Algebraic Problems in Computational Complexity

A thesis submitted to the University of Mumbai
for the degree of
Doctor of Philosophy in Computer Science

by

Pranab Sen

School of Technology and Computer Science
Tata Institute of Fundamental Research
Mumbai 400005, India

2001

Statutory Declarations

Name of the Candidate	:	Pranab Sen
Title of the Thesis	:	Algebraic Problems in Computational Complexity
Degree	:	Doctor of Philosophy in the Faculty of Sciences
Subject	:	Computer Science
Name of the Guide	:	Prof. R .K .Shyamasundar
Registration Number and Date	:	TIFR171, January 23, 1998
Place of Research	:	School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai 400005, India

STATEMENT BY THE CANDIDATE

As required by the University Ordinances 770 and 771, I wish to state that the work embodied in this thesis titled "**Algebraic Problems in Computational Complexity**" forms my own contribution to the research work carried out under the guidance of **Prof. R. K. Shyamasundar** at the Tata Institute of Fundamental Research. This work has not been submitted for any other degree of this or any other University. Whenever references have been made to previous works of others, it has been clearly indicated as such and included in the Bibliography.

Certified by

Signature of Guide	Signature of Candidate
Prof. R. K. Shyamasundar Name of Guide	Pranab Sen Name of Candidate

To Ma and Baba

Acknowledgements

I am deeply indebted to my adviser, Jaikumar Radhakrishnan, for his support and guidance during the course of this thesis. Learning from him and working with him has been an immensely satisfying experience. His insights and clarity of thought have been present at every moment of this work, and I owe a great intellectual debt to him. He has been a friend and guide throughout my stay at TIFR, always encouraging me and believing in me, even in those times when I did not do so myself! I want to thank him for giving me a lot of freedom, academic and otherwise, to study what I want, to pursue my non-academic interests, and to fool around!

I thank R. K. Shyamasundar for serving as my official guide, and giving me freedom to pursue my research interests in TIFR.

Part of the work in this thesis was done during my visit to UC Berkeley and DIMACS, under a Sarojini Damodaran International Fellowship grant. I am grateful to Umesh Vazirani for supporting my visit to Berkeley, and to Eric Allender and Mike Saks for supporting my visit to DIMACS. I also thank Ashwin Nayak for the many stimulating discussions on quantum computing that I had with him in Berkeley and DIMACS, which have helped me a lot, and directly influenced part of this work.

I thank Amir Shpilka for sending me a preliminary version of his paper "Affine projections of symmetric polynomials" which directly inspired part of the work in this thesis. I also thank Hartmut Klauck and Peter Bro Miltersen for useful discussions, which have influenced part of this work.

I am grateful to Ajit Diwan, my B.Tech. adviser at IIT Bombay, for encouraging me to take up a research career in theoretical computer science. His clear thinking and attitude to problem solving will always be an inspiration. I also thank Sundar Vishwanathan for his wonderful courses during my B.Tech. days, which inspired me to take up theoretical computer science. He has also been a collaborator for part of this work.

I thank V. Arvind for supporting my visits to IMSc., and for the interesting discussions that I had with him. I also thank Ravi Rao, B. Sury and R. Sridharan of the School of Mathematics at TIFR for their courses on algebra and analysis which I took during my second year here. I learnt a lot of mathematics in those courses, some of which helped me directly in this work.

I would like to thank all the members of the School of Technology and Computer Science, past and present, for their encouragement and help that they extended to me at various stages of my stay here. I wish to thank R. K. Shyamasundar, P. S. Subramanian,

Paritosh Pandya, Subir Ghosh, N. Raja, Y. S. Ramakrishna, Purandar Bhaduri, Milind Sohoni, Abhiram Ranade and Vivek Borkar for the courses that they have given, and all that I have learnt from them. John Barretto and the other office staff deserve a special word of thanks for their excellent administrative support, which has really smoothened the life of a research scholar here. John has often gone out of his way to help me.

TIFR has been a great place to live in, mainly because of the many friends I have had here over the years. Kumar and Basant have been great seniors and I have learnt a lot from them. I have had wild and wonderful times with Venks, Karri, Holla and Amalendu. Venks has also been a collaborator for much of this work. Kavitha has been a close friend all these years. The atmosphere in the group really livened up with the arrival of the three *chotus*—Krishnan, Amitava, and the one and only Rahul Jain! I thank the other research scholars in STCS, Anoop, Aghav and Narayanan, for their enjoyable company. I also thank Anjali for the great time we had when she was a visiting student here.

I have been fortunate to have had many friends in TIFR outside the department—IG, Jishnu, Maneesh, Siddhartha, Pralay, Preeti, Keshari, Debu, Arvind, Tomás, Rajesh, Arun, Sanjib, Manojendu, Tirtha, Santosh, Surjeet, Yeshpal and Ashok. The long and hearty conversations in McRajan and the TIFR colonnade that I have had with them, their company in music concerts and treks—these memories shall remain with me for a long time. I also thank Ravindra for his great company and help during my visits to IMSc.

In TIFR, I have been extremely fortunate to have got the opportunity to learn Hindustani classical music. I express my deep sense of gratitude to *Guruji* for teaching me how to sing (though some people still harbour some doubts)! Thanks to him, music has become a very important part of my life, and it shall remain so always.

And finally, I express my heartfelt thanks to *Ma* and *Baba* for their patience, love and support all these long years. I dedicate this thesis to them.

Synopsis

Introduction

Given a computational task, we can ask the following question: what is the amount of resources we need to carry out this task? Computational complexity theory aims at determining the exact amount of resources required to solve a problem in a mathematical model of computation.

In this thesis we study some problems in computational complexity, where the models of computation have an algebraic flavour. Specifically, we study the computational complexity of some problems in the *arithmetic circuit*, *quantum cell probe* and *quantum two-party communication* models.

This synopsis is organised as follows. In the next section, we formally define the computational models and the problems therein, which have been studied in this thesis. We outline the main results obtained in the section after that.

Computational models and problems studied

$\Sigma\Pi\Sigma$ arithmetic circuits

By a $\Sigma\Pi\Sigma$ arithmetic circuit over a field $\mathbb{F}$, we mean an expression of the form

$$\sum_{i=1}^{r}\prod_{j=1}^{s_i} L_{ij}(X)$$

where each L_{ij} is a (possibly inhomogeneous) linear form in variables $X_1, \ldots, X_n$. The above expression is to be treated as over the field $\mathbb{F}$. Such 'depth-three' circuits play an important role in the study of arithmetic complexity [GR00, SW99]. If each linear form $L_{ij}(X)$ is homogeneous (i.e. has constant term zero), then the circuit is said to be *homogeneous*, or else, it is said to be *inhomogeneous*. We also define a restricted homogeneous model, the *graph model*, where all the coefficients of the variables in the linear forms have to be 0 or 1, and for a given i, no variable can occur (with coefficient 1) in more than one L_{ij}. Although depth-three circuits appear to be rather restrictive, these are the strongest model of arithmetic circuits for which super polynomial lower bounds are known; no such lower bounds are known at present for depth-four circuits.

The degree two elementary symmetric polynomial on n variables is defined by

$$S_n^2(X_1, \ldots, X_n) \triangleq \sum_{1 \le i < j \le n} X_i X_j$$

In this thesis, we study the problem of computing $S_n^2(X_1, \ldots, X_n)$ using $\Sigma\Pi\Sigma$ arithmetic circuits over several fields, with the aim of obtaining tight bounds on the number of multiplication gates required. Many of the techniques developed earlier (e.g. Nisan and Wigderson's method of partial derivatives [NW96]), in fact, give lower bounds on the number of multiplication gates. We show our upper bounds in the graph and the homogeneous model; our lower bounds hold even in the stronger inhomogeneous model. We obtain matching exact bounds for infinitely many n, for various fields.

Bounds on the number of multiplication gates required for computing $S_n^2(X)$ over the field $\mathbb{R}$ in the graph model imply the same bounds for the problem of covering the complete graph on n vertices K_n by complete bipartite graphs, such that each edge is covered exactly once. This problem was first solved by Graham and Pollack [GP72], who showed the tight bound of $n - 1$ for all n. Bounds on the number of multiplication gates required for computing $S_n^2(X)$ over the field GF(2) in the graph model imply the same bounds for the *odd cover problem*. In the odd cover problem, we want to cover K_n using complete bipartite graphs, such that each edge is covered an odd number of times. The connection to combinatorial problems is one more reason why we are interested in the number of multiplication gates in $\Sigma\Pi\Sigma$ circuits computing $S_n^2(X)$. The odd cover problem was stated by Babai and Frankl [BF92], who also observed a lower bound of $\lfloor n/2 \rfloor$. But the problem of finding matching upper bounds was left open. In this thesis, we obtain a tight matching bound of $\lceil n/2 \rceil$ for infinitely many odd and even n.

The quantum cell probe model

A static data structure problem consists of a set of data D, a set of queries Q, a set of answers A, and a function $f : D \times Q \to A$. The aim is to store the data efficiently and succinctly, so that any query can be answered with only a few probes to the data structure. Yao [Yao81] started the study of static data structure problems in the classical cell probe model. A classical (s, w, t) *cell probe scheme* for f has two components: a *storage scheme* and a *query scheme*. Given the data to be stored, the storage scheme stores it as a table of s cells, each cell w bits long. The query scheme has to answer queries about the data stored. Given a query, the query scheme computes the answer to that query by making at most t probes to the stored table, where each probe reads one cell at a time. The storage scheme is deterministic whereas the query scheme can be deterministic or randomised. The goal is to study tradeoffs between s, t and w.

In this thesis, we initiate the study of static data structure problems in the quantum setting. To this end, we define the *quantum cell probe model*. A quantum (s, w, t) *cell probe scheme* for f has two components: a classical deterministic *storage scheme* that stores the data $d \in D$ in a table T_d using s cells each containing w bits, and a quantum *query scheme*

that answers queries by 'quantumly probing a cell at a time' at most t times. Formally speaking, the table T_d for the stored data is made available to the query algorithm in the form of an oracle unitary transform O_d. To define O_d formally, we represent the basis states of the query algorithm as $|j, b, z\rangle$, where $j \in [s-1]$ is a binary string of length $\log s$, b is a binary string of length w, and z is a binary string of some fixed length. Here, j denotes the address of a cell in the table T_d, b denotes the qubits which will hold the contents of a cell and z stands for the rest of the qubits ('work qubits') in the query algorithm. O_d maps $|j, b, z\rangle$ to $|j, b \oplus (T_d)_j, z\rangle$, where $(T_d)_j$ is a bit string of length w and denotes the contents of the jth cell in T_d. A quantum query scheme with t probes is just a sequence of unitary transformations

$$U_0 \to O_d \to U_1 \to O_d \to \ldots U_{t-1} \to O_d \to U_t$$

where U_j's are arbitrary unitary transformations that do not depend on the data stored (representing the internal computations of the query algorithm). For a query $q \in Q$, the computation starts in a computational basis state $|q\rangle|0\rangle$, where we assume that the ancilla qubits are initially in the basis state $|0\rangle$. Then we apply in succession, the operators $U_0, O_d, U_1, \ldots, U_{t-1}, O_d, U_t$, and measure the final state. The answer consists of the values on some of the output wires of the circuit. We say that the scheme has worst case error probability less than ϵ if the answer is equal to $f(d, q)$, for every $(d, q) \in D \times Q$, with probability greater than $1 - \epsilon$. The term 'exact quantum scheme' means that $\epsilon = 0$, and the term 'bounded error quantum scheme' means that $\epsilon = 1/3$.

In this thesis, we study the *static membership* problem. Here one has to store a subset S of size at most n from a universe $\mathbf{U}$ of size m, such that, given any query element $x \in \mathbf{U}$, one can quickly decide whether $x \in S$. 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] and Pagh [Pag01]). Most of these studies were in the classical deterministic cell probe model. Yao [Yao81] showed that if the storage scheme is restricted to be *implicit*, that is, the storage scheme can either store a member of S in a cell or a 'pointer value' (the family of 'pointer values' is a set disjoint from the universe U), then any deterministic query algorithm requires $\Omega(\log n)$ probes in the worst case, provided that the universe U is large enough. Recently, this problem was considered by Buhrman, Miltersen, Radhakrishnan and Venkatesh [BMRV00] in the *classical bit probe model* (cell probe model where the cell size is only one bit), 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. In this work, we study this problem in the quantum bit probe model and show tradeoffs between storage space and the number of probes for exact quantum bit probe schemes and lower bounds on the storage space for ϵ-error quantum bit probe schemes making a given number of probes. We also study this problem in the bounded error quantum cell probe model with implicit storage schemes, and extend Yao's result to the quantum setting.

We also study the *static predecessor problem*. Here one has to store a subset S of size at most n from the universe $[m]$, such that, given any query element $x \in [m]$, one can quickly

find the predecessor of x in S. We show lower bounds for this problem in a restricted version of the quantum cell probe model viz. the *address-only* quantum cell probe model. Here the storage scheme is as in the general model, but the query scheme is restricted to be 'address-only'. This means that the state vector before a query to the oracle O_d is always a *tensor product* of a state vector on the address and work qubits (the (j, z) part in (j, b, z) above), and a state vector on the data qubits (the b part in (j, b, z) above). The state vector on the data qubits before a query to the oracle O_d is *independent of the query element q and the data d* but can vary with the probe number. Intuitively, we are only making use of quantum parallelism over the address lines. This mode of querying a table subsumes classical querying, and also many non-trivial quantum algorithms like Grover's algorithm [Gro96], Farhi *et al.*'s algorithm [FGGS99], Høyer *et al.*'s algorithm [HNS01] etc. satisfy this condition. For classical querying, the state vector on the data qubits is $|0\rangle$, independent of the probe number. For Grover and Farhi *et al.*, the state vector on the data qubit is $(|0\rangle - |1\rangle)/\sqrt{2}$, independent of the probe number. For Høyer *et al.*, the state vector on the data qubit is $|0\rangle$ for some probe numbers, and $(|0\rangle - |1\rangle)/\sqrt{2}$ for the other probe numbers.

The two-party quantum communication model

This model was defined by Yao [Yao93] to study communication as a resource in quantum computation. Let E, F, G be arbitrary finite sets and $f : E \times F \to G$ be a function. There are two players Alice and Bob, who hold qubits. When the communication game starts, Alice holds $|x\rangle$ where $x \in E$ together with some ancilla qubits in the state $|0\rangle$, and Bob holds $|y\rangle$ where $y \in F$ together with some ancilla qubits in the state $|0\rangle$. Thus the qubits of Alice and Bob are initially in computational basis states, and the initial superposition is simply $|x\rangle_A|0\rangle_A|y\rangle_B|0\rangle_B$. Here the subscripts denote the ownership of the qubits by Alice and Bob. The players take turns to communicate to compute $f(x, y)$. Suppose it is Alice's turn. Alice can make an arbitrary unitary transformation on her qubits and then send one or more qubits to Bob. Sending qubits does not change the overall superposition, but rather changes the ownership of the qubits, allowing Bob to apply his next unitary transformation on his original qubits plus the newly received qubits. At the end of the protocol, the last recipient of qubits performs a measurement on the qubits in her possession to output an answer. We say a quantum protocol computes f with ϵ-error in the worst case, if for any input $(x, y) \in E \times F$, the probability that the protocol outputs the correct result $f(x, y)$ is greater than $1 - \epsilon$. The term 'bounded error quantum protocol' means that $\epsilon = 1/3$.

We require that Alice and Bob make a secure copy of their inputs before beginning the protocol. This is possible since the inputs to Alice and Bob are in computational basis states. Thus, without loss of generality, the input qubits of Alice and Bob are never sent as messages, their state remains unchanged throughout the protocol, and they are never measured i.e. some work qubits are measured to determine the result of the protocol. We call such protocols *secure*. We will assume henceforth that all our protocols are secure.

To state our round elimination lemma in quantum communication, we have to define the concept of a *safe* quantum communication protocol.

Definition (Safe quantum protocol) *A $[t, c, l_1, \ldots, l_t]^A$ ($[t, c, l_1, \ldots, l_t]^B$) safe quantum protocol is a secure quantum protocol where Alice (Bob) starts the communication, the first message is $l_1 + c$ qubits long, the ith message, for $i \geq 2$, is l_i qubits long, and the communication goes on for t rounds. We think of the first message as having two parts: the 'main part' which is l_1 qubits long, and the 'safe overhead part' which is c qubits long. The density matrix of the 'safe overhead' is independent of the inputs to Alice and Bob.*

For the round elimination lemma, we also need to define the concept of a quantum protocol with *public coins*. Intuitively, a public coin quantum protocol is a probability distribution over (*coinless*) quantum protocols. We shall henceforth call the standard definition of a quantum protocol as *coinless*. Our definition is similar to the classical scenario, where a randomised protocol with public coins is a probability distribution over deterministic protocols. We note however, that our definition of a public coin quantum protocol is *not* the same as that of a quantum protocol with prior entanglement, which has been studied previously (see e.g. [CvDNT98]). Our definition is weaker, in that it does not allow the unitary transformations of Alice and Bob to alter the 'public coin'.

Definition (Public coin quantum protocol) *In a quantum protocol with a public coin, there is, before the start of the protocol, a quantum state called a* public coin*, of the form $\sum_c \sqrt{p_c}|c\rangle_A|c\rangle_B$, where the subscripts denote ownership of qubits by Alice and Bob, p_c are finitely many non-negative real numbers and $\sum_c p_c = 1$. Alice and Bob make (entangled) copies of their respective halves of the public coin using* CNOT *gates before commencing the protocol. The unitary transformations of Alice and Bob during the protocol do not touch the public coin. The public coin is never measured, nor is it ever sent as a message.*

Hence, one can think of the public coin quantum protocol to be a probability distribution, with probability p_c, over finitely many coinless quantum protocols indexed by the coin basis states $|c\rangle$. A *safe public coin* quantum protocol is similarly defined as a probability distribution over finitely many safe coinless quantum protocols.

In this thesis, we prove a round elimination lemma in quantum communication complexity. Suppose $f : E \times F \to G$ is a function. In the communication game corresponding to f, Alice gets a string $x \in E$, Bob gets a string $y \in F$, and they have to communicate and compute $f(x, y)$. In the communication game $f^{(n)}$, Alice gets n strings $x_1, \ldots, x_n \in E$; Bob gets an integer $i \in [n]$, a string $y \in F$, and a copy of the strings $x_1, \ldots, x_{i-1}$. Their aim is to communicate and compute $f(x_i, y)$. Suppose a quantum protocol for $f^{(n)}$ is given where Alice starts, and her first message is much smaller than n qubits. Intuitively, it would seem that since Alice does not know i, the first round of communication cannot give much information about x_i, and thus, would not be very useful to Bob. Hence it should be possible to eliminate the first round of communication, giving a quantum protocol for computing $f(x_i, y)$ where Bob starts, with one less round of communication, and having the same message complexity and similar error probability. The round elimination lemma

justifies this intuition. Moreover, we show that this is true even if Bob also gets copies of $x_1, \ldots, x_{i-1}$, a case which is needed in many applications.

Main results

Bounds for computing $S_n^2(X)$

We prove the following results about the number of multiplication gates required to compute $S_n^2(X)$ using $\Sigma\Pi\Sigma$ arithmetic circuits over various fields. In each case, our upper and lower bounds match for infinitely many n. The proofs use linear algebraic techniques. Our upper bounds are in the homogeneous model; our lower bounds hold in the inhomogeneous model too. Our bounds in the graph model of $\Sigma\Pi\Sigma$ arithmetic circuits for the field GF(2) translate to the same bounds for the odd cover problem too.

Result *For infinitely many odd and even n, $\lceil n/2 \rceil$ complete bipartite graphs are necessary and sufficient to cover each edge of the complete graph on n vertices an odd number of times. A similar result also holds for the number of multiplication gates required to compute $S_n^2(X_1, \ldots, X_n)$ over the field GF(2), using $\Sigma\Pi\Sigma$ arithmetic circuits.*

Result *For all n, $\lceil n/2 \rceil$ multiplication gates are necessary and sufficient to compute $S_n^2(X_1, \ldots, X_n)$ over complex numbers, using $\Sigma\Pi\Sigma$ arithmetic circuits.*

The above results are joint work with Jaikumar Radhakrishnan and Sundar Vishwanathan [RSV00b].

Static membership in quantum bit probe model

We show a general time-space tradeoff for the static membership problem in the exact quantum bit probe model, using linear algebraic techniques.

Result *Suppose there exists an exact quantum bit probe scheme for storing subsets S of size at most n from a universe of size m that uses s bits of storage and answers membership queries 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 with exact quantum querying. 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 with exact quantum querying—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 with exact quantum querying, 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 use $O(nt(m/n)^{2/(t+1)})$ 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 the above result. 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 result 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, but always answers 'No' for a query x that is not a member of S.

We also give a simplified linear algebraic proof of the above theorem for deterministic and positive error classical bit probe schemes. This theorem is in fact stronger than the tradeoff results known previously for such schemes.

We then consider ϵ-error quantum bit probe schemes. We show the following lower bound on the space requirement of an ϵ-error quantum bit probe scheme for the static membership problem making p probes. The proof again uses linear algebraic methods.

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

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

Buhrman *et al.* [BMRV00] showed the existence of two-sided ϵ-error classical bit probe schemes which solve the static membership problem, making only one bit probe and using space $O(\frac{n \log m}{\epsilon^2})$, for $\epsilon < 1/16$. From this result, 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 above storage scheme 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.

We also improve the lower bound in the result above on the space requirement of ϵ-error bit probe schemes for the static membership problem making p probes, when the query schemes are classical randomised.

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

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

These results are joint work with Jaikumar Radhakrishnan and S.Venkatesh [RSV00a].

Static membership in implicit storage quantum cell probe model

In this thesis, we generalise the $\Omega(\log n)$ lower bound of Yao on the number of probes required in any classical deterministic cell probe solution to the static membership problem with implicit storage schemes, to the quantum setting. Consider the problem of storing a subset S of size at most n of the universe $[m]$ in a table with q cells, so that membership queries can be answered efficiently. We restrict the storage scheme to be *implicit*, using at most p 'pointer values'. A 'pointer value' is a member of a set of size p (the set of 'pointers') disjoint from the universe. The term implicit means that the storage scheme can store either a 'pointer value' or a member of S in a cell. In particular, the storage scheme is not allowed to store an element of the universe which is not a member of S. The query algorithm answers membership queries by performing t (general) quantum cell probes. We call such schemes (p, q, t) *implicit storage quantum cell probe schemes.*

Result *For every n, p, q, there exists an $N(n, p, q)$ such that for all $m \geq N(n, p, q)$, the following holds: Consider any bounded error (p, q, t) implicit storage quantum cell probe scheme for the static membership problem with universe size m and size of the stored subset at most n. Then the quantum query scheme must make $t = \Omega(\log n)$ probes.*

This result is joint work with S.Venkatesh [SV01].

Static predecessor in address-only quantum cell probe model

To show lower bounds for the static predecessor problem in the address-only quantum cell probe model, we use a connection between quantum cell probe schemes for static data structure problems and two-party quantum communication complexity. This connection similar to that in Miltersen, Nisan, Safra and Wigderson [MNSW98], who exploited it in the classical setting. Using this connection, we can convert an address-only quantum cell probe solution for the predecessor problem into a particular kind of quantum communication game. The quantum round elimination lemma is then used to prove lower bounds on the rounds complexity of this game. Using this approach, we prove the following theorem.

Result *Suppose we have a* $(n^{O(1)}, (\log m)^{O(1)}, t)$ *bounded error quantum address-only cell probe solution to the static predecessor problem, where the universe size is* m *and the subset size is at most* n*. Then the number of queries* t *is at least* $\Omega\left(\frac{\log\log m}{\log\log\log m}\right)$ *as a function of* m*, and at least* $\Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$ *as a function of* n*.*

Since our *address-only* quantum cell probe model subsumes the classical cell probe model with randomised query schemes, our lower bound for the static predecessor problem also holds in this classical randomised setting. This improves the previous lower bound of $\Omega(\sqrt{\log\log m})$ as a function of m and $\Omega(\log^{1/3} n)$ as a function of n for this setting, shown by Miltersen, Nisan, Safra and Wigderson [MNSW98]. Beame and Fich [BF99] have shown an upper bound matching our lower bound up to constant factors, which uses $n^{O(1)}$ cells of storage of word size $O(\log m)$ bits. In fact, both the storage and the query schemes are classical deterministic in Beame and Fich's solution. In the classical deterministic cell probe model, Beame and Fich show a lower bound of $t = \Omega\left(\frac{\log\log m}{\log\log\log m}\right)$ as a function of m for $(n^{O(1)}, 2^{(\log m)^{1-\Omega(1)}}, t)$ cell probe schemes, and a lower bound of $t = \Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$ as a function of n for $(n^{O(1)}, (\log m)^{O(1)}, t)$ cell probe schemes. But Beame and Fich's lower bound proof breaks down if the query scheme is randomised. Our result thus shows that the upper bound scheme of Beame and Fich is optimal all the way up to the bounded error address-only quantum cell probe model. Also, our proof is substantially simpler than that of Beame and Fich.

This result is joint work with S.Venkatesh [SV01].

Round elimination in quantum and classical communication

We prove a round elimination lemma for quantum communication complexity in this thesis. This result can be viewed as a quantum analogue of the round elimination lemma of Miltersen, Nisan, Safra and Wigderson [MNSW98] for classical communication complexity. Our quantum round elimination lemma is in fact stronger (!) than the classical round elimination lemma of [MNSW98], and it allows us to show a quantum lower bound for the static predecessor problem matching Beame and Fich's upper bound, which the classical round elimination lemma of [MNSW98] was unable to do. The quantum round elimination lemma can be used to prove similar lower bounds for many other static data structure problems in the address-only quantum cell probe model. It also finds applications to various problems in quantum communication complexity (e.g. the 'greater-than' problem), which are interesting on their own. Our quantum round elimination lemma is proved using quantum information theoretic techniques, and builds on the work of Klauck *et al.* [KNTZ01].

Result *Suppose* $f : E \times F \to G$ *is a function. Suppose the communication game* $f^{(n)}$ *has a* $[t, c, l_1, \ldots, l_t]^A$ *safe public coin quantum protocol with worst case error less than* δ*.*

Then there is a $[t-1, c+l_1, l_2, \ldots, l_t]^B$ *safe public coin quantum protocol for* f *with worst case error less than* $\epsilon \stackrel{\Delta}{=} \delta + (4l_1 \ln 2/n)^{1/4}$.

In the classical setting, we can refine our information theoretic techniques to prove an even stronger round elimination lemma for classical communication complexity.

Result *Suppose* $f : E \times F \to G$ *is a function. Suppose the communication game* $f^{(n)}$ *has a* $[t, 0, l_1, \ldots, l_t]^A$ *public coin classical randomised protocol with worst case error less than* δ. *Then there is a* $[t-1, 0, l_2, \ldots, l_t]^B$ *public coin classical randomised protocol for* f *with worst case error less than* $\epsilon \stackrel{\Delta}{=} \delta + (1/2)(2l_1 \ln 2/n)^{1/2}$.

These results are joint work with S.Venkatesh [SV01].

Communication complexity of the 'greater-than' problem

As an application of our round elimination lemmas, we prove rounds versus communication tradeoffs for the 'greater-than' problem. In the 'greater-than' problem GT_n, Alice is given $x \in \{0,1\}^n$, Bob is given $y \in \{0,1\}^n$, and they have to communicate and decide whether $x > y$ (treating x, y as integers).

Result *The* t *round bounded error quantum (classical randomised) communication complexity of* GT_n *is* $\Omega(n^{1/t}t^{-3})$ *(*$\Omega(n^{1/t}t^{-2})$*).*

There exists a bounded error classical randomised protocol for GT_n using t rounds of communication and having a complexity of $O(n^{1/t} \log n)$. Hence, for a constant number of rounds, our quantum lower bound matches the classical upper bound to within logarithmic factors. For one round quantum protocols, our result implies an $\Omega(n)$ lower bound for GT_n (which is optimal to within constant factors), improving upon the previous $\Omega(n/\log n)$ lower bound of Klauck [Kla00]. No rounds versus communication tradeoff for this problem, for more than one round, was known earlier in the quantum setting. For classical randomised protocols, Miltersen *et al.* [MNSW98] showed a lower bound of $\Omega(n^{1/t}2^{-O(t)})$ using their round elimination lemma. If the number of rounds is unbounded, then there is a classical randomised protocol for GT_n using $O(\log n)$ rounds of communication and having a complexity of $O(\log n)$ [Nis93]. An $\Omega(\log n)$ lower bound for the bounded error quantum communication complexity of GT_n (irrespective of the number of rounds) follows from Kremer's result [Kre95] that the bounded error quantum communication complexity of a function is lower bounded (up to constant factors) by the logarithm of the one round (classical) deterministic communication complexity.

These results are joint work with S.Venkatesh [SV01].

List of Publications

[RSV00a] J. Radhakrishnan, P. Sen, and S. Venkatesh. The quantum complexity of set membership. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, pages 554–562, 2000. Full version to appear in *Special issue of Algorithmica on Quantum Computation and Quantum Cryptography*. Also quant-ph/0007021.

[RSV00b] J. Radhakrishnan, P. Sen, and S. Vishwanathan. Depth-3 arithmetic circuits for $S_n^2(X)$ and extensions of the Graham-Pollack theorem. In *Proceedings of the 20th conference on the Foundations of Software Technology and Theoretical Computer Science*, Lecture Notes in Computer Science, vol. 1974, pages 176–187. Springer-Verlag, 2000. Also cs.DM/0110031.

[SV01] P. Sen and S. Venkatesh. Lower bounds in the quantum cell probe model. In *Proceedings of the 28th International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science, vol. 2076, pages 358–369. Springer-Verlag, 2001. Also quant-ph/0104100.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Given a computational task, we can ask the following question: what is the amount of resources we need to carry out this task? Computational complexity theory is an area of research in theoretical computer science that aims at determining the exact amount of resources required to solve a problem in a model of computation.

Determining the exact computational complexity of a problem involves two notions. The first is to define a *mathematical model* of computation. The second notion is to define the *computational resources* used to solve a problem in this model. Once these are defined, understanding the complexity of any problem involves establishing upper and lower bounds on the amount of resources required to solve the problem. Tradeoffs between various resources are also studied.

In recent years, a lot of excitement has been generated by a new model of computation viz. quantum computation. In this thesis, the term "classical" refers to traditional non-quantum models of computation. The *quantum computation model* aims to exploit the quantum mechanical behaviour of nature for information processing purposes. The most striking example of the power of this model, so far, has been Shor's polynomial time algorithm for prime factorisation of integers on a quantum computer [Sho97]. Another notable example is Grover's quantum algorithm for searching an unstructured database using $O(\sqrt{n})$ queries.

In this thesis, we study some problems in computational complexity where the models of computation have an algebraic flavour. Specifically, we study the computational complexity of some problems in the *arithmetic circuit*, *quantum cell probe* and *quantum two-party communication* models.

In this chapter, we describe the above computational models and the problems we study in these models. We also describe the results obtained in the course of this work.

1.1 The arithmetic circuit model

Boolean circuits as a model of computation have been studied since the 1980s. Upper and lower bounds for many problems in this model have been discovered. In particular,

constant depth boolean circuits with gates of unbounded fanin have been studied with great success, and many strong lower bounds are known for various boolean functions (e.g. PARITY) in this model (see e.g. [Hås89, Smo87]).

For functions with an algebraic flavour, it is natural to consider other models of computation also. One of these is the *arithmetic circuit model*. An arithmetic circuit over a field $\mathbb{F}$ computes a polynomial in variables $X_1, \ldots, X_n$ over $\mathbb{F}$. It is a directed acyclic graph with a single node of out-degree 0, representing the 'output' of the circuit. Nodes of in-degree 0 are labelled by variables from $X_1, \ldots, X_n$. The rest of the nodes (the 'internal nodes') are labelled either by addition gates, or by multiplication gates. Here, addition and multiplication are to be understood as being over $\mathbb{F}$. The addition gate computes the sum, and the multiplication gate computes the product of the polynomials at its inputs. The edges of the graph (the 'wires' of the circuit) are labelled by scalars from $\mathbb{F}$. They are to be thought of as multiplying the polynomial at the tail of the edge, to get the polynomial at the head of the edge. Thus, every node of the circuit naturally computes a polynomial in $X_1, \ldots, X_n$ over $\mathbb{F}$. The 'output' of the circuit is the polynomial computed at the output node.

Though the arithmetic circuit model is less general than the boolean circuit model, and it may seem more amenable to mathematical study, fewer and weaker lower bounds are known for explicit polynomials in this model. In particular, lower bounds for explicit polynomials are known only if we allow polynomials with large degree or large coefficients (see e.g. [Str73, BS82]). However, if we limit the degree and size of coefficients to be $O(1)$, then no non-trivial lower bound is known for general arithmetic circuits. For constant depth circuits, exponential lower bounds are only known for fields $\mathbb{F}$ with characteristic 2 [Raz87, Smo87]. For finite fields of odd characteristic, exponential lower bounds are only known for depth 3 [GK98, GR00]; no super polynomial lower bounds are known at present for circuits of depth 4 and more. For characteristic zero, no super polynomial lower bounds are known, even for depth-3 circuits. The best lower bounds for depth-3 circuits over fields of characteristic zero are the almost quadratic lower bounds of [SW99].

By a $\Sigma\Pi\Sigma$ arithmetic circuit over a field $\mathbb{F}$, we mean an expression of the form

$$\sum_{i=1}^{r} \prod_{j=1}^{s_i} L_{ij}(X) \tag{1.1}$$

where each L_{ij} is a (possibly inhomogeneous) linear form in variables $X_1, \ldots, X_n$. The above expression is to be treated as over the field $\mathbb{F}$. If each linear form $L_{ij}(X)$ is homogeneous (i.e. has constant term zero), then the circuit is said to be *homogeneous*, or else, it is said to be *inhomogeneous*. In this thesis, we also define a restricted homogeneous model, the *graph model*, where all the coefficients of the variables in the linear forms have to be 0 or 1, and for a given i, no variable can occur (with coefficient 1) in more than one L_{ij}.

The k-th elementary symmetric polynomial on n variables is defined by

$$S_n^k(X) \triangleq \sum_{T \in \binom{[n]}{k}} \prod_{i \in T} X_i.$$

Elementary symmetric polynomials are the most commonly studied candidates for showing lower bounds in arithmetic circuits. Nisan and Wigderson [NW96] showed that any homogeneous $\Sigma\Pi\Sigma$ circuit for computing $S_n^{2k}(X)$ has size $\Omega((n/4k)^k)$. In their paper, they explicitly stated the method of partial derivatives (but see also Alon [Alo86]). Although a super polynomial lower-bound was obtained in [NW96], the lower bound applied only to homogeneous circuits. Indeed, Ben-Or (see [NW96]) showed that any elementary symmetric polynomial can be computed by an inhomogeneous $\Sigma\Pi\Sigma$ formula of size $O(n^2)$ (contrast this with super polynomial lower bounds for computing MAJORITY using constant depth boolean circuits). Thus inhomogeneous circuits are significantly more powerful than homogeneous circuits. Shpilka and Wigderson [SW99] (and later, Shpilka [Shp01]) addressed this shortcoming of the Nisan-Wigderson result and showed an $\Omega(n^2)$ lower bound on the size of inhomogeneous formulae computing certain elementary symmetric polynomials, thus showing that Ben-Or's construction is optimal.

1.1.1 Computing $S_n^2(X)$ using $\Sigma\Pi\Sigma$ arithmetic circuits

In this thesis, we study the problem of computing $S_n^2(X_1, \ldots, X_n)$, the degree two elementary symmetric polynomial in $X_1, \ldots, X_n$, using $\Sigma\Pi\Sigma$ arithmetic circuits over several fields, with the aim of obtaining tight bounds on the number of multiplication gates required. Many of the techniques developed earlier (e.g. Nisan and Wigderson's method of partial derivatives [NW96]), in fact, give lower bounds on the number of multiplication gates. We show our upper bounds in the graph and the homogeneous model; our lower bounds hold even in the stronger inhomogeneous model. We obtain matching exact bounds for infinitely many n, for various fields.

