Lower Bounds for Generalized Quantum Finite Automata

Mark Mercer

Département d'Informatique,
Université de Sherbrooke, QC, Canada

Abstract. We obtain several lower bounds on the language recognition power of Nayak's *generalized quantum finite automata (GQFA)* [12]. Techniques for proving lower bounds on Kondacs and Watrous' *one-way quantum finite automata (KWQFA)* were introduced by Ambainis and Freivalds [2], and were expanded in a series of papers. We show that many of these techniques can be adapted to prove lower bounds for GQFAs. Our results imply that the class of languages recognized by GQFAs is not closed under union. Furthermore, we show that there are languages which can be recognized by GQFAs with probability $p > 1/2$, but not with $p > 2/3$.

Quantum finite automata (QFA) are online, space-bounded models of quantum computation. Similar to *randomized finite automata* [16] where the state is a random variable over a finite set, the state of a QFA is a quantum superposition of finite dimension. The machine processes strings $w \in \Sigma^*$ by applying a sequence of state transformations specified by the sequence of letters in w, and the output of the machine is determined by a measurement of the machine state. A central problem is to characterize the language recognition power of QFAs.

Most quantized versions of classical computation devices (such as quantum circuits [17]) are at least as powerful as their classical counterparts. It is not clear that this should be the case for quantum finite automata. Typically, the execution of classical computation on a quantum device is performed by converting classical computation into reversible computation using standard techniques such as in [5]. The most general definitions of QFAs [8] are equal in language recognition power to deterministic finite automata. However, such definitions require a nonconstant sized (but not directly accessible) memory for bookkeeping. This in some sense violates the spirit of the definition of a finite machine.

For this reason, most QFA research has been focused on the case where the transformations are limited to various combinations of unitary transformations and projective measurements on a finite dimensional state. In this case, the class of languages recognized by these QFAs is a strict subset of the regular languages. It is important to note that, despite this limit on language recognition power, there is a sense in which QFAs can be more powerful than their deterministic counterparts. In particular, there are languages which can be recognized by QFAs using exponentially fewer states than the smallest deterministic or randomized finite automaton [2,6].

C. Martín-Vide, F. Otto, and H. Fernau (Eds.): LATA 2008, LNCS 5196, pp. 373–384, 2008.

The simplest type of QFA is the *measure-once QFA (MOQFA)* model of Moore and Crutchfield [11]. These QFAs are limited to recognizing those languages whose minimal automaton is such that each letter induces a permutation on the states. Two types of generalizations of the MOQFA model have been considered. In the first type, the machine is allowed to halt before reading the entire input word. This corresponds to *Kondacs and Watrous' one-way QFAs (KWQFAs)* [10]. The second type allows state transformations to include the application of quantum measurements, which generates some classical randomness in the system. This corresponds to Ambainis et. al's *Latvian QFAs (LQFAs)* [1].

Nayak [12] investigated a model called *generalized QFAs (GQFAs)*, which generalize both KWQFAs and LQFAs. This paper introduced new entropy-based techniques which were used to show that GQFAs cannot recognize the language $\Sigma^* a$. These techniques have since been used to obtain lower bounds on quantum random access codes [12] and quantum communication complexity [13]. However, no further lower bounds have been shown for GQFAs.

In a series of papers [2,7,4,3], a number lower bounds on the power of KWQFA were shown. These results identify limits on the computational advantage of KWQFAs over MOQFAs. The main tool used in these results was a technical lemma which is used to decompose the state space of a KWQFA into two subspaces (called the *ergodic* and *transient* subspaces) in which the state transitions have specific behaviors. In this paper, we show that this lemma, and many of the same results, can be adapted to the case of GQFA. The framework of our proof follows the basic outline of [2], however we must overcome a number of technical hurdles which arise from allowing classical randomness in the state.

Following [4], we can use the lemma to show that a certain property of the minimal automaton for L implies that L is not recognizable by a GQFA. We use this result to show that the class of languages recognized by this model is not closed under union. Furthermore, we show the existence of languages which can be recognized by GQFA with probability $p = 2/3$ but not $p > 2/3$. These results highlight the key similarities and differences between KWQFA and GQFA.

The paper is organized as follows. In Section 1 we give definitions and basic properties of GQFA and we review the necessary background. In Section 2 we prove the main technical lemma and in Section 3 we apply this lemma to prove the remaining results. We conclude with a brief discussion of open problems and future work.

1 Introduction

Let us review some concepts from quantum mechanics. See e.g. [14] for more details on the mathematics of quantum computation. We use the notation $|\psi\rangle$ to denote vectors in $\mathbb{C}^n$, and we denote by $\langle\psi|$ the dual of $|\psi\rangle$.

Let Q be a finite set with $|Q| = n$, and let $\{|q\rangle\}_{q\in Q}$ be an orthonormal basis for $\mathbb{C}^n$. Then a *superposition* over Q is a vector $|\psi\rangle = \sum_q \alpha_q |q\rangle$ which satisfies $\langle\psi|\psi\rangle = \sum_q |\alpha_q|^2 = 1$. We say α_q is the *amplitude* with which $|\psi\rangle$ is in state q. The state space of a QFA will be a superposition over a finite set Q.

