zero numbers at the positions  $j$  where  $A[u, j] \neq 0$  for some  $u \in U$ . The length  $\ell$  of such a block is first found by iterated quantum counting (with number of queries specified in the proof below) and the non-zero input numbers are then found by an iterated Grover search. For each such a number, we update all counters  $y[u]$  and close all rows that have exceeded their threshold  $b[u]$ .

*MatrixProduct* (fixed matrix  $A_{N \times N}$  and threshold  $t$ , input vectors  $x$  and  $b$  of length  $N$ ) returns output vector  $y = (Ax)_{\leq b}$ :

- For  $i = 1, 2, \dots, \frac{N}{S'}$ , where  $S' = S / \log N$ :
  1. Run *SmallMatrixProduct* on the  $i$ -th block of  $S'$  rows of  $A$ .
  2. Output the  $S'$  obtained results for those rows.

*SmallMatrixProduct* (fixed  $A_{S' \times N}$ , input  $x_{N \times 1}$  and  $b_{S' \times 1}$ ) returns  $y_{S' \times 1} = (Ax)_{\leq b}$ :

1. Read  $b$ , initialize  $y := (0, 0, \dots, 0)$ ,  $p := 1$ , and  $U := \{1, \dots, S'\}$ .  
Let  $B_{1 \times N}^U$  denote the row-vector  $B^U[j] = \sum_{u \in U} A[u, j]$ ; it is computed on-line.
2. While  $p \leq N$  and  $U \neq \emptyset$ , do the following:
  - (a) Let  $C_{p,k}^U$  denote the scalar product  $B^U[p, \dots, p+k-1] \cdot x[p, \dots, p+k-1]$ ; it is estimated by quantum counting. Initialize  $k = S'$ .  
First, while  $p+k-1 < N$  and  $C_{p,k}^U < S'$ , double  $k$ . Second, find by binary search the maximal  $\ell \in [\frac{k}{2}, k]$  such that  $p+\ell-1 \leq N$  and  $C_{p,\ell}^U \leq 2S'$ .
  - (b) Use quantum search to find the set  $J$  of all positions  $j \in \{p, \dots, p+\ell-1\}$  such that  $B^U[j]x[j] > 0$ .
  - (c) For all  $j \in J$ , read  $x[j]$ , and then do the following for all  $u \in U$ :
    - Increase  $y[u]$  by  $A[u, j]x[j]$ .
    - If  $y[u] \geq b[u]$ , set  $y[u] := b[u]$  and remove  $u$  from  $U$ .
  - (d) Increase  $p$  by  $\ell$ .
3. Return  $y$ .

**Theorem 6** *MatrixProduct has bounded error probability, its space complexity is  $O(S)$ , and its query complexity is  $T = O(N^{3/2} \sqrt{t} \cdot (\log N)^{5/2} / \sqrt{S})$ .*

**Proof.** In Appendix C. Basically, one just uses quantum counting with the number of queries  $M = \sqrt{\text{length of the interval}}$  and applies the Cauchy-Schwarz inequality a few times.  $\square$

## 5.2 Lower bound

Here we use our direct product theorems to lower-bound the quantity  $T^2S$  for  $T$ -query,  $S$ -space quantum algorithms deciding systems of linear inequalities. The lower bound even holds if we fix  $b$  to the all- $t$  vector  $\vec{t}$  and let  $A$  and  $x$  be Boolean.

**Theorem 7** *Let  $S = \min(O(N/t), o(N/\log N))$ . There exists an  $N \times N$  Boolean matrix  $A$  such that every 2-sided error quantum algorithm that uses  $T$  queries and  $S$  qubits of space to decide a system  $Ax \geq \vec{t}$  of  $N$  inequalities, satisfies  $T^2S = \Omega(tN^3)$ .*

**Proof.** The proof is a modification of Theorem 22 of [KŠW04] (quant-ph version). They use the probabilistic method to establish the following

**Fact:** For every  $k = o(N/\log N)$ , there exists an  $N \times N$  Boolean matrix  $A$ , such that all rows of  $A$  have weight  $N/2k$ , and every set of  $k$  rows of  $A$  contains a set  $R$  of  $k/2$  rows with the following property: each row in  $R$  contains at least  $n = N/6k$  ones that occur in no other row of  $R$ .

Fix a matrix  $A$  for  $k = cS$ , for some constant  $c$  to be chosen later. Consider a quantum circuit with  $T$  queries and space  $S$  that solves the problem with success probability at least  $2/3$ . We “slice” the quantum circuit into disjoint consecutive slices, each containing  $Q = \alpha\sqrt{tNS}$  queries, where  $\alpha$  is the constant from our direct product theorem (Theorem 1). The total number of slices is  $L = T/Q$ . Together, these disjoint slices contain all  $N$  output gates. Our aim below is to show that with sufficiently small constant  $\alpha$  and sufficiently large constant  $c$ , no slice can produce more than  $k$  outputs. This will imply that the number of slices is  $L \geq N/k$ , hence

$$T = LQ \geq \frac{\alpha N^{3/2} \sqrt{t}}{c\sqrt{S}}.$$

Now consider any slice. It starts with an  $S$ -qubit state, delivered by the previous slice and depending on the input, then it makes  $Q$  queries and outputs some  $\ell$  results that are jointly correct with probability at least  $2/3$ . Suppose, by way of contradiction, that  $\ell \geq k$ . Then there exists a set of  $k$  rows of  $A$  such that our slice produces the  $k$  corresponding results ( $t$ -threshold functions) with probability at least  $2/3$ . By the above Fact, some set  $R$  of  $k/2$  of those rows has the property that each row in  $R$  contains a set of  $n = N/6k = \Theta(N/S)$  ones that do not occur in any of the  $k/2 - 1$  other rows of  $R$ . By setting all other  $N - kn/2$  bits of  $x$  to 0, we naturally get that our slice, with the appropriate  $S$ -qubit starting state, solves  $k/2$  independent  $t$ -threshold functions  $T_t$  on  $n$  bits each. (Note that we need  $t \leq n/2 = O(N/S)$ ; this follows from our assumption  $S = O(N/t)$  with appropriately small constant in the  $O(\cdot)$ .) Now we replace the initial  $S$ -qubit state by the completely mixed state, which has “overlap”  $2^{-S}$  with every  $S$ -qubit state. This turns the slice into a stand-alone algorithm solving  $T_t^{(k/2)}$  with success probability

$$\sigma \geq \frac{2}{3} 2^{-S}.$$

But this algorithm uses only  $Q = \alpha\sqrt{tNS} = O(\alpha k \sqrt{tn})$  queries, so our direct product theorem (Theorem 1) with sufficiently small constant  $\alpha$  implies

$$\sigma \leq 2^{-\Omega(k/2)} = 2^{-\Omega(cS/2)}.$$

Choosing  $c$  a sufficiently large constant (independent of this specific slice), our upper and lower bounds on  $\sigma$  contradict. Hence the slice must produce fewer than  $k$  outputs.  $\square$

It is easy to see that the case  $S \geq N/t$  (equivalently,  $t \geq N/S$ ) is at least as hard as the  $S = N/t$  case, for which we have the lower bound  $T^2S = \Omega(tN^3) = \Omega(N^4/S)$ , hence  $TS = \Omega(N^2)$ . But that lower bound matches the *classical* deterministic upper bound up to a logarithmic factor and hence is essentially tight also for quantum. We thus have two different regimes for space: for small space, a quantum computer is faster than a classical one in solving systems of linear inequalities, while for large space it is not.

A similar slicing proof using Theorem 5 (with each slice of  $Q = \alpha\sqrt{NS}$  queries producing at most  $S/t$  outputs) gives the following lower bound on time-space tradeoffs for 1-sided error algorithms.

**Theorem 8** *Let  $t \leq S \leq \min(O(N/t^2), o(N/\log N))$ . There exists an  $N \times N$  Boolean matrix  $A$  such that every 1-sided error quantum algorithm that uses  $T$  queries and  $S$  qubits of space to decide a system  $Ax \geq \vec{t}$  of  $N$  inequalities, satisfies  $T^2S = \Omega(t^2N^3)$ .*

Note that our lower bound  $\Omega(t^2N^3)$  for 1-sided error algorithms is higher by a factor of  $t$  than the best upper bounds for 2-sided error algorithms. This lower bound is probably not optimal. If  $S > N/t^2$  then the essentially optimal classical tradeoff  $TS = \Omega(N^2)$  takes over.

## 6 Summary

In this paper we proved two new quantum direct product theorems:

- For every symmetric function  $f$ , every 2-sided error quantum algorithm for  $f^{(k)}$  using fewer than  $\alpha k Q_2(f)$  queries has success probability at most  $2^{-\Omega(k)}$ .
- For every  $t$ -threshold function  $f$ , every 1-sided error quantum algorithm for  $f^{(k)}$  using fewer than  $\alpha k Q_2(f)$  queries has success probability at most  $2^{-\Omega(kt)}$ .

