Theoretical Computer Science 410 (2009) 1923–1931

Contents lists available at ScienceDirect

Theoretical Computer Science

journal homepage: www.elsevier.com/locate/tcs

Improved constructions of mixed state quantum automata

Rūsiņš Freivalds *, Māris Ozols, Laura Mančinska

Institute of Mathematics and Computer Science, University of Latvia, Raiņa bulv. 29, Rīga, Latvia

ARTICLE INFO

Keywords:
Finite automata
Quantum algorithms
Permutation groups

ABSTRACT

Quantum finite automata with mixed states are proved to be super-exponentially more concise rather than quantum finite automata with pure states. It was proved earlier by A. Ambainis and R. Freivalds that quantum finite automata with pure states can have an exponentially smaller number of states than deterministic finite automata recognizing the same language. There was an unpublished "folk theorem" proving that quantum finite automata with mixed states are no more super-exponentially more concise than deterministic finite automata. It was not known whether the super-exponential advantage of quantum automata is really achievable.

We prove that there is an infinite sequence of distinct integers n, languages L_n, and quantum finite automata with mixed states with $5n$ states recognizing language L_n with probability $\frac{3}{4}$, while any deterministic finite automaton recognizing L_n needs at least $e^{O(n \ln n)}$ states.

Unfortunately, the alphabet for these languages grows with n. In order to prove a similar result for languages in a fixed alphabet we consider a counterpart of Hamming codes for permutations of finite sets, i.e. sets of permutations such that any two distinct permutations in the set have Hamming distance at least d. The difficulty arises from the fact that in the traditional Hamming codes for binary strings, positions in the string are independent while positions in a permutation are not independent. For instance, any two permutations of the same set either coincide or their Hamming distance is at least 2. The main combinatorial problem still remains open.

1. Introduction

A. Ambainis and R. Freivalds proved in [4] that for the recognition of some languages the quantum finite automata can have a smaller number of states than deterministic ones, and this difference can even be exponential. The proof contained a slight non-constructiveness, and the exponent was not shown explicitly. For probabilistic finite automata exponentiality of such a distinction was not yet proved. The best (smaller) gap was proved by Ambainis [2]. The languages recognized by automata in [4] were presented explicitly but the exponent was not. In a very recent paper by Freivalds [8] the non-constructiveness is modified, and an explicit (and seemingly much better) exponent is obtained at the expense of having only a non-constructive description of the languages used. Moreover, the best estimate proved in this paper is proved under assumption of the well-known Artin's Conjecture (1927) in Number Theory [5]. [8] contains also a theorem that does not depend on any open conjectures but the estimate is worse, and the description of the languages used is even less constructive. This seems to be the first result in finite automata depending on open conjectures in Number Theory.

* Corresponding author.
E-mail address: rusinsf@latnet.lv (R. Freivalds).

0304-3975/$ – see front matter © 2009 Published by Elsevier B.V.
doi:10.1016/j.tcs.2009.01.028

The following two theorems are proved in [8]:

Theorem 1. *Assume Artin's Conjecture. There exists an infinite sequence of regular languages $L_1, L_2, L_3, \ldots$ in a 2-letter alphabet and an infinite sequence of positive integers $z(1), z(2), z(3), \ldots$ such that for arbitrary j:*

(1) *there is a probabilistic reversible automaton with $z(j)$ states recognizing L_j with the probability $\frac{19}{36}$,*
(2) *any deterministic finite automaton recognizing L_j has at least $(2^{1/4})^{z(j)} == (1.1892071115\ldots)^{z(j)}$ states.*

Theorem 2. *There exists an infinite sequence of regular languages $L_1, L_2, L_3, \ldots$ in a 2-letter alphabet and an infinite sequence of positive integers $z(1), z(2), z(3), \ldots$ such that for arbitrary j:*

(1) *there is a probabilistic reversible automaton with $z(j)$ states recognizing L_j with the probability $\frac{68}{135}$,*
(2) *any deterministic finite automaton recognizing L_j has at least $(7^{\frac{1}{14}})^{z(j)} == (1.1149116725\ldots)^{z(j)}$ states,*

The two theorems above are formulated in [8] as assertions about reversible probabilistic automata. For probabilistic automata (reversible or not) it was unknown before the paper [8] whether the gap between the size of probabilistic and deterministic automata can be exponential. It is easy to rewrite the proofs in order to prove counterparts of Theorems 1 and 2 for quantum finite automata with pure states. The aim of this paper is to prove a counterpart of these theorems for quantum finite automata with mixed states.

Quantum algorithms with mixed states were first considered by Aharonov, Kitaev, and Nisan [1]. More detailed descriptions of quantum finite automata with mixed states can be found in Ambainis, Beaudry, Golovkins, Ķikusts, Mercer, and Thérien [3].