Bounds on the number of multiplication gates required for computing $S_n^2(X)$ over the field $\mathbb{R}$ in the graph model imply the same bounds for the problem of covering the complete graph on n vertices K_n by complete bipartite graphs, such that each edge is covered exactly once. This problem was first solved by Graham and Pollack [GP72], who showed the tight bound of $n-1$ for all n. Bounds on the number of multiplication gates required for computing $S_n^2(X)$ over the field GF(2) in the graph model imply the same bounds for the *odd cover problem*. In the odd cover problem, we want to cover K_n using complete bipartite graphs, such that each edge is covered an odd number of times. A similar connection holds between computing $S_n^2(X)$ over the field GF(p), p an odd prime in the graph model, and the 1 mod p cover problem (where we want to cover K_n using complete bipartite graphs, such that each edge is covered 1 mod p times). The connection to combinatorial problems is one more reason why we are interested in the number of multiplication gates in $\Sigma\Pi\Sigma$ circuits computing $S_n^2(X)$. The odd cover problem was stated by Babai and Frankl [BF92], who also observed a lower bound of $\lfloor n/2 \rfloor$. But the problem of finding matching upper bounds was left open. In this thesis, we obtain a tight matching bound of $\lceil n/2 \rceil$ for infinitely many odd and even n.

Result 1 *For infinitely many odd and even n, $\lceil n/2 \rceil$ complete bipartite graphs are necessary and sufficient to cover each edge of the complete graph on n vertices an odd number*

of times. A similar result also holds for the number of multiplication gates required to compute $S_n^2(X_1, \ldots, X_n)$ *over the field* $GF(2)$*, using* $\Sigma\Pi\Sigma$ *arithmetic circuits.*

Result 2 *For infinitely many odd and even* n*,* $\lceil n/2 \rceil$ *complete bipartite graphs are necessary and sufficient to cover each edge of the complete graph on* n *vertices* $1 \bmod p$ *times.*

Result 3 *For all* n*,* $\lceil n/2 \rceil$ *multiplication gates are necessary and sufficient to compute* $S_n^2(X_1, \ldots, X_n)$ *over complex numbers, using* $\Sigma\Pi\Sigma$ *arithmetic circuits. Similar, but weaker, results hold for computing* $S_n^2(X)$ *over finite fields of odd characteristic.*

The above results are joint work with Jaikumar Radhakrishnan and Sundar Vishwanathan [RSV00b].

1.2 The quantum cell probe model

The *classical cell probe model* is a combinatorial model for studying static and dynamic data structure problems. This model (or rather a variant, the *classical bit probe model*) was first defined in the book *Perceptrons* by Minsky and Papert [MP69]. They studied average case upper bounds for the *static membership* problem in this model. But it was Yao [Yao81], who first took up the worst-case complexity study of static data structure problems in the classical cell probe model.

A static data structure problem consists of a set of data D, a set of queries Q, a set of answers A, and a function $f : D \times Q \to A$. The aim is to store the data efficiently and succinctly, so that any query can be answered with only a few probes to the data structure. A classical (s, w, t) *cell probe scheme* for f has two components: a *storage scheme* and a *query scheme*. Given the data to be stored, the storage scheme stores it as a table of s cells, each cell w bits long. The query scheme has to answer queries about the data stored. Given a query, the query scheme computes the answer to that query by making at most t probes to the stored table, where each probe reads one cell at a time. The storage scheme is deterministic whereas the query scheme can be deterministic or randomised. The goal is to study tradeoffs between s, t and w. A crucial aspect of the cell probe model is that we only charge a scheme for the number of probes made to memory cells, and for the total number of cells of storage used. All other computation is for free. Thus lower bounds in the cell probe model are lower bounds on the complexity of any implementation of the problem on a unit cost RAM with the same word size. An important variation of the classical cell probe model is the classical bit probe model, where each cell holds just a single bit. Thus, in this model, the query algorithm is allowed to probe only one bit of the memory at a time. Arguably, the bit probe complexity of a data structure problem is a fundamental measure; this, in particular, applies to *decision problems* where the final answer to a query is a single bit.

An important static data structure problem is the *static membership* problem.

> *Let $U = \{1, 2, \ldots, m\}$. Given a subset $S \subseteq U$ of at most n keys, store it efficiently and succinctly so that queries of the form "Is x in S?" can be answered with only a few probes to the data structure.*

When the static membership problem is usually studied in the classical cell probe model, the set S is stored as a table of cells, each capable of holding one element of the universe; that is, if the universe has size m then each cell holds $O(\log m)$ bits. Queries are to be answered by probing a cell of the table at a time adaptively; that is, each probe can depend on the results of earlier probes and the query element x. The goal is to process membership queries with as few probes as possible, and at the same time keep the size of the table small. The static membership problem has a long history of study in this model. Yao [Yao81] showed that if the storage scheme is restricted to be *implicit*, that is, the storage scheme can either store a member of S in a cell or a 'pointer value' (the family of 'pointer values' is a set disjoint from the universe U), then any deterministic query algorithm requires $\Omega(\log n)$ probes in the worst case, provided that the universe U is large enough. Fredman, Komlós and Szemerédi [FKS84] gave a solution for the static membership problem in the cell probe model that used a constant number of probes and a table of size $O(n)$. Their storage scheme is not implicit though; in fact, it can store in a cell an element of the universe which is not a member of S. Note that if one is required to store sets of size at most n, then there is an information theoretic lower bound of $\left\lceil \log \sum_{i \leq n} \binom{m}{i} \right\rceil$ on the number of bits used. For $n \leq m^{1-\Omega(1)}$, this implies that the data structure must store $\Omega(n \log m)$ bits (and must, therefore, use $\Omega(n)$ cells). Thus, up to constant factors, the above scheme uses optimal space and number of cell probes. Recently, this problem was considered by Buhrman, Miltersen, Radhakrishnan and Venkatesh [BMRV00] in the classical *bit probe model*; 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 each case, their lower bounds roughly matched the upper bounds. Also recently, Pagh [Pag01] has shown classical deterministic schemes using the information-theoretic minimum space $\left\lceil \log \sum_{i \leq n} \binom{m}{i} \right\rceil$ and making $O(\log(m/n))$ bit probes. This matches the lower bound for classical deterministic schemes in [BMRV00].

Another important static data structure problem is the *static predecessor* problem.

> *Let $U = \{1, 2, \ldots, m\}$. Given a subset $S \subseteq U$ of at most n keys, store it efficiently and succinctly so that queries of the form "What is the predecessor of x in S?" can be answered with only a few probes to the data structure.*

The static predecessor problem too has a long history of study in the classical deterministic $(n^{O(1)}, O(\log m), t)$-cell probe model. Ajtai [Ajt88] was the first to show a super constant lower bound on t. The lower bounds were later improved by various people [Xia92, Mil94]. Miltersen, Nisan, Safra and Wigderson [MNSW98] showed that any classical $(n^{O(1)}, (\log m)^{O(1)}, t)$-cell probe solution to the predecessor problem with randomised query schemes requires $t = \Omega(\sqrt{\log \log m})$ as a function of m, and $t = \Omega(\log^{1/3} n)$ as a

function of n. Recently, Beame and Fich [BF99] gave a $(n^{O(1)}, O(\log m), t)$ classical deterministic cell probe solution for the predecessor problem where

$$t = O\left(\min\left(\frac{\log\log m}{\log\log\log m}, \sqrt{\frac{\log n}{\log\log n}}\right)\right)$$

Beame and Fich [BF99] also showed a lower bound of $t = \Omega\left(\frac{\log\log m}{\log\log\log m}\right)$ as a function of m for $(n^{O(1)}, 2^{(\log m)^{1-\Omega(1)}}, t)$ classical deterministic cell probe schemes for predecessor, and a lower bound of $t = \Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$ as a function of n for $(n^{O(1)}, (\log m)^{O(1)}, t)$ classical deterministic cell probe schemes for predecessor. But their lower bound proof breaks down if the query algorithm is randomised; for such schemes, the best lower bound known till now was that of Miltersen *et al.* [MNSW98]. Also, no upper bound better than that of [BF99] was known for such schemes. Thus, there was a gap between upper and lower bounds when the query scheme was randomised. For an account of many interesting results in the classical cell probe model, see the recent survey of Miltersen [Mil99].

In this thesis, we initiate the study of static data structure problems in the quantum setting. To that end, we define the *quantum cell probe model*. A quantum (s, w, t) *cell probe scheme* for a static data structure problem f has two components: a classical deterministic *storage scheme* that stores the data $d \in D$ in a table T_d using s cells each containing w bits, and a quantum *query scheme* that answers queries by 'quantumly probing a cell at a time' at most t times. Thus, our quantum cell probe model is basically the quantum black box query model (see e.g. [BBC+98]) applied to the table of cells created by the storage scheme. Formally speaking, the table T_d for the stored data is made available to the query algorithm in the form of an oracle unitary transform O_d. To define O_d formally, we represent the basis states of the query algorithm as $|j, b, z\rangle$, where $j \in [s-1]$ is a binary string of length $\log s$, b is a binary string of length w, and z is a binary string of some fixed length. Here, j denotes the address of a cell in the table T_d, b denotes the qubits which will hold the contents of a cell and z stands for the rest of the qubits ('work qubits') in the query algorithm. O_d maps $|j, b, z\rangle$ to $|j, b \oplus (T_d)_j, z\rangle$, where $(T_d)_j$ is a bit string of length w and denotes the contents of the jth cell in T_d. In most previous work on the quantum black box model, the data b was only one bit long. But in keeping with the analogy to the classical cell probe model, we allow the data here to be w bits long. A quantum query scheme with t probes is just a sequence of unitary transformations

$$U_0 \to O_d \to U_1 \to O_d \to \ldots U_{t-1} \to O_d \to U_t$$

where U_j's are arbitrary unitary transformations that do not depend on the data stored (representing the internal computations of the query algorithm). For a query $q \in Q$, the computation starts in a computational basis state $|q\rangle|0\rangle$, where we assume that the ancilla qubits are initially in the basis state $|0\rangle$. Then we apply in succession, the operators $U_0, O_d, U_1, \ldots, U_{t-1}, O_d, U_t$, and measure the final state. The answer consists of the values on some of the output wires of the circuit. We say that the scheme has worst case error

probability less than ϵ if the answer is equal to $f(d,q)$, for every $(d,q) \in D \times Q$, with probability greater than $1-\epsilon$. The term 'exact quantum scheme' means that $\epsilon = 0$, and the term 'bounded error quantum scheme' means that $\epsilon = 1/3$.

Remark: Our model for storage does not permit O_d to be any arbitrary unitary transformation. However, this restricted form of the oracle is closer to the way data is stored and 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]).

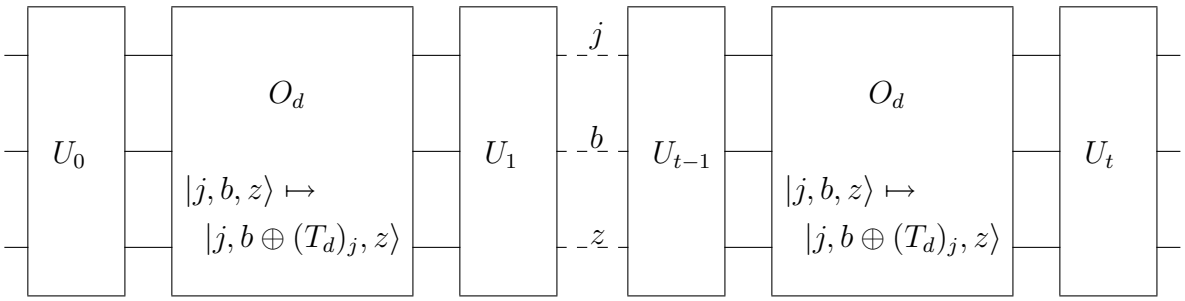


Figure 1.1: The query algorithm in a quantum cell probe scheme.

We also study a restricted version of the quantum cell probe model, which we call the *address-only* quantum cell probe model. Here the storage scheme is as in the general model, but the query scheme is restricted to be 'address-only'. This means that the state vector before a query to the oracle O_d is always a *tensor product* of a state vector on the address and work qubits (the (j,z) part in (j,b,z) above), and a state vector on the data qubits (the b part in (j,b,z) above). The state vector on the data qubits before a query to the oracle O_d is *independent of the query element q and the data d* but can vary with the probe number. Intuitively, we are only making use of quantum parallelism over the address lines. This mode of querying a table subsumes classical querying, and also many non-trivial quantum algorithms like Grover's algorithm [Gro96], Farhi *et al.*'s algorithm [FGGS99], Høyer *et al.*'s algorithm [HNS01] etc. satisfy this condition. For classical querying, the state vector on the data qubits is $|0\rangle$, independent of the probe number. For Grover and Farhi *et al.*, the state vector on the data qubit is $(|0\rangle - |1\rangle)/\sqrt{2}$, independent of the probe number. For Høyer *et al.*, the state vector on the data qubit is $|0\rangle$ for some probe numbers, and $(|0\rangle - |1\rangle)/\sqrt{2}$ for the other probe numbers.

1.2.1 Static membership in the quantum bit probe model

In this thesis, we study the static membership problem in the *quantum bit probe* model, which is the quantum cell probe model with cell size w equal to one. We show tradeoffs between storage space and the number of probes for exact quantum bit probe schemes and lower bounds on the storage space for ϵ-error quantum bit probe schemes making a given number of probes. 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. Thus, our quantum lower bounds almost match the appropriate classical upper bounds.

Our investigations into the quantum bit probe complexity of set membership 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 (in particular, bounds on the sizes of r-cover-free families [NW94, EFF85, DR82]), seem to be powerless in giving the results in the quantum model. Instead, 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].

For classical deterministic query algorithms, Buhrman et al. [BMRV00] showed 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 4 *Suppose there exists an exact quantum bit probe scheme for storing subsets S of size at most n from a universe of size m that uses s bits of storage and answers membership queries 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 with exact quantum querying. 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 with exact quantum querying—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 with exact quantum querying, 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 use $O(nt(m/n)^{2/(t+1)})$ 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 the above result. 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 result 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, but always answers 'No' for a query x that is not a member of S.

We also give a simplified linear algebraic proof of the above theorem for deterministic and positive error classical bit probe schemes. This theorem is in fact stronger than the tradeoff results known previously for such schemes.

In the classical setting, 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 5 *For any $p \geq 1$ and $n/m < \epsilon < 2^{-3p}$, suppose there is a quantum bit probe scheme with two-sided error ϵ which stores subsets of size at most n from a universe of size m and answers membership queries 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. 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.

We also improve the lower bound in the result above on the space requirement of ϵ-error bit probe schemes for the static membership problem making p probes, when the query schemes are classical randomised.

Result 6 *Let $p \geq 1$, $18^{-p} > \epsilon > 1/m^{1/3}$ and $m^{1/3} > 18n$. Define $\delta \stackrel{\Delta}{=} \epsilon^{1/p}$. Any two-sided ϵ-error classical randomised scheme which stores subsets of size at most n from a universe of size m and answers membership queries using at most p bit probes must use space*

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

These results are joint work with Jaikumar Radhakrishnan and S.Venkatesh [RSV00a].

1.2.2 Static membership in the implicit storage quantum cell probe model

In this thesis, we generalise the $\Omega(\log n)$ lower bound of Yao on the number of probes required in any classical deterministic cell probe solution to the static membership problem with implicit storage schemes, to the quantum setting. Consider the problem of storing a subset S of size at most n of the universe $[m]$ in a table with q cells, so that membership queries can be answered efficiently. We restrict the storage scheme to be *implicit*, using at most p 'pointer values'. A 'pointer value' is a member of a set of size p (the set of 'pointers') disjoint from the universe. The term implicit means that the storage scheme can store either a 'pointer value' or a member of S in a cell. In particular, the storage scheme is not allowed to store an element of the universe which is not a member of S. The query algorithm answers membership queries by performing t (general) quantum cell probes. We call such schemes (p, q, t) *implicit storage quantum cell probe schemes*

Result 7 *For every n, p, q, there exists an $N(n, p, q)$ such that for all $m \geq N(n, p, q)$, the following holds: Consider any bounded error (p, q, t) implicit storage quantum cell probe scheme for the static membership problem with universe size m and size of the stored subset at most n. Then the quantum query scheme must make $t = \Omega(\log n)$ probes.*

This result is joint work with S.Venkatesh [SV01].

1.2.3 Static predecessor in the address-only quantum cell probe model

In this thesis, we also study the static predecessor problem. However, our lower bounds are not in the most general quantum cell probe model, but in a restricted version viz. the *address-only* quantum cell probe model. To show the lower bound for the static predecessor problem in the address-only quantum cell probe model, we use a connection between quantum cell probe schemes for static data structure problems and two-party quantum communication complexity. This connection similar to that in Miltersen, Nisan, Safra and Wigderson [MNSW98], who exploited it in the classical setting. Using this connection, we can convert an address-only quantum cell probe solution for the predecessor problem into a particular kind of quantum communication game. We then use a round elimination lemma

in quantum communication complexity to show lower bounds on the rounds complexity of this game. Using this approach, we prove the following theorem.

Result 8 *Suppose we have a* $(n^{O(1)}, (\log m)^{O(1)}, t)$ *bounded error quantum address-only cell probe solution to the static predecessor problem, where the universe size is* m *and the subset size is at most* n*. Then the number of queries* t *is at least* $\Omega\left(\frac{\log\log m}{\log\log\log m}\right)$ *as a function of* m*, and at least* $\Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$ *as a function of* n*.*

Since our *address-only* quantum cell probe model subsumes the classical cell probe model with randomised query schemes, our lower bound for the static predecessor problem also holds in this classical randomised setting. This improves the previous lower bound $\Omega(\sqrt{\log\log m})$ as a function of m and $\Omega(\log^{1/3} n)$ as a function of n for this setting, shown by Miltersen, Nisan, Safra and Wigderson [MNSW98]. Beame and Fich [BF99] have shown an upper bound matching our lower bound up to constant factors, which uses $n^{O(1)}$ cells of storage of word size $O(\log m)$ bits. In fact, both the storage and the query schemes are classical deterministic in Beame and Fich's solution. In the classical deterministic cell probe model, Beame and Fich show a lower bound of $t = \Omega\left(\frac{\log\log m}{\log\log\log m}\right)$ as a function of m for $(n^{O(1)}, 2^{(\log m)^{1-\Omega(1)}}, t)$ cell probe schemes, and a lower bound of $t = \Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$ as a function of n for $(n^{O(1)}, (\log m)^{O(1)}, t)$ cell probe schemes. But Beame and Fich's lower bound proof breaks down if the query scheme is randomised. Our result thus shows that the upper bound scheme of Beame and Fich is optimal all the way up to the bounded error address-only quantum cell probe model. Also, our proof is substantially simpler than that of Beame and Fich.

This result is joint work with S.Venkatesh [SV01].

1.3 The two-party quantum communication model

Classical communication complexity aims at studying the number of (classical) bits of communication that the components of a communication system need to exchange to perform certain tasks. Yao [Yao79] defined a very simple model for studying communication as a resource in the classical setting—the *two-party (classical) communication model*. In this model, there are two parties, Alice and Bob, and their task is to evaluate a function $f(x, y)$, where x is Alice's input and y is Bob's input. The computation of $f(x, y)$ is done according to a *(classical) communication protocol* P. During the execution of the protocol, the two parties alternately send messages as strings of bits. The protocol P is a set of rules specifying the player who starts the protocol, the player whose turn it is to send a message (based on the communication so far), what the players send (based on their inputs and the communication so far) and when a run terminates. At the end of the run, the last recipient of a message announces the output of the protocol. If the action of Alice is entirely a function of x and the communication which she has seen so far, and the same holds for the case of Bob, the protocol is called (classical) deterministic. The communication complexity of

a deterministic protocol P is the number of bits exchanged by the two parties in protocol P for the worst case input (x, y). A deterministic communication protocol for function f always outputs the correct value $f(x, y)$, given the input x to Alice and the input y to Bob. The *deterministic communication complexity* of f is the communication complexity of the best classical deterministic protocol computing f.

We can strengthen the two-party deterministic model by allowing Alice and Bob to 'toss coins' during the execution of the communication protocol. We assume that the coin tosses are done in 'public', that is, the action of Alice is a functions of x, the communication which she has seen so far, and the 'public coin tosses', and the same holds for Bob. We allow the protocol to make errors. A *public coins randomised protocol* for function f outputs the correct answer $f(x, y)$, when Alice is given x and Bob is given y, with probability at least $2/3$. The communication complexity of protocol P means the worst-case complexity, over every input (x, y) and coin toss sequence. The *randomised communication complexity* of f is the communication complexity of the best public coins randomised protocol computing f. Similar definitions can be given for *private coins randomised protocols*, where the coin tosses are done in 'private'.

The two-party classical communication model has been extensively studied in the past, and a rich theory has been built on it. For a comprehensive introduction, see the book by Kushilevitz and Nisan [KN96].

We consider the following *round elimination problem* in communication complexity. Suppose $f : E \times F \to G$ is a function. In the communication game corresponding to f, Alice gets a string $x \in E$, Bob gets a string $y \in F$, and they have to compute $f(x, y)$. In the communication game $f^{(n)}$, Alice gets n strings $x_1, \ldots, x_n \in E$; Bob gets an integer $i \in [n]$, a string $y \in F$, and a copy of the strings $x_1, \ldots, x_{i-1}$. Their aim is to communicate and compute $f(x_i, y)$. Suppose a protocol for $f^{(n)}$ is given where Alice starts, and her first message is a bits long, where a is much smaller than n. Intuitively, it would seem that since Alice does not know i, the first round of communication cannot give much information about x_i, and thus, would not be very useful to Bob. The *round elimination lemma* of Miltersen, Nisan, Safra and Wigderson [MNSW98] for classical communication complexity justifies this intuition. It says, informally speaking, that a public coins randomised protocol P for $f^{(n)}$ with t rounds of communication and Alice starting, gives rise to a public coins randomised protocol Q for f with $t-1$ rounds of communication and Bob starting, and the message complexity and error probability of Q are comparable to those of P. Moreover, we show that this is true even if Bob also gets copies of $x_1, \ldots, x_{i-1}$, a case which is needed in many applications of the round elimination lemma, for example, in proving lower bounds for many static data structure problems in the classical setting. In fact, Miltersen *et al.* [MNSW98] exploit the round elimination lemma in various ways to prove lower bounds for the static predecessor and other static data structure problems. They also use it to prove lower bounds for some communication complexity problems.

To study communication as a resource in quantum computation, Yao [Yao93] defined the *two-party quantum communication model*, similar to the the two-party classical communication model. Let E, F, G be arbitrary finite sets and $f : E \times F \to G$ be a function. There are two players Alice and Bob, who hold qubits. When the communication game

starts, Alice holds $|x\rangle$ where $x \in E$ together with some ancilla qubits in the state $|0\rangle$, and Bob holds $|y\rangle$ where $y \in F$ together with some ancilla qubits in the state $|0\rangle$. Thus the qubits of Alice and Bob are initially in computational basis states, and the initial superposition is simply $|x\rangle_A|0\rangle_A|y\rangle_B|0\rangle_B$. Here the subscripts denote the ownership of the qubits by Alice and Bob. The players take turns to communicate to compute $f(x, y)$. Suppose it is Alice's turn. Alice can make an arbitrary unitary transformation on her qubits and then send one or more qubits to Bob. Sending qubits does not change the overall superposition, but rather changes the ownership of the qubits, allowing Bob to apply his next unitary transformation on his original qubits plus the newly received qubits. At the end of the protocol, the last recipient of qubits performs a measurement on the qubits in her possession to output an answer. We say a quantum protocol computes f with ϵ-error in the worst case, if for any input $(x, y) \in E \times F$, the probability that the protocol outputs the correct result $f(x, y)$ is greater than $1-\epsilon$. The term 'bounded error quantum protocol' means that $\epsilon = 1/3$.

We require that Alice and Bob make a secure copy of their inputs before beginning the protocol. This is possible since the inputs to Alice and Bob are in computational basis states. Thus, without loss of generality, the input qubits of Alice and Bob are never sent as messages, their state remains unchanged throughout the protocol, and they are never measured i.e. some work qubits are measured to determine the result of the protocol. We call such protocols *secure*. We will assume henceforth that all our protocols are secure.

To state our round elimination lemma in quantum communication, we have to define the concept of a *safe* quantum communication protocol.

Definition 1.1 (Safe quantum protocol) *By a $[t, c, l_1, \ldots, l_t]^A$ ($[t, c, l_1, \ldots, l_t]^B$) safe quantum protocol, we mean a secure quantum protocol where Alice (Bob) starts the communication, the first message is $l_1 + c$ qubits long, the ith message, for $i \geq 2$, is l_i qubits long, and the communication goes on for t rounds. We think of the first message as having two parts: the 'main part' which is l_1 qubits long, and the 'safe overhead part' which is c qubits long. The density matrix of the 'safe overhead' is independent of the inputs to Alice and Bob.*

For the round elimination lemma, we also need to define the concept of a quantum protocol with *public coins*. Intuitively, a public coin quantum protocol is a probability distribution over finitely many (*coinless*) quantum protocols. We shall henceforth call the standard definition of a quantum protocol as *coinless*. Our definition is similar to the classical scenario, where a randomised protocol with public coins is a probability distribution over finitely many deterministic protocols. We note however, that our definition of a public coin quantum protocol is *not* the same as that of a quantum protocol with prior entanglement, which has been studied previously (see e.g. [CvDNT98]). Our definition is weaker, in that it does not allow the unitary transformations of Alice and Bob to alter the 'public coin'.

Definition 1.2 (Public coin quantum protocol) *In a quantum protocol with a public coin, there is, before the start of the protocol, a quantum state called a* public coin*, of*

the form $\sum_c \sqrt{p_c}|c\rangle_A|c\rangle_B$*, where the subscripts denote ownership of qubits by Alice and Bob,* p_c *are finitely many non-negative real numbers and* $\sum_c p_c = 1$*. Alice and Bob make (entangled) copies of their respective halves of the public coin using* CNOT *gates before commencing the protocol. The unitary transformations of Alice and Bob during the protocol do not touch the public coin. The public coin is never measured, nor is it ever sent as a message.*

Hence, one can think of the public coin quantum protocol to be a probability distribution, with probability p_c, over finitely many coinless quantum protocols indexed by the coin basis states $|c\rangle$. A *safe public coin* quantum protocol is similarly defined as a probability distribution over finitely many safe coinless quantum protocols.

1.3.1 Round elimination lemmas in quantum and classical communication

We prove a round elimination lemma for quantum communication complexity in this thesis. This result can be viewed as a quantum analogue of the round elimination lemma of Miltersen, Nisan, Safra and Wigderson [MNSW98] for classical communication complexity. Our quantum round elimination lemma is in fact stronger (!) than the classical round elimination lemma of [MNSW98], and it allows us to show a quantum lower bound for the static predecessor problem matching Beame and Fich's upper bound, which the classical round elimination lemma of [MNSW98] was unable to do. The quantum round elimination lemma can be used to prove similar lower bounds for many other static data structure problems in the address-only quantum cell probe model. It also finds applications to various problems in quantum communication complexity (e.g. the 'greater-than' problem), which are interesting on their own. Our quantum round elimination lemma is proved using quantum information theoretic techniques, and builds on the work of Klauck *et al.* [KNTZ01].

Result 9 *Suppose* $f : E \times F \to G$ *is a function. Suppose the communication game* $f^{(n)}$ *has a* $[t, c, l_1, \ldots, l_t]^A$ *safe public coin quantum protocol with worst case error less than* δ*. Then there is a* $[t-1, c+l_1, l_2, \ldots, l_t]^B$ *safe public coin quantum protocol for* f *with worst case error less than* $\epsilon \stackrel{\Delta}{=} \delta + (4l_1 \ln 2/n)^{1/4}$*.*

In the classical setting, we can refine our information theoretic techniques to prove an even stronger round elimination lemma for classical communication complexity.

Result 10 *Suppose* $f : E \times F \to G$ *is a function. Suppose the communication game* $f^{(n)}$ *has a* $[t, 0, l_1, \ldots, l_t]^A$ *public coin classical randomised protocol with worst case error less than* δ*. Then there is a* $[t-1, 0, l_2, \ldots, l_t]^B$ *public coin classical randomised protocol for* f *with worst case error less than* $\epsilon \stackrel{\Delta}{=} \delta + (1/2)(2l_1 \ln 2/n)^{1/2}$*.*

These results are joint work with S.Venkatesh [SV01].

1.3.2 Rounds versus communication tradeoffs for the 'greater-than' problem

As an application of our round elimination lemmas, we prove rounds versus communication tradeoffs for the 'greater-than' problem. In the 'greater-than' problem GT_n, Alice is given $x \in \{0,1\}^n$, Bob is given $y \in \{0,1\}^n$, and they have to communicate and decide whether $x > y$ (treating x, y as integers).

Result 11 *The t round bounded error quantum (classical randomised) communication complexity of GT_n is $\Omega(n^{1/t}t^{-3})$ ($\Omega(n^{1/t}t^{-2})$).*

There exists a bounded error classical randomised protocol for GT_n using t rounds of communication and having a complexity of $O(n^{1/t}\log n)$. Hence, for a constant number of rounds, our quantum lower bound matches the classical upper bound to within logarithmic factors. For one round quantum protocols, our result implies an $\Omega(n)$ lower bound for GT_n (which is optimal to within constant factors), improving upon the previous $\Omega(n/\log n)$ lower bound of Klauck [Kla00]. No rounds versus communication tradeoff for this problem, for more than one round, was known earlier in the quantum setting. For classical randomised protocols, Miltersen *et al.* [MNSW98] showed a lower bound of $\Omega(n^{1/t}2^{-O(t)})$ using their round elimination lemma. If the number of rounds is unbounded, then there is a classical randomised protocol for GT_n using $O(\log n)$ rounds of communication and having a complexity of $O(\log n)$ [Nis93]. An $\Omega(\log n)$ lower bound for the bounded error quantum communication complexity of GT_n (irrespective of the number of rounds) follows from Kremer's result [Kre95] that the bounded error quantum communication complexity of a function is lower bounded (up to constant factors) by the logarithm of the one round (classical) deterministic communication complexity.

These results are joint work with S.Venkatesh [SV01].

1.4 Organisation of the thesis

In Chapter 2, we present our results on the computation of $S_n^2(X)$ using $\Sigma\Pi\Sigma$ arithmetic circuits. We talk about our results on the static membership problem in the quantum bit probe model, and in the quantum cell probe model with implicit storage schemes, in Chapter 3. A complete proof of a weaker lower bound in the implicit storage quantum cell probe model can be found in the appendix. We then discuss the earlier round elimination based approach of Miltersen *et al.* [MNSW98], as well as our improved round elimination based approach, to the static predecessor problem in the classical setting, in Chapter 4. In Chapter 5, we prove our quantum round elimination lemma, and use it to prove a lower bound for predecessor in the address-only quantum cell probe model. This chapter also contains an application of the quantum round elimination lemma to the communication complexity of the 'greater-than' problem. To avoid congesting Chapters 4 and 5, the proofs of some technical lemmas in those chapters have been moved to the appendix. We end with a brief conclusion and a list of some open problems in Chapter 6.

Chapter 2

Depth-3 arithmetic circuits for $S_n^2(X)$

In this chapter, we present our results on computing $S_n^2(X)$ using $\Sigma\Pi\Sigma$ arithmetic circuits (defined in Section 1.1 over various fields. We first recall Graham and Pollack's theorem [GP72] on covering the complete graph on n vertices by complete bipartite graphs, such that each edge is covered exactly once. We then state the connections between the Graham-Pollack problem and computing $S_n^2(X)$ in the $\Sigma\Pi\Sigma$ model, and after that, go on to prove our bounds on computing $S_n^2(X)$ in this model.

The main new results in this chapter are

- For infinitely many odd and even n, $\lceil n/2 \rceil$ complete bipartite graphs are necessary and sufficient to cover each edge of the complete graph on n vertices an odd number of times (Theorem 2.2, Corollary 2.2 and Theorem 2.8). A similar result also holds for the number of multiplication gates required to compute $S_n^2(X)$ over the field GF(2), using $\Sigma\Pi\Sigma$ arithmetic circuits (Theorems 2.3 and 2.8).
- For any odd prime p, for infinitely many odd and even n, $\lceil n/2 \rceil$ complete bipartite graphs are sufficient to cover each edge of the complete graph on n vertices 1 mod p times (Theorem 2.4).
- For all n, $\lceil n/2 \rceil$ multiplication gates are necessary and sufficient to compute $S_n^2(X)$ over complex numbers, using $\Sigma\Pi\Sigma$ arithmetic circuits (Theorems 2.5 and 2.9). Similar, but weaker, results hold for computing $S_n^2(X)$ over finite fields of odd characteristic (Theorems 2.6, 2.7 and 2.10).

2.1 The Graham-Pollack theorem

Let K_n denote the complete graph on n vertices. By a *decomposition* of K_n, we mean a set $\{G_1, G_2, \ldots, G_r\}$ of subgraphs of K_n such that

1. Each G_i is a complete bipartite graph (on some subset of the vertex set of K_n); and
2. Each edge of K_n appears in precisely one of the G_i's.

It is easy to see that there is such a decomposition of the complete graph with $n-1$ complete bipartite graphs. Graham and Pollack [GP72] showed that this is tight.

Theorem *If $\{G_1, G_2, \ldots, G_r\}$ is a decomposition of K_n, then $r \geq n-1$.*

The original proof of this theorem, and other proofs discovered since then [dCH89, Pec84, Tve82], used algebraic reasoning in one form or another; no combinatorial proof of this fact is known.

One of the goals of this work is to obtain extensions of this theorem. To better motivate the problems we study, we first present a proof of this theorem. This will also help us explain how algebraic reasoning enters the picture. Consider polynomials in variables $X = X_1, X_2, \ldots, X_n$ with rational coefficients. Let

$$\begin{aligned} S_n^2(X) &\triangleq \sum_{1 \leq i < j \leq n} X_i X_j; \\ T_n^2(X) &\triangleq \sum_{i=1}^{n} X_i^2. \end{aligned}$$

Then, we can reformulate the question as follows. What is the smallest r for which there exist sets $L_i, R_i \subseteq [n], L_i \cap R_i = \emptyset$, for $i = 1, 2, \ldots, r$, such that

$$S_n^2(X) = \sum_{i=1}^{r} (\sum_{j \in L_i} X_j) \times (\sum_{j \in R_i} X_j) \tag{2.1}$$

Notice that the two sums in the product on the right are homogeneous linear forms i.e. linear forms in $X_1, \ldots, X_n$ with constant term 0. One may generalise this question, and ask: What is the smallest r for which there exist homogeneous linear forms $L_i(X), R_i(X)$ for $i = 1, 2 \ldots, r$, such that

$$S_n^2(X) = \sum_{i=1}^{r} L_i(X) R_i(X) \tag{2.2}$$

Tverberg [Tve82] gave the following elegant argument to show that r must be at least $n-1$. Observe that $T_n^2(X) = (\sum_{i=1}^{n} X_i)^2 - 2S_n^2(X)$. Thus, (2.2) implies

$$T_n^2(X) = (\sum_{i=1}^{n} X_i)^2 - 2 \sum_{i=1}^{r} L_i(X)\, R_i(X) \tag{2.3}$$

Now if r is less than $n-1$, then there exists a non-zero $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{Q}^n$ such that $L_i(\alpha) = 0$ for $i = 1, 2 \ldots, r$ and $\sum_{i=1}^{n} \alpha_i = 0$ (because at most $n-1$ homogeneous equations in n variables always have a non-zero solution). Under this assignment to the variables, the right hand side of (2.3) is zero but the left hand side is not.

With this introduction to the Graham-Pollack theorem and its proof, we are now ready to state the questions we consider in this chapter. Observe that the lower bound for r in (2.2) depended crucially on the field being $\mathbb{Q}$, and there are two main difficulties

in generalising it to other fields. First, over fields of characteristic two, the relationship between $S_n^2(X)$ and $T_n^2(X)$ does not hold, for we cannot divide by 2. Second, even if we are not working over fields of characteristic two, $T_n^2(X)$ can vanish at some non-zero points. Equations similar to (2.2) have been studied in the past in at least two different contexts viz. covering a complete graph by complete bipartite graphs such that each edge is covered an odd number of times (the *odd cover problem*), and depth–3 arithmetic circuits for $S_n^2(X)$. In the following sections, we will study some of these questions in detail.

2.2 At a glance: The bounds for computing $S_n^2(X)$

We study the computation of $S_n^2(X)$ in three different flavours of $\Sigma\Pi\Sigma$ arithmetic circuits.

1. *The graph model:* This is the weakest model. Here, the linear forms $L_i(X)$ and $R_i(X)$ (see equation (2.2) above) must correspond to bipartite graphs; that is, all coefficients must be 1 (or 0), no variable can appear in both L_i and R_i (with coefficient 1), and no constant term is allowed in these linear forms. This is the setting for the Graham-Pollack theorem and its generalisations.