We consider two types of operations on superpositions. First, a *unitary transformation* U is a linear operator on $\mathbb{C}^n$ such that the conjugate transpose $U^\dagger$ of U satisfies $U^\dagger U = UU^\dagger = I$. Unitary operators are exactly those which preserve the inner product, thus unitary matrices map superpositions to superpositions. The second type of operation is projective measurements. Such measurements are specified by a set $\mathcal{M} = \{P_i\}$ of orthonormal projectors on $\mathbb{C}^n$ satisfying $\sum_i P_i = I$. The *outcome* of is the measurement $\mathcal{M}$ on state $|\psi\rangle$ is the random variable which takes the value i with probability $\||P_i|\psi\rangle\|^2$. If the outcome of the measurement is i, the state is transformed to $|\psi'\rangle = P_i|\psi\rangle / \||P_i|\psi\rangle\|$. Note that measurement induces a probabilistic transformation on the state. Measurements describe the interface by which we obtain observations from a quantum system, but they also model *decoherence*, the process by which a quantum system becomes a probabilistic system through interaction with the environment (c.f. Chapter 8 of [14]).

A *generalized QFA* (GQFA) [12] is given by a tuple of the form:

$$M = (\Sigma, Q, q_0, \{U_a\}_{a\in\Gamma}, \{\mathcal{M}_a\}_{a\in\Gamma}, Q_{acc}, Q_{rej}).$$

The set Σ is the input alphabet. The working alphabet will be $\Gamma = \Sigma \cup \{¢, \$\}$. The set Q is finite set of state indices with $q_0 \in Q$, $Q_{acc}, Q_{rej} \subseteq Q$. On input $w \in \Sigma^*$, M will process the letters of the string $¢w\$$ from left to right. The ¢ and \$ characters are present to allow for pre- and post- processing of the state. The sets $\{U_a\}_{a\in\Gamma}$ and $\{\mathcal{M}_a\}_{a\in\Gamma}$ are collections of unitary transformations and projective measurements.

The state of the machine is expressed as a superposition over Q, and the initial state is $|q_0\rangle$. When a letter $a \in \Gamma$ is read, a state transformation is made in the manner we describe below. After each letter is read, the machine may decide to halt and accept the input, to halt and reject the input, or to continue processing the string. The set Q is partitioned into three parts: an *accepting* set (Q_{acc}), a *rejecting* set (Q_{rej}) and a *nonhalting* set ($Q_{non} = Q - Q_{acc} \cup Q_{rej}$). We define $P_{acc} = \sum_{q\in Q_{acc}} |q\rangle\langle q|$ and we likewise define P_{rej} and P_{non}. Finally, we define $\mathcal{M}_H = \{P_{acc}, P_{rej}, P_{non}\}$.

Suppose that after reading some input prefix the machine is in state $|\psi\rangle$. To process $a \in \Gamma$, we first apply the unitary U_a, then the measurement $\mathcal{M}_a$ (recall that this is a probabilistic transformation), then the measurement $\mathcal{M}_H$. If the outcome of the measurement $\mathcal{M}_H$ is *acc* or *rej*, then the machine halts and accepts or rejects accordingly. Otherwise, the outcome of the $\mathcal{M}_H$ was *non* and the machine reads the next symbol in the string[1].

The GQFA defined above will behave stochastically. We will be interested in what languages can be recognized by this machine with bounded error. For $p > \frac{1}{2}$ we say that language $L \subseteq \Sigma^*$ is *recognized by M with probability p* if all

[1] The original definition allowed ℓ alternations of unitary operators and measurements per letter. However, such alternations can be simulated by a single transformation and measurement (Claim 1 of [1]) and so this change does not limit the class of transformations allowed by GQFAs.

words are correctly distinguished with probability at least p. We say that L is *recognized with bounded error* if there is a $p > \frac{1}{2}$ such that L is recognized with probability p.

Here are some basic facts about GQFAs. For all p, the class of languages recognized by GQFA with probability p is closed under complement, inverse morphisms, and word quotient. We also make note of the relationship between GQFAs and other QFA definitions. Firstly, in the case that each $\mathcal{M}_a$ is equal to the trivial measurement $\{I\}$ (i.e. so that $\mathcal{M}_H$ is the only measurement applied to the state), we obtain KWQFAs as a special case. Second, in the case that we are promised that the machine does not halt until the entire input is read, then we have the special case of Ambainis et al's LQFAs. If both of these conditions hold, we obtain MOQFAs.

In this paper we will see that many of the lower bounds for KWQFAs apply also to GQFAs. It should be noted, however, that GQFA are strictly more powerful than KWQFA. In [1] it was shown that any language L whose transition monoid is a *block group* [15] can be recognized by an LQFA with probability $1 - \varepsilon$ for any $\varepsilon > 0$. This language class corresponds exactly to the boolean closure of languages of the form $L_0 a_1 L_1 \dots a_k L_k$, where the a_i's are letters and the L_i's are languages recognized by permutation automata. On the other hand, KWQFA cannot recognize $\Sigma^* a \Sigma^* b \Sigma^*$ with probability more than 7/9 [2]. It was moreover shown in [1] that LQFA cannot recognize the languages $a\Sigma^*$ or $\Sigma^* a$. We will need these properties in order to prove our results.

Furthermore it is known that KWQFA, and hence GQFA, can recognize languages which cannot be recognized by LQFA. For example KWQFA can simulate a certain type of reversible automaton where $\delta(q_1, x) = \delta(q_2, x) = q_2$ is permitted only when q_2 is a sink. These machines, and the class of languages which they recognize, were considered in [9]. Machines of this type can recognize $a\Sigma^*$, so KWQFA can recognize languages which cannot be recognized by LQFA.

