On quantum and probabilistic communication: Las Vegas and one-way protocols*

Hartmut Klauck
Johann Wolfgang Goethe-Universität Frankfurt
60054 Frankfurt am Main, Germany
klauck@thi.informatik.uni-frankfurt.de

Abstract

We investigate the power of quantum communication protocols compared to classical probabilistic protocols. In our first result we describe a total Boolean function that has a quantum Las Vegas protocol communicating at most $O(N^{10/11+\epsilon})$ qubits for all $\epsilon > 0$, while any classical probabilistic protocol (with bounded error) needs $\Omega(N/\log N)$ bits. Then we investigate quantum one-way communication complexity. First we show that the VC-dimension lower bound on one-way probabilistic communication of [26] holds for quantum protocols, too. Then we prove that for one-way protocols computing total functions quantum Las Vegas communication is asymptotically as efficient as exact quantum communication, which is exactly as efficient as deterministic communication. We describe applications of the lower bounds for one-way communication complexity to quantum finite automata and quantum formulae.

1 Introduction

The investigation of the power of quantum computation has attracted much interest recently, mainly due to the breakthrough results of Shor on factoring integers in quantum polynomial time [36] and Grover on searching an item in an unordered database of n elements with $O(\sqrt{n})$ queries [15]. To understand what possibilities and restrictions lie in quantum mechanical computing, several classical restricted models of computation have been redefined as quantum computing systems, e.g. quantum finite automata, quantum decision trees (aka black box computations), quantum formulae, and quantum communication protocols (see [24], [15], [38]).
The nice applications of communication complexity results to lower bound proofs are motivation to study this model also in the quantum case (introduction to (classical) communication complexity can be found in [27] and [19],

*This paper has appeared in the Proceedings of the 32nd ACM Symposium on Theory of Computing, 2000.

an overview on quantum communication complexity in [37]). In a quantum protocol (as defined in [38]) the players exchange qubits rather than bits. Another scenario, where the players may also possess some (input independent) qubits that are entangled with the other players qubits is proposed in [11] and [12]. Due to the "superdense coding" technique of [5] in this model 2 classical bits can be communicated by transmitting one qubit (and using one entangled qubit). The main objective of quantum communication complexity theory is to determine the maximum speedup one can get in comparison to classical communication for the different modes of acceptance: exact protocols (without error), bounded error, and Las Vegas (no error, but only the expected communication is measured).

While quite a lot of lower bound methods are known for classical communication complexity, so far only few lower bound methods for quantum protocols are known: the rank lower bound is known to hold for exact quantum communication [7], the discrepancy lower bound for bounded error protocols [25]. But the latter technique is not strong enough to prove e.g. a superlogarithmic lower bound on the bounded error quantum communication complexity of the disjointness problem DISJ (both players receive an incidence vector of a subset of $\{1, \ldots, n\}$ and have to decide whether the sets are not disjoint: $\bigvee(x_i \wedge y_i)$), maybe the most important communication problem. In [7] an upper bound of $O(\sqrt{n} \log n)$ is shown for this problem by an application of Grover's quantum search algorithm. This yields the largest known gap between quantum and classical communication for a total function, the probabilistic communication of disjointness is $\Omega(n)$ [21]. Note that the exponential gap between bounded error quantum and classical communication of [34] holds for a partial function only. An exponential gap between exact quantum and deterministic classical communication is also known for a partial function [7].

The generalization of Grover's black box search problem (corresponding to an OR problem) to constant depth trees of ANDs and ORs is solvable with almost quadratic speedup [7]. [8] prove that this technique can be extended to get an almost quadratic gap between quantum Las Vegas and classical probabilistic black box-computations. By a reduction to communication complexity problems this yields a polynomial gap between quantum Las Vegas and probabilistic communication for some specific relation. This result inspired the first result of our paper: a polynomial gap for for the same modes of computation, but this time for a total function. Whereas the upper bound uses the tree evaluation of [8] the lower bound employs the heavy machinery of [20], where protocols with limited nondeterminism are investigated. In fact we prove that a small, but still hard subset of the language considered in [20] can be decided by an efficient quantum protocol. The complicated lower bound method is needed, because we have to prove a lower bound on probabilistic communication while nondeterministic communication is small (since we want to employ Grover search). Apart from [20] the only result separating these complexity modes is given in [4] (for which we describe an improvement).

The quantum protocols discovered so far that achieve a polynomial speedup for total functions compared to classical protocols share the feature of using a

large amount of communication rounds (which comes from using Grover search). We turn our interest to the question how efficient communication problems can be solved in the quantum world when the number of rounds is restricted. A very restricted model is one-way communication, where only a monologue is transmitted from one player to the other, who decides the function value. This model has important applications for automata [14] and also allows to rephrase the Nečiporuk lower bound on formula size nicely [22]. Kremer [25] investigates quantum one-way communication and exhibits a complete problem for polylog quantum communication (with bounded error). We show that the VC-dimension lower bound of [26] can be extended to the (bounded error) quantum case. We get a tight bound by using the results of [3] on random access quantum coding. We furthermore show that for total functions and one-way protocols exact quantum communication is exactly as powerful as deterministic communication and asymptotically as efficient as quantum Las Vegas communication. Similar results hold in the model with prior entanglement.

We conclude from these results that quantum communication is dependent on rounds. Good speedups for total function seem to use rounds, this is provably true for all nonconstant gaps for the Las Vegas mode and for the disjointness problem in bounded error mode. It is not clear, whether quantum compared to classical communication ever helps to save rounds for total function, e.g. if there is a total function with no efficient classical probabilistic one-way protocol, but an efficient bounded error quantum one-way protocol.

Quantum finite automata (qfa's) are defined in [24]. The following is known about one-way qfa's (with bounded error): they may be exponentially more succinct than classical automata [2], but cannot compute every regular language [24], furthermore they can be exponentially larger than dfa's for some finite languages [3],[28]. We provide a combinatorial tool to prove lower bounds on qfa's by considering the VC-dimension of a matrix associated with the function. Another corollary is that for total function exact qfa's are never more succinct than dfa's and Las Vegas qfa's (qfa's without error that may "give up" with probability ϵ) have size at least $D^{1-\epsilon}$ for the minimal dfa size D. We show these results even for a generalized version of quantum finite automata which are at least as strong as qfa and as probabilistic finite automata.

