Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 182 (2019)

Lower Bounds on the Running Time of Two-Way Quantum Finite Automata and Sublogarithmic-Space Quantum Turing Machines

Zachary Remscrim
Department of Mathematics
MIT
remscrim@mit.edu

Abstract

The two-way finite automaton with quantum and classical states (2QCFA), defined by Ambainis and Watrous, is a model of quantum computation whose quantum part is extremely limited; however, as they showed, 2QCFA are surprisingly powerful: a 2QCFA with only a single-qubit can recognize the language $L_{pal} = \{w \in \{a,b\}^* : w \text{ is a palindrome}\}$ with bounded error in expected time $2^{O(n)}$, on inputs of length n.

We prove that their result essentially cannot be improved upon: a 2QCFA (of any size) cannot recognize L_{pal} with bounded error in expected time $2^{o(n)}$. To our knowledge, this is the first example of a language that can be recognized with bounded error by a 2QCFA in exponential time but not in subexponential time. Moreover, we prove that a quantum Turing machine (QTM) running in space $o(\log n)$ and expected time $2^{n^{1-\Omega(1)}}$ cannot recognize L_{pal} with bounded error; again, this is the first lower bound of its kind.

Far more generally, we establish a lower bound on the running time of any 2QCFA or $o(\log n)$-space QTM that recognizes any language L in terms of a natural "hardness measure" of L. This allows us to exhibit a large family of languages for which we have asymptotically matching lower and upper bounds on the running time of any such 2QCFA or QTM recognizer.

1 Introduction

Quantum algorithms, such as Shor's quantum polynomial time integer factorization algorithm [39], Grover's algorithm for unstructured search [15], and the linear system solver of Harrow, Hassidim, and Lloyd [16], provide examples of natural problems on which quantum computers seem to have an advantage over their classical counterparts. However, these algorithms are designed to be run on a quantum computer that has the full power of a quantum Turing machine, whereas current experimental quantum computers only possess a rather limited quantum part.

This naturally motivates the study of models of quantum computation that are far weaker than a polynomial time quantum Turing machine, such as the two-way finite automaton with quantum and classical states (2QCFA), originally defined by Ambainis and Watrous [2]. Informally, a 2QCFA is a two-way deterministic finite automaton (2DFA) that has been augmented by a quantum register of constant size; we define the 2QCFA model formally in Section 2.2. 2QCFA are surprisingly powerful, as originally demonstrated by Ambainis and Watrous, who showed that a 2QCFA, with only a single-qubit quantum register, can recognize, with bounded error, the language $L_{eq} = \{a^m b^m : m \in \mathbb{N}\}$ in expected time $O(n^4)$ and the language $L_{pal} = \{w \in \{a,b\}^* : w \text{ is a palindrome}\}$ in expected time $2^{O(n)}$. In a recent paper [33], we presented further evidence of the power of few qubits by showing that 2QCFA are capable of recognizing many group word problems with bounded error.

ISSN 1433-8092

It is known that 2QCFA are more powerful than 2DFA and two-way probabilistic finite automata (2PFA). A 2DFA can only recognize regular languages [32]. A 2PFA can recognize some nonregular languages with bounded error, given sufficient running time: in particular, a 2PFA can recognize L_{eq} with bounded error in expected time $2^{O(n)}$ [11]. However, a 2PFA cannot recognize L_{eq} with bounded error in expected time $2^{o(n)}$, by a result of Greenberg and Weiss [12]; moreover, a 2PFA cannot recognize L_{pal} with bounded error in any time bound [10]. More generally, the landmark result of Dwork and Stockmeyer [9] showed that a 2PFA cannot recognize any nonregular language in expected time $2^{n^{o(1)}}$. In order to prove this statement, they defined a particular "hardness measure" $D_L : \mathbb{N} \to \mathbb{N}$ of a language L. They showed that, if a 2PFA recognizes some language L with bounded error in expected time at most $T(n)$ on all inputs of length at most n, then there is a positive real number a (that depends only on the number of states of the 2PFA), such that $T(n) = \Omega\left(2^{D_L(n)^a}\right)$ [9, Lemma 4.3]; we will refer to this statement as the "Dwork-Stockmeyer lemma."

Very little was known about the limitations of 2QCFA. Are there any languages that a single-qubit 2QCFA can recognize with bounded error in expected exponential time but not in expected subexponential time? In particular, is it possible for a single-qubit 2QCFA to recognize L_{pal} with bounded error in expected subexponential time, or perhaps even in expected polynomial time? More generally, are there any languages that a 2QCFA (that is allowed to have a quantum register of any constant size) can recognize with bounded error in expected exponential time but not in expected subexponential time? These natural questions, to our knowledge, were all open (see, for instance, [2, 3, 46] for previous discussions of these questions).

In this paper, we answer these and other related questions. In particular, we show that a 2QCFA (of any size) cannot recognize L_{pal} with bounded error in expected time $2^{o(n)}$. Far more generally, we prove an analogue of the Dwork-Stockmeyer lemma for 2QCFA: if a 2QCFA recognizes some language L with bounded error in expected time at most $T(n)$ on all inputs of length at most n, then there a positive real number a (that depends only on the number of states of the 2QCFA), such that $T(n) = \Omega\left(D_L(n)^a\right)$. We note that, while our lower bound on the running time of a 2QCFA is exponentially weaker than the lower bound on the running time of a 2PFA provided by the Dwork-Stockmeyer lemma, both lower bounds are in fact (asymptotically) tight; the exponential difference provides yet another example of a situation in which quantum computers have an exponential advantage over their classical counterparts. We also establish a lower bound on the expected running time of a 2QCFA recognizer of L in terms of the one-way deterministic communication complexity of testing membership in L.

Furthermore, we show that the class of languages recognizable with bounded error by a 2QCFA in expected polynomial time is contained in $\mathsf{L/poly}$. This result, which shows that the class of languages recognizable by a particular quantum model is contained in the class of languages recognizable by a particular classical model, is a type of *dequantumization* result. It is (qualitatively) similar to the Adleman-type [1] *derandomization* result $\mathsf{BPL} \subseteq \mathsf{L/poly}$, where BPL denotes the class of languages recognizable with bounded error by a probabilistic Turing machine (PTM) that uses $O(\log n)$ space and runs in expected polynomial time. The only previous dequantumization result that we are aware of was of a very different type: the class of languages recognizable by a 2QCFA, or more generally a quantum Turing machine (QTM) that uses $O(\log n)$ space, with algebraic number transition amplitudes (even with unbounded error and with no time bound), is contained in $\mathsf{DSPACE}(O(\log^2 n))$ [43]. This dequantumization results is analogous to the derandomization result: the class of languages recognizable by a PTM that uses $O(\log n)$ space (even with unbounded error and with no time bound), is contained in $\mathsf{DSPACE}(O(\log^2 n))$ [6].

We then generalize our results to prove a lower bound on the expected running time $T(n)$ of a

QTM that uses sublogarithmic space (i.e., $o(\log n)$ space) and recognizes a language L with bounded error, where this lower bound is also in terms of $D_L(n)$. In particular, we show that L_{pal} cannot be recognized with bounded error by a QTM that uses sublogarithmic space and runs in expected time $2^{n^{1-\Omega(1)}}$. This result is particularly intriguing as L_{pal} can be recognized by a *deterministic* TM in $O(\log n)$ space (and, trivially, polynomial time); therefore, L_{pal} provides an example of a natural problem for which polynomial time *quantum* TMs have no (asymptotic) advantage over polynomial time *deterministic* TMs in terms of the needed amount of space.

We also investigate which group word problems can be recognized by 2QCFA, or sublogarithmic-space QTM, with particular resource bounds. Informally, the word problem of a finitely generated group is the problem of determining if the product of a sequence of elements of that group is equal to the identity element. There is a deep connection between the algebraic properties of a finitely generated group G and the complexity of its word problem W_G, as has been demonstrated by many famous results; for example, $W_G \in \mathsf{REG} \Leftrightarrow G$ is finite [4], $W_G \in \mathsf{CFL} \Leftrightarrow G$ is virtually free [8, 28], $W_G \in \mathsf{NP} \Leftrightarrow G$ is a subgroup of a finitely presented group with polynomial Dehn function [5]. We have recently shown that if G is virtually abelian, then W_G may be recognized with bounded error by a single-qubit 2QCFA in expected polynomial time, and that, for any group G in a certain broad class of groups of exponential growth, W_G may be recognized with bounded error by a 2QCFA (in many cases a single-qubit 2QCFA) in expected time $2^{O(n)}$ [33].

We now show that, if G has exponential growth, then W_G cannot be recognized by a 2QCFA with bounded error in expected time $2^{o(n)}$, thereby providing a broad and natural class of languages that may be recognized with bounded error by a 2QCFA in expected time $2^{O(n)}$ but not $2^{o(n)}$. We also show that, if W_G is recognizable by a 2QCFA with bounded error in expected polynomial time, then G must be virtually nilpotent (i.e., G must have polynomial growth), thereby obtaining progress towards an exact classification of those word problems recognizable by a 2QCFA in expected polynomial time. Furthermore, we show analogous results for sublogarithmic-space QTMs.

One of the key tools used in our proof is a quantum version of Hennie's [17] notion of a crossing sequence, which may be of independent interest. Crossing sequences played an important role in the aforementioned 2PFA results of Dwork and Stockmeyer [9] and of Greenberg and Weiss [12]. In particular, we show that the computation of a 2QCFA on a particular portion of the input string can be modeled by an operator that is, in fact, a *quantum channel*. This allows us to bring the tools of quantum information theory to bear to analyze the behavior of a 2QCFA.

The remainder of this paper is organized as follows. In Section 2, we briefly recall the fundamentals of quantum computation and the definition of 2QCFA. In Section 3, we develop our notion of a quantum crossing sequence. The Dwork-Stockmeyer hardness measure D_L of a language L, as well as several other related hardness measures of L, play a key role in our lower bounds; we recall the definitions of these hardness measures in Section 4.1. Then, in Section 4.2, using our notion of a quantum crossing sequence, we prove an analogue of the Dwork-Stockmeyer lemma for 2QCFA. Using this lemma, in Section 4.3, we establish various lower bounds on the expected running time of 2QCFA for particular languages and prove certain complexity class separations and inclusions. In Section 5, we establish lower bounds on the expected running time of sublogarithmic-space QTMs. In Section 6, we study group word problems and establish lower bounds on the expected running time of 2QCFA and sublogarithmic-space QTMs that recognize certain word problems. Finally, in Section 7, we discuss several interesting open problems related to our work.

2 Preliminaries

2.1 Quantum Computation

In this section, we briefly recall the fundamentals of quantum computation needed in this paper (see, for instance, [23, 30, 44] for a more detailed presentation of the material in this section). We begin by establishing some notation. Let V denote a finite-dimensional complex Hilbert space with inner product $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$. The dual space V^* of V is the $\mathbb{C}$-vector space consisting of all linear functionals on V (i.e., all $\mathbb{C}$-linear maps of the form $f : V \to \mathbb{C}$). We use the standard Dirac bra-ket notation throughout this paper. We denote elements of V by *kets*: $|\psi\rangle$, $|\varphi\rangle$, $|q\rangle$, etc. For the *ket* $|\psi\rangle \in V$, we define the corresponding *bra* $\langle\psi| \in V^*$ to be the linear functional on V given by $\langle |\psi\rangle, \cdot \rangle : V \to \mathbb{C}$ (i.e, for any $|\varphi\rangle \in V$, we have $\langle\psi|(|\varphi\rangle) = \langle |\psi\rangle, |\varphi\rangle\rangle$). For notational clarity and brevity, we write $\langle\psi|\varphi\rangle$ in place of $\langle\psi|(|\varphi\rangle)$.