Both results are tight up to constant factors. The first is proved using the adversary method, the second using the polynomial method. From these results we derived the following time-space tradeoffs for quantum algorithms that decide a system  $Ax \geq b$  of  $N$  linear inequalities (where  $A$  is a fixed  $N \times N$  matrix of nonnegative integers,  $x, b$  are variable, and  $b_i \leq t$  for all  $i$ ):

- Every  $T$ -query,  $S$ -space 2-sided error quantum algorithm for deciding  $Ax \geq b$  satisfies  $T^2S = \Omega(tN^3)$  if  $S \leq N/t$ , and satisfies  $TS = \Omega(N^2)$  if  $S > N/t$ . We gave an algorithm matching these bounds up to polylog factors.
- Every  $T$ -query,  $S$ -space 1-sided error quantum algorithm for deciding  $Ax \geq b$  satisfies  $T^2S = \Omega(t^2N^3)$  if  $t \leq S \leq N/t^2$ , and satisfies  $TS = \Omega(N^2)$  if  $S > N/t^2$ . We do not have a matching algorithm in the first case and conjecture that this bound is not tight.

## References

- [Aar02] S. Aaronson. Quantum lower bound for the collision problem. In *Proceedings of 34th ACM STOC*, pages 635–642, 2002. quant-ph/0111102.
- [Aar04] S. Aaronson. Limitations of quantum advice and one-way communication. In *Proceedings of 19th IEEE Conference on Computational Complexity*, pages 320–332, 2004. quant-ph/0402095.
- [Amb00] A. Ambainis. Quantum lower bounds by quantum arguments. In *Proceedings of 32nd ACM STOC*, pages 636–643, 2000. quant-ph/0002066.
- [Amb03] A. Ambainis. Polynomial degree vs quantum query complexity. In *Proceedings of 44th IEEE FOCS*, pages 230–239, 2003. quant-ph/0305028.
- [Amb05] A. Ambainis. A new quantum lower bound method, with an application to strong direct product theorem for quantum search. quant-ph/0508200, 26 Aug 2005.
- [BBC<sup>+</sup>01] R. Beals, H. Buhrman, R. Cleve, M. Mosca, and R. de Wolf. Quantum lower bounds by polynomials. *Journal of the ACM*, 48(4):778–797, 2001. Earlier version in FOCS’98. quant-ph/9802049.
- [BBHT98] M. Boyer, G. Brassard, P. Høyer, and A. Tapp. Tight bounds on quantum searching. *Fortschritte der Physik*, 46(4–5):493–505, 1998. Earlier version in Physcomp’96. quant-ph/9605034.
- [BCWZ99] H. Buhrman, R. Cleve, R. de Wolf, and Ch. Zalka. Bounds for small-error and zero-error quantum algorithms. In *Proceedings of 40th IEEE FOCS*, pages 358–368, 1999. cs.CC/9904019.
- [BHMT02] G. Brassard, P. Høyer, M. Mosca, and A. Tapp. Quantum amplitude amplification and estimation. In *Quantum Computation and Quantum Information: A Millennium Volume*, volume 305 of *AMS Contemporary Mathematics Series*, pages 53–74. 2002. quant-ph/0005055.
- [BV97] E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM Journal on Computing*, 26(5):1411–1473, 1997. Earlier version in STOC’93.
- [BW02] H. Buhrman and R. de Wolf. Complexity measures and decision tree complexity: A survey. *Theoretical Computer Science*, 288(1):21–43, 2002.
- [CR92] D. Coppersmith and T. J. Rivlin. The growth of polynomials bounded at equally spaced points. *SIAM Journal on Mathematical Analysis*, 23(4):970–983, 1992.
- [FGGS98] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. A limit on the speed of quantum computation in determining parity. *Physical Review Letters*, 81:5442–5444, 1998. quant-ph/9802045.
- [GKP94] R. L. Graham, D. E. Knuth, and O. Patashnik. *Concrete Mathematics: A Foundation for Computer Science*. Addison-Wesley, second edition, 1994.

- [Gro96] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of 28th ACM STOC*, pages 212–219, 1996. quant-ph/9605043.
- [HMW03] P. Høyer, M. Mosca, and R. de Wolf. Quantum search on bounded-error inputs. In *Proceedings of 30th International Colloquium on Automata, Languages and Programming (ICALP'03)*, volume 2719 of *Lecture Notes in Computer Science*, pages 291–299. Springer, 2003. quant-ph/0304052.
- [KŠW04] H. Klauck, R. Špalek, and R. de Wolf. Quantum and classical strong direct product theorems and optimal time-space tradeoffs. In *Proceedings of 45th IEEE FOCS*, pages 12–21, 2004. quant-ph/0402123.
- [NC00] M. A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
- [NS94] N. Nisan and M. Szegedy. On the degree of Boolean functions as real polynomials. *Computational Complexity*, 4(4):301–313, 1994. Earlier version in STOC'92.
- [Pat92] R. Paturi. On the degree of polynomials that approximate symmetric Boolean functions. In *Proceedings of 24th ACM STOC*, pages 468–474, 1992.
- [Raz03] A. Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya of the Russian Academy of Science, mathematics*, 67(1):159–176, 2003. quant-ph/0204025.
- [Riv90] T. J. Rivlin. *Chebyshev Polynomials: From Approximation Theory to Algebra and Number Theory*. Wiley-Interscience, second edition, 1990.
- [Sha01] R. Shaltiel. Towards proving strong direct product theorems. In *Proceedings of 16th IEEE Conference on Computational Complexity*, pages 107–119, 2001.
- [Shi02] Y. Shi. Quantum lower bounds for the collision and the element distinctness problems. In *Proceedings of 43rd IEEE FOCS*, 2002. quant-ph/0112086.
- [ŠS05] R. Špalek and M. Szegedy. All quantum adversary methods are equivalent. In *Proceedings of 32nd ICALP*, volume 3580 of *Lecture Notes in Computer Science*, pages 1299–1311, 2005. quant-ph/0409116.

## A Proofs from Section 3

### A.1 Proof of Lemma 2

The measurement of  $\mathcal{H}_A$  decomposes the state in the  $\mathcal{H}_I$  register as follows:

$$\rho = \sum_{a_1, \dots, a_k \in \{0,1\}} p_{a_1, \dots, a_k} \sigma_{a_1, \dots, a_k},$$

with  $p_{a_1, \dots, a_k}$  being the probability of the measurement giving the answer  $(a_1, \dots, a_k)$  (where  $a_j = 1$  means the algorithm outputs—not necessarily correctly—that  $|x^j| = t$  and  $a_j = 0$  means  $|x^j| = t-1$ ) and  $\sigma_{a_1, \dots, a_k}$  being the density matrix of  $\mathcal{H}_I$ , conditional on this outcome of the measurement. Since

the support of  $\rho$  is contained in  $S_{0-} \oplus \dots \oplus S_{m-}$ , the support of states  $\sigma_{a_1, \dots, a_k}$  is also contained in  $S_{0-} \oplus \dots \oplus S_{m-}$ . The probability that the answer  $(a_1, \dots, a_k)$  is correct is equal to

$$\text{Tr} \Pi_{\bigotimes_{j=1}^k \bigoplus_{l=0}^{t-1+a_j} S_{l, a_j}} \sigma_{a_1, \dots, a_k}. \quad (2)$$

We show that, for any  $\sigma_{a_1, \dots, a_k}$  with support contained in  $S_{0-} \oplus \dots \oplus S_{m-}$ , (2) is at most  $\frac{\sum_{m'=0}^m \binom{k}{m'}}{2^k}$ .

For brevity, we now write  $\sigma$  instead of  $\sigma_{a_1, \dots, a_k}$ . A measurement w.r.t.  $\bigotimes_{j=1}^k \bigoplus_l S_{l, a_j}$  and its orthogonal complement commutes with a measurement w.r.t. the collection of subspaces

$$\bigotimes_{j=1}^k (S_{l_j, 0} \oplus S_{l_j, 1}),$$

where  $l_1, \dots, l_k$  range over  $\{0, \dots, t\}$ . Therefore,

$$\text{Tr} \Pi_{\bigotimes_{j=1}^k \bigoplus_l S_{l, a_j}} \sigma = \sum_{l_1, \dots, l_k} \text{Tr} \Pi_{\bigotimes_{j=1}^k \bigoplus_l S_{l, a_j}} \Pi_{\bigotimes_{j=1}^k (S_{l_j, 0} \oplus S_{l_j, 1})} \sigma.$$

This means that, to bound (2), it suffices to prove the same bound with

$$\sigma' = \Pi_{\bigotimes_{j=1}^k (S_{l_j, 0} \oplus S_{l_j, 1})} \sigma.$$

instead of  $\sigma$ . Since

$$\left( \bigotimes_{j=1}^k (S_{l_j, 0} \oplus S_{l_j, 1}) \right) \cap \left( \bigotimes_{j=1}^k (\bigoplus_l S_{l, a_j}) \right) = \bigotimes_{j=1}^k S_{l_j, a_j},$$

we have

$$\text{Tr} \Pi_{\bigotimes_{j=1}^k (\bigoplus_l S_{l, a_j})} \sigma' = \text{Tr} \Pi_{\bigotimes_{j=1}^k S_{l_j, a_j}} \sigma'. \quad (3)$$

We prove this bound for the case when  $\sigma'$  is a pure state:  $\sigma' = |\psi\rangle\langle\psi|$ . Then, equation (3) is equal to

$$\|\Pi_{\bigotimes_{j=1}^k S_{l_j, a_j}} \psi\|^2. \quad (4)$$

The bound for mixed states  $\sigma'$  follows by decomposing  $\sigma'$  as a mixture of pure states  $|\psi\rangle$ , bounding (4) for each of those states and then summing up the bounds.