A second application of one-way complexity is the Nečiporuk lower bound on formula size. Recently [35] proved that the classical Nečiporuk function is essentially also a lower bound on quantum formula size (within a logarithmic factor). This is astonishing, since the bound does not hold for probabilistic formulae [22]. Therefore we can conclude that the model of quantum formulae considered in [38] is restricted and we propose to consider more general possibilities of measurements as in the definitions of quantum circuits with mixed states defined in [1] and additionally consider additional input variables for the formulae to allow the simulation of random inputs. For the generalized model of quantum formulae a lower bound based on the VC-dimension is proved.

In the last section we take a look at communication with less severe interaction limitations. A series of papers (see [13], [17], [31], [32], [23]) give round hierarchies for classical protocols of the following form: A function f_k (usually the so

called pointer jumping function) can be computed in, say, $k \log n$ communication in k rounds when A starts, but needs large communication when B starts. We note that proving a lower bound of $c(n)$ for quantum communication would imply via reductions also lower bounds for the k round bounded error quantum communication complexity of the disjointness problem.

2 Preliminaries

In this section we provide definitions of the models considered.

Definition 1 *Let $f : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$ be a function. In a communication protocol player A and B receive x and y and compute $f(x,y)$. Classically the players exchange binary messages. The communication complexity of a protocol is the worst case number of bits exchanged. The deterministic communication complexity $D(f)$ of f is the complexity of an optimal protocol for f.*
In a randomized protocol both players have access to public random bits. The output is required to be correct with probability $1-\epsilon$ for some constant ϵ. The randomized communication complexity of a function $R_\epsilon(f)$ is then defined analogously to the deterministic communication complexity. A protocol is called Monte Carlo, if the error is one-sided (no input with $f(x,y) = 0$ is accepted), and called Las Vegas, if no error is allowed, but only expected communication is measured, the notation is $R_0(f)$. Alternatively we consider protocols that may give up without answer with probability at most ϵ (but never err), the notation is $R_{0,\epsilon}(f)$.
A protocol has k rounds, if the players exchange k messages with A and B alternating as speakers. A protocol is called one-way if only one players sends a message. The complexity notations are superscripted with the number of allowed rounds and eventually with the player starting, like $C^{k,B}$ or C^1 (usually A starts).
The communication matrix of f is a matrix with $M[x,y] = f(x,y)$. Furthermore we will consider nondeterministic protocols, which are defined in the standard way.

Now we define quantum communication protocols. For general information on quantum computation see [16] and [33].

Definition 2 *In a quantum protocol both players have a private set of qubits. Some of the qubits are initialized to the input before the start of the protocol, the other qubits are in state $|0\rangle$. In a communication round each player can perform some unitary transformation on his qubits and then one sends some of the qubits in his possession to the other player (at the cost of the number of qubits sent). After the end of the protocol the state of some qubits is measured and (a part of) the result is taken as the output.*
In an exact quantum protocol the result must be computed correctly with probability 1. The exact quantum complexity of a function $Q_E(f)$ is the minimal

complexity of an exact quantum protocol for f. In a (bounded error) quantum protocol the correct answer must be given with probability $1-\epsilon$ for some $1/2 > \epsilon > 0$. The (bounded error) quantum complexity of a function is called $Q_\epsilon(f)$. For a quantum Las Vegas protocol acceptance in defined as for classical Las Vegas protocols, the notations are $Q_0(f)$ and $Q_{0,\epsilon}(f)$. Bounded rounds are indicated as above.
[11] and [12] propose a different model of quantum communication: A and B may possess an arbitrary input-independent set of (entangled) qubits. Then they communicate, where we consider qubit communication. This model can be simulated by allowing first an input-independent communication (to set up the entanglement) with no cost and then a communication with cost. The superdense coding technique of [5] allows to transmit n bits of information with $\lceil n/2 \rceil$ qubits in this model. We denote the complexity in this model by the superscript e.

For the definition of one-way qfa's we refer to [24] and [28]. Our results hold for the models defined in both papers. We furthermore consider exact qfa's (no error allowed) and Las Vegas qfa's (that may give up with some probability ϵ, but never err).
Nayak [28] has proposed to consider a model of quantum automata in which arbitrary measurements are allowed. Here we propose a formal definition of such a model, in which those measurements are controlled by a classical finite automaton.

Definition 3 *A quantum/classical (finite) automaton is desribed by a tuple $(Q, \Sigma, F_a, F_r, q_0, k, S, \delta)$, where Q is a finite set of classical states, Σ is an alphabet, $F_a \subseteq Q$ the set of accepting states, $F_r \subseteq Q$ the set of rejecting states, $F_a \cap F_r = \emptyset$, q_0 is the starting state, k is the dimension of the Hilbert space $\mathbb{C}^k$ (which describes the quantum memory of the automaton). The set S consists of pairs of unitary transformation and observables for the Hilbert space $\mathbb{C}^k$, $\delta : Q \times \{1, \ldots, k\} \times \Sigma \to Q \times S$ is the transition function.*
A computation starts in state q_0. Also in the beginning set $m = 1$ and the initial quantum state is $|0\rangle$. When a letter $a \in \Sigma$ is read, the transition function is applied to the state, m and a, and results in a new state plus a superoperator (consisting of a unitary transformation and a measurement). The superoperator is applied to the quantum state and yields a new quantum state plus a measurement result, which is stored in m (there are at most k different results). The computation ends, after the input has been read. The automaton accepts, if the (classical) state is in F_a, and rejects, if the state is in F_r, otherwise the automaton gives up without result.
The size of a quantum/classical automaton is $k + |Q|$.
Acceptance modi are defined as usual.

Obviously a quantum/classical automaton can efficiently simulate both a qfa in the Kondacs/Watrous definition, and a classical probabilistic automaton: random bits can be provided by measuring input independent quantum states. Due to the results of [24], [3], [28] this model can recognize languages which

qfa cannot, and can be exponentially smaller than both qfa and pfa for some languages.
Classical formulae are defined as usual, probabilistic formulae are formulae with additional random inputs. Quantum circuits are defined by Yao in [38]. A quantum formula is a quantum circuit with the additional requirement that for every input there is at most one unique path that connects it to the output wire. The size of the quantum formula is the number of gates without predecessors. Acceptance is defined as usual.
The definition by Yao allows classical constants 0,1 to be read by the formula besides its (Boolean) input. Only one final measurement may be applied to the output qubits of the formula. [1] introduces a more general model of quantum circuits working on mixed states, where density matrices are transformed by superoperator gates.
We call the model defined in [38] *pure* quantum formulae and the mixed state model *generalized* quantum formulae. There is, however, one more point: to simulate random input variables we allow the formulae to read additional input variables, where each of these is a (mixed) ensemble of pure states (not entangled with the states of the other additional input variables). Each of these quantum random variables can be read arbitrarily often by the formula. For each quantum random variable in the beginning of the computation a pure state is chosen from the corresponding ensemble and each occurence of the quantum random variable is replaced by the pure state.

