Size Lower Bounds for Quantum Automata*

Maria Paola Bianchi, Carlo Mereghetti, and Beatrice Palano

Dip. Informatica, Univ. degli Studi di Milano, v. Comelico 39, 20135 Milano, Italy
{bianchi,mereghetti,palano}@di.unimi.it

Abstract. We compare the descriptional power of quantum finite automata with control language (QFCs) and deterministic finite automata (DFAs). By suitably adapting Rabin's technique, we show how to convert any given QFC to an equivalent DFA, incurring in an at most exponential size increase. This enables us to state a lower bound on the size of QFCs, which is logarithmic in the size of equivalent minimal DFAs. In turn, this result yields analogous size lower bounds for several models of quantum finite automata in the literature.

Keywords: quantum finite automata, descriptional complexity.

1 Introduction

While we can hardly expect to see a full-featured quantum computer in the near future, it is reasonable to envision classical computing devices incorporating small quantum components. Since the physical realization of quantum systems has proved to be a complex task, it is reasonable to keep quantum components as "small" as possible. Thus, it is well worth investigating, from a theoretical point of view, *lower limits* to the size of quantum devices when performing certain tasks, also emphasizing *trade-offs* with the size of equivalent classical devices.

Small size quantum devices are modeled by *quantum finite automata* (QFAs), a theoretical model for quantum machines with finite memory. Originally, two models of QFAs are proposed: *measure-once* QFAs (MO-QFAs) [8,14], where the probability of accepting words is evaluated by "observing" just once, at the end of input processing, and *measure-many* QFAs (MM-QFAs) [2,13], having such an observation performed after each move. Results in the literature (see, e.g., [4] for a survey) show that MO-QFAs are strictly less powerful than MM-QFAs which, in turn, are strictly less powerful than classical (deterministic or probabilistic) automata. Several modifications to these two original models of QFAs are then proposed, in order to tune computational power and motivated by different possible physical realizations. Thus, e.g., enhanced [16], reversible [9], Latvian [1] QFAs, and QFAs with quantum and classical states [21] are introduced.

Along this line of research, the model of *quantum finite automata with control language* (QFCs) is proposed in [4], as a hybrid system featuring both a quantum

* Partially supported by MIUR under the project "PRIN: Automi e Linguaggi Formali: Aspetti Matematici e Applicativi."

G. Mauri et al. (Eds.): UCNC 2013, LNCS 7956, pp. 19–30, 2013.

and a classical component. In [4,15], it is proved that the class of languages accepted with isolated cut point by QFCs coincides with regular languages, and that QFCs can be exponentially smaller than equivalent classical automata.

A relevant feature of QFCs, of interest in this paper, is that they can naturally and directly simulate several models of QFAs by preserving the size. This property makes QFCs a general unifying framework within which to investigate size results for different quantum paradigms: size lower bounds or size trade-offs proved for QFCs may directly apply to simulated types of QFAs as well. In fact, the need for a general quantum framework is witnessed by several results in the literature (see, e.g., [2,3,5,7]), showing that QFAs can be exponentially more succinct than equivalent classical automata, by means of techniques which are typically targeted on the particular type of QFA and not easily adaptable to other paradigms. So, to cope with this specialization problem, here we study size lower bounds and trade-offs for QFCs.

After introducing some basic notions, we show in Section 3 how to build from a given QFC an equivalent DFA. To this aim, we must suitably modify classical Rabin's technique [17], since the equivalence relation we choose to define the state set of the DFA is not a congruence. On the other hand, this relation – based on the classical Euclidean norm – allows us to directly estimate the cost of the conversion QFC $\to$ DFA by a geometrical argument on compact spaces. We obtain that the size of the resulting DFA is at most exponentially larger than the size of the QFC. Stated in other terms in Section 4, this latter result directly implies that QFCs are at most exponentially more succinct than classical equivalent devices. Indeed, due to QFCs generality, this succinctness result carries over other models of QFAs, such as MO-QFAs, MM-QFAs, and reversible QFAs.

2 Preliminaries

We quickly recall some notions of linear algebra, useful to describe the quantum world. For more details, we refer the reader to, e.g., [12,19]. The fields of real and complex numbers are denoted by $\mathbb{R}$ and $\mathbb{C}$, respectively. Given a complex number $z = a + ib$, we denote its *real part*, *conjugate*, and *modulus* by $z_R = a$, $z^* = a - ib$, and $|z| = \sqrt{zz^*}$, respectively. We let $\mathbb{C}^{n\times m}$ denote the set of $n \times m$ matrices with entries in $\mathbb{C}$. Given a matrix $M \in \mathbb{C}^{n\times m}$, for $1 \leq i \leq n$ and $1 \leq j \leq m$, we let M_{ij} denote its (i, j)th entry. The *transpose* of M is the matrix $M^T \in \mathbb{C}^{m\times n}$ satisfying ${M^T}_{ij} = M_{ji}$, while we let M^* be the matrix satisfying ${M^*}_{ij} = (M_{ij})^*$. The *adjoint* of M is the matrix $M^\dagger = (M^T)^*$.

For matrices $A, B \in \mathbb{C}^{n\times m}$, their *sum* is the $n\times m$ matrix $(A+B)_{ij} = A_{ij}+B_{ij}$. For matrices $C \in \mathbb{C}^{n\times m}$ and $D \in \mathbb{C}^{m\times r}$, their *product* is the $n \times r$ matrix $(CD)_{ij} = \sum_{k=1}^{m} C_{ik}D_{kj}$. For matrices $A \in \mathbb{C}^{n\times m}$ and $B \in \mathbb{C}^{p\times q}$, their *Kronecker (or tensor) product* is the $np \times mq$ matrix defined as

$$A \otimes B = \begin{pmatrix} A_{11}B & \cdots & A_{1m}B \\ \vdots & \ddots & \vdots \\ A_{n1}B & \cdots & A_{nm}B \end{pmatrix}.$$