We have

$$(S_{0-} \oplus \dots \oplus S_{m-}) \cap \left( \bigotimes_{j=1}^k (S_{l_j, 0} \oplus S_{l_j, 1}) \right) = \bigoplus_{\substack{r_1, \dots, r_k \in \{+, -\}, \\ |\{i: r_i = -\}| \leq m}} \bigotimes_{j=1}^k S_{l_j, r_j}.$$

We express

$$|\psi\rangle = \sum_{\substack{r_1, \dots, r_k \in \{+, -\}, \\ |\{i: r_i = -\}| \leq m}} \alpha_{r_1, \dots, r_k} |\psi_{r_1, \dots, r_k}\rangle,$$

with  $|\psi_{r_1, \dots, r_k}\rangle \in \bigotimes_{j=1}^k S_{l_j, r_j}$ . Therefore,

$$\begin{aligned} \|\Pi_{\bigotimes_{j=1}^k S_{l_j, a_j}} \psi\|^2 &\leq \left( \sum_{r_1, \dots, r_k} |\alpha_{r_1, \dots, r_k}| \cdot \|\Pi_{\bigotimes_{j=1}^k S_{l_j, a_j}} \psi_{r_1, \dots, r_k}\| \right)^2 \\ &\leq \sum_{r_1, \dots, r_k} \|\Pi_{\bigotimes_{j=1}^k S_{l_j, a_j}} \psi_{r_1, \dots, r_k}\|^2, \end{aligned} \quad (5)$$

where the second inequality follows from Cauchy-Schwarz and

$$\|\psi\|^2 = \sum_{r_1, \dots, r_k} |\alpha_{r_1, \dots, r_k}|^2 = 1.$$

We have

**Claim 9**

$$\|\Pi_{\otimes_{j=1}^k S_{l_j, a_j}} \psi_{r_1, \dots, r_k}\|^2 \leq \frac{1}{2^k}.$$

**Proof.** Let  $|\varphi_i^{j,0}\rangle$ ,  $i \in [\dim S_{l_j,0}]$  be a basis for the subspace  $S_{l_j,0}$ . Define a map  $U_j : S_{l_j,0} \rightarrow S_{l_j,1}$  by  $U_j|\tilde{\psi}_{i_1, \dots, i_{l_j}}^0\rangle = |\tilde{\psi}_{i_1, \dots, i_{l_j}}^1\rangle$ . Then  $U_j$  is a multiple of a unitary transformation:  $U_j = c_j U'_j$  for some unitary  $U'_j$  and a constant  $c_j$ . (This follows from Claim 11 in Section A.3.)

Let  $|\varphi_i^{j,1}\rangle = U'_j|\varphi_i^{j,0}\rangle$ . Since  $U'_j$  is a unitary transformation,  $|\varphi_i^{j,1}\rangle$  is a basis for  $S_{l_j,1}$ . Therefore,

$$\bigotimes_{i=1}^k |\varphi_{i_j}^{j, a_j}\rangle \tag{6}$$

is a basis for  $\otimes_{j=1}^k S_{l_j, a_j}$ . Moreover, the states

$$|\varphi_i^{j,+}\rangle = \frac{1}{\sqrt{2}}|\varphi_i^{j,0}\rangle + \frac{1}{\sqrt{2}}|\varphi_i^{j,1}\rangle, \quad |\varphi_i^{j,-}\rangle = \frac{1}{\sqrt{2}}|\varphi_i^{j,0}\rangle - \frac{1}{\sqrt{2}}|\varphi_i^{j,1}\rangle$$

are a basis for  $S_{l_j,+}$  and  $S_{l_j,-}$ , respectively. Therefore,

$$|\psi_{r_1, \dots, r_k}\rangle = \sum_{i_1, \dots, i_k} \alpha_{i_1, \dots, i_k} \bigotimes_{j=1}^k |\varphi_{i_j}^{j, r_j}\rangle. \tag{7}$$

The inner product between  $\bigotimes_{i=1}^k |\varphi_{i'_j}^{j, a_j}\rangle$  and  $\bigotimes_{j=1}^k |\varphi_{i_j}^{j, r_j}\rangle$  is

$$\prod_{j=1}^k \langle \varphi_{i_j}^{j, r_j} | \varphi_{i'_j}^{j, a_j} \rangle.$$

Note that  $r_j \in \{+, -\}$  and  $a_j \in \{0, 1\}$ . The terms in this product are  $\pm \frac{1}{\sqrt{2}}$  if  $i'_j = i_j$  and 0 otherwise.

This means that  $\bigotimes_{j=1}^k |\varphi_{i_j}^{j, r_j}\rangle$  has inner product  $\pm \frac{1}{2^{k/2}}$  with  $\bigotimes_{i=1}^k |\varphi_{i_j}^{j, a_j}\rangle$  and inner product 0 with all other basis states (6). Therefore,

$$\Pi_{\otimes_{j=1}^k S_{l_j, a_j}} \bigotimes_{j=1}^k |\varphi_{i_j}^{j, r_j}\rangle = \pm \frac{1}{2^{k/2}} \bigotimes_{i=1}^k |\varphi_{i_j}^{j, a_j}\rangle.$$

Together with equation (7), this means that

$$\|\Pi_{\otimes_{j=1}^k S_{l_j, a_j}} \psi_{r_1, \dots, r_k}\| \leq \frac{1}{2^{k/2}} \|\psi_{r_1, \dots, r_k}\| = \frac{1}{2^{k/2}}.$$

Squaring both sides completes the proof of the claim.  $\square$

Since there are  $\binom{k}{m'}$  tuples  $(r_1, \dots, r_k)$  with  $r_1, \dots, r_k \in \{+, -\}$  and  $|\{i : r_i = -\}| = m'$ , Claim 9 together with equation (5) implies

$$\|\Pi_{\otimes_{j=1}^k S_{l_j, a_j}} \psi\|^2 \leq \frac{\sum_{m'=0}^m \binom{k}{m'}}{2^k}.$$

## A.2 Proof of Corollary 3

Let  $|\psi\rangle$  be a purification of  $\rho$  in  $\mathcal{H}_A \otimes \mathcal{H}_I$ . Let

$$|\psi\rangle = \sqrt{1-\delta}|\psi'\rangle + \sqrt{\delta}|\psi''\rangle$$

where  $|\psi'\rangle$  is in the subspace  $\mathcal{H}_A \otimes (S_{0-} \oplus S_{1-} \oplus \dots \oplus S_{m-})$  and  $|\psi''\rangle$  is in the subspace  $\mathcal{H}_A \otimes (S_{0-} \oplus S_{1-} \oplus \dots \oplus S_{m-})^\perp$ . Then,  $\delta = \text{Tr } \Pi_{(S_{0-} \oplus \dots \oplus S_{m-})^\perp} \rho$ .

The success probability of  $\mathcal{A}$  is the probability that, if we measure both the register of  $\mathcal{H}_A$  containing the result of the computation and  $\mathcal{H}_I$ , then, we get  $a_1, \dots, a_k$  and  $x^1, \dots, x^k$  such that, for every  $j \in \{1, \dots, k\}$ ,  $x^j$  contains  $t-1+a_j$  ones.

Consider the probability of getting  $a_1, \dots, a_k \in \{0, 1\}$  and  $x^1, \dots, x^k \in \{0, 1\}^n$  with this property, when measuring  $|\psi'\rangle$  (instead of  $|\psi\rangle$ ). By Lemma 2, this probability is at most  $\frac{\sum_{m'=0}^m \binom{k}{m'}}{2^k}$ . We have

$$\|\psi - \psi'\| \leq (1 - \sqrt{1-\delta})\|\psi'\| + \sqrt{\delta}\|\psi''\| = (1 - \sqrt{1-\delta}) + \sqrt{\delta} \leq 2\sqrt{\delta}.$$

We now apply

**Lemma 10 ([BV97])** *For any states  $|\psi\rangle$  and  $|\psi'\rangle$  and any measurement  $M$ , the variational distance between the probability distributions obtained by applying  $M$  to  $|\psi\rangle$  and  $|\psi'\rangle$  is at most  $2\|\psi - \psi'\|$ .*

Lemma 10 implies that the success probability of  $\mathcal{A}$  is at most

$$\frac{\sum_{m'=0}^m \binom{k}{m'}}{2^k} + 4\sqrt{\delta} = \frac{\sum_{m'=0}^m \binom{k}{m'}}{2^k} + 4\sqrt{\text{Tr } \Pi_{(S_{0-} \oplus \dots \oplus S_{m-})^\perp} \rho}.$$