3 Quantum Las Vegas communication

In this section we prove a polynomial gap between quantum Las Vegas and bounded error probabilistic communication complexity for a total Boolean function. This function accepts a subset of the language defined now.
$\mathcal{P}_{m,n}$ is the set of all size n subsets of $\{1,\ldots,m\}$. Let $D_{n,s}$ denote the following language:

$$\begin{aligned} D_{n,s} = \quad \{ \quad & (x_1, x_2, \ldots, x_s; y_1, y_2, \ldots, y_s) | \\ & \forall i : x_i, y_i \in \mathcal{P}_{n^{32},n} \text{ and } x_i \cap y_i \neq \emptyset \} \end{aligned}$$

Note that the input length is $\Theta(ns \log n)$ and that nondeterministic protocols can easily compute $D_{n,s}$ and its complement with communication $O(s \log n)$ resp. $O(n \log n)$. Hromkovič and Schnitger [20] consider the question how efficient a nondeterministic protocol with less than s nondeterministic guess bits can be and prove

Fact 1 *There is an $\epsilon > 0$ such that any deterministic protocol which computes $D_{n,s}$ correctly on a fraction of $1/2^{\epsilon s}$ of the ones without accepting any zero must exchange at least $\Omega(ns)$ bits.*

Any Monte Carlo protocol (and any Las Vegas protocol) for a function induces a deterministic protocol for any distribution, that must compute a function

correctly on 1/2 of its ones without accepting any zero. This follows from the easy part of the minimax principle [27]. So the above fact holds for Monte Carlo protocols as well. We will describe a large enough subset of the ones of the function that can be computed by a quantum Las Vegas protocol. But first we generalize the above result to allow two-sided error.

Lemma 1 *There is an $\epsilon > 0$ such that any probabilistic protocol which decides a subset of size $1/2^{\epsilon s}$ of the ones of $D_{n,s}$ against the zeroes with error δ must exchange at least $\Omega(ns)$ bits.*

Proof Sketch: This proof needs some modifications of the proof in [20]. We describe what has to be done, but leave out any details due to the length of that proof. In the deduction of theorem 2.1 the authors argue that accepting a large fraction of 1-inputs by a deterministic errorless protocol implies the existence of many 1-chromatic submatrices in the communication matrix. Now with two-sided error there are many almost monochromatic submatrices, that is matrices with $1 - \delta$ ones (according to the uniform distribution). Lemma 2.1 can be proved approximately with such matrices. Then the proof of lemma 2.4 works within minor modifications and theorem 2.1 follows also for probabilistic protocols as stated above. □

Now we turn to the quantum protocol. We consider black-box quantum computations as in [15], [7], [8] and derive a communication upper bound via a lemma in [7]. For an input assignment $a : \{0,1\}^{\log l} \to \{0,1\}$ an *a-gate* is an unitary mapping:

$$U_a : |i\rangle|b\rangle \mapsto |i\rangle|a(i) \oplus b\rangle,$$

where $i \in \{0,1\}^{\log l}$ and $b \in \{0,1\}$. A quantum circuit with a blackbox input possesses a set of qubits in an initial state $|0,\ldots,0\rangle$. Now unitary transformations from the basis of gates as well as a-gates are applied to the qubits in a fixed way. At the end the state of the output qubit is measured. The circuit computes a function $f(a)$ in Las Vegas mode (i.e., the circuit may give up with probability ϵ). Given an algorithm for the black-box problem G we can solve a related communication problem by the following result of [7].

Fact 2 *Let $x, y \in K^l$. Let G be a black box problem (with input assignment $a : \{0,1\}^{\log l} \to \{0,1\}$) solvable by an algorithm making at most t calls to an a-gate in any computation. Let L be a function $K \times K \to \{0,1\}$. Then there is a communication protocol where A receives x, B receives y and both have to compute $G(L(x_1, y_1), \ldots, L(x_l, y_l))$. The protocol communicates at most $O(t(\log l + \log K))$ qubits.*

The function L in our setting is the comparison of two numbers from $K = \{1,\ldots,m = n^{32}\}$.

We are interested in evaluating the following black-box problem: an AND with fan-in s of ORs with fan-in r of Boolean variables. [8] gave an errorless quantum black box algorithm for the evaluation of depth 2 AND-OR trees. This algorithm consists of 2 algorithms that accept ones or give up resp. accept zeroes or give

up. Both algorithms return a certificate for their answer, i.e., the first algorithms finds a input variables that are one for every OR, while the second finds an OR where all input variables are 0.
A subroutine used in both algorithms is Grover's quantum search algorithm, where the error probability is reduced to at most ϵ/s and more than one input may be a solution [6]. The cost of the algorithm is $O(\sqrt{n}\log s)$ for finding a one input in a fan-in n OR of inputs if one exists.
The first algorithm works as follows: for all of the s ORs do a Grover search with success probability amplified to $1-\epsilon/s$. This gives us a certificate for all of the ORs within $O(s\sqrt{r}\log s)$ queries with probability $1-\epsilon$ if such a certificate exists. If no certificate is found for some OR then give up.
The second algorithm starts with multi-level Grover search as described in [7] and finds an unsatisfied OR (if one exists) with $O(\sqrt{sr}\log(sr))$ with constant probability. Then on the unsatisfied OR all r inputs are queried. Again the algorithm gives up if no certificate is found.
Running both algorithms in parallel gives us the correct answer after at most $O(s\sqrt{r}\log(sr)+r)$ queries with probability 1/2.
Now by fact 2 we can find a quantum protocol for the following problem: A and B each receive a vector of sets (we assume that the elements of the sets are listed in ascending order). The protocol accepts if for all pairs x_i, y_i of sets r specific comparisons of elements have at least one success, that is we accept a subset of $D_{n,s}$.
Our goal is to compute a large subset of the ones of $D_{n,s}$. A first idea would be to solve the whole problem: we want to compare all pairs in the two sets for all s positions. For this we need a tree with $r = n^2$. A quick calculation shown that our algorithm is not good enough. But intuitively for two random sets fewer comparisons should be enough: the ith largest element of the set should be somewhere near $i \cdot m/n$ (the inputs are drawn according to the uniform distribution). For any i and a random set we have

$$\begin{aligned} &Pr(\text{There are not } i \pm c \cdot n^{2/3} \text{ elements} \\ &\text{in the interval } [1,\ldots,im/n]) \leq O(n)/(c^2 n^{4/3}). \end{aligned}$$