2. *The homogeneous model:* Here the linear forms are required to be homogeneous, that is, no constant term is allowed in them. However, any element from the field is allowed as a coefficient in the linear forms. This model was studied by Nisan and Wigderson [NW96], using the method of partial derivatives.

3. *The inhomogeneous model:* This is the most general model; there is no restriction on the coefficients or the constant term.

We show our upper bounds in the graph and the homogeneous model; our lower bounds hold even in the stronger inhomogeneous model. We juxtapose our results against the previously known results and also briefly mention the proof technique used, highlighting our contribution. Note that the previous lower bounds were for the homogeneous circuit model only, and were proved using the method of partial derivatives [NW96] (but see also the rank arguments of Babai and Frankl [BF92] for the graph model).

The notation $\exists^\infty n$ used below means *'for infinitely many n'* and the notation $\forall n$ means *'for all n'*.

In the rest of the chapter, the *odd cover problem* shall refer to the problem of covering a complete graph by complete bipartite graphs such that each edge is covered an odd number of times. For an odd prime p, the 1 mod p *cover problem* shall refer to the problem of covering a complete graph by complete bipartite graphs such that each edge is covered 1 mod p times. Observe that the odd cover problem corresponds to the graph model of $\Sigma\Pi\Sigma$ circuits over GF(2), and the 1 mod p cover problem corresponds to the graph model of $\Sigma\Pi\Sigma$ circuits over GF(p) (p an odd prime).

2.2.1 The odd cover problem and computing $S^2_n(X)$ over GF(2)

Bounds:

	Our Bounds			Previous Bounds	
	Upper Bounds		Lower Bounds	Upper Bounds	Lower Bounds
	Graph	Hom.	Inhom.	Graph	Hom.
$n \equiv 0 \bmod 4$	$\frac{n}{2}\exists^\infty n$	$\frac{n}{2}\exists^\infty n$	$\frac{n}{2}\forall n$	$n-1\forall n$	$\frac{n}{2}\forall n$
$n \equiv 2 \bmod 4$	$\frac{n}{2}\exists^\infty n$	$\frac{n}{2}\exists^\infty n$	$\frac{n}{2}\forall n$	$n-1\forall n$	$\frac{n}{2}\forall n$
$n \equiv 3 \bmod 4$	$\lceil\frac{n}{2}\rceil\ \exists^\infty n$	$\lceil\frac{n}{2}\rceil\ \exists^\infty n$	$\lceil\frac{n}{2}\rceil\ \forall n$	$n-1\forall n$	$\lfloor\frac{n}{2}\rfloor\ \forall n$
$n \equiv 1 \bmod 4$	$\lceil\frac{n}{2}\rceil\ \exists^\infty n$	$\lfloor\frac{n}{2}\rfloor\ \exists^\infty n$	$\lfloor\frac{n}{2}\rfloor\ \forall n$	$n-1\forall n$	$\lfloor\frac{n}{2}\rfloor\ \forall n$

Table 2.1: Bounds for the odd cover problem and computing $S^2_n(X)$ over GF(2).

Proof Methods. For the upper bound in the graph model, we restrict our attention to a class of schemes, which we call *pairs constructions*, for constructing odd covers of K_n . We relate the pairs construction to the existence of certain kinds of *good* matrices. We then give two different constructions of *good* matrices. The first construction is based on *conference matrices*, which are related to *Hadamard matrices*. The second construction is based on *symmetric designs*, and uses some elementary properties about quadratic residues. The first construction gives optimal odd covers for infinitely many n of the form 0 mod 4; the second gives optimal odd covers for infinitely many n of the form 2 mod 4. We get $\lceil\frac{n}{2}\rceil$ sized odd covers for infinitely many n of the forms $n = 1, 3 \bmod 4$ from odd covers of K_{n+1} of optimal size.

The $\lfloor\frac{n}{2}\rfloor$ upper bound in the homogeneous model for $n \equiv 1 \bmod 4$ is got by locally transforming a homogeneous circuit computing $S^2_{n-1}(X)$ using $\frac{n-1}{2}$ multiplication gates to a homogeneous circuit computing $S^2_n(X)$ using the same number of multiplication gates.

For the lower bound, we use the method of substitution used by Shpilka and Wigderson [SW99], and subsequently refined by Shpilka [Shp01]. However, the proof is not a straightforward application of earlier methods. Technical difficulties arise because we are working over GF(2) and not over fields of characteristic zero. Almost all the earlier lower bound proofs used partial derivatives in some way or the other. Over GF(2), most of these approaches fail to work. Thus, we have to exploit the method of substitution in ways which do not use partial derivatives.

In fact, we place the method of substitution in a general framework and recast it to obtain a family of equations. We then exploit the family of equations depending upon the field in question, to obtain different lower bounds for different fields.

2.2.2 1 mod p cover problem, p an odd prime

Bounds:

	Our Bounds	Previous Bounds	
	Upper Bounds Graph	Upper Bounds Graph	Lower Bounds Hom.
n even	$\frac{n}{2}\exists^{\infty}n$	$n-1\forall n$	$\frac{n}{2}\forall n$
n odd	$\left\lceil\frac{n}{2}\right\rceil\exists^{\infty}n$	$n-1\forall n$	$\left\lfloor\frac{n}{2}\right\rfloor\forall n$

Table 2.2: Bounds for the 1 mod p cover problem.

Proof Methods. The upper bound follows by a *pairs construction* argument (refer Section 2.2.1). We reduce the problem of existence of a pairs construction to the existence of certain kinds of matrices *good for p*. By a modification of the *symmetric designs* construction (refer Section 2.2.1), we construct an infinite family of matrices *good for p*. This suffices to show the upper bounds for the 1 mod p cover problem. We use the same lower bounds as those known earlier for homogeneous circuits.

2.2.3 Computing $S^2_n(X)$ over $\mathbb{C}$

Bounds:

	Our Bounds		Previous Bounds	
	Upper Bounds Hom.	Lower Bounds Inhom.	Upper Bounds Hom.	Lower Bounds Hom.
$\forall n$	$\left\lceil\frac{n}{2}\right\rceil$	$\left\lceil\frac{n}{2}\right\rceil$	$\left\lceil\frac{n+1}{2}\right\rceil$	$\left\lceil\frac{n}{2}\right\rceil$

Table 2.3: Bounds for computing $S^2_n(X)$ over $\mathbb{C}$.

Proof Methods. For the upper bound, we reformulate the algebraic problem and arrive at a suitable bilinear form. Then, if the notion of "distance" between vectors is defined using this bilinear form, the problem reduces to finding suitably spaced vectors with complex coordinates. We then show the existence of such a suitably spaced family of vectors. The proof has a geometric flavour.

For the lower bound, we now use the general framework mentioned in Section 2.2.1. This time however, the way we exploit the family of equations is very different; in particular,

we view the constraints geometrically and arrive at a (different) bilinear form. Then, if the notion of "distance" between vectors is defined using this bilinear form, the problem reduces to placing a certain number of points on a sphere of a certain radius such that all the points are equidistant with a certain common distance. We then show that such a placement of points is impossible.

2.2.4 Computing $S_n^2(X)$ over GF(p^r), p odd

Bounds:

Field		Our Bounds		Previous Bounds	
		Upper Bnds. Hom.	Lower Bnds. Inhom.	Upper Bnds. Hom.	Lower Bnds. Hom.
GF(p^r) r even $p > 3$	n even	$\frac{n}{2}\forall n$	$\frac{n}{2}\forall n$	$\frac{n}{2}+1\forall n$	$\frac{n}{2}\forall n$
	n odd	$\lceil\frac{n}{2}\rceil\ \forall n$	$\lceil\frac{n}{2}\rceil\ \exists^\infty n$ $\lfloor\frac{n}{2}\rfloor\ \forall n$	$\lceil\frac{n}{2}\rceil\ \forall n$	$\lceil\frac{n}{2}\rceil\ \exists^\infty n$ $\lfloor\frac{n}{2}\rfloor\ \forall n$
GF(3^r) r even	n even	$\frac{n}{2}\forall n$	$\frac{n}{2}\forall n$	$\frac{n}{2}+1\forall n$	$\frac{n}{2}\forall n$
	n odd	$\lceil\frac{n}{2}\rceil\ \forall n$	$\lfloor\frac{n}{2}\rfloor\ \forall n$	$\lceil\frac{n}{2}\rceil\ \forall n$	$\lceil\frac{n}{2}\rceil\ \exists^\infty n$ $\lfloor\frac{n}{2}\rfloor\ \forall n$
GF(p^r) r odd $p \equiv 1 \bmod 4$	n even	$\frac{n}{2}\exists^\infty n$	$\frac{n}{2}\forall n$	$\frac{n}{2}+1\forall n$	$\frac{n}{2}\forall n$
	n odd	$\lceil\frac{n}{2}\rceil\ \forall n$	$\lceil\frac{n}{2}\rceil\ \exists^\infty n$ $\lfloor\frac{n}{2}\rfloor\ \forall n$	$\lceil\frac{n}{2}\rceil\ \forall n$	$\lceil\frac{n}{2}\rceil\ \exists^\infty n$ $\lfloor\frac{n}{2}\rfloor\ \forall n$
GF(p^r) r odd $p \equiv 3 \bmod 4$	n even	$\frac{n}{2}\exists^\infty n$	$\frac{n}{2}\forall n$	$n-1\forall n$	$\frac{n}{2}\forall n$
	n odd	$\lceil\frac{n}{2}\rceil\ \exists^\infty n$	$\lfloor\frac{n}{2}\rfloor\ \forall n$	$n-1\forall n$	$\lfloor\frac{n}{2}\rfloor\ \forall n$

Table 2.4: Bounds for computing $S_n^2(X)$ over GF(p^r), p an odd prime.

Proof Methods. For GF(p^r), r even and GF$(p^r), p \equiv 1 \bmod 4$, r odd, the proof of the upper bound is very similar to our upper bound proof for complex numbers. The technical

reason behind this is that these fields have square roots of -1.

The upper bound for $\mathrm{GF}(p^r), p \equiv 3 \bmod 4$, r odd, follows from our upper bound for the $1 \bmod p$ cover problem. Since these fields do not have square roots of -1, we cannot mimic the upper bound arguments for complex numbers for these fields.

The proof of the lower bound for finite fields of odd characteristic is similar to the lower bound proof for complex numbers, though, because of technical difficulties, the results are not as tight for some values of n, as they were in the case of complex numbers.

2.2.5 Computing $S_n^2(X)$ over $\mathbb{R}$ and $\mathbb{Q}$

Bounds:

	Our Bounds		Previous Bounds	
	Upper Bounds Graph	Lower Bounds Inhom.	Upper Bounds Graph	Lower Bounds Hom.
$\forall n$	$n-1$	$n-1$	$n-1$	$n-1$

Table 2.5: Bounds for computing $S_n^2(X)$ over $\mathbb{R}$ and $\mathbb{Q}$.

Proof Methods. In this case, we show that the trivial upper bound of $n-1$ is tight even for inhomogeneous circuits. The proof of the Graham-Pollack theorem works only for homogeneous circuits. To extend the result to inhomogeneous circuits, we need to use the method of substitution. The result is relatively straightforward once the problem is placed in this framework. We state the result for completeness.

2.3 Upper bounds

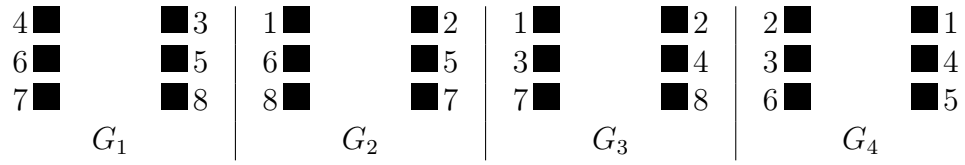
2.3.1 The odd cover problem and computing $S_n^2(X)$ over $\mathbf{GF}(2)$

In this section, we will show that there is an odd cover of K_{2n} by n complete bipartite graphs whenever there exists a $n \times n$ matrix satisfying certain properties. We describe a particular scheme for producing an odd cover of K_{2n}, which we call a *pairs construction*. We express the requirements for a pairs construction in the language of matrices, and then give sufficient conditions for a matrix to encode a pairs construction. We call a matrix satisfying these sufficient conditions a *good* matrix.

We want to cover the edges of K_{2n} with n complete bipartite graphs such that each edge is covered an odd number of times. A complete bipartite graph is fully described by specifying its two colour classes A and B. Partition the vertex set $[2n]$ (of K_{2n}) into ordered pairs $(1,2),(3,4),\ldots,(2n-1,2n)$. In a *pairs construction* of an odd cover of K_{2n}, if one element of a pair does not participate in a complete bipartite graph G in the

odd cover decomposition, then the other element of the pair does not participate in G either, and also, both the elements of a pair do not appear in the same colour class in G. Hence, to describe a complete bipartite graph G in a pairs construction of an odd cover decomposition, it suffices to specify for each pair $(2i-1, 2i)$, whether the pair participates in the bipartite graph, and when it does, whether $2i$ appears in colour class A or B. We specify the n complete bipartite graphs in the odd cover decomposition by a $n \times n$ matrix $\mathbf{M}$ with entries in $\{-1, 0, 1\}$. The rows of the matrix are indexed by pairs; the ith row is for the pair $(2i-1, 2i)$. The columns are indexed by the complete bipartite graphs of the odd cover decomposition. If $\mathbf{M}_{ij} = 0$, the pair $(2i-1, 2i)$ does not participate in the jth bipartite graph G_j; if $\mathbf{M}_{ij} = 1$, $2i$ appears in colour class B of G_j; if $\mathbf{M}_{ij} = -1$, $2i$ appears in colour class A of G_j.

$$\mathbf{M} = \begin{array}{c} \\ (1,2) \\ (3,4) \\ (5,6) \\ (7,8) \end{array} \begin{array}{c} \begin{array}{cccc} G_1 & G_2 & G_3 & G_4 \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & -1 \\ -1 & 0 & 1 & 1 \\ -1 & -1 & 0 & -1 \\ 1 & -1 & 1 & 0 \end{bmatrix} \end{array}$$



The matrix $\mathbf{M}$ describes a pairs construction of an odd cover of K_8 by complete bipartite graphs G_1, G_2, G_3, G_4.

Figure 2.1: An example of a pairs construction.

We now identify properties of the matrix $\mathbf{M}$ which ensure that the complete bipartite graphs arising from it form an odd cover of K_{2n}.

Definition 2.1 *A $n \times n$ matrix with entries from $\{-1, 0, 1\}$ is* good *if it satisfies the following conditions:*

1. *In every row, the number of non-zero entries is odd.*
2. *For every pair of distinct rows, the number of columns where they both have non-zero entries is congruent to* 2 mod 4*.*
3. *Any two distinct rows are orthogonal over the integers.*

Lemma 2.1 *If an $n \times n$ matrix is good, then the n complete bipartite graphs that arise from it form an odd cover of K_{2n}.*

Proof: Since the number of non-zero entries in a row is odd, the number of times the corresponding edge $\{2i-1, 2i\}$ is covered is odd. Next, consider edges whose vertices come from different pairs: say, the edge $\{1, 3\}$. We need to show that the number of bipartite graphs where 1 and 3 are placed on opposite sides is odd. Consider the rows of the matrix corresponding to pairs $(1, 2)$ and $(3, 4)$. Since these rows are orthogonal over the integers, the number of times 1 appears on the opposite side of 3 must be equal to the number of times 1 appears on the opposite side of 4. Since the number of columns where both rows have non-zero entries is congruent to 2 mod 4, the number of times 1 appears on the opposite side of 3 (as well as the number of times 1 appears on the opposite side of 4) must be odd. Thus, given a good matrix, we can construct n complete bipartite graphs covering each edge of K_{2n} an odd number of times. ■

Thus, to obtain odd covers, it is enough to construct good matrices. We now give two methods for constructing such matrices.

Construction 1: Skew symmetric conference matrices

A *Hadamard matrix* $\mathbf{H}_n$ is an $n \times n$ matrix with entries in $\{-1, 1\}$ such that $\mathbf{H}_n\mathbf{H}_n^T = n\mathbf{I}_n$, where $\mathbf{I}_n$ is the $n \times n$ identity matrix. A *conference matrix* $\mathbf{C}_n$ is an $n \times n$ matrix, with 0's on the diagonal and $-1, +1$ elsewhere, such that $\mathbf{C}_n\mathbf{C}_n^T = (n-1)\mathbf{I}_n$. The following fact can be verified easily.

Lemma 2.2 *$n \times n$ conference matrices, where $n \equiv 0 \bmod 4$, are good matrices.*

Skew symmetric conference matrices can be obtained from *skew Hadamard matrices*. A skew Hadamard matrix is defined as a Hadamard matrix that one gets by adding the identity matrix to a skew symmetric conference matrix. Several constructions of skew Hadamard matrices can be found in [Hal86, p. 247]. In particular, the following theorem is proved there.

Theorem 2.1 *There is a skew Hadamard matrix of order n if $n = 2^t k_1 \cdots k_s$, where $n \equiv 0 \bmod 4$, each $k_i \equiv 0 \bmod 4$ and each k_i is of the form $p^r + 1$, p an odd prime.*

Corollary 2.1 *There is a good matrix of order n if n satisfies the conditions in the above theorem. Note that the conditions hold for infinitely many n.*

As an illustrative example, we show the existence of skew Hadamard matrices $\mathbf{F}_n$ when n is a power of 2. To do this, we modify the well-known recursive construction for Hadamard matrices. For $n = 2$, set $(\mathbf{F}_2)_{21} = -1$ and the rest of the entries 1. Suppose now that we have constructed $\mathbf{F}_n$. To construct $\mathbf{F}_{2n}$, place a copy of $\mathbf{F}_n$ in the top left corner, a copy of $-\mathbf{F}_n$ in the bottom left corner, and copies of $\mathbf{F}_n^T$ in the top right and bottom right corners. It is easy to check that $\mathbf{F}_{2n}$ so constructed is skew Hadamard. In fact, the matrix $\mathbf{M}$ in Figure 2.1 is nothing but $\mathbf{F}_4 - \mathbf{I}_4$.

Construction 2: Symmetric designs

The matrices $\mathbf{M}$ that we now construct are based on a well-known construction for symmetric designs. These matrices are not conference matrices; in fact, they have more than one zero in every row.

Let q be a prime power congruent to 3 mod 4. Let $\mathbb{F} = \mathrm{GF}(q)$ be the finite field of q elements. Index the rows of $\mathbf{M}$ with lines and the columns with points of the projective 2-space over $\mathbb{F}$. That is, the projective points and lines are the one dimensional and two dimensional subspaces respectively, of $\mathbb{F}^3$. A projective point is represented by a vector in $\mathbb{F}^3$ (out of $q-1$ possible representatives) in the one dimensional subspace corresponding to it. A projective line is also represented by a vector in $\mathbb{F}^3$ (out of $q-1$ possible representatives). The representative for a projective line can be thought of as a 'normal vector' to the two dimensional subspace corresponding to it. We associate with each projective line L a linear form on the vector space $\mathbb{F}^3$, given by $L(w) = v^T w$, where $w \in \mathbb{F}^3$ and v is the chosen representative for L. For a projective line L and a projective point Q, let $L(Q) \triangleq L(w)$, where w is the chosen representative for Q. Now the matrix $\mathbf{M}$ is defined as follows. If $L(Q) = 0$ (i.e. projective point Q lies on projective line L), we set $\mathbf{M}_{L,Q} = 0$; if $L(Q)$ is a (non-zero) square in $\mathbb{F}$, set $\mathbf{M}_{L,Q} = 1$; otherwise, set $\mathbf{M}_{L,Q} = -1$.

We now check that $\mathbf{M}$ is a good matrix. M is a $n \times n$ matrix, where $n = q^2+q+1$, q a prime power congruent to 3 mod 4. The number of non-zero entries per row is $q^2+q+1-(q+1) = q^2$, which is odd. The number of columns where two distinct rows have non-zero entries is $q^2+q+1-2(q+1)+1 = q^2-q$. This number is 2 mod 4 since $q \equiv 3 \bmod 4$. Recall that in the projective 2-space over $\mathrm{GF}(q)$, each line contains $q+1$ points, and two distinct lines intersect in a single point. Now we only need to check that any two distinct rows (corresponding to distinct projective lines L, L') are orthogonal over the integers. We first observe that the following equality holds over the integers.

$$\sum_P \eta(L(P))\eta(L'(P)) = \frac{1}{q-1} \sum_{v \neq (0,0,0)} \eta(L(v))\eta(L'(v)) \tag{2.4}$$

where,

$$\eta(x) \;=\; \begin{cases} 0 & \text{if } x = 0 \\ 1 & \text{if } x \text{ is a (non-zero) square} \\ -1 & \text{if } x \text{ is not a square} \end{cases}.$$

[The first sum is over all points P of the projective 2-space. The second is over all non-zero triples v in $\mathbb{F}^3$.] The equality holds because if we take two non-zero triples u and $w = \alpha u$ ($\alpha \neq 0$) corresponding to the same projective point, then

$$\begin{aligned} \eta(L(w))\eta(L'(w)) &= \eta(L(\alpha u))\eta(L'(\alpha u)) \\ &= \eta(\alpha L(u))\eta(\alpha L'(u)) \\ &= \eta(\alpha)\eta(L(u))\eta(\alpha)\eta(L'(u)) \\ &= \eta(L(u))\eta(L'(u)) \end{aligned}$$

Now consider the sum on the right hand side of (2.4). We have

$$\sum_{v \neq (0,0,0)} \eta(L(v))\eta(L'(v)) \;=\; \sum_{a,b \in \mathbb{F}; a,b \neq 0} \; \sum_{\substack{v : L(v)=a, L'(v)=b \\ v \neq (0,0,0)}} \eta(a)\eta(b)$$

The linear forms corresponding to two distinct projective lines are linearly independent; i.e., L and L' are linearly independent. Hence, for every pair (a, b) in the sum above, there are exactly q triples v such that $L(v) = a$ and $L'(v) = b$. Thus,

$$\begin{aligned}\sum_{v\neq(0,0,0)} \eta(L(v))\eta(L'(v)) &= q \cdot \sum_{a,b\in\mathbb{F};\ a,b\neq 0} \eta(a)\eta(b) \\ &= q \cdot \sum_{a,b\in\mathbb{F};\ a,b\neq 0} \eta(ab) \\ &= q(q-1) \cdot \sum_{c\in\mathbb{F};\ c\neq 0} \eta(c) \\ &= 0\end{aligned}$$

The last equality holds because there are exactly $(q-1)/2$ squares and the same number of non–squares in $\mathbb{F} - \{0\}$. We conclude that the left hand side of (2.4) is 0; hence, the rows corresponding to distinct projective lines are orthogonal over the integers.

We have thus proved the following lemma.

Lemma 2.3 *If $q \equiv 3 \bmod 4$ is a prime power then there is a good matrix of order q^2+q+1. Note that infinitely many such q exist.*

We can now easily prove the following theorem and its corollary.

Theorem 2.2 *For infinitely many $n \equiv 0, 2 \bmod 4$ we have an odd cover of K_n using $\frac{n}{2}$ complete bipartite graphs.*

Proof: We use $\frac{n}{2} \times \frac{n}{2}$ *good* matrices to construct an odd cover of K_n using $\frac{n}{2}$ complete bipartite graphs(see Lemma 2.1). For infinitely many $n \equiv 0 \bmod 4$, we can use the *good* matrices of Corollary 2.1. For infinitely many $n \equiv 2 \bmod 4$, we can use the *good* matrices of Lemma 2.3. ∎

Corollary 2.2 *For infinitely many $n \equiv 1, 3 \bmod 4$ we have an odd cover of K_n using $\lceil \frac{n}{2} \rceil$ complete bipartite graphs.*

Proof: For odd n, any odd cover of K_{n+1} using $\frac{n+1}{2}$ complete bipartite graphs gives us an odd cover for K_n too. The corollary now follows from the above theorem. ∎

We also prove the following lemma, which allows us to construct homogeneous $\Sigma\Pi\Sigma$ circuits for $S^2_n(X)$ with $\lfloor \frac{n}{2} \rfloor$ multiplication gates, for infinitely many $n \equiv 1 \bmod 4$.

Lemma 2.4 *If $S^2_n(X), n \equiv 0 \bmod 4$, can be computed over GF(2) by a homogeneous $\Sigma\Pi\Sigma$ circuit using $\frac{n}{2}$ multiplication gates, then $S^2_{n+1}(X)$ can be computed over GF(2) by a homogeneous $\Sigma\Pi\Sigma$ circuit using $\frac{n}{2}$ multiplication gates.*

Proof: Consider a homogeneous circuit over GF(2)

$$\sum_{i=1}^{r} L_i(X_1, \ldots, X_n) R_i(X_1, \ldots, X_n) \tag{2.5}$$

for $S^2_n(X_1, \ldots, X_n)$, $n \equiv 0 \bmod 4$, where $r = \frac{n}{2}$. Define for $1 \leq i \leq r$, homogeneous linear forms $L'_i(X_1, \ldots, X_{n+1})$, $R'_i(X_1, \ldots, X_{n+1})$ over GF(2) as follows.

$$\begin{array}{rcll} L'_i(X_1, \ldots, X_{n+1}) & \stackrel{\Delta}{=} & L_i(X_1, \ldots, X_n) + X_{n+1} & \text{if } L_i \text{ has an odd number of terms} \\ & \stackrel{\Delta}{=} & L_i(X_1, \ldots, X_n) & \text{otherwise} \\ R'_i(X_1, \ldots, X_{n+1}) & \stackrel{\Delta}{=} & R_i(X_1, \ldots, X_n) + X_{n+1} & \text{if } R_i \text{ has an odd number of terms} \\ & \stackrel{\Delta}{=} & R_i(X_1, \ldots, X_n) & \text{otherwise} \end{array}$$

We have the following equality over GF(2).

Claim

$$S^2_{n+1}(X_1, \ldots, X_{n+1}) = \sum_{i=1}^{r} L'_i(X_1, \ldots, X_{n+1}) R'_i(X_1, \ldots, X_{n+1})$$

Proof: Define homogeneous linear forms over $\mathbb{Z}$, $L''_i(X_1, \ldots, X_{n+1})$, $R''_i(X_1, \ldots, X_{n+1})$, for $1 \leq i \leq r$, as follows.

$$\begin{array}{rcl} L''_i(X_1, \ldots, X_{n+1}) & \stackrel{\Delta}{=} & L_i(X_1, \ldots, X_n) + a_i X_{n+1} \\ R''_i(X_1, \ldots, X_{n+1}) & \stackrel{\Delta}{=} & R_i(X_1, \ldots, X_n) + b_i X_{n+1} \end{array}$$

where a_i, b_i denote the number of (non-zero) terms in L_i, R_i respectively. Consider the following formula over $\mathbb{Z}$.

$$\sum_{i=1}^{r} L''_i(X_1, \ldots, X_{n+1}) R''_i(X_1, \ldots, X_{n+1}) \tag{2.6}$$

Let $c_{jk}, 1 \leq j \leq k \leq n$ denote the coefficient of $X_j X_k$ in (2.5), treating (2.5) as a formula over $\mathbb{Z}$ instead of over GF(2). Since formula (2.5) computes $S^2_n(X)$ over GF(2), $c_{jk}, 1 \leq j < k \leq n$ are odd, and $c_{jj}, 1 \leq j \leq n$ are even. Let $c''_{jk}, 1 \leq j \leq k \leq n+1$ denote the coefficient of $X_j X_k$ in (2.6) (note that c''_{jk} is an integer). For $1 \leq j \leq k \leq n$, $c''_{jk} = c_{jk}$. We will now show that $c''_{j,n+1}, 1 \leq j \leq n$ are odd, and $c''_{n+1,n+1}$ is even. This suffices to prove the claim, since $L''_i \equiv L'_i \bmod 2$ and $R''_i \equiv R'_i \bmod 2$.

For any $1 \leq j \leq n$, it can be easily checked that

$$\begin{array}{rcl} c''_{j,n+1} & = & \displaystyle\sum_{\substack{k:1 \leq k \leq n \\ k \neq j}} c_{jk} + 2c_{jj} \\ & \equiv & \displaystyle\sum_{\substack{k:1 \leq k \leq n \\ k \neq j}} 1 + 0 \pmod 2 \\ & \equiv & 1 \pmod 2 \end{array}$$

The last equivalence follows from the fact that, for any fixed j, the number of monomials $X_j X_k, 1 \leq k \leq n, k \neq j$ is odd, since n is even.

$$c''_{n+1,n+1} = \sum_{1 \leq j \leq k \leq n} c_{jk}$$

$$\begin{aligned} &= \sum_{1\le j<k\le n} c_{jk} + \sum_{1\le j\le n} c_{jj} \\ &\equiv \left(\sum_{1\le j<k\le n} 1 + \sum_{1\le j\le n} 0\right) \pmod 2 \\ &\equiv 0 \pmod 2 \end{aligned}$$

The last equivalence follows from the fact that the number of monomials $X_jX_k, 1 \le j < k \le n$ is even, since $n \equiv 0 \bmod 4$.

Hence the claim is proved. ■

The lemma now follows from the above claim. ■

We can now prove the following theorem.

Theorem 2.3 *For infinitely many $n \equiv 0, 2, 3 \bmod 4$ we have homogeneous $\Sigma\Pi\Sigma$ circuits computing $S_n^2(X)$ over GF(2) using $\lceil \frac{n}{2} \rceil$ multiplication gates. For infinitely many $n \equiv 1 \bmod 4$ we can compute $S_n^2(X)$ over GF(2) using homogeneous $\Sigma\Pi\Sigma$ circuits having $\lfloor \frac{n}{2} \rfloor$ multiplication gates.*

Proof: The first part of the theorem follows from Theorem 2.2 and Corollary 2.2. To prove the second part, consider a homogeneous circuit for $S_{n-1}^2(X_1, \ldots, X_{n-1})$, $n \equiv 1 \bmod 4$, using $r = \frac{n-1}{2}$ multiplication gates. Such circuits exist for infinitely many $n \equiv 1 \bmod 4$ by the first part of the theorem. We now invoke Lemma 2.4 to complete the proof. ■

2.3.2 $1 \bmod p$ cover problem, p an odd prime

In this subsection we will in fact show, for any odd number p (not necessarily prime), that there is a $1 \bmod p$ cover of K_{2n} by n complete bipartite graphs whenever there exists an $n \times n$ matrix *good for p* (defined below). Also, from a $1 \bmod p$ cover of K_{2n+2} by $n+1$ bipartite graphs, we get a $1 \bmod p$ cover of K_{2n+1} by $n+1$ bipartite graphs. We note that the skew Hadamard matrix construction of Section 2.3.1 does not generalise to give us matrices *good for p*, when p is odd.

Definition 2.2 *Let p be an odd number. A matrix with entries from $\{-1, 0, 1\}$ is called a* good matrix for p *if it satisfies the following conditions:*

1. *In every row, the number of non-zero entries is* $1 \bmod p$.
2. *For every pair of distinct rows, the number of columns where they both have non-zero entries is congruent to* $2 \bmod 2p$.
3. *Any two distinct rows are orthogonal over the integers.*

Lemma 2.5 *Let p be an odd number. If an $n \times n$ matrix is good for p, then the n complete bipartite graphs that arise from it form a $1 \bmod p$ cover of K_{2n}. If $n = q^2 + q + 1$ where q is a prime power and $q \equiv -1 \bmod 2p$, then an $n \times n$ good matrix for p exists. Note that infinitely many such q exist, by a result of Dirichlet.*

Proof: The proof of the fact that an $n \times n$ good matrix for p gives us a 1 mod p cover of K_{2n} by n complete bipartite graphs, is similar to the proof of Lemma 2.1. The construction of an $n \times n$ good matrix for p when n is of the given form is similar to the symmetric designs construction of Section 2.3.1. ∎

From the lemma, we can now prove the following theorem.

Theorem 2.4 *Given an odd number p, for infinitely many odd and even n, we have a 1 mod p cover of K_n using $\lceil \frac{n}{2} \rceil$ bipartite graphs.*

2.3.3 Fields of characteristic different from 2

Now we give the proofs for the upper bounds in the homogeneous circuit model for computing $S^2_n(X)$ over various fields of characteristic different from 2. We start by proving two lemmas.

Lemma 2.6 *$S^2_{2k+1}(X_1, \ldots, X_{2k+1})$ can be computed by a homogeneous $\Sigma\Pi\Sigma$ circuit using $k+1$ multiplication gates over any field of characteristic not equal to 2 which has square roots of -1.*

Proof: This result has been observed implicitly by Shpilka [Shp01]. We give a proof here for completeness. Let i denote a square root of -1.

$$\begin{aligned}
& S^2_{2k+1}(X_1, \ldots, X_{2k+1}) \\
&= \frac{1}{2}((\sum_{j=1}^{2k+1} X_j)^2 - \sum_{j=1}^{2k+1} X_j^2) \\
&= \frac{1}{2}(((\sum_{j=1}^{2k+1} X_j)^2 - X_1^2) - \sum_{j=2}^{2k+1} X_j^2) \\
&= \frac{1}{2}((\sum_{j=2}^{2k+1} X_j)(2X_1 + \sum_{j=2}^{2k+1} X_j) - \sum_{j=1}^{k}(X_{2j}^2 + X_{2j+1}^2)) \\
&= \frac{1}{2}((\sum_{j=2}^{2k+1} X_j)(2X_1 + \sum_{j=2}^{2k+1} X_j) - \sum_{j=1}^{k}(X_{2j} + iX_{2j+1})(X_{2j} - iX_{2j+1}))
\end{aligned}$$

This shows that $S^2_{2k+1}(X_1, \ldots, X_{2k+1})$ can be done with $k+1$ multiplication gates. ∎

Lemma 2.7 *$S^2_{2k}(X_1, \ldots, X_{2k})$ can be computed by a homogeneous $\Sigma\Pi\Sigma$ circuit using k multiplication gates over any field $\mathbb{F}$ of characteristic not equal to 2 which has square roots of -1, 2 and $2k-1$.*

Proof: Let $a_m(X_1, \ldots, X_{2k})$ and $b_m(X_1, \ldots, X_{2k})$ denote the two homogeneous linear forms feeding into the mth multiplication gate, $1 \leq m \leq k$. Let

$$\left.\begin{array}{rcl} a_m(X_1, \ldots, X_{2k}) & \triangleq & \sum_{n=1}^{2k} a_{mn} X_{mn} \\ b_m(X_1, \ldots, X_{2k}) & \triangleq & \sum_{n=1}^{2k} b_{mn} X_{mn} \end{array}\right\} \quad 1 \leq m \leq k$$

Since the circuit computes $S^2_{2k}(X_1, \ldots, X_{2k})$, equating the coefficients of $X_j^2, 1 \leq j \leq 2k$ we get

$$\sum_{m=1}^{k} a_{mj} b_{mj} = 0 \qquad 1 \leq j \leq 2k$$

Since the characteristic is not equal to 2, we can get an equivalent equation by multiplying both sides by 2.

$$\sum_{m=1}^{k} (a_{mj} b_{mj} + a_{mj} b_{mj}) = 0 \qquad 1 \leq j \leq 2k \tag{2.7}$$

Equating the coefficients of $X_j X_l, 1 \leq j < l \leq 2k$ we get

$$\sum_{m=1}^{k} (a_{mj} b_{ml} + a_{ml} b_{mj}) = 1 \qquad 1 \leq j < l \leq 2k \tag{2.8}$$

Let us define vectors $y_j \in \mathbb{F}^{2k}, 1 \leq j \leq 2k$ as follows

$$y_j^T \triangleq (a_{1j}, b_{1j}, a_{2j}, b_{2j}, \ldots, a_{kj}, b_{kj})$$

We can write (2.7), (2.8) in a succinct matrix form as

$$\left.\begin{array}{rcll} y_j^T \mathbf{A} y_j & = & 0 & 1 \leq j \leq 2k \\ y_j^T \mathbf{A} y_l & = & 1 & 1 \leq j < l \leq 2k \end{array}\right\} \tag{2.9}$$

where the $2k \times 2k$ matrix $\mathbf{A}$ consists of k blocks of the 2×2 matrix

$$M \triangleq \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

arranged along the diagonal. M has two eigenvalues 1 and -1, with corresponding eigenvectors $u_1^T = (1, 1)$ and $u_{-1}^T = (1, -1)$ (note that $1 \neq -1$ in $\mathbb{F}$). It will be convenient to scale these vectors to obtain alternate eigenvectors $v_1^T = \frac{1}{\sqrt{2}}(1, 1)$ and $v_{-1}^T = \frac{1}{\sqrt{2}}(i, -i)$, where i denotes a square root of -1 in $\mathbb{F}$ (note that $2 \neq 0$ in $\mathbb{F}$ and 2 and -1 have square roots in $\mathbb{F}$). Now,

$$v_1^T M v_1 = v_{-1}^T M v_{-1} = 1$$
$$v_1^T M v_{-1} = 0$$

The 2×2 matrix

$$N \triangleq \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix}$$

is the change of basis matrix for going from the basis $\{v_1, v_{-1}\}$ of $\mathbb{F}^2$ to the standard basis $\{(1, 0)^T, (0, 1)^T\}$ of $\mathbb{F}^2$. We define another $2k \times 2k$ matrix $\mathbf{B}$, which consists of k blocks of the 2×2 matrix N arranged along the diagonal. $\mathbf{B}$ is a change of basis matrix from a

basis of $\mathbb{F}^{2k}$ consisting of eigenvectors of $\mathbf{A}$, to the standard basis of $\mathbb{F}^{2k}$. If $z_j, 1 \leq j \leq 2k$ are the representations of the vectors $y_j, 1 \leq j \leq 2k$ in the eigenbasis of $\mathbf{A}$, then

$$y_j = \mathbf{B}z_j \qquad 1 \leq j \leq 2k$$

Since

$$\mathbf{B}^T\mathbf{A}\mathbf{B} = \mathbf{I}_{2k}$$

where $\mathbf{I}_{2k}$ is the $2k \times 2k$ identity matrix, (2.9) now becomes

$$\begin{aligned} z_j^T z_j &= 0 \qquad 1 \leq j \leq 2k \\ z_j^T z_l &= 1 \qquad 1 \leq j < l \leq 2k \end{aligned}$$

We can write a set of equations equivalent to the above as follows (since $2 \neq 0$ in $\mathbb{F}$)

$$\left.\begin{aligned} z_j^T z_j &= 0 \qquad 1 \leq j \leq 2k \\ (z_j - z_l)^T(z_j - z_l) &= -2 \qquad 1 \leq j < l \leq 2k \end{aligned}\right\} \tag{2.10}$$

The second equation above can be thought as finding vectors $z_j \in \mathbb{F}^{2k}, 1 \leq j \leq 2k$ such that the "distance" between any two of them is $\sqrt{-2}$. The following set of vectors meets this requirement

$$z'_j = ie_j \qquad 1 \leq j \leq 2k$$

where $e_j, 1 \leq j \leq 2k$ are the standard basis vectors in $\mathbb{F}^{2k}$. We now have to ensure that the "length" of each vector is 0. For this shift the origin to a point $p \triangleq (w, w, \ldots, w)$, where w will be determined later. Note that this operation does not change the "distance" between any pair of vectors. To determine w we have to solve the following equation

$$(i - w)^2 + (2k - 1)w^2 = 0$$

which can be solved whenever $2k - 1$ has a square root in the field. We now define

$$z_j \triangleq z'_j - p \qquad 1 \leq j \leq 2k$$

The vectors $z_j, 1 \leq j \leq 2k$ are a solution to (2.10) which in turn implies a solution to (2.9) which proves the existence of a homogeneous circuit for the polynomial $S^2_{2k}(X_1, \ldots, X_{2k})$ using k multiplication gates. ∎

Using Lemmas 2.6 and 2.7, we can prove our upper bound results for complex numbers and for finite fields of odd characteristic.