Let $\mathrm{L}(V)$ denote the $\mathbb{C}$-vector space consisting of all $\mathbb{C}$-linear maps of the form $A : V \to V$. For $|\psi\rangle, |\varphi\rangle \in V$, we define $|\psi\rangle\langle\varphi| \in \mathrm{L}(V)$ in the natural way: for $|\rho\rangle \in V$, $|\psi\rangle\langle\varphi|(|\rho\rangle) = |\psi\rangle\langle\varphi|\rho\rangle = \langle\varphi|\rho\rangle|\psi\rangle$. For $A, A' \in \mathrm{L}(V)$ and $|\psi\rangle \in V$, we, again for the sake of notational clarify and brevity, write $A|\psi\rangle$ to denote the element $A(|\psi\rangle) \in V$ obtained by applying the map A to the element $|\psi\rangle$ and write AA' to denote the composition $A \circ A'$. Let $1_V \in \mathrm{L}(V)$ denote the identity operator on V (i.e., $1_V|\psi\rangle = |\psi\rangle$, $\forall |\psi\rangle \in V$) and let $0_V \in \mathrm{L}(V)$ denote the zero operator on V (i.e., $0_V|\psi\rangle = 0$ (the zero vector in V), $\forall|\psi\rangle \in V$). For $A \in \mathrm{L}(V)$, we define $A^\dagger \in \mathrm{L}(V)$, the *Hermitian transpose* of A, to be the unique element of $\mathrm{L}(V)$ such that $\langle A|\psi_1\rangle, |\psi_2\rangle\rangle = \langle |\psi_1\rangle, A^\dagger|\psi_2\rangle$, $\forall |\psi_1\rangle, |\psi_2\rangle \in V$. Let $\mathrm{Herm}(V) = \{A \in \mathrm{L}(V) : A = A^\dagger\}$ denote the set of *Hermitian operators* on V, let $\mathrm{Pos}(V) = \{A^\dagger A : A \in \mathrm{L}(V)\} \subseteq \mathrm{Herm}(V)$ denote the set of *positive semi-definite operators* on V, let $\mathrm{Proj}(V) = \{A \in \mathrm{Pos}(V) : A^2 = A\}$ denote the set of *projection operators* on V, let $\mathrm{U}(V) = \{A \in \mathrm{L}(V) : AA^\dagger = 1_V\}$ denote the set of *unitary operators* on V, and let $\mathrm{Den}(V) = \{A \in \mathrm{Pos}(V) : \mathrm{Tr}(A) = 1\}$ denote the set of *density operators* on V.

A *quantum register* is specified by a finite set of *quantum basis states* $Q = \{q_0, \ldots, q_{k-1}\}$. Corresponding to these k quantum basis states is an orthonormal basis $\{|q_0\rangle, \ldots, |q_{k-1}\rangle\}$ of the finite-dimensional complex Hilbert space $\mathbb{C}^k$. The quantum register stores a *superposition* $|\psi\rangle = \sum_{q \in Q} \alpha_q |q\rangle \in \mathbb{C}^k$, where each $\alpha_q \in \mathbb{C}$ and $\sum_{q \in Q} |\alpha_q|^2 = 1$; in other words, a superposition $|\psi\rangle$ is simply an element of $\mathbb{C}^k$ of norm 1. Let $\mathbb{C}^Q$ denote the $\mathbb{C}$-vector space consisting of all functions from Q to $\mathbb{C}$. Of course, $\mathbb{C}^Q \cong \mathbb{C}^k$; it will often be more convenient to think of superpositions as being elements of $\mathbb{C}^Q$ of norm 1.

A 2QCFA may only interact with its quantum register in two ways: by applying a *unitary transformation* or performing a *quantum measurement*. If the quantum register is currently in the superposition $|\psi\rangle \in \mathbb{C}^Q$, then after applying the unitary transformation $T \in \mathrm{U}(\mathbb{C}^Q)$, the quantum register will be in the superposition $T|\psi\rangle$. A *von Neumann measurement* is specified by some $P_1, \ldots, P_l \in \mathrm{Proj}(\mathbb{C}^Q)$, such that $P_i P_j = 0_{\mathbb{C}^Q}$, $\forall i, j$ with $i \neq j$, and $\sum_j P_j = 1_{\mathbb{C}^Q}$. Quantum measurement is a probabilistic process where, if the quantum register is currently in the superposition $|\psi\rangle$, then the *result* of the measurement has the value $r \in \{1, \ldots, l\}$ with probability $\|P_r|\psi\rangle\|^2$; if the result is r, then the quantum register collapses to the superposition $\frac{1}{\|P_r|\psi\rangle\|} P_r|\psi\rangle$. We emphasize that performing a quantum measurement changes the state of the quantum register. We note that all results in this paper would also follow if we allowed the more general notion of quantum measurement where we now only require that $P_1, \ldots, P_l \in \mathrm{L}(\mathbb{C}^Q)$ and $\sum_j P_j = 1_{\mathbb{C}^Q}$; see, for instance, [30, Section 2.2.3] for a more detailed discussion of the varying types of quantum measurements).

Let V and V' denote a pair of finite-dimensional complex Hilbert spaces. Let $\mathrm{T}(V, V')$ denote the $\mathbb{C}$-vector space consisting of all $\mathbb{C}$-linear maps (i.e., operators) of the form $\Phi : \mathrm{L}(V) \to \mathrm{L}(V')$. Define $\mathrm{T}(V) = \mathrm{T}(V, V)$ and let $1_{\mathrm{L}(V)} \in \mathrm{T}(V)$ denote the identity operator. Consider some $\Phi \in \mathrm{T}(V, V')$. We say that Φ is *positive* if, $\forall A \in \mathrm{Pos}(V)$, we have $\Phi(A) \in \mathrm{Pos}(V')$. We say that Φ is *completely-positive* if, for every finite-dimensional complex Hilbert space W, $\Phi \otimes 1_{\mathrm{L}(W)}$ is positive, where $\otimes$ denotes the tensor product. We say that Φ is *trace-preserving* if, $\forall A \in \mathrm{L}(V)$, we have $\mathrm{Tr}(\Phi(A)) = \mathrm{Tr}(A)$. If Φ is both completely-positive and trace-preserving, then we say Φ is a *quantum channel* (what some call a completely-positive superoperator). Let $\mathrm{Chan}(V, V') = \{\Phi \in \mathrm{T}(V, V') : \Phi \text{ is a quantum channel}\}$ denote the set of all such channels, and define $\mathrm{Chan}(V) = \mathrm{Chan}(V, V)$.

Throughout the paper, we write $\mathbb{N}_{\geq 1}$ to denote the positive natural numbers, $\mathbb{R}_{\geq 0}$ to denote the nonnegative real numbers, and so on. For $p \in \mathbb{N}_{\geq 1}$, we define the *Schatten p-norm* $\|\cdot\|_p : \mathrm{L}(V) \to \mathbb{R}_{\geq 0}$, where $\|Z\|_p = (\mathrm{Tr}((Z^\dagger Z)^{\frac{p}{2}}))^{\frac{1}{p}}$, $\forall Z \in \mathrm{L}(V)$. Observe that the Schatten p-norm is indeed a norm, for every p. We also use the term *trace norm* to refer to the Schatten 1-norm, and we note that $\|Z\|_1$ is given by the sum of the singular values of the operator $Z \in \mathrm{L}(V)$. Similarly, we use the term *Hilbert-Schmidt norm* to refer to the Schatten 2-norm, and we note that $\|Z\|_2 = \sqrt{\sum_{i,j \in B} |\langle e_i|Z|e_j\rangle|^2}$, $\forall Z \in \mathrm{L}(V)$, where $\{|e_i\rangle : i \in B\}$ is an orthonormal basis of V. We define the *induced trace norm* $\|\cdot\|_1 : \mathrm{T}(V, V') \to \mathbb{R}_{\geq 0}$, where $\|\Phi\|_1 = \sup\{\|\Phi(Z)\|_1 : Z \in \mathrm{L}(V), \|Z\|_1 \leq 1\}$, for any $\Phi \in \mathrm{T}(V, V')$. Observe that the induced trace norm is also a norm.

2.2 Definition of the 2QCFA Model

In this section, we define two-way finite automata with quantum and classical states (2QCFA), essentially following the original definition given by Ambainis and Watrous [2]. Informally, a 2QCFA is a two-way DFA that has been augmented with a quantum register of constant size; the machine may apply unitary transformations to the quantum register and perform (perhaps many) measurements of its quantum register during its computation. Formally, a 2QCFA is a 10-tuple,

$$N = (Q, C, \Sigma, \delta_{\mathrm{type}}, \delta_{\mathrm{transform}}, \delta_{\mathrm{measure}}, q_{\mathrm{start}}, c_{\mathrm{start}}, c_{\mathrm{acc}}, c_{\mathrm{rej}}),$$

where Q is a finite set of quantum basis states, C is a finite set of classical states, Σ is a finite input alphabet, $\delta_{\mathrm{type}}, \delta_{\mathrm{transform}}$, and $\delta_{\mathrm{measure}}$ collectively specify the transition function, $q_{\mathrm{start}} \in Q$ is the quantum start state, $c_{\mathrm{start}} \in C$ is the classical start state, and $c_{\mathrm{acc}}, c_{\mathrm{rej}} \in C$, with $c_{\mathrm{acc}} \neq c_{\mathrm{rej}}$, specify the classical accept and reject states, respectively. We define $\#_L, \#_R \notin \Sigma$, with $\#_L \neq \#_R$, to be special symbols that serve as a left and right end-marker, respectively; we then define the tape alphabet $\Sigma_+ = \Sigma \sqcup \{\#_L, \#_R\}$. Let $\widehat{C} = C \setminus \{c_{\mathrm{acc}}, c_{\mathrm{rej}}\}$ denote the non-halting classical states. The components of the transition function are specified as follows. Firstly, $\delta_{\mathrm{type}} : \widehat{C} \times \Sigma_+ \to \{transform, measure\}$ specifies whether N performs a unitary transformation or a quantum measurement when reading the symbol $\sigma \in \Sigma_+$ while in classic state $c \in \widehat{C}$. In the cases in which N performs a unitary transformation, $\delta_{\mathrm{transform}} : \delta_{\mathrm{type}}^{-1}(transform) \to \mathrm{U}(\mathbb{C}^Q) \times C \times \{-1, 0, 1\}$ specifies the particular transformation to be performed to the quantum register, the new classical state, and the direction in which the head is to move. If, instead, $\delta_{\mathrm{type}}(c, \sigma) = measure$, then $\delta_{\mathrm{measure}}(c, \sigma)$ is of the form $(P_1, \ldots, P_l, f)$ where $P_1, \ldots, P_l$ specifies some von Neumann measurement and $f : \{1, \ldots, l\} \to C \times \{-1, 0, 1\}$ is a function that specifies the new classical state and the direction in which the head is to move for each possible outcome of that measurement.

On an input $w = w_1 \cdots w_n \in \Sigma^*$, with each $w_i \in \Sigma$, the 2QCFA N operates as follows. The machine has a read-only tape that contains the string $\#_L w_1 \cdots w_n \#_R$. Initially, the classic state of N is c_{start}, the quantum register is in the superposition $|q_{\mathrm{start}}\rangle$, and the head is at the left end of the tape, over the left end-marker $\#_L$. On each step of the computation, if the classic state is currently $c \in \widehat{C}$ and the head is over the symbol $\sigma \in \Sigma_+$, N behaves as follows. First, suppose $\delta_{\text{type}}(c, \sigma) = \textit{transform}$ and $\delta_{\text{transform}}(c, \sigma) = (t, c', d)$, for some $t \in \mathrm{U}(\mathbb{C}^Q)$, $c' \in C$, and $d \in \{-1, 0, 1\}$; then N applies the transformation t to its quantum register, enters the classic state c', and moves its head left (resp. right) if $d = -1$ (resp. $d = 1$), keeping its head stationary if $d = 0$. If, instead, $\delta_{\text{type}}(c, \sigma) = \textit{measure}$, then if $\delta_{\text{measure}}(c, \sigma) = (P_1, \ldots, P_l, f)$, N performs the quantum measurement specified by $P_1, \ldots, P_l$, producing the result $r \in \{1, \ldots, l\}$; if $f(r) = (c', d)$, then N enters the classic state $c' \in C$ and moves its head according to $d \in \{-1, 0, 1\}$. We assume that $\delta_{\text{transform}}$ and δ_{measure} are both defined such that N will never attempt to move its head off the tape (i.e., N will never move its head left when reading $\#_L$ or right when reading $\#_R$) and that N will keep its head stationary when transitioning to either c_{acc} or c_{rej}. If, at any point in the computation, N enters the classical state c_{acc} (resp. c_{rej}), then (that branch of the computation) halts and immediately accepts (resp. rejects) its input.

Due to the fact that quantum measurement is a probabilistic process, the computation of N on an input w is probabilistic. For any language L and any $\epsilon \in [0, \frac{1}{2})$, we say that a 2QCFA N recognizes L with *two-sided bounded error* ϵ if, $\forall w \in L$, $\Pr[N \text{ accepts } w] \geq 1 - \epsilon$, and, $\forall w \notin L$, $\Pr[N \text{ accepts } w] \leq \epsilon$. Then, for any function $T : \mathbb{N} \to \mathbb{N}$, we define $\mathsf{B2QCFA}(k, d, T(n), \epsilon)$ as the class of languages L for which there is a 2QCFA, with at most k quantum basis states and at most d classical states, that recognizes L with two-sided bounded error ϵ, and has expected running time at most $T(n)$ on all inputs of length at most n.