When operations can be performed, we have that $(A \otimes B) \cdot (C \otimes D) = AC \otimes BD$. A *Hilbert space* of dimension n is the linear space $\mathbb{C}^{1\times n}$ of n-dimensional complex row vectors equipped with sum and product by elements in $\mathbb{C}$, in which the *inner product* $\langle \varphi, \psi \rangle = \varphi\psi^\dagger$ is defined. From now on, for the sake of simplicity, we will write $\mathbb{C}^n$ instead of $\mathbb{C}^{1\times n}$. The *norm* of a vector $\varphi \in \mathbb{C}^n$ is given by $\|\varphi\| = \sqrt{\langle \varphi, \varphi \rangle}$. We recall the following properties, for $\varphi, \psi, \xi, \zeta \in \mathbb{C}^n$ and $r \in \mathbb{R}$:

$$\begin{array}{cc} \langle \varphi, \psi \rangle = \langle \psi, \varphi \rangle^* = \langle \psi^*, \varphi^* \rangle, & \langle \varphi + \psi, \xi \rangle = \langle \varphi, \xi \rangle + \langle \psi, \xi \rangle, \\ \langle r\varphi, \psi \rangle = r\langle \varphi, \psi \rangle = \langle \varphi, r\psi \rangle, & \langle \varphi \otimes \psi, \xi \otimes \zeta \rangle = \langle \varphi, \xi \rangle \langle \psi, \zeta \rangle, \\ \|\varphi - \psi\|^2 = \|\varphi\|^2 + \|\psi\|^2 - 2\langle \varphi, \psi \rangle_R, & \|\varphi \otimes \psi\| = \|\varphi\| \|\psi\|. \end{array}$$

The *angle between complex vectors* φ *and* ψ is defined as (see, e.g., [18]):

$$\mathrm{ang}(\varphi, \psi) = \arccos \frac{\langle \varphi, \psi \rangle_R}{\|\varphi\| \|\psi\|}.$$

If $\langle \varphi, \psi \rangle = 0$, we say that φ is *orthogonal* to ψ. Two subspaces $X, Y \subseteq \mathbb{C}^n$ are orthogonal if any vector in X is orthogonal to any vector in Y. In this case, the linear space generated by $X \cup Y$ is denoted by $X \oplus Y$. A matrix $M \in \mathbb{C}^{n\times n}$ is said to be *unitary* whenever $MM^\dagger = I = M^\dagger M$, where $I \in \mathbb{C}^{n\times n}$ is the identity matrix. Equivalently, M is unitary if and only if it preserves the norm, i.e., $\|\varphi M\| = \|\varphi\|$ for any $\varphi \in \mathbb{C}^n$. M is said to be *Hermitian* whenever $M = M^\dagger$. For a Hermitian matrix $\mathcal{O} \in \mathbb{C}^{n\times n}$, let $c_1, \ldots, c_s$ be its eigenvalues and $E_1, \ldots, E_s$ the corresponding eigenspaces. It is well known that each eigenvalue c_k is real, that E_i is orthogonal to E_j, for every $1 \leq i \neq j \leq s$, and that $E_1 \oplus \cdots \oplus E_s = \mathbb{C}^n$. So, every vector $\varphi \in \mathbb{C}^n$ can be uniquely decomposed as $\varphi = \varphi_1 + \cdots + \varphi_s$, for unique $\varphi_j \in E_j$. The linear transformation $\varphi \mapsto \varphi_j$ is the *projector* P_j onto the subspace E_j. Actually, the Hermitian matrix $\mathcal{O}$ is biunivocally determined by its eigenvalues and projectors as $\mathcal{O} = \sum_{i=1}^s c_i P_i$. We recall that a matrix $P \in \mathbb{C}^{n\times n}$ is a projector if and only if P is Hermitian and idempotent (i.e., $P^2 = P$). As we will see, unitary matrices describe evolution in quantum systems, while Hermitian matrices represent observables to be measured.

We recall that $S \subseteq \mathbb{C}^n$ is a *compact set* if and only if every infinite sequence of elements in S contains a convergent subsequence, whose limit lies in S. For a given vector $\varphi \in \mathbb{C}^n$ and a real positive value r, we define the set $\mathcal{B}_r(\varphi) = \{v \in \mathbb{C}^n \mid \|v - \varphi\| \leq r\}$ as the *ball of radius* r *centered in* φ. The balls $\mathcal{B}_r(\varphi)$ are examples of compact sets in $\mathbb{C}^n$.

We assume the reader is familiar with basic notions on formal language theory (see, e.g., [11]). The set of all words (including the empty word ε) over a finite alphabet Σ is denoted by Σ^*, and with Σ^n we denote the set of words of length n.

A *deterministic finite state automaton* (DFA) is a 5-tuple $D = \langle Q, \Sigma, \tau, q_1, F \rangle$, where Q is the finite set of states, Σ the finite input alphabet, $q_1 \in Q$ the initial state, $F \subseteq Q$ the set of final (accepting) states, and $\tau : Q \times \Sigma \to Q$ is the transition function. An input word is *accepted*, if the induced computation starting from the state q_1 ends in some final state $q \in F$ after consuming the whole input. The set of all words accepted by D is denoted by L_D and called the accepted language. An alternative equivalent representation for D is by the

3-tuple $D = \langle \alpha, \{M(\sigma)\}_{\sigma\in\Sigma}, \beta\rangle$, where $\alpha \in \{0,1\}^{|Q|}$ is the characteristic row vector of the initial state, $M(\sigma) \in \{0,1\}^{|Q|\times|Q|}$ is the boolean matrix satisfying $(M(\sigma))_{ij} = 1$ if and only if $\tau(q_i,\sigma) = q_j$, and $\beta \in \{0,1\}^{|Q|\times 1}$ is the characteristic column vector of the final states. The accepted language can now be defined as $L_D = \{\sigma_1\cdots\sigma_n \in \Sigma^* \mid \alpha M(\sigma_1)\cdots M(\sigma_n)\beta = 1\}$.

Let us now introduce the model of quantum finite automata with control language [4,15].