Finally, a few notes about density matrices. Recall that the state of a GQFA after reading some input prefix is a random variable. In other words, the state is taken from a probability distribution $\mathcal{E} = \{(p_j, |\psi_j\rangle)\}$ of superpositions, where $|\psi_j\rangle$ occurs with probability p_j. Such systems are called *mixed states*. The measurement statistics which can be obtained from transforming and measuring a mixed state can be described succinctly in terms of *density matrices*. In our case it will be sufficient to identify a mixed state with its density matrix.

The density matrix corresponding to $\mathcal{E}$ is $\rho = \sum_j p_j |\psi_j\rangle\langle\psi_j|$. Density matrices are positive operators so their eigenvalues are nonnegative real. For a operator M we denote by $Tr(M)$ the *trace*, or the sum of the eigenvalues, of M. In the case of density matrices we have $Tr(\rho) = 1$. Unitary operators U transform density matrices according to the rule $\rho \mapsto U^\dagger \rho U$. A measurement $\mathcal{M} = \{P_i\}$ will transform the states by the rule $\rho \mapsto \sum_i P_i \rho P_i$ in the case that the outcome is unknown, or by $\rho \mapsto P_i \rho P_i / Tr(P_i \rho)$ if the outcome is known to be i.

Density matrices are examples of normal matrices. The spectral decomposition theorem states that every normal matrix can be decomposed as $\rho = \sum_i \lambda_i |\phi_i\rangle\langle\phi_i|$, where $\{|\phi_i\rangle\}$ is a set of orthonormal eigenvectors of ρ and λ_i

is the eigenvalue corresponding to $|\phi_i\rangle$. We say that the *support* of ρ, or $supp(\rho)$, is the space spanned by the nonzero eigenvectors of ρ.

2 Technical Results

Fix a GQFA M. We will be using density matrices weighted by a factor $p \in [0,1]$ to describe the state of M on reading some prefix $\mathfrak{c}w$. Let A_a be the mapping $\rho \mapsto \sum_i P_{a,i} U_a \rho U_a^\dagger P_{a,i}$, and let $A'_a = P_{non}(A_a\rho)P_{non}$. Furthermore for $w = w_1 \dots w_n \in \Sigma^*$, we define $A'_w = A'_{w_n} \cdots A'_{w_1}$. Then $A'_w\rho$ is a scaled density matrix such that $Tr(A'_w\rho) = pTr(\rho)$, where p is the probability of not halting in the process of reading w while in state ρ. Let $\rho_w = A'_{\mathfrak{c}w}|q_0\rangle\langle q_0|$. Then $Tr(\rho_w)$ is the probability of not halting while processing $\mathfrak{c}w$, and the density matrix describing the machine state in the case that it has not halted is $\rho_w/Tr(\rho_w)$.

We first state a technical lemma which gives an important characterization of the behaviour of a GQFA machine. It is the counterpart to Lemma 1 of [2]. This, along with its extension (Lemma 2), will be instrumental in proving the later results.

Lemma 1. *For every $w \in \Sigma^*$ there exists a pair E_1, E_2 of orthonormal subspaces of $\mathbb{C}^n$ such that $\mathbb{C}^n = E_1 \oplus E_2$ and for all weighted density matrices ρ over $\mathbb{C}^n$ we have:*

1. *If $supp(\rho) \subseteq E_1$, then $supp(A'_w\rho) \subseteq E_1$ and $Tr(A'_w\rho) = Tr(\rho)$.*
2. *If $supp(\rho) \subseteq E_2$, then $supp(A'_w\rho) \subseteq E_2$ and $lim_{k\to\infty} Tr((A'_w)^k\rho) = 0$.*

The E_1 and E_2 parts of the state are called the *ergodic* and *transient* parts. Suppose M is in state ρ, and suppose that ρ satisfies $supp(\rho) \subseteq E_1$. Then $Tr(A'_w\rho) = Tr(\rho)$ would imply that M did not halt in the process of reading w. Thus, M is behaving exactly as an LQFA. Suppose now that M is in state ρ, then the fact $lim_{k\to\infty} Tr((A'_w)^k\rho) = 0$ implies that the probability that M does not halt after reading w^k tends to 0 as $k \to \infty$. In general $supp(\rho)$ will be partially in E_1 and partially in E_2.

Proof: The proof proceeds as in [2]. We first show how to do this for the case that $|w| = 1$, and then we sketch how to extend it to arbitrary length words. Let $w = a$. We first construct the subspace E_1 of $\mathbb{C}^n$. E_2 will be the orthogonal complement of E_1. Let

$$E_1^1 = span(\{|\psi\rangle : Tr(A'_a|\psi\rangle\langle\psi|) = Tr(|\psi\rangle\langle\psi|)\}).$$