This is true by the Chebychef inequality, since the variance of the random variable is bounded by $O(i) = O(n)$: the distribution is hypergeometric.

$$\begin{aligned} &Pr(\forall i \in [1,\ldots,n] \text{ there are} \\ &i \pm c \cdot n^{2/3} \text{ elements in } [1,\ldots,im/n]) \\ &\geq Pr(\forall i \in [1, n^{2/3}, 2n^{2/3}, \ldots, n] \text{ there are} \\ &i \pm c/2 \cdot n^{2/3} \text{ elements in } [1,\ldots,im/n]) \\ &\geq (1 - \tfrac{O(1)}{c^2 n^{1/3}})^{n^{1/3}} \geq 1-\epsilon \end{aligned}$$

for some large enough constant c. The first inequality holds, because if the condition holds for i and $i + n^{2/3}$, then for all intermediate j with larger constants. The second inequality holds, because the probability that the condition

holds for i is at most the probability that the condition holds for i under the restriction that it holds for all $j < i$. Note that the above inequality also holds if we fix some element and choose only $n-1$.
For a random vector of sets we can assume in the following that with probability $1/2^{\epsilon' s}$ all sets have the property that for all i between the ith element in the set and im/n there are only $O(n^{2/3})$ elements. This holds also for a pair of random vectors chosen under the restriction that they are a one of $D_{n,s}$, since we can fix one intersecting element for all positions and choose the rest randomly.
Assume that two sets x, y intersect in an element a, which is x^i and y^j. If $j > i + 8cn^{2/3}$ then let $\lceil i + (j-i)/2 \rceil = k$. Now if $a \leq km/n$ then we have at most $cn^{2/3}$ elements y^l between a and jm/n and thus at most $cn^{2/3}$ elements y^l between km/m and jm/n although $j - k > 4cn^{2/3}$ and get the contradiction that up to jm/n there are at most $k+2cn^{2/3} \leq j-2cn^{2/3}$ of the y^l. A symmetric contradiction holds for $a \geq km/n$. Thus $|j-i| = O(n^{2/3})$.
So if we compare for all s pairs $t = 1, \ldots, s$ of sets x_t, y_t the ith largest element x_t^i with the elements y_t^k from $i - cn^{2/3}$ up to $i + cn^{2/3}$, then we can decide intersection from non-intersection for a fraction of $1/2^{\epsilon' s}$ of all inputs in $D_{n,s}$ without accepting a nonintersection. This gives us a large fraction of the ones of the function. By fact 2 we have found a quantum Las Vegas protocol for this subset, for which the lower bound of lemma 1 clearly holds. What is the complexity of the quantum protocol? We evaluate a tree of an s fan-in AND and r fan-in ORs. r was chosen to be $n \cdot O(n^{2/3})$. By the previous calculation we need communication $O(s\sqrt{r}\log(sr)\log n + r\log n) = O(sn^{5/6}\log(sr)\log n + n^{5/3}\log n)$. Choosing $s = n^{5/6}$ the input length is $N = \Theta(n^{11/6}\log n)$, so we get the following theorem.

Theorem 1 *There is a total Boolean function f such that $Q_0(f) = O(N^{10/11+\epsilon})$ for all $\epsilon > 0$, whereas $R_\epsilon(f) = \Omega(N/\log N)$.*

Beame and Lawry [4] describe a function for which the nondeterministic and co-nondeterministic complexity is $O(\log n)$, but the bounded error randomized complexity is $\Omega(\log^2 n)$. Lemma 1 implies

Theorem 2 *$N(D_{n,n}), N(\neg D_{n,n}) \leq O(\sqrt{m}\log m)$, but*
$R_\epsilon(D_{n,n}) = \Omega(m/\log m)$ for input length $m = \Theta(n^2 \log n)$.

4 Information Theory

Our results in the next section use entropy and information theory arguments.

Definition 4 *Let Ω be a finite set, $X \subseteq \Omega$, $Pr : X \to [0,1]$ a probability distribution, and $x \in X$ be a random variable distributed with Pr. Subsets of X are events.*
The entropy of X is $H(X) = -\sum_{x \in X} Pr(x)\log Pr(x)$.
The conditional entropy $H(X|Y)$ of X given Y is $E_{y \in Y}[-\sum_{x \in X} Pr(x|y)\log Pr(x|y)]$.
The information between X and Y is $H(X:Y)$ is $H(X) - H(X|Y)$.

Definition 5 *Let $\{p_i, |\phi_i\rangle\}$ be a mixed state of a quantum system, i.e., with all $p_i \geq 0$, and $\sum_i p_i = 1$, for pure states $|\phi_i\rangle$. Then $\rho_i = |\phi_i\rangle\langle\phi_i|$ is the density matrix of the pure state $|\phi_i\rangle$. $\sum_i p_i \rho_i$ is the density matrix of the mixed state. All possible measurements of the mixed state are completely determined by ρ. In quantum mechanics the density matrix plays an analogous role to random variables in classical probability theory.*
The von Neumann entropy of a density matrix ρ_X is defined by $S(X) = S(\rho_X) = -trace(\rho_X \log \rho_X)$.
The conditional von Neumann entropy $S(X|Y)$ for a bipartite system with density matrix ρ_{XY} is defined by $S(XY) - S(Y)$, where the Y subsystem is the result of a partial trace (see [10]). The von Neumann information is $S(X : Y) = S(X) + S(Y) - S(XY)$. The conditional information is $S(X : Y|Z) = S(XZ) + S(YZ) - S(Z) - S(XYZ)$.
Let $\mathcal{E} = (\rho_i, p_i)_i$ be an ensemble of density matrices with their probabilities. The Holevo information of this ensemble is $\chi(\mathcal{E}) = S(\rho) - \sum_i p_i S(\rho_i)$.