Definition 1. *Given an input alphabet Σ and an endmarker symbol $\sharp \notin \Sigma$, a q-state* quantum finite automaton with control language (QFC) *is a system $\mathcal{A} = \langle \phi, \{U(\gamma)\}_{\gamma\in\Gamma}, \mathcal{O}, \mathcal{L}\rangle$, for $\Gamma = \Sigma \cup \{\sharp\}$, where*

- *$\phi \in \mathbb{C}^q$ is the initial amplitude vector satisfying $\|\phi\| = 1$,*
- *$U(\gamma) \in \mathbb{C}^{q\times q}$ is a unitary matrix, for any $\gamma \in \Gamma$,*
- *$\mathcal{O} = \sum_{c\in C} cP(c)$ is a Hermitian matrix representing an* observable *where C, the set of eigenvalues of $\mathcal{O}$, is the set of all possible outcomes of measuring $\mathcal{O}$, and $P(c)$ denotes the projector onto the eigenspace corresponding to $c \in C$,*
- *$\mathcal{L} \subseteq C^*$ is a regular language, called the* control language.

An input for $\mathcal{A}$ is any word from Σ^* closed by the symbol $\sharp$. The behavior of $\mathcal{A}$ on $x_1\cdots x_n\sharp \in \Sigma^*\sharp$ is as follows. At any time, the state of $\mathcal{A}$ is a vector $\xi \in \mathbb{C}^q$ with $\|\xi\| = 1$. The computation starts in the state ϕ, then transformations associated with the symbols in $x_1\cdots x_n\sharp$ are applied in succession. Precisely, the transformation corresponding to a symbol $\gamma \in \Gamma$ consists of two steps:

(i) EVOLUTION: the unitary operator $U(\gamma)$ is applied to the current state ξ of the automaton, leading to the new state ξ'.
(ii) MEASURING: the observable $\mathcal{O}$ is measured on ξ'. According to quantum mechanics principles, the result of measurement is c_k with probability $\|\xi' P(c_k)\|^2$, and the state of the automaton "collapses" to $\frac{\xi' P(c_k)}{\|\xi' P(c_k)\|}$.

So, the computation on $x_1\cdots x_n\sharp$ yields a given sequence $y_1\cdots y_n y_\sharp$ of results of the measurements of $\mathcal{O}$ with probability $p_{\mathcal{A}}(y_1\cdots y_n y_\sharp; x_1\cdots x_n\sharp)$ defined as

$$p_{\mathcal{A}}(y_1\cdots y_n y_\sharp; x_1\cdots x_n\sharp) = \left\| \phi \left(\prod_{i=1}^{n} U(x_i)P(y_i) \right) U(\sharp)P(y_\sharp) \right\|^2 .$$

A computation yielding the word $y_1\cdots y_n y_\sharp$ of measure outcomes is *accepting* whenever $y_1\cdots y_n y_\sharp \in \mathcal{L}$, otherwise it is rejecting. Hence, the probability that the QFC $\mathcal{A}$ exhibits an accepting computation on input $x_1\cdots x_n\sharp$ is

$$\mathcal{E}_{\mathcal{A}}(x_1\cdots x_n) = \sum_{y_1\cdots y_n y_\sharp \in \mathcal{L}} p_{\mathcal{A}}(y_1\cdots y_n y_\sharp; x_1\cdots x_n\sharp).$$

The function $\mathcal{E}_{\mathcal{A}} : \Sigma^* \to [0,1]$ is the *stochastic event induced by* $\mathcal{A}$.

The *language accepted by $\mathcal{A}$ with cut point $\lambda \in [0,1]$* is the set of words $L_{\mathcal{A},\lambda} = \{x \in \Sigma^* \mid \mathcal{E}_{\mathcal{A}}(x) > \lambda\}$. The cut point is said to be *isolated* whenever there exists $\delta \in (0, \frac{1}{2}]$ such that $|\mathcal{E}_{\mathcal{A}}(x) - \lambda| \geq \delta$, for any $x \in \Sigma^*$.

When referring to the size of a QFC, we must account for both the quantum and the classical component. Hence, in what follows, we say that $\mathcal{A}$ has q quantum states and k classical states whenever it is a q-state QFC and the control language $\mathcal{L}$ is recognized by a k-state DFA.

Throughout the paper, we say that two automata are *equivalent* whenever they accept the same language.

3 Converting QFCs to DFAs

We start by defining a matrix representation for QFCs. Then, for any given QFC, we construct an equivalent DFA by suitably generalizing Rabin's technique. Finally, we analyze the state complexity of the resulting DFA with respect to the size of the original QFC.

3.1 Linear Representation of QFCs

A convenient way to work with QFCs is by using their linear representation [4]. Let $\mathcal{A} = \langle \phi,\ \{U(\sigma)\}_{\sigma\in\Gamma},\ \mathcal{O} = \sum_{c\in C} cP(c),\ \mathcal{L}\rangle$ be a QFC with δ-isolated cut point λ, and let $D = \langle \alpha,\ \{M(c)\}_{c\in C},\ \beta\rangle$ be the minimal DFA recognizing $\mathcal{L}$. Denote by q and k the number of quantum and classical states of $\mathcal{A}$. We define the *linear representation* of $\mathcal{A}$ as the 3-tuple $\mathrm{Li}(\mathcal{A}) = \langle \varphi_0,\ \{V(\sigma)\}_{\sigma\in\Gamma},\ \eta\rangle$ with

- $\varphi_0 = (\phi \otimes \phi^* \otimes \alpha)$, a vector in $\mathbb{C}^{q^2k}$,
- $V(\sigma) = (U(\sigma)\otimes U^\dagger(\sigma)\otimes I)\cdot\sum_{c\in C} P(c)\otimes P(c)\otimes M(c)$, a matrix in $\mathbb{C}^{q^2k\times q^2k}$,
- $\eta = \sum_{j=1}^{q} e_j \otimes e_j \otimes \beta$, a vector in $\mathbb{C}^{q^2k}$,

where e_j is the vector with 1 in its jth component and 0 elsewhere. The main point, not so hard to verify, is that $\mathrm{Li}(\mathcal{A})$ *enables us to represent the stochastic event induced by* $\mathcal{A}$ *as* $\mathcal{E}_{\mathcal{A}}(x) = \varphi_0 V(x^\sharp)\,\eta$, where we let $V(\omega) = \prod_{i=1}^{n} V(\sigma_i)$ for any $\omega = \sigma_1\cdots\sigma_n \in \Gamma^*$. In addition, as shown in [4], we have $\|\varphi_0 V(\omega)\| \leq 1$ for any $\omega \in \Gamma^*$. Therefore, all the state vectors of $\mathrm{Li}(\mathcal{A})$ belong to the unitary ball $\mathcal{B}_1(\mathbf{0}) \subset \mathbb{C}^{q^2k}$ centered in the zero-vector $\mathbf{0}$.