The automaton is defined by the initial density matrix ρ_0. Every symbol a_i in the input alphabet is associated with a unitary matrix A_i. When the automaton reads the symbol a_i the current density matrix ρ is transformed into $A_i^* \rho A_i$. When the reading of the input word is finished and the end-marker \$ is read, the current density matrix ρ is transformed into $A_{end}^* \rho A_{end}$ and separate measurements of all states are performed. After that the probabilities of all the accepting states are totalled, and the probabilities of all the rejecting states are totalled.

Like quantum finite automata with pure states described by Kondacs and Watrous [10] we allow measurement of the accepting states and rejecting states after every step of the computation.

The main result in our paper is:

Theorem 3. *There is an infinite sequence of distinct integers n such that there are languages L'_n such that there are quantum finite automata with mixed states with $5n$ states recognizing the language L'_n with probability $\frac{3}{4}$ while any deterministic finite automaton recognizing L'_n needs to have at least $e^{O(n \ln n)}$ states.*

Proof. The proof is delayed till Section 5. □

Since the number of the states for deterministic automata and quantum automata with pure states differ no more than exponentially, we have

Theorem 4. *There is an infinite sequence of distinct integers n such that there are languages L_n in a 2-letter alphabet such that there are quantum finite automata with mixed states with $2n$ states recognizing the language L_n with probability $\frac{3}{4}$ while any quantum finite automaton with pure states recognizing L_n with bounded error needs to have at least $e^{O(n \ln n)}$ states.*

Unfortunately, the alphabet of the languages of Theorem 3 grows unbounded with n. It is only natural to try to prove a counterpart of Theorem 3 in 2- or 3-letter alphabet. We have developed a methodology for such a proof based on combining ideas of Theorem 3 and the results of [8]. However we need a notion similar to Hamming codes for permutations. Since the Hamming distance between permutations is already considered in several well-known textbooks (e.g. [6]) it seemed natural that the corresponding theory might be already published. Very far from truth!

2. Permutations

The permutation of the set N_n is a 1-1 correspondence from N_n onto itself. Let f be such a permutation. The fact that it is *onto* means that for any $k \in N_n$ there exists $i \in N_n$ such that $f(i) = k$.

If we think of a permutation that "changes" the position of the first element to the first element, the second to the second, and so on, we really have not changed the positions of the elements at all. Because of its action, we describe it as the identity permutation because it acts as an identity function.

There are two main notations for such permutations. In relation notation, one can just arrange the "natural" ordering of the elements being permuted on a row, and the new ordering on another row:

$$\begin{Bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{Bmatrix}$$

stands for the permutation s of the set $\{1, 2, 3, 4, 5\}$ defined by $s(1) = 2, s(2) = 5, s(3) = 4, s(4) = 3, s(5) = 1$.

Rather often in the literature permutations are described by the string $s(1) = 2, s(2) = 5, s(3) = 4, s(4) = 3, s(5) = 1$ only. Alternatively, we can write the permutation in terms of how the elements change when the permutation is successively applied. This is referred to as the permutation's decomposition in a product of disjoint cycles. It works as follows: starting from one element x, we write the sequence $(xs(x)s^2(x)\cdots)$ until we get back the starting element (at which point we close the parenthesis without writing it for a second time). This is called the cycle associated to x's orbit following s. Then we take an element we did not write yet and do the same thing, until we have considered all elements. In the above example, we get: $s = (125)(34)$.

Every fixed point is a cycle with length 1.

If we have a finite set E of n elements, it is by definition in bijection with the set $1, \ldots, n$, where this bijection f corresponds just to numbering the elements. Once they are numbered, we can identify the permutations of the set E with permutations of the set $\{1, \ldots, n\}$.

If one has some permutation, called P, one may describe a permutation, written P^{-1}, which undoes the action of applying P. In essence, performing P then P^{-1} is equivalent to performing the identity permutation. One always has such a permutation since a permutation is a bijective map. Such a permutation is called the inverse permutation.

One can define the product of two permutations. If we have two permutations, P and Q, the action of performing P and Q will be the same as performing some other permutation, $R = P \circ Q$, itself. The product of P and Q is defined to be the permutation R. An even permutation is a permutation which can be expressed as the product of an even number of transpositions, and the identity permutation is an even permutation as it equals $(12) \circ (12)$. An odd permutation is a permutation which can be expressed as the product of an odd number of transpositions. It can be shown that every permutation is either odd or even and cannot be both.

The set of all permutations of the set $\{1, \ldots, n\}$ with algebraic operation "product of permutations" can be considered as a group G_n. This group has two generating elements, the permutations $(123\cdots n)$ and $(12)(3)(4)\cdots(n)$.