Complex numbers

Theorem 2.5 *$S^2_n(X_1, \ldots, X_n)$ can be computed by a homogeneous $\Sigma\Pi\Sigma$ circuit using $\lceil \frac{n}{2} \rceil$ multiplication gates over the field of complex numbers.*

Proof: Follows directly from Lemmas 2.6 and 2.7. ∎

GF(p^r), r even, p odd and GF(p^r), r odd, $p \equiv 1 \bmod 4$

Theorem 2.6 *Let p be an odd prime. $S^2_n(X)$ can be computed by a homogeneous $\Sigma\Pi\Sigma$ circuit using $\lceil \frac{n}{2} \rceil$ multiplication gates over GF(p^r), r even. Over GF(p^r), r odd, $p \equiv 1 \bmod 4$, $S^2_n(X)$ can be computed using $\lceil \frac{n}{2} \rceil$ multiplication gates if n is odd, $\frac{n}{2}$ multiplication gates for infinitely many even n, and $\frac{n}{2}+1$ multiplication gates for all even n.*

Proof: If $p \equiv 1 \bmod 4$ then -1 and 2 have square roots in GF(p) (see e.g. [NZM91, Chapter 3]). Hence using Lemmas 2.6 and 2.7, over GF(p^r), r odd, $p \equiv 1 \bmod 4$ $S^2_n(X)$ can be computed using $\lceil \frac{n}{2} \rceil$ multiplication gates if n is odd, and using $\frac{n}{2}$ multiplication gates for even n such that $n-1$ has a square root in GF(p^r), which holds for infinitely many even n. For all even n, $S^2_n(X)$ can be computed using $\frac{n}{2}+1$ multiplication gates by taking a circuit with that many gates for $S^2_{n+1}(X_1, \ldots, X_{n+1})$, and setting X_{n+1} to 0. Over GF(p^r), r even every element of GF(p) has a square root (see e.g. [Art91, Chapter 13]). Hence, using Lemmas 2.6 and 2.7 again, $S^2_n(X)$ can be computed using $\lceil \frac{n}{2} \rceil$ multiplication gates for all n. ∎

GF(p^r), r odd, $p \equiv 3 \bmod 4$

Theorem 2.7 *Let $p \equiv 3 \bmod 4$ be a prime. For infinitely many even and odd n, $S^2_n(X)$ can be computed by a homogeneous $\Sigma\Pi\Sigma$ circuit using $\lceil \frac{n}{2} \rceil$ multiplication gates over GF(p^r), r odd.*

Proof: Such fields do not have a square root of -1. Hence we cannot use either of the Lemmas 2.6 and 2.7. To get upper bounds of $\lceil \frac{n}{2} \rceil$ for infinitely many even and odd n, we have to make use of the fact that upper bounds for the $1 \bmod p$ cover problem (Theorem 2.4) give us upper bounds for computing $S^2_n(X)$ in the homogeneous circuit model. ∎

2.4 Lower bounds

2.4.1 Preliminaries

In this subsection, we develop a framework for proving lower bounds for computing $S^2_n(X)$ in the inhomogeneous $\Sigma\Pi\Sigma$ model, based on the method of substitution [SW99, Shp01]. Suppose that over a field $\mathbb{F}$

$$S^2_n(X) = \sum_{i=1}^{r} \prod_{j=1}^{s_i} L_{ij}(X) \tag{2.11}$$

where each $L_{ij}(X)$ is a linear form over $X_1, \ldots, X_n$, not necessarily homogeneous. We wish to show that r must be large. Following the proof of the Graham-Pollack theorem that was sketched in the introduction, we could try to force some of the L_{ij}'s to zero by setting the variables to appropriate field elements. There are two difficulties with this plan. First, since the L_{ij}'s are not necessarily homogeneous, we may not be able to set all of them to zero; we can do so if the linear forms have linearly independent homogeneous parts. The second difficulty arises from the nature of the underlying field: as remarked in the introduction, $S^2_n(X)$ might vanish on non-trivial subspaces of $\mathbb{F}^n$.

In this subsection, our goal is to first show that if r is small, then $S_n^2(X)$ must be zero over a linear subspace of $\mathbb{F}^n$ of large dimension. Similar observations have been used by Shpilka and Wigderson [SW99, Lemma 3.3] and Shpilka [Shp01, Claim 4.6]. Our second goal is to examine linear subspaces of $\mathbb{F}^n$ over which $S_n^2(X)$ is forced to be zero. We derive conditions on such subspaces, and relate them to the existence of a certain family of vectors. Later on, we will exploit these equations based on the field in question, and derive our lower bounds for r.

Goal 1: Obtaining the subspace.

Lemma 2.8 *If $S_n^2(X)$ can be written in the form of (2.11) over a field $\mathbb{F}$, then there exist homogeneous linear forms $\ell_1, \ell_2, \ldots, \ell_r$ in variables $X_1, X_2, \ldots, X_{n-r}$ such that*

$$S_n^2(X_1, X_2, \ldots, X_{n-r}, \ell_1, \ell_2, \ldots, \ell_r) = 0 \tag{2.12}$$

Proof: We implement the idea discussed at the beginning of Section 2.4.1. Given an expression of the form (2.11), we collect a maximal consistent set of equations of the form $L_{ij}(X) = 0$, with at most one equation for each i. We write these equations in the form

$$\mathbf{A}X = b \tag{2.13}$$

where $\mathbf{A}$ is an $r' \times n$ matrix and $b \in \mathbb{F}^{r'}$ for some $r' \leq r$. Since (2.13) has a solution, and the rank of $\mathbf{A}$ is at most r, there is an affine subspace of solutions Γ of dimension $n-r$ in $\mathbb{F}^n$. (If the actual solution set is an affine subspace of dimension greater than $n-r$, then we let Γ be an affine subspace of the solution space of dimension exactly $n-r$.) We can view this set of solutions as follows (see e.g. [Art91, Chapter 1]): there are $n-r$ 'free variables,' and the values of the remaining r variables are given by (possibly inhomogeneous) linear forms in these $n-r$ variables. Since $S_n^2(X)$ is symmetric, we may assume that the $n-r$ 'free variables' are $X_1, X_2, \ldots, X_{n-r}$; for $i = 1, 2, \ldots, r$, let $\tilde{\ell}_i$ be the (possibly inhomogeneous) linear form in $X_1, X_2, \ldots, X_{n-r}$ that determines the value of X_{n-r+i} once the values for $X_1, X_2, \ldots, X_{n-r}$ are fixed.

Observe that $S_n^2(X)$ is constant over Γ. To see this, consider the right hand side of (2.11). If for some i an L_{ij} participates in (2.13), then that product contributes zero to the sum. Otherwise, since the chosen set of equations is maximal, for this i, the homogeneous part of each L_{ij} is in the row span of the matrix $\mathbf{A}$. That is, once $\mathbf{A}X$ has been fixed to b, the homogeneous part, and hence the entire linear form, is fixed. We conclude that

$$S_n^2(X_1, X_2, \ldots, X_{n-r}, \tilde{\ell}_1, \tilde{\ell}_2, \ldots, \tilde{\ell}_r) = \textit{constant}$$

Now comparing the coefficients of monomials of degree two on both sides of the above equation, we see that

$$S_n^2(X_1, X_2, \ldots, X_{n-r}, \ell_1, \ell_2, \ldots, \ell_r) = 0$$

where ℓ_i is the homogeneous part of $\tilde{\ell}_i$. ∎

Goal 2: The nature of the subspace. Our goal now is to understand the algebraic structure of the coefficients that appear in the linear forms $\ell_1, \ell_2, \ldots, \ell_r$ promised by Lemma 2.8. Let $\ell_i = \sum_{j=1}^{n-r} \ell_{ij} X_j, \ell_{ij} \in \mathbb{F}$, and let $\mathbf{L}$ be the $r \times (n-r)$ matrix (ℓ_{ij}). Let $y_1, y_2, \ldots, y_{n-r} \in \mathbb{F}^r$ be the $n-r$ columns of $\mathbf{L}$. We will obtain conditions on the columns by computing the coefficients of monomials X_j^2 for $1 \leq j \leq n-r$, and $X_i X_j$ for $1 \leq i < j \leq n-r$, in equation (2.12). For X_j^2 $(1 \leq j \leq n-r)$, we obtain the following equation over $\mathbb{F}$.

$$\sum_{k=1}^{r} \ell_{kj} + \sum_{1 \leq k < k' \leq r} \ell_{kj} \ell_{k'j} = 0 \quad 1 \leq j \leq r \tag{2.14}$$

For monomials of the form $X_i X_j$ $(1 \leq i < j \leq n-r)$, we obtain the following equation over $\mathbb{F}$.

$$1 + \sum_{k=1}^{r} \ell_{ki} + \sum_{k=1}^{r} \ell_{kj} + \sum_{1 \leq k < k' \leq r} (\ell_{ki} \ell_{k'j} + \ell_{k'i} \ell_{kj}) = 0 \quad 1 \leq i < j \leq n-r \tag{2.15}$$

For a positive integer m, let $\mathbf{1}_m$ be the all 1's column vector and $\mathbf{0}_m$ be the all 0's column vector of dimension m. Let $\mathbf{U}_m$ be the $m \times m$ matrix with 1's above the diagonal and zero elsewhere. Let $\mathbf{J}_m$ be the $m \times m$ matrix with all 1's, and let $\mathbf{I}_m$ be the $m \times m$ identity matrix. Using this notation, we can rewrite (2.14) and (2.15) as follows.

$$\mathbf{1}_r^T y_j + y_j^T \mathbf{U}_r y_j = 0 \quad 1 \leq j \leq n-r \tag{2.16}$$

$$1 + \mathbf{1}_r^T y_i + \mathbf{1}_r^T y_j + y_i^T (\mathbf{J}_r - \mathbf{I}_r) y_j = 0 \quad 1 \leq i < j \leq n-r \tag{2.17}$$

If the characteristic of $\mathbb{F}$ is not two, we may rewrite (2.16) as

$$2\mathbf{1}_r^T y_j + y_j^T (\mathbf{J}_r - \mathbf{I}_r) y_j = 0 \quad 1 \leq j \leq n-r \tag{2.18}$$

With this, we are now ready to prove lower bounds. We will exploit (2.16), (2.17) and (2.18) (if the characteristic is not 2) to derive lower bounds for various fields.

2.4.2 Lower bounds for GF(2)

Let $\mathbb{Z}$ stand for the integers. For $y \in \mathbb{Z}^r$, let $|y|$ denote the number of odd components in y. For $y, y' \in \mathbb{Z}^r$, let $y \cdot y' \stackrel{\Delta}{=} \sum_{m=1}^{r} y_m y'_m$ be the dot product of y and y' over $\mathbb{Z}$.

Lemma 2.9 *Suppose $\ell_1, \ldots, \ell_r$ are homogeneous linear forms in variables $X_1, \ldots, X_{n-r}$ such that $S_n^2(X_1, \ldots, X_{n-r}, \ell_1, \ldots, \ell_r) = 0$ over GF(2). Then $r \geq \lfloor \frac{n}{2} \rfloor$. If $n \equiv 3 \bmod 4$, then $r \geq \lceil \frac{n}{2} \rceil$.*

Proof: We use the arguments of Section 2.4.1. If there exist homogeneous linear forms $\ell_1, \ldots, \ell_r$ over variables $X_1, \ldots, X_{n-r}$ so that $S_n^2(X_1, \ldots, X_{n-r}, \ell_1, \ldots, \ell_r) = 0$ over GF(2), we have, from (2.16) and (2.17), vectors $y_j \in \mathrm{GF}(2)^r, 1 \leq j \leq n-r$ such that the following

equations hold over $GF(2)$ (recall that J_r denotes the $r \times r$ all 1's matrix, and I_r denotes the $r \times r$ identity matrix).

$$\mathbf{1}_r^T y_j + y_j^T \mathbf{U}_r y_j = 0 \quad 1 \le j \le n-r \tag{2.19}$$

$$1 + \mathbf{1}_r^T y_i + \mathbf{1}_r^T y_j + y_i^T (\mathbf{J}_r - \mathbf{I}_r) y_j = 0 \quad 1 \le i < j \le n-r \tag{2.20}$$

Instead of thinking of the above equations as holding over $GF(2)$, it will help for this proof to treat the vectors y_j as elements of $\mathbb{Z}^r$ and the equations (2.19) and (2.20) as equivalences over the integers mod2.

By counting the number of odd components (i.e. 1's) on the left and right hand side of (2.19), we obtain

$$|y_j| + \binom{|y_j|}{2} \equiv 0 \pmod 2 \quad 1 \le j \le n-r$$

From this it follows that

$$|y_j| \equiv 0 \text{ or } 3 \pmod 4 \quad 1 \le j \le n-r \tag{2.21}$$

Since $y_i^T(\mathbf{J}_r - \mathbf{I}_r)y_j = |y_i|\,|y_j| - y_i \cdot y_j$ over $\mathbb{Z}$, by counting the number of odd components (i.e. 1's) on both sides of (2.20), we get

$$|y_i| + |y_j| + |y_i|\,|y_j| + y_i \cdot y_j \equiv 1 \pmod 2 \quad 1 \le i < j \le n-r$$

In other words,

$$y_i \cdot y_j \equiv (1 + |y_i|)(1 + |y_j|) \pmod 2 \quad 1 \le i < j \le n-r \tag{2.22}$$

Let $w_1, \ldots, w_s$ be the vectors among $y_1, \ldots, y_{n-r}$ with $|y_j|$ odd, and let $e_1, \ldots, e_t$ be the remaining $t = n - r - s$ vectors, with $|y_j|$ even.

Claim If $y_1, y_2, \ldots, y_{n-r}$ are not linearly independent over GF(2), then the only dependency over GF(2) among them is $\sum_{k=1}^{t} e_k = \mathbf{0}_r$. Also, in that case, t is odd.

Proof: Let

$$\sum_{i=1}^{s} \alpha_i w_i + \sum_{k=1}^{t} \beta_k e_k \equiv \mathbf{0}_r \pmod 2$$

In the above equation, we think of w_i, e_k as vectors in $\mathbb{Z}^r$, α_i, β_k as integers, and the equality as an equivalence over the integers mod2. We take dot products of the two sides above with w_i and conclude, using (2.22), that $\alpha_i \equiv 0 \bmod 2$, for $1 \le i \le s$. Similarly, taking dot products with e_k, we obtain the system of equations $(\mathbf{J}_t - \mathbf{I}_t)\beta \equiv \mathbf{0}_t \bmod 2$, where $\beta \in \mathbb{Z}^t$ and the kth component of β is β_k. If t is even, $(\mathbf{J}_t - \mathbf{I}_t)$ is full-rank over GF(2), so $\beta \equiv \mathbf{0}_t \bmod 2$. So the y_j's are linearly independent over GF(2), which is a contradiction.

Now, if the y_j's are not linearly independent, then t must be odd, and the only dependency among them corresponds to β such that $(\mathbf{J}_t - \mathbf{I}_t)\beta \equiv \mathbf{0}_t \bmod 2$. The only non–trivial solution mod2 for this equation is $\beta \equiv \mathbf{1}_t \bmod 2$. ∎

By the claim above, we see that there are at least $n-r-1$ linearly independent vectors over GF(2) among the y_j's. Since the y_j's are r-dimensional vectors, we get $r \geq n-r-1$ i.e. $r \geq \left\lfloor \frac{n}{2} \right\rfloor$. This proves the first part of the lemma.

To obtain a better bound for r when $n \equiv 3 \bmod 4$, we make better use of our equations, especially (2.21), which we have neglected so far. So suppose $n = 2r+1$ and $n \equiv 3 \bmod 4$. We shall derive a contradiction.

If $n = 2r+1$, then $n - r > r$, and since the y_j are r-dimensional vectors, y_j are not linearly independent over GF(2). Then by the claim above, t is odd, $\sum_{k=1}^{t} e_k \equiv \mathbf{0}_r \bmod 2$, and $w_1, \ldots, w_s, e_1, \ldots, e_{t-1}$ are linearly independent over GF(2). Since $s+t-1 = n-r-1 = r$, these vectors form a basis (over GF(2)) of the vector space $\mathrm{GF}(2)^r$; in particular $\mathbf{1}_r$ is in their span, that is

$$\sum_{i=1}^{s} \alpha_i w_i + \sum_{k=1}^{t-1} \beta_k e_k \equiv \mathbf{1}_r \pmod 2$$

for some $\alpha_i, \beta_k \in \mathbb{Z}$. Taking dot products with w_i and e_k, we conclude (using (2.22)) that $\alpha_i \equiv 1 \bmod 2$ for $1 \leq i \leq s$, and $(\mathbf{J}_{t-1} - \mathbf{I}_{t-1})\beta \equiv \mathbf{0}_{t-1} \bmod 2$, where $\beta \in \mathbb{Z}^{t-1}$ and the kth component of β is β_k. Since t is odd, $\mathbf{J}_{t-1} - \mathbf{I}_{t-1}$ is full rank over GF(2), and $\beta \equiv \mathbf{0}_{t-1} \bmod 2$. Thus

$$\sum_{i=1}^{s} w_i \equiv \mathbf{1}_r \pmod 2 \tag{2.23}$$

It is easy to verify that for all integer vectors y

$$|y| \equiv y \cdot y \pmod 4 \tag{2.24}$$

Using (2.23) and (2.24), $(\sum_{i=1}^{s} w_i) \cdot (\sum_{i=1}^{s} w_i) \equiv |\sum_{i=1}^{s} w_i| \equiv r \bmod 4$, that is

$$\sum_{i=1}^{s} w_i \cdot w_i + 2 \sum_{1 \leq i < j \leq s} w_i \cdot w_j \equiv r \pmod 4$$

By (2.21) and (2.24), $w_i \cdot w_i \equiv |w_i| \equiv 3 \bmod 4$, and by (2.22), $w_i \cdot w_j \equiv 0 \bmod 2$ for $i \neq j$. Thus

$$\sum_{i=1}^{s} 3 + \sum_{1 \leq i < j \leq s} 0 \equiv r \pmod 4$$

$$\Rightarrow 3s \equiv r \pmod 4 \tag{2.25}$$

Similarly, starting with $\sum_{k=1}^{t} e_k \equiv \mathbf{0}_r \bmod 2$ and using (2.24) we get, $(\sum_{k=1}^{t} e_k) \cdot (\sum_{k=1}^{t} e_k) \equiv |\sum_{k=1}^{t} e_k| \equiv 0 \bmod 4$, that is

$$\sum_{i=1}^{t} e_i \cdot e_i + 2 \sum_{1 \leq i < j \leq t} e_i \cdot e_j \equiv 0 \pmod 4$$

By (2.21) and (2.24), $e_i \cdot e_i \equiv 0 \bmod 4$, and by (2.22), $e_i \cdot e_j \equiv 1 \bmod 2$ for $i \neq j$. Thus

$$\sum_{i=1}^{t} 0 + \sum_{1 \leq i < j \leq t} 2 \equiv 0 \pmod 4$$

$$\Rightarrow \frac{t(t-1)}{2}\, 2 \equiv 0 \bmod 4$$

Since t is odd, we conclude that $t \equiv 1 \bmod 4$. But then, using (2.25),

$$n \equiv r + s + t \equiv 3s + s + 1 \equiv 1 \pmod 4$$

which is a contradiction.

Since $r \geq \lfloor \frac{n}{2} \rfloor$ holds for all n, we have shown that if $n \equiv 3 \bmod 4$, then $r \geq \lceil \frac{n}{2} \rceil$. ∎

Using Lemmas 2.8 and 2.9, we can now prove the following theorem.

Theorem 2.8 *Any (not necessarily homogeneous) ΣΠΣ circuit computing $S^2_n(X_1, \ldots, X_n)$ over GF(2) requires at least $\lceil \frac{n}{2} \rceil$ multiplication gates if $n \equiv 0, 2, 3 \bmod 4$, and at least $\lfloor \frac{n}{2} \rfloor$ multiplication gates if $n \equiv 1 \bmod 4$.*

2.4.3 Fields of characteristic different from 2

In this subsection, we give the proofs of our lower bounds for computing $S^2_n(X)$ using (not necessarily homogeneous) ΣΠΣ arithmetic circuits over various fields of characteristic different from 2. Lemma 2.10 proves an upper bound on the dimension of a subspace over which $S^2_{2k}(X_1, \ldots, X_{2k})$ vanishes. The proof uses Nisan and Wigderson's method of partial derivatives.

Lemma 2.10 *If $k \neq 0$ in the field $\mathbb{F}$ then $S^2_{2k}(X_1, \ldots, X_{k+1}, \ell_1, \ldots, \ell_{k-1}) \neq 0$ for any $k-1$ homogeneous linear forms $\ell_1, \ldots, \ell_{k-1}$ in the variables $X_1, \ldots, X_{k+1}$ over $\mathbb{F}$.*

Proof: This lemma is in fact a special case of a more general result due to Shpilka [Shp01]. We give a short proof of it here, which is essentially Shpilka's proof restricted to our special case. We have the identity

$$\begin{aligned} S^2_{2k}(X_1, \ldots, X_{k+1}, \ell_1, \ldots, \ell_{k-1}) &= S^2_{k+1}(X_1, \ldots, X_{k+1}) + \\ &\quad (X_1 + \cdots + X_{k+1})(\ell_1 + \cdots + \ell_{k-1}) + \\ &\quad S^2_{k-1}(\ell_1, \ldots, \ell_{k-1}) \end{aligned}$$

Assuming for the sake of contradiction that the left hand side of the above equation is zero, we get

$$\begin{aligned} &S^2_{k+1}(X_1, \ldots, X_{k+1}) \\ &\quad = -(X_1 + \cdots + X_{k+1})(\ell_1 + \cdots + \ell_{k-1}) - S^2_{k-1}(\ell_1, \ldots, \ell_{k-1}) \end{aligned}$$

We take the first order partial derivatives with respect to $X_1, \ldots, X_{k+1}$ of both the sides of the above equation. Since $k \neq 0$ in $\mathbb{F}$, the vector space spanned by the set of first-order partial derivatives of $S^2_{k+1}(X_1, \ldots, X_{k+1})$ is of dimension $k+1$. This follows from the fact that the matrix $\mathbf{J}_{k+1} - \mathbf{I}_{k+1}$ is of full rank if $k \neq 0$ in $\mathbb{F}$, where $\mathbf{J}_{k+1}$ is the $(k+1) \times (k+1)$ all 1's matrix and $\mathbf{I}_{k+1}$ is the $(k+1) \times (k+1)$ identity matrix. The vector space spanned by the first order partial derivatives of the right hand side of the above equation lies in the span of the linear forms $(X_1 + \cdots + X_{k+1})$ and $\ell_1, \ldots, \ell_{k-1}$. Hence its dimension is at most k, which results in a contradiction. This proves the lemma. ■

Lemma 2.11 also proves upper bounds on the dimension of a subspace over which $S^2_{2k}(X_1, \ldots, X_{2k})$ vanishes, but the proof does not use partial derivatives.

Lemma 2.11 *Suppose $k \neq -1$ in the field $\mathbb{F}$ and $\mathbb{F}$ is not of characteristic 2. Then $S^2_{2k}(X_1, \ldots, X_{k+1}, \ell_1, \ldots, \ell_{k-1}) \neq 0$ for any $k-1$ homogeneous linear forms $\ell_1, \ldots, \ell_{k-1}$ in the variables $X_1, \ldots, X_{k+1}$ over $\mathbb{F}$.*

Proof: Using the arguments of Section 2.4.1 (in particular (2.17) and (2.18)), we assume (using the notation of that section) for the sake of contradiction that there exist vectors $y_j \in \mathbb{F}^{k-1}, 1 \leq j \leq k+1$, such that the following equations hold (note that the characteristic of $\mathbb{F}$ is not 2).

$$\left.\begin{array}{rcll} \langle y_j, y_j \rangle + 2\mathbf{1}^T_{k-1} y_j & = & 0 & 1 \leq j \leq k+1 \\ \langle y_j, y_l \rangle + \mathbf{1}^T_{k-1} y_j + \mathbf{1}^T_{k-1} y_l & = & -1 & 1 \leq j < l \leq k+1 \end{array}\right\} \tag{2.26}$$

where $\langle v, w \rangle \stackrel{\Delta}{=} v^T(\mathbf{J}_{k-1} - \mathbf{I}_{k-1})w$ is a symmetric bilinear form on vectors in $\mathbb{F}^{k-1}$.

From the above equation, we get

$$\langle y_j - y_l, y_j - y_l \rangle = 2 \qquad 1 \leq j < l \leq k+1 \tag{2.27}$$

We can think of equation (2.27) as placing $k+1$ points with pairwise "distance" $\sqrt{2}$ in $\mathbb{F}^{k-1}$. We now show that if $k \neq -1$ in $\mathbb{F}$, this is impossible.

We have, for $1 < j < l \leq k+1$

$$\begin{array}{rcl} 2 & = & \langle y_j - y_l, y_j - y_l \rangle \qquad \ldots \text{u}sing\ (2.27) \\ & = & \langle (y_j - y_1) - (y_l - y_1), (y_j - y_1) - (y_l - y_1) \rangle \\ & = & \langle y_j - y_1, y_j - y_1 \rangle - 2\langle y_j - y_1, y_l - y_1 \rangle + \langle y_l - y_1, y_l - y_1 \rangle \\ & = & 2 + 2 - 2\langle y_j - y_1, y_l - y_1 \rangle \qquad \ldots \text{u}sing\ (2.27) \end{array}$$

Hence, since $2 \neq 0$ in $\mathbb{F}$,

$$\langle y_j - y_1, y_l - y_1 \rangle = 1 \qquad 1 < j < l \leq k+1 \tag{2.28}$$

Now define a $k \times k$ matrix $\mathbf{A}$ where

$$a_{jl} \stackrel{\Delta}{=} \langle y_{j+1} - y_1, y_{l+1} - y_1 \rangle \qquad 1 \leq j, l \leq k$$

Using (2.27) and (2.28), we see that the matrix $\mathbf{A}$ has 2's on the main diagonal and 1's in other places. Since $k \neq -1$ in $\mathbb{F}$, $\mathbf{A}$ is of full rank. This implies that the vectors $y_2 - y_1, y_3 - y_1, \ldots, y_{k+1} - y_1$ are linearly independent. In fact we have shown that the vectors $y_1, \ldots, y_{k+1}$ are affinely independent. Since these vectors lie in $\mathbb{F}^{k-1}$, we have arrived at a contradiction. Hence the lemma is proved. ∎

We can now prove the following lemma. This lemma allows us to prove lower bounds for computing $S_n^2(X)$ using (not necessarily homogeneous) $\Sigma\Pi\Sigma$ arithmetic circuits over $\mathbb{F}$ when $\mathbb{F}$ is not of characteristic 2 and n is even.

Lemma 2.12 *$S_{2k}^2(X_1, \ldots, X_{k+1}, \ell_1, \ldots, \ell_{k-1}) \neq 0$ for any $k-1$ homogeneous linear forms $\ell_1, \ldots, \ell_{k-1}$ in the variables $X_1, \ldots, X_{k+1}$ over a field $\mathbb{F}$, if $\mathbb{F}$ is not of characteristic 2.*

Proof: Follows from Lemmas 2.10 and 2.11. ∎

We also prove the following lemma. This lemma allows us to prove lower bounds for computing $S_n^2(X)$ using (not necessarily homogeneous) $\Sigma\Pi\Sigma$ arithmetic circuits over $\mathbb{F}$ when $\mathbb{F}$ is not of characteristic 2 and n is odd.

Lemma 2.13 *Suppose $k \neq 0, \pm 1$ in the field $\mathbb{F}$ and $\mathbb{F}$ is not of characteristic 2. Then $S_{2k+1}^2(X_1, \ldots, X_{k+1}, \ell_1, \ldots, \ell_k) \neq 0$ for any k homogeneous linear forms $\ell_1, \ldots, \ell_k$ in the variables $X_1, \ldots, X_{k+1}$ over $\mathbb{F}$.*

Proof: Using the arguments of Section 2.4.1 (in particular (2.17) and (2.18)), we assume (using the notation of that section) for the sake of contradiction that there exist vectors $y_j \in \mathbb{F}^k, 1 \leq j \leq k+1$, such that the following equations hold (note that the characteristic of $\mathbb{F}$ is not 2).

$$\left.\begin{array}{rcll} \langle y_j, y_j \rangle + 2\mathbf{1}_k^T y_j & = & 0 & 1 \leq j \leq k+1 \\ \langle y_j, y_l \rangle + \mathbf{1}_k^T y_j + \mathbf{1}_k^T y_l & = & -1 & 1 \leq j < l \leq k+1 \end{array}\right\} \tag{2.29}$$

where $\langle v, w \rangle \stackrel{\Delta}{=} v^T(\mathbf{J}_k - \mathbf{I}_k)w$ is a symmetric bilinear form on vectors in $\mathbb{F}^k$.

We can similarly show, as in the proof of Lemma 2.11, that the vectors $y_2 - y_1, y_3 - y_1, \ldots, y_{k+1} - y_1$ are linearly independent (since $k \neq -1$ and $2 \neq 0$ in $\mathbb{F}$). Also

$$\langle y_j - y_l, y_j - y_l \rangle = 2 \qquad 1 \leq j < l \leq k+1 \tag{2.30}$$

Since $k \neq 1$ in $\mathbb{F}$, let us define a vector $c \in \mathbb{F}^k$, $c \stackrel{\Delta}{=} \frac{-1}{k-1}\mathbf{1}_k$. Now $(\mathbf{J}_k - \mathbf{I}_k)c = -\mathbf{1}_k$ and $c^T(\mathbf{J}_k - \mathbf{I}_k)c = \frac{k}{k-1}$. Hence we have, for $1 \leq j \leq k+1$

$$\begin{aligned} \langle y_j - c, y_j - c \rangle &= \langle y_j, y_j \rangle - 2\langle y_j, c \rangle + \langle c, c \rangle \\ &= \langle y_j, y_j \rangle + 2\mathbf{1}_k^T y_j + \frac{k}{k-1} \end{aligned}$$

Using the first equation in (2.29) and above equation, we get the following equation

$$\langle y_j - c, y_j - c \rangle = \frac{k}{k-1} \qquad 1 \leq j \leq k+1 \tag{2.31}$$

Shifting the origin to the vector c and using (2.30) and (2.31) we have (using the same letters $y_j, 1 \leq j \leq k+1$ to denote the new vectors)

$$\left.\begin{array}{rcll} \langle y_j, y_j \rangle & = & \frac{k}{k-1} & 1 \leq j \leq k+1 \\ \langle y_j - y_l, y_j - y_l \rangle & = & 2 & 1 \leq j < l \leq k+1 \end{array}\right\} \tag{2.32}$$

We can think of equations (2.32) as placing $k+1$ points of pairwise "distance" $\sqrt{2}$ on the surface of a sphere of "radius" $\sqrt{\frac{k}{k-1}}$ in $\mathbb{F}^k$. We now show that if $k \neq 0, \pm 1$ in $\mathbb{F}$, this is impossible.

Using (2.32) we get, for $1 \leq j < l \leq k+1$

$$\begin{aligned} 2 &= \langle y_j - y_l, y_j - y_l \rangle \\ &= \langle y_j, y_j \rangle - 2\langle y_j, y_l \rangle + \langle y_l, y_l \rangle \\ &= \frac{2k}{k-1} - 2\langle y_j, y_l \rangle \end{aligned}$$

Since $2 \neq 0$ in $\mathbb{F}$, we get

$$\langle y_j, y_l \rangle = \frac{1}{k-1} \qquad 1 \leq j < l \leq k+1 \tag{2.33}$$

Using (2.32) and (2.33) we have, for $1 < j \leq k+1$

$$\begin{aligned} \langle \sum_{i=1}^{k+1} y_i, y_j - y_1 \rangle &= \langle \sum_{i=1}^{k+1} y_i, y_j \rangle - \langle \sum_{i=1}^{k+1} y_i, y_1 \rangle \\ &= 0 \end{aligned}$$

Since $y_2 - y_1, y_3 - y_1, \ldots, y_{k+1} - y_1$ are k linearly independent vectors in $\mathbb{F}^k$, we conclude that

$$\sum_{i=1}^{k+1} y_i = 0 \tag{2.34}$$

as only the zero vector is orthogonal to all vectors in $\mathbb{F}^k$ under the bilinear map induced by the full rank matrix $\mathbf{J}_k - \mathbf{I}_k$ (since $k \neq 1$ in $\mathbb{F}$, $\mathbf{J}_k - \mathbf{I}_k$ is of full rank). Using (2.32), (2.33) and (2.34) and the fact that $2 \neq 0$ in $\mathbb{F}$, we get

$$\begin{aligned} 0 &= \langle \sum_{j=1}^{k+1} y_j, \sum_{j=1}^{k+1} y_j \rangle \\ &= \sum_{j=1}^{k+1} \langle y_j, y_j \rangle + 2 \sum_{1 \leq j < l \leq k+1} \langle y_j, y_l \rangle \\ &= (k+1)\frac{k}{k-1} + 2\frac{(k+1)k}{2}\frac{1}{k-1} \\ &= \frac{2k(k+1)}{k-1} \end{aligned}$$

We have thus come to a contradiction since $k \neq 0, \pm 1$ and $2 \neq 0$ in $\mathbb{F}$. Hence the lemma is proved. ∎

We can now prove our lower bound results for complex numbers and for finite fields of odd characteristic.