We are going to show a crucial result saying, roughly speaking, that any word ω induces an evolution in $\mathrm{Li}(\mathcal{A})$ which increases the distance between two different starting vectors only by a constant factor *not depending on the length of* ω. To this aim, we need some technical lemmas, the first one shown in [4]:

Lemma 1. *For any $\sigma \in \Sigma$, let $U(\sigma)$ be a unitary matrix, and let an observable $\mathcal{O} = \sum_{c\in C} cP(c)$. Then, for any complex vector φ and word $\sigma_1\cdots\sigma_n \in \Gamma^*$, we have $\sum_{y=y_1\cdots y_n\in C^n} \|\varphi \prod_{j=1}^{n} U(\sigma_j)P(y_j)\|^2 = \|\varphi\|^2$.*

The next lemma states a property of vectors lying within unitary balls. From now on, for the sake of brevity, we will simply write $\mathcal{B}_1$ to denote a unitary ball centered in $\mathbf{0}$, regardless the dimension of the space within which such a ball is embedded.

Lemma 2. *For any $v, v' \in \mathcal{B}_1$ satisfying $\|v'\| \geq \|v\|$ and $\cos(\mathrm{ang}(v', v)) \geq 0$, we have $\cos(\mathrm{ang}(v' - v, v)) \geq -\frac{1}{\sqrt{2}}$.*

We are now ready to prove the crucial result on the distance between trajectories in $\mathrm{Li}(\mathcal{A})$:

Lemma 3. *For any state vectors $\varphi = v \otimes v^* \otimes a$ and $\varphi' = v' \otimes v'^* \otimes a'$ of $\mathrm{Li}(\mathcal{A})$, and any $\omega \in \Gamma^*$, we have*

$$\|\varphi' V(\omega) - \varphi V(\omega)\| \leq 4\|\varphi' - \varphi\|. \tag{1}$$

Proof. We consider the case in which $a = a'$, and quickly address the opposite case at the end of the proof. Without loss of generality, we can assume that $\|v'\| \geq \|v\|$. Moreover, we assume that $\cos(\mathrm{ang}(v', v)) \geq 0$. Otherwise, we can consider the vector $-v'$ instead of v', for which it holds $\cos(\mathrm{ang}(-v', v)) \geq 0$, and the proof works unchanged since $(-v') \otimes (-v')^* \otimes a = v' \otimes v'^* \otimes a = \varphi'$.

By letting $\Delta = v' - v$, we have $\varphi' - \varphi = v \otimes \Delta^* \otimes a + \Delta \otimes v^* \otimes a + \Delta \otimes \Delta^* \otimes a$. So, we can rewrite the left side of Inequality (1) as

$$\begin{aligned}\|(\varphi' - \varphi)V(\omega)\| &= \|(v \otimes \Delta^* \otimes a)V(\omega) + (\Delta \otimes v^* \otimes a)V(\omega) + (\Delta \otimes \Delta^* \otimes a)V(\omega)\| \\ &\leq \|(v \otimes \Delta^* \otimes a)V(\omega)\| + \|(\Delta \otimes v^* \otimes a)V(\omega)\| + \|(\Delta \otimes \Delta^* \otimes a)V(\omega)\|. \end{aligned} \tag{2}$$

To simplify Inequality (2), we analyze the generic form $\|(v_1 \otimes {v_2}^* \otimes a)V(\omega)\|$, which can be written as

$$\left\| \sum_{y = y_1 \cdots y_n \in C^n} v_1 \prod_{j=1}^{n} U(\sigma_j)P(y_j) \otimes {v_2}^* \prod_{j=1}^{n} U^\dagger(\sigma_j)P(y_j) \otimes aM(y) \right\|.$$

Since D, the automaton for the control language $\mathcal{L} \subseteq C^*$ in $\mathcal{A}$, is a DFA, we have $\|aM(y)\| = 1$ for every $y \in C^*$. So, we can write

$$\|(v_1 \otimes {v_2}^* \otimes a)V(\omega)\| \leq \sum_{y = y_1 \cdots y_n \in C^n} \left\| v_1 \prod_{j=1}^{n} U(\sigma_j)P(y_j) \right\| \cdot \left\| {v_2}^* \prod_{j=1}^{n} U^\dagger(\sigma_j)P(y_j) \right\|.$$

The right side of this inequality can be seen as the inner product between two vectors $\hat{v}_1, \hat{v}_2$ of dimension $|C|^n$, with the yth component of $\hat{v}_1$ (resp., $\hat{v}_2$) being $\left\| v_1 \prod_{j=1}^{n} U(\sigma_j)P(y_j) \right\|$ (resp., $\left\| {v_2}^* \prod_{j=1}^{n} U^\dagger(\sigma_j)P(y_j) \right\|$). By Cauchy-Schwarz inequality, we have $|\langle \hat{v}_1, \hat{v}_2 \rangle| \leq \|\hat{v}_1\| \|\hat{v}_2\|$. So, by Lemma 1, we can write

$$\begin{aligned}\|(v_1 \otimes {v_2}^* \otimes a)V(\omega)\| &\leq \sqrt{\sum_{y_1 \cdots y_n} \|v_1 \prod_{j=1}^{n} U(\sigma_j)P(y_j)\|^2} \cdot \sqrt{\sum_{y_1 \cdots y_n} \|{v_2}^* \prod_{j=1}^{n} U^\dagger(\sigma_j)P(y_j)\|^2} \\ &= \sqrt{\|v_1\|^2 \|v_2\|^2} = \|v_1\| \|v_2\|. \end{aligned}$$

By replacing v_1 and v_2 with the vectors involved in Inequality (2), we obtain

$$\|\varphi' V(\omega) - \varphi V(\omega)\| \leq 2\|v\|\|\Delta\| + \|\Delta\|^2. \tag{3}$$