We can also represent a permutation in matrix form – the resulting matrix is known as a permutation matrix.

A permutation matrix is a matrix obtained by permuting the rows of an $n \times n$ identity matrix according to some permutation of the numbers 1 to n. Every row and column therefore contains precisely a single 1 with 0s everywhere else, and every permutation corresponds to a unique permutation matrix. There are therefore $n!$ permutation matrices of size n, where $n!$ is a factorial.

3. Hamming distance

The Hamming distance between two objects is the number of changes one needs to perform to obtain an object from another. The Hamming distance between two binary words (of the same length) is defined as the number of positions at which they differ. For instance, we consider a set of three binary words {0011, 0110, 1100}. The first word is at Hamming distance 3 from the other two. Additionally, every word in this set is at Hamming distance at least 2 from any other. Such systems of words are called codes. They are important because they allow us to eliminate accidental errors when transmitting the words through noisy information channels.

We consider the Hamming distances between permutations. The Hamming distance between the permutation s of the set $\{1, 2, 3, \ldots, n\}$ and the permutation r of the same set is the number of distinct numbers i such that $s(i) \neq r(i)$. For instance, let s be a permutation of the set $\{1, 2, 3, \ldots, n\}$ and the number of it's fixed points be p. Then the Hamming distance between the permutation s and the identity permutation is the number $n - p$.

It would be interesting to develop a theory of Hamming codes for permutations, i.e. sets of permutations such that any two distinct permutations in the set have a Hamming distance at least d. Unfortunately, we were not able to find in the literature a solution to this problem. The difficulty is in the fact that in the traditional Hamming codes for binary strings positions in the string are independent while positions in a permutation are not independent. For instance, any two permutations of the same set either coincide or their Hamming distance is at least 2.

For an arbitrary n, there is a set P_3 of n-permutations such that the Hamming distance between any two distinct permutations in this set is at least 3: Take the set of all even permutations as P_3. It is easy to see that there is no bigger set of n-permutations with this property.

For an arbitrary n, there is a set P_n of n-permutations such that the Hamming distance between any two distinct permutations in this set is at least n: Take the set of all cyclic permutations of type $s = x + d \pmod{n}$ where $d \in \{0, 1, 2, \ldots, n-1\}$ as P_n.

It is easy to see that there is no bigger set of n-permutations with this property. It is easy to see that both P_3 and P_n are groups with the operation "product of permutations". It is more difficult to construct maximum cardinality sets P_d for d between 3 and n. We have only partial results for this. However, our main goal is the complexity of quantum finite automata, not permutations. The subsequent sections contain results on the Hamming distance between permutations sufficient for our goal.

Lemma 5. *Let d be an arbitrary real number such that $0 \leq d \leq 1$. No more than $2^{dn \ln n}$ permutations can be of a Hamming distance less or equal than dn from the identity permutation.*

Proof. By the Stirling formula, $n! = e^{n \ln n - o(n \ln n)}$. Let π be an arbitrary n-permutation. How distinct n-permutations are there that differ from the permutation π in no more than dn positions? The differing positions can be chosen in

$$\leq \binom{n}{d} < 2^n$$

ways and these $\leq dn$ positions are permuted. Hence there are no more than $2^n \cdot 2^{dn \ln n - o(n \ln n)} \leq 2^{dn \ln n}$ permutations of this type. □

Theorem 6. *For arbitrary constant $c < 1$ such that for arbitrary n there is a set G_n of n-permutations containing $e^{\Omega(n \log n)}$ permutations such that the pairwise Hamming distance of permutations is at least $c \cdot n$.*

Proof. Immediately from Lemma 1. □

4. Permutations and automata

Definition 7. The *Hamming distance* or simply *distance* $d(r, s)$ between two n-permutations r and s on the set S is the number of elements $x \in S$ such that $r(x) \neq s(x)$. The *similarity* $e(r, s)$ is the number of $x \in S$ such that $r(x) = s(x)$. Note that $d(r, s) + e(r, s) = |S| = n$.

Theorem 8. *Let c be a fixed constant. Assume that there is an infinite sequence of distinct integers n such that for each n there exists a group G_n of permutations of the set $\{1, 2, \ldots, n\}$ with k generating elements. Assume further that the pairwise Hamming distance of permutations is at least $c \cdot n$. Then there is an infinite sequence of distinct integers n so that for each n there is a language L_n in a k-letter alphabet that can be recognized with probability $\frac{c}{2}$ by a quantum finite automata with mixed states having $2n$ states, while any deterministic finite automaton recognizing L_n must have at least $|G_n|$ states.*

Proof. For each permutation group G_n we define the language L_n as follows:

> *The letters of L_n are the k generators of the group G_n and it consists of words $s_1 s_2 s_3 \ldots s_m$ such that the product $s_1 \circ s_2 \circ s_3 \circ \cdots \circ s_m$ differs from the identity permutation.*

(A) Any deterministic automaton recognizing L_n is to remember the first input letter by a specific state.