Complex numbers

Theorem 2.9 *Any (not necessarily homogeneous) $\Sigma\Pi\Sigma$ circuit computing $S_n^2(X_1, \ldots, X_n)$ over the field of complex numbers requires at least $\lceil \frac{n}{2} \rceil$ multiplication gates.*

Proof: Since $S_3^2(X_1, X_2, X_3)$ is an irreducible polynomial, any $\Sigma\Pi\Sigma$ circuit computing it should have at least 2 multiplication gates. For larger values of n, we invoke Lemmas 2.8, 2.12 and 2.13 to complete the proof. ∎

GF(p^r), p odd

Theorem 2.10 *Any (not necessarily homogeneous) $\Sigma\Pi\Sigma$ circuit computing $S_n^2(X)$ over GF(p^r) where p is an odd prime, requires at least*

1. *$\lceil \frac{n}{2} \rceil$ multiplication gates if n is even*
2. *$\lceil \frac{n}{2} \rceil$ multiplication gates if n is odd and $n \not\equiv \pm 1, 3 \bmod p$*
3. *$\lfloor \frac{n}{2} \rfloor$ multiplication gates if n is odd and $n \equiv \pm 1, 3 \bmod p$*

Thus, as long as p is an odd prime, we have a lower bound of $\lfloor \frac{n}{2} \rfloor$ for all n. If $p > 3$, we have a $\lceil \frac{n}{2} \rceil$ lower bound for all even and infinitely many odd n.

Proof: The lower bounds in parts 1 and 2 follow from Lemmas 2.8, 2.12 and 2.13. Suppose n is odd. Since a $\Sigma\Pi\Sigma$ circuit computing $S_n^2(X_1, \ldots, X_n)$ also gives us a $\Sigma\Pi\Sigma$ circuit computing $S_{n-1}^2(X_1, \ldots, X_{n-1})$ for which we have a lower bound of $\frac{n-1}{2}$, we get the lower bound in part 3. ∎

Rational and real numbers

Finally, we show that the $n-1$ lower bound of Graham and Pollack also extends to inhomogeneous $\Sigma\Pi\Sigma$ circuits over rational and real numbers.

Theorem 2.11 *Any (not necessarily homogeneous) $\Sigma\Pi\Sigma$ circuit computing $S_n^2(X)$ over reals or rationals requires at least $n-1$ multiplication gates.*

Proof: As observed in Section 2.1

$$T_n^2(X_1, \ldots, X_n) = (\sum_{j=1}^{n} X_j)^2 - 2S_n^2(X_1, \ldots, X_n)$$

Hence, any $\Sigma\Pi\Sigma$ circuit computing $S_n^2(X_1, \ldots, X_n)$ with less than $n-1$ multiplication gates gives us a $\Sigma\Pi\Sigma$ circuit computing $T_n^2(X_1, \ldots, X_n)$ with less than n multiplication gates. This implies, from the ideas of Section 2.4.1, that there are $n-1$ homogeneous linear forms $\ell_1, \ldots, \ell_{n-1}$ in the variable X_1 such that $T_n^2(X_1, \ell_1, \ldots, \ell_{n-1}) = 0$. This is clearly impossible over rationals and reals, since the coefficient of X_1^2 will not vanish. ∎

Chapter 3

The static membership problem

In this chapter, we outline our results on the set membership problem (defined in Section 1.2) in the quantum bit probe model. We also give simplified proofs of lower bounds for set membership in the classical bit probe model. We then discuss the set membership problem in the quantum cell probe model with implicit storage schemes, and prove a $\Omega(\log n)$ lower bound on the number of queries, generalising a result of Yao [Yao81].

The main new results in the chapter are

- A tradeoff between space s and number of probes t in the exact quantum bit probe model, for the static membership problem with universe size m and size of stored subset at most n. The tradeoff is captured by the following inequality (Theorem 3.2).

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

  We also give a simplified proof (Theorem 3.6) of the above inequality for classical deterministic bit probe schemes for static membership.

- For quantum bit probe schemes with two-sided error at most ϵ, storing subsets of size at most n from a universe of size m, and answering membership queries using p quantum probes, the following lower bound (Theorem 3.5) on the space s.

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

  Above, $\delta \stackrel{\Delta}{=} \epsilon^{1/p}$. For classical randomised schemes with two-sided error at most ϵ, we prove a slightly improved lower bound on the space s (Theorem 3.9).

- An $\Omega(\log n)$ lower bound for bounded error implicit storage quantum cell probe schemes for static membership (Theorem 3.10), if the size of the universe (m) is sufficiently large as compared to the size of the stored set (n), the size of the 'pointer space', and the number of cells of storage used.

3.1 Definitions and notations

In this section we give an alternate, but for our purposes more useful, definition of the *quantum bit probe model* and then give some formal definitions and notations which will be used in the proofs of the lower bounds on set membership in the quantum bit probe model.

3.1.1 The quantum bit probe model

A quantum (s,t)-bit probe scheme for a static data structure problem $f : D \times Q \to A$ is basically a quantum $(s,t,1)$-cell probe scheme for f. It has two components: a classical deterministic storage scheme that stores the data $d \in D$ using s bits, and a quantum query scheme that answers queries $q \in Q$ by 'quantumly probing a bit at a time' at most t times.

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 Section 1.2 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 $[x]_T \stackrel{\Delta}{=} \sum_{i \in T} x_i \pmod 2$. 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 $b = 0$, $l_i = \emptyset$).

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.

3.1.2 Framework for the lower bound proofs in the quantum bit probe model

We now describe the general framework in which the various proofs of lower bounds in the quantum bit probe model are presented, and also give some definitions and notations which will be used in those proofs.

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.

3.2 Quantum bit probe schemes

We illustrate the linear algebraic approach to proving space-time tradeoffs for set membership in the quantum bit probe model by giving a simple proof of the fact that if the

quantum query scheme is exact and makes only one probe, then the characteristic vector representation of the stored set is optimal.

Theorem 3.1 *Suppose there exists a scheme for storing subsets S of size at most n, $n \geq 1$, from a universe $\mathbf{U}$ of size m that uses s bits of storage and answers membership queries, with zero error probability, with only one quantum probe. Then, $s \geq m$.*

Proof: We use the notation of Section 3.1. For any subset $S \subseteq \mathbf{U}$, $|S| \leq n$, let us define

$$W_S \stackrel{\Delta}{=} U_1 O_S U_0$$

Claim 3.1 $\{(O_S)_{S \subseteq \mathbf{U}:|S| \leq 1}\}$ *are linearly independent.*

Proof: Since $W_S \stackrel{\Delta}{=} U_1 O_S U_0$ where U_1 and U_0 are unitary, it suffices to show that $\{(W_S)_{S \subseteq \mathbf{U}:|S| \leq 1}$ are linearly independent. Suppose there is a nontrivial linear combination

$$\sum_{S \subseteq \mathbf{U}:|S| \leq 1} \alpha_S W_S = 0$$

Consider the situation when the singleton set $\{i\}, i \in [m]$ is stored. Applying $|\phi_{\{i\}}\rangle$ to the linear combination above, we have

$$\sum_{S \subseteq \mathbf{U}:|S| \leq 1, S \neq \{i\}} \alpha_S W_S |\phi_{\{i\}}\rangle + \alpha_{\{i\}} W_{\{i\}} |\phi_{\{i\}}\rangle = 0$$

In the above equation, $W_S|\phi_{\{i\}}\rangle \in A_0$ for $S \neq \{i\}$ and $W_{\{i\}}|\phi_{\{i\}}\rangle \in A_1$. Hence $\alpha_{\{i\}} = 0$ for all $i \in [m]$. This implies that $\alpha_\emptyset = 0$ as well, leading to a contradiction.

Hence, the claim is proved. ∎

We also prove the following claim.

Claim 3.2 $\{(O_S)_{S \subseteq \mathbf{U}:|S| \leq 1}\}$ *lie in a vector space of dimension at most $s + 1$.*

Proof: O_S is completely determined by the bit string $x(S)$ stored by the storage scheme for set S i.e.

$$O_S = \sum_{j \in [s], z} (-1)^{x(S)_j} |j, 1, z\rangle\langle j, 1, z| + \sum_{j \in [s], z} |j, 0, z\rangle\langle j, 0, z|$$

Define linear operators $A_{\{j\}}$, $j \in [s]$

$$A_{\{j\}} \stackrel{\Delta}{=} \sum_z |j, 1, z\rangle\langle j, 1, z|$$

Also define a linear operator $A_\emptyset$

$$A_\emptyset \stackrel{\Delta}{=} \sum_{j \in [s], z} |j, 0, z\rangle\langle j, 0, z|$$

Then

$$O_S = A_\emptyset + \sum_{j\in[s]} (-1)^{x(S)_j} A_{\{j\}}$$

Hence we see that $A_{\{j\}}, j \in [s]$ and $A_\emptyset$ together span $\{(O_S)_{S\subseteq \mathbf{U}:|S|\leq 1}\}$. So, $\{(O_S)_{S\subseteq \mathbf{U}:|S|\leq 1}\}$ lie in a vector space of dimension at most $s+1$. ■

From the above two claims, $s+1 \geq m+1$ i.e. $s \geq m$. This proves the theorem. ■

The proofs for quantum schemes making more than one probe are refinements of the basic idea of the proof of Theorem 3.1.

Theorem 3.2 *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 3.1. 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 3.3 $\{W_S^{\otimes n}\}_{S\in\binom{\mathbf{U}}{\leq n}}$ *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

$$\sum_{S\in\binom{\mathbf{U}}{\leq k}, S\neq T} \alpha_S W_S^{\otimes n}|\phi_T\rangle + \alpha_T W_T^{\otimes n}|\phi_T\rangle = 0 \tag{3.1}$$

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, $1 \leq l \leq k$ 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 3.1), 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 3.4 *$\{W_S^{\otimes n}\}_{S\in\binom{\mathrm{U}}{\leq n}}$ lie in a vector space of dimension at most $\sum_{i=0}^{nt}\binom{s}{i}$.*

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

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

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

For any pair of indices (i, j),

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

where, recalling the notation of Section 3.1, $x(S)$ 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 $T_{k_{t-1},\ldots,k_0} \stackrel{\Delta}{=} l_{k_{t-1}} \triangle l_{k_{t-2}} \triangle \cdots \triangle l_{k_0}$ and $[x(S)]_T$ to be the parity of the bits stored in $x(S)$ at the locations of T, we have

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

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

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

Then we have,

$$W_S = \sum_{T\in\binom{[s]}{\leq t}} (-1)^{[x(S)]_T} A_T \tag{3.2}$$

Hence,

$$(W_S)^{\otimes n} = \left(\sum_{T_1\in\binom{[s]}{\leq t}} (-1)^{[x(S)]_{T_1}} A_{T_1}\right) \otimes \cdots \otimes \left(\sum_{T_n\in\binom{[s]}{\leq t}} (-1)^{[x(S)]_{T_n}} A_{T_n}\right)$$

$$= \sum_{\substack{T_i \in \binom{[s]}{\leq t} \\ 1 \leq i \leq n}} (-1)^{[x(S)]_{T_1}} \cdots (-1)^{[x(S)]_{T_n}} (A_{T_1} \otimes \cdots \otimes A_{T_n})$$

$$= \sum_{\tilde{T} \in \binom{[s]}{\leq nt}} (-1)^{[x(S)]_{\tilde{T}}} B_{\tilde{T}}$$

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

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

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

Now the theorem is an easy consequence of the above two claims. ∎

Remark: Equation 3.2 in the proof of Claim 3.4 above is similar to the statement of a lemma of Shi.

Lemma 3.1 ([Shi00, Lemma 2.4] rephrased) *Consider a quantum query algorithm, having 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,*

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

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 3.3 *The tradeoff result of Theorem 3.2 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 3.2 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 3.1) 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 bit probe scheme for the static membership problem.

Theorem 3.4 *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

$$|F| \geq \frac{\binom{m}{n}}{\binom{n}{\frac{n}{2}}\binom{m-\frac{n}{2}}{\frac{n}{2}}} \geq \frac{\frac{m}{n}\frac{m-1}{n-1}\cdots\frac{m-n/2+1}{n-n/2+1}}{2^n} \geq \frac{\left(\frac{m}{n}\right)^{n/2}}{2^n} = \left(\frac{m}{4n}\right)^{n/2} \tag{3.3}$$

Let $t \stackrel{\Delta}{=} \left\lceil \frac{4\log|F|}{n\log(1/(4\epsilon))} \right\rceil$. 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,

$$\frac{4\log|F|}{n\log(1/(4\epsilon))} \leq t \leq \frac{4\log|F|}{n\log(1/(4\epsilon))} + 1 \leq \frac{6\log|F|}{n\log(1/(4\epsilon))} \tag{3.4}$$

Claim 3.5 *$\{W_S^{\otimes nt}\}_{S\in F}$ 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

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

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

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

$$\Rightarrow \sum_{S\in F} \alpha_S (\langle \phi_{i_1}|W_S^\dagger W_T|\phi_{i_1}\rangle \cdots \langle \phi_{i_n}|W_S^\dagger W_T|\phi_{i_n}\rangle)^t = 0 \tag{3.5}$$

- 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,

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

We now note that for every $T \in F$, we have a linear combination as in equation 3.5 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 (3.4), 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 3.6 *$\left\{W_S^{\otimes nt}\right\}_{S\in F}$ lie in a vector space of dimension at most $\sum_{j=0}^{nt}\binom{s}{j}$.*

Proof: Similar to proof of Claim 3.4 in Theorem 3.2. ∎

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 (3.4), 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 (3.3) and the lower bound on t from (3.4), 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 3.5 *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 3.4. 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 3.5 in Theorem 3.4 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 3.4 in Theorem 3.2 was proved. Therefore, by a argument similar to that at the end of the proof of Theorem 3.4, we get a lower bound

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

∎

3.3 Classical bit probe schemes

We now give a simpler proof for the space v/s probes tradeoff equation for classical deterministic bit probe schemes solving the static membership problem.

Theorem 3.6 *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 $f_i : \{0,1\}^s \rightarrow \mathbb{R}$ 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 $\Phi : \binom{\mathbf{U}}{\leq n} \rightarrow (\{0,1\}^s \rightarrow \mathbb{R})$ 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

$$\Phi(\{\}) \stackrel{\Delta}{=} \text{constant 1 function}$$

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

Claim 3.7 $\{\Phi(S)\}_{S \in \binom{\mathbf{U}}{\leq n}}$ *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 3.8 $\{\Phi(S)\}_{S \in \binom{\mathbf{U}}{\leq n}}$ *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 3.7 *The tradeoff result of Theorem 3.6 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 3.6 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 $f_i \stackrel{\Delta}{=} \sum_{D \in F_i} g_D$. 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 3.8 *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 $k \stackrel{\Delta}{=} \left\lceil \frac{4\log(27m)}{3\log(1/4e\epsilon)} \right\rceil$. Since $m^{1/3} > 1/\epsilon$, $\frac{4\log(27m)}{3\log(1/4e\epsilon)} \geq 4$. Therefore,

$$\frac{4\log(27m)}{3\log(1/4e\epsilon)} \leq k \leq \frac{4\log(27m)}{3\log(1/4e\epsilon)} + 1 \leq \frac{5\log(27m)}{3\log(1/4e\epsilon)} \tag{3.6}$$

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 (3.6), we get

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

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 3.6, 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 (3.6) 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 3.4, 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 3.9 *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 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 3.8 above. We repeat the query scheme $k \stackrel{\Delta}{=} \left\lceil \frac{4\log(27m)}{3\log(1/4e\epsilon)} \right\rceil$ 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 3.8. 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 3.8. ■

3.4 Quantum cell probe model with implicit storage schemes

In this section, we show that for universe sizes m that are 'large' compared to n, p, q, $\Omega(\log n)$ is a lower bound on the number of quantum probes t required to solve the static membership problem with (p, q, t) implicit storage quantum cell probe schemes. We start with the following lemma. We only a sketch the proof of the lemma. A complete proof of a weaker version of the lemma, which nevertheless illustrates the main idea of a "logical interval" (defined below), is given in the appendix.

Lemma 3.2 *Suppose S is an n element subset of the universe $[m]$, where $m \geq 2n$. If the storage scheme is implicit, always stores the same 'pointer' values in the same locations, and in the remaining locations, stores the elements of S in a fixed order (repetitions of an*

element are allowed, but all elements have to be stored) based on their relative ranking in S, then $\Omega(\log n)$ probes are needed by any bounded error quantum cell query strategy to answer membership queries.

Proof: **(Sketch)** The proof follows by modifying Ambainis's lower bound proof for quantum ordered searching [Amb99]. There, it was shown that if S is stored in sorted order in a table T then, given any query element q, $\Omega(\log n)$ probes are required by any quantum search strategy to find out the smallest index i, $1 \leq i \leq n$, such that $q \leq T(i)$. We observe that the lemma above does not follow directly from the result of Ambainis, since we only need to decide if q is present in the table or not, and this is a weaker requirement. To prove the lemma, we follow the adversary strategy of [Amb99] with some minor changes. We study the behaviour of the quantum query scheme with query element n. The proof of Ambainis is based on a clever strategy of subdividing "intervals" (an interval is a contiguous set of locations in the sorted table). We work instead with "logical intervals", where a logical interval denotes the set of locations in the table where elements contiguous in the natural ordering are stored (as determined by the fixed storing order). After this definition, one can easily show that the same subdivision strategy as in [Amb99] goes through. In Ambainis's proof, the adversary constructs inputs by padding with zeros from the beginning up to the left of an interval, and with ones from the end up to the right of the interval. Instead, we pad with *small* numbers $(1, 2, 3, \ldots)$ from the logical beginning up to the logical left of a logical interval, and with *large* numbers $(m, m-1, m-2, \ldots)$ from the logical end up to the logical right of the logical interval. We store the appropriate 'pointer values' in the 'pointer locations' (predetermined by the storing strategy). After doing this, one can easily show that the same error analysis of [Amb99] goes through. Thus, the adversary finally can produce two inputs, one of them containing n and the other not, such that the behaviour of the query scheme is very similar on both. This is a contradiction. ■

Remark: Høyer *et al.* also prove an $\Omega(\log n)$ lower bound for quantum ordered searching [HNS01]. But their approach, which is based on "distinguishing oracles", does not seem to be suitable for proving lower bounds for boolean valued functions. Hence to prove Lemma 3.2, we modify the older $\Omega(\log n)$ lower bound of Ambainis for quantum ordered searching.

Theorem 3.10 *For every n, p, q, there exists an $N(n, p, q)$ such that for all $m \geq N(n, p, q)$, the following holds: Consider any bounded error (p, q, t) implicit storage quantum cell probe scheme for the static membership problem with universe size m and size of the stored subset at most n. Then the quantum query scheme must make $t = \Omega(\log n)$ probes.*

Proof: **(Sketch)** Our proof follows from the Ramsey theoretic arguments of Yao [Yao81] together with Lemma 3.2. The details are omitted. ■

Chapter 4

Static predecessor: Classical case

In this chapter, we present our lower bound for the query complexity of the static predecessor problem (defined in Section 1.2) in the classical cell probe model with randomised query schemes. We first recall the approach of Miltersen, Nisan, Safra and Wigderson for proving lower bounds for the predecessor problem via their round elimination lemma [MNSW98] in Sections 4.1 and 4.2. In Section 4.3, we discuss how improving the round elimination lemma improves the lower bound for predecessor. After that, we discuss some preliminaries from (classical) information theory in Section 4.4, which will be required in the proof of our improved (classical) round elimination lemma. We present the proof of a technical lemma in Section 4.5, which is required in the proof of our improved round elimination lemma. Finally, we present our improved round elimination lemma in Section 4.6, and use it to prove lower bounds for the predecessor problem matching the classical deterministic upper bounds of Beame and Fich [BF99] in Section 4.7. We also use our round elimination lemma to prove improved rounds versus communication tradeoffs for the 'greater-than' problem in Section 4.8. Sections 4.6, 4.7 and 4.8 contain new results.

The main new results in this chapter are

- An improved round elimination lemma (Lemma 4.5) for classical communication protocols. This improves on the earlier round elimination lemma (Lemma 4.3) of Miltersen, Nisan, Safra and Wigderson [MNSW98].

- Optimal lower bound of

$$t = \Omega\left(\min\left(\frac{\log\log m}{\log\log\log m}, \sqrt{\frac{\log n}{\log\log n}}\right)\right)$$

  on the number of queries t required to solve the static predecessor problem with universe size m and size of stored subset at most n, in the classical cell probe model with randomised query schemes, with word size $(\log m)^{O(1)}$ and number of cells $n^{O(1)}$. The reason the above lower bound is optimal is because Beame and Fich [BF99] have shown matching classical cell probe solutions for predecessor. In fact, in their solution

the query schemes are deterministic. Our lower bound improves the previous lower bound of

$$t = \Omega\left(\min\left(\sqrt{\log\log m}, (\log n)^{1/3}\right)\right)$$

of Miltersen *et al.* [MNSW98] in the classical cell probe model with randomised query schemes. In the case of the classical cell probe model with deterministic query schemes, Beame and Fich [BF99] had already proved an optimal lower bound; however, our lower bound proof is substantially simpler than the lower bound proof of Beame and Fich.

- Improved lower bound of $\Omega(n^{1/t}t^{-2})$ for the t round public coins classical randomised communication complexity of the 'greater-than' problem on n bit integers. The previous lower bound was $\Omega(n^{1/t}2^{-O(t)})$ by Miltersen *et al.* [MNSW98]. A t-round private coin $O(n^{1/t}\log n)$ classical randomised protocol is known for this problem. Thus, for a fixed number of rounds, our lower bound is optimal to within logarithmic factors.

4.1 Cell probe complexity and communication: The classical case

In this section, we describe a connection between the classical cell probe complexity of a static data structure problem and the classical communication complexity of an associated communication game. This connection is the standard approach for showing lower bounds for static data structure problems in the classical cell probe model. It was first observed by Miltersen [Mil94].

Definition 4.1 *A $[t, l_1, \ldots, l_t]^A$ ($[t, l_1, \ldots, l_t]^B$) classical communication protocol is a protocol where Alice (Bob) starts the communication, the ith message is l_i bits long, and the communication goes on for t rounds. A $(t, a, b)^A$ ($(t, a, b)^B$) classical communication protocol is a protocol where Alice (Bob) starts the communication, each message from Alice to Bob is a bits long, each message from Bob to Alice is b bits long, and the communication goes on for t rounds.*

Let $f : D \times Q \to A$ be a static data structure problem. Consider a two-party communication problem where Alice is given a query $q \in Q$, Bob is given data $d \in D$, and they have to communicate and find out the answer $f(d, q)$. We have the following lemma.

Lemma 4.1 ([Mil94]) *Suppose there is a classical (s, w, t) cell probe solution with randomised query schemes for the static data structure problem f. Then there is a private coin $(2t, \log s, w)^A$ classical randomised protocol for the corresponding communication problem. The error probability of the communication protocol is the same as that of the cell probe scheme.*

Proof: **(Sketch)** The communication protocol just simulates the cell probe solution. ■

In many natural data structure problems $\log s$ is much smaller than w. This asymmetry in message lengths is crucial in proving non-trivial lower bounds on the number of rounds t, as we shall see below.

4.2 Predecessor: Earlier round elimination approach

In this section, we recall the approach of Miltersen, Nisan, Safra and Wigderson [MNSW98] for proving lower bounds for the query complexity of the static predecessor problem in the classical cell probe model with randomised query schemes. Their main technical tool was a round elimination lemma for public coin classical randomised communication protocols.

Definition 4.2 (Rank parity communication games, [MNSW98]) *In the* rank parity *communication game* $\mathsf{PAR}_{p,q}$*, Alice is given a bit string x of length p, Bob is given a set S of bit strings of length p, $|S| \leq q$, and they have to communicate and decide whether the rank of x in S (treating the bit strings as integers) is odd or even. By the rank of x in S, we mean the cardinality of the set $\{y \in S \mid y \leq x\}$. In the game $\mathsf{PAR}^{(k),A}_{p,q}$, Alice is given k bit strings $x_1, \ldots, x_k$ each of length p, Bob is given a set S of bit strings of length p, $|S| \leq q$, an index $i \in [k]$, and a copy of $x_1, \ldots, x_{i-1}$; they have to communicate and decide whether the rank of x_i in S is odd or even. In the game $\mathsf{PAR}^{(k),B}_{p,q}$, Alice is given a bit string x of length p and an index $i \in [k]$, Bob is given k sets $S_1, \ldots, S_k$ of bit strings of length p, $|S_j| \leq q$; they have to communicate and decide whether the rank of x in S_i is odd or even.*

Proposition 4.1 ([MNSW98]) *Let there be a $(n^{O(1)}, (\log m)^{O(1)}, t)$ classical cell probe solution with randomised query schemes to the static predecessor problem, where the universe size is m and the subset size is at most n. Then there is a private coin (and hence, public coin) $\left(2t + O(1), O(\log n), (\log m)^{O(1)}\right)^A$ classical randomised protocol for the rank parity communication game $\mathsf{PAR}_{\log m,n}$. The error probability of the communication protocol is the same as that of the cell probe scheme.*

Proof: Consider the *static rank parity* data structure problem where the storage scheme has to store a set $S \subseteq [m]$, $|S| \leq n$, and the query scheme, given a query $x \in [m]$, has to decide whether the rank of x in S is odd or even. Fredman, Komlós and Szemerédi [FKS84] have shown the existence of two-level perfect hash tables containing, for each member y of the stored subset S, y's rank in S, and using $O(n)$ cells of word size $O(\log m)$ and requiring only $O(1)$ deterministic cell probes. Combining a $(n^{O(1)}, (\log m)^{O(1)}, t)$ classical cell probe solution to the static predecessor problem with such a perfect hash table, gives us a $(n^{O(1)} + O(n), \max((\log m)^{O(1)}, O(\log m)), t + O(1))$ classical cell probe solution to the static rank parity problem. The error probability of the cell probe scheme for the rank parity problem is the same as the error probability of the cell probe scheme for the predecessor problem. By Lemma 4.1, we get a $(2t + O(1), O(\log n), (\log m)^{O(1)})^A$ private coin classical randomised protocol for the rank parity communication game $\mathsf{PAR}_{\log m,n}$. The

error probability of the communication protocol is the same as that of the cell probe scheme for the predecessor problem. ∎

Proposition 4.2 ([MNSW98]) *Suppose k divides p. A communication protocol with Alice starting for $\mathsf{PAR}_{p,q}$, gives us a communication protocol with Alice starting for $\mathsf{PAR}^{(k),A}_{p/k,q}$, with the same message complexity, number of rounds and error probability.*

Proof: Consider the problem $\mathsf{PAR}^{(k),A}_{p/k,q}$. Alice, who is given $x_1, \ldots, x_k$, computes the concatenation $\hat{x} \triangleq x_1 \cdot x_2 \cdots x_k$. Bob, who is given S, i and $x_1, \ldots, x_{i-1}$, computes

$$\hat{S} \triangleq \left\{ x_1 \cdot x_2 \cdots x_{i-1} \cdot y \cdot 0^{p(1-i/k)} \mid y \in S \right\}$$

After this, Alice and Bob run the protocol for $\mathsf{PAR}_{p,q}$ on inputs $\hat{x}$, $\hat{S}$ to solve the problem $\mathsf{PAR}^{(k),A}_{p/k,q}$. ∎

Proposition 4.3 ([MNSW98]) *Suppose k divides q, and k is a power of 2. A communication protocol for $\mathsf{PAR}_{p,q}$ with Bob starting, gives us a communication protocol for $\mathsf{PAR}^{(k),B}_{p-\log k-1,q/k}$ with Bob starting, with the same message complexity, number of rounds and error probability.*

Proof: Consider the problem $\mathsf{PAR}^{(k),B}_{p-\log k-1,q/k}$. Alice, given x and i, computes $\hat{x} \triangleq (i-1) \cdot 0 \cdot x$. Bob, given $S_1, \ldots, S_k$, computes the sets $S'_1, \ldots, S'_k$ where

$$S'_j \triangleq \begin{cases} \{(j-1) \cdot 0 \cdot y \mid y \in S_j\} & \text{if } |S_j| \text{ is even} \\ \{(j-1) \cdot 0 \cdot y \mid y \in S_j\} \bigcup \left\{(j-1) \cdot 1^{p-\log k}\right\} & \text{if } |S_j| \text{ is odd} \end{cases}$$

Above, the integers $(i-1), (j-1)$ are to be thought of as bit strings of length $\log k$. Bob also computes $\hat{S} \triangleq \bigcup_{j=1}^{k} S'_j$. Alice and Bob then run the protocol for $\mathsf{PAR}_{p,q}$ on inputs $\hat{x}$, $\hat{S}$ to solve the problem $\mathsf{PAR}^{(k),B}_{p-\log k-1,q/k}$. ∎

We now state the round elimination lemma of Miltersen *et al.* in the 'fixed error' form.

Lemma 4.2 (Round elimination, fixed error form, [MNSW98]) *Let $C = 99$ and $R = 4256$. Suppose $f : E \times F \to G$ is a function. Suppose there is a $[t, l_1, \ldots, l_t]^A$ public coin classical randomised protocol with worst case error less than $1/3$ for the communication game $f^{(Rl_1)}$. Then there is a $[t-1, Cl_2, \ldots, Cl_t]^B$ public coin classical randomised protocol for f with worst case error less than $1/3$.*

We now prove Miltersen *et al.*'s lower bound for the classical query complexity of static predecessor.

Theorem 4.1 *Suppose we have a $(n^{O(1)}, (\log m)^{O(1)}, t)$ bounded error classical cell probe solution to the static predecessor problem, where the universe size is m and the subset size is at most n. Then the number of queries t is at least $\Omega\left(\sqrt{\log \log m}\right)$ as a function of m, and at least $\Omega\left((\log n)^{1/3}\right)$ as a function of n.*

Proof: By Proposition 4.1, it suffices to consider communication protocols for the rank parity communication game $\mathsf{PAR}_{\log m,n}$. Let $n = 2^{(\log\log m)^2}$. Let R, C be as in Lemma 4.2. For constants $c_2, c_3 \geq 1$, define

$$a \triangleq c_2 \log n \qquad b \triangleq (\log m)^{c_3} \qquad t \triangleq (R + C + c_2 + c_3)^{-1}\sqrt{\log\log m}$$

We shall show that $\mathsf{PAR}_{\log m,n}$ does not have bounded error $(2t, a, b)^A$ public coin classical randomised communication protocols, thus proving the desired lower bounds on the query complexity of static predecessor.

Given a $(2t, a, b)^A$ bounded error public coin classical protocol for $\mathsf{PAR}_{\log m,n}$, we get a $(2t, a, b)^A$ bounded error public coin classical protocol for

$$\mathsf{PAR}^{(Ra),A}_{\frac{\log m}{Ra},n}$$

by Proposition 4.2. Using Lemma 4.2, we get a $(2t-1, Ca, Cb)^B$ bounded error public coin classical protocol for

$$\mathsf{PAR}_{\frac{\log m}{Ra},n}$$

Using the reduction of Proposition 4.3, we get a $(2t-1, Ca, Cb)^B$ bounded error public coin classical protocol for

$$\mathsf{PAR}^{(RCb),B}_{\frac{\log m}{Ra}-\log(RCb)-1,\frac{n}{RCb}}$$

From the given values of the parameters, we see that

$$\frac{\log m}{(2RC^{t-1}a)^t} \geq \log(RC^{2t-1}b) + 1$$

This implies that we also have a $(2t-1, Ca, Cb)^B$ bounded error public coin classical protocol for

$$\mathsf{PAR}^{(RCb),B}_{\frac{\log m}{2Ra},\frac{n}{RCb}}$$

Using Lemma 4.2 again, we get a $(2t-2, C^2a, C^2b)^A$ bounded error public coin classical protocol for

$$\mathsf{PAR}_{\frac{\log m}{2Ra},\frac{n}{RCb}}$$

We do the above steps repeatedly. After applying the above steps i times, we get a $(2t-2i, C^{2i}a, C^{2i}b)^A$ bounded error public coin classical protocol for

$$\mathsf{PAR}_{\frac{\log m}{(2RC^{i-1}a)^i},\frac{n}{(RC^ib)^i}}$$

By applying the above steps t times, we finally get a $(0, C^{2t}a, C^{2t}b)^A$ bounded error public coin classical protocol for

$$\mathsf{PAR}_{\frac{\log m}{(2RC^{t-1}a)^t},\frac{n}{(RC^tb)^t}}$$

From the given values of the parameters, we see that

$$\frac{\log m}{(2RC^{t-1}a)^t} \geq (\log m)^{\Omega(1)} \qquad \frac{n}{(RC^t b)^t} \geq n^{\Omega(1)}$$

Thus we get a bounded error zero round protocol for a rank parity problem on a non-trivial domain, which is a contradiction.

In the above proof, we are tacitly ignoring "rounding off" problems. We remark that this does not affect the correctness of the proof. ∎

4.3 Improving lower bounds for predecessor

There is a gap between the lower bound for static predecessor proved in Theorem 4.1 and the upper bound of Beame and Fich [BF99]. Beame and Fich describe $(n^{O(1)}, O(\log m), t)$ classical cell probe solutions with deterministic query schemes for predecessor where

$$t = O\left(\min\left(\frac{\log\log m}{\log\log\log m}, \sqrt{\frac{\log n}{\log\log n}}\right)\right) \tag{4.1}$$

Beame and Fich [BF99] also prove a lower bound matching their upper bound to within constant factors, but their lower bound holds only if the query scheme is deterministic. The lower bound of Theorem 4.1 holds even if the query scheme is randomised. The round elimination lemma for fixed error (Lemma 4.2) is not strong enough to prove a lower bound matching (4.1), unless one can make $C = 1$ in that lemma. Such a statement looks unlikely to be true, though we have no counterexample for it as yet.

Miltersen *et al.* [MNSW98] prove the 'fixed error' form of their round elimination lemma (Lemma 4.2) by first proving a 'variable error' form of the round elimination lemma. We state the 'variable error' round elimination lemma below.

Lemma 4.3 (Round elimination, variable error form, [MNSW98]) *Let $\epsilon, \delta > 0$ be such that $\delta \leq \epsilon^2(100\ln(8/\epsilon))^{-1}$. Suppose $f : E \times F \to G$ is a function. Suppose the communication game $f^{(n)}$ has a $[t, l_1, \ldots, l_t]^A$ public coin classical randomised protocol with worst case error less than δ. Also suppose that $n \geq 20(l_1 \ln 2 + \ln 5)\epsilon^{-1}$. Then there is a $[t-1, l_2, \ldots, l_t]^B$ public coin classical randomised protocol for f with worst case error less than ϵ.*

Lemma 4.3 is also not strong enough to prove a lower bound matching (4.1). This is because the dependence between δ and ϵ in Lemma 4.3 is quadratic.

Suppose one can prove a variable error round elimination lemma where the error of the new protocol is within a *additive term* of the error of the old protocol. Also suppose that the additive term is bounded by a polynomial in l_1/n (refer the notation of Lemma 4.3). Then one can prove a lower bound for predecessor matching Beame and Fich's upper bound, even though one still does not prove a fixed error round elimination lemma with $C = 1$. In the remaining part of this chapter, we prove just such a lemma.

Our proof of the improved variable error round elimination lemma uses some results from (classical) information theory. We believe that the information-theoretic approach brings out naturally the intuition behind why such a result should be true. The advantages of using an improved round elimination lemma to prove lower bounds for predecessor are that the lower bounds obtained hold for the classical cell probe model with randomised query schemes, are optimal, and also, the proofs are substantially simpler than the lower bound proofs of Beame and Fich.

4.4 Information theoretic preliminaries

In this section, we discuss the information theoretic lemmas which will be used in the proof of our improved classical round elimination lemma. For a good account of (classical) information theory, see the book by Cover and Thomas [CT91].

Let A be a classical random variable taking values in a finite set S. Let A take the value $x \in S$ with probability p_x. Then, the *Shannon entropy* of A is defined as $S(A) \stackrel{\Delta}{=} -\sum_{x \in S} p_x \log p_x$. Let A, B be classical random variables. Then, their *mutual information* is defined as $I(A:B) \stackrel{\Delta}{=} S(A) + S(B) - S(AB)$.

Suppose X, M are classical random variables taking values in finite sets S, T respectively. For $x \in S, m \in T$, let $\sigma(m \mid x)$ denote the conditional probability that $T = m$ given that $X = x$. In what follows, we shall sometimes think of M as a classical randomised encoding $x \mapsto \sigma_x$ of X, where σ_x denotes the conditional distribution of M given that $X = x$. We shall define $\sigma \stackrel{\Delta}{=} \sum_x p_x \sigma_x$ to be the distribution of the average encoding. Then, $S(XQ) = S(X) + \sum_x p_x S(\sigma_x)$, and hence, $I(X:Q) = S(\sigma) - \sum_x p_x S(\sigma_x)$.