## A.3 Proof of Lemma 4

Let  $|\psi_d\rangle$  be the state of  $\mathcal{H}_A \otimes \mathcal{H}_I$  after  $d$  queries. We decompose

$$|\psi_d\rangle = \sum_{i=0}^{kn} a_i |\psi_{d,i}\rangle,$$

with  $|\psi_{d,i}\rangle$  being the part in which the query register contains  $|i\rangle$ . Let  $\rho_{d,i} = |\psi_{d,i}\rangle\langle\psi_{d,i}|$ . Then,

$$\rho_d = \sum_{i=0}^{kn} a_i^2 \rho_{d,i}. \tag{8}$$

Because of

$$\text{Tr } \Pi_{V_m} \rho_d = \sum_{i=0}^{kn} a_i^2 \text{Tr } \Pi_{V_m} \rho_{d,i},$$

we have  $P(\rho_d) = \sum_{i=0}^{kn} a_i^2 P(\rho_{d,i})$ . Let  $\rho'_d$  be the state after the  $d^{\text{th}}$  query and let  $\rho'_d = \sum_{i=0}^{kn} a_i^2 \rho'_{d,i}$  be a decomposition similar to equation (8). Lemma 4 follows by showing

$$P(\rho'_{d,i}) \leq \left(1 + \frac{C}{\sqrt{tn}}(q^{t/2} - 1) + \frac{C\sqrt{t}}{\sqrt{n}}(q - 1)\right) P(\rho_{d,i}) \tag{9}$$

for each  $i$ . For  $i = 0$ , the query does not change the state if the query register contains  $|i\rangle$ . Therefore,  $\rho'_{d,0} = \rho_{d,0}$  and  $P(\rho'_{d,0}) = P(\rho_{d,0})$ . This means that the equation (9) is true for  $i = 0$ . To prove the  $i \in \{1, \dots, kn\}$  case, it suffices to prove the  $i = 1$  case (because of symmetry).

Let  $|\psi_{i_1, \dots, i_j}^{a,b}\rangle$  (where  $a, b \in \{0, 1\}$  and  $i_1, \dots, i_j \in \{2, \dots, n\}$ ) be the uniform superposition over basis states  $|b, x_2, \dots, x_n\rangle$  (of  $\mathcal{H}_{one}$ ) with  $b + x_2 + \dots + x_n = t - 1 + a$  and  $x_{i_1} = \dots = x_{i_j} = 1$ . Let  $T_{j,a,b}$  be the space spanned by all states  $|\psi_{i_1, \dots, i_j}^{a,b}\rangle$  and let  $S_{j,a,b} = T_{j,a,b} \cap T_{j-1,a,b}^\perp$ . Let  $|\tilde{\psi}_{i_1, \dots, i_j}^{a,b}\rangle = \Pi_{T_{j-1,a,b}^\perp} |\psi_{i_1, \dots, i_j}^{a,b}\rangle$ .

**Claim 11** *The maps  $U_{01} : T_{j,0,0} \rightarrow T_{j,0,1}$ ,  $U_{10} : T_{j,0,0} \rightarrow T_{j,1,0}$  and  $U_{11} : T_{j,0,0} \rightarrow T_{j,1,1}$  defined by  $U_{ab}|\tilde{\psi}_{i_1, \dots, i_j}^{0,0}\rangle = |\tilde{\psi}_{i_1, \dots, i_j}^{a,b}\rangle$  are multiples of unitary transformations:  $U_{ab} = c_{ab}U'_{ab}$  for some unitary  $U'_{ab}$  and some constant  $c_{ab}$ .*

**Proof.** We define  $M : T_{j,0,0} \rightarrow T_{j,0,1}$  by

$$M|0x_2 \dots x_n\rangle = \sum_{l:x_l=1} |1x_2 \dots x_{l-1}0x_{l+1} \dots x_n\rangle.$$

As shown in [Amb05, the proof of Claim 10],

$$M|\tilde{\psi}_{i_1, \dots, i_j}^{0,0}\rangle = c|\tilde{\psi}_{i_1, \dots, i_j}^{0,1}\rangle, \quad (10)$$

$$M^\dagger|\tilde{\psi}_{i_1, \dots, i_j}^{0,1}\rangle = c'|\tilde{\psi}_{i_1, \dots, i_j}^{0,0}\rangle,$$

for some constants  $c$  and  $c'$ . Therefore,

$$\langle \tilde{\psi}_{i_1, \dots, i_j}^{0,0} | M^\dagger M | \tilde{\psi}_{i'_1, \dots, i'_j}^{0,0} \rangle = c \cdot c' \langle \tilde{\psi}_{i_1, \dots, i_j}^{0,0} | \tilde{\psi}_{i'_1, \dots, i'_j}^{0,0} \rangle.$$

Therefore,  $M$  is a multiple of a unitary transformation. By equation (10),  $U_{01} = M/c$  and, therefore,  $U_{01}$  is also a multiple of a unitary transformation.

Next, we define  $M$  by  $M|0x_2 \dots x_n\rangle = |1x_2 \dots x_n\rangle$ . Then,  $M$  is a unitary transformation from the space spanned by  $|0x_2 \dots x_n\rangle$ ,  $x_2 + \dots + x_n = t - 1$  to the space spanned by  $|1x_2 \dots x_n\rangle$ ,  $1 + x_2 + \dots + x_n = t$ . We claim that  $U_{11} = M$ . To prove that, we first observe that

$$M|\psi_{i_1, \dots, i_j}^{0,0}\rangle = \frac{1}{\sqrt{\binom{n-j-1}{t-j-1}}} \sum_{\substack{x_2, \dots, x_n: \\ x_{i_1} = \dots = x_{i_j} = 1}} M|0x_2 \dots x_n\rangle = \frac{1}{\sqrt{\binom{n-j-1}{t-j-1}}} \sum_{\substack{x_2, \dots, x_n: \\ x_{i_1} = \dots = x_{i_j} = 1}} |1x_2 \dots x_n\rangle = |\psi_{i_1, \dots, i_j}^{1,1}\rangle.$$

Since  $T_{j,a,b}$  is defined as the subspace spanned by all  $|\psi_{i_1, \dots, i_j}^{a,b}\rangle$ , this means that  $M(T_{j,0,0}) = T_{j,1,1}$  and, similarly,  $M(T_{j-1,0,0}) = T_{j-1,1,1}$ . Since  $M$  is unitary, this implies  $M(T_{j-1,0,0}^\perp) = T_{j-1,1,1}^\perp$  and

$$M|\tilde{\psi}_{i_1, \dots, i_j}^{0,0}\rangle = M\Pi_{(T_{j-1,0,0})^\perp} |\psi_{i_1, \dots, i_j}^{0,0}\rangle = \Pi_{(T_{j-1,1,1})^\perp} |\psi_{i_1, \dots, i_j}^{1,1}\rangle = |\tilde{\psi}_{i_1, \dots, i_j}^{1,1}\rangle.$$

Finally, we have  $U_{10} = U'_{10}U_{11}$ , where  $U'_{10}$  is defined by  $U'_{10}|\tilde{\psi}_{i_1, \dots, i_j}^{1,1}\rangle = |\tilde{\psi}_{i_1, \dots, i_j}^{1,0}\rangle$ . Since  $U_{11}$  is unitary, it suffices to prove that  $U'_{10}$  is a multiple of a unitary transformation and this follows similarly to  $U_{01}$  being a multiple of a unitary transformation.  $\square$

Let  $|\psi_{00}\rangle$  be an arbitrary state in  $T_{j,0,0}$  for some  $j \in \{0, \dots, t-1\}$ . Define  $|\psi_{ab}\rangle = U_{ab}|\psi_{00}\rangle$  for  $ab \in \{01, 10, 11\}$ . Let  $|\psi_2\rangle, \dots, |\psi_k\rangle$  be vectors from subspaces  $S'_{j_2}, \dots, S'_{j_k}$ , for some  $j_2, \dots, j_k$ . We first analyze the case when  $\rho_{d,1}$  belongs to the subspace  $\mathcal{H}_4$  spanned by

$$|\psi_{ab}\rangle \otimes |\psi_2\rangle \otimes \dots \otimes |\psi_k\rangle.$$

We have

**Claim 12** Let  $\alpha'_a = \sqrt{\frac{n-(t-1+a)}{n-j}} \|\tilde{\psi}_{i_1, \dots, i_j}^{a,0}\|$  and  $\beta'_a = \sqrt{\frac{(t-1+a)-j}{n-j}} \|\tilde{\psi}_{i_1, \dots, i_j}^{a,1}\|$ . Let  $\alpha_a = \frac{\alpha'_a}{\sqrt{(\alpha'_a)^2 + (\beta'_a)^2}}$ ,  $\beta_a = \frac{\beta'_a}{\sqrt{(\alpha'_a)^2 + (\beta'_a)^2}}$ . Then,

1.  $|\phi_1\rangle = \alpha_0|\psi_{00}\rangle + \beta_0|\psi_{01}\rangle + \alpha_1|\psi_{10}\rangle + \beta_1|\psi_{11}\rangle$  belongs to  $S_{j,+}$ ;
2.  $|\phi_2\rangle = \beta_0|\psi_{00}\rangle - \alpha_0|\psi_{01}\rangle + \beta_1|\psi_{10}\rangle - \alpha_1|\psi_{11}\rangle$  belongs to  $S_{j+1,+}$ ;
3. Any linear combination of  $|\psi_{00}\rangle, |\psi_{01}\rangle, |\psi_{10}\rangle$  and  $|\psi_{11}\rangle$  which is orthogonal to  $|\phi_1\rangle$  and  $|\phi_2\rangle$  belongs to  $S_- = \bigoplus_{j=0}^t S_{j,-}$ .