(B) We will construct a quantum automaton with mixed states. It has $4n$ states and the initial density matrix ρ_0 is a diagonal block-matrix that consists of n blocks $\widetilde{\rho}_0$:

$$\widetilde{\rho}_0 = \frac{1}{2n} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

For each of the k generators $g_i \in G_n$ we will construct the corresponding unitary matrix U_i as follows – it is a $2n \times 2n$ permutation matrix, that permutes the elements in the even positions according to permutation g_i, but leaves the odd positions unpermuted.

For example, $g = 3241$ can be expressed as the following permutation matrix that acts on a column vector:

$$g = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

The initial density matrix ρ_0 for $n = 4$ and the unitary matrix U that corresponds to the permutation matrix (4) of permutation g are as follows:

$$\rho_0 = \frac{1}{8} \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix}$$

$$U = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The unitary matrix $U_\$$ for the end-marker is also a diagonal block-matrix. It consists of n blocks that are the *Hadamard matrices*

$$\widetilde{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Notice how the Hadamard matrix $\widetilde{H}$ acts on two specific 2×2 density matrices:

$$\text{if } \rho = \frac{1}{2n}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad \text{then } \widetilde{H}\rho\widetilde{H}^\dagger = \frac{1}{2n}\begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix},$$
$$\text{if } \rho = \frac{1}{2n}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \text{then } \widetilde{H}\rho\widetilde{H}^\dagger = \frac{1}{2n}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

For example, when the letter g is read, the unitary matrix U is applied to the density matrix ρ_0 (both are given in Eq. (1)) and the density matrix $\rho_1 = U\rho_0 U^\dagger$ is obtained. When the endmarker "\$" is read, the density matrix becomes $\rho_\$ = U_\$\rho_1 U_\$^\dagger$. Matrices ρ_1 and $\rho_\$$ are as follows:

$$\rho_1 = \frac{1}{8}\begin{pmatrix} 1&0&0&0&0&0&0&1 \\ 0&1&0&0&1&0&0&0 \\ 0&0&1&1&0&0&0&0 \\ 0&0&1&1&0&0&0&0 \\ 0&1&0&0&1&0&0&0 \\ 0&0&0&0&0&1&1&0 \\ 0&0&0&0&0&1&1&0 \\ 1&0&0&0&0&0&0&1 \end{pmatrix}, \qquad \rho_\$ = \frac{1}{8}\begin{pmatrix} 1&0&0&0&\frac{1}{2}&\frac{1}{2}&\frac{1}{2}&-\frac{1}{2} \\ 0&1&0&0&-\frac{1}{2}&-\frac{1}{2}&\frac{1}{2}&-\frac{1}{2} \\ 0&0&2&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0 \\ \frac{1}{2}&-\frac{1}{2}&0&0&1&0&\frac{1}{2}&\frac{1}{2} \\ \frac{1}{2}&-\frac{1}{2}&0&0&0&1&-\frac{1}{2}&-\frac{1}{2} \\ \frac{1}{2}&\frac{1}{2}&0&0&\frac{1}{2}&-\frac{1}{2}&1&0 \\ -\frac{1}{2}&-\frac{1}{2}&0&0&\frac{1}{2}&-\frac{1}{2}&0&1 \end{pmatrix}.$$

Finally, we declare the states in the even positions to be accepted, but the states in the odd positions to be rejected. Therefore one must sum up the diagonal entries that are in the even positions of the final density matrix to find the probability that a given word is accepted.

In our example the final density matrix $\rho_\$$ is given in (4). It corresponds to the input word "g\$", which is accepted with probability $\frac{1}{8}(1+0+1+1) = \frac{3}{8}$ and rejected with probability $\frac{1}{8}(1+2+1+1) = \frac{5}{8}$. Note that the accepting and rejecting probabilities sum up to 1.

It is easy to see, that the words that do not belong to the language L_n are rejected with certainty, because the matrix $U_\$\rho_0 U_\$^\dagger$ has all zeros in the even positions on the main diagonal. However, the words that belong to L_n are accepted with the probability at least $\frac{d}{2n} = \frac{cn}{2n} = \frac{c}{2}$, because all permutations are at least at the distance d from the identity permutation.