In order to make our lower bound as strong as possible, we do *not* require N to halt with probability 1 on all $w \in \Sigma^*$ (i.e., we permit N to reject an input by looping) and we permit language recognition under the more relaxed condition of *two-sided* bounded error. The bounds that we show for this 2QCFA model of course also apply to the 2QCFA model as originally defined by Ambainis and Watrous [2], which required N to halt with probability 1 on all inputs and operated under the more restrictive *negative one-sided* bounded error recognition condition.

3 2QCFA Crossing Sequences

In this section, we develop a generalization of Hennie's [17] notion of crossing sequences to 2QCFA, in which we make use of several ideas from the 2PFA results of Dwork and Stockmeyer [9] and Greenberg and Weiss [12]. This notion will play a key role in our proof of a lower bound on the expected running time of a 2QCFA.

Consider a 2QCFA $N = (Q, C, \Sigma, \delta_{\text{type}}, \delta_{\text{transform}}, \delta_{\text{measure}}, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$. Let $\Psi = \{|\psi\rangle \in \mathbb{C}^Q : \||\psi\rangle\| = 1\}$ denote the set of possible superpositions of the quantum register of N. Consider an input $w = w_1 \cdots w_n \in \Sigma^*$, where each $w_i \in \Sigma$. When N is run on input w, the tape consists of $\#_L w_1 \cdots w_n \#_R$; for convenience, we define $w_0 = \#_L$ and $w_{n+1} = \#_R$. One may describe the total configuration of *a single probabilistic branch* of N at any particular point in time by a triple $(|\psi\rangle, c, h)$, where the quantum register is currently in the superposition $|\psi\rangle \in \Psi$, the classical state is currently $c \in C$, and the head is currently over tape cell $h \in \{0, \ldots, n+1\}$.

We partition the input as $w = xy$, where $x = w_1 \cdots w_{n'}$ and $y = w_{n'+1} \cdots w_n$ for some $n' \in \{0, \ldots, n\}$. We then imagine running N beginning in the configuration $(|\psi\rangle, c, n')$, for some $|\psi\rangle \in \Psi$ and $c \in \widehat{C} = C \setminus \{c_{\text{acc}}, c_{\text{rej}}\}$ (i.e., the head is initially over the rightmost symbol of $\#_L x$). We wish to describe the configuration (or, more accurately, ensemble of configurations) that N will be in when it "finishes computing" on the prefix $\#_L x$, either by "leaving" the string $\#_L x$ (where here we say that N "leaves" $\#_L x$ if N moves its head right when over the rightmost symbol of $\#_L x$), or by accepting or rejecting its input. Of course, N may leave $\#_L x$, then later reenter $\#_L x$, then later leave $\#_L x$ again, and so on, which will naturally lead to our notion of a crossing sequence. Note that the particular choice of the string y does not affect this subcomputation as it occurs entirely within the prefix $\#_L x$.

More generally, we consider the case in which N is run on the prefix $\#_L x$, where N starts in some ensemble of configurations $\{(p_i, (|\psi_i\rangle, c_i, n')) : i \in I\}$, where the probability of being in configuration $(|\psi_i\rangle, c_i, n')$ is given by p_i (note that the head position in each configuration is over the rightmost symbol of $\#_L x$); we call this ensemble a *starting ensemble*. We then wish to describe the ensemble of configurations that N will be in when it "finishes computing" on the prefix $\#_L x$, (essentially) as defined above; we call this ensemble a *stopping ensemble*[1]. However, we do not wish to use the large (potentially infinite) ensemble of configurations as the basis of our definition of a 2QCFA crossing sequence, as they would not be suitable for the type of analysis we wish to perform. Instead, we will describe an ensemble of configurations using density operators.

3.1 Describing Ensembles of Configurations of 2QCFA

Let $|x|$ denote the length of a string x, let $\widehat{H}_x = \{0, \ldots, |x|\}$ denote the head positions corresponding to the prefix $\#_L x$, and let $H_x = \{0, \ldots, |x| + 1\}$ denote the set of possible positions the head of N may be in when N is run on the prefix $\#_L x$ until N "finishes computing" on the prefix $\#_L x$. We now establish some notation that will allow us to (non-uniquely) describe ensembles of configurations of N.

We first consider an *ensemble of pure states* of the quantum register of N. In particular, we consider the ensemble $\{(p_i, |\psi_i\rangle) : i \in I\}$, for some index set I, where $p_i \in [0, 1]$ denotes the probability of the quantum register of N being in the superposition $|\psi_i\rangle \in \Psi$, and $\sum_i p_i = 1$. This ensemble corresponds to the density operator $A = \sum_i p_i |\psi_i\rangle\langle\psi_i| \in \mathrm{Den}(\mathbb{C}^Q)$. Of course, many distinct ensembles correspond to the density operator A; however, all ensembles that correspond to a particular density operator will behave the same, for our purposes (see, for instance, [30, Section 2.4] for a detailed discussion of this phenomenon, and of the following claims). That is to say, for any ensemble described by a density operator $A \in \mathrm{Den}(\mathbb{C}^Q)$, applying the transformation $T \in \mathrm{U}(\mathbb{C}^Q)$ produces an ensemble described by the density operator $TAT^\dagger$. Similarly, consider the von Neumann measurement specified by some $P_1, \ldots, P_l \in \mathrm{Proj}(\mathbb{C}^Q)$. Then for any ensemble described by the density operator A, the probability that the result of this measurement is r is given by $\mathrm{Tr}(P_r A P_r^\dagger)$, and if the result is r then the ensemble collapses to an ensemble described by the density operator $\frac{1}{\mathrm{Tr}(P_r A P_r^\dagger)} P_r A P_r^\dagger$. As N performs only a (classically controlled) sequence of unitary transformations and quantum measurements of its quantum register, the behavior of N is well-defined on density operators.

Remark. We note that the quantum register of any 2QCFA at any particular point in time is described by an ensemble of pure states (i.e., the quantum register is in a *mixed state*). However, the ensembles of pure states of the quantum register that we will consider when defining the crossing sequence of a 2QCFA *do not* describe the quantum register at a particular *time*; instead, the ensembles that we study consist of all the possible states of the quantum register at particular important *events* (such as the j^{th} time the head crosses the boundary between $\#_L x$ and $y\#_R$). We elaborate on this issue in Section 3.3.

We then consider an *ensemble of configurations* $\{(p_i, (|\psi_i\rangle, c_i, h_i)) : i \in I\}$, for some index set I, where $|\psi_i\rangle \in \Psi$, $c_i \in C$, and $h_i \in H_x$, $\forall i \in I$, and where the probability of N being in configuration $(|\psi_i\rangle, c_i, h_i)$ is given by p_i. Let $\widehat{i}(c, h) = \{i \in I : c_i = c \text{ and } h_i = h\}$ denote the indices of those

[1] We use the terms "starting ensemble" and "stopping ensemble" to make clear the similarity to the notion of a "starting condition" and of a "stopping condition" used by Dwork and Stockmeyer [9] in their 2PFA result.

configurations in classical state c and with head position h. We describe the ensemble by means of the pair of functions $p : C \times H_x \to [0,1]$ and $A : C \times H_x \to \mathrm{Den}(\mathbb{C}^Q)$, where $p(c,h)$ denotes the probability of the classical state being c and the head position being h, and $A(c,h)$ is a density operator that describes the ensemble of quantum register superpositions restricted to configurations in classical state c and head position h, where we assign an arbitrary value to $A(c,h)$ if there are no such configurations. To be precise, we have

$$p(c,h) = \sum_{i \in \widehat{i}(c,h)} p_i \quad \text{and} \quad A(c,h) = \begin{cases} \sum_{i \in \widehat{i}(c,h)} \frac{p_i}{p(c,h)} |\psi_i\rangle\langle\psi_i|, & \text{if } p(c,h) \neq 0 \\ |q_{\mathrm{start}}\rangle\langle q_{\mathrm{start}}|, & \text{if } p(c,h) = 0. \end{cases}$$

The 2QCFA N posseses both a constant-sized *quantum register*, that stores a superposition $|\psi\rangle \in \mathbb{C}^Q$, and a constant-sized *classical register*, that stores a classical state $c \in C$. We can naturally interpret each $c \in C$ as an element $|c\rangle \in \mathbb{C}^C$, of a special type; that is to say, each classical state c corresponds to some element $|c\rangle$ in the natural orthonormal basis of $\mathbb{C}^C$, whereas each superposition $|\psi\rangle$ of the quantum register corresponds to an element of $\mathbb{C}^Q$ of norm 1. One may also view N as possessing a *head register* that stores a (classical) head position $h \in H_x$ (when computing on the prefix $\#_L x$); of course, the size of this pseudo-register grows with the input prefix x. We analogously interpret a head position $h \in H_x$ as being the "classical" element $|h\rangle \in \mathbb{C}^{H_x}$, in the same way as we have done for the classical state $c \in C$. A configuration $(|\psi\rangle, c, h)$ of N is then simply a state of the *combined register*, which consists of the quantum, classical, and head registers; we then naturally interpret a configuration as an element of $\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}$, of a special form, in the obvious way. Let $\mathrm{Den}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ denote the set of all density operators on the combined space $\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}$. For a pair (p, A) that describes an ensemble of configurations, the element $Z = \sum_{c \in C, h \in H_x} p(c,h) A(c,h) \otimes |c\rangle\langle c| \otimes |h\rangle\langle h| \in \mathrm{Den}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ describes the same ensemble. Let $\widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ denote the set of all density operators given by some Z of the above form (i.e., those density operators that respect the fact that both the classical state and head position are classical). We write $(p, A) \leftrightarrow Z$ to denote this correspondence between a pair (p, A) that describes some ensemble and the element $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ that describes the same ensemble. We use these two types of notation interchangeably.

We also consider the case in which the head position does not need to be recorded and we are only interested in the combined state of the quantum register and classical register. We then analogously describe an ensemble $\{(p_i, (|\psi_i\rangle, c_i)) : i \in I\}$ by a pair of functions $p : C \to [0,1]$ and $A : C \to \mathrm{Den}(\mathbb{C}^Q)$, where $p(c)$ denotes the probability of the classical state being c and $A(c)$ is a density operator that describes the ensemble of quantum register superpositions restricted to configurations in classical state c. We similarly consider the set $\mathrm{Den}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ of density operators on the space $\mathbb{C}^Q \otimes \mathbb{C}^C$, and we define $\widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ to be those density operators that describe a valid ensemble of configurations.

In a starting ensemble, as defined above, all configurations have the same head position: $|x|$. We define the map $I_x : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ such that $I_x(Z) = Z \otimes ||x|\rangle\langle|x||$, $\forall Z \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. Notice that, for any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, if $\{(p_i, (|\psi_i\rangle, c_i)) : i \in I\}$ is any ensemble of states of the quantum register and classical register of N that is described by Z, then the ensemble $\{(p_i, (|\psi_i\rangle, c_i, |x|)) : i \in I\}$ of configurations of N is described by $I_x(Z) \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$.

Similarly, in a stopping ensemble, all configurations either have head position $|x| + 1$ or are accepting or rejecting configurations (in which the head position is not relevant). Let $1_{\mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)} \otimes \mathrm{Tr} : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ denote the unique element of $\mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}, \mathbb{C}^Q \otimes \mathbb{C}^C)$ such that $(1_{\mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)} \otimes \mathrm{Tr})(Z_{QC} \otimes Z_H) = \mathrm{Tr}(Z_H) Z_{QC}$, $\forall Z_{QC} \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, $\forall Z_H \in \mathrm{L}(\mathbb{C}^{H_x})$.

We call the operator $1_{\mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)} \otimes \mathrm{Tr}$ the *partial trace with respect to* $\mathbb{C}^{H_x}$ and we use $\mathrm{Tr}_{\mathbb{C}^{H_x}}$ as a shorthand notation for this operator. Notice that, for any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$, if $\{(p_i, (|\psi_i\rangle, c_i, h_i)) : i \in I\}$ is any ensemble of configurations of N described by Z, then the ensemble $\{(p_i, (|\psi_i\rangle, c_i)) : i \in I\}$ of states of the quantum register and classical register of N is described by $\mathrm{Tr}_{\mathbb{C}^{H_x}}(Z) \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$.