If the vectors $|\phi_i\rangle$ span a Hilbert space of dimension d then the von Neumann entropy of the density matrix is bounded by $\log d$. We also need the Holevo-bound [18].

Fact 3 *Let X be a classical random variable with $Pr(X = x) = p_x$. Suppose that for each x a quantum state with density matrix ρ_x is prepared in a quantum register Z, i.e., we have the ensemble $\mathcal{E} = (\rho_x, p_x)_x$. Let Y be the random variable induced by a measurement on the quantum state with density matrix $\rho_z = \sum_x p_x \rho_x$. Then*

$$H(X : Y) \leq \chi(\mathcal{E}) = S(X : Z).$$

We need the following lemma (a Las Vegas version of a lemma in [28]) in section 5.

Lemma 2 *Let X be a classical random variable with $Pr(X = x) = p_x$. Let $\mathcal{E} = \{(p_x, \sigma_x)|x = 0, \ldots, k\}$ be an ensemble of density matrices on a quantum register Z and let $\sigma = \sum_x p_x \sigma_x$ be the density matrix of the mixed state of the ensemble. Assume there is an observable with possible measurement results x and ?, so that for all x measuring the observable on σ_x yields x with probability at least $1 - \epsilon$, the result ? with probability at most ϵ, and a result $x' \neq x$ with probability 0, then*

$$S(\sigma) \geq \sum_x p_x S(\sigma_x) + (1 - \epsilon)H(X), \text{ i.e., } S(X : Z) = \chi(\mathcal{E}) \geq (1 - \epsilon)H(X).$$

Proof: We code classical states from a random source X, which produces x with probability p_x. The code of x has the density matrix σ_x. The density matrix of the coding is σ with von Neumann entropy $S(\sigma)$. By Holevo's theorem the information that can be obtained on X is for any measurement with outcome Y bounded $H(X : Y) \leq S(\sigma) - \sum_x p_x S(\sigma_x)$. But there is a measurement

as described in the lemma with $H(X:Y) = H(X) - H(X|Y)$, where for $\delta = \Pr(Y =?) \leq \epsilon$ and $\epsilon_x = Pr(Y =?|X = x) \leq \epsilon$

$$\begin{aligned} H(X|Y) \quad & \leq (1-\delta)H(X|Y \neq ?) + \delta H(X|Y =?) \\ & = \delta H(X|Y =?) \\ & = -\delta \sum_x Pr(X = x|Y =?) \log(Pr(X = x|Y =?)) \\ & = -\delta \sum_x (\epsilon_x p_x/\delta) \log(\epsilon_x p_x/\delta) \\ & \leq -\epsilon \sum_x p_x \log p_x + \delta \sum_x (\epsilon_x p_x/\delta) \log(\delta/\epsilon_x) \\ & \leq -\epsilon H(X) + \delta \log \sum_x p_x \\ & \leq \epsilon H(X). \end{aligned}$$

Thus the lemma follows. □

5 Quantum one-way communication

[26] proves a lower bound on randomized one-way communication based on the notion of the VC-dimension.

Definition 6 *For a function $f(x,y)$ a set S is shattered, if for all $R \subseteq S$ there is an x such that $f(x,y) = 1 \iff y \in R$ for all $y \in S$. The VC-dimension of a function $f(x,y)$ is the maximal size of a shattered set.*

Fact 4 *For all $f : R^1_\epsilon(f) \geq \Omega(VC(f))$.*

As an example look at the disjointness problem: the set of singleton sets (or incidence vectors with only one 1) is clearly shattered und we have $VC(DISJ) = n$.
It is known that this lower bound is not always tight, though sometimes simple reductions can still prove good lower bounds, see [26].
We want to show that the lower bound holds for one-way quantum protocols with bounded error, too. To do this we give a reduction from the Index function to any function with high VC-dimension. The Index function is defined as follows: $IX_n : \{0,1\}^n \times \{0,1\}^{\log n} \to \{0,1\}$ and $f(x,y) = 1 \iff x_y = 1$. Note that the player possessing x has to send the message in a one-way protocol. It is easy to see that $VC(IX_n) = n$ and so the classical bounded error one-way communication is high.
Recently *random access quantum coding* was analyzed in [3], [28]. In n, m, ϵ-random access quantum coding all n-bit strings have to be coded in m qubits, such that for $i = 1, \ldots, n$ there is a measurement of the qubits that yields the desired bit x_i with probability $1 - \epsilon$. Nayak [28] shows

Fact 5 *For any random access n, m, ϵ-quantum coding $m \geq (1 - H(\epsilon))n$.*

The reader might have observed that the problem of random access quantum coding is equivalent to the problem of finding a quantum one-way protocol for IX. If a protocol exists, then the messages serve as mixed state codes, and if coding is possible, then codes may be used as messages.

Theorem 3 *For all* $f : Q^1_\epsilon(f) \geq (1-H(\epsilon))VC(f)$ *and* $Q^{1,e}_\epsilon(f) \geq (1-H(\epsilon))VC(f)/2$.

PROOF: For the first statement we give a reduction from the Index function. If $VC(f) = d$ then there is a set $S = \{s_1, \ldots, s_d\}$ which is shattered by f. We reduce IX_d to f. For every subset $R \subseteq S$ the length d incidence vector c_R of R is used as a name for R. For each c_R (interpreted as input of A to IX_d) we choose an input x_R that separates this subset from the rest of S. The reduction maps c_R to these corresponding x_R inputs. B's inputs i of IX_d are mapped to the s_i. So $f(x_R, s_i) = 1 \iff s_i \in R \iff c_R(i) = 1$.
In this way a quantum protocol for f solves IX_d. By fact 5 the first lower bound follows.
For the second statement we use the same reduction as above, but we have to reanalyze the complexity of IX_n in the entanglement model. We sketch the proof.
Denote the density matrix of the state of the message and those qubits in B's possession, which are entangled with some of A's qubits, by σ_{ME} (induced by uniformly random inputs X). Now if each bit has to be decodable with probability $1-\epsilon$ then $S(X_i : ME) \geq 1 - H(\epsilon)$ for all i. This holds because the Holevo information of the following ensemble of density matrices equals $S(X_i : ME)$, and is thus a bound on the accessible classical information: one matrix for the codes (and entangled bits) of words with $X_i = 0$, probability $1/2$ and one for the codes (and entangled bits) of words with $X_i = 1$, probability $1/2$. The accessible information has to be at least $1 - H(\epsilon)$.
But then $S(X : ME) \geq (1 - H(\epsilon)n$ (because the X_i are independent). $S(X : ME) = S(X : E) + S(X : M|E) \leq 2S(M)$ using the Araki-Lieb inequality. Note that $S(X : E) = 0$. So the number of qubits sent must be at least $(1-H(\epsilon))n/2$.
□