Equivalently, $E_1^1 = span\{|\psi\rangle : supp(A_a(|\psi\rangle\langle\psi|)) \subseteq S_{non}\}$ where S_{non} is the nonhalting subspace. We claim that $supp(\rho) \in E_1^1$ implies that $supp(A_a(\rho)) \in S_{non}$. By linearity it is sufficient to show this for $\rho = |\psi\rangle\langle\psi|$. Essentially, we need to show that the condition of $|\psi\rangle$ satisfying $Tr(A'|\psi\rangle\langle\psi|) = Tr(|\psi\rangle\langle\psi|)$ is closed under linear combinations. Suppose that $|\psi\rangle = \sum_j \alpha_j|\psi_j\rangle$, with $|\psi_j\rangle$ satisfying $supp(A_a(|\psi_j\rangle\langle\psi_j|)) \in S_{non}$ and $\sum_j |\alpha_j|^2 = 1$. Then:

$$\| \sum_i P_{halt} P_{a,i} U_a (\sum_j \alpha_j |\psi_j\rangle)\|^2 \leq \sum_{i,j} \|\alpha_j P_{halt} P_{a,i} U_a |\psi_j\rangle\|^2 = 0,$$

and thus $supp(A_a|\psi\rangle\langle\psi|) \in S_{non}$. Thus, for mixed states ρ we have $supp(A_a\rho) \in S_{non}$ if and only if $supp(\rho) \in E_1^1$. For general $i > 2$, let:

$$E_1^i = span(\{|\psi\rangle : supp(A_a|\psi\rangle\langle\psi|) \in E_1^{i-1} \wedge Tr(A'_a|\psi\rangle\langle\psi|) = Tr(|\psi\rangle\langle\psi|)\}).$$

As before, for weighted density matrices ρ, we can interchange the condition $Tr(A'_a\rho) = Tr(\rho)$ for $supp(A_a\rho) \subseteq S_{non}$.

Observe that $E_1^i \subseteq E_1^{i+1}$ for all i. Since the dimension of each of these spaces is finite, there must be an i_0 such that $E_1^{i_0} = E_1^{i_0+j}$ for all $j > 0$. We define $E_1 = E_1^{i_0}$, and set E_2 to be the orthogonal complement of E_1.

It is clear that the first condition of the lemma is true for mixed states with support in E_1. For the second part, it will be sufficient to show the following proposition, which implies that the probability with which the machine will halt while reading a^j is bounded by a constant.

Proposition 1. *Let $j \in \{1, \ldots, i_0\}$. There is a constant $\delta_j > 0$ such that for any $|\psi\rangle \in E_2^j$ there is an $l \in \{0, \ldots, j-1\}$ such that $Tr(P_{halt}A_a(A'_a)^l(|\psi\rangle\langle\psi|)) \geq \delta_j$.*

Proof: We proceed by induction on j. Let $\mathcal{H} = \bigoplus_{k=1}^{m_a} \mathbb{C}^n$. Let $P_k : E_2^1 \to \mathcal{H}$ be the projector into the kth component of $\mathcal{H}$, and let $T_1 : E_2^1 \to \mathcal{H}$ be the function $T_1|\psi\rangle = \sum_k P_k P_{halt} P_{a,k} A_a|\psi\rangle$. Observe that $\|T_1|\psi\rangle\|^2$ is the probability of halting when a is read while the machine is in state $|\psi\rangle\langle\psi|$. By the previous discussion, $Tr(A'_a|\psi\rangle\langle\psi|) = 1 - \|T_1|\psi\rangle\|^2$. Define $\|T_1\| = min_{\||\psi\rangle\|=1}\|T_1|\psi\rangle\|$. Note that the minimum exists since the set of unit vectors in $\mathbb{C}^n$ is a compact space. Also, let $\delta_1 = \|T_1\|^2$. Then $\delta_1 > 0$, otherwise there would be a vector $|\psi\rangle \in E_2^1$ such that $supp(A_a|\psi\rangle\langle\psi|) \in S_{non}$, a contradiction.

Now assume that δ_{j-1} has been found. We need to show that, for $|\psi\rangle \in E_2^j$, either a constant sized portion of $|\psi\rangle$ is sent into the halting subspace, or it is mapped to a vector on which we can apply the inductive assumption. We construct two functions $T_{j,halt}, T_{j,non} : E_2^j \to \mathcal{H}$ defined by:

$$T_{j,halt}|\psi\rangle = \sum_{k=1}^{m_a} P_k P_{halt} P_{a,k} A_a|\psi\rangle,$$

$$T_{j,non}|\psi\rangle = \sum_{k=1}^{m_a} P_k P_{E_2^{j-1}} P_{non} P_{a,k} A_a|\psi\rangle.$$

Then the quantity $\|T_{j,halt}|\psi\rangle\|^2$ is the probability of halting while reading a, and $\|T_{j,non}|\psi\rangle\|^2 = Tr(P_{E_2^{j-1}}A'_a|\psi\rangle\langle\psi|)$. Note that for all vectors $|\psi\rangle \in E_2^j$ we must have either $\|T_{j,halt}|\psi\rangle\| \neq 0$ or $\|T_{j,non}|\psi\rangle\| \neq 0$, otherwise $|\psi\rangle$ is in E_1^j, a contradiction. This implies that $\|T_{j,non} \oplus T_{j,halt}\| > 0$. Note also that $\|T_{j,non} \oplus T_{j,halt}\| \leq 1$.