It is also easy to see that any deterministic automaton that recognizes the language L_n must have at least $|G_n|$ states. If the number of states is less than $|G_n|$, then there are two distinct words u and v such that the deterministic automaton ends up in the same state no matter which one of the two words it reads. Since G_n is a group, for each word we can find an inverse that returns the automaton in the initial state (the only rejecting state). Since u and v are different, they have different inverses and $u \circ u^{-1}$ is the identity permutation and must be rejected, but $v \circ u^{-1}$ is not the identity permutation and must be accepted — a contradiction. □

5. Super-exponential size advantage

Now we wish to prove Theorem 3.

Consider the following infinite sequence of languages. For every n take the set G_n considered in Theorem 5. The language L'_n consists of all the words aa (of the length 2) where a is a symbol for an arbitrary element from G_n. Hence there are $e^{\Omega(n \log n)}$ letters in the alphabet of the language L'_n and equally many words in L'_n.

Theorem 3. *There is an infinite sequence of distinct integers n such that there are languages L'_n such that there are quantum finite automata with mixed states with $5n$ states recognizing the language L'_n with probability $\frac{3}{4}$ while any deterministic finite automaton recognizing L'_n needs to have at least $e^{O(n \ln n)}$ states.*

Proof. The proof is similar to the proof of Theorem 6. When constructing the quantum automaton for Theorem 6 we used two distinct sets of states $\{q_1, q_3, \ldots, q_{2n-1}\}$ and $\{q_2, q_4, \ldots, q_{2n}\}$. The unitary transformations corresponding to all generators in the group G_n permuted some states in the first set and it left all the states in the second set unpermuted.

Now we have five such sets of n-tuples of states. Let g_i be any one of the permutations in G_n. (Remember that the cardinality of G_n equals $e^{\Omega(n \log n)}$.) Let the inverse permutation of g_i be denoted as h_i. The unitary transformation corresponding to g_i leaves the states from the first set unpermuted. It takes the states from the second set into the third one performing the permutation g_i. For instance, if $n = 5$ and g_i is (35421) then $f(6) = 13, f(7) = 15, f(8) = 14, f(9) = 12, f(10) = 11$. The unitary transformation takes the third set into the fourth one performing the permutation h_i. The unitary transformation takes the fourth set into the fifth one performing no permutation. The unitary transformation takes the fifth set into the first one performing no permutation. After this transformation a measurement is performed which measures all the states in the fifth set considering them as rejecting states. No measurement of accepting states is performed. The unitary transformation corresponding to the end-marker uses n instances of the *Hadamard matrices*

$$\widetilde{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

This transformation has the following property. If the state q_{3n+j} has been the result of permutation from the state q_{n+j}, then the amplitude of the state q_{4n+j} becomes double the amplitude of the state q_{3n+j} on the preceding step. If the state q_{3n+j} has been the result of permutation from the state q_{n+j}, then the amplitude of the state q_{4n+j} becomes equal to the amplitude of the state q_{3n+j} on the preceding step. After this transformation a measurement is performed which measures all the states in the first set considering them as accepting states. All the other states are measured as rejecting states.

It is easy to see that if the input word is *aa*, then all the amplitudes of the states q_{3n+j} are doubled at the moment when the end-marker is read. The word is accepted with the probability 1. If the input word is shorter than two letters, then the probability to accept this word equals zero. If the input word is longer than two letters, then the probability of acceptance cannot exceed $\frac{1}{2}$. If the length of the input word is two letters but the letters are not equal, then by the property of G_n described in Theorem 5, the probability of acceptance is less than $\frac{1}{2}$. □

6. Smaller alphabets

We were interested in permutation groups such that distinct permutations have large Hamming distance.

Definition 9. A group G of permutations on the set S is called *k-transitive* if for every two k-tuples $(x_1, x_2, \ldots, x_k)$ and $(y_1, y_2, \ldots, y_k)$ of distinct elements of S, there is a permutation $p \in G$ such that $p(x_i) = y_i$ for all $i \in \{1, 2, \ldots, k\}$. If there is exactly one such permutation p, then G is called *sharply k-transitive*. Note that a sharply k-transitive group is also sharply $(k-1)$-transitive.

Lemma 10. *If G is a sharply k-transitive set of n-permutations, then for any distinct $r, s \in G$:*

$$d(r, s) \geq n - k + 1. \tag{1}$$

Proof. Let us assume that $d(r, s) < n - k + 1$ for some $r, s \in G$. It means, that either the similarity $e(r, s)$ is no less than k or the two permutations act on some t-tuple ($t \geq k$) in the same way. This is a contradiction, since G is sharply k-transitive. □

Definition 11. $G(n, d)$ denotes the size of the largest group of n-permutations with the pairwise distance at least d ($d \leq n$).

Cameron [7] cites the following.

Lemma 12. *The following upper bound holds:*

$$G(n, d) \leq \underbrace{n(n-1)(n-2)\cdots(d+1)d}_{n-d+1 \text{ multipliers}} \tag{2}$$

with equality if and only if there is a sharply $(n - d + 1)$-transitive group of permutations.