3.2 Overview of 2QCFA Crossing Sequences

We now sketch our definition of the *crossing sequence* of the 2QCFA N on the partitioned input xy. Consider running N on the prefix $\#_L x$ beginning in some starting ensemble $\{(p_i, (|\psi_i\rangle, c_i, |x|)) : i \in I\}$. To avoid unnecessary cases later, we also allow N to start in a configuration of the form $(|\psi\rangle, c, |x|)$, where $c \in \{c_{\mathrm{acc}}, c_{\mathrm{rej}}\}$, where we adopt the convention that in such a case N immediately leaves $\#_L x$ in the configuration $(|\psi\rangle, c, |x|+1)$. For any $m \in \mathbb{N}$, we define the *m-truncated stopping ensemble* as the ensemble of configurations (of the quantum register and classical register, we ignore the head position here) N will be in when it "finishes computing" on $\#_L x$, as defined above, with the modification that if any particular branch of N attempts to perform more than m quantum measurements, the computation of that branch will be "interrupted" immediately before it attempts to perform the $m+1^{\mathrm{st}}$ quantum measurement and instead immediately reject; we also adopt a special convention to deal with branches that, after some point, run forever without leaving $\#_L x$ or ever performing any measurements (and so are rejecting by looping), where we consider these branches to be in a configuration with classical state c_{rej}. To be clear, both the truncation of branches that perform many measurements and this convention concerning looping branches occur only in the analysis of N; we do not modify the 2QCFA N.

We then define the *m-truncated transfer operator* $N^{\hookrightarrow}_{x,m} : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ such that, for any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, if N is run on the prefix $\#_L x$ beginning in an ensemble of configurations described by $I_x(Z)$, then the m-truncated stopping ensemble will be described by $N^{\hookrightarrow}_{x,m}(Z)$. For m sufficiently large, with respect to the expected running time of N on the (total) input xy, this operator accurately describes the behavior of N when computing on the prefix $\#_L x$. This follows from the fact that, if a particular branch of N runs for s steps, that branch cannot possibly make more than s quantum measurements; therefore, interrupting branches that perform an extremely large number of quantum measurements will have a negligible impact on the behavior of N. Symmetrically, we define the operator $N^{\hookleftarrow}_{y,m} : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ that defines the behavior of N when computing on the suffix $y\#_R$. The *m-truncated crossing sequence* will then consist of the sequence of density operators obtained by beginning with the simple density operator that describes the ensemble of configurations of (a slightly modified version of) N when it first crosses between $\#_L x$ and $y\#_R$, and then alternately applying the operators $N^{\hookrightarrow}_{x,m}$ and $N^{\hookleftarrow}_{y,m}$ in an infinite sequence.

Crucially, we will observe that $N^{\hookrightarrow}_{x,m}, N^{\hookleftarrow}_{y,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, $\forall x, y \in \Sigma^*$, $\forall m \in \mathbb{N}$. This will allow us to make use of the machinery of quantum channels to analyze the behavior of a 2QCFA. In fact, the analysis that we perform on the m-truncated transfer operators, which allows us to exhibit a lower bound on the expected running time of a 2QCFA, only requires a somewhat weaker property than being a quantum channel; we prove this stronger property as these notions of transfer operators and crossing sequences may be of use in proving other properties of 2QCFA in the future.

Remark. While the m-truncated crossing operator $N^{\hookrightarrow}_{x,m}$ completely suffices for our analysis, one could also define a *non-truncated transfer operator* $N^{\hookrightarrow}_{x} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ as an accumulation point of the sequence $(N^{\hookrightarrow}_{x,m})_{m \in \mathbb{N}}$; such an accumulation point exists due to the fact that $\mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$

is compact (see, for instance, [44, Proposition 2.28]). Using $N_x^{\hookrightarrow}$ and the symmetrically defined $N_y^{\hookleftarrow}$, one could then define the *non-truncated crossing sequence* of N on xy. The resulting analyses of these two types of crossing sequences would essentially be identical, and so we do not consider this definition further here; however, the (somewhat cleaner) non-truncated crossing sequence may be more useful in other applications.

3.3 Definition and Properties of 2QCFA Crossing Sequences

We now formally define the notion of a crossing sequence of a 2QCFA, sketched in the previous section, and prove certain needed properties. We begin by considering running N on the prefix $\#_L x$ beginning in any configuration of the more general form $(|\psi\rangle, c, h)$, for some $|\psi\rangle \in \Psi$, $c \in \widehat{C}$, and $h \in \widehat{H}_x$. Note that, while this computation is a probabilistic process, it is entirely deterministic until N makes its first quantum measurement; in particular, the decision of when to perform a quantum measurement is entirely deterministic. Therefore, if we run N starting in the configuration $(|\psi\rangle, c, h)$, then eventually one of the following three disjoint events will occur: (1) N leaves $\#_L x$ before ever performing a quantum measurement, (2) N accepts or rejects its input before leaving $\#_L x$ or performing a quantum measurement, (3) N performs a quantum measurement. Recall that, by our definition of the 2QCFA model, N may not move its head when transitioning to c_{acc} or c_{rej}, and so N may not leave $\#_L x$ in the same step in which it accepts or rejects its input. Note that case (2) includes the possibility that N never leaves $\#_L x$ and never performs a quantum measurement, in which case N is looping and so N has rejected its input. We define subcases $(2)_{\mathrm{halt}}$ and $(2)_{\mathrm{loop}}$ corresponding to N halting within some finite number of steps and N running forever, respectively. Furthermore, note that the particular case that occurs depends exclusively on x, c, and h (i.e., $|\psi\rangle$ is not relevant).

We will refer to the above events (1), $(2)_{\mathrm{halt}}$, $(2)_{\mathrm{loop}}$, and (3) as *key-events*. We define $\mathrm{keyEv}_x : \widehat{C} \times \widehat{H}_x \to \{(1), (2)_{\mathrm{halt}}, (2)_{\mathrm{loop}}, (3)\}$ such that $\mathrm{keyEv}_x(c, h)$ is the first key-event that occurs when running N on prefix $\#_L x$, beginning in the configuration $(|\psi\rangle, c, h)$, for some (irrelevant) $|\psi\rangle \in \Psi$. We now define the functions $t_x : C \times H_x \to \mathrm{U}(\mathbb{C}^Q)$, $\gamma_x : C \times H_x \to C$, and $h_x : C \times H_x \to H_x$, which describe the behavior of N until the first key-event, as follows.

First, consider $c \in \widehat{C}$ and $h \in \widehat{H}_x$ such that $\mathrm{keyEv}_x(c, h) \in \{(1), (2)_{\mathrm{halt}}, (3)\}$. As noted above, the computation of N is completely deterministic before the first quantum measurement is performed, and depends only on x, c, and h. Define $\widehat{s}_{x,c,h} \in \mathbb{N}_{\geq 1}$ such that the first time that a key-event occurs is on step $\widehat{s}_{x,c,h}$ of the computation (of this single branch of N, where the first step occurs when N is in the configuration $(|\psi\rangle, c, h)$). If $\mathrm{keyEv}_x(c, h) \in \{(1), (2)_{\mathrm{halt}}\}$, let $s_{x,c,h} = \widehat{s}_{x,c,h}$, if $\mathrm{keyEv}_x(c, h) = (3)$, let $s_{x,c,h} = \widehat{s}_{x,c,h} - 1$. We define $t_x(c, h)$, $\gamma_x(c, h)$, and $h_x(c, h)$, such that, immediately after performing step $s_{x,c,h}$, N is in the single configuration $(t_x(c, h)|\psi\rangle, \gamma_x(c, h), h_x(c, h))$. Note that if (1) or $(2)_{\mathrm{halt}}$ occurs, then $(t_x(c, h)|\psi\rangle, \gamma_x(c, h), h_x(c, h))$ is the configuration of N *immediately after* the step in which the key-event occurs, and if (3) occurs, then $(t_x(c, h)|\psi\rangle, \gamma_x(c, h), h_x(c, h))$ is the configuration of N *immediately before* the first key-event occurs. To be precise, for $i \in \{1, \ldots, s_{x,c,h}\}$, let $T_{x,c,h,i} \in \mathrm{U}(\mathbb{C}^Q)$ denote the unitary transformation that N applies to its quantum register on the i^{th} step. Let $t_x(c, h) = T_{x,c,h,s_{x,c,h}} \circ \cdots \circ T_{x,c,h,1} \in \mathrm{U}(\mathbb{C}^Q)$ denote the total unitary transformation applied to the quantum register (recall that we apply transformations on the left), let $\gamma_x(c, h) \in C$ denote the classical state that N enters on step $s_{x,c,h}$, and let $h_x(c, h) \in H_x$ be the position the head of N moves to on step $s_{x,c,h}$.

Next, consider $c \in \widehat{C}$ and $h \in \widehat{H}_x$ such that $\mathrm{keyEv}_x(c, h) = (2)_{\mathrm{loop}}$. In this case, we have a branch of the computation of N that runs forever without ever leaving $\#_L x$ or performing a quantum measurement. As such a branch corresponds to the case in which N is rejecting its input by looping, we will simply consider such a branch to be in the classical state c_{rej}, to avoid unnecessary cases in our analysis later. In particular, we define $t_x(c,h) = 1_{\mathbb{C}^Q}$ (the identity map), $\gamma_x(c,h) = c_{\text{rej}}$, and $h_x(c,h) = h$. Of course, we are not modifying the machine N such that these branches halt; this convention is used only in our analysis of N.

Notice that, if N is run on the prefix $\#_L x$ beginning in the single configuration $(|\psi\rangle, c, h)$, for some $|\psi\rangle \in \Psi$, $c \in \widehat{C}$, and $h \in \widehat{H}_x$, then when the first key-event occurs (with the conventions stated above), N will be in the single configuration $(t_x(c,h)|\psi\rangle, \gamma_x(c,h), h_x(c,h))$, which satisfies the following properties. If $\text{keyEv}_x(c,h) = (1)$, then N has just left $\#_L x$ for the first time; in particular, $h_x(c,h) = |x| + 1$ (i.e., the head is one cell to the right of the rightmost symbol of $\#_L x$). If $\text{keyEv}_x(c,h) = (2)_{\text{halt}}$, then N has just halted, accepting or rejecting the input (on this branch); in particular, $\gamma_x(c,h) \in \{c_{\text{acc}}, c_{\text{rej}}\}$. If $\text{keyEv}_x(c,h) = (2)_{\text{loop}}$, then N is rejecting its input by looping (on this branch); in particular, $\gamma_x(c,h) = c_{\text{rej}}$. If $\text{keyEv}_x(c,h) = (3)$, then $h_x(c,h) \in \widehat{H}_x$ and $\delta_{\text{type}}(\gamma_x(c,h), x_{h_x(c,h)}) = \textit{measure}$; in particular, N will perform a quantum measurement of its quantum register at step $\widehat{s}_{x,c,h} = s_{x,c,h} + 1$, after having performed exclusively unitary transformations of its quantum register within the first $s_{x,c,h}$ steps. In particular, if $\text{keyEv}_x(c,h) \in \{(1), (2)_{\text{halt}}, (2)_{\text{loop}}\}$, then the ensemble of configurations that N is in when it "finishes computing" on the prefix $\#_L x$ (where N begins in the single configuration $(|\psi\rangle, c, h)$) is given by the single configuration $(t_x(c,h)|\psi\rangle, \gamma_x(c,h), h_x(c,h))$. Of course, if $\text{keyEv}_x(c,h) = (3)$, then N will perform a quantum measurement on its next step, after which point N will be in an ensemble of configurations. After completing our definition and analysis of t_x, γ_x, and h_x, we will subsequently define functions that describe the behavior of N when it performs a quantum measurement; this will ultimately allow us to describe the m-truncated stopping ensemble.

We have, so far, defined $t_x(c,h)$, $\gamma_x(c,h)$, and $h_x(c,h)$, $\forall c \in \widehat{C}$, $\forall h \in \widehat{H}_x$. For any other pair (c,h) (i.e., if $c \in \{c_{\text{acc}}, c_{\text{rej}}\}$ or $h = |x| + 1$), we define $t_x(c,h) = 1_{\mathbb{C}^Q}$, $\gamma_x(c,h) = c$, and $h_x(c,h) = h$. That is to say, we define these functions such that they leave configurations $(|\psi\rangle, c, h)$, with $c \in \{c_{\text{acc}}, c_{\text{rej}}\}$ or $h = |x| + 1$ unchanged; we do this as we want to group together the different branches of the computation of N when each branch "finishes computing" on $\#_L x$ for the first time. This will be explained more fully when we formally define crossing sequences. This completes our definition of the functions t_x, γ_x, and h_x, which describe the behavior of N until the first key event.