Corollary 1 $Q^1_\epsilon(DISJ) \geq (1 - H(\epsilon))n$.
$Q^1_\epsilon(GT) = \Omega(n/\log n)$ *for the function* $GT(x, y) = 1 \iff x \geq y$.

The latter follows from a reduction as in [26]. The first result has been observed independently in [9].
Now turn to exact and Las Vegas quantum one-way communication complexity for total functions. Note that for classical one-way protocols and total functions [14] show that Las Vegas communication is at most a constant factor smaller than deterministic communication.

Theorem 4 *Let* $row(f)$ *denote the number of different rows in a communication matrix for* $f(x, y)$. *Then for all total functions* f:
$Q^1_E(f) = \lceil \log row(f) \rceil = D^1(f)$, $Q^1_{0,\epsilon}(f) \geq (1 - \epsilon)D^1(f)$.

PROOF: We show that a Las Vegas protocol with give-up probability $\epsilon \geq 0$ for f with $row(f) = R$ must use messages with von Neumann entropy at least $(1-\epsilon)\log R$ when run on the mixed state of uniformly random inputs to A (identified with rows). Then we conclude that the message space must have dimension at least $R^{1-\epsilon}$ and thus at least $(1-\epsilon)\log R$ qubits have to be sent. This yields both lower bounds (the upper bound is obvious).
We describe a random process of choosing rows bit per bit starting in column 1. Call the probability of a 0 in column 1 p_1. Then choose a 0 with probability p_1 and a 1 with probability $1-p_1$. Afterward the set of rows is partitioned into I_1 and I_0 according to the value of $M(i,1)$. If $x_1 = b$ continue with I_b as described. If a complete row is chosen take the density matrix of the message on that row.
Let ρ_y denote the density matrix of the rows starting with the prefix y. Then the probability that a 0 is chosen after choosing y is called p_y, and the number of the different rows in the submatrix of the rows starting with y is called row_y. Clearly $S(\rho_x) \geq 0$ for any completely chosen x. By lemma 2 we get $S(\rho_y) \geq p_y S(\rho_{y0}) + (1-p_y)S(\rho_{y1}) + (1-\epsilon)H(p_y)$. So via induction

$$\begin{aligned} S(\rho_y) &\geq p_y((1-\epsilon)\log row_{y0}) \\ &+(1-p_y)((1-\epsilon)\log row_{y1}) + (1-\epsilon)H(p_y) \\ &= (1-\epsilon)(p_y \log(p_y row_y) \\ &+(1-p_y)\log((1-p_y)row_y) + H(p_y)) \\ &= (1-\epsilon)\log row_y. \end{aligned}$$

We conclude that $S(\rho) \geq (1-\epsilon)\log row(f)$. □
Now we consider the same problem in the entanglement model. The proof methods of theorem 4 and theorem 3 part 2 yield

Theorem 5 *For all total functions f:*
$Q_E^{1,e}(f) = \lceil D^1(f)/2 \rceil, Q_{0,\epsilon}^{1,e}(f) \geq D^1(f)(1-\epsilon)/2$

6 Quantum one-way automata

The common way to prove lower bounds on classical deterministic finite automata size is by means of the Nerode index. A generalization of one-way communication (see [14]) captures the Nerode index.

Definition 7 *In a one-way communication problem A receives an input $x \in \Sigma^*$, B receives $y \in \Sigma^*$ and they decide the Boolean function $f(xy)$. Deterministic and quantum one-way communication complexity of f as well as the VC-dimension are defined as above.*

The number of different rows of the (infinite) communication matrix (defined by $M(x,y) = f(xy)$ with rows and columns for all $x, y \in \Sigma^*$) equals the Nerode index [14].

Theorem 6 *For all total functions $f : \Sigma^* \to \{0,1\}$ with minimal dfa's of size D the following holds:*
Every exact quantum/classical automaton for f has at least size D.
Every quantum/classical automaton with error ϵ for f has size at least $2^{(1-H(\epsilon))VC(f)}$.
Every Las Vegas quantum/classical automaton with give-up probability ϵ has at least size $D(f)^{1-\epsilon}$.

PROOF: As argued in the proof of theorem 4 the von Neumann entropy of the messages in a protocol with a communication matrix having R rows and give up probability ϵ is at least $(1-\epsilon)\log R$. Since the protocol can simulate an automaton of size q using messages in a q-dimensional Hilbert space we have that $q > R^{1-\epsilon}$. R equals the Nerode index and we get the result for exact and Las Vegas automata.
For the case of bounded error simulate the automaton with size q by a protocol with q-dimensional messages. Since the von Neumann entropy of the message space must be at least $(1-H(\epsilon))VC(f)$ and at most $\log q$ the lower bound follows. □

Corollary 2 *The size of a quantum/classical automaton for IX is $2^{\Omega(n)}$. The size of a quantum/classical for DISJ is $2^{\Omega(n)}$.*

7 Quantum formulae

Let us rephrase the Nečiporuk lower bound [29] on formula size in terms of one-way communication as done in [22].

Definition 8 *Let f be a function and $y_1 \dots y_k$ a disjoint partion of the n inputs. Player B knows the inputs in y_i and A all other inputs.*
The deterministic one-way communication complexity of f under this partition is called $c_i(f)$. $vc_i(f)$ denotes the VC-dimension of the communication problem. We call $\sum_i vc_i(f)$ the VC-Nečiporuk function.