Define $\delta_j = \delta_{j-1}\frac{\|T_{j,non} \oplus T_{j,halt}\|^2}{2m_a}$. Take any unit vector $\rangle\psi \in E_2^j$. Then $\|(T_{j,non} \oplus T_{j,halt})|\psi\rangle\| \geq \|T_{j,non} \oplus T_{j,halt}\|$. Recall that the range of $T_{j,non} \oplus T_{j,halt}$ is $\bigoplus_{k=1}^{m_a} \mathbb{C}^n \oplus \bigoplus_{k=1}^{m_a} \mathbb{C}^n$. In one of these subspaces, $(T_{j,non} \oplus T_{j,halt})|\psi\rangle$ has size at least $\frac{1}{\sqrt{2 \cdot m_a}}$. If it is in one of the last m_a subspaces, corresponding to $T_{j,halt}$ part,

then there is nothing further to prove. Otherwise, assume that this component is in one of the subspaces corresponding to the $T_{j,non}$ part. In particular, there is a k such that $|\phi\rangle = P_{non}P_{a,k}A_a|\psi\rangle$ satisfies:

$$\|P_{E_2^{j-1}}|\phi\rangle\|^2 \geq \frac{1}{2 \cdot m_a}.$$

We can split $|\phi\rangle$ into $|\phi_1\rangle + |\phi_2\rangle$, with $|\phi_i\rangle \in E_i^{j-1}$. By the inductive hypothesis, there is an $l < j-1$ such that $Tr(P_{halt}A_a(A'_a)^l(|\phi_2\rangle\langle\phi_2|)) \geq \delta_{j-1}Tr(|\phi_2\rangle\langle\phi_2|)$. Furthermore, the first condition of the lemma implies that for every choice of $(k_1, \ldots, k_l) \in [m^a]^l$,

$$P_{halt}P_{a,k_l}U_aP_{a,k_{l-1}}U_a \cdots P_{a,k_1}U_a|\phi_1\rangle = \mathbf{0}.$$

This implies $Tr(P_{halt}A_a(A'_a)^l(|\phi_1\rangle\langle\phi_1|)) = 0$ and $Tr(P_{halt}A_a(A'_a)^l(|\phi_1\rangle\langle\phi_2|)) = Tr(P_{halt}A_a(A'_a)^l(|\phi_2\rangle\langle\phi_1|)) = 0$. Together, we obtain:

$$\begin{aligned}
&Tr(P_{halt}A_a(A'_a)^l|\phi\rangle\langle\phi|) \\
&= Tr(P_{halt}(A'_a)^l(|\phi_1\rangle\langle\phi_1| + |\phi_1\rangle\langle\phi_2| + |\phi_2\rangle\langle\phi_1| + |\phi_2\rangle\langle\phi_2|)) \\
&= Tr(P_{halt}A_a(A'_a)^l(|\phi_1\rangle\langle\phi_1|)) + Tr(P_{halt}A_a(A'_a)^l(|\phi_1\rangle\langle\phi_2|)) \\
&\qquad + Tr(P_{halt}A_a(A'_a)^l(|\phi_2\rangle\langle\phi_1|)) + Tr(P_{halt}A_a(A'_a)^l(|\phi_2\rangle\langle\phi_2|)) \\
&= Tr(P_{halt}A_a(A'_a)^l(|\phi_2\rangle\langle\phi_2|)) \geq \delta_{j-1}\frac{\|T_{j,non} \oplus T_{j,halt}\|^2}{2m_a}.
\end{aligned}$$

This concludes the proof of the proposition. □

Proposition 2. *Let U_a be the unitary transformation that is applied when a is read. Then $U_a = U_a^1 \oplus U_a^2$, where U_a^i acts unitarily on subspace E_i.*

Proof: By the unitarity of U_a, it is sufficient to show that $|\psi\rangle \in E_1$ implies $U_a|\psi\rangle \in E_1$. By definition of E_1, $|\psi\rangle \in E_1$ implies that all of the vectors $P_{a,i}U_a|\psi\rangle$ are in E_1. But $U_a|\psi\rangle = \sum_i P_{a,i}U_a|\psi\rangle$, and thus $U_a|\psi\rangle \in E_1$ since E_1 is a subspace. □

We are now ready to prove the second part of the lemma. We first show that $|\psi\rangle \in E_2$ implies $supp(A_a|\psi\rangle\langle\psi|) \subseteq E_2$. Let $|\psi'\rangle = U_a|\psi\rangle$. Then $A_a|\psi\rangle\langle\psi| = \sum_i |\psi_i\rangle\langle\psi_i|$, where $|\psi_i\rangle = P_{a,i}U_a|\psi\rangle$. Split $|\psi_i\rangle$ into vectors $|\psi_{i,1}\rangle + |\psi_{i,2}\rangle$, with $|\psi_{i,1}\rangle \in E_1$ and $|\psi_{i,2}\rangle \in E_2$. We claim that either $|\psi_{i,1}\rangle$ or $|\psi_{i,2}\rangle$ are trivial vectors. Suppose $\||\psi_{i,1}\rangle\| \neq 0$, and consider the intersection of the image of $P_{a,i}$ in the space spanned by $|\psi_{i,1}\rangle$ and $|\psi_{i,2}\rangle$. Now $|\psi_{i,1}\rangle$ implies that $U_a^{-1}|\psi_{i,1}\rangle \in E_1$ and thus $P_{a,i}|\psi_{i,1}\rangle \in E_1$, which implies $|\psi_i\rangle \in E_1$.

Now since each $|\psi_i\rangle$ satisfies $|\psi_i\rangle \in E_1$ or $|\psi_i\rangle \in E_2$, then we are done since the fact that the $|\psi_i\rangle$'s are orthonormal and sum to $U_a|\psi\rangle \in E_2$ implies that $|\psi_i\rangle \in E_2$ for all i. Thus, $|\psi\rangle \in E_2$ implies $span(A_a|\psi\rangle\langle\psi|) \subseteq E_2$.

Now supposing $supp(\rho) \in E_2$, we can repeatedly apply Proposition 1 to show that $Tr((A'_a)^k(\rho)) \to 0$ as $k \to \infty$. To apply the claim to a general mixed state, we first use the spectral decomposition to show that the mixed state is equivalent to an ensemble of at most n pure states.