It is clear that the *symmetric group* S_n of all permutations on the set $\{1, 2, \ldots, n\}$ is sharply n-transitive, because there is exactly one permutation that sends any n-tuple to any other n-tuple. However, S_n is also sharply $(n-1)$-transitive, because if the action of the permutation on $n - 1$ elements is known, the action on the last element is uniquely determined. From Lemma 10 we obtain that the distance between distinct permutations of S_n is at least 2. It is clear, because distance 1 is not possible for permutations.

The group S_n can be generated by two generators (in cycle notation):

$$g_1 = (12)(3)(4)\ldots(n), \tag{3}$$

$$g_2 = (123\ldots n). \tag{4}$$

The first one corresponds to a transposition of first two elements, but the second one — to a cyclic shift of all elements. The group S_n consists of $n!$ permutations.

The *signature* or *sign* of a permutation s is defined as the parity of the number of *inversions* in s, i.e., pairs i, j such that $i < j$, but $s(i) > s(j)$. For example, $s = 3241$ has 4 inversions, namely 32, 31, 21, and 41 thus it is an even permutation. It is easy to show that a *transposition* (a permutation that swaps to elements) changes the sign of a permutation to the opposite. In fact the signature "sgn" of a permutation is a group homomorphism from S_n to $\{-1, 1\}$, because for any two permutations r and s we have $\operatorname{sgn}(s \circ r) = \operatorname{sgn} s \cdot \operatorname{sgn} r$. Therefore it is not hard to see that the set of all even permutations of the set $\{1, 2, \ldots, n\}$ forms a group — the *alternating group* A_n.

It is easy to show that A_n is sharply $(n-2)$-transitive – if we know the action of a permutation on $n-2$ elements, the remaining two elements can be either swapped or remain in the same order. One of these cases corresponds to an even permutation, but the other – to odd. From Lemma 10 it follows that the pairwise distance of distinct permutations of A_n is at least 3. This can also be obtained directly – even permutations can not have a distance of 2, because then they differ only by one transposition and therefore have different signs. Since the number of odd and even permutations is the same, A_n consists of $n!/2$ permutations.

An example of a sharply 1-transitive group is the cyclic group C_n that consists of n permutations and is generated by a cyclic shift

$$g_1 = (123\ldots n). \tag{5}$$

C_n is clearly sharply 1-transitive, because there is exactly one way how to shift any element to any other. The pairwise distance between distinct elements of C_n is exactly n.

An infinite sequence of sharply 2-transitive groups can be constructed using an *affine transformation* $y(x) = ax+b$, where $a, b \in \mathbb{F}_n$, $a \neq 0$. Here $\mathbb{F}_n$ denotes the *finite field* of order n, i.e., a set of n elements together with two binary operations – addition and multiplication, such that $(\mathbb{F}_n, +)$ and $(\mathbb{F}_n\backslash\{0\}, *)$ are Abelian groups and both distribute laws hold. Such a field $\mathbb{F}_n$ exists if and only if n is a power of a prime number.

The function $y(x)$ acts on the elements of the field $\mathbb{F}_n$ as a permutation, because $ax_1 + b = ax_2 + b$ implies $x_1 = x_2$. There are in total $n(n-1)$ such permutations and they form a group: if $y_1(x) = a_1x + b_1$ and $y_2(x) = a_2x + b_2$, then $(y_1 \circ y_2)(x) = y_1(y_2(x)) = (a_1a_2)x + (a_1b_2 + b_1)$ which is also an affine transformation. To prove that the group is sharply 2-transitive, we have to show that there is a unique solution to the following system of two linear equations:

$$\begin{cases} y_1 = ax_1 + b \\ y_2 = ax_2 + b \end{cases} \tag{6}$$

where $x_1 \neq x_2$ and $y_1 \neq y_2$. This solution is

$$a = \frac{y_1 - y_2}{x_1 - x_2}, \quad b = \frac{x_1y_2 - x_2y_1}{x_1 - x_2}. \tag{7}$$

Thus the group is sharply 2-transitive. In a similar way one can explicitly show that the pairwise Hamming distance between distinct permutations is at least $n-1$, but it follows from Lemma 10 as well.