Let X, Y, M be classical random variables taking values in finite sets. Define the random variable $Z \stackrel{\Delta}{=} XY$. Let p_{xy} denote the (marginal) probability that $(X, Y) = (x, y)$. We shall sometimes think of M to be a classical randomised encoding $(x, y) \mapsto \sigma_{xy}$ of Z, where σ_{xy} is the conditional distribution of M given that $(X, Y) = (x, y)$. Define q_y^x to be the (conditional) probability that $Y = y$ given that $X = x$. $y \mapsto \sigma_{xy}$ can be thought of as a classical randomised encoding M^x of Y given that $X = x$. We let $I((Y:Q)|X = x)$ denote the mutual information of this encoding.

The following proposition has been observed by Klauck *et al.* [KNTZ01].

Proposition 4.4 *Suppose M is a classical randomised encoding of a classical random variable X. Suppose $X = X_1 X_2 \ldots X_n$, where the X_i are classical independent random variables. Then, $I(X_1 \ldots X_n : M) = \sum_{i=1}^{n} I(X_i : M X_1 \ldots X_{i-1})$.*

Proof: We use the following information theoretic identity, which follows easily from the definitions.

$$I(A:BC) = I(A:B) + I(AB:C) - I(B:C)$$

The proof of the proposition now follows by induction on n, using the above identity repeatedly. We also use the fact that $I(X_i : X_{i+1} \ldots X_n) = 0$ for $1 \leq i < n$, since $X_1, \ldots, X_n$ are independent classical random variables. ∎

We shall also require the following property of mutual information.

Proposition 4.5 *Let X, Y be classical random variables and M be a classical randomised encoding of (X, Y). Then, $I(Y : MX) = I(X : Y) + E_X[I((Y : M)|X = x)]$.*

Proof: Let σ_{xy} be the distribution of M when $X, Y = x, y$. Let p_x be the (marginal) probability that $X = x$ and q_y^x the (conditional) probability that $Y = y$ given $X = x$. Define $\sigma_x \stackrel{\Delta}{=} \sum_y q_y^x \sigma_{xy}$. We now have

$$\begin{aligned} I(Y : MX) &= S(Y) + S(MX) - S(MXY) \\ &= S(Y) + S(X) + \sum_x p_x S(\sigma_x) - (S(XY) + \sum_{x,y} p_x q_y^x S(\sigma_{xy})) \\ &= I(X : Y) + \sum_x p_x (S(\sigma_x) - \sum_y q_y^x S(\sigma_{xy})) \\ &= I(X : Y) + \sum_x p_x I((Y : M)|X = x) \\ &= I(X : Y) + E_X[I((Y : M)|X = x)] \end{aligned}$$

■

We will need the following "average encoding theorem" of Klauck *et al.* [KNTZ01]. Klauck *et al.* actually prove a quantum version of this theorem in their paper, but we will use the classical version in the proof of the classical round elimination lemma. Intuitively speaking, the theorem says that if the mutual information between a (classical) random variable and its (classical) encoding is small, then the various probability distributions on the codewords are close to the average probability distribution on the codewords.

Definition 4.3 (Total variation distance) *Let P, Q be probability distributions on the same finite sample space Ω. Let p_x (q_x) denote the probability of the sample point $x \in \Omega$ under P (Q). The* total variation distance *(also known as the ℓ_1-distance) between P and Q, denoted by $\|P - Q\|_1$, is defined as*

$$\|P - Q\|_1 \stackrel{\Delta}{=} \sum_{x \in \Omega} |p_x - q_x|$$

Theorem 4.2 (Average encoding, classical version, [KNTZ01]) *Let X be a classical random variable which takes value x with probability p_x, and M be a classical randomised encoding $x \mapsto \sigma_x$ of X, where σ_x is a probability distribution over the sample space of codewords. The probability distribution of the average encoding is $\sigma \stackrel{\Delta}{=} \sum_x p_x \sigma_x$. Then*

$$\sum_x p_x \|\sigma_x - \sigma\|_1 \leq \sqrt{(2 \ln 2) I(X : M)}$$

A proof of this theorem can be found in the appendix.

4.5 A classical round reduction lemma

In this section, we prove a classical round reduction lemma (Lemma 4.4), which will be required to prove the classical round elimination lemma. The proof of Lemma 4.4 is somewhat similar to the proof of Lemma 4.4 in Klauck *et al.* [KNTZ01], but much simpler since we are in the classical setting. Intuitively speaking, the lemma says that if the first message of Alice carries little information about her input, under some probability distribution on inputs, then it can be eliminated, giving rise to a protocol where Bob starts, with one less round of communication, and the same message complexity and similar error probability, with respect to the same probability distribution on inputs.

Consider communication protocols computing a function $f : E \times F \to G$. For an input $(x, y) \in E \times F$, we define the error $\epsilon^P_{x,y}$ of the protocol P on (x, y), to be the probability that the result of P on input (x, y) is not equal to $f(x, y)$. For a protocol P, given a probability distribution D on $E \times F$, we define the average error ϵ^P_D of P with respect to D as the expectation over D of the error of P on inputs $(x, y) \in E \times F$. We define ϵ^P to be worst case error of P on inputs $(x, y) \in E \times F$.

Lemma 4.4 (Classical round reduction lemma) *Suppose $f : E \times F \to G$ is a function. Let D be a probability distribution on $E \times F$, and P be a $[t, l_1, \ldots, l_t]^A$ private coin classical randomised protocol for f. Let X stand for the classical random variable denoting Alice's input (under distribution D), M be the first message of Alice in the protocol P, and $I(X : M)$ denote the mutual information between X and M under distribution D. Then there exists a $[t-1, l_2, \ldots, l_t]^B$ classical deterministic protocol Q for f, such that*

$$\epsilon^Q_D \leq \epsilon^P_D + \frac{1}{2}((2 \ln 2) I(X : M))^{1/2}$$

Proof: We first give an overview of the plan of the proof, before getting down to the details. The proof proceeds in stages.

Stage 1: Starting from P, we construct a $[t, l_1, \ldots, l_t]^A$ private coin protocol P', where the first message is independent of Alice's input, and $\epsilon^{P'}_D \leq \epsilon^P_D + (1/2)((2 \ln 2) I(X : M))^{1/2}$. The important idea in this step is to first generate Alice's message using a new private coin without "looking" at her input, and after that, to adjust Alice's old private coin in a suitable manner so as to be consistent with her message and input.

Stage 2: Since the first message of P' is independent of Alice's input, Bob can generate it himself. Doing this and setting coin tosses appropriately, gives us a $[t-1, l_2, \ldots, l_t]^B$ deterministic protocol Q such that $\epsilon^Q_D \leq \epsilon^{P'}_D$.

The protocol Q of Stage 2 is our desired $[t-1, l_2, \ldots, l_t]^B$ classical deterministic protocol for f. We have

$$\epsilon^Q_D = \epsilon^{P'}_D \leq \epsilon^P_D + \frac{1}{2}((2 \ln 2) I(X : M))^{1/2}$$

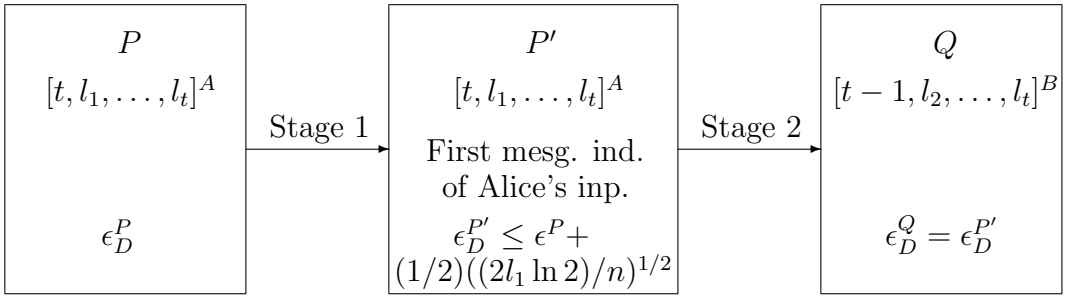


Figure 4.1: The various stages in the proof of Lemma 4.4.

We now give the details of the proof. Let σ_x be the probability distribution of the first message M of protocol P when Alice's input $X = x$. Let Y denote Bob's input register. Define $\sigma \stackrel{\Delta}{=} \sum_x p_x \sigma_x$, where p_x is the (marginal) probability of x under distribution D. σ is the probability distribution of the average first message under distribution D. By Theorem 4.2, we get that

$$\sum_x p_x \|\sigma_x - \sigma\|_1 \leq \sqrt{(2 \ln 2) I(X : M)}$$

For $x \in E$ and an instance m of the first message of Alice, let q_r^{xm} denote the (conditional) probability that the private coin toss of Alice results in r, given that Alice's input is x and her first message in protocol P is m. Let $\sigma(m \mid x)$ denote the probability that the first message of Alice in P is m, given that her input is x. Let $\sigma(m)$ denote the probability of m occurring in the average first message of Alice. Then, $\sigma(m) = \sum_x p_x \sigma(m \mid x)$.

Stage 1: We construct a $[t, l_1, \ldots, l_t]^A$ private coin classical randomised protocol P' for f with average error under distribution D, $\epsilon_D^{P'} \leq \epsilon_D^P + (1/2)((2 \ln 2) I(X : M))^{1/2}$, and where the probability distribution of the first message is independent of the input to Alice. Suppose Alice is given $x \in E$ and Bob is given $y \in F$. Alice tosses a fresh private coin to pick m with probability $\sigma(m)$. She then sets her old private coin to r with probability q_r^{xm}. (If in P, message m cannot occur when Alice's input is x, we say that protocol P' gives an error if such a thing happens.) After this, Alice and Bob behave as in protocol P (henceforth, Alice ignores the new private coin which she had tossed to generate her first message m). Hence in P', the probability distribution of the first message is independent of Alice's input.

Let us now compare the situations in protocols P and P' when Alice's input is x, Bob's input is y, Alice has finished tossing her private coins, but no communication has taken place as yet. In protocol P, the probability that Alice's private coin toss results in r is

$$\sum_m \sigma(m \mid x) q_r^{xm}$$

In protocol P', the probability that Alice's (old) private coin toss results in r is

$$\sum_m \sigma(m) q_r^{xm}$$

Thus, the ℓ_1 distance between the probability distributions on Alice's (old) private coin toss is

$$\begin{aligned}
&\sum_r \left| \sum_m q_r^{xm} (\sigma(m \mid x) - \sigma(m)) \right| \\
&\leq \sum_r \sum_m q_r^{xm} \left| \sigma(m \mid x) - \sigma(m) \right| \\
&= \sum_m \left(\left| \sigma(m \mid x) - \sigma(m) \right| \sum_r q_r^{xm} \right) \\
&= \sum_m \left| \sigma(m \mid x) - \sigma(m) \right| \\
&= \|\sigma_x - \sigma\|_1
\end{aligned}$$

Hence, the error probability of P' on input x, y

$$\epsilon_{x,y}^{P'} \leq \epsilon_{x,y}^{P} + \frac{1}{2} \|\sigma_x - \sigma\|_1$$

Let q_{xy} be the probability that $(X, Y) = (x, y)$ under distribution D. Then, the average error of P' under distribution D, $\epsilon_D^{P'}$, is bounded by

$$\begin{aligned}
\epsilon_D^{P'} &= \sum_{x,y} q_{xy} \epsilon_{x,y}^{P'} \\
&\leq \sum_{x,y} q_{xy} \left(\epsilon_{x,y}^{P} + \frac{1}{2} \|\sigma_x - \sigma\|_1 \right) \\
&= \epsilon_D^P + \frac{1}{2} \sum_x p_x \|\sigma_x - \sigma\|_1 \\
&\leq \epsilon_D^P + \frac{1}{2} ((2 \ln 2) I(X : M))^{1/2}
\end{aligned}$$

The last inequality follows from the "average encoding theorem" (Theorem 4.2).

Stage 2: We now construct our desired $[t-1, l_2, \ldots, l_t]^B$ classical deterministic protocol Q for f with $\epsilon_D^Q \leq \epsilon_D^{P'}$. Suppose all the coin tosses of Alice and Bob in P' were done publicly before any communication takes place. Now there is no need for the first message from Alice to Bob, because Bob can reconstruct the message by looking at the public coin tosses. This gives us a $[t-1, l_2, \ldots, l_t]^B$ public coin protocol Q', such that $\epsilon_{x,y}^{Q'} = \epsilon_{x,y}^{P'}$ for

every $(x, y) \in E \times F$. By setting the public coin of Q' to an appropriate value, we get a $[t-1, l_2, \ldots, l_t]^B$ deterministic protocol Q such that $\epsilon_D^Q \leq \epsilon_D^{Q'} = \epsilon_D^{P'}$. We have

$$\epsilon_D^Q = \epsilon_D^{P'} \leq \epsilon_D^P + \frac{1}{2}((2 \ln 2) I(X : M))^{1/2}$$

This completes the proof of Lemma 4.4. ∎

4.6 The classical round elimination lemma

We now prove the improved classical round elimination lemma (for the communication game $f^{(n)}$). The round elimination lemma is stated for public coin classical randomised protocols only. Since a public coin protocol can be converted to a private coin protocol at the expense of an additive increase in the communication complexity by at most logarithm of the total bit size of the inputs [New91], we also get a similar round elimination lemma for private coin protocols. But since the statement of the round elimination lemma is cleanest for public coin protocols, we give it below for such protocols only.

Lemma 4.5 (Classical round elimination lemma) *Suppose $f : E \times F \to G$ is a function. Suppose the communication game $f^{(n)}$ has a $[t, l_1, \ldots, l_t]^A$ public coin classical randomised protocol with worst case error less than δ. Then there is a $[t-1, l_2, \ldots, l_t]^B$ public coin classical randomised protocol for f with worst case error less than $\epsilon \stackrel{\Delta}{=} \delta + (1/2)(2l_1 \ln 2/n)^{1/2}$.*

Proof: Suppose the given protocol for $f^{(n)}$ has worst case error $\tilde{\delta} < \delta$. Define $\tilde{\epsilon} \stackrel{\Delta}{=} \tilde{\delta} + (1/2)(2l_1 \ln 2/n)^{1/2}$. To prove the classical round elimination lemma it suffices to give, by the harder direction of the minimax lemma, for any probability distribution D on $E \times F$, a $[t-1, l_2, \ldots, l_t]^B$ classical deterministic protocol P for f with average distributional error $\epsilon_D^P \leq \tilde{\epsilon} < \epsilon$. To this end, we will first construct a probability distribution D^* on $E^n \times [n] \times F$ as follows. Choose $i \in [n]$ uniformly at random. Choose independently, for each $j \in [n]$, $(x_j, y_j) \in E \times F$ according to distribution D. Set $y = y_i$ and throw away $y_j, j \neq i$. By the easier direction of the minimax lemma, we get a $[t, l_1, \ldots, l_t]^A$ classical deterministic protocol P^* for $f^{(n)}$ with distributional error, $\epsilon_{D^*}^{P^*} \leq \tilde{\delta} < \delta$. In P^*, Alice gets $x_1, \ldots, x_n \in E$, Bob gets $i \in [n]$, $y \in F$ and a copy of $x_1, \ldots, x_{i-1}$. We shall construct the desired protocol P from the protocol P^*.

Let M be the first message of Alice in P^*. Let the input to Alice be denoted by the classical random variable $X = X_1 X_2 \ldots X_n$ where X_i is the classical random variable corresponding to the ith input to Alice. Let the classical random variable Y denote the input y of Bob. Define $\epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*}$ to be the average error of P^* under distribution D^* when i is fixed and $X_1, \ldots, X_{i-1}$ are fixed to $x_1, \ldots, x_{i-1}$. Using Propositions 4.4, 4.5 and the fact that under distribution D^*, $X_1, \ldots, X_n$ are independent classical random variables, we get that

$$\begin{aligned} \frac{l_1}{n} &\geq \frac{I(X:M)}{n} \\ &= E_i[I(X_i : MX_1, \ldots, X_{i-1})] \\ &= E_{i,X}[I((X_i : M)|X_1, \ldots, X_{i-1} = x_1, \ldots, x_{i-1})] \end{aligned} \tag{4.2}$$

Also

$$\tilde{\delta} \geq \epsilon_{D^*}^{P^*} = E_{i,X}\left[\epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*}\right] \tag{4.3}$$

The expectations above are under distribution D^*.

For any $i \in [n]$, $x_1, \ldots, x_{i-1} \in E$, define the $[t, l_1, \ldots, l_t]^A$ private coin classical randomised protocol $P'_{i;x_1,\ldots,x_{i-1}}$ for the function f as follows. Alice is given $x \in E$ and Bob is given $y \in F$. Bob sets i to the given value, and both Alice and Bob set $X_1, \ldots, X_{i-1}$ to the values $x_1, \ldots, x_{i-1}$. Alice tosses her private coin to choose $X_{i+1}, \ldots, X_n \in E$, where each $X_j, i+1 \leq j \leq n$ is chosen independently according to the (marginal) distribution on E induced by D. Alice sets $X_i = x$ and Bob sets $Y = y$. Then they run protocol P^* on these inputs. The probability that $P'_{i;x_1,\ldots,x_{i-1}}$ makes an error for an input (x, y), $\epsilon_{x,y}^{P'_{i;x_1,\ldots,x_{i-1}}}$, is the average probability of error of P^* under distribution D^* when i is fixed to the given value, $X_1, \ldots, X_{i-1}$ are fixed to $x_1, \ldots, x_{i-1}$, and X_i, Y are fixed to x, y. Hence, the average probability of error of $P'_{i;x_1,\ldots,x_{i-1}}$ under distribution D

$$\epsilon_D^{P'_{i;x_1,\ldots,x_{i-1}}} = \epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*} \tag{4.4}$$

Let M' denote the first message of $P'_{i;x_1,\ldots,x_{i-1}}$ and X' denote the register X_i holding the input x to Alice. Then

$$I(X' : M') = I((X_i : M)|X_1, \ldots, X_{i-1} = x_1, \ldots, x_{i-1}) \tag{4.5}$$

Using Lemma 4.4 and equations (4.4) and (4.5), we get a $[t-1, l_2, \ldots, l_t]^B$ classical deterministic protocol $P_{i;x_1,\ldots,x_{i-1}}$ for f with

$$\begin{aligned} \epsilon_D^{P_{i;x_1,\ldots,x_{i-1}}} &\leq \epsilon_D^{P'_{i;x_1,\ldots,x_{i-1}}} + \tfrac{1}{2}((2\ln 2) I(X' : M'))^{1/2} \\ &= \epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*} + \tfrac{1}{2}((2\ln 2) I((X_i : M)|X_1, \ldots, X_{i-1} = x_1, \ldots, x_{i-1}))^{1/2} \end{aligned} \tag{4.6}$$

We have that (note that the expectations below are under distribution D^*)

$$\begin{aligned} E_{i,X}\left[\epsilon_D^{P_{i;x_1,\ldots,x_{i-1}}}\right] &\leq E_{i,X}\left[\epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*}\right] + \\ &\quad E_{i,X}\left[\tfrac{1}{2}\left((2\ln 2) I((X_i : M)|X_1, \ldots, X_{i-1} = x_1, \ldots, x_{i-1})\right)^{1/2}\right] \\ &\leq E_{i,X}\left[\epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*}\right] + \\ &\quad (1/2)\left((2\ln 2) E_{i,X}\left[I((X_i : M)|X_1, \ldots, X_{i-1} = x_1, \ldots, x_{i-1})\right]\right)^{1/2} \\ &\leq \tilde{\delta} + \tfrac{1}{2}\left(\tfrac{2l_1 \ln 2}{n}\right)^{1/2} \\ &= \tilde{\epsilon} \end{aligned} \tag{4.7}$$

The first inequality follows from (4.6), the second inequality follows from the concavity of the fourth root function and the last inequality from from (4.2) and (4.3).

From (4.7), we see that there exist $i \in [n]$ and $x_1, \ldots, x_{i-1} \in E$ such that $\epsilon_D^{P_{i;x_1,\ldots,x_{i-1}}} \leq \tilde{\epsilon}$. Let $P \triangleq P_{i;x_1,\ldots,x_{i-1}}$. P is our desired $[t-1, l_2, \ldots, l_t]^B$ classical deterministic protocol for f with $\epsilon_D^P \leq \tilde{\epsilon}$, thus completing the proof of the classical round elimination lemma. ∎

4.7 Predecessor: Optimal classical lower bounds

In this section, we prove our (optimal) lower bounds on the query complexity of static predecessor in the classical cell probe model with randomised query schemes.

Theorem 4.3 *Suppose we have a $(n^{O(1)}, (\log m)^{O(1)}, t)$ bounded error classical cell probe solution to the static predecessor problem, where the universe size is m and the subset size is at most n. Then the number of queries t is at least $\Omega\left(\frac{\log\log m}{\log\log\log m}\right)$ as a function of m, and at least $\Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$ as a function of n.*

Proof: The proof is similar to the proof of Theorem 4.1, but with different parameters, and using the stronger round elimination lemma (Lemma 4.5).

By Proposition 4.1, it suffices to consider communication protocols for the rank parity communication game $\mathsf{PAR}_{\log m,n}$. Let $n = 2^{(\log\log m)^2/\log\log\log m}$. Let $c_1 \stackrel{\Delta}{=} (2\ln 2)6^2$. For any given constants $c_2, c_3 \geq 1$, define

$$a \stackrel{\Delta}{=} c_2 \log n \qquad b \stackrel{\Delta}{=} (\log m)^{c_3} \qquad t \stackrel{\Delta}{=} \frac{\log\log m}{(c_1 + c_2 + c_3)\log\log\log m}$$

We shall show that $\mathsf{PAR}_{\log m,n}$ does not have bounded error $(2t, a, b)^A$ public coin classical randomised communication protocols, thus proving the desired lower bounds on the query complexity of static predecessor.

Given a $(2t, a, b)^A$ public coin protocol for $\mathsf{PAR}_{\log m,n}$ with error probability δ ($\delta < 1/3$), we get a $(2t, a, b)^A$ public coin protocol for

$$\mathsf{PAR}^{(c_1at^2),A}_{\frac{\log m}{c_1at^2},n}$$

with the same error probability δ, by Proposition 4.2. Using Lemma 4.5, we get a $(2t-1, a, b)^B$ public coin protocol for

$$\mathsf{PAR}_{\frac{\log m}{c_1at^2},n}$$

but the error probability increases to at most $\delta + (12t)^{-1}$. Using the reduction of Proposition 4.3, we get a $(2t-1, a, b)^B$ public coin protocol for

$$\mathsf{PAR}^{(c_1bt^2),B}_{\frac{\log m}{c_1at^2} - \log(c_1bt^2) - 1, \frac{n}{c_1bt^2}}$$

with error probability at most $\delta + (12t)^{-1}$. From the given values of the parameters, we see that

$$\frac{\log m}{(2c_1at^2)^t} \geq \log(c_1bt^2) + 1$$

This implies that we also have a $(2t-1, a, b)^B$ public coin protocol for

$$\mathsf{PAR}^{(c_1bt^2),B}_{\frac{\log m}{2c_1at^2}, \frac{n}{c_1bt^2}}$$

with error probability at most $\delta + (12t)^{-1}$. Using Lemma 4.5 again, we get a $(2t-2, a, b)^A$ public coin protocol for

$$\mathsf{PAR}_{\frac{\log m}{2c_1at^2}, \frac{n}{c_1bt^2}}$$

but the error probability increases to at most $\delta + 2(12t)^{-1}$.

We do the above steps repeatedly. After applying the above steps i times, we get a $(2t-2i, a, b)^A$ public coin protocol for

$$\mathsf{PAR}_{\frac{\log m}{(2c_1at^2)^i}, \frac{n}{(c_1bt^2)^i}}$$

with error probability at most $\delta + 2i(12t)^{-1}$.

By applying the above steps t times, we finally get a $(0, a, b)^A$ public coin protocol for

$$\mathsf{PAR}_{\frac{\log m}{(2c_1at^2)^t}, \frac{n}{(c_1bt^2)^t}}$$

with error probability at most $\delta + 2t(12t)^{-1} < 1/2$. From the given values of the parameters, we see that

$$\frac{\log m}{(2c_1at^2)^t} \geq (\log m)^{\Omega(1)} \qquad \frac{n}{(c_1bt^2)^t} \geq n^{\Omega(1)}$$

Thus we get a zero round protocol for a rank parity problem on a non-trivial domain with error probability less than $1/2$, which is a contradiction.

In the above proof, we are tacitly ignoring "rounding off" problems. We remark that this does not affect the correctness of the proof. ∎

4.8 The 'greater-than' problem

We illustrate another application of the round elimination lemma to communication complexity by proving improved rounds versus communication tradeoffs for the 'greater-than' problem.

Theorem 4.4 *The t round bounded error classical randomised communication complexity of GT_n is $\Omega(n^{1/t}t^{-2})$.*

Proof: We recall the following reduction from $GT^{(k)}_{n/k}$ to GT_n (see [MNSW98]): In $GT^{(k)}_{n/k}$, Alice is given $x_1, \ldots, x_k \in \{0,1\}^{n/k}$, Bob is given $i \in [k]$, $y \in \{0,1\}^{n/k}$, and copies of $x_1, \ldots, x_{i-1}$, and they have to communicate and decide if $x_i > y$. To reduce $GT^{(k)}_{n/k}$ to GT_n, Alice constructs $\tilde{x} \in \{0,1\}^n$ by concatenating $x_1, \ldots, x_k$, Bob constructs $\tilde{y} \in \{0,1\}^n$ by concatenating $x_1, \ldots, x_{i-1}, y, 1^{n(1-i/k)}$. It is easy to see that $\tilde{x} > \tilde{y}$ iff $x_i > y$.

Suppose there is a t round bounded error public coins protocol for GT_n with communication complexity l. We can think of the protocol as a $[t, l, \ldots, l]^A$ public coin protocol with worst case error probability less than $1/3$. Suppose

$$n \geq (Ct^2l)^t$$

where $C \stackrel{\Delta}{=} (2\ln 2)3^2$. Define $k \stackrel{\Delta}{=} Ct^2l$. For $1 \leq i \leq t$, define

$$n_i \stackrel{\Delta}{=} \frac{n}{k^i} \qquad \epsilon_i \stackrel{\Delta}{=} \frac{1}{3} + \frac{i}{2}\left(\frac{(2\ln 2)l}{k}\right)^{1/2}$$

Also define $n_0 \stackrel{\Delta}{=} n$ and $\epsilon_0 \stackrel{\Delta}{=} 1/3$. Then

$$\epsilon_t = \frac{1}{3} + \frac{t}{2}\left(\frac{(2\ln 2)l}{k}\right)^{1/2} = 1/2$$

and

$$n_t = \frac{n}{k^t} = \frac{n}{(Ct^2l)^t} \geq 1$$

We now apply the above self-reduction and Lemma 4.5 alternately. Before the ith stage, we have a $[t-i+1, l, \ldots, l]^Z$ public coin protocol for $GT_{n_{i-1}}$ with worst case error probability less than ϵ_{i-1}. Here $Z = A$ if i is odd, $Z = B$ otherwise. For the ith stage, we apply the self-reduction to get a $[t-i+1, l, \ldots, l]^Z$ public coin protocol for $GT_{n_i}^{(k)}$ with the same error probability. We then apply Lemma 4.5 to get a $[t-i, l, \ldots, l]^{Z'}$ public coin protocol for GT_{n_i} with worst case error probability less than ϵ_i. Here $Z' = B$ if $Z = A$ and $Z' = A$ if $Z = B$. This completes the ith stage.

Applying the self-reduction and the round elimination lemma alternately for t stages gives us a zero round protocol for the 'greater-than' problem on a domain of size $n_t \geq 1$ with worst case error probability less than $\epsilon_t = 1/2$, which is a contradiction.

In the above proof, we are tacitly ignoring "rounding off" problems. We remark that this does not affect the correctness of the proof.

This proves the classical lower bound of $\Omega(n^{1/t}t^{-2})$ on the message complexity. ∎

Remark: In the above proof, we think of a t round public coin protocol with communication complexity l as a $[t, l, \ldots, l]^A$ public coin protocol. But, suppose we are promised that every run of the public coin protocol uses l_i bits in the ith round, $l_1 + \cdots + l_t = l$, where l_i depends only on n. In other words, we are promised a $[t, l_1, \ldots, l_t]^A$ public coin protocol. Then one can do a more refined argument, where in the ith stage one does the self-reduction with $k = Ct^2l_i$, to show a stronger lower bound of $l = \Omega(n^{1/t}t^{-1})$. Such a refined argument, but for quantum protocols, is given in the proof of the quantum version of the above theorem (Theorem 5.5). Notice that the definition of quantum protocols requires that l_i be a function of n only.

Miltersen *et al.* [MNSW98] also use their round elimination lemma (Lemma 4.2) to prove (classical) lower bounds for other static data structure and communication complexity problems. We remark that all those results can be improved by using Lemma 4.5 in place of Lemma 4.2.

Chapter 5

Static predecessor: Quantum case

In this chapter, we present our lower bound for the query complexity of the static predecessor problem (defined in Section 1.2) in the bounded error address-only quantum cell probe model. The arguments in this chapter can be largely viewed as quantum generalisations of the arguments of Chapter 4.

We first discuss the connection between quantum cell probe complexity and quantum communication, paying special attention to address-only quantum cell probe schemes, in Section 5.1. We then delve into some results from quantum information theory in Section 5.2, which will be required in the proof of our quantum round elimination lemma. In Section 5.3, we prove a technical lemma which will be used in the proof of the quantum round elimination lemma. Finally, we present our quantum round elimination lemma in Section 5.4, and use it to prove lower bounds for the predecessor problem in the address-only quantum cell probe model in Section 5.5. Our lower bounds match the classical deterministic upper bounds of Beame and Fich [BF99], thus showing that Beame and Fich's scheme is optimal all the way up to address-only quantum. We also use the quantum round elimination lemma to prove the first rounds versus communication tradeoffs for the 'greater-than' problem in the quantum setting, in Section 5.6. Sections 5.4, 5.5 and 5.6 contain new results.

The main new results in this chapter are

- A round elimination lemma (Lemma 5.4) for quantum communication protocols.
- Optimal lower bound of

$$t = \Omega\left(\min\left(\frac{\log\log m}{\log\log\log m}, \sqrt{\frac{\log n}{\log\log n}}\right)\right)$$

  on the number of queries t required to solve the static predecessor problem with universe size m and size of stored subset at most n, in the bounded error address-only quantum cell probe model, with word size $(\log m)^{O(1)}$ and number of cells $n^{O(1)}$. The reason the above lower bound is optimal is because Beame and Fich [BF99] have shown matching classical deterministic cell probe solutions for predecessor.

- A lower bound of $\Omega(n^{1/t}t^{-3})$ for t round bounded error quantum communication protocols for the 'greater-than' problem on n bit integers. These bounds are the first rounds versus communication tradeoffs for the 'greater-than' problem in the quantum setting.

5.1 Cell probe complexity and communication: The quantum case

The lower bounds for the static membership problem in the quantum bit probe model, proved in Chapter 3, relied on linear algebraic techniques. Unfortunately, these techniques appear to be powerless in the quantum cell probe model. To prove a lower bound for the predecessor problem, we use a connection between the quantum cell probe complexity of a static data structure problem and the quantum communication complexity of an associated communication game. This connection can be thought of as a quantum analogue of Lemma 4.1. Below, the notation $(t, c, a, b)^A$ ($(t, c, a, b)^B$) denotes a $[t, c, l_1, \ldots, l_t]^A$ ($[t, c, l_1, \ldots, l_t]^B$) safe quantum protocol, where the per round message lengths of Alice and Bob are a and b qubits respectively i.e. if Alice (Bob) starts, $l_i = a$ for i odd and $l_i = b$ for i even ($l_i = b$ for i odd and $l_i = a$ for i even).

Let $f : D \times Q \rightarrow A$ be a static data structure problem. Consider a two-party communication problem where Alice is given a query $q \in Q$, Bob is given data $d \in D$, and they have to communicate and find out the answer $f(d, q)$. We have the following lemma.

Lemma 5.1 *Suppose we have a quantum (s, w, t) cell probe solution to the static data structure problem f. Then we have a $(2t, 0, \log s + w, \log s + w)^A$ safe coinless quantum protocol for the corresponding communication problem. If the query scheme is address-only, we can get a $(2t, 0, \log s, \log s + w)^A$ safe coinless quantum protocol. The error probability of the communication protocol is the same as that of the cell probe scheme.*

Proof: Given a quantum (s, w, t) cell probe solution to the static data structure problem f, we can get a $(2t, 0, \log s+w, \log s+w)^A$ safe coinless quantum protocol for the corresponding communication problem by just simulating the cell probe solution. If in addition, the query scheme is address-only, the messages from Alice to Bob need consist only of the 'address' part. This can be seen as follows. Let the state vector of the data qubits before the ith query be $|\theta_i\rangle$. $|\theta_i\rangle$ is independent of the query element and the stored data. Bob keeps t special ancilla registers in states $|\theta_i\rangle, 1 \leq i \leq t$ at the start of the protocol P. These special ancilla registers are in tensor with the rest of the qubits of Alice and Bob at the start of P. Protocol P simulates the cell probe solution, but with the following modification. To simulate the ith query of the cell probe solution, Alice prepares her 'address' and 'data' qubits as in the query scheme, but sends the 'address' qubits only. Bob treats those 'address' qubits together with $|\theta_i\rangle$ in the ith special ancilla register as Alice's query, and performs the oracle table transformation on them. He then sends these qubits (both the 'address' as well as the ith special register qubits) to Alice. Alice exchanges the contents

of the ith special register with her 'data' qubits (i.e. exchanges the basis states), and proceeds with the simulation of the query scheme. This gives us a $(2t, 0, \log s, \log s + w)^A$ safe coinless quantum protocol with the same error probability as that of the cell probe query scheme. ■

In many natural data structure problems $\log s$ is much smaller than w and thus, in the address-only quantum case, we get a $(2t, 0, \log s, O(w))^A$ safe protocol. This asymmetry in message lengths is crucial in proving non-trivial lower bounds on t. The concept of a safe quantum protocol helps us in exploiting this asymmetry. The reason, intuitively speaking, is as follows. In the previous quantum round reduction arguments (e.g. those of Klauck *et al.* [KNTZ01]), the complexity of the first message in the protocol increases quickly as the number of rounds is reduced and the asymmetry gets lost. This leads to a problem where the first message soon gets big enough to potentially convey substantial information about the input of one player to the other, destroying any hope of proving strong lower bounds on the number of rounds. But in a safe quantum protocol one can show through a careful quantum information theoretic analysis of the round reduction process, that though the complexity of the first message increases a lot, this increase is confined to the safe overhead and so, the information content does not increase much. This is the key property which allows us to prove a round elimination lemma for safe quantum protocols.

To prove lower bounds for the query complexity of data structure problems in the address-only quantum cell probe model via communication complexity, we need to define public coin quantum protocols and make use of Yao's minimax lemma. The reason is as follows. The minimax lemma is the main tool which allows one to convert 'average case' round reduction arguments to 'worst case' arguments. But this conversion is at the expense of a 'public coin'. We need 'worst case' round reduction arguments to prove lower bounds for the rounds complexity of communication games arising from data structure problems. This is because many of these lower bound proofs use some notion of "self-reducibility" arising from the original data structure problem which fails to hold in the 'average case', but holds for the 'worst case'. The quantum round reduction arguments of Klauck *et al.* [KNTZ01] are 'average case' arguments, and this is one of the reasons why they do not suffice to prove lower bounds for the rounds complexity of communication games arising from data structure problems.

Let us see what happens for the particular example of the rank parity communication game which is used to prove lower bounds for static predecessor. Recall the notation of Theorem 4.1 and its proof. Suppose we have a $(2t, a, b)^A$ communication protocol for the rank parity problem with small worst case error. Suppose we apply the self-reduction of Proposition 4.2, and then an 'average case' round reduction argument (e.g. a round reduction argument *à la* Klauck *et al*). After this, we get a $(2t-1, a', b')$ protocol, for some a', b', for the rank parity problem on a smaller domain. But now we can only guarantee that the *average error* of this protocol, for the uniform distribution on inputs, is small. In particular, when we try to apply the self-reduction of Proposition 4.3 next, we cannot guarantee that the average error, under the uniform distribution, on the kinds of inputs constructed in the proof of Proposition 4.3 is small. Hence, one needs 'worst case' round reduction arguments to prove lower bounds for the rounds complexity of the rank parity

communication game. 'Average case' round reduction arguments do not suffice.