To construct E_1 and E_2 for $w = w_1 \dots w_n$, we define $E_1^0 = S_{non}$ and E_1^k to be the set of all vectors $|\psi\rangle$ such that $Tr(A'_{w_k \mod n+1}|\psi\rangle\langle\psi|) = 1$ and $supp(A'_{w_k \mod n+1}|\psi\rangle\langle\psi|) \in E_1^{k-1}$, and we follow the proof as above. The proof of the first part of the theorem and of the claim will generalize since the proof does not make use of the fact that the transformation and measurement defining E_1^j is the same as that of E_1^{j+1}. Proposition 2 will apply to w_i for all i. □

Lemma 2. *Let M be an n-state GQFA over alphabet Σ, and let $x, y \in \Sigma^*$. Then there exists a pair E_1, E_2 of orthonormal subspaces of $\mathbb{C}^n$ such that $\mathbb{C}^n = E_1 \oplus E_2$ and for all weighted density matrices ρ over $\mathbb{C}^n$ we have:*

1. *If $supp(\rho) \subseteq E_1$, then for all $w \in (x \cup y)^*$, $supp(A'_w\rho) \subseteq E_1$, and $Tr(A'_w\rho) = Tr(\rho)$.*
2. *If $supp(\rho) \subseteq E_2$, then $supp(A'_w\rho) \subseteq E_2$ and for all $\varepsilon > 0$ there exists a word $w \in (x \cup y)^*$ such that $Tr(A'_w\rho) \leq \varepsilon$.*

Proof: This is the counterpart of Lemma 2.3 of [4]. Let E_1^w be the subspace constructed as in Lemma 1. Define $E_1 = \cap_{w \in (x \cup y)^*} E_1^w$, and let E_2 be the orthogonal complement of E_1.

Suppose $supp(\rho) \subseteq E_2$. If there is a $w \in (x \cup y)^*$ such that $supp(\rho) \subseteq E_2^w$, we can directly apply the argument from the previous lemma to show that $Tr((A'_w)^j\rho) \to 0$ as $j \to \infty$. However such a w may not exist so a stronger argument is necessary. As the application of an A'_w transformation can only decrease the trace of ρ, for any ε there exists a $t \in (x \cup y)^*$ such that for all $w \in (x \cup y)^*$, $Tr(A'_t\rho) - Tr(A'_{tw}) \leq \varepsilon$. For all i let t_i be a such a string for $\varepsilon = \frac{1}{2^i}$. Consider the sequence $\rho_1, \rho_2, \dots$ defined by $\rho_i = A'_{t_i}\rho$. The set of weighted density matrices form a compact, closed space with respect to the trace metric, and so this sequence of must have a limit point ρ.

We claim that $Tr(\rho) = 0$. Suppose not. The support of ρ is in E_2, so there must be some word $w \in (x \cup y)^*$ such that $Tr(A'_w\rho) < Tr(\rho)$. This contradicts the assumption that ρ is a limit point. □

Finally we note a very simple fact that will allow us to extend impossibility results for LQFA to GQFA:

Fact 1. *Let M be a GQFA. Let E_1 be the subspace defined as in Lemma 2, and suppose that the state of the machine ρ on reading the ¢ character satisfies $supp(\rho) \in E_1$. Then there is an LQFA M' such that, for all $w \in (x \cup y)^*$ the state of M on reading w is isomorphic to the state of M' on reading w.*

3 Applications

We now apply the results of the previous section to prove several fundamental properties of GQFAs. The first result is a formal condition for recognizability by GQFAs:

Theorem 1. *Let M_L be the minimal automaton for $L \subseteq \Sigma^*$ and let F be the accepting set. If there exists words $x, y, z_1, z_2 \subseteq \Sigma^*$ and states q_0, q_1, q_2 such that $\delta(q_0, x) = q_1$, $\delta(q_0, y) = q_2$, $\delta(q_1, x) = \delta(q_1, y) = q_1$, $\delta(q_2, x) = \delta(q_2, y) = q_2$, $\delta(q_1, z_1) \in F$, $\delta(q_2, z_1) \notin F$, $\delta(q_1, z_2) \notin F$, $\delta(q_2, z_2) \in F$, then L cannot be recognized by GQFA with probability $p > \frac{1}{2}$.*

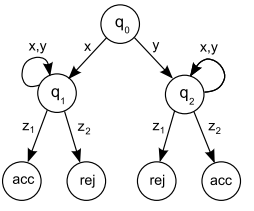

Fig. 1. The forbidden construction of Theorem 1

Proof: Suppose that L satisfies the conditions of the theorem, and suppose that M recognizes L with probability $p > \frac{1}{2}$. By closure under left quotient, we can assume that the state q_0 in the forbidden construction is also the initial state of the minimal automaton for L.

Let $\rho_w = A'_{\mathbb{C}w}|q_0\rangle\langle q_0|$. The basic outline of the proof is that we will use Lemma 2 to find two words $w_1 \in x(x \cup y)^*$, $w_2 \in y(x \cup y)^*$ such that ρ_{w_1} and ρ_{w_2} have similar output behavior. We then analyze the acceptance probabilities of the words w_1z_1, w_1z_2, w_2z_1, and w_2z_2 to arrive at a contradiction.