We now analyze the right side of Inequality (1). We first observe that

$$\begin{aligned}\|\varphi' - \varphi\|^2 =& \|v \otimes \Delta^* + \Delta \otimes v^* + \Delta \otimes \Delta^*\|^2 \qquad \text{(since } \|a\| = 1\text{)}\\ =& \|v\|^2\|\Delta\|^2 + \|\Delta\|^2\|v\|^2 + \|\Delta\|^2\|\Delta\|^2 + 2(\langle v, \Delta\rangle\langle \Delta^*, v^*\rangle)_R + \\ &+ 2(\langle v, \Delta\rangle\langle \Delta^*, \Delta^*\rangle)_R + 2(\langle \Delta, \Delta\rangle\langle v^*, \Delta^*\rangle)_R \\ =& \|v\|^2\|\Delta\|^2 + \|\Delta\|^2\|v\|^2 + \|\Delta\|^2\|\Delta\|^2 + \\ &+ 2|\langle v, \Delta\rangle|^2 + 2(\langle v, \Delta\rangle\|\Delta\|^2)_R + 2(\|\Delta\|^2\langle v^*, \Delta^*\rangle)_R \\ \geq& 2\|v\|^2\|\Delta\|^2 + \|\Delta\|^4 + 2(\langle v, \Delta\rangle_R)^2 + 4\|\Delta\|^2\langle v, \Delta\rangle_R .\end{aligned}$$

By letting $\theta = \text{ang}(v, \Delta)$, we have

$$\|\varphi' - \varphi\|^2 \geq 2\|v\|^2\|\Delta\|^2 + \|\Delta\|^4 + 2\|v\|^2\|\Delta\|^2(\cos(\theta))^2 + 4\|v\|\|\Delta\|^3\cos(\theta). \quad (4)$$

By joining Inequalities (3) and (4), in order to prove the desired Inequality (1) it is enough to show that

$$(2\|v\|\|\Delta\| + \|\Delta\|^2)^2 \leq 16(\|\Delta\|^4 + 4\|v\|\|\Delta\|^3\cos(\theta) + 2\|v\|^2\|\Delta\|^2(1 + (\cos(\theta))^2)).$$

We can divide both sides by $\|\Delta\|^2$, since for $\|\Delta\| = 0$ the inequality is trivially verified. By solving with respect to $\|\Delta\|$, we get that the inequality is always true if it holds $4\|v\|^2(16\cos(\theta) - 1)^2 - 60\|v\|^2(8\ (\cos(\theta))^2 + 7) \leq 0$. If $\|v\| = 0$, this is clearly verified. Otherwise, dividing by $\|v\|^2$ and routine manipulation lead us to study the equivalent inequality

$$17(\cos(\theta))^2 - 4\cos(\theta) - 13 \leq 0. \quad (5)$$

Recall that, at the beginning of the proof, we assumed that $\|v'\| \geq \|v\|$ and $\cos(\text{ang}(v, v')) \geq 0$. So, by Lemma 2, we get $-\frac{1}{\sqrt{2}} \leq \cos(\theta) \leq 1$. Within this interval, the left side of Inequality (5) is never positive, whence the result follows.

We conclude by quickly noticing that in the case $a \neq a'$, we have $\langle a, a'\rangle = 0$. So, one may easily obtain $\|\varphi' - \varphi\|^2 = \|v'\|^4 + \|v\|^4$ and $\|(\varphi' - \varphi)V(\omega)\| \leq \|v'\|^2 + \|v\|^2$, and the claimed result again follows. □

3.2 Conversion to DFAs

We are now ready to construct a DFA $D_{\mathcal{A}}$ equivalent to the QFC $\mathcal{A}$, by using the linear representation $\text{Li}(\mathcal{A}) = \langle \varphi_0, \{V(\sigma)\}_{\sigma\in\Gamma}, \eta\rangle$.

For any word $\omega \in \Sigma^*$, let $\varphi_\omega = \varphi_0 V(\omega)$ be the state vector reached by $\text{Li}(\mathcal{A})$ after reading ω. We define the relation $\sim$ on the set $\{\varphi_\omega \mid \omega \in \Sigma^*\} \subseteq \mathcal{B}_1$ as:

$$\varphi_\omega \sim \varphi_{\omega'} \iff \begin{array}{l}\text{there exists a sequence of words } \omega_1, \omega_2, \ldots, \omega_n \in \Sigma^* \\ \text{satisfying } \omega = \omega_1,\ \omega' = \omega_n,\ \text{and } \|\varphi_{\omega_i} - \varphi_{\omega_{i+1}}\| < \frac{\delta}{2\sqrt{qk}}.\end{array}$$

It is easy to verify that $\sim$ is an equivalence relation, and that the distance between two vectors belonging to different equivalence classes is at least $\frac{\delta}{2\sqrt{qk}}$. This latter fact shows that $\sim$ is of *finite* index, since otherwise, by taking one

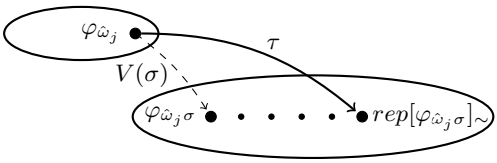

Fig. 1. The transition τ on a symbol σ. The dots represent state vectors of $\mathrm{Li}(\mathcal{A})$, while the ellipses indicate equivalence classes of $\sim$. The smaller points between $\varphi_{\hat{\omega}_j\sigma}$ and $rep[\varphi_{\hat{\omega}_j\sigma}]_\sim$ represent the state vectors at distance smaller than $\frac{\delta}{2\sqrt{qk}}$ witnessing the relation $\sim$ between them. The dashed arrow indicates the original evolution on $\mathrm{Li}(\mathcal{A})$, while the full arrow represents the behavior of the DFA $D_{\mathcal{A}}$.

vector from each class, one could construct an infinite sequence of elements in $\mathcal{B}_1$ which cannot have any convergent subsequence, against the compactness of $\mathcal{B}_1$. Therefore, by letting s be the index of $\sim$, we choose a representative for each equivalence class, and call them $\varphi_{\hat{\omega}_1}, \varphi_{\hat{\omega}_2}, \ldots, \varphi_{\hat{\omega}_s}$. In addition, for any word $\omega \in \Sigma^*$, we let $rep[\varphi_\omega]_\sim$ denote the representative of the equivalence class the state vector φ_ω belongs to.