**Proof.** Let  $i_1, \dots, i_j$  be  $j$  distinct elements of  $\{2, \dots, n\}$ . As shown in [Amb05, beginning of the proof of Claim 11],

$$|\tilde{\psi}_{i_1, \dots, i_j}^a\rangle = \sqrt{\frac{n-(t-1+a)}{n-j}} |\tilde{\psi}_{i_1, \dots, i_j}^{a,0}\rangle + \sqrt{\frac{(t-1+a)-j}{n-j}} |\tilde{\psi}_{i_1, \dots, i_j}^{a,1}\rangle.$$

Thus,

$$|\tilde{\psi}_{i_1, \dots, i_j}^a\rangle = \alpha'_a \frac{|\tilde{\psi}_{i_1, \dots, i_j}^{a,0}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{a,0}\|} + \beta'_a \frac{|\tilde{\psi}_{i_1, \dots, i_j}^{a,1}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{a,1}\|},$$

This means that  $\|\tilde{\psi}_{i_1, \dots, i_j}^a\| = \sqrt{(\alpha'_a)^2 + (\beta'_a)^2}$  and

$$\frac{|\tilde{\psi}_{i_1, \dots, i_j}^a\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^a\|} = \alpha_a \frac{|\tilde{\psi}_{i_1, \dots, i_j}^{a,0}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{a,0}\|} + \beta_a \frac{|\tilde{\psi}_{i_1, \dots, i_j}^{a,1}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{a,1}\|}.$$

Since the states  $|\tilde{\psi}_{i_1, \dots, i_j}^0\rangle$  span  $S_{j,0}$ , the state  $|\psi_{00}\rangle$  is a linear combination of states  $\frac{|\tilde{\psi}_{i_1, \dots, i_j}^{0,0}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{0,0}\|}$ . By

Claim 11, the states  $|\psi_{ab}\rangle$  are linear combinations of  $\frac{|\tilde{\psi}_{i_1, \dots, i_j}^{a,b}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{a,b}\|}$  with the same coefficients. Therefore,  $|\phi_1\rangle$  is a linear combination of

$$\alpha_0 \frac{|\tilde{\psi}_{i_1, \dots, i_j}^{0,0}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{0,0}\|} + \beta_0 \frac{|\tilde{\psi}_{i_1, \dots, i_j}^{0,1}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{0,1}\|} + \alpha_1 \frac{|\tilde{\psi}_{i_1, \dots, i_j}^{1,0}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{1,0}\|} + \beta_1 \frac{|\tilde{\psi}_{i_1, \dots, i_j}^{1,1}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{1,1}\|} = \frac{|\tilde{\psi}_{i_1, \dots, i_j}^0\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^0\|} + \frac{|\tilde{\psi}_{i_1, \dots, i_j}^1\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^1\|},$$

each of which, by definition, belongs to  $S_{j,+}$ .

Let  $i_1, \dots, i_j$  be distinct elements of  $\{2, \dots, n\}$ . We claim that

$$\frac{|\tilde{\psi}_{1, i_1, \dots, i_j}^a\rangle}{\|\tilde{\psi}_{1, i_1, \dots, i_j}^a\|} = \beta_a \frac{|\tilde{\psi}_{i_1, \dots, i_j}^{a,0}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{a,0}\|} - \alpha_a \frac{|\tilde{\psi}_{i_1, \dots, i_j}^{a,1}\rangle}{\|\tilde{\psi}_{i_1, \dots, i_j}^{a,1}\|}. \quad (11)$$

As shown in [Amb05, Claim 11], the right hand side of (11) belongs to  $S_{j+1,a}$ . We need to show that it is equal to  $|\tilde{\psi}_{1,i_1,\dots,i_j}^a\rangle$ . We have

$$\begin{aligned} |\tilde{\psi}_{1,i_1,\dots,i_j}^a\rangle &= \Pi_{T_{j,a}^\perp} |\psi_{1,i_1,\dots,i_j}^a\rangle = \Pi_{T_{j,a}^\perp} |\psi_{i_1,\dots,i_j}^{a,1}\rangle \\ &= \Pi_{T_{j,a}^\perp} \Pi_{T_{j-1,a,1}^\perp} |\psi_{i_1,\dots,i_j}^{a,1}\rangle = \Pi_{T_{j,a}^\perp} |\tilde{\psi}_{i_1,\dots,i_j}^{a,1}\rangle, \end{aligned}$$

where the third equality follows from  $T_{j-1,a,1} \subseteq T_{j,a}$  (which is shown in [Amb05, equation 15]). We express

$$|\tilde{\psi}_{i_1,\dots,i_j}^{a,1}\rangle = c_1|\varphi_1\rangle + c_2|\varphi_2\rangle$$

where

$$|\varphi_1\rangle = \alpha_a \frac{|\tilde{\psi}_{i_1,\dots,i_j}^{a,0}\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^{a,0}\|} + \beta_a \frac{|\tilde{\psi}_{i_1,\dots,i_j}^{a,1}\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^{a,1}\|}, \quad |\varphi_2\rangle = \beta_a \frac{|\tilde{\psi}_{i_1,\dots,i_j}^{a,0}\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^{a,0}\|} - \alpha_a \frac{|\tilde{\psi}_{i_1,\dots,i_j}^{a,1}\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^{a,1}\|}.$$

We have  $|\varphi_1\rangle \in S_{j,a} \subseteq T_{j,a}$ ,  $|\varphi_2\rangle \in S_{j+1,a} \subseteq T_{j,a}^\perp$ . Therefore,  $\Pi_{T_{j,a}^\perp} |\tilde{\psi}_{i_1,\dots,i_j}^{a,1}\rangle = c_2|\varphi_2\rangle$  and

$$\frac{|\tilde{\psi}_{1,i_1,\dots,i_j}^a\rangle}{\|\tilde{\psi}_{1,i_1,\dots,i_j}^a\|} = |\varphi_2\rangle = \beta_a \frac{|\tilde{\psi}_{i_1,\dots,i_j}^{a,0}\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^{a,0}\|} - \alpha_a \frac{|\tilde{\psi}_{i_1,\dots,i_j}^{a,1}\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^{a,1}\|},$$

proving (11).

Similarly to the argument for  $|\phi_1\rangle$ , equation (11) implies that  $|\phi_2\rangle$  is a linear combination of

$$\beta_0 \frac{|\tilde{\psi}_{i_1,\dots,i_j}^{0,0}\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^{0,0}\|} - \alpha_0 \frac{|\tilde{\psi}_{i_1,\dots,i_j}^{0,1}\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^{0,1}\|} + \beta_1 \frac{|\tilde{\psi}_{i_1,\dots,i_j}^{1,0}\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^{1,0}\|} - \alpha_1 \frac{|\tilde{\psi}_{i_1,\dots,i_j}^{1,1}\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^{1,1}\|} = \frac{|\tilde{\psi}_{1,i_1,\dots,i_j}^0\rangle}{\|\tilde{\psi}_{1,i_1,\dots,i_j}^0\|} + \frac{|\tilde{\psi}_{1,i_1,\dots,i_j}^1\rangle}{\|\tilde{\psi}_{1,i_1,\dots,i_j}^1\|}$$

and each of those states belongs to  $S_{j+1,+}$ .

To prove the third part of Claim 12, we observe that any vector orthogonal to  $|\phi_1\rangle$  and  $|\phi_2\rangle$  is a linear combination of

$$\begin{aligned} |\phi_3\rangle &= \alpha_0|\psi_{00}\rangle + \beta_0|\psi_{01}\rangle - \alpha_1|\psi_{10}\rangle - \beta_1|\psi_{11}\rangle = \frac{|\tilde{\psi}_{i_1,\dots,i_j}^0\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^0\|} - \frac{|\tilde{\psi}_{i_1,\dots,i_j}^1\rangle}{\|\tilde{\psi}_{i_1,\dots,i_j}^1\|}, \\ |\phi_4\rangle &= \beta_0|\psi_{00}\rangle - \alpha_0|\psi_{01}\rangle - \beta_1|\psi_{10}\rangle + \alpha_1|\psi_{11}\rangle = \frac{|\tilde{\psi}_{1,i_1,\dots,i_j}^0\rangle}{\|\tilde{\psi}_{1,i_1,\dots,i_j}^0\|} - \frac{|\tilde{\psi}_{1,i_1,\dots,i_j}^1\rangle}{\|\tilde{\psi}_{1,i_1,\dots,i_j}^1\|}. \end{aligned}$$

We have  $|\phi_3\rangle \in S_{j,-}$  and  $|\phi_4\rangle \in S_{j+1,-}$ . □

**Claim 13** ([Amb05, Claim 14])  $\beta_0 \leq \sqrt{\frac{4(t-1-j)}{n+3(t-1)-4j}}$  and  $\beta_1 \leq \sqrt{\frac{4(t-j)}{n+3t-4j}}$ .