Let $\{(p_i, (|\psi_i\rangle, c_i, h_i)) : i \in I\}$ be any ensemble of configurations where $|\psi_i\rangle \in \Psi$, $c_i \in C$, and $h_i \in H_x$, $\forall i \in I$. We define the *ensemble of configurations at the next key-event* to be the ensemble $\{(p_i, (t_x(c_i,h_i)|\psi_i\rangle, \gamma_x(c_i,h_i), h_x(c_i,h_i))) : i \in I\}$. In other words, for each i with $c_i \in \widehat{C}$ and $h_i \in \widehat{H}_x$, we replace the configuration $(|\psi_i\rangle, c_i, h_i)$ by the configuration $(t_x(c_i,h_i)|\psi_i\rangle, \gamma_x(c_i,h_i), h_x(c_i,h_i))$ that N is in when the first key-event occurs, with the above conventions; for any other i, we leave the configuration unchanged. We now define an operator K_x that encapsulates the above computation in a useful way. In particular, consider any $Z \in \widehat{\text{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ and let $\{(p_i, (|\psi_i\rangle, c_i, h_i)) : i \in I\}$ be any ensemble of configurations described by Z. We define K_x such that the ensemble of configurations at the next key-event is described by $K_x(Z)$.

Definition 3.1. Consider a 2QCFA $N = (Q, C, \Sigma, \delta_{\text{type}}, \delta_{\text{transform}}, \delta_{\text{measure}}, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$ and input prefix $x \in \Sigma^*$. Define the functions t_x, γ_x, and h_x as above. For each $c \in C$ and each $h \in H_x$, let $E_{x,c,h} = t_x(c,h) \otimes |\gamma_x(c,h)\rangle\langle c| \otimes |h_x(c,h)\rangle\langle h| \in \text{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$. We then define the operator $K_x : \text{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \text{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ such that $K_x(Z) = \sum_{c \in C, h \in H_x} E_{x,c,h} Z E_{x,c,h}^\dagger$, $\forall Z \in \text{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$.

We next observe that K_x operates as described on density operators, and that K_x is a quantum channel.

Lemma 3.2. *Using the notation of Definition 3.1, the following statements hold.*

(i) For any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$*, if* $\{(p_i, (|\psi_i\rangle, c_i, h_i)) : i \in I\}$ *is any ensemble of configurations described by* Z*, then the ensemble of configurations at the next key-event is described by* $K_x(Z)$*.*

(ii) We have $K_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$*.*

Proof. (i) Any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ is of the form $Z = \sum_{\widehat{c} \in C, \widehat{h} \in H_x} p(\widehat{c}, \widehat{h}) A(\widehat{c}, \widehat{h}) \otimes |\widehat{c}\rangle\langle\widehat{c}| \otimes |\widehat{h}\rangle\langle\widehat{h}|$, for some $p : C \times H_x \to [0,1]$ and $A : C \times H_x \to \mathrm{Den}(\mathbb{C}^Q)$. We then have

$$K_x(Z) = \sum_{\substack{c \in C \\ h \in H_x}} E_{x,c,h} \Bigg(\sum_{\substack{\widehat{c} \in C \\ \widehat{h} \in H_x}} p(\widehat{c}, \widehat{h}) A(\widehat{c}, \widehat{h}) \otimes |\widehat{c}\rangle\langle\widehat{c}| \otimes |\widehat{h}\rangle\langle\widehat{h}| \Bigg) E^{\dagger}_{x,c,h}$$

$$= \sum_{\substack{c, \widehat{c} \in C \\ h, \widehat{h} \in H_x}} p(\widehat{c}, \widehat{h}) t_x(c,h) A(\widehat{c}, \widehat{h}) t_x(c,h)^{\dagger} \otimes |\gamma_x(c,h)\rangle\langle c|\widehat{c}\rangle\langle\widehat{c}|c\rangle\langle\gamma_x(c,h)| \otimes |h_x(c,h)\rangle\langle h|\widehat{h}\rangle\langle\widehat{h}|h\rangle\langle h_x(c,h)|$$

$$= \sum_{\substack{c \in C \\ h \in H_x}} p(c,h) t_x(c,h) A(c,h) t_x(c,h)^{\dagger} \otimes |\gamma_x(c,h)\rangle\langle\gamma_x(c,h)| \otimes |h_x(c,h)\rangle\langle h_x(c,h)|.$$

As noted previously, if the unitary transformation $T \in \mathrm{U}(\mathbb{C}^Q)$ is applied to any ensemble of superpositions of the quantum register described by some density operator $A \in \mathrm{Den}(\mathbb{C}^Q)$, the result is an ensemble described by the density operator $TAT^{\dagger}$. The claim is then immediate from definitions.

(ii) The family $\{E_{x,c,h} : c \in C, h \in H_x\}$ is a *Kraus representation* of the operator K_x (see, for instance, [44, Section 2.2] for a formal definition). It is straightforward to see that $K_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ if an only if $\sum_{c \in C, h \in H_x} E^{\dagger}_{x,c,h} E_{x,c,h} = 1_{\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}}$ (see, for instance, [44, Corollary 2.27]). For any $c \in C$ and $h \in H_x$, we have

$$E^{\dagger}_{x,c,h} E_{x,c,h} = \Big(t_x(c,h)^{\dagger} \otimes |c\rangle\langle\gamma_x(c,h)| \otimes |h\rangle\langle h_x(c,h)| \Big) \Big(t_x(c,h) \otimes |\gamma_x(c,h)\rangle\langle c| \otimes |h_x(c,h)\rangle\langle h| \Big)$$

$$= t_x(c,h)^{\dagger} t_x(c,h) \otimes |c\rangle\langle\gamma_x(c,h)|\gamma_x(c,h)\rangle\langle c| \otimes |h\rangle\langle h_x(c,h)|h_x(c,h)\rangle\langle h|$$

$$= 1_{\mathbb{C}^Q} \otimes |c\rangle\langle c| \otimes |h\rangle\langle h|.$$

Therefore,

$$\sum_{\substack{c \in C \\ h \in H_x}} E^{\dagger}_{x,c,h} E_{x,c,h} = \sum_{\substack{c \in C \\ h \in H_x}} 1_{\mathbb{C}^Q} \otimes |c\rangle\langle c| \otimes |h\rangle\langle h| = 1_{\mathbb{C}^Q} \otimes 1_{\mathbb{C}^C} \otimes 1_{\mathbb{C}^{H_x}} = 1_{\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}}.$$

□

We next consider the behavior of N when it performs a quantum measurement. Suppose N is in the configuration $(|\psi\rangle, c, h)$, for some $|\psi\rangle \in \Psi$, $c \in \widehat{C}$, and $h \in \widehat{H}_x$, where $\delta_{\mathrm{type}}(c, x_h) = \mathit{measure}$ (i.e., N will perform a quantum measurement on the next step of its computation). Define $P_{x,c,h,1}, \ldots, P_{x,c,h,l_{x,c,h}} \in \mathrm{Proj}(\mathbb{C}^Q)$, $R_{x,c,h} = \{1, \ldots, l_{x,c,h}\}$, and the function $f_{x,c,h} : R_{x,c,h} \to C \times \{-1, 0, 1\}$ such that $\delta_{\mathrm{measure}}(c, x_h) = (P_{x,c,h,1}, \ldots, P_{x,c,h,l_{x,c,h}}, f_{x,c,h})$. For each $r \in R_{x,c,h}$, define $\widetilde{\gamma}_x(c,h,r) \in C$ and $d_{x,c,h,r} \in \{-1,0,1\}$ such that $f_{x,c,h}(r) = (\widetilde{\gamma}_x(c,h,r), d_{x,c,h,r})$ and define $\widetilde{h}_x(c,h,r) = h + d_{x,c,h,r}$. The outcome of the measurement is $r \in R_{x,c,h}$ with probability $\|P_{x,c,h,r}|\psi\rangle\|^2$; if the outcome is r, then the quantum register of N collapses to the superposition $\frac{1}{\|P_{x,c,h,r}|\psi\rangle\|} P_{x,c,h,r}|\psi\rangle$. Therefore, after performing the above measurement, N is in the ensemble $\left\{\left(\|P_{x,c,h,r}|\psi\rangle\|^2, \left(\frac{1}{\|P_{x,c,h,r}|\psi\rangle\|} P_{x,c,h,r}|\psi\rangle, \widetilde{\gamma}_x(c,h,r), \widetilde{h}_x(c,h,r)\right)\right) : r \in R_{x,c,h}, \|P_{x,c,h,r}|\psi\rangle\| \neq 0\right\}$.

We have made the above definitions of $R_{x,c,h}$, $\widetilde{\gamma}_x(c,h,r)$, etc., for all cases in which N performs a quantum measurement on the next step of its computation while N is computing within the prefix $\#_L x$ (i.e., when $c \in \widehat{C}$, $h \in \widehat{H}_x$, and $\delta_{\text{type}}(c, x_h) = \textit{measure}$). Otherwise, we define $R_{x,c,h} = \{1\}$, $P_{x,c,h,1} = 1_{\mathbb{C}^Q}$, $\widetilde{\gamma}_x(c,h,1) = c$, and $\widetilde{h}_x(c,h,1) = h$; this will assure that all other configurations are left unchanged (again, we do this as we want to group together the different branches of the computation of N when each branch "finishes computing" on $\#_L x$ for the first time). We now define an operator M_x that performs at most one quantum measurement.

Definition 3.3. Consider a 2QCFA $N = (Q, C, \Sigma, \delta_{\text{type}}, \delta_{\text{transform}}, \delta_{\text{measure}}, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$ and input prefix $x \in \Sigma^*$. Using the above notation, for each $c \in C$, $h \in H_x$, and $r \in R_{x,c,h}$, let $\widetilde{E}_{x,c,h,r} = P_{x,c,h,r} \otimes |\widetilde{\gamma}_x(c,h,r)\rangle\langle c| \otimes |\widetilde{h}_x(c,h,r)\rangle\langle h| \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$. We then define the operator $M_x : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ such that $M_x(Z) = \sum_{c \in C, h \in H_x} \sum_{r \in R_{x,c,h}} \widetilde{E}_{x,c,h,r} Z \widetilde{E}^\dagger_{x,c,h,r}$, $\forall Z \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$.

Lemma 3.4. *Using the notation of Definition 3.3, the following statements hold.*

(i) For any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$, if $\{(p_i, (|\psi_i\rangle, c_i, h_i)) : i \in I\}$ is any ensemble of configurations described by Z, then $M_x(Z)$ describes an ensemble of configurations for which each configuration with $c_i \in \widehat{C}$, $h_i \in \widehat{H}_x$, and $\delta_{type}(c_i, x_{h_i}) = measure$ is replaced by the ensemble of configurations obtained by performing a single quantum measurement and all other configurations are left unchanged.

(ii) We have $M_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$.

Proof. (i) This follows immediately from the fact that quantum measurement is well defined on density operators. For the sake of completeness, we now exhibit the straightforward proof.

Any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ is of the form $Z = \sum_{\widehat{c} \in C, \widehat{h} \in H_x} p(\widehat{c}, \widehat{h}) A(\widehat{c}, \widehat{h}) \otimes |\widehat{c}\rangle\langle\widehat{c}| \otimes |\widehat{h}\rangle\langle\widehat{h}|$, for some $p : C \times H_x \to [0,1]$ and $A : C \times H_x \to \mathrm{Den}(\mathbb{C}^Q)$. For each $\widehat{c} \in C$ and each $\widehat{h} \in H_x$, let $Z_{\widehat{c},\widehat{h}} = A(\widehat{c}, \widehat{h}) \otimes |\widehat{c}\rangle\langle\widehat{c}| \otimes |\widehat{h}\rangle\langle\widehat{h}| \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$, and for each $r \in R_{x,c,h}$, let $D_{\widetilde{\gamma}_x(\widehat{c},\widehat{h},r)} = |\widetilde{\gamma}_x(\widehat{c}, \widehat{h}, r)\rangle\langle\widetilde{\gamma}_x(\widehat{c}, \widehat{h}, r)| \in \widehat{\mathrm{Den}}(\mathbb{C}^C)$ and let $D_{\widetilde{h}_x(\widehat{c},\widehat{h},r)} = |\widetilde{h}_x(\widehat{c}, \widehat{h}, r)\rangle\langle\widetilde{h}_x(\widehat{c}, \widehat{h}, r)| \in \widehat{\mathrm{Den}}(\mathbb{C}^{H_x})$.