Let E_1 and E_2 be subspaces which meet the conditions of Lemma 2 with respect to x and y. Note that if the support of ρ is in E_1, M will not halt while reading $w \in (x \cup y)^*$, and in this case M can be simulated by an LQFA. Let P_{E_i} be the projection onto subspace E_i. We claim that for all $\varepsilon > 0$ there exists $u, v \in (x \cup y)^*$ such that $\|Tr(P_{E_1}\rho_{xu} - P_{E_1}\rho_{yv})\|_t \leq \varepsilon$. Suppose to the contrary that there exists $\varepsilon > 0$ such that $\|Tr(P_{E_1}\rho_{xu} - P_{E_1}\rho_{yv})\|_t > \varepsilon$ for all u, v. Then there exists an LQFA which can recognize the language $x(x \cup y)^*$ with bounded error, contradicting the fact that LQFA is closed under inverse morphisms and cannot recognize $a\Sigma^*$ [1]. Let $\delta = p - \frac{1}{2}$ and let $\varepsilon = \frac{\delta}{4}$.

By Lemma 2, for all ε' we can find $u' \in (x \cup y)^*$ such that $Tr(P_{E_2}\rho_{xuu'}) < \varepsilon'$. Furthermore we can find $v' \in (x \cup y)^*$ such that $Tr(P_{E_2}\rho_{xuu'v'}) < \varepsilon'$ and $Tr(P_{E_2}\rho_{yvu'v'}) < \varepsilon'$. Let $w_1 = xuu'v'$ and $w_2 = yvu'v'$, and let $\varepsilon' = \frac{\delta}{4}$.

Let $p_{i,acc}$ ($p_{i,rej}$) be the probability with which M accepts (rejects) while reading w_i. Furthermore let $q_{ij,acc}$ (resp $q_{ij,rej}$) be the probability that M accepts if the state of the machine is ρ_{w_1} and the string $z_j\$$ is read. Since $\|\rho_{w_1} - \rho_{w_2}\|_t \leq \|\rho_{xu} - \rho_{yv}\|_t = \frac{\delta}{2} \leq \varepsilon$, $q_{1j,acc}$ (and likewise $q_{1j,rej}$) can be different from $q_{2j,acc}$ by a factor of at most $\frac{\delta}{2}$. As a consequence, one of the words w_1z_1, w_1z_2, w_2z_1, or w_2z_2 must not be classified correctly. Suppose e.g. that w_1z_1, w_1z_2, and w_2z_1 are classified correctly. Since $q_{11,rej}$ differs from $q_{21,rej}$ by a factor of at most $\frac{\delta}{2}$, the fact that w_1z_1 is accepted and w_2z_1 is rejected implies that $p_{2,rej} > p_{1,rej} + \delta$. since $q_{12,rej}$ differs from $q_{22,rej}$ by at most a factor of $\frac{\delta}{2}$, will be rejected with probability greater than $1 - p$, a contradiction. The other cases are similar. □

We now apply Theorem 1 to prove nonclosure under union.

Theorem 2. *The class of languages recognized by GQFA with bounded error is not closed under union.*

Proof: Let A, B_0, B_1 be languages over $\Sigma = \{a, b\}$ defined as follows. Let $A = \{w : |w|_a \mod 2 = 0\}$, $B_0 = (aa)^*b\Sigma^*$, and $B_1 = a(aa)^*b\Sigma^*$. Finally, let $L_1 = (\overline{A} \cap a^*) \cup (A \cap B_1)$, and let $L_2 = (A \cap a^*) \cup (\overline{A} \cap B_0)$. The union $L_1 \cup L_2$ consists of the strings containing either no b's or an odd number of a's after the first b.

In Theorem 3.2 of [4], the languages L_1 and L_2 were shown to be recognizable by KWQFAs with probability of correctness 2/3, thus they can also be recognized by GQFA with this probability of correctness. On the other hand, the minimal automaton of $L_1 \cup L_2$ contains the forbidden construction of Theorem 1. □

In [2] it was shown that there exists languages L and constants $p > \frac{1}{2}$ such that L can be recognized by KWQFA with bounded probability, but not with probability p. Furthermore, it was demonstrated that certain properties of the minimal automaton for L would imply that L is not recognized with probability p. We will show that a similar situation holds for GQFAs.

Theorem 3. *If the minimal DFA M_L for L contains states q_0, q_1, q_2, such that for some words x, y, z_1, z_2 we have $\delta(q_0, x) = \delta(q_1, x) = \delta(q_1, y) = q_1$, $\delta(q_0, y) = \delta(q_2, y) = \delta(q_2, x) = q_2$, $\delta(q_2, z_2) \in F$, $\delta(q_2, z_1) \notin F$, then L cannot be recognized by GQFA with probability $p > \frac{2}{3}$.*

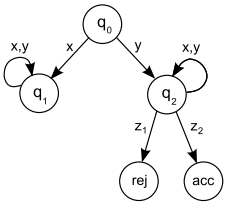

Fig. 2. The forbidden construction of Theorem 3

Proof: Suppose that the GQFA M recognizes L with probability $p > 2/3$. Since $q_2 \neq q_3$ and by closure under complement, there exists a word z_3 such that $xz_3 \in L$ and $yz_3 \notin L$. We can also assume by closure under left quotient that q_1 is the initial state. As in Lemma 2, split $\mathbb{C}^n$ into subspaces E_1 and E_2 with respect to x and y.