It is easy to see that $\sum_i c_i(f)$ coincides asymptotically with the (standard) Nečiporuk function and is thus an asymptotical lower bound for deterministic formula size [29]. [35] considers pure quantum formulae (which are allowed to read only inputs and Boolean constants, and which may only measure one output qubit). The result is:

Fact 6 *Any pure quantum formula computing f has size $\Omega(\sum_i c_i(f)/\log c_i(f))$.*

We now consider a function, for which probabilistic formulae are more succinct than pure quantum formulae (see [22]). The matrix product function MP receives $n \times n$-matrices T_1, T_2 over $GF(2)$ as inputs and accepts if $T_1 \cdot T_2 \neq 0$.

Fact 7 *The MP function can be computed in size $O(n^2)$ by a probabilistic formula.*
There is a partition of the inputs such that the Nečiporuk function for the partition and the MP function is $\Omega(n^3)$.

This leads to the following

Corollary 3 *There is a function with size $O(N)$ probabilistic formulae that requires size $\Omega(N^{3/2}/\log N)$ for pure quantum formulae.*

Clearly any probabilistic formula can be simulated efficiently by a generalized quantum formula. A different lower bound method is needed.

Theorem 7 *The VC-Nečiporuk function is an asymptotic lower bound on the size of generalized quantum formulae with bounded error.*
The Nečiporuk function is an asymptotic lower bound on the size of generalized quantum Las Vegas formulae.

Proof Sketch: For a given partition of the inputs we show how a quantum formula F can be simulated by the k communication games, such that the quantum one-way communication of game i is asymptotically bounded by the number of leaves in the subtree of F that contains exactly the variables belonging to B.
Measurements can be deferred to the end of the computation of the formula. The argument of [35] cannot be used directly, because of the additional inputs. For the simulation the communication model with entanglement is used. Due to theorems 3 and 5 this does not affect the lower bound by more than a constant factor.
Let an additional input read by the formula be in the state $\alpha_i|0\rangle + \beta_i|1\rangle$ with probability p_i. Using measurements on the shared entangled qubits one player produces the distribution p_i. Then with probability p_i the state $\alpha_i|0\rangle + \beta_i|1\rangle$ is produced by both players as input to the formula.
In all communication games player B evaluates the formula as far as possible without the help of A. By a path squeezing argument similar to that in [35] $O(1)$ qubits communication suffice for evaluating a path in the formula with the following property: all inner gates have one input from A and one input from the predecessor in the path.
By standard arguments the number of occurrences of input variables of B in the formula is lower bounded by the number of such paths and thus by the communication.
For this path squeezing we employ a version of programmable quantum gates [30] with amplified success probability. Informally speaking, player A sends the program of a unitary operation corresponding to the path to player B using an average of $O(1)$ qubits. Player B can perform that operation with high probability in the Las Vegas sense. If B's operation fails he sends a note about this to A (such a note bears no information and can be tolerated in the lower bound argument for quantum one-way communication complexity).
Overall the size as the number of leaves is lower bounded by the sum of the communications in the games, which is lower bounded by the VC Nečiporuk function due to theorem 3.
For Las Vegas formulae we get a Las Vegas one-way protocol, which is bounded by theorem 5. □

It is well known that the Indirect Storage Access function (ISA) has deterministic formula size $\Theta(n^2/\log n)$. The ISA function is defined as follows: there are inputs U, X, Y, with $|U| = \log n - \log\log n$, $|X| = |Y| = n$, $ISA(U, X, Y) = Y_{X_U}$.

Corollary 4 *A generalized quantum formula computing ISA with bounded error has size* $\Omega(n^2/\log n)$.

Considering matrix multiplication again we get

Corollary 5 *There is a function which can be computed in size* $O(N)$ *by a generalized quantum formula with bounded error, but requires size* $\Omega(N^{3/2})$ *for generalized quantum Las Vegas formulae, i.e., there is a gap of* $\Omega(N^{1/2})$ *between Las Vegas and bounded error formula size.*

8 Rounds in quantum communication?

It is well known that for deterministic, probabilistic, (and even limited nondeterministic) communication complexity there are functions which can be computed much more efficiently in k rounds than in $k-1$ rounds (see [13], [17], [31], [32], [23]). In most of these these results the pointer jumping function is considered.

Definition 9 *Let* V_A *and* V_B *be disjoint sets of* n *vertices each.*
Let $F_A = \{f_A | f_A : V_A \to V_B\}$, *and* $F_B = \{f_B | f_B : V_B \to V_A\}$.
Then let $f(v) = f_A(v)$ *(resp.* $f_B(v)$*) if* $v \in V_A$ *(resp.* $v \in V_B$*).*
Define $f^{(0)}(v) = v$ *and* $f^{(k)}(v) = f(f^{(k-1)}(v))$.
Then $g_k : F_A \times F_B \to (V_A \cup V_B)$ *is defined by* $g_k(f_A, f_B) = (f^{(k+1)}(v_1))$, *where* v_1 *is fix.*
The function $f_k : F_A \times F_B \to \{0, 1\}$ *is the XOR of all bits in the binary code of the output of* g_k.

Nisan and Wigderson proved in [31] that f_k has a randomized k round communication complexity of $\Omega(n/k^2 - k\log n)$ if B starts communicating and a deterministic k round communication complexity of $k\log n$ if A starts. They also describe a randomized protocol computing g_k with communication $O((n/k)\log n)$ in the "bad" situation, [32] show that deterministic communication is $\Theta(n)$ then.
The following is however easy to deduce from the proof of [31].

Observation 1 $D^{k,A}(g_k) \leq k\log n$.
$R_\epsilon^{k,B}(f_k) \geq \Omega(n/k + k)$.

Proof Idea: The first statement is trivial.
In the conclusion of their theorem 2 the authors show that deterministic protocols with communication $\delta n - k\log n$ have error at least $1/2^k$ on the uniform distribution for some small constant δ. But then protocols with $o(n/k - \log n)$ communication cannot have error separated from $1/2$ by a constant. The lower bound k is trivial for k round protocols. □
By a simple reduction we get the following result.

Observation 2 *If the $Q_\epsilon^{k,B}(f_k) \geq c(n)$ then any k-round quantum protocol for the disjointness problem needs communication $\Omega(c(n^{1/k}))$.*

The best upper bound known is $O(\sqrt{n}\log n)$ communication in $O(\sqrt{n})$ rounds [7].
The main difficulties in analyzing the quantum complexity of pointer jumping lie in the fact that the information of the messages on the next pointer may increase very fast (if one consider a suitable adaption of the protocol in [32]), while still for every measurement only with small probability large information is obtained.