**Claim 14**

$$|\alpha_0\beta_1 - \alpha_1\beta_0| = O\left(\frac{1}{\sqrt{tn}}\right).$$

**Proof.** We first estimate

$$\frac{\alpha'_0 \beta'_1}{\alpha'_1 \beta'_0} = \frac{\sqrt{(n-t+1)(t-j)}}{\sqrt{(n-t)(t-1-j)}} \cdot \frac{\|\tilde{\psi}_{i_1, \dots, i_j}^{0,0}\| \|\tilde{\psi}_{i_1, \dots, i_j}^{1,1}\|}{\|\tilde{\psi}_{i_1, \dots, i_j}^{0,1}\| \|\tilde{\psi}_{i_1, \dots, i_j}^{1,0}\|}.$$

We start by showing that

$$\frac{\|\tilde{\psi}_{i_1, \dots, i_j}^{0,0}\| \|\tilde{\psi}_{i_1, \dots, i_j}^{1,1}\|}{\|\tilde{\psi}_{i_1, \dots, i_j}^{0,1}\| \|\tilde{\psi}_{i_1, \dots, i_j}^{1,0}\|} = 1 \pm O\left(\frac{1}{t}\right). \quad (12)$$

From [Amb05, equation 22], we have  $\|\tilde{\psi}_{i_1, \dots, i_j}^{a,b}\| = \frac{1}{\sqrt{A_{a,b}}}$  where

$$A_{a,b} = \sum_{m=0}^j \binom{j}{m} \frac{(t-m+a-b-1)^{j-m}}{(n-t-a+b)^{j-m}},$$

and  $x^{\underline{j}} = x(x-1)\dots(x-j+1)$ . To prove (12), it suffices to show  $\frac{A_{0,0}A_{1,1}}{A_{1,0}A_{0,1}} = 1 \pm O(\frac{1}{t})$ .

We compute the combinatorial sum  $A_{a,b}$  using hypergeometric series [GKP94, Section 5.5]. Let

$$C_{a,b}(m) = \binom{j}{m} \frac{(t-m+a-b-1)^{j-m}}{(n-t-a+b)^{j-m}}.$$

Then  $A_{a,b} = \sum_{m=0}^{\infty} C_{a,b}(m)$ , because  $\binom{j}{m} = 0$  for  $m > j$ . Since

$$\frac{C_{a,b}(m+1)}{C_{a,b}(m)} = \frac{(m-j)(m+n-t-j-a+b+1)}{(m+1)(m-t-a+b+1)}$$

is a rational function of  $m$ ,  $A_{a,b}$  is a hypergeometric series and its value is

$$A_{a,b} = \sum_{m=0}^{\infty} C_{a,b}(m) = C_{a,b}(0) \cdot F\left(\begin{matrix} -j, & n-t-j-a+b+1 \\ & -t-a+b+1 \end{matrix} \middle| 1\right).$$

We apply Vandermonde's convolution  $F\left(\begin{matrix} -j, & x \\ & y \end{matrix} \middle| 1\right) = (x-y)^{\underline{j}}/(-y)^{\underline{j}}$  [GKP94, Equation 5.93 on page 212], which holds for every integer  $j \geq 0$ , and obtain

$$A_{a,b} = \frac{(t+a-b-1)^{\underline{j}}}{(n-t-a+b)^{\underline{j}}} \cdot \frac{(n-j)^{\underline{j}}}{(t+a-b-1)^{\underline{j}}} = \frac{(n-j)^{\underline{j}}}{(n-t-a+b)^{\underline{j}}}.$$

It follows that

$$\frac{A_{0,0}A_{1,1}}{A_{1,0}A_{0,1}} = \frac{(n-t+1)^{\underline{j}}(n-t-1)^{\underline{j}}}{((n-t)^{\underline{j}})^2} = \frac{(n-t+1)(n-t-j)}{(n-t)(n-t-j+1)} = 1 - \frac{j}{(n-t)(n-t-j+1)},$$

which is  $1 - \Theta(j/n^2) = 1 - O(1/t)$  for  $t < n/2$  and  $j < t/2$ .

Therefore, we have

$$\frac{\alpha'_0 \beta'_1}{\alpha'_1 \beta'_0} = \left(1 \pm O\left(\frac{1}{t}\right)\right) \frac{\sqrt{(n-t+1)(t-j)}}{\sqrt{(n-t)(t-1-j)}}.$$

Since  $\frac{\sqrt{t-j}}{\sqrt{t-1-j}} = \sqrt{1 + \frac{1}{t-1-j}} = 1 + O(\frac{1}{t-1-j}) = 1 + O(\frac{1}{t})$  and, similarly,  $\frac{\sqrt{n-t+1}}{\sqrt{n-t}} = 1 + O(\frac{1}{n-t}) = 1 + O(\frac{1}{t})$ , we have shown that  $\frac{\alpha'_0 \beta'_1}{\alpha'_1 \beta'_0}$  is of order  $1 \pm O(\frac{1}{t})$ .

We thus have

$$|\alpha_0 \beta_1 - \beta_0 \alpha_1| = O\left(\frac{1}{t}\right) |\beta_0 \alpha_1| = O\left(\frac{1}{t}\right) \frac{\beta'_0 \alpha'_1}{\sqrt{((\alpha'_0)^2 + (\beta'_0)^2)((\alpha'_1)^2 + (\beta'_1)^2)}}. \quad (13)$$

For the numerator, we have

$$\beta'_0 \alpha'_1 = \frac{\sqrt{t-1-j}}{\sqrt{n-j}} \|\tilde{\psi}_{i_1, \dots, i_j}^{0,1}\| \frac{\sqrt{n-t}}{\sqrt{n-j}} \|\tilde{\psi}_{i_1, \dots, i_j}^{1,0}\|.$$

Since  $\|\tilde{\psi}_{i_1, \dots, i_j}^{0,1}\|$  and  $\|\tilde{\psi}_{i_1, \dots, i_j}^{1,0}\|$  are bounded from above by 1, we have

$$\beta'_0 \alpha'_1 = O\left(\frac{\sqrt{t-1-j}}{\sqrt{n-j}} \frac{\sqrt{n-t}}{\sqrt{n-j}}\right) = O\left(\frac{\sqrt{t-1-j}}{\sqrt{n-j}}\right) = O\left(\frac{\sqrt{t}}{\sqrt{n}}\right).$$

To bound the denominator of (13), we recall that  $\|\psi_{i_1, \dots, i_j}^{a,b}\|$  are all bounded from below by  $1/\sqrt{e}$ . Thus,

$$\sqrt{(\alpha'_0)^2 + (\beta'_0)^2} \geq \sqrt{\frac{1}{\sqrt{e}} \left( \frac{n-t+1}{n-j} + \frac{t-1-j}{n-j} \right)} = \frac{1}{\sqrt[4]{e}}.$$

This means that

$$|\alpha_0 \beta_1 - \beta_0 \alpha_1| = O\left(\frac{1}{t}\right) O\left(\frac{\sqrt{t}}{\sqrt{n}}\right) = O\left(\frac{1}{\sqrt{tn}}\right).$$

□

We pick an orthonormal basis for  $\mathcal{H}_4$  that has  $|\phi_1\rangle$  and  $|\phi_2\rangle$  as its first two vectors. Let  $|\phi_3\rangle$  and  $|\phi_4\rangle$  be the other two basis vectors. We define

$$|\chi_i\rangle = |\phi_i\rangle \otimes |\psi_2\rangle \otimes \dots \otimes |\psi_k\rangle. \quad (14)$$

By Claim 12,  $|\chi_1\rangle$  belongs to  $S_{j,+} \otimes S'_{j_2} \otimes \dots \otimes S'_{j_k}$  which is contained in  $V_{\min(j, t/2) + j_2 + \dots + j_k}$ . Similarly,  $|\chi_2\rangle$  belongs to  $V_{\min(j+1, t/2) + j_2 + \dots + j_k}$  and  $|\chi_3\rangle$  and  $|\chi_4\rangle$  belong to  $V_{t/2 + j_2 + \dots + j_k}$ . If  $j < t/2$ , this means that

$$P(\rho_{d,1}) = q^{j_2 + \dots + j_k} \cdot \left( q^j \langle \chi_1 | \rho_{d,1} | \chi_1 \rangle + q^{j+1} \langle \chi_2 | \rho_{d,1} | \chi_2 \rangle + q^{\frac{t}{2}} \langle \chi_3 | \rho_{d,1} | \chi_3 \rangle + q^{\frac{t}{2}} \langle \chi_4 | \rho_{d,1} | \chi_4 \rangle \right) \quad (15)$$

(If  $j \geq t/2$ , then  $|\chi_1\rangle, |\chi_2\rangle, |\chi_3\rangle, |\chi_4\rangle$  all belong to  $V_{t/2 + j_2 + \dots + j_k}$ . This means that  $P(\rho_{d,1}) = q^{t/2 + j_2 + \dots + j_k}$  and it remains unchanged by a query.)