First, suppose $\widehat{c} \in \widehat{C}$, $\widehat{h} \in \widehat{H}_x$, and $\delta_{\text{type}}(\widehat{c}, x_{\widehat{h}}) = \textit{measure}$. If N is in an ensemble of configurations described by $Z_{\widehat{c},\widehat{h}}$, then all configurations in that ensemble are in classic state $\widehat{c}$ and have head position $\widehat{h}$, $A(\widehat{c}, \widehat{h})$ describes the ensemble of superpositions of the quantum register, and N will perform the same quantum measurement in its next computational step on all configurations in the ensemble. As noted earlier, when performing this quantum measurement, the probability of outcome $r \in R_{x,\widehat{c},\widehat{h}}$ is given by $\mathrm{Tr}\left(P_{x,\widehat{c},\widehat{h},r} A(\widehat{c}, \widehat{h}) P^\dagger_{x,\widehat{c},\widehat{h},r}\right)$; if the outcome is r, the ensemble of configurations of the quantum register will collapse to an ensemble described by $\frac{1}{\mathrm{Tr}\left(P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^\dagger_{x,\widehat{c},\widehat{h},r}\right)} P_{x,\widehat{c},\widehat{h},r} A(\widehat{c}, \widehat{h}) P^\dagger_{x,\widehat{c},\widehat{h},r}$.

Let $\widetilde{R}_{x,\widehat{c},\widehat{h},A(\widehat{c},\widehat{h})} = \left\{r \in R_{x,\widehat{c},\widehat{h}} : \operatorname{Tr}\left(P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r}\right) \neq 0\right\}$ denote those measurement outcomes that occur with non-zero probability. Note that $P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r} \in \operatorname{Den}(\mathbb{C}^Q) \subseteq \operatorname{Pos}(\mathbb{C}^Q)$, and so all eigenvalues of the operator $P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r}$ are non-negative real numbers. If we have $\operatorname{Tr}\left(P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r}\right) = 0$, then the operator $P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r}$ has only the eigenvalue 0 (with multiplicity $|Q|$), which then implies that $P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r} = 0_{\mathbb{C}^Q}$ (as $P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r} \in \operatorname{Pos}(\mathbb{C}^Q)$). Therefore, after performing the above quantum measurement, N is in an ensemble of configurations described by the density operator $Z'_{\widehat{c},\widehat{h}}$, where

$$\begin{aligned}
Z'_{\widehat{c},\widehat{h}} &= \sum_{r \in \widetilde{R}_{x,\widehat{c},\widehat{h},A(\widehat{c},\widehat{h})}} \frac{\operatorname{Tr}\left(P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r}\right)}{\operatorname{Tr}\left(P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r}\right)} P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r} \otimes D_{\widetilde{\gamma}_x(\widehat{c},\widehat{h},r)} \otimes D_{\widetilde{h}_x(\widehat{c},\widehat{h},r)} \\
&= \sum_{r \in \widetilde{R}_{x,\widehat{c},\widehat{h},A(\widehat{c},\widehat{h})}} P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r} \otimes D_{\widetilde{\gamma}_x(\widehat{c},\widehat{h},r)} \otimes D_{\widetilde{h}_x(\widehat{c},\widehat{h},r)} \\
&= \sum_{r \in R_{x,\widehat{c},\widehat{h}}} P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r} \otimes D_{\widetilde{\gamma}_x(\widehat{c},\widehat{h},r)} \otimes D_{\widetilde{h}_x(\widehat{c},\widehat{h},r)}.
\end{aligned}$$

Next, suppose instead that it is *not* the case that $\widehat{c} \in \widehat{C}$, $\widehat{h} \in \widehat{H}_x$, and $\delta_{\text{type}}(\widehat{c}, x_{\widehat{h}}) = \textit{measure}$. We then define $Z'_{\widehat{c},\widehat{h}} = \sum_{r \in R_{x,\widehat{c},\widehat{h}}} P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r} \otimes D_{\widetilde{\gamma}_x(\widehat{c},\widehat{h},r)} \otimes D_{\widetilde{h}_x(\widehat{c},\widehat{h},r)}$, as in the previous case. Note that, for $\widehat{c}$ and $\widehat{h}$ of this form, we have $Z'_{\widehat{c},\widehat{h}} = Z_{\widehat{c},\widehat{h}}$.

By the above, after performing quantum measurements for all appropriate configurations (i.e., for all configurations on which N will perform a quantum measurement in its next computational step), and leaving all other configurations unchanged, N will be an ensemble of configurations described by Z', where

$$Z' = \sum_{\substack{\widehat{c} \in C \\ \widehat{h} \in H_x}} p(\widehat{c},\widehat{h}) Z_{\widehat{c},\widehat{h}} = \sum_{\substack{\widehat{c} \in C \\ \widehat{h} \in H_x}} \sum_{r \in R_{x,\widehat{c},\widehat{h}}} p(\widehat{c},\widehat{h}) P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r} \otimes D_{\widetilde{\gamma}_x(\widehat{c},\widehat{h},r)} \otimes D_{\widetilde{h}_x(\widehat{c},\widehat{h},r)}.$$

Let $F(\widehat{c},\widehat{h}) = p(\widehat{c},\widehat{h}) P_{x,\widehat{c},\widehat{h},r} A(\widehat{c},\widehat{h}) P^{\dagger}_{x,\widehat{c},\widehat{h},r}$. We then have

$$\begin{aligned}
M_x(Z) &= \sum_{\substack{c \in C \\ h \in H_x \\ r \in R_{x,c,h}}} \widetilde{E}_{x,c,h,r} Z \widetilde{E}^{\dagger}_{x,c,h,r} \\
&= \sum_{\substack{c \in C \\ h \in H_x \\ r \in R_{x,c,h}}} \widetilde{E}_{x,c,h,r} \left(\sum_{\substack{\widehat{c} \in C \\ \widehat{h} \in H_x}} p(\widehat{c},\widehat{h}) A(\widehat{c},\widehat{h}) \otimes |\widehat{c}\rangle\langle\widehat{c}| \otimes |\widehat{h}\rangle\langle\widehat{h}| \right) \widetilde{E}^{\dagger}_{x,c,h,r} \\
&= \sum_{\substack{c,\widehat{c} \in C \\ h,\widehat{h} \in H_x \\ r \in R_{x,c,h}}} F(\widehat{c},\widehat{h}) \otimes |\widetilde{\gamma}_x(\widehat{c},\widehat{h},r)\rangle\langle c|\widehat{c}\rangle\langle\widehat{c}|c\rangle\langle\widetilde{\gamma}_x(\widehat{c},\widehat{h},r)| \otimes |\widetilde{h}_x(\widehat{c},\widehat{h},r)\rangle\langle h|\widehat{h}\rangle\langle\widehat{h}|h\rangle\langle\widetilde{h}_x(\widehat{c},\widehat{h},r)|
\end{aligned}$$

$$= \sum_{\substack{\widehat{c}\in C\\ \widehat{h}\in H_x\\ r\in R_{x,\widehat{c},\widehat{h}}}} F(\widehat{c},\widehat{h}) \otimes |\widetilde{\gamma}_x(\widehat{c},\widehat{h},r)\rangle\langle\widetilde{\gamma}_x(\widehat{c},\widehat{h},r)| \otimes |\widetilde{h}_x(c',h',r)\rangle\langle\widetilde{h}_x(c',h',r)| = Z'.$$

(ii) We proceed analogously to the proof of Lemma 3.2(ii). For any $c \in C$, $h \in H$, and $r \in R_{x,c,h}$, recall that $P_{x,c,h,r} \in \mathrm{Proj}(\mathbb{C}^Q)$, which implies $P^{\dagger}_{x,c,h,r}P_{x,c,h,r} = P_{x,c,h,r}P_{x,c,h,r} = P_{x,c,h,r}$; we then have

$$\begin{aligned}\widetilde{E}^{\dagger}_{x,c,h,r}\widetilde{E}_{x,c,h,r} &= (P^{\dagger}_{x,c,h,r}\otimes|c\rangle\langle\widetilde{\gamma}_x(c,h,r)|\otimes|h\rangle\langle\widetilde{h}_x(c,h,r)|)(P_{x,c,h,r}\otimes|\widetilde{\gamma}_x(c,h,r)\rangle\langle c|\otimes|\widetilde{h}_x(c,h,r)\rangle\langle h|)\\ &= P^{\dagger}_{x,c,h,r}P_{x,c,h,r} \otimes |c\rangle\langle\widetilde{\gamma}_x(c,h,r)|\widetilde{\gamma}_x(c,h,r)\rangle\langle c| \otimes |h\rangle\langle\widetilde{h}_x(c,h,r)|\widetilde{h}_x(c,h,r)\rangle\langle h|\\ &= P_{x,c,h,r} \otimes |c\rangle\langle c| \otimes |h\rangle\langle h|.\end{aligned}$$

As $\{P_{x,c,h,r} : r \in R_{x,c,h}\}$ specifies a quantum measurement, we have $\sum_{r\in R_{x,c,h}} P_{x,c,h,r} = 1_{\mathbb{C}^Q}$, which implies

$$\sum_{\substack{c\in C\\ h\in H_x\\ r\in R_{x,c,h}}} \widetilde{E}^{\dagger}_{x,c,h,r}\widetilde{E}_{x,c,h,r} = \sum_{\substack{c\in C\\ h\in H_x\\ r\in R_{x,c,h}}} P_{x,c,h,r}\otimes|c\rangle\langle c|\otimes|h\rangle\langle h| = \sum_{\substack{c\in C\\ h\in H_x}} 1_{\mathbb{C}^Q}\otimes|c\rangle\langle c|\otimes|h\rangle\langle h| = 1_{\mathbb{C}^Q\otimes\mathbb{C}^C\otimes\mathbb{C}^{H_x}}.$$

By [44, Corollary 2.27], $M_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$.

□

We next define a truncation operator T_x.

Definition 3.5. Consider a 2QCFA $N = (Q, C, \Sigma, \delta_{\text{type}}, \delta_{\text{transform}}, \delta_{\text{measure}}, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$ and input prefix $x \in \Sigma^*$. For each $c \in C$ and each $h \in H_x$, we define $\widehat{E}_{x,c,h} \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ as follows. If $c \in \widehat{C}$ and $h \in \widehat{H}_x$, then $\widehat{E}_{x,c,h} = 1_{\mathbb{C}^Q} \otimes |c_{\text{rej}}\rangle\langle c| \otimes |h\rangle\langle h|$, otherwise, $\widehat{E}_{x,c,h} = 1_{\mathbb{C}^Q} \otimes |c\rangle\langle c| \otimes |h\rangle\langle h|$. We then define the operator $T_x : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ such that $T_x(Z) = \sum_{c\in C, h\in H_x} \widehat{E}_{x,c,h}Z\widehat{E}^{\dagger}_{x,c,h}$, $\forall Z \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$.

Lemma 3.6. *Using the notation of Definition 3.5, the following statements hold.*

(i) For any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$, if $\{(p_i, (|\psi_i\rangle, c_i, h_i)) : i \in I\}$ is any ensemble of configurations described by Z, then $T_x(Z)$ describes an ensemble of configurations for which each configuration with both $c_i \in \widehat{C}$ and $h_i \in \widehat{H}_x$ is replaced by the configuration $(|\psi_i\rangle, c_{rej}, h_i)$ and all other configurations are left unchanged. In other words, all configurations that correspond to the case in which N has "finished computing" on $\#_L x$ are left unchanged, and all other configurations become rejecting configurations.

(ii) We have $T_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$.

Proof. (i) Immediate from definitions.

(ii) As in the proof of Lemma 3.2(ii), we may straightforwardly show $\sum_{c\in C,h\in H_x} \widehat{E}^{\dagger}_{x,c,h}\widehat{E}_{x,c,h} = 1_{\mathbb{C}^Q\otimes\mathbb{C}^C\otimes\mathbb{C}^{H_x}}$, which implies $T_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ [44, Corollary 2.27].