Finally, note that Yao's minimax lemma is traditionally used in the context of public coin versus deterministic classical protocols. But it holds in the context of bounded error public coin versus coinless quantum protocols too.

5.2 Quantum information theoretic preliminaries

In this section, we discuss some basic facts from quantum information theory that will be used in the proof of the quantum round elimination lemma. We follow the notation of Klauck, Nayak, Ta-Shma and Zuckerman's paper [KNTZ01]. For a good account of quantum information theory, see the book by Nielsen and Chuang [NC00].

If A is a quantum system with density matrix ρ, then $S(A) \stackrel{\Delta}{=} S(\rho) \stackrel{\Delta}{=} -\mathrm{Tr}\ \rho \log \rho$ is the *von Neumann entropy* of A. If A, B are two disjoint quantum systems, their *mutual information* is defined as $I(A:B) \stackrel{\Delta}{=} S(A)+S(B)-S(AB)$. We now state some properties about von Neumann entropy and mutual information which will be useful later. The proofs follow easily from the definitions, using basic properties of von Neumann entropy like subadditivity and triangle inequality (see e.g. [NC00, Chapter 11]).

Lemma 5.2 *Suppose A, B, C are disjoint quantum systems. Then*

$$I(A:BC) = I(A:B) + I(AB:C) - I(B:C)$$

$$0 \leq I(A:B) \leq 2S(A)$$

If the Hilbert space of A has dimension d, then

$$0 \leq S(A) \leq \log d$$

Suppose X, Q are disjoint quantum systems with finite dimensional Hilbert spaces $\mathcal{H}, \mathcal{K}$ respectively. For every computational basis state $|x\rangle \in \mathcal{H}$, suppose σ_x is a density matrix in $\mathcal{K}$. Suppose the density matrix of (X, Q) is $\sum_x p_x |x\rangle\langle x| \otimes \sigma_x$, where $p_x > 0$ and $\sum_x p_x = 1$. Thus X is in a mixed state $\{p_x, |x\rangle\}$, and we shall say that X is a classical random variable and that Q is a quantum encoding $|x\rangle \mapsto \sigma_x$ of X. Define $\sigma \stackrel{\Delta}{=} \sum_x p_x \sigma_x$. σ is the reduced density matrix of Q, and we shall say that σ is the the density matrix of the average encoding. Then, $S(XQ) = S(X) + \sum_x p_x S(\sigma_x)$, and hence, $I(X:Q) = S(\sigma) - \sum_x p_x S(\sigma_x)$.

Let X, Y, Q be disjoint quantum systems with finite dimensional Hilbert spaces $\mathcal{H}, \mathcal{K}, \mathcal{L}$ respectively. Let $x \in \mathcal{H}$, $y \in \mathcal{K}$ be computational basis vectors. For every $|x\rangle|y\rangle \in \mathcal{H} \otimes \mathcal{K}$, suppose σ_{xy} is a density matrix in $\mathcal{L}$. Let Z refer to the quantum system (X, Y). Suppose (X, Y, Z) has density matrix $\sum_{x,y} p_{xy} |x\rangle\langle x| \otimes |y\rangle\langle y| \otimes \sigma_{xy}$, where $p_{xy} > 0$ and $\sum_{x,y} p_{xy} = 1$. Thus, X and Y are classical random variables, and $Z = XY$ is in a mixed state $\{p_{xy}, |x\rangle|y\rangle\}$. Q is a quantum encoding $|xy\rangle \mapsto \sigma_{xy}$ of Z. Define q_y^x to be the (conditional) probability that $Y = y$ given that $X = x$. $|y\rangle \mapsto \sigma_{xy}$ can be thought of as a quantum encoding Q^x of Y given that $X = x$. The joint density matrix of (Y, Q^x) is $\sum_y q_y^x |y\rangle\langle y| \otimes \sigma_{xy}$. We let $I((Y:Q)|X = x)$ denote the mutual information of this encoding.

We now prove the following propositions.

Proposition 5.1 *Let M_1, M_2 be disjoint finite dimensional quantum systems. Suppose $M \stackrel{\Delta}{=} (M_1, M_2)$ is a quantum encoding $|x\rangle \mapsto \sigma_x$ of a classical random variable X. Suppose the density matrix of M_2 is independent of X i.e. $Tr_{M_1} \sigma_x$ is the same for all x. Let M_1 be supported on a qubits. Then, $I(X : M) \leq 2a$.*

Proof: By Lemma 5.2, $I(X : M) = I(X : M_1M_2) = I(X : M_2) + I(XM_2 : M_1) - I(M_2 : M_1)$. But since the density matrix of M_2 is independent of X, $I(X : M_2) = 0$. Hence, by again using Lemma 5.2, we get that $I(X : M) \leq I(XM_2 : M_1) \leq 2S(M_1) \leq 2a$. ∎

Remark: This proposition is the key observation allowing us to "ignore" the size of the "safe" overhead M_2 in the round elimination lemma. It will be very useful in the applications of the round elimination lemma, where the complexity of the first message in the protocol increases quickly, but the blow up is confined to the "safe" overhead. Earlier round reduction arguments were unable to handle this large blow up in the complexity of the first message.

The next proposition has been observed by Klauck *et al.* [KNTZ01].

Proposition 5.2 *Suppose M is a quantum encoding of a classical random variable X. Suppose $X = X_1X_2 \dots X_n$, where the X_i are classical independent random variables. Then, $I(X_1 \dots X_n : M) = \sum_{i=1}^{n} I(X_i : MX_1 \dots X_{i-1})$.*

Proof: **(Sketch)** Similar to that of Proposition 4.4. ∎

Proposition 5.3 *Let X, Y be classical random variables and M be a quantum encoding of (X, Y). Then, $I(Y : MX) = I(X : Y) + E_X[I((Y : M)|X = x)]$.*

Proof: **(Sketch)** Similar to that of Proposition 4.5. ∎

For a linear operator A on a finite dimensional Hilbert space, the *trace norm* of A is defined as $\|A\|_t \stackrel{\Delta}{=} \text{Tr}\sqrt{A^\dagger A}$. The following fundamental theorem (see [AKN98]) shows that the trace distance between two density matrices ρ_1, ρ_2, $\|\rho_1 - \rho_2\|_t$, bounds how well one can distinguish between ρ_1, ρ_2 by a measurement.

Theorem 5.1 ([AKN98]) *Let ρ_1, ρ_2 be two density matrices on the same Hilbert space. Let $\mathcal{M}$ be a general measurement (i.e. a POVM), and $\mathcal{M}\rho_i$ denote the probability distributions on the (classical) outcomes of $\mathcal{M}$ got by performing measurement $\mathcal{M}$ on ρ_i. Let the ℓ_1 distance (total variation distance) between $\mathcal{M}\rho_1$ and $\mathcal{M}\rho_2$ be denoted by $\|\mathcal{M}\rho_1 - \mathcal{M}\rho_2\|_1$. Then*

$$\|\mathcal{M}\rho_1 - \mathcal{M}\rho_2\|_1 \leq \|\rho_1 - \rho_2\|_t$$

In fact the above upper bound is tight, and measuring in the orthonormal eigenbasis of $\rho_1 - \rho_2$ attains equality above.

Remark: This theorem will be used in the proof of the quantum round reduction lemma (Lemma 5.3). In the proof of the classical round reduction lemma (Lemma 4.4), we tacitly used the argument that if the total variation distance between the global states of Alice and Bob in two protocols is close, then the error probabilities of the two protocols have to be close. The above theorem can be thought of as the quantum version of this argument.

We will also need the following "local transition theorem" of Klauck *et al.* [KNTZ01].

Theorem 5.2 (Local transition, [KNTZ01]) *Let ρ_1, ρ_2 be two mixed states with support in a Hilbert space $\mathcal{H}$, $\mathcal{K}$ any Hilbert space of dimension at least the dimension of $\mathcal{H}$, and $|\phi_i\rangle$ any purifications of ρ_i in $\mathcal{H} \otimes \mathcal{K}$. Then, there is a local unitary transformation U on $\mathcal{K}$ that maps $|\phi_2\rangle$ to $|\phi_2'\rangle \stackrel{\Delta}{=} (I \otimes U)|\phi_2\rangle$ (I is the identity operator on $\mathcal{H}$) such that*

$$\||\phi_1\rangle\langle\phi_1| - |\phi_2'\rangle\langle\phi_2'|\|_t \leq 2\sqrt{\|\rho_1 - \rho_2\|_t}$$

Remark: In the proof of the classical round reduction lemma (Lemma 4.4), we created an intermediate protocol where the first message of Alice was independent of her input. This was done by generating Alice's message using a new private coin without "looking" at her input, and after that, adjusting Alice's old private coin in a suitable manner so as to be consistent with her message and input. In the proof of the quantum round reduction lemma (Lemma 5.3), we have to do a similar "blind" generation and "adjusting" procedure. The above theorem will be used in the "adjusting" procedure.

And finally, we will need the "average encoding theorem" of Klauck *et al.* [KNTZ01]. Intuitively speaking, it says that if the mutual information between a classical random variable and its quantum encoding is small, then the various quantum "codewords" are close to the "average codeword".

Theorem 5.3 (Average encoding, quantum version, [KNTZ01]) *Suppose that X, Q are two disjoint quantum systems, where X is a classical random variable, which takes value x with probability p_x, and Q is a quantum encoding $x \mapsto \sigma_x$ of X. Let the density matrix of the average encoding be $\sigma \stackrel{\Delta}{=} \sum_x p_x \sigma_x$. Then*

$$\sum_x p_x \|\sigma_x - \sigma\|_t \leq \sqrt{(2 \ln 2) I(X : Q)}$$

A proof of this theorem can be found in the appendix.

5.3 A quantum round reduction lemma

In this section, we prove a quantum round reduction lemma (Lemma 5.3), which will be required to prove the quantum round elimination lemma. The proof of Lemma 5.3 is similar to the proof of Lemma 4.4 in Klauck *et al.* [KNTZ01], but with a careful accounting of "safe" overheads in the messages communicated by Alice and Bob. Intuitively speaking, the lemma says that if the first message of Alice carries little information about her input,

under some probability distribution on inputs, then it can be eliminated, giving rise to a protocol where Bob starts, with one less round of communication, and the same message complexity and similar error probability, with respect to the same probability distribution on inputs. We observe, in the lemma below, that though there is a overhead of $l_1 + c$ qubits on the first message of Bob, it is a "safe" overhead.

For an input $(x, y) \in E \times F$, we define the error $\epsilon^P_{x,y}$ of the protocol P on (x, y), to be the probability that the result of P on input (x, y) is not equal to $f(x, y)$. For a protocol P, given a probability distribution D on $E \times F$, we define the average error ϵ^P_D of P with respect to D as the expectation over D of the error of P on inputs $(x, y) \in E \times F$. We define ϵ^P to be worst case error of P on inputs $(x, y) \in E \times F$.

Lemma 5.3 (Quantum round reduction lemma) *Suppose $f : E \times F \to G$ is a function. Let D be a probability distribution on $E \times F$, and P be a $[t, c, l_1, \ldots, l_t]^A$ safe coinless quantum protocol for f. Let X stand for the classical random variable denoting Alice's input (under distribution D), M be the first message of Alice in the protocol P, and $I(X : M)$ denote the mutual information between X and M under distribution D. Then there exists a $[t - 1, c + l_1, l_2, \ldots, l_t]^B$ safe coinless quantum protocol Q for f, such that*

$$\epsilon^Q_D \leq \epsilon^P_D + ((2 \ln 2) I(X : M))^{1/4}$$

Proof: We first give an overview of the plan of the proof, before getting down to the details. The proof proceeds in stages. We remark on the similarities between the stages in the quantum proof, and the stages in the classical proof (Lemma 4.4). Stages 1A and 1B of the quantum proof together correspond to Stage 1 of the classical proof, and Stages 2A and 2B of the quantum proof together correspond to Stage 2 of the classical proof.

Stage 1A: Starting from the $[t, c, l_1, \ldots, l_t]^A$ safe coinless protocol P, we construct a $[t, c, l_1, \ldots, l_t]^A$ safe coinless protocol $\tilde{P}$ with $\epsilon^{\tilde{P}}_{x,y} = \epsilon^P_{x,y}$ for every $(x, y) \in E \times F$. $\tilde{P}$ contains an extra "secure" copy of Alice's input $x \in E$, but is otherwise the same as P.

Stage 1B: Starting from $\tilde{P}$, we construct a $[t, c, l_1, \ldots, l_t]^A$ safe coinless protocol P', where the first message is independent of Alice's input, and $\epsilon^{P'}_D \leq \epsilon^{\tilde{P}}_D + ((2 \ln 2) I(X : M))^{1/4}$. The important idea in this step is to first generate Alice's average message (which is independent of her input), and after that, use the extra "secure" copy of Alice's input x to apply a unitary transformation U_x on some of her qubits without touching her message. U_x is used to adjust Alice's state in a suitable manner so as to be consistent with her input and message. This "adjustment" step requires the use of the "local transition theorem" (Theorem 5.2).

Stage 2A: Since in P' the first message is independent of Alice's input, Bob can generate it himself. But it is also necessary to achieve the correct entanglement between Alice's qubits and the first message (This is a uniquely quantum problem; in the classical setting we got away by requiring that the coin toss be done in public; the quantum solution to this

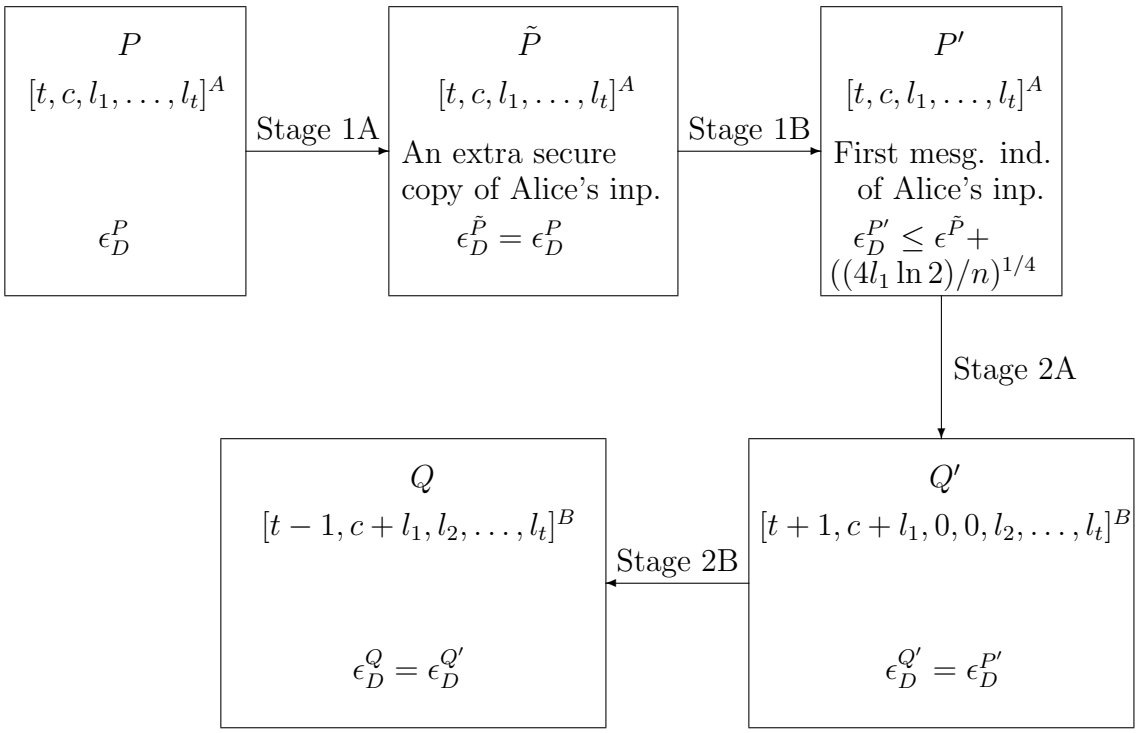


Figure 5.1: The various stages in the proof of Lemma 5.3.

problem lies in the "safe" overhead instead). Bob does this by first sending a safe message of $l_1 + c$ qubits. Alice then applies a unitary transformation V_x on some of her qubits, using the extra "secure" copy of her input x, to achieve the correct entanglement. The existence of such a V_x follows from Theorem 5.2. Doing all this gives us a $[t+1, c+l_1, 0, 0, l_2, \ldots, l_t]^B$ safe coinless protocol Q', such that $\epsilon_{x,y}^{Q'} = \epsilon_{x,y}^{P'}$ for every $(x, y) \in E \times F$.

Stage 2B: Since the first message of Alice in Q' is zero qubits long, Bob can concatenate his first two messages, giving us a $[t-1, c+l_1, l_2, \ldots, l_t]^B$ safe coinless protocol Q, such that $\epsilon_{x,y}^{Q} = \epsilon_{x,y}^{Q'}$ for every $(x, y) \in E \times F$. The technical reason behind this is that unitary transformations on disjoint sets of qubits commute.

The protocol Q of Stage 2B is our desired $[t-1, c+l_1, l_2, \ldots, l_t]^B$ safe coinless quantum protocol for f. We have

$$\epsilon_D^Q = \epsilon_D^{Q'} = \epsilon_D^{P'} \leq \epsilon_D^{\tilde{P}} + ((2 \ln 2) I(X : M))^{1/4} = \epsilon_D^P + ((2 \ln 2) I(X : M))^{1/4}$$

We now give the details of the proof. Let σ_x be the density matrix of the first message M of protocol P when Alice's input $X = x$. Let Y denote Bob's input register. Define $\sigma \stackrel{\Delta}{=} \sum_x p_x \sigma_x$, where p_x is the (marginal) probability of x under distribution D. σ is the density matrix of the average first message under distribution D. By the "secureness" of P, σ is also the density matrix of the first message when $|\psi\rangle$ is fed to Alice's input register

X, where $|\psi\rangle \stackrel{\Delta}{=} \sum_x \sqrt{p_x}|x\rangle$. By Theorem 5.3, we get that

$$\sum_x p_x \|\sigma_x - \sigma\|_t \leq \sqrt{(2\ln 2) I(X:M)}$$

Stage 1A: We first construct a $[t, c, l_1, \ldots, l_t]^A$ safe coinless quantum protocol $\tilde{P}$ for f such that $\epsilon^{\tilde{P}}_{x,y} = \epsilon^P_{x,y}$, for every $(x, y) \in E \times F$. Let X be Alice's input register in P. In $\tilde{P}$, Alice has an additional register C, and the input x to Alice is fed to register C, instead of X. X is initialised to $|0\rangle$ in $\tilde{P}$. In protocol $\tilde{P}$, Alice first copies the contents of C to X. After that, things in $\tilde{P}$ proceed as in P. Register C is not touched henceforth, and thus, C holds an extra "secure" copy of x throughout the run of protocol $\tilde{P}$.

Stage 1B: We now construct a $[t, c, l_1, \ldots, l_t]^A$ safe coinless quantum protocol P' for f with average error under distribution D, $\epsilon^{P'}_D \leq \epsilon^{\tilde{P}}_D + ((2\ln 2) I(X:M))^{1/4}$, and where the density matrix of the first message is independent of the input x to Alice. Alice is given $x \in E$ and Bob is given $y \in F$. Consider the situation in $\tilde{P}$ after the first message has been prepared by Alice, but before it is sent to Bob. Let register A denote Alice's qubits excluding the message qubits M and the qubits of the "secure" copy C (in particular, A includes the qubits of register X). Without loss of generality, one can assume that register A has at least $l_1 + c$ qubits, because one can initially pad up A with ancilla qubits set to $|0\rangle$. Let $|x\rangle_C \otimes |\theta_x\rangle_{AM}$ be the state vector of CAM in $\tilde{P}$ at this point, where the subscripts denote the registers. $|\theta_x\rangle_{AM}$ is a purification of σ_x. We note that $|\theta_x\rangle$ is also the state vector of AM in protocol P at this point. P' is similar to $\tilde{P}$ except for the following. Alice puts $|\psi\rangle$ in register X (instead of copying C to X as in $\tilde{P}$) to create the first message in register M with density matrix σ. AM now contains a purification $|\theta\rangle$ of σ. Then Alice applies a unitary transformation U_x depending upon x (which is available "securely" in register C) on A, so that $|\theta'_x\rangle_{AM} \stackrel{\Delta}{=} (U_x \otimes I)|\theta\rangle_{AM}$ is "close" to $|\theta_x\rangle_{AM}$. Here I stands for the identity transformation on M. Theorem 5.2 tells us that there exists a unitary transformation U_x on A such that

$$\||\theta_x\rangle\langle\theta_x| - |\theta'_x\rangle\langle\theta'_x|\|_t \leq 2\sqrt{\|\sigma_x - \sigma\|_t}$$

Thus, $|x\rangle_C \otimes |\theta'_x\rangle_{AM}$ is the state vector of CAM in P' after the application of U_x. Alice then sends register M to Bob and after this, Alice and Bob behave as in $\tilde{P}$. Application of U_x does not affect the density matrix of register M, which continues to be σ. Hence in P', the density matrix of the first message is independent of Alice's input.

Let us now compare the situations in protocols $\tilde{P}$ and P' when Alice's input is x, Bob's input is y, Alice has prepared her first message, but no communication has taken place as yet. At this point, in both protocols $\tilde{P}$ and P', the state vector of Bob's qubits is the same, and in tensor with the state vector of Alice's qubits. Let B denote the register of Bob's qubits (including his input qubits Y) and let $|\eta\rangle_B$ denote the state vector of B at this point. Hence the global state of protocol $\tilde{P}$ at this point is $|x\rangle_C \otimes |\theta_x\rangle_{AM} \otimes |\eta\rangle_B$, and

the global state of P' is $|x\rangle_C \otimes |\theta'_x\rangle_{AM} \otimes |\eta\rangle_B$. Therefore, the global states of protocols $\tilde{P}$ and P' at this point differ in trace distance by the quantity

$$\||x\rangle\langle x|\otimes|\theta_x\rangle\langle\theta_x|\otimes|\eta\rangle\langle\eta|-|x\rangle\langle x|\otimes|\theta'_x\rangle\langle\theta'_x|\otimes|\eta\rangle\langle\eta|\|_t = \||\theta_x\rangle\langle\theta_x|-|\theta'_x\rangle\langle\theta'_x|\|_t \leq 2\sqrt{\|\sigma_x-\sigma\|_t}$$

Using Theorem 5.1, we see that the error probability of P' on input x, y

$$\epsilon^{P'}_{x,y} \leq \epsilon^{\tilde{P}}_{x,y} + \frac{1}{2}\||x\rangle\langle x|\otimes|\theta_x\rangle\langle\theta_x|\otimes|\eta\rangle\langle\eta| - |x\rangle\langle x|\otimes|\theta'_x\rangle\langle\theta'_x|\otimes|\eta\rangle\langle\eta|\|_t \leq \epsilon^{\tilde{P}}_{x,y} + \sqrt{\|\sigma_x-\sigma\|_t}$$

Let q_{xy} be the probability that $(X, Y) = (x, y)$ under distribution D. Then, the average error of P' under distribution D, $\epsilon^{P'}_D$, is bounded by

$$\begin{aligned}
\epsilon^{P'}_D &= \sum_{x,y} q_{xy}\epsilon^{P'}_{x,y} \\
&\leq \sum_{x,y} q_{xy}\left(\epsilon^{\tilde{P}}_{x,y} + \sqrt{\|\sigma_x-\sigma\|_t}\right) \\
&\leq \epsilon^{\tilde{P}}_D + \sqrt{\sum_{x,y} q_{xy}\|\sigma_x-\sigma\|_t} \\
&= \epsilon^{\tilde{P}}_D + \sqrt{\sum_{x} p_x\|\sigma_x-\sigma\|_t} \\
&\leq \epsilon^{\tilde{P}}_D + ((2\ln 2)I(X:M))^{1/4}
\end{aligned}$$

For the second inequality above, we use the concavity of the square root function. The last inequality follows from the "average encoding theorem" (Theorem 5.3).

Stage 2A: We now construct a $[t+1, c+l_1, 0, 0, l_2, \ldots, l_t]^B$ safe coinless quantum protocol Q' for f with $\epsilon^{Q'}_{x,y} = \epsilon^{P'}_{x,y}$, for all $(x, y) \in E \times F$. Alice is given $x \in E$ and Bob is given $y \in F$. The protocol Q' will be constructed from P'. The input x is fed to register C of Alice, and the input y is fed to register Y of Bob. Let register G denote all the qubits of register A, except the last $l_1 + c$ qubits. In protocol Q' the registers initially in Alice's possession are C and G, and the registers initially in Bob's possession are B, M, and a new register R, where R is $l_1 + c$ qubits long. The qubits of G are initially set to $|0\rangle$. Bob first prepares the state vector $|\eta\rangle$ in register B as in protocol P'. He then constructs a canonical purification of σ in registers MR. The density matrix of M is σ. Bob then sends R to Alice. The density matrix of R is independent of the inputs x, y (in fact, if the canonical purification in MR is the Schmidt purification, then the density matrix of R is also σ). After receiving R, Alice treats GR as the register A in the remainder of the protocol. AM now contains a purification of σ. Alice applies a unitary transformation V_x depending upon x (which is available "securely" in register C) on A, so that the state vector of AM becomes $|\theta'_x\rangle_{AM}$. The existence of such a V_x follows from Theorem 5.2. At this point, the global state vector (over all the qubits of Alice and Bob) in Q' is the same

as the global state vector in P' viz. $|x\rangle_C \otimes |\theta'_x\rangle_{AM} \otimes |\eta\rangle_B$. Bob now treats register M as if it were the first message of Alice in P', and proceeds to compute his response N of length l_2. Bob sends N to Alice and after this protocol Q' proceeds as in P'. In Q' Bob starts the communication, the communication goes on for $t+1$ rounds, the first message of Bob of length $l_1 + c$ (i.e. register R) is a safe message, and the first message of Alice is zero qubits long.

Stage 2B: We finally construct a $[t-1, c+l_1, l_2, \ldots, l_t]^B$ safe coinless quantum protocol Q for f with $\epsilon^Q_{x,y} = \epsilon^{Q'}_{x,y}$, for all $(x, y) \in E \times F$. In protocol Q, Bob (after doing the same computations as in Q') first sends as a single message register RN of length $(l_1+c)+l_2$, and after that Alice applies V_x on A followed by her appropriate unitary transformation on AN (the unitary transformation of Alice in Q' on her qubits AN after she has received the first two messages of Bob). At this point, the global state vector (over all the qubits of Alice and Bob) in Q is the same as the global state vector in Q', since unitary transformations on disjoint sets of qubits commute. After this, things in Q proceed as in Q'. In protocol Q Bob starts the communication, the communication goes on for $t-1$ rounds, and the first message of Bob of length $(l_1 + c) + l_2$ contains a safe overhead (the register R) of $l_1 + c$ qubits.

This completes the proof of Lemma 5.3. ∎

5.4 The quantum round elimination lemma

We now prove the quantum round elimination lemma (for the communication game $f^{(n)}$). The proof of this lemma is similar to the proof of its classical twin (Lemma 4.5), but using the quantum round reduction lemma (Lemma 5.3) instead of the classical one (Lemma 4.4).

The round elimination lemma is stated for safe public coin quantum protocols only. Since a public coin quantum protocol can be converted to a coinless quantum protocol at the expense of an additional "safe" overhead in the first message, we also get a similar round elimination lemma for coinless protocols. We can decrease the overhead to logarithmic in the total bit size of the inputs by a technique similar to the public to private coins conversion for classical randomised protocols [New91]. But since the statement of the round elimination lemma is cleanest for safe public coin quantum protocols, we give it below for such protocols only.

Lemma 5.4 (Quantum round elimination lemma) *Suppose $f : E \times F \to G$ is a function. Suppose the communication game $f^{(n)}$ has a $[t, c, l_1, \ldots, l_t]^A$ safe public coin quantum protocol with worst case error less than δ. Then there is a $[t-1, c+l_1, l_2, \ldots, l_t]^B$ safe public coin quantum protocol for f with worst case error less than $\epsilon \stackrel{\Delta}{=} \delta + (4l_1 \ln 2/n)^{1/4}$.*

Proof: Suppose the given protocol for $f^{(n)}$ has worst case error $\tilde{\delta} < \delta$. Define $\tilde{\epsilon} \stackrel{\Delta}{=} \tilde{\delta} + (4l_1 \ln 2/n)^{1/4}$. To prove the quantum round elimination lemma it suffices to give, by the harder direction of the minimax lemma, for any probability distribution D on $E \times F$, a

$[t-1, c+l_1, l_2, \ldots, l_t]^B$ safe coinless quantum protocol P for f with average distributional error $\epsilon_D^P \leq \tilde{\epsilon} < \epsilon$. To this end, we will first construct a probability distribution D^* on $E^n \times [n] \times F$ as follows. Choose $i \in [n]$ uniformly at random. Choose independently, for each $j \in [n]$, $(x_j, y_j) \in E \times F$ according to distribution D. Set $y = y_i$ and throw away $y_j, j \neq i$. By the easier direction of the minimax lemma, we get a $[t, c, l_1, \ldots, l_t]^A$ safe coinless quantum protocol P^* for $f^{(n)}$ with distributional error, $\epsilon_{D^*}^{P^*} \leq \tilde{\delta} < \delta$. In P^*, Alice gets $x_1, \ldots, x_n$, Bob gets i, y and $x_1, \ldots, x_{i-1}$. We shall construct the desired protocol P from the protocol P^*.

Let M be the first message of Alice in P^*. By the definition of a safe protocol, M has two parts: M_1 l_1 qubits long, and the "safe" overhead M_2, c qubits long. Let the input to Alice be denoted by the classical random variable $X = X_1 X_2 \ldots X_n$ where X_i is the classical random variable corresponding to the ith input to Alice. Let the classical random variable Y denote the input y of Bob. Define $\epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*}$ to be the average error of P^* under distribution D^* when i is fixed and $X_1, \ldots, X_{i-1}$ are fixed to $x_1, \ldots, x_{i-1}$. Using Propositions 5.1, 5.2, 5.3 and the fact that under distribution D^*, $X_1, \ldots, X_n$ are independent classical random variables, we get that

$$\begin{aligned} \frac{2l_1}{n} &\geq \frac{I(X:M)}{n} \\ &= E_i[I(X_i : MX_1, \ldots, X_{i-1})] \\ &= E_{i,X}[I((X_i : M)|X_1, \ldots, X_{i-1} = x_1, \ldots, x_{i-1})] \end{aligned} \tag{5.1}$$

Also

$$\tilde{\delta} \geq \epsilon_{D^*}^{P^*} = E_{i,X}\left[\epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*}\right] \tag{5.2}$$

The expectations above are under distribution D^*.

For any $i \in [n]$, $x_1, \ldots, x_{i-1} \in E$, define the $[t, c, l_1, \ldots, l_t]^A$ safe coinless quantum protocol $P'_{i;x_1,\ldots,x_{i-1}}$ for the function f as follows. Alice is given $x \in E$ and Bob is given $y \in F$. Bob sets i to the given value, and both Alice and Bob set $X_1, \ldots, X_{i-1}$ to the values $x_1, \ldots, x_{i-1}$. Alice puts an independent copy of a pure state $|\psi\rangle$ (defined below) for each of the inputs $X_{i+1}, \ldots, X_n$. She sets $X_i = x$ and Bob sets $Y = y$. Then they run protocol P^* on these inputs. Here $|\psi\rangle \triangleq \sum_{x \in E} \sqrt{p_x}|x\rangle$, where p_x is the (marginal) probability of x under distribution D. Since P^* is a safe coinless quantum protocol, so is $P'_{i;x_1,\ldots,x_{i-1}}$. Because P^* is a secure protocol, the probability that $P'_{i;x_1,\ldots,x_{i-1}}$ makes an error for an input (x, y), $\epsilon_{x,y}^{P'_{i;x_1,\ldots,x_{i-1}}}$, is the average probability of error of P^* under distribution D^* when i is fixed to the given value, $X_1, \ldots, X_{i-1}$ are fixed to $x_1, \ldots, x_{i-1}$, and X_i, Y are fixed to x, y. Hence, the average probability of error of $P'_{i;x_1,\ldots,x_{i-1}}$ under distribution D

$$\epsilon_D^{P'_{i;x_1,\ldots,x_{i-1}}} = \epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*} \tag{5.3}$$

Let M' denote the first message of $P'_{i;x_1,\ldots,x_{i-1}}$ and X' denote the register X_i holding the input x to Alice. Because of the "secureness" of P^*, the density matrix of (X', M') in protocol $P'_{i;x_1,\ldots,x_{i-1}}$ is the same as the density matrix of (X_i, M) in protocol P^* when $X_1, \ldots, X_{i-1}$ are set to $x_1, \ldots, x_{i-1}$. Hence

$$I(X' : M') = I((X_i : M)|X_1, \ldots, X_{i-1} = x_1, \ldots, x_{i-1}) \tag{5.4}$$

Using Lemma 5.3 and equations (5.3) and (5.4), we get a $[t-1, c+l_1, l_2, \ldots, l_t]^B$ safe coinless quantum protocol $P_{i;x_1,\ldots,x_{i-1}}$ for f with

$$\begin{aligned} \epsilon_D^{P_{i;x_1,\ldots,x_{i-1}}} &\leq \epsilon_D^{P'_{i;x_1,\ldots,x_{i-1}}} + ((2\ln 2) I(X' : M'))^{1/4} \\ &= \epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*} + ((2\ln 2) I((X_i : M) | X_1, \ldots, X_{i-1} = x_1, \ldots, x_{i-1}))^{1/4} \end{aligned} \tag{5.5}$$

We have that (note that the expectations below are under distribution D^*)

$$\begin{aligned} E_{i,X}\left[\epsilon_D^{P_{i;x_1,\ldots,x_{i-1}}}\right] &\leq E_{i,X}\left[\epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*}\right] + \\ &\quad E_{i,X}\left[((2\ln 2) I((X_i : M) | X_1, \ldots, X_{i-1} = x_1, \ldots, x_{i-1}))^{1/4}\right] \\ &\leq E_{i,X}\left[\epsilon_{D^*;i;x_1,\ldots,x_{i-1}}^{P^*}\right] + \\ &\quad ((2\ln 2) E_{i,X}\left[I((X_i : M) | X_1, \ldots, X_{i-1} = x_1, \ldots, x_{i-1})\right])^{1/4} \\ &\leq \tilde{\delta} + \left(\tfrac{4 l_1 \ln 2}{n}\right)^{1/4} \\ &= \tilde{\epsilon} \end{aligned} \tag{5.6}$$

The first inequality follows from (5.5), the second inequality follows from the concavity of the fourth root function and the last inequality from from (5.1) and (5.2).

From (5.6), we see that there exist $i \in [n]$ and $x_1, \ldots, x_{i-1} \in E$ such that $\epsilon_D^{P_{i;x_1,\ldots,x_{i-1}}} \leq \tilde{\epsilon}$. Let $P \triangleq P_{i;x_1,\ldots,x_{i-1}}$. P is our desired $[t-1, c+l_1, l_2, \ldots, l_t]^B$ safe coinless quantum protocol for f with $\epsilon_D^P \leq \tilde{\epsilon}$, thus completing the proof of the quantum round elimination lemma. ■

5.5 Static predecessor: Optimal address-only quantum lower bounds

In this section, we prove our (optimal) lower bounds on the query complexity of static predecessor in the address-only quantum cell probe model.

Theorem 5.4 *Suppose we have a $(n^{O(1)}, (\log m)^{O(1)}, t)$ bounded error quantum address-only cell probe solution to the static predecessor problem, where the universe size is m and the subset size is at most n. Then the number of queries t is at least $\Omega\left(\frac{\log\log m}{\log\log\log m}\right)$ as a function of m, and at least $\Omega\left(\sqrt{\frac{\log n}{\log\log n}}\right)$ as a function of n.*

Proof: The proof is very similar to the proof of Theorem 4.3, but using the quantum round elimination lemma (Lemma 5.4).

By Proposition 4.1 (which continues to hold in the quantum setting by virtue of Lemma 5.1, it suffices to consider communication protocols for the rank parity communication game $\mathsf{PAR}_{\log m,n}$. Let $n = 2^{(\log\log m)^2/\log\log\log m}$. Let $c_1 \triangleq (4\ln 2)12^4$. For any given constants $c_2, c_3 \geq 1$, define

$$a \triangleq c_2 \log n \qquad b \triangleq (\log m)^{c_3} \qquad t \triangleq \frac{\log\log m}{(c_1 + c_2 + c_3)\log\log\log m}$$

We shall show that the rank parity communication game $\mathsf{PAR}_{\log m,n}$ does not have bounded error $(2t, 0, a, b)^A$ safe public coin quantum protocols, thus proving the desired lower bounds on the query complexity of static rank parity (and hence, static predecessor) by Lemma 5.1.