We define  $\gamma_j = \langle \chi_j | \rho_{d,1} | \chi_j \rangle$ . Since the support of  $\rho_{d,1}$  is contained in the subspace spanned by  $|\chi_j\rangle$ , we have  $\text{Tr } \rho_{d,1} = \gamma_1 + \gamma_2 + \gamma_3 + \gamma_4$  and, therefore  $\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4 = 1$ . This means that equation (15) can be rewritten as

$$\begin{aligned} P(\rho_{d,1}) &= q^{j+j_2+\dots+j_k} \gamma_1 + q^{j+j_2+\dots+j_k+1} \gamma_2 + q^{t/2+j_2+\dots+j_k} (\gamma_3 + \gamma_4) \\ &= q^{t/2+j_2+\dots+j_k} + q^{j_2+\dots+j_k} (q^{j+1} - q^{t/2}) (\gamma_1 + \gamma_2) + q^{j_2+\dots+j_k} (q^j - q^{j+1}) \gamma_1 \end{aligned} \quad (16)$$

$P(\rho'_{d,1})$  can be also expressed in a similar way, with  $\gamma'_j = \langle \chi_j | \rho'_{d,1} | \chi_j \rangle$  instead of  $\gamma_j$ . By combining equations (16) for  $P(\rho_{d,1})$  and  $P(\rho'_{d,1})$ , we get

$$P(\rho'_{d,1}) - P(\rho_{d,1}) = q^{j+j_2+\dots+j_k+1}(q^{t/2-j-1} - 1)(\gamma_1 + \gamma_2 - \gamma'_1 - \gamma'_2) + q^{j+j_2+\dots+j_k}(q - 1)(\gamma_1 - \gamma'_1).$$

Therefore, it suffices to bound  $|\gamma_1 + \gamma_2 - \gamma'_1 - \gamma'_2|$  and  $|\gamma_1 - \gamma'_1|$ . W.l.o.g., we can assume that  $\rho_{d,1}$  is a pure state  $|\varphi\rangle\langle\varphi|$ . Let

$$|\varphi\rangle = (a|\psi_{00}\rangle + b|\psi_{01}\rangle + c|\psi_{10}\rangle + d|\psi_{11}\rangle) \otimes |\psi_2\rangle \otimes \dots \otimes |\psi_k\rangle.$$

Then, the state after a query is

$$|\varphi'\rangle = (a|\psi_{00}\rangle - b|\psi_{01}\rangle + c|\psi_{10}\rangle - d|\psi_{11}\rangle) \otimes |\psi_2\rangle \otimes \dots \otimes |\psi_k\rangle$$

and we have to bound

$$\gamma_j - \gamma'_j = |\langle \chi_j | \varphi \rangle|^2 - |\langle \chi_j | \varphi' \rangle|^2$$

for  $j \in \{1, 2\}$ . For  $j = 1$ , we have

$$\langle \chi_1 | \varphi \rangle = \frac{1}{\sqrt{2}}(a\alpha_0 + b\beta_0 + c\alpha_1 + d\beta_1).$$

The expression for  $\varphi'$  is similar, with minus signs in front of  $b\beta_0$  and  $d\beta_1$ . Therefore,

$$||\langle \chi_1 | \varphi \rangle|^2 - |\langle \chi_1 | \varphi' \rangle|^2| \leq 2|a||b|\alpha_0\beta_0 + 2|c||d|\alpha_1\beta_1 + 2|a||d|\alpha_0\beta_1 + 2|b||c|\alpha_1\beta_0. \quad (17)$$

Since  $|a|, |b|, |c|, |d|$  are all at most  $\|\varphi\| = 1$  and  $\alpha_0, \alpha_1$  are less than 1, equation (17) is at most  $4\beta_0 + 4\beta_1$ . By Claim 13, we have

$$4\beta_0 + 4\beta_1 \leq 4\sqrt{\frac{4(t-1-j)}{n+3(t-1)-4j}} + 4\sqrt{\frac{4(t-j)}{n+3t-4j}} \leq 16\sqrt{\frac{t}{n}}.$$

We also have

$$\begin{aligned} |\gamma_1 + \gamma_2 - \gamma'_1 - \gamma'_2| &= ||\langle \chi_1 | \varphi \rangle|^2 + |\langle \chi_2 | \varphi \rangle|^2 - |\langle \chi_1 | \varphi' \rangle|^2 - |\langle \chi_2 | \varphi' \rangle|^2| \\ &\leq |a||d||\alpha_0\beta_1 - \alpha_1\beta_0| + |b||c||\alpha_1\beta_0 - \alpha_0\beta_1| \leq 2|\alpha_0\beta_1 - \alpha_1\beta_0| \leq \frac{2C}{\sqrt{tn}} \end{aligned} \quad (18)$$

where  $C$  is the big-O constant from Claim 14. By taking into account that  $P(\rho_{d,1}) \geq q^{j+j_2+\dots+j_k}$ ,

$$\begin{aligned} P(|\varphi'\rangle\langle\varphi'|) - P(|\varphi\rangle\langle\varphi|) &\leq \left( (q-1)\frac{16\sqrt{t}}{\sqrt{n}} + (q^{t/2-j} - q)\frac{2C}{\sqrt{tn}} \right) P(|\varphi\rangle\langle\varphi|) \\ &\leq \left( (q-1)\frac{16\sqrt{t}}{\sqrt{n}} + (q^{t/2} - 1)\frac{2C}{\sqrt{tn}} \right) P(|\varphi\rangle\langle\varphi|). \end{aligned} \quad (19)$$

This proves the lemma for the case when the support of  $\rho_{d,1}$  is contained in  $\mathcal{H}_4$ . (If  $\rho_{d,1}$  is a mixed state, we just express it as a mixture of pure states  $|\varphi\rangle$ . The bound for  $\rho_{d,1}$  follows by summing equations (19) for every  $|\varphi\rangle$ .)

For the general case, we divide the entire state space  $\mathcal{H}_I$  into 4-dimensional subspaces. To do that, we first subdivide  $\mathcal{H}_I$  into subspaces

$$(T_{j,0,0} \oplus T_{j,0,1} \oplus T_{j,1,0} \oplus T_{j,1,1}) \otimes S'_{j_2} \otimes \dots \otimes S'_{j_k}. \quad (20)$$

Let  $|\psi_{1,i}^{0,0}\rangle$ ,  $i \in [\dim T_{j,0,0}]$  be a basis for  $T_{j,0,0}$  and let  $|\psi_{1,i}^{a,b}\rangle = U_{ab}|\psi_{1,i}^{0,0}\rangle$  for  $(a,b) \in \{(0,1), (1,0), (1,1)\}$ . Then,  $|\psi_{1,i}^{a,b}\rangle$  is a basis for  $T_{j,a,b}$ , by Claim 11.

Let  $|\psi_{l,i}\rangle$ ,  $i \in [\dim S'_{j_l}]$  be a basis for  $S'_{j_l}$ ,  $l \in \{2, \dots, k\}$ . We subdivide (20) into 4-dimensional subspaces  $H_{i_1, \dots, i_k}$  spanned by

$$|\psi_{1,i}^{a,b}\rangle \otimes |\psi_{2,i}\rangle \otimes \dots \otimes |\psi_{k,i}\rangle,$$

where  $a, b$  range over  $\{0, 1\}$ . Let  $\mathcal{H}_{all}$  be the collection of all  $H_{i_1, \dots, i_k}$  obtained by subdividing all subspaces (20). We claim that

$$P(\rho) = \sum_{H \in \mathcal{H}_{all}} P(\Pi_H \rho). \quad (21)$$

The equation (21), together with equation (19) implies the lemma. Since  $P(\rho)$  is defined as a weighted sum of traces  $\text{Tr } \Pi_{V_m} \rho$ , we can prove equation (21) by showing

$$\text{Tr } \Pi_{V_m} \rho_{d,1} = \sum_{H \in \mathcal{H}_{all}} \text{Tr } \Pi_{V_m} \Pi_H \rho_{d,1}. \quad (22)$$

To prove (22), we define a basis for  $\mathcal{H}_I$  by first decomposing  $\mathcal{H}_I$  into subspaces  $H \in \mathcal{H}_{all}$  and, then, for each subspace, taking the basis consisting of  $|\chi_1\rangle$ ,  $|\chi_2\rangle$ ,  $|\chi_3\rangle$  and  $|\chi_4\rangle$  defined by equation (14). By Claim 12, each of the basis states belongs to one of the subspaces  $V_m$ . This means that each  $V_m$  is spanned by some subset of this basis.

The left hand side of (22) is equal to the sum of squared projections of  $\rho_{d,1}$  to basis states  $|\chi_j\rangle$  that belong to  $V_m$ . Each of the terms  $\text{Tr } \Pi_{V_m} \Pi_H \rho_{d,1}$  on the right hand side is equal to the sum of squared projections to basis states  $|\chi_j\rangle$  that belong to  $V_m \cap H$ . Summing over all  $H$  gives the sum of squared projections of  $\rho_{d,1}$  to all  $|\chi_j\rangle$  that belong to  $V_m$ . Therefore, the two sides of (22) are equal.