We construct our DFA $D_{\mathcal{A}}$ as follows:

- the *set of states* coincides with the set of representatives $\{\varphi_{\hat{\omega}_1}, \varphi_{\hat{\omega}_2}, \ldots, \varphi_{\hat{\omega}_s}\}$,
- the *input alphabet* is Σ,
- the *initial state* is the vector $rep[\varphi_\varepsilon]_\sim$, which we assume to be $\varphi_{\hat{\omega}_1}$,
- the *transition function* is defined, for any $\sigma \in \Sigma$, as $\tau(\varphi_{\hat{\omega}_j}, \sigma) = rep[\varphi_{\hat{\omega}_j\sigma}]_\sim$; a step of τ is intuitively shown in Fig. 1,
- the *final states* are the representatives $\{\varphi_{\hat{\omega}_j} \mid \varphi_{\hat{\omega}_j} V(\sharp)\, \eta \geq \lambda + \delta\}$ associated with words accepted in the original QFC $\mathcal{A}$; equivalently, $\varphi_{\hat{\omega}_j}$ is final if and only if its equivalence class contains φ_ω for some word $\omega\sharp$ accepted by $\mathcal{A}$.

Before showing the correctness of our construction, we stress the fact that the equivalence relation $\sim$ *is not a congruence* (in fact, $\varphi_\omega \sim \varphi_{\omega'}$ does not necessarily implies $\varphi_{\omega\sigma} \sim \varphi_{\omega'\sigma}$ for $\sigma \in \Sigma$, as the reader may easily verify). So, the correctness does not come straightforwardly as in Rabin's setting, but we need an explicit proof:

Theorem 1. *$D_{\mathcal{A}}$ is equivalent to $\mathcal{A}$.*

Proof. We begin by introducing some notation:

- For a word $z = z_1 z_2 \cdots z_n \in \Sigma^*$, we let $z_{\{j\}} = z_1 z_2 \cdots z_j$ be the prefix of z of length j, and $z_{\{-j\}} = z_{j+1} z_{i+2} \cdots z_n$ the remaining suffix.
- We let $\rho_j = \tau(\varphi_{\hat{\omega}_1}, z_{\{j\}})$ be the state reached by $D_{\mathcal{A}}$ after reading the first j symbols of z. So, $\rho_0 = \varphi_{\hat{\omega}_1}$ is the initial state of $D_{\mathcal{A}}$.
- We let $\psi_j = \rho_{j-1} V(z_j)$ be the state vector reached by $j-1$ steps of $D_{\mathcal{A}}$ followed by one step of $\mathrm{Li}(\mathcal{A})$. So, $\psi_0 = \varphi_0$ is the initial state of $\mathrm{Li}(\mathcal{A})$.

Note that, for each $0 \leq j \leq n$, we have $\psi_j \sim \rho_j$ since $\rho_j = rep[\psi_j]_\sim$. Moreover, by definition, the vectors witnessing $\psi_j \sim \rho_j$ are reachable in $\mathrm{Li}(\mathcal{A})$. Formally: there

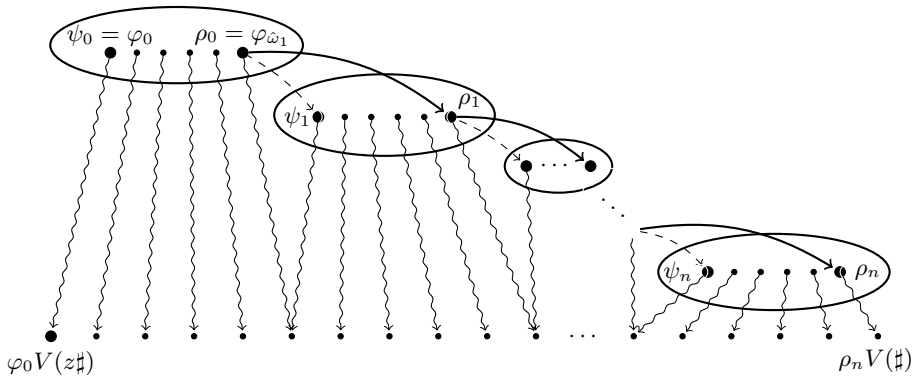

Fig. 2. Evolution scheme of the computation over the word $z\sharp$. The full arrows describe the transitions of the DFA $D_{\mathcal{A}}$, while the snake arrows denote the evolution in $\mathrm{Li}(\mathcal{A})$ from each vector $\gamma_{j,t}$ in the equivalence class reached after j symbols, through the dynamic V over the remaining suffix $z_{\{-j\}}\sharp$, leading to the vector $\gamma_{j,t}V(z_{\{-j\}}\sharp)$ in the bottom chain. In this bottom chain, the leftmost point denotes the vector reached by $\mathrm{Li}(\mathcal{A})$ after reading $z\sharp$, while the rightmost point is the state reached by $D_{\mathcal{A}}$ after reading z, with a final transition of $\mathrm{Li}(\mathcal{A})$ on $\sharp$. Intuitively, the correctness of $D_{\mathcal{A}}$ comes from the fact that all the vectors in the bottom chain are sufficiently close to their neighbors to represent either all accepting or all rejecting quantum states in the original QFC $\mathcal{A}$.

exists a sequence $\psi_j = \gamma_{j,1}, \gamma_{j,2}, \ldots \gamma_{j,\ell_j} = \rho_j$ satisfying $\|\gamma_{j,i} - \gamma_{j,i+1}\| < \frac{\delta}{2\sqrt{qk}}$, and there exist $x_{j,t} \in \Sigma^*$ such that $\varphi_0 V(x_{j,t}) = \gamma_{j,t}$ for $1 \leq t \leq \ell_j$. As a consequence of Lemma 3, for every $0 \leq j \leq n$ and $1 \leq t \leq \ell_j$, we have

$$\|\gamma_{j,t}V(z_{\{-j\}}\sharp) - \gamma_{j,(t+1)}V(z_{\{-j\}}\sharp)\| < 4 \cdot \frac{\delta}{2\sqrt{qk}} = \frac{2\delta}{\sqrt{qk}}. \tag{6}$$

In addition, since

$$\rho_j V(z_{\{-j\}}\sharp) = \psi_{j+1} V(z_{\{-(j+1)\}}\sharp),$$

for all j's, Inequality (6) implies that the vectors $\rho_j V(z_{\{-j\}}\sharp)$ form a chain of vectors from the final state vector $\varphi_0 V(z\sharp)$ of $\mathrm{Li}(\mathcal{A})$ to the vector $\rho_n V(\sharp)$, where the distance between each pair of consecutive vectors is strictly smaller than $\frac{2\delta}{\sqrt{qk}}$. This is intuitively shown in Fig. 2.