Sharply 3-transitive groups can be described using a different formalism — the *linear fractional transformation* (also called *Möbius transformation*):

$$y(x) = \frac{ax + b}{cx + d}, \tag{8}$$

where $a, b, c, d \in \mathbb{F}_q$ and $ad - bc \neq 0$ (otherwise $a/c = b/d = \alpha$ and $y(x) = \alpha$). It acts on the set $\mathbb{F}_q \cup \{\infty\}$ as a permutation. The following conventions regarding the element ∞ are used:

$$y\left(-\frac{d}{c}\right) = \infty, \qquad y(\infty) = \begin{cases} \infty & \text{if } c = 0, \\ \frac{a}{c} & \text{otherwise.} \end{cases} \tag{9}$$

In fact, the element ∞ corresponds to the vector $\binom{1}{0}$ in the above construction. Note that the inverse of (8) is also a linear fractional transformation:

$$x(y) = \frac{-dy + b}{cy - a}. \tag{10}$$

The same stands for the composition of two linear fractional transformations.

It is known that the *Mathieu group* M_{11} is sharply 4-transitive and therefore the pairwise Hamming distance of distinct elements is at least 8. It consists of $11 \cdot 10 \cdot 9 \cdot 8 = 7920$ elements. It is generated by

$$g_1 = (2, 10)(4, 11)(5, 7)(8, 9)(1)(3)(6), \tag{11}$$

$$g_2 = (1, 4, 3, 8)(2, 5, 6, 9)(7)(10)(11). \tag{12}$$

The *Mathieu group* M_{12} is sharply 5-transitive and the pairwise Hamming distance of distinct elements is also at least 8. It has $12 \cdot 11 \cdot 10 \cdot 9 \cdot 8 = 95040$ elements and is generated by:

$$g_1 = (1, 2)(3, 4)(5, 6)(7, 8)(9, 10)(11, 12), \tag{13}$$

$$g_2 = (1, 2, 3)(4, 5, 7)(8, 9, 11)(6)(10)(12). \tag{14}$$

However, no other sharply transitive groups exist:

Theorem 13 (*See [6]*). *A sharply k-transitive group ($k \geq 4$) is isomorphic either to S_n ($n \geq 4$), A_n ($n \geq 6$) or one of the Mathieu groups M_{11} or M_{12}.*

Table 1
Experimentally obtained results for $G(n, d)$. The columns have the following meaning: n — the size of the set S, d — the pairwise Hamming distance, $G(n, d)$ — the size of the group obtained, "Bound" — the upper bound for $G(n, d)$, "Generators" — the two generators of the group.

n	d	$G(n, d)$	Bound	Generators
7	4	168	840	6, 4, 3, 2, 5, 1, 7 6, 1, 7, 5, 2, 3, 4
8	5	336	1 680	3, 8, 6, 2, 4, 5, 1, 7 7, 4, 6, 3, 1, 5, 2, 8
8	4	1 344	6 720	2, 6, 8, 4, 5, 7, 1, 3 7, 4, 3, 5, 1, 8, 6, 2
9	6	1 512	3 024	4, 5, 1, 8, 3, 7, 6, 2, 9 3, 4, 8, 5, 7, 1, 6, 9, 2
9	5	1 512	15 120	9, 4, 1, 6, 5, 2, 7, 8, 3 1, 4, 5, 3, 7, 9, 8, 2, 6
9	4	1 512	60 480	7, 2, 8, 3, 5, 6, 9, 4, 1 6, 1, 3, 8, 2, 4, 9, 5, 7
10	7	720	5 040	3, 9, 5, 7, 4, 8, 10, 6, 1, 2 7, 9, 4, 5, 3, 6, 8, 1, 10, 2
10	6	1 512	30 240	8, 2, 10, 7, 4, 3, 1, 6, 5, 9 1, 2, 8, 5, 10, 6, 3, 7, 9, 4
10	5	1 512	151 200	1, 10, 3, 9, 6, 8, 5, 4, 7, 2 1, 10, 8, 3, 2, 4, 5, 7, 6, 9
10	4	1 920	604 800	5, 1, 4, 8, 9, 7, 6, 10, 2, 3 7, 8, 2, 1, 10, 3, 9, 6, 4, 5
15	12	2 520	32 760	7, 2, 4, 5, 11, 10, 13, 15, 3, 9, 6, 8, 14, 12, 1 9, 15, 11, 6, 4, 2, 10, 13, 7, 12, 8, 1, 14, 3, 5
16	12	40 320	524 160	16, 5, 6, 12, 14, 13, 11, 1, 10, 3, 7, 4, 15, 8, 9, 2 6, 7, 14, 8, 15, 3, 12, 2, 9, 10, 13, 11, 4, 16, 1, 5

Of course, the property of a group to be sharply transitive would be desirable but not necessary for our needs. We performed computer experiments to find permutation groups with pairwise Hamming distance in the region between $d \geq 4$ and $d \leq n - 3$. The obtained results for $n = 7, 8, 9, 10$ are shown in Table 1. In addition we mention also two large groups for $n = 15$ and $n = 16$.