References

[1] D. Aharonov, A. Kitaev, N. Nisan. Quantum Circuits with Mixed States. *Proc. 30th ACM Symp. on Theory of Comp.*, pp.20–30,1998.

[2] A.Ambainis, R.Freivalds. 1-way quantum finite automata: strengths, weaknesses and generalizations. *Proc. 39th Symp. Found. of Computer Science*, pp. 332–341, 1998.

[3] A.Ambainis, A.Nayak, A.Ta-Shma, U.Vazirani. Dense quantum coding and a lower bound for 1-way quantum finite automata. *Proc. 31th ACM Symp. on Theory of Comp.*, pp.376–383, 1999.

[4] P.Beame, J.Lawry. Randomized versus Communication Nondeterministic Complexity. *Proc. 24th ACM Symp. on Theory of Comp.*, pp.188–199, 1992.

[5] C.H.Bennett, S.J.Wiesner. Communication via one- and two-particle operators on Einstein-Podolsky-Rosen states. *Phys. Rev. Lett.*, vol.69, pp.2881–2884, 1992.

[6] M.Boyer, G.Brassard, P.Hoyer, A.Tapp. Tight bounds on quantum searching. *4th Workshop on Physics and Computation*, pp.36-43, 1996.

[7] H.Buhrman, R.Cleve, A.Wigderson. Quantum vs. classical communication and comp. *Proc. 30th ACM Symp. on Theory of Comp.*, pp.63–68,1998.

[8] H.Buhrman, R.Cleve, R.de Wolf, C.Zalka. Reducing error probability in quantum algorithms. *Proc. 40th Symp. Found. of Computer Science*, 1999.

[9] H.Buhrman, R.de Wolf. Communication Complexity Lower Bounds by Polynomials. quant-ph 991001.

[10] N.Cerf, C.Adami. Quantum information theory of entanglement and measurement. *Proc. of Physis and Computation PhysComp*, pp.65–71, 1996.

[11] R.Cleve, H.Buhrman. Substituting quantum entanglement for communication. *Physical Review A*, vol.56, pp.1201-1204, 1997.

[12] R.Cleve, W.van Dam, M.Nielsen, A.Tapp. Quantum Entanglement and the Communication Complexity of the Inner Product Function. quant-ph 9708019.

[13] P.Duris, Z.Galil, G.Schnitger. Lower Bounds on Communication Complexity. *Information & Computation*, vol.73, pp.1–22, 1987.

[14] P.Duriš, J.Hromkovič, J.D.P.Rolim, G.Schnitger. Las Vegas Versus Determinism for One-way Communication Complexity, Finite Automata, and Polynomial-time Computations. *Symp. on Theoretical Aspects of Computer Science*, pp.117–128, 1997.

[15] L.K.Grover. A fast quantum mechanical algorithm for database search. *28th ACM Symposium on Theory of Computing*, pp.212-219, 1996.

[16] J.Gruska. Quantum Computing. Wiley Interscience. 1999.

[17] B.Halstenberg, R.Reischuk. Different Modes of Communication. *SIAM Journal Comput.*, vol.22, pp.913–934, 1993.

[18] A.S. Holevo. Some estimates on the information transmitted by quantum communication channels. *Problems of Information Transmission*, vol.9, pp.177–183, 1973.

[19] J.Hromkovič. Communication Complexity and Parallel Computing. Springer, 1997.

[20] J.Hromkovič, G.Schnitger. Nondeterministic Communication with a Limited Number of Advice Bits. *Proc. 28th ACM Symp. on Theory of Comp.*, pp.451–560, 1996.

[21] B.Kalyanasundaram, G.Schnitger. The Probabilistic Communication Complexity of Set Intersection. *SIAM Journ.Discr.Math.,* vol.5, pp.545-557, 1992.

[22] H.Klauck. On the Size of Probabilistic Formulae. *8th Int. Symp. on Algorithms and Computation*, pp.243–252, 1997.

[23] H.Klauck. Lower bounds for computation with limited Nondeterminism. *13th IEEE Conference on Computational Complexity*, pp.141–153, 1998.

[24] A.Kondacs, J.Watrous. On the power of quantum finite state automata. *38th Symp. Found. Computer Science*, pp.66–75, 1997.

[25] I. Kremer. Quantum Communication. Master's thesis (Hebrew University), 1995.

[26] I. Kremer, N. Nisan, D. Ron. On Randomized One-Round Communication Complexity. *27th Symp. Theory of Comp.*, pp. 596–605, 1995

[27] E.Kushilevitz, N.Nisan. Communication Complexity. Cambridge Univ. Press, 1996.

[28] A.Nayak. Optimal lower bounds for quantum automata and random access codes. *Proc. 40th Symp. Found. of Computer Science*, 1999.

[29] E.I. Nečiporuk. A Boolean function. *Sov.Math.Dokl.*, vol.7, pp. 999-1000, 1966.

[30] M.A. Nielsen, I. Chuang. *Programmable quantum gate arrays,* Phys. Rev. Letters, pp. 321–324, 1997.

[31] N.Nisan, A.Wigderson. Rounds in communication complexity revisited. *SIAM Journ. Comput.*, vol.22, pp.211-219, 1993.

[32] S.J.Ponzio, J.Radhakrishnan, S.Venkatesh. The communication complexity of pointer chasing: applications of entropy and sampling *Proc. 31th ACM Symp. on Theory of Comp.*, pp. 602-611, 1999.

[33] J. Preskill. Lecture notes on quantum information and quantum computation. Web address: www.theory.caltech.edu/people/preskill/ph229.

[34] R. Raz. Exponential Separation of Quantum and Classical Communication Complexity. *Proc. 31th ACM Symp. on Theory of Comp.*, pp. 358–367, 1999.

[35] V.P. Roychowdhury, F. Vatan. An Almost-Quadratic Lower Bound for Quantum Formula Size. quant-ph 9903042, 1999.

[36] P.W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal Comput.*, vol.26, pp.1484–1509, 1997.

[37] A. Ta-Shma. Classical versus Quantum Communication Complexity. *SIGACT News*, vol.30(3), pp.25-34, 1999.

[38] A.C. Yao. Quantum Circuit Complexity. *34th Symp. Found. of Computer Science*, pp. 352–361, 1993.