We first show that $z \in L_{\mathcal{A},\lambda} \;\Rightarrow\; \tau(\varphi_{\hat{\omega}_1}, z) \in F$, which is equivalent to showing

$$\varphi_0 V(z\sharp)\,\eta \geq \lambda + \delta \;\Rightarrow\; \rho_n V(\sharp)\,\eta \geq \lambda + \delta. \tag{7}$$

Note that $\varphi_0 = \gamma_{0,1}$, $\rho_n = \gamma_{n,\ell_n}$, and that, for $0 \leq j \leq n$ and $1 \leq t \leq \ell_j$, all $\gamma_{j,t}$'s witnessing the relation $\sim$ are reachable in $\mathrm{Li}(\mathcal{A})$ through some word $x_{j,t} \in \Sigma^*$, i.e., $\gamma_{j,t}V(z_{\{-j\}}\sharp) = \varphi_0 V(x_{j,t} \cdot z_{\{-j\}}\sharp)$. Since λ is a δ-isolated cut point, we have

$$\gamma_{j,t}V(z_{\{-j\}}\sharp)\,\eta \begin{cases} \geq \lambda + \delta & \text{if } x_{j,t}z_{\{-j\}} \in L_{\mathcal{A},\lambda}, \\ \leq \lambda - \delta & \text{if } x_{j,t}z_{\{-j\}} \notin L_{\mathcal{A},\lambda}. \end{cases}$$

Assume, by contradiction, that Inequality (7) does not hold. Then, there exists a position in the bottom chain of Fig. 2 where the acceptance probability associated with a state vector in the chain is above the cut point, while the acceptance probability associated to its right neighbor is below the cut point. More formally, there must exist ι, κ such that:

$$\gamma_{\iota,\kappa} V(z_{\{-\iota\}}\sharp)\,\eta \geq \lambda + \delta \quad \text{and} \quad \gamma_{\iota,(\kappa+1)} V(z_{\{-\iota\}}\sharp)\,\eta \leq \lambda - \delta,$$

From these two inequalities and by observing that $\|\eta\| \leq \sqrt{qk}$, we get

$$\begin{aligned} 2\delta \leq & \|(\gamma_{\iota,\kappa} V(z_{\{-\iota\}}\sharp) - \gamma_{\iota,(\kappa+1)} V(z_{\{-\iota\}}\sharp))\eta\| \\ \leq & \|\gamma_{\iota,\kappa} V(z_{\{-\iota\}}\sharp) - \gamma_{\iota,(\kappa+1)} V(z_{\{-\iota\}}\sharp)\|\|\eta\| \\ \leq & \|\gamma_{\iota,\kappa} V(z_{\{-\iota\}}\sharp) - \gamma_{\iota,(\kappa+1)} V(z_{\{-\iota\}}\sharp)\| \cdot \sqrt{qk} \\ < & \frac{2\delta}{\sqrt{qk}} \cdot \sqrt{qk} = 2\delta \qquad \text{(by Inequality 6).} \end{aligned}$$

which is an *absurdum.*

Symmetrically, one can show that $z \notin L_{\mathcal{A}} \Rightarrow \tau(\varphi_{\hat{\omega}_1}, z) \notin F$, and this completes the proof.

3.3 Size Cost of the Conversion

We now analyze the cost, in terms of number of states, of the above conversion from QFCs to DFAs. This will enable us to obtain a general gap at most exponential between the succinctness of the quantum and classical paradigm.

Theorem 2. *For any given* QFC $\mathcal{A}$ *with q quantum states, k classical states, and δ-isolated cut point, there exists an equivalent* DFA $D_{\mathcal{A}}$ *with s states satisfying*

$$s \leq \left(1 + \frac{4\sqrt{qk}}{\delta}\right)^{q^2 k}.$$

Proof. Let $\mathrm{Li}(\mathcal{A}) = \langle \varphi_0, \{V(\sigma)\}_{\sigma \in \Gamma}, \eta \rangle$ be the linear representation of $\mathcal{A}$. As observed in Section 3.1, its state vectors lies within $\mathcal{B}_1(\mathbf{0}) \subset \mathbb{C}^d$, for $d = q^2 k$. When constructing the equivalent DFA $D_{\mathcal{A}}$ as described in Section 3.2, the number s of states of $D_{\mathcal{A}}$ coincides with the number of equivalence classes of the relation $\sim$.

To estimate s, consider the ball $\mathcal{B}_{\frac{\delta}{4\sqrt{qk}}}(\varphi_{\hat{\omega}_i}) \subset \mathbb{C}^d$, for each representative $\varphi_{\hat{\omega}_i}$. Clearly, such a ball is disjoint from the analogous ball centered in $\varphi_{\hat{\omega}_j}$, for every $1 \leq i \neq j \leq s$. Moreover, all such balls are contained in $\mathcal{B}_{1+\frac{\delta}{4\sqrt{qk}}}(\mathbf{0}) \subset \mathbb{C}^d$, and their number is exactly the number s of equivalence classes of $\sim$. Since the volume of a d-dimensional ball of radius r is Kr^d, for a suitable constant K depending on d, there exist at most

$$\frac{K(1 + \delta/4\sqrt{qk})^d}{K(\delta/4\sqrt{qk})^d} = \left(1 + \frac{4\sqrt{qk}}{\delta}\right)^{q^2 k}$$

balls of radius $\frac{\delta}{4\sqrt{qk}}$ in $\mathcal{B}_{1+\frac{\delta}{4\sqrt{qk}}}(\mathbf{0})$. So, this number is an upper bound for s. □

4 Size Lower Bound for Quantum Paradigms

By using the inequality of Theorem 2 "the other way around", we are able to state lower limits to the descriptional power of QFCs:

Theorem 3. *Any* QFC *with q quantum states, k classical states, and δ-isolated cut point accepting a regular language whose minimal* DFA *has μ states, satisfies*

$$qk \geq \left(\frac{\log(\mu)}{\log\left(\frac{5}{\delta}\right)}\right)^{\frac{4}{9}}.$$