Given a $(2t, 0, a, b)^A$ safe public coin quantum protocol for $\mathsf{PAR}_{\log m,n}$ with error probability δ ($\delta < 1/3$), we get a $(2t, 0, a, b)^A$ safe public coin quantum protocol for

$$\mathsf{PAR}^{(c_1at^4),A}_{\frac{\log m}{c_1at^4},n}$$

with the same error probability δ, by Proposition 4.2. Using the quantum round elimination lemma (Lemma 5.4), we get a $(2t-1, a, a, b)^B$ safe public coin quantum protocol for

$$\mathsf{PAR}_{\frac{\log m}{c_1at^4},n}$$

but the error probability increases to at most $\delta + (12t)^{-1}$. Using the reduction of Proposition 4.3, we get a $(2t-1, a, a, b)^B$ safe public coin quantum protocol for

$$\mathsf{PAR}^{(c_1bt^4),B}_{\frac{\log m}{c_1at^4}-\log(c_1bt^4)-1,\frac{n}{c_1bt^4}}$$

with error probability at most $\delta + (12t)^{-1}$. From the given values of the parameters, we see that

$$\frac{\log m}{(2c_1at^4)^t} \geq \log(c_1bt^4) + 1$$

This implies that we also have a $(2t-1, a, a, b)^B$ safe public coin quantum protocol for

$$\mathsf{PAR}^{(c_1bt^4),B}_{\frac{\log m}{2c_1at^4},\frac{n}{c_1bt^4}}$$

with error probability at most $\delta + (12t)^{-1}$. Using the quantum round elimination lemma (Lemma 5.4) again, we get a $(2t-2, a+b, a, b)^A$ safe public coin quantum protocol for

$$\mathsf{PAR}_{\frac{\log m}{2c_1at^4},\frac{n}{c_1bt^4}}$$

but the error probability increases to at most $\delta + 2(12t)^{-1}$.

We do the above steps repeatedly. After applying the above steps i times, we get a $(2t-2i, i(a+b), a, b)^A$ safe public coin quantum protocol for

$$\mathsf{PAR}_{\frac{\log m}{(2c_1at^4)^i},\frac{n}{(c_1bt^4)^i}}$$

with error probability at most $\delta + 2i(12t)^{-1}$.

By applying the above steps t times, we finally get a $(0, t(a+b), a, b)^A$ safe public coin quantum protocol for

$$\mathsf{PAR}_{\frac{\log m}{(2c_1at^4)^t},\frac{n}{(c_1bt^4)^t}}$$

with error probability at most $\delta + 2t(12t)^{-1} < 1/2$. From the given values of the parameters, we see that

$$\frac{\log m}{(2c_1at^4)^t} \geq (\log m)^{\Omega(1)} \qquad \frac{n}{(c_1bt^4)^t} \geq n^{\Omega(1)}$$

Thus we get a zero round protocol for a rank parity problem on a non-trivial domain with error probability less than $1/2$, which is a contradiction.

In the above proof, we are tacitly ignoring "rounding off" problems. We remark that this does not affect the correctness of the proof. ∎

5.6 The 'greater-than' problem

We illustrate another application of the quantum round elimination lemma to quantum communication complexity by proving the first rounds versus communication tradeoffs for the 'greater-than' problem in the quantum setting.

Theorem 5.5 *The t round bounded error quantum communication complexity of GT_n is $\Omega(n^{1/t}t^{-3})$.*

Proof: We recall the following reduction from $GT^{(k)}_{n/k}$ to GT_n (see [MNSW98]): In $GT^{(k)}_{n/k}$, Alice is given $x_1, \ldots, x_k \in \{0,1\}^{n/k}$, Bob is given $i \in [k]$, $y \in \{0,1\}^{n/k}$, and copies of $x_1, \ldots, x_{i-1}$, and they have to communicate and decide if $x_i > y$. To reduce $GT^{(k)}_{n/k}$ to GT_n, Alice constructs $\tilde{x} \in \{0,1\}^n$ by concatenating $x_1, \ldots, x_k$, Bob constructs $\tilde{y} \in \{0,1\}^n$ by concatenating $x_1, \ldots, x_{i-1}, y, 1^{n(1-i/k)}$. It is easy to see that $\tilde{x} > \tilde{y}$ iff $x_i > y$.

Suppose GT_n has a $[t, 0, l_1, \ldots, l_t]^A$ safe public coin quantum protocol with worst case error probability less than $1/3$. Suppose

$$n \geq \left(Ct^3(l_1 + \cdots + l_t)\right)^t$$

where $C \triangleq (4\ln 2)6^4$. For $1 \leq i \leq t$, define

$$k_i \triangleq Ct^4 l_i \qquad n_i \triangleq \frac{n}{\prod_{j=1}^{i} k_j} \qquad \epsilon_i \triangleq \frac{1}{3} + \sum_{j=1}^{i} \left(\frac{(4\ln 2)l_j}{k_j}\right)^{1/4}$$

Also define $n_0 \triangleq n$ and $\epsilon_0 \triangleq 1/3$. Then

$$\epsilon_t \triangleq \frac{1}{3} + \sum_{j=1}^{t} \left(\frac{(4\ln 2)l_j}{k_j}\right)^{1/4} = \frac{1}{3} + \frac{t}{6t} = 1/2$$

and

$$n_t = \frac{n}{\prod_{j=1}^{t} k_j} = \frac{n}{(Ct^4)^t l_1 \cdots l_t} \geq \frac{nt^t}{C^t t^{4t}(l_1 + \cdots + l_t)^t} \geq 1$$

We now apply the above self-reduction and the quantum round elimination lemma (Lemma 5.4) alternately. Before the ith stage, we have a $[t-i+1, \sum_{j=1}^{i-1} l_j, l_i, \ldots, l_t]^Z$ safe public coin quantum protocol for $GT_{n_{i-1}}$ with worst case error probability less than ϵ_{i-1}. Here $Z = A$ if i is odd, $Z = B$ otherwise. For the ith stage, we apply the self-reduction with $k = k_i$. This gives us a $[t-i+1, \sum_{j=1}^{i-1} l_j, l_i, \ldots, l_t]^Z$ safe public coin quantum protocol for $GT_{n_i}^{(k_i)}$ with the same error probability. We then apply the quantum round elimination lemma (Lemma 5.4) to get a $[t-i, \sum_{j=1}^{i} l_j, l_{i+1}, \ldots, l_t]^{Z'}$ safe public coin quantum protocol for GT_{n_i} with worst case error probability less than ϵ_i. Here $Z' = B$ if $Z = A$ and $Z' = A$ if $Z = B$. This completes the ith stage.

Applying the self-reduction and the round elimination lemma alternately for t stages gives us a zero round quantum protocol for the 'greater-than' problem on a domain of size $n_t \geq 1$ with worst case error probability less than $\epsilon_t = 1/2$, which is a contradiction.

In the above proof, we are tacitly ignoring "rounding off" problems. We remark that this does not affect the correctness of the proof.

This proves the quantum lower bound of $\Omega(n^{1/t}t^{-3})$ on the message complexity. ■

Miltersen *et al.* [MNSW98] also use their round elimination lemma (Lemma 4.2) to prove lower bounds for other static data structure and communication complexity problems in the classical setting. We remark that all those results can be extended to the quantum setting by using the quantum round elimination lemma (Lemma 5.4).

Chapter 6

Conclusions and open problems

In this thesis, we have studied some problems in computational complexity in models of computation with an algebraic flavour. We have investigated the complexity of computing the degree two elementary symmetric polynomial $S_n^2(X)$ using $\Sigma\Pi\Sigma$ arithmetic circuits. We have studied the complexity of static membership and static predecessor in the quantum bit probe and quantum cell probe models. In the process, we have obtained a round elimination lemma in quantum communication complexity, which has implications to the complexity of some quantum communication problems, like the 'greater-than' problem. In this chapter, we conclude with a brief discussion of the results obtained and point out some open problems which arise naturally out of this work.

6.1 Computing $S_n^2(X)$ using $\Sigma\Pi\Sigma$ arithmetic circuits

6.1.1 Results

- We show an exact bound of $\lceil n/2 \rceil$, for infinitely many n, for the odd cover problem. We also show similar bounds on the number of multiplication gates in $\Sigma\Pi\Sigma$ arithmetic circuits computing $S_n^2(X)$ over GF(2).
- For any odd prime p, we show an upper bound of $\lceil n/2 \rceil$, for infinitely many n, for the 1 mod p cover problem.
- We show an exact bound of $\lceil n/2 \rceil$, for all n, on the number of multiplication gates in $\Sigma\Pi\Sigma$ arithmetic circuits computing $S_n^2(X)$ over $\mathbb{C}$. We also show similar, but weaker, bounds on the number of multiplication gates in $\Sigma\Pi\Sigma$ arithmetic circuits computing $S_n^2(X)$ over finite fields of odd characteristic.

6.1.2 Open problems

- In most of the cases, our exact bounds for computing $S_n^2(X)$ hold only for infinitely many n, but not for all n. Can this shortcoming be removed?

- Give tight bounds for computing the degree k elementary symmetric polynomial, $S_n^k(X)$, in the $\Sigma\Pi\Sigma$ model, for $k > 2$, and over various fields. In particular, can one prove a quadratic lower bound for $S_n^k(X)$ over $\mathbb{C}$ when $k = \sqrt{n}$?

- Give super polynomial lower bounds for inhomogeneous $\Sigma\Pi\Sigma$ circuits computing an explicit polynomial (e.g. determinant, permanent) over fields of characteristic zero.

6.2 Static membership problem

6.2.1 Results

- We show a tradeoff between space and the number of probes for any exact quantum bit probe scheme solving the static membership problem. The lower bounds obtained from this tradeoff match, within polynomials, to known upper bounds in the classical deterministic bit model.

- We show lower bounds on the storage space used by any two-sided ϵ-error quantum bit probe schemes making p probes. These bounds are almost matched by upper bounds in the classical bit probe model with two-sided error randomised query schemes.

- We show a $\Omega(\log n)$ lower bound on the number of probes made by any quantum cell probe solution of the static membership problem, with implicit storage schemes. This generalises a result of Yao [Yao81] to the bounded error quantum setting.

6.2.2 Open problems

- Buhrman et al. [BMRV00] consider 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, which make only one bit probe, 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.

6.3 Static predecessor problem

6.3.1 Results

- We prove a lower bound for the static predecessor problem in the bounded error address-only quantum cell probe model, matching the upper bound of Beame and Fich [BF99] for this problem in the classical deterministic cell probe model.

6.3.2 Open problems

- Our lower bound for static predecessor holds only in the address-only quantum cell probe model. Extending this result to the general quantum cell probe model, or showing that there are efficient schemes in this model, is an important open problem. The naive connection between quantum cell probe data structure problems and quantum communication complexity does not give us any hope for proving strong lower bounds in the general quantum cell probe model. Maybe, a new lower bound technique in quantum black box complexity is required for this.

6.4 Quantum communication complexity

6.4.1 Results

- We prove a round elimination lemma in classical communication complexity similar, but stronger, than the round elimination lemma of Miltersen *et al.* [MNSW98].

- We also prove a round elimination lemma in quantum communication complexity. The quantum round elimination lemma too is stronger than the round elimination lemma of Miltersen *et al.* [MNSW98].

- We use our round elimination lemmas to prove rounds versus communication tradeoffs for the 'greater-than' problem, in both quantum and classical settings. The quantum round elimination lemma should find application to other problems in quantum communication complexity as well.

6.4.2 Open problems

- The quantum round elimination lemma allows us to prove rounds-communication tradeoffs for various quantum communication complexity problems. *Pointer chasing* is a popular communication complexity problem to show rounds-communication tradeoffs. Optimal (or nearly optimal) rounds-communication tradeoffs are known for this problem in the classical deterministic and randomised setting, for both the full pointer and the bit versions [PRV01]. Recently, Klauck, Nayak, Ta-Shma and Zuckerman [KNTZ01] have shown a lower bound for the quantum communication complexity of pointer chasing, with the wrong player starting the communication. This bound is stronger than what can be proved using the quantum round elimination lemma (which is the bound Klauck *et al.* [KNTZ01] prove as their 'tree pointer jumping' result). But the lower bound of Klauck still does not match the classical upper bound. Also, the best quantum upper bound known is nothing but the classical upper bound. Thus, there is a gap here, and resolving it is an important open problem.

- Improve the rounds-communication tradeoffs for other problems in quantum communication complexity e.g. set disjointness. Rounds-communication tradeoffs for pointer chasing imply lower bounds on the bounded round communication complexity of set disjointness (see [KNTZ01]), but this method is insufficient to give lower bounds matching the best quantum upper bound of $O(\sqrt{n}c^{\log^* n})$ by Høyer and de Wolf [HdW01] for this problem. Høyer and de Wolf [HdW01] have also shown an $\Omega(\sqrt{n})$ lower bound for a restricted class of bounded error quantum protocols for the set disjointness problem. This restricted class of protocols encompasses their protocol and the protocol of Buhrman, Cleve and Wigderson [BCW98]. For general bounded error quantum protocols, the best lower bound known is $\Omega(\log n)$, arising from Kremer's result [Kre95] that the bounded error quantum communication complexity of a function is lower bounded (up to constant factors) by the logarithm of the one round (classical) deterministic communication complexity. Improving either the upper bound or the lower bound for set disjointness seems to require new ideas.

Bibliography

[Ajt88] M. Ajtai. A lower bound for finding predecessors in Yao's cell probe model. *Combinatorica*, 8(3):235–247, 1988.

[AKN98] D. Aharonov, A. Kitaev, and N. Nisan. Quantum circuits with mixed states. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 20–30, 1998. Also quant-ph/9806029.

[Alo86] N. Alon. Decomposition of the complete r-graph into complete r-partite r-graphs. *Graphs and Combinatorics*, 2:95–100, 1986.

[Amb99] A. Ambainis. A better lower bound for quantum algorithms searching an ordered list. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 352–357, 1999. Also quant-ph/9902053.

[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.

[Art91] M. Artin. *Algebra*. Prentice-Hall India Private Limited, 1991.

[AST+98] A. Ambainis, L. 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.

[BCW98] H. Buhrman, R. Cleve, and A. Wigderson. Quantum vs classical communication and computation. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 63–68, 1998. Also quant-ph/9802040.

[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.

[BF99] P. Beame and F. Fich. Optimal bounds for the predecessor problem. In *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 295–304, 1999.

[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.

[BS82] W. Baur and V. Strassen. The complexity of partial derivatives. *Theoretical Computer Science*, 22:317–330, 1982.

[CT91] T. Cover and J. Thomas. *Elements of Information Theory*. Wiley Series in Telecommunications. John Wiley and Sons, 1991.

[CvDNT98] R. Cleve, W. van Dam, M. Nielsen, and A. Tapp. Quantum entanglement and the communication complexity of the inner product function. In *Proceedings of the 1st NASA International Conference on Quantum Computing and Quantum Communications*, Lecture Notes in Computer Science, vol. 1509, pages 61–74. Springer-Verlag, 1998. Also quant-ph/9708019.

[dCH89] D. de Caen and D. Hoffman. Impossibility of decomposing the complete graph on n points into $n-1$ isomorphic complete bipartite graphs. *SIAM Journal of Discrete Mathematics*, 2:48–50, 1989.

[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.

[FGGS99] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. Invariant quantum algorithms for insertion into an ordered list. Manuscript at quant-ph/9901059, January 1999.

[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.

[GK98] D. Grigoriev and M. Karpinski. An exponential lower bound for depth-3 arithmetic circuits. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 577–582, 1998.

[GP72] R. Graham and H. Pollack. On embedding graphs in squashed cubes. In *Graph Theory and Applications*, Lecture Notes in Mathematics, volume 303, pages 99–110. Springer-Verlag, 1972.

[GR00] D. Grigoriev and A. Razborov. Exponential lower bounds for depth-3 arithmetic circuits in algebras of functions over finite fields. *Applicable Algebra in Engineering, Communication and Computing*, 10(6):465–487, 2000.

[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.

[Hal86] M. Hall Jr. *Combinatorial Theory*. Wiley Interscience series in Discrete Mathematics, 1986.

[Hås89] J. Håstad. Almost optimal lower bounds for small depth circuits. In S. Micali, editor, *Randomness and Computation*, volume 5 of *Advances in Computing Research*, pages 143–170. JAI Press, 1989.

[HdW01] P. Høyer and R. de Wolf. Improved quantum communication complexity bounds for disjointness and equality. Manuscript at quant-ph/0109068, September 2001.

[HNS01] P. Høyer, J. Neerbek, and Y. Shi. Quantum complexities of ordered searching, sorting, and element distinctness. In *Proceedings of the 28th International Colloquium on Automata, Languages and Programming*, pages 346–357, 2001. Also quant-ph/0102078.

[Kla00] H. Klauck. Quantum communication complexity. In *Proceedings of the Satellite Workshops at the 27th International Colloquium on Automata, Languages and Programming, Workshop on Boolean Functions and Applications (invited lecture)*, pages 241–252. Carleton Scientific, Waterloo, Ontario, Canada, 2000. Also quant-ph/0005032.

[KN96] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1996.

[KNTZ01] H. Klauck, A. Nayak, A. Ta-Shma, and D. Zuckerman. Interaction in quantum communication and the complexity of set disjointness. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing*, pages 124–133, 2001.

[Kre95] I. Kremer. Quantum communication. Master's thesis, Hebrew University, 1995.

[Mil94] P. B. Miltersen. Lower bounds for union-split-find related problems on random access machines. In *Proceedings of the 26th Annual ACM Symposium on Theory of Computing*, pages 625–634, 1994.

[Mil99] P. B. Miltersen. Cell probe complexity — a survey. In *Pre-conference workshop on Advances in Data Structures at the 19th conference on Foundations of Software Technology and Theoretical Computer Science (invited talk)*, 1999. Also available from http://www.daimi.au.dk/~bromille/Papers/survey3.ps.

[MNSW98] P. B. Miltersen, N. Nisan, S. Safra, and A. Wigderson. On data structures and asymmetric communication complexity. *Journal of Computer and System Sciences*, 57(1):37–49, 1998.

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

[NC00] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[New91] I. Newman. Private vs common random bits in communication complexity. *Information Processing Letters*, 39:67–71, 1991.

[Nis93] N. Nisan. The communication complexity of threshold gates. In *Combinatorics, Paul Erdős is Eighty (Vol. 1)*, pages 301–315. Janos Bolyai Mathematical Society, Budapest, Hungary, 1993.

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

[NW96] N. Nisan and A. Wigderson. Lower bounds on arithmetic circuits via partial derivatives. *Computational Complexity*, 6:217–234, 1996.

[NZM91] I. Niven, H. Zuckerman, and H. Montgomery. *An introduction to the theory of numbers*. John Wiley & Sons, Inc., 1991. Fifth edition.

[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.

[Pec84] G. Peck. A new proof of a theorem of Graham and Pollack. *Discrete Mathematics*, 49:327–328, 1984.

[PRV01] S. Ponzio, J. Radhakrishnan, and S. Venkatesh. The communication complexity of pointer chasing. *Journal of Computer and System Sciences*, 62(2):323–355, 2001.

[Raz87] A. Razborov. Lower bounds on the dimension of schemes of bounded depth in a complete basis containing the logical addition function. *Matematicheskie Zametki*, 41(4):598–607, 1987. (In Russian). English translation in *Mathematical Notes*, 41(3–4):333–338, 1987.

[RSV00a] J. Radhakrishnan, P. Sen, and S. Venkatesh. The quantum complexity of set membership. In *Proceedings of the 41st Annual IEEE Symposium on Foundations of Computer Science*, pages 554–562, 2000. Full version to appear in *Special issue of Algorithmica on Quantum Computation and Quantum Cryptography*. Also quant-ph/0007021.

[RSV00b] J. Radhakrishnan, P. Sen, and S. Vishwanathan. Depth-3 arithmetic circuits for $S_n^2(X)$ and extensions of the Graham-Pollack theorem. In *Proceedings of the 20th conference on the Foundations of Software Technology and Theoretical Computer Science*, Lecture Notes in Computer Science, vol. 1974, pages 176–187. Springer-Verlag, 2000. Also cs.DM/0110031.

[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.

[Sho97] P. Shor. Polynomial time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[Shp01] A. Shpilka. Affine projections of symmetric polynomials. In *Proceedings of the 16th Annual IEEE Conference on Computational Complexity*, pages 160–171, 2001.

[Smo87] R. Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proceedings of the 19th Annual ACM Symposium on Theory of Computing*, pages 77–82, 1987.

[Str73] V. Strassen. Die berechnungskomplexitat von elementarsymmetrischen funktionen und von interpolationskoefizienten. *Numerische Mathematik*, 20:238–251, 1973. (In German).

[SV01] P. Sen and S. Venkatesh. Lower bounds in the quantum cell probe model. In *Proceedings of the 28th International Colloquium on Automata, Languages and Programming*, Lecture Notes in Computer Science, vol. 2076, pages 358–369. Springer-Verlag, 2001. Also quant-ph/0104100.

[SW99] A. Shpilka and A. Wigderson. Depth-3 arithmetic formulae over fields of characteristic zero. In *Proceedings of the 14th Annual IEEE Conference on Computational Complexity*, pages 87–96, 1999.

[Tve82] H. Tverberg. On the decomposition of K_n into complete bipartite graphs. *Journal of Graph Theory*, 6:493–494, 1982.

[Xia92] B. Xiao. *New bounds in cell probe model*. PhD thesis, University of California at San Diego, 1992.

[Yao79] A. C-C. Yao. Some complexity questions related to distributed computing. In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing*, pages 209–213, 1979.

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

[Yao93] A. C-C. Yao. Quantum circuit complexity. In *Proceedings of the 34th Annual IEEE Symposium on Foundations of Computer Science*, pages 352–361, 1993.

Appendix A

A weaker version of Lemma 3.2

In this chapter, we give a complete proof of a weaker version of Lemma 3.2. In this version, we only get an $\Omega\left(\frac{\log n}{\log\log n}\right)$ lower bound, instead of the $\Omega(\log n)$ lower bound claimed in Lemma 3.2. The proof of the weaker version is given to illustrate the idea of using "logical intervals". By using "logical intervals", one can similarly modify Ambainis's $\Omega(\log n)$ lower bound for ordered searching [Amb99] to prove Lemma 3.2.

Remark: Combining the weaker version of Lemma 3.2 with the Ramsey theoretic arguments of Yao [Yao81], gives us a weaker $\Omega\left(\frac{\log n}{\log\log n}\right)$ version of Theorem 3.10.

A.1 A folklore proposition

We will require the following folklore proposition in what follows.

Proposition A.1 *Suppose $|\phi\rangle$, $|\psi\rangle$ are two state vectors. Suppose there is a boolean valued measurement $\mathcal{M}$ which gives 1 with probability at least $1-\epsilon$ if the state vector is $|\phi\rangle$, and with probability at most ϵ if the state vector is $|\psi\rangle$. Then*

$$\||\phi\rangle - |\psi\rangle\| \geq \sqrt{2(1-2\epsilon)}$$

Proof: Let V_1, V_0 denote the orthogonal subspaces for $\mathcal{M}$ corresponding to measurement outcomes 1, 0 respectively. Let $|\phi_1\rangle$, $|\psi_1\rangle$ denote the projections of $|\phi\rangle$, $|\psi\rangle$ respectively onto V_1. Let $|\phi_0\rangle$, $|\psi_0\rangle$ denote the respective projections onto V_0. Then $\||\phi_0\rangle\|, \||\psi_1\rangle\| \leq \sqrt{\epsilon}$. Hence

$$\begin{aligned} |\langle\phi|\psi\rangle| &= |\langle\phi_0|\psi_0\rangle + \langle\phi_1|\psi_1\rangle| \\ &\leq \||\phi_0\rangle\|\||\psi_0\rangle\| + \||\phi_1\rangle\|\||\psi_1\rangle\| \\ &\leq 2\sqrt{\epsilon} \end{aligned}$$

Therefore

$$\begin{aligned} \| |\phi\rangle - |\psi\rangle \|^2 &= \| |\phi\rangle \|^2 + \| |\psi\rangle \|^2 - \langle\phi|\psi\rangle - \langle\psi|\phi\rangle \\ &= 2 - 2 \cdot \text{Re}\left(\langle\phi|\psi\rangle\right) \\ &\geq 2 - 2 \cdot |\langle\phi|\psi\rangle| \\ &\geq 2 - 4\sqrt{\epsilon} \end{aligned}$$

■

A.2 Proof of the weaker version of Lemma 3.2

We now prove the weaker version of Lemma 3.2.

Lemma 3.2 (weak version) *Suppose S is an n element subset of the universe $[m]$, where $m \geq 2n$. If the storage scheme is implicit, always stores the same 'pointer' values in the same locations, and in the remaining locations, stores the elements of S in a fixed order (repetitions of an element are allowed, but all elements have to be stored) based on their relative ranking in S, then $\Omega\left(\frac{\log n}{\log\log n}\right)$ probes are needed by any bounded error quantum cell query strategy to answer membership queries.*

Proof: The proof is via a 'hybrid' adversary argument. Consider the behaviour of the quantum query scheme with query element n. Suppose the query scheme uses less than $t \triangleq \frac{\log n}{2\log\log n}$ cell queries. The adversary shall construct two sets $A, B \subseteq [m]$, $|A| = |B| = n$, such that $n \in A$, $n \notin B$, but the query scheme gives the same answer for A and B, which is a contradiction.

The adversary's strategy is as follows. In the first stage, he partitions the "logical interval" $I_0 \triangleq [1, \ldots, n]$ into $\log^2 n$ "logical subintervals" of length $n/\log^2 n$ each. He simulates the query scheme up to the first query. Let $|\phi_0\rangle$ be the state vector of the query scheme before the first query. There is a "logical subinterval"

$$I_1 \triangleq \left[\frac{(l-1)n}{\log^2 n} + 1, \ldots, \frac{ln}{\log^2 n}\right]$$

where $1 \leq l \leq \log^2 n$, that is queried by $|\phi_0\rangle$ with probability at most $1/\log^2 n$. The adversary answers the first query according to the oracle for the set

$$T_1 \triangleq \left\{1, \ldots, \frac{(l-1)n}{\log^2 n}\right\} \bigcup \left\{m - n + \frac{(l-1)n}{\log^2 n} + 1, \ldots, m\right\}$$

In the second stage, the adversary splits the "logical interval" I_1 into $\log^2 n$ "logical subintervals" of length $n/\log^4 n$ each. He simulates the query scheme up to the second query. Let $|\phi_1\rangle$ be the state vector of the query scheme before the second query. There is a "logical subinterval"

$$I_2 \triangleq \left[\frac{(l-1)n}{\log^2 n} + \frac{(k-1)n}{\log^4 n} + 1, \ldots, \frac{(l-1)n}{\log^2 n} + \frac{kn}{\log^4 n}\right]$$

where $1 \leq k \leq \log^2 n$, that is queried by $|\phi_1\rangle$ with probability at most $1/\log^2 n$. The adversary answers the second query according to the oracle for the set

$$T_2 \stackrel{\Delta}{=} \left\{1, \ldots, \frac{(l-1)n}{\log^2 n} + \frac{(k-1)n}{\log^4 n}\right\} \bigcup \left\{m - n + \frac{(l-1)n}{\log^2 n} + \frac{(k-1)n}{\log^4 n} + 1, \ldots, m\right\}$$

The adversary repeats the splitting in this fashion until the "logical interval" is smaller than $\log^2 n$ in length. This means that he can do up to $t \stackrel{\Delta}{=} \frac{\log n}{2 \log \log n}$ splittings. Let $|\phi_{i-1}\rangle$ denote the state vector of the query scheme before the ith query, and T_i be the set according to whose oracle the adversary answers the ith query, in this simulation.

Let $[i+1, \ldots, j]$ be the final "logical interval", at the end of the adversary's simulation. Define two sets $A, B \subseteq [m]$ as follows.

$$A \stackrel{\Delta}{=} \{1, \ldots, i\} \cup \{n\} \cup \{m - n + i + 2, \ldots, m\}$$

$$B \stackrel{\Delta}{=} \{1, \ldots, i\} \cup \{n+1\} \cup \{m - n + i + 2, \ldots, m\}$$

We have that $|A| = |B| = n$, $n \in A$ and $n \notin B$.

We now do a standard 'hybrid' argument. The quantum query scheme is a sequence of unitary transformations

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

where U_j's are arbitrary unitary transformations that do not depend on the set stored (representing the internal computations of the query algorithm), and O_S represents the oracle for the stored set S. Define $|\alpha_{i-1}\rangle$, $|\beta_{i-1}\rangle$ to be the state vectors of the query scheme before the ith query when sets A, B respectively are stored. We shall show that

$$\||\phi_i\rangle - |\alpha_i\rangle\| \leq \frac{2i}{\log n} \qquad \||\phi_i\rangle - |\beta_i\rangle\| \leq \frac{2i}{\log n} \tag{A.1}$$

The proof of (A.1) is by induction on i. It is true for $i = 0$, since $|\phi_0\rangle = |\alpha_0\rangle = |\beta_0\rangle$. Suppose it is true for $i-1$. We prove it for i as follows. Let O_{T_i}, O_A be the oracle unitary transformations for sets T_i, A respectively.

$$\begin{aligned}
\||\phi_i\rangle - |\alpha_i\rangle\| &= \|U_i O_{T_i} |\phi_{i-1}\rangle - U_i O_A |\alpha_{i-1}\rangle\| \\
&= \|O_{T_i} |\phi_{i-1}\rangle - O_A |\alpha_{i-1}\rangle\| \\
&\leq \|O_{T_i} |\phi_{i-1}\rangle - O_A |\phi_{i-1}\rangle\| + \|O_A |\phi_{i-1}\rangle - O_A |\alpha_{i-1}\rangle\| \\
&\leq \frac{2}{\log n} + \||\phi_{i-1}\rangle - |\alpha_{i-1}\rangle\| \\
&\leq \frac{2}{\log n} + \frac{2(i-1)}{\log n} \\
&= \frac{2i}{\log n}
\end{aligned}$$

The second inequality above follows from the fact that T_i and A differ only in the "logical interval" I_i, which is queried with probability at most $1/\log^2 n$ by $|\phi_{i-1}\rangle$. The third inequality follows from the induction hypothesis. Thus, we have proved the first inequality in (A.1). The proof of the second inequality in (A.1) is similar.

By plugging in $i = t$ in (A.1) we get

$$\begin{aligned} \||\alpha_t\rangle - |\beta_t\rangle\| &\leq \||\alpha_t\rangle - |\phi_t\rangle\| + \||\phi_t\rangle - |\beta_t\rangle\| \\ &\leq \left(\frac{2}{\log n}\right)\left(\frac{\log n}{2\log\log n}\right) + \left(\frac{2}{\log n}\right)\left(\frac{\log n}{2\log\log n}\right) \\ &= \frac{2}{\log\log n} \end{aligned}$$

Since the quantum query scheme errs with probability at most $1/3$, by Proposition A.1, we also get that $\||\alpha_t\rangle - |\beta_t\rangle\| \geq \sqrt{2/3}$, which is a contradiction. This finishes the proof of the lemma. ■

Appendix B

The average encoding theorem

In this chapter, we give a proof of the quantum average encoding theorem (Theorem 5.3). We also show how one can prove the classical average encoding theorem (Theorem 4.2) without appealing to quantum mechanics.

B.1 The classical average encoding theorem

We require a non-trivial theorem from classical information theory. To state the theorem, we need the following definition of *information divergence*. A proof of the theorem can be found in the book by Cover and Thomas [CT91].

Definition B.1 (Information divergence) *Let P, Q be probability distributions on the same finite sample space Ω. Let p_x (q_x) denote the probability of the sample point $x \in \Omega$ under P (Q). The* information divergence *between P and Q, denoted by $D(P : Q)$, is defined as*

$$D(P : Q) \stackrel{\Delta}{=} \sum_{x \in \Omega} p_x \log \left(\frac{p_x}{q_x} \right)$$

Theorem B.1 ([CT91, Lemma 12.6.1]) *Let P and Q be probability distributions on the same finite sample space Ω. Then*

$$D(P : Q) \geq \frac{1}{2 \ln 2} \|P - Q\|_1^2$$

We can now prove the classical average encoding theorem.

Theorem 4.2 (Average encoding, classical version, [KNTZ01]) *Let X be a classical random variable which takes value x with probability p_x, and M be a classical randomised encoding $x \mapsto \sigma_x$ of X, where σ_x is a probability distribution over the sample space of codewords. The probability distribution of the average encoding is $\sigma \stackrel{\Delta}{=} \sum_x p_x \sigma_x$. Then*

$$\sum_x p_x \|\sigma_x - \sigma\|_1 \leq \sqrt{(2 \ln 2) I(X : M)}$$

Proof: Let S, T be the (finite) ranges of random variables X, M respectively. We define two probability distributions P, Q on $S \times T$. In distribution P, the probability of $(x, m) \in S \times T$ is $p_x \cdot \sigma(m \mid x)$, where $\sigma(m \mid x)$ is the probability that $M = m$ given that $X = x$. In distribution Q, the probability of $(x, m) \in S \times T$ is $p_x \cdot \sigma(m)$, where $\sigma(m)$ is the probability of message m in the average encoding i.e. $\sigma(m) \stackrel{\Delta}{=} \sum_x p_x \sigma(m \mid x)$.

One can easily check that

$$D(P : Q) = I(X : M) \qquad \|P - Q\|_1 = \sum_x p_x \|\sigma_x - \sigma\|_1$$

The result now follows by applying Theorem B.1 to P and Q. ∎

B.2 The quantum average encoding theorem

To prove the quantum average encoding theorem, we need to define the quantum analogue of information divergence, called the *relative von Neumann entropy*.

Definition B.2 (Relative von Neumann entropy) *Let ρ, σ be density matrices on the same finite dimensional Hilbert space. The* relative von Neumann entropy *between ρ and σ, denoted by $S(\rho|\sigma)$, is defined as*

$$S(\rho|\sigma) \stackrel{\Delta}{=} \text{Tr } (\rho(\log \rho - \log \sigma))$$

We also need a quantum analogue of Theorem B.1, which has been proved by Klauck *et al.* [KNTZ01].

Theorem B.2 ([KNTZ01]) *Let ρ, σ be density matrices over the same finite dimensional Hilbert space $\mathcal{H}$. Then*

$$S(\rho|\sigma) \geq \frac{1}{2 \ln 2} \|\rho - \sigma\|_t^2$$

Proof: Let $\mathcal{M}$ be a measurement operator measuring in the orthonormal eigenbasis of $\rho - \sigma$. Then, by Theorem 5.1

$$\|\mathcal{M}\rho - \mathcal{M}\sigma\|_1 = \|\rho - \sigma\|_t$$

where $\mathcal{M}\rho$, $\mathcal{M}\sigma$ denote the probability distributions on the (classical) outcomes of $\mathcal{M}$ got by performing measurement $\mathcal{M}$ on ρ, σ respectively. By the Lindblad-Uhlmann monotonicity theorem (see e.g. [NC00, Theorem 11.17])

$$S(\rho|\sigma) \geq D(\mathcal{M}\rho : \mathcal{M}\sigma)$$

We complete the proof by invoking Theorem B.1. ∎

We can now prove the quantum average encoding theorem in a similar fashion as its classical twin.

Theorem 5.3 (Average encoding, quantum version, [KNTZ01]) *Suppose that X, Q are two disjoint quantum systems, where X is a classical random variable, which takes value x with probability p_x, and Q is a quantum encoding $x \mapsto \sigma_x$ of X. Let the density matrix of the average encoding be $\sigma \stackrel{\Delta}{=} \sum_x p_x \sigma_x$. Then*

$$\sum_x p_x \|\sigma_x - \sigma\|_t \leq \sqrt{(2 \ln 2) I(X:Q)}$$

Proof: Let the joint density matrix of (X, Q) be $\rho_1 \stackrel{\Delta}{=} \sum_x p_x |x\rangle\langle x| \otimes \sigma_x$. Define another density matrix $\rho_2 \stackrel{\Delta}{=} (\sum_x p_x |x\rangle\langle x|) \otimes \sigma$.

One can easily check that

$$S(\rho_1|\rho_2) = I(X:M) \qquad \|\rho_1 - \rho_2\|_t = \sum_x p_x \|\sigma_x - \sigma\|_t$$

The result now follows by applying Theorem B.2 to ρ_1 and ρ_2. ∎