We performed computer experiments to find permutation groups with pairwise Hamming distance in the region between $d \geq 4$ and $d \leq n - 3$. The obtained results for $n = 7, 8, 9, 10$ are shown in Table 1. In addition we mention also two large groups for $n = 15$ and $n = 16$. Some of the groups obtained in this way have very interesting properties:

(1) $G(7, 4)$ has $168 = 7 \cdot 6 \cdot 4$ elements and is isomorphic to the automorphism group of the *Fano plane*.
(2) $G(8, 4)$ has $1344 = 8 \cdot 168 = 8 \cdot 7 \cdot 6 \cdot 4$ elements. This group has the property, that the stabilizers of any element form a group that is isomorphic to the automorphism group of the Fano plane. This group also has a property that for any 3-tuples x and y of distinct elements there are exactly 4 permutations that send x to y. It is isomorphic to the automorphism group of the *octonion* multiplication table.
(3) $G(8, 4)$ has $1512 = 9 \cdot 168 = 9 \cdot 8 \cdot 7 \cdot 3$ elements and it has the same stabilizer property, but for each 3-tuples x and y there are exactly 3 permutations that send x to y.
(4) $G(15, 12)$ has $2520 = 15 \cdot 168 = 15 \cdot 14 \cdot 12$ elements and it also has the stabilizer property, but for each 2-tuples x and y there are exactly 12 permutations that send x to y.
(5) $G(16, 12)$ has $40\,320 = 16 \cdot 15 \cdot 168 = 16 \cdot 15 \cdot 14 \cdot 12$ elements. The stabilizers of any two elements form a group that is isomorphic to the automorphism group of the Fano plane. For any 3-tuples x and y there are exactly 12 permutations that send x to y.

Unfortunately, when this paper was already submitted, an anonymous referee pointed out the following

Theorem 14 ([9]). (1) *If the Hamming distance m between any two n-permutations in the group G_n is greater than* $\log_2 n$, *then* $|G_n| \leq 2^{10n}$.
(2) *If the Hamming distance m between any two n-permutations in the group G_n is smaller than* $\log_2 n$, *then* $|G_n| \leq n^{10n/m}$.

Hence our methodology cannot prove the existence of languages in a fixed alphabet for which quantum finite automata with mixed states have a super-exponential size advantage over deterministic finite automata. On the other hand, Theorem 14 provides a hope that a super-exponential size advantage over deterministic finite automata can be proved for a sequence of languages L_n where the language L_n is in an alphabet of no more than $\log_2 n$ letters. This would be a serious improvement of our Theorem 3 where the language is in an alphabet consisting of $e^{O(n \ln n)}$ letters.

Acknowledgments

We would like to acknowledge Simone Severini for drawing our attention to the Mathieu groups, Andris Ambainis for pointing at related results in [6] and the anonymous referee who informed us about [9], therefore closing one direction of our attempts to prove the version of Theorem 3 for languages with a constant number of letters.

This research is supported by Grant No. 05.1528 from the Latvian Council of Science.

References

[1] Dorit Aharonov, Alexei Kitaev, Noam Nisan, Quantum circuits with mixed states, in: Proc. STOC 1998, 1998, pp. 20–30.
[2] Andris Ambainis, The complexity of probabilistic versus deterministic finite automata, in: Lecture Notes in Computer Science, vol. 1178, Springer, 1996, pp. 233–237.
[3] Andris Ambainis, Martin Beaudry, Marats Golovkins, Arnolds Ķikusts, Mark Mercer, Denis Thérien, Algebraic results on quantum automata, Theory Comput. Syst. 39 (1) (2006) 165–188.
[4] Andris Ambainis, Rūsiņš Freivalds, 1-way quantum finite automata: Strengths, weaknesses and generalizations, in: Proc. IEEE FOCS'98, 1998, pp. 332–341.
[5] Emil Artin, Beweis des allgemeinen Reziprozitätsgesetzes, Math. Sem. Univ. Hamburg B.5 (1927) 353–363.
[6] Peter Cameron, Permutation Groups, in: London Mathematical Society Student Texts Series, Cambridge University Press, 1999.
[7] Peter Cameron, Permutation codes, in: Talk at the International Combinatorics, Geometry and Computer Science Conference, Luminy, 2007.
[8] Rūsiņš Freivalds, Non-constructive methods for finite probabilistic automata, in: Lecture Notes in Computer Science, vol. 4588, Springer, 2007, pp. 169–180.
[9] Julia Kempe, Laszlo Pyber, Aner Shalev, Permutation groups, minimal degrees and quantum computing, Groups Geom. Dynam. 1 (4) (2007) 553–584 (See also arXiv:quant-ph/0607204v1).
[10] Attila Kondacs, John Watrous, On the power of quantum finite state automata, in: Proc. IEEE FOCS'97, 1997, pp. 66–75.