□

We now formally define the notion of a *m-truncated transfer operator* and of a *m-truncated crossing sequence*. Firstly, given a 2QCFA N, we produce an equivalent N' of a certain convenient form, in much the same way that Dwork and Stockmeyer [9] converted a 2PFA to an equivalent 2PFA of a convenient form. The 2QCFA N' is identical to N, except for the addition of two new classical states: c'_{start} and c', where c'_{start} will be the classical start state of N'. On any input w, N' will move its head continuously to the right until it reaches $\#_R$, remaining in state c'_{start} and performing the trivial transformation to its quantum register along the way. When the head reaches $\#_R$, N' will enter c' and perform the trivial transformation to its quantum register; then, N' will move its head continuously to the left until it reaches $\#_L$, remaining in state c' and performing the trivial transformation to its quantum register along the way. When the head reaches $\#_L$, N' will enter the original classical start state c_{start} and perform the trivial transformation to its quantum register. After this point, N' behaves identically to N. For the remainder of the paper, we assume all 2QCFA under consideration have this form.

Definition 3.7. Consider a 2QCFA $N = (Q, C, \Sigma, \delta_{\text{type}}, \delta_{\text{transform}}, \delta_{\text{measure}}, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$.

(i) For any $x \in \Sigma^*$, define $I_x : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ and $\mathrm{Tr}_{\mathbb{C}^{H_x}} : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ as in Section 3.1, define $K_x, M_x, T_x : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ as above. For each $m \in \mathbb{N}$, we define the *m-truncated transfer operator* $N^{\hookrightarrow}_{x,m} : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ by $N^{\hookrightarrow}_{x,m} = \mathrm{Tr}_{\mathbb{C}^{H_x}} \circ T_x \circ (K_x \circ M_x)^m \circ K_x \circ I_x$.

(ii) For any $y \in \Sigma^*$, we next consider the "dual case" of running N on the suffix $y\#_R$ beginning in some ensemble of configurations $\{(p_i, (|\psi_i\rangle, c_i, |x| + 1)) : i \in I\}$ (i.e., the head position of every configuration is over the leftmost symbol of $y\#_R$). We define the notion of an m-truncated stopping ensemble, and all other notions, symmetrically. That is to say, a branch of N "finishes computing" on $y\#_R$ when it either "leaves" $y\#_R$ (by moving its head left from the leftmost symbol of $y\#_R$), or accepts or rejects the input, or attempts to perform $m+1$ quantum measurements. We then define $N^{\hookleftarrow}_{y,m} : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ as the corresponding "dual" m-truncated transfer operator for y.

(iii) For any $x, y \in \Sigma^*$ and any $m \in \mathbb{N}$, we then define the *m-truncated crossing sequence* of N with respect to the (partitioned) input xy to be the sequence $Z_1, Z_2, \ldots \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, defined as follows. The density operator Z_1 describes the ensemble consisting of the single configuration (of the quantum register and classical register) $(|q_{\text{start}}\rangle, c_{\text{start}})$ that N is in when it first crosses from $\#_L x$ into $y\#_R$, which is of this simple form due to the assumed form of N. The sequence $Z_1, Z_2, \ldots$ is then obtained by starting with Z_1 and alternately applying $N^{\hookleftarrow}_{y,m}$ and $N^{\hookrightarrow}_{x,m}$. To be precise,

$$Z_i = \begin{cases} |q_{\text{start}}\rangle\langle q_{\text{start}}| \otimes |c_{\text{start}}\rangle\langle c_{\text{start}}|, & i = 1 \\ N^{\hookleftarrow}_{y,m}(Z_{i-1}), & i > 1, i \text{ is even} \\ N^{\hookrightarrow}_{x,m}(Z_{i-1}), & i > 1, i \text{ is odd.} \end{cases}$$

Lemma 3.8. *Using the notation of Definition 3.7, the following statements hold.*

(i) For any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, if N is run on the prefix $\#_L x$ beginning in any ensemble of configurations described by $I_x(Z)$ (i.e., the head position of every configuration is over the rightmost symbol of $\#_L x$), then the m-truncated stopping ensemble is described by $N^{\hookrightarrow}_{x,m}(Z)$.

(ii) *Symmetrically, for any* $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$*, if* N *is run on the suffix* $y\#_R$ *beginning in any ensemble of configurations described by* $\widetilde{I}_y(Z)$ *(i.e., the head position of every configuration is over the leftmost symbol of* $y\#_R$*), then the* m*-truncated stopping ensemble is described by* $N^{\hookleftarrow}_{y,m}(Z)$*.*

(iii) *We have* $N^{\hookrightarrow}_{x,m}, N^{\hookleftarrow}_{y,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$*,* $\forall x, y \in \Sigma^*$*,* $\forall m \in \mathbb{N}$*.*

Proof. (i) For any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, let $\{(p_i, (|\psi_i\rangle, c_i, |x|)) : i \in I\}$ be any ensemble of configurations described by $I_x(Z)$. By Lemma 3.2(i), $\{(p_i, (t_x(c_i, |x|)|\psi_i\rangle, \gamma_x(c_i, |x|), h_x(c_i, |x|))) : i \in I\}$, the ensemble of configurations at the first key-event, is described by $K_x(I_x(Z))$. For any $i \in I$ such that $c_i \in \{c_{\mathrm{acc}}, c_{\mathrm{rej}}\}$ or $\mathrm{keyEv}_x(c_i, |x|) \in \{(1), (2)_{\mathrm{halt}}, (2)_{\mathrm{loop}}\}$, the configuration $(t_x(c_i, |x|)|\psi_i\rangle, \gamma_x(c_i, |x|), h_x(c_i, |x|)))$ is one on which N has "finished computing" on $\#_L x$. For any other $i \in I$ (i.e., $c_i \in \widehat{C}$ and $\mathrm{keyEv}_x(c_i, |x|) = (3)$), the configuration $(t_x(c_i, |x|)|\psi_i\rangle, \gamma_x(c_i, |x|), h_x(c_i, |x|)))$ is one on which N will perform a quantum measurement in the next step of its computation.

First, suppose $m = 0$. Then terminating these configurations on which N is about to perform a quantum measurement (by replacing the classic state of each such configuration by c_{rej}), would yield an ensemble of configurations that, after ignoring the head position, is the 0-truncated stopping ensemble. By Lemma 3.6(i), we then conclude that $\mathrm{Tr}_{\mathbb{C}^{H_x}}(T_x(K_x(I_x(Z)))) = N^{\hookrightarrow}_{x,0}(Z)$ describes the 0-truncated stopping ensemble, as desired.

Next, suppose $m > 0$. Let $\{(p'_i, (|\psi'_i\rangle, c'_i, h'_i)) : i \in I'\}$ denote the ensemble of configurations obtained from $\{(p_i, (t_x(c_i, |x|)|\psi_i\rangle, \gamma_x(c_i, |x|), h_x(c_i, |x|))) : i \in I\}$ by performing a single quantum measurement on appropriate configurations (i.e., for each $i \in I$ such that $\gamma_x(c_i, |x|) \in \widehat{C}$, $h_x(c_i, |x|) \in \widehat{H}_x$, and $\delta_{\mathrm{type}}(\gamma_x(c_i, |x|), x_{h_x(c_i,|x|)}) = \mathit{measure}$, we replace the configuration $(t_x(c_i, |x|)|\psi_i\rangle, \gamma_x(c_i, |x|), h_x(c_i, |x|))$ by the ensemble of configurations that result from applying a single quantum measurement) and leaving all other configurations unchanged. By Lemma 3.4(i), $M_x(K_x(I_x(Z))$ describes the ensemble $\{(p'_i, (|\psi'_i\rangle, c'_i, h'_i)) : i \in I'\}$. By another application of Lemma 3.2(i), $K_x(M_x(K_x(I_x(Z)))$ describes the ensemble $\{(p'_i, (t_x(c'_i, h'_i)|\psi'_i\rangle, \gamma_x(c'_i, h'_i), h_x(c'_i, h'_i))) : i \in I'\}$ obtained by running N on the ensemble $\{(p'_i, (|\psi'_i\rangle, c'_i, h'_i)) : i \in I'\}$ until the next key-event occurs (where configurations on which N has already "finished computing" on $\#_L x$ (by having accepted or rejected the input, or by having left $\#_L x$ once) are left unchanged).

If $m = 1$, then, as argued above, terminating all those configurations in the ensemble $\{(p'_i, (t_x(c'_i, h'_i)|\psi'_i\rangle, \gamma_x(c'_i, h'_i), h_x(c'_i, h'_i))) : i \in I'\}$ on which N is about to perform a quantum measurement, would yield an ensemble of configurations that, after ignoring the head position, is the 1-truncated stopping ensemble. By Lemma 3.6(i), we then conclude that $\mathrm{Tr}_{\mathbb{C}^{H_x}}(T_x(K_x(M_x(K_x(I_x(Z)))))) = N^{\hookrightarrow}_{x,0}(Z)$ describes the 1-truncated stopping ensemble, as desired. If $m > 1$, then by continuing in this fashion, we conclude that $N^{\hookrightarrow}_{x,m}(Z)$ describes the m-truncated stopping ensemble, as desired.

(ii) Immediate by Definition 3.7(ii), and analogous versions of Lemma 3.2(i), Lemma 3.4(i), and Lemma 3.6(i).

(iii) By Definition 3.7(i), $N^{\hookrightarrow}_{x,m} = \mathrm{Tr}_{\mathbb{C}^{H_x}} \circ T_x \circ (K_x \circ M_x)^m \circ K_x \circ I_x$. By Lemma 3.2(ii), Lemma 3.4(ii), and Lemma 3.6(ii), we have $K_x, M_x, T_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$. It is straightforward to see that $I_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C, \mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ and $\mathrm{Tr}_{\mathbb{C}^{H_x}} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}, \mathbb{C}^Q \otimes \mathbb{C}^C)$ and that the composition of quantum channels is a quantum channel (see, for instance, [44, Section 2.2]).The claim for $N^{\hookleftarrow}_{y,m}$ follows by an analogous argument.

□

Note that the $\{Z_i\}$ that comprise a crossing sequence do *not* describe the ensemble of configurations of N at particular points in time during its computation on the input xy; instead, Z_i describes the ensemble of configurations of the set of all the probabilistic branches of N at the i^{th} time each branch crosses between $\#_L x$ and $y\#_R$ (with the convention stated above of considering a branch that has accepting or rejected its input to "cross" in classic state c_{acc} or c_{rej}, respectively, indefinitely; as well as the convention that if a given branch of N attempts to perform more than m quantum measurements within the prefix $\#_L x$ or within the suffix $y\#_R$, that branch is interrupted and immediately forced to reject). Of course, a given branch may not cross between $\#_L x$ and $y\#_R$ more than i times within the first i steps of the computation, nor may a given branch perform more than i quantum measurements within i steps of computation; this will allow us to use such crossing sequences to prove a lower bound on the expected running-time of N.

4 Lower Bounds on the Running Time of 2QCFA

Dwork and Stockmeyer proved a lower bound [9, Lemma 4.3] on the expected running time $T(n)$ of any 2PFA that recognizes any language L with bounded error, where the lower bound is in terms of their hardness measure $D_L(n)$. We prove that an analogous claim holds for any 2QCFA. The preceding quantum generalization of a crossing sequence plays a key role in the proof, essentially taking the place of the Markov chains used both in the aforementioned result of Dwork and Stockmeyer and in the earlier result of Greenberg and Weiss [12] that showed that a 2PFA cannot recognize $L_{eq} = \{a^m b^m : m \in \mathbb{N}\}$ with bounded error in subexponential time.

4.1 Nonregularity, Automaticity, and Similar Hardness Measures

For any language L, Dwork and Stockmeyer [9] defined a particular "hardness measure" $D_L : \mathbb{N} \to \mathbb{N}$, which they called the *nonregularity* of L. We begin by recalling this definition. Let Σ be a finite alphabet, $L \subseteq \Sigma^*$ a language, and $n \in \mathbb{N}$. For a string $w \in \Sigma^*$, we use $|w|$ to denote its length. Let $\Sigma^{\leq n} = \{w \in \Sigma^* : |w| \leq n\}$ denote the set of all strings over Σ of length at most n and consider some $x, x' \in \Sigma^{\leq n}$. We say that x and x' are (L, n)-*dissimilar*, which we denote by writing $x \not\sim_{L,n} x'$, if $\exists y \in \Sigma^{\leq n'}$, where $n' = n - \max(|x|, |x'|)$, such that $xy \in L \Leftrightarrow x'y \notin L$. Recall the classic Myhill-Nerode inequivalence relation, in which $x, x' \in \Sigma^*$ are L-dissimilar if $\exists y \in \Sigma^*$, such that $xy \in L \Leftrightarrow x'y \notin L$. Then $x, x' \in \Sigma^{\leq n}$ are (L, n)-dissimilar precisely when they are L-dissimilar, and the dissimilarity is witnessed by a "short" string y. We then define the function $D_L : \mathbb{N} \to \mathbb{N}$ such that $D_L(n)$ is the largest $h \in \mathbb{N}$ such that $\exists x_1, \ldots, x_h \in \Sigma^{\leq n}$ that are pairwise (L, n)-dissimilar (i.e., $\forall i, j$ with $i \neq j$, $x_i \not\sim_{L,n} x_j$).