## B Main Polynomial Lemma

In this appendix we supply the main lemma about polynomials used in the proof of the direct product theorem for threshold functions of Section 4. We use three results about polynomials, also used in [BCWZ99, KŠW04]. The first is by Coppersmith and Rivlin [CR92, p. 980] and gives a general bound for polynomials bounded by 1 at integer points:

**Theorem 15 (Coppersmith & Rivlin [CR92])** *Every polynomial  $p$  of degree  $d \leq n$  that has absolute value*

$$|p(i)| \leq 1 \text{ for all integers } i \in [0, n],$$

*satisfies*

$$|p(x)| < ae^{bd^2/n} \text{ for all real } x \in [0, n],$$

*where  $a, b > 0$  are universal constants (no explicit values for  $a$  and  $b$  are given in [CR92]).*

The other two results concern the Chebyshev polynomials  $T_d$ , defined as in [Riv90]:

$$T_d(x) = \frac{1}{2} \left( \left( x + \sqrt{x^2 - 1} \right)^d + \left( x - \sqrt{x^2 - 1} \right)^d \right).$$

$T_d$  has degree  $d$  and its absolute value  $|T_d(x)|$  is bounded by 1 if  $x \in [-1, 1]$ . On the interval  $[1, \infty)$ ,  $T_d$  exceeds all others polynomials with those two properties ([Riv90, p.108] and [Pat92, Fact 2]):

**Theorem 16** *If  $q$  is a polynomial of degree  $d$  such that  $|q(x)| \leq 1$  for all  $x \in [-1, 1]$  then  $|q(x)| \leq |T_d(x)|$  for all  $x \geq 1$ .*

Paturi [Pat92, before Fact 2] proved

**Lemma 17 (Paturi [Pat92])**  $T_d(1 + \mu) \leq e^{2d\sqrt{2\mu + \mu^2}}$  for all  $\mu \geq 0$ .

**Proof.** For  $x = 1 + \mu$ :  $T_d(x) \leq (x + \sqrt{x^2 - 1})^d = (1 + \mu + \sqrt{2\mu + \mu^2})^d \leq (1 + 2\sqrt{2\mu + \mu^2})^d \leq e^{2d\sqrt{2\mu + \mu^2}}$  (using that  $1 + z \leq e^z$  for all real  $z$ ).  $\square$

The following lemma is key. It analyzes polynomials that are 0 on the first  $m$  integer points, and that significantly “jump” a bit later.

**Lemma 18** *Suppose  $E, N, m$  are integers satisfying  $10 \leq E \leq \frac{N}{2m}$ , and let  $p$  be a degree- $D$  polynomial such that*

$$\begin{aligned} p(i) &= 0 \text{ for all } i \in \{0, \dots, m-1\}, \\ p(8m) &= \sigma, \\ p(i) &\in [0, 1] \text{ for all } i \in \{0, \dots, N\}. \end{aligned}$$

*Then  $\sigma \leq 2^{O(D^2/N + D\sqrt{Em/N} - m \log E)}$ .*

**Proof.** Divide  $p$  by  $\prod_{j=0}^{m-1} (x - j)$  to obtain

$$p(x) = q(x) \prod_{j=0}^{m-1} (x - j),$$

where  $d = \deg(q) = D - m$ . This implies the following about the values of the polynomial  $q$ :

$$\begin{aligned} |q(8m)| &\geq \sigma / (8m)^m \\ |q(i)| &\leq 1 / ((E - 1)m)^m \text{ for } i \in \{Em, \dots, N\} \end{aligned}$$

Theorem 15 implies that there are constants  $a, b > 0$  such that

$$|q(x)| \leq \frac{a}{((E - 1)m)^m} e^{bd^2/(N - Em)} = B \text{ for all real } x \in [Em, N].$$

We now divide  $q$  by  $B$  to normalize it, and rescale the interval  $[Em, N]$  to  $[1, -1]$  to get a degree- $d$  polynomial  $t$  satisfying

$$\begin{aligned} |t(x)| &\leq 1 \text{ for all } x \in [-1, 1] \\ t(1 + \mu) &= q(8m)/B \text{ for } \mu = 2(E - 8)m/(N - Em) \end{aligned}$$

Since  $t$  cannot grow faster than the degree- $d$  Chebyshev polynomial, we have

$$t(1 + \mu) \leq e^{2d\sqrt{2\mu+\mu^2}}.$$

Combining our upper and lower bounds on  $t(1 + \mu)$  gives

$$\frac{\sigma}{(8m)^m} \cdot \frac{((E-1)m)^m}{ae^{O(d^2/N)}} \leq e^{O(d\sqrt{Em/N})},$$

which implies the lemma.  $\square$

## C Query complexity of MatrixProduct

**Proof of Theorem 6.** The space complexity of *SmallMatrixProduct* is  $O(S' \log N) = O(S)$ , because it stores a subset  $U \subseteq \{1, \dots, S'\}$ , integer vectors  $y, b$  of length  $S'$  with numbers at most  $t \leq N$ , the set  $J$  of size  $O(S')$  with numbers at most  $N$ , and a few counters. Let us compute its query complexity.

Consider the  $i$ -th block found by *SmallMatrixProduct*; let  $p_i$  be its left column and let  $\ell_i$  be its length. The scalar product  $C_{p_i, \ell_i}^U$  is estimated by quantum counting<sup>4</sup> with the number of queries  $M = \sqrt{\ell_i}$ . Finding a proper  $\ell_i$  requires  $O(\log \ell_i)$  iterations. Let  $r_i$  be the number of rows closed during the processing of this block and let  $s_i$  be the total number added to the counters for other (still open) rows in this block. The numbers  $\ell_i, r_i, s_i$  are random variables. If we instantiate them at the end of the quantum subroutine, the following inequalities hold:

$$\sum_i \ell_i \leq N, \quad \sum_i r_i \leq S', \quad \text{and} \quad \sum_i s_i \leq tS'.$$

The iterated Grover search finds ones for two purposes: closing rows and increasing counters. Since each  $b_i \leq t$ , the total cost in the  $i$ -th block is at most  $\sum_{j=1}^{r_i t} O(\sqrt{\ell_i/j}) + \sum_{j=1}^{s_i} O(\sqrt{\ell_i/j}) = O(\sqrt{\ell_i r_i t} + \sqrt{\ell_i s_i})$ . By a Cauchy-Schwarz inequality, the total number of queries that *SmallMatrixProduct* spends in the Grover searches is at most

$$\sum_{i=1}^{\text{\#blocks}} (\sqrt{\ell_i r_i t} + \sqrt{\ell_i s_i}) \leq \sqrt{\sum_i \ell_i} \sqrt{t \sum_i r_i} + \sqrt{\sum_i \ell_i} \sqrt{\sum_i s_i} \leq \sqrt{N} \sqrt{tS'} + \sqrt{N} \sqrt{tS'} = O(\sqrt{NS't}).$$

The error probability of the Grover searches can be made polynomially small in a logarithmic overhead. It remains to analyze the outcome and error probability of quantum counting. Let  $c_i = C_{p_i, \ell_i}^U \in [S', 2S']$ . One quantum counting call with  $M = \sqrt{\ell_i}$  gives a  $w$  such that

$$|w - c_i| = O\left(\sqrt{\frac{c_i(\ell_i - c_i)}{\ell_i}} + \frac{\ell_i}{\ell_i}\right) = O(\sqrt{c_i}) = O(\sqrt{S'})$$

---

<sup>4</sup>Quantum counting works on bits and we need to compute a scalar product of two *integer* vectors. It is clear that  $\{0, 1\}$ -numbers are the worst case for us. Hence we can change all non-zeroes to 1 for the purposes of quantum counting.

with probability at least  $8/\pi^2 \approx 0.8$ . We do it  $O(\log N)$  times and take the median, hence we obtain an estimate of  $c_i$  with accuracy  $O(\sqrt{S'})$  with polynomially small error probability. The result of quantum counting is compared with the given threshold, that is with  $S'$  or  $2S'$ . Binary search for  $\ell \in [\frac{k}{2}, k]$  costs another factor of  $\log k \leq \log N$ . By a Cauchy-Schwarz inequality, the total number of queries spent in the quantum counting is at most  $(\log N)^2$  times

$$\sum_i \sqrt{\ell_i} \leq \sqrt{\sum_i \ell_i} \sqrt{\sum_i 1} \leq \sqrt{N} \sqrt{\# \text{blocks}} \leq \sqrt{N} \sqrt{S' + t} \leq \sqrt{NS't},$$

because in every block the algorithm closes a row or adds  $\Theta(S')$  in total to the counters. The number of closed rows is at most  $S'$  and the number  $S'$  can be added at most  $t$  times.

The total query complexity of *SmallMatrixProduct* is thus  $O(\sqrt{NS't} \cdot (\log N)^2)$  and the query complexity of *MatrixProduct* is  $N/S'$ -times bigger. The overall error probability is at most the sum of the polynomially small error probabilities of the different subroutines, hence it can be kept below  $1/3$ .  $\square$