Proof. From our QFC, we can obtain an equivalent DFA with a number of states bounded as in Theorem 2. Thus, for $\delta \in (0, \frac{1}{2}]$ and $q, k \geq 1$, we have

$$\mu \leq \left(1 + \frac{4\sqrt{qk}}{\delta}\right)^{q^2 k} \leq \left(\frac{5\sqrt{qk}}{\delta}\right)^{q^2 k} \leq \left(\frac{5}{\delta}\right)^{\sqrt[4]{qk}\cdot q^2 k} \leq \left(\frac{5}{\delta}\right)^{(qk)^{\frac{9}{4}}},$$

whence the result follows. □

The lower bound in Theorem 3 is not only interesting in the world of QFCs, but it turns out to have several applications in the world of quantum automata. In fact, as recalled in the Introduction, QFCs represent a general unifying framework within which several types of quantum automata may directly and naturally be represented. In particular, in [4] it is proved that: (i) Any q-state measure-once quantum finite automaton (MO-QFA) can be simulated by a QFC with $2q$ quantum states and 1 classical state. (ii) Any q-state measure-many quantum finite automaton (MM-QFA) can be simulated by a QFC with q quantum states and 3 classical states. (iii) Any q-state quantum reversible automaton (QRA) can be simulated by a QFC with q quantum states and 2 classical states. So, by such simulation results and Theorem 3, one immediately gets

Theorem 4. *To accept a regular language having a μ-state minimal* DFA *by a* MO-QFA*,* MM-QFA *or* QRA*, at least $\kappa\left(\log(\mu)/\log\left(\frac{5}{\delta}\right)\right)^{4/9}$ states are necessary, with $\kappa = 1/2$ for* MO-QFA *and* QRA*, and $\kappa = 1/3$ for* MM-QFA*.*

A better asymptotically optimal lower bound of $\log(\mu)/(2\log(1 + 2/\delta))$ is obtained in [6] for MO-QFAs. There, however, Rabin's approach has a more direct application since the equivalence relation yielding the states of the equivalent DFA is in fact a congruence, so the correctness of the DFA is straightforward. In the case of QFCs, instead, the equivalence relation $\sim$ is not a congruence, so we had to ensure that, starting from two different state vectors in the same equivalence class, after the evolution on the same word, the two resulting vectors are still either both accepting or both rejecting, even if they belong to different classes. This was possible because of the property proved in Lemma 3.

As natural open problems, it remains either to witness the optimality of our size lower bound for QFCs, or to improve it, especially for the particular cases of simulated machines such as, e.g., MM-QFAs and QRAs. Moreover, one of the

anonymous referees pointed out another general framework, namely *quantum automata with open time evolution* [10], which may be worth investigating by the same geometrical approach, since the computation of such devices on a given input can also be linearized [20].

Acknowledgements. The authors wish to thank Alberto Bertoni for useful discussions, and the anonymous referees for their comments.

References

1. Ambainis, A., Beaudry, M., Golovkins, M., Kikusts, A., Mercer, M., Thérien, D.: Algebraic results on quantum automata. Th. Comp. Sys. 39, 165–188 (2006)
2. Ambainis, A., Freivalds, R.: 1-way quantum finite automata: strengths, weaknesses and generalizations. In: Proc. 39th Symp. Found. Comp. Sci., pp. 332–342 (1998)
3. Ambainis, A., Yakaryilmaz, A.: Superiority of exact quantum automata for promise problems. Information Processing Letters 112, 289–291 (2012)
4. 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)
5. Bertoni, A., Mereghetti, C., Palano, B.: Small size quantum automata recognizing some regular languages. Theoretical Computer Science 340, 394–407 (2005)
6. Bertoni, A., Mereghetti, C., Palano, B.: Some formal tools for analyzing quantum automata. Theoretical Computer Science 356, 14–25 (2006)
7. Bianchi, M.P., Palano, B.: Events and languages on unary quantum automata. Fundamenta Informaticae 104, 1–15 (2010)
8. Brodsky, A., Pippenger, N.: Characterizations of 1-way quantum finite automata. SIAM J. Comput. 5, 1456–1478 (2002)
9. Golovkins, M., Kravtsev, M.: Probabilistic reversible automata and quantum automata. In: Ibarra, O.H., Zhang, L. (eds.) COCOON 2002. LNCS, vol. 2387, pp. 574–583. Springer, Heidelberg (2002)
10. Hirvensalo, M.: Quantum automata with open time evolution. Int. J. Nat. Comp. Res. 1, 70–85 (2010)
11. Hopcroft, J.E., Motwani, R., Ullman, J.D.: Introduction to Automata Theory, Languages, and Computation. Addison-Wesley, Reading (2001)
12. Hughes, R.I.G.: The Structure and Interpretation of Quantum Mechanics. Harvard University Press, Cambridge (1992)
13. Kondacs, A., Watrous, J.: On the power of quantum finite state automata. In: Proc. 38th Annual Symposium on Foundations of Computer Science, pp. 66–75 (1997)
14. Moore, C., Crutchfield, J.: Quantum automata and quantum grammars. Theoretical Computer Science 237, 275–306 (2000)
15. Mereghetti, C., Palano, B.: Quantum finite automata with control language. Theoretical Informatics and Applications 40, 315–332 (2006)
16. Nayak, A.: Optimal lower bounds for quantum automata and random access codes. In: Proc. 40th Symposium on Foundations of Computer Science, pp. 369–376 (1999)
17. Rabin, M.O.: Probabilistic automata. Information and Control 6, 230–245 (1963)
18. Scharnhorst, K.: Angles in complex vector spaces. Act. Ap. Math. 69, 95–103 (2001)
19. Shilov, G.: Linear Algebra. Prentice Hall (1971); Reprinted by Dover (1977)
20. Yakaryilmaz, A., Cem Say, A.C.: Unbounded-error quantum computation with small space bounds. Information & Computation 209, 873–892 (2011)
21. Zheng, S., Qiu, D., Li, L., Gruska, J.: One-way finite automata with quantum and classical states. In: Bordihn, H., Kutrib, M., Truthe, B. (eds.) Languages Alive. LNCS, vol. 7300, pp. 273–290. Springer, Heidelberg (2012)