For all ε, we can find $w_1 \in x(x \cup y)^*$ and $w_2 \in y(x \cup y)^*$ such that $\|\rho_{w_1} - \rho_{w_2}\|_t \leq \varepsilon$, $Tr(P_{E_2}\rho_w) < \varepsilon$, $Tr(P_{E_2}\rho_w) < \varepsilon$. let p_i be the probability that M rejects while reading w_i, and let p_{i3} be the probability of rejecting when M is in state q_i and reads z_3. By setting ε, the difference between p_{13} and p_{23} can be made arbitrarily small, so that $p_1 + p_{13} \leq (1 - p) < 1/3$ and $p_2 + p_{23} \geq p > 2/3$ imply that $p_2 - p_1 > 1/3$. Thus M rejects while reading w_2 with probability greater

than 1/3, contradicting the assumption that $w_2 z_2$ is accepted with probability greater than 2/3. □

Corollary 1. *There is a language L which can be recognized by GQFAs with probability $p = 2/3$, but not with $p > 2/3$.*

To see this, note that the constructions for L_1 and L_2 in [4] achieve the optimal probability of correctness.

4 Discussion

We have shown that several of the known lower proofs for KWQFA can be adapted to the case of GQFA. In particular, we have shown that the class of languages recognized by GQFA is not closed under union, and there exists languages which can be recognized by GQFA with probability $p = 2/3$ but not $p > 2/3$. Both KWQFA and GQFA are permitted to halt before the end, and the lack of robustness in these models seems to arise from this feature. By comparison, the classes of languages recognized by MOQFA and LQFA respectively are closed under union, and any language recognized with probability $p > 1/2$ by these machines can be recognized with probability $1 - \varepsilon$ for any $\varepsilon > 0$.

We note here that not all of the KWQFA lower bound results hold for GQFA. For example, it was shown that a^*b^* can be recognized by KWQFA with probability $p \approx 0.68$ but not $p > 7/9$, while this language can be recognized by GQFA with probability $1 - \varepsilon$ for any $\varepsilon > 0$. Several other KWQFA lower bounds were shown in [4,3], and we can clarify the relationship between the two models by identifying which of these results extend to GQFAs. It is still not known whether the class of languages recognized with bounded error by GQFA is strictly larger than the class recognized by KWQFA. We conjecture that the language class is indeed larger and that a proof would involve the fact that the probability with which KWQFAs can recognize $\Sigma^* a_1 \Sigma^* \dots a_k \Sigma^*$ tends to 1/2 as $k \to \infty$.

References

1. Ambainis, A., Beaudry, M., Golovkins, M., Kikusts, A., Mercer, M., Thérien, D.: Algebraic results on quantum automata. Theory of Computing Systems 38, 165–188 (2006)
2. Ambainis, A., Freivalds, R.: 1-way quantum finite automata: strengths, weaknesses and generalizations. In: 39th Annual Symposium on Foundations of Computer Science, pp. 332–341 (1998)
3. Ambainis, A., Ķikusts, A.: Exact results for accepting probabilities of quantum automata. Theoretical Computer Science 295(1–3), 3–25 (2003)
4. Ambainis, A., Ķikusts, A., Valdats, M.: On the class of languages recognizable by 1-way quantum finite automata. In: Ferreira, A., Reichel, H. (eds.) STACS 2001. LNCS, vol. 2010, pp. 75–86. Springer, Heidelberg (2001)
5. Bennett, C.H.: Logical reversibility of computation. IBM Journal of Research and development 6, 525–532 (1973)

6. Bertoni, A., Mereghetti, C., Palano, B.: Quantum computing: 1-way quantum automata. In: Ésik, Z., Fülöp, Z. (eds.) DLT 2003. LNCS, vol. 2710, pp. 1–20. Springer, Heidelberg (2003)
7. Brodsky, A., Pippenger, N.: Characterizations of 1-way quantum finite automata. SIAM Journal on Computing 31(5), 1456–1478 (2002)
8. Ciamarra, M.P.: Quantum reversibility and a new model of quantum automaton. Fundamentals of Computation Theory 13, 376–379 (2001)
9. Golovkins, M., Pin, J.-É.: Varieties generated by certain models of reversible finite automata. In: Chen, D.Z., Lee, D.T. (eds.) COCOON 2006. LNCS, vol. 4112, pp. 83–93. Springer, Heidelberg (2006)
10. Kondacs, A., Watrous, J.: On the power of quantum finite state automata. In: 38th Annual Symposium on Foundations of Computer Science, pp. 66–75. IEEE Computer Society Press, Los Alamitos (1997)
11. Moore, C., Crutchfield, J.: Quantum automata and quantum grammars. Theoretical Computer Science 237(1-2), 275–306 (2000)
12. Nayak, A.: Optimal lower bounds for quantum automata and random access codes. In: 40th Annual Symposium on Foundations of Computer Science, pp. 369–377 (1999)
13. Nayak, A., Salzman, J.: On communication over an entanglement-assisted quantum channel. In: Proceedings of the Thirty-Fourth Annual ACM Symposium on the Theory of Computing, pp. 698–704 (2002)
14. Nielsen, M., Chuang, I.: Quantum Computation and Quantum Information. Cambridge University Press, Cambridge (2000)
15. Pin, J.-É.: k BG=PG, a success story. In: Fountain, J. (ed.) NATO Advanced Study Institute Semigroups, Formal Languages, and Groups, pp. 33–47. Kluwer Academic Publishers, Dordrecht (1995)
16. Rabin, M.: Probabilistic automata. Information and Control 6(3), 230–245 (1963)
17. Yao, A.C.-C.: Quantum circuit complexity. In: Proceedings of the 36th annual Symposium on Foundations of Computer Science, pp. 352–361 (1993)