In fact, the hardness measure D_L of a language L has been defined by many authors, both before and after Dwork and Stockmeyer, who gave many different names to D_L and who (repeatedly) rediscovered certain basic facts about D_L; we refer the reader to the excellent paper of Shallit and Breitbart [36] for a detailed history of the study of D_L and related hardness measures. In the remainder of this section, we briefly recall two crucial equivalent definitions of D_L, as well as the definition of a certain related (inequivalent) hardness measure, which we will need in order to prove our various lower bounds in their full generality.

For some DFA (one-way deterministic finite automaton) M, let $|M|$ denote the number of states of M and let $L(M)$ denote the language of M (i.e., the set of strings accepted by M). The earliest definition of a hardness measure equivalent to Dwork-Stockmeyer nonregularity was given by Karp [22], who defined $A_L(n) = \min\{|M| : M \text{ is a DFA and } L(M) \cap \Sigma^{\leq n} = L \cap \Sigma^{\leq n}\}$ to be the minimum number of states of a DFA that agrees with L on all strings of length at most n; Shallit and Breitbart use the term *deterministic automaticity* to refer to A_L. It is immediately obvious that $A_L(n) \geq D_L(n), \forall n$; somewhat less obviously, $A_L(n) = D_L(n), \forall n$ [21, 22, 36], and so the notions of nonregularity and deterministic automaticity coincide.

Consider a language $L \subseteq \Sigma^*$ and two communicating parties: Alice, who knows some string $x \in \Sigma^*$, and Bob, who knows some string $y \in \Sigma^*$. Alice sends some message $A(x) \in \{0,1\}^*$ to Bob, after which Bob must be able to determine, using $A(x)$ and y, if the string $w = xy$ is in L. Let $C_L(n)$ denote the maximum, taken over all $x, y \in \Sigma^*$ such that $|xy| \leq n$, of the number of bits sent from Alice to Bob by the optimal such (deterministic one-way) protocol. This quantity, the *one-way deterministic communication complexity* of testing membership in L, is related to the nonregularity of L; in particular, $C_L(n) = \log D_L(n), \forall n$ [7].

Lastly, we recall the definition of a related (but inequivalent) hardness measure used by Ibarra and Ravikumar [20] in their study of non-uniform small-space DTMs (deterministic Turing machines). Let $\Sigma^n = \{w \in \Sigma^* : |w| = n\}$. We then consider 2DFA (two-way deterministic finite automata), and use the same notation as was used above for DFA. For a language L, define $A^{2DFA}_{L,=}(n) = \min\{|M| : M \text{ is a 2DFA and } L(M) \cap \Sigma^n = L \cap \Sigma^n\}$ to be the minimum number of states of a 2DFA that agrees with L on all strings of length exactly n. Clearly, for any language L, $A^{2DFA}_{L,=}(n) \leq A_L(n)$, $\forall n$. They then defined $\mathsf{NUDSPACE}(O(S(n)))$ (non-uniform deterministic space $O(S(n))$) to be the class of languages L such that $A^{2DFA}_{L,=}(n) = 2^{O(S(n))}$. Note that $\mathsf{NUDSPACE}(O(S(n))) = \mathsf{DSPACE}(O(S(n)))/2^{O(S(n))}$, the class of languages recognizable by a DTM that, on any input w, uses space $O(S(|w|))$, and has access to an "advice" string $y_{|w|}$, which depends only on the length $|w|$ of the input and is itself of length $|y_n| = 2^{O(S(n))}$. In particular, $\mathsf{L/poly} := \mathsf{DSPACE}(O(\log n))/2^{O(\log n)} = \mathsf{NUDSPACE}(O(\log n)) = \{L : A^{2DFA}_{L,=}(n) = n^{O(1)}\}$.

4.2 A 2QCFA Analogue of the Dwork-Stockmeyer Lemma

In the Dwork and Stockmeyer [9] lower bound on the expected running time of any 2PFA that recognizes a language L, the function D_L played an important role, as, intuitively, D_L measures the number of strings which must be distinguished, in a certain sense, by any 2PFA that recognizes L. This function also plays an important role in our result, as we shall now demonstrate that an analogous statement holds for 2QCFA.

The main idea is as follows. Consider a 2QCFA N, with quantum basis states Q an classical states C, that recognizes $L \subseteq \Sigma^*$, with two-sided bounded error $\epsilon \in \mathbb{R}_{>0}$, in expected time at most $T(n)$ on all inputs of length at most n. For any $n \in \mathbb{N}$, consider $x, x' \in \Sigma^{\leq n}$ such that $x \not\sim_{L,n} x'$. By definition, $\exists y \in \Sigma^{\leq n'}$, where $n' = n - \max(|x|, |x'|)$, such that $xy \in L \Leftrightarrow x'y \notin L$; note that $xy, x'y \in \Sigma^{\leq n}$. Without loss of generality, we assume $xy \in L$, and hence $x'y \notin L$. We consider running N on the partitioned input xy as well as on the partitioned input $x'y$. For $m \in \mathbb{N}$, we define the m-truncated transfer operators $N^{\hookrightarrow}_{x,m}$, $N^{\hookrightarrow}_{x',m}$, and $N^{\hookleftarrow}_{y,m}$ as in Definition 3.7. By Lemma 3.8(iii), $N^{\hookrightarrow}_{x,m}, N^{\hookrightarrow}_{x',m}, N^{\hookleftarrow}_{y,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. We define a distance metric on $\mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. We show that, if $D_L(n)$ is "large", then, for any m, we can find $x, x' \in \Sigma^{\leq n}$ such that $x \not\sim_{L,n} x'$ and the distance between $N^{\hookrightarrow}_{x,m}$ and $N^{\hookrightarrow}_{x',m}$ is "small." We also show that, for m sufficiently large, if the distance between $N^{\hookrightarrow}_{x,m}$ and $N^{\hookrightarrow}_{x',m}$ is "small," then the behavior of N on the inputs xy and $x'y$ will be similar; in particular, if $T(n)$ is "small" compared to a suitable function of $D_L(n)$, then $\Pr[N \text{ accepts } xy] \approx \Pr[N \text{ accepts } x'y]$. However, as $xy \in L$, we must have $\Pr[N \text{ accepts } xy] \geq 1-\epsilon$, and as $x'y \notin L$, we must have $\Pr[N \text{ accepts } x'y] \leq \epsilon$, which is impossible. This contradiction allows us to establish a lower bound on $T(n)$ in terms of $D_L(n)$. In this section, we formalize this idea.

For density operators $Z, Z' \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, we use $\|Z - Z'\|_1$, the distance metric induced by the trace norm, to measure the distance between Z and Z'. For $x, x' \in \Sigma^*$ and $m \in \mathbb{N}$, we use $\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1$, the distance metric induced by the induced trace norm, to measure the distance between $N^{\hookrightarrow}_{x,m}$ and $N^{\hookrightarrow}_{x',m}$. Suppose N is run on two distinct partitioned inputs xy and $x'y$, producing two distinct m-truncated crossing sequences, following Definition 3.7(iii). We first show that if $\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1$ is "small", then these crossing sequences are similar.

Lemma 4.1. *Consider a 2QCFA N with quantum basis states Q, classical states C, and input alphabet Σ. For $x, x', y \in \Sigma^*$ and $m \in \mathbb{N}$, let $Z_1, Z_2, \ldots \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ (resp. $Z'_1, Z'_2, \ldots \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$) denote the m-truncated crossing sequence obtained when N is run on xy (resp. $x'y$). Then $\|Z_i - Z'_i\|_1 \le \lfloor \frac{i-1}{2} \rfloor \|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1$, $\forall i \in \mathbb{N}_{\ge 1}$.*

Proof. Note that $\|\Phi(Z)\|_1 \le \|Z\|_1$, $\forall Z \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, $\forall \Phi \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ (see, for instance, [44, Corollary 3.40]). Therefore, for any $\Phi \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ and any $Z, Z' \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, we have

$$\|\Phi(Z) - \Phi(Z')\|_1 = \|\Phi(Z - Z')\|_1 \le \|Z - Z'\|_1.$$

That is to say, the distance metric on $\mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ induced by the trace norm is *contractive* under any map $\Phi \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. By Lemma 3.8(iii), $N^{\hookrightarrow}_{x,m}, N^{\hookrightarrow}_{x',m}, N^{\hookleftarrow}_{y,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$.

By definition, $Z_1 = |q_{\mathrm{start}}\rangle\langle q_{\mathrm{start}}| \otimes |c_{\mathrm{start}}\rangle\langle c_{\mathrm{start}}| = Z'_1$, and so $\|Z_1 - Z'_1\|_1 = 0$. For i even, we have, by definition, $Z_i = N^{\hookleftarrow}_{y,m}(Z_{i-1})$ and $Z'_i = N^{\hookleftarrow}_{y,m}(Z'_{i-1})$. By the above observation concerning the contractivity of the trace norm, we then have

$$\|Z_i - Z'_i\|_1 = \|N^{\hookleftarrow}_{y,m}(Z_{i-1}) - N^{\hookleftarrow}_{y,m}(Z'_{i-1})\|_1 \le \|Z_{i-1} - Z'_{i-1}\|_1, \quad \text{if } i \text{ is even.}$$

For i odd, with $i > 1$, we have, by definition $Z_i = N^{\hookrightarrow}_{x,m}(Z_{i-1})$ and $Z'_i = N^{\hookrightarrow}_{x',m}(Z'_{i-1})$. Note that, for any $Z \in \mathrm{Den}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, we have $\|Z\|_1 = 1$, which implies $\|\Phi(Z)\|_1 \le \|\Phi\|_1$, $\forall \Phi \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C)$; of course, $N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m} \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. By this observation and the earlier observation concerning the contractivity of the trace norm, we have

$$\|Z_i - Z'_i\|_1 = \|N^{\hookrightarrow}_{x,m}(Z_{i-1}) - N^{\hookrightarrow}_{x',m}(Z'_{i-1})\|_1 \le \|N^{\hookrightarrow}_{x,m}(Z_{i-1}) - N^{\hookrightarrow}_{x,m}(Z'_{i-1})\|_1 + \|N^{\hookrightarrow}_{x,m}(Z'_{i-1}) - N^{\hookrightarrow}_{x',m}(Z'_{i-1})\|_1$$

$$\le \|N^{\hookrightarrow}_{x,m}(Z_{i-1} - Z'_{i-1})\|_1 + \|(N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m})(Z'_{i-1})\|_1 \le \|Z_{i-1} - Z'_{i-1}\|_1 + \|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1, \ \text{ if } i \text{ odd}, i > 1.$$

The claim then follows by induction on $i \in \mathbb{N}_{\ge 1}$. □

Lemma 4.2. *Consider a language L over some finite alphabet Σ. Suppose $L \in \mathsf{B2QCFA}(k, d, T(n), \epsilon)$, for some $k, d \in \mathbb{N}_{\ge 2}$, $T : \mathbb{N} \to \mathbb{N}$, and $\epsilon \in [0, \frac{1}{2})$. If, for some $n \in \mathbb{N}$, $\exists x, x' \in \Sigma^{\le n}$ such that $x \not\sim_{L,n} x'$, then $T(n) \ge \frac{(1-2\epsilon)^2}{2} \|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1^{-1}$, $\forall m \ge \lceil \frac{2}{1-2\epsilon} T(n) \rceil$.*

Proof. By definition, $x \not\sim_{L,n} x'$ precisely when $\exists y \in \Sigma^*$ such that $xy, x'y \in \Sigma^{\le n}$, and $xy \in L \Leftrightarrow x'y \notin L$. Fix such a y, and assume, without loss of generality, that $xy \in L$ (and hence $x'y \notin L$). For $m \in \mathbb{N}$, suppose that, when N is run on the partitioned input xy (resp. $x'y$), we obtain the m-truncated crossing sequence $Z_{m,1}, Z_{m,2}, \ldots \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ (resp. $Z'_{m,1}, Z'_{m,2}, \ldots \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$). For $s \in \mathbb{N}_{\ge 1}$, define $p_{m,s}, p'_{m,s} : C \to [0, 1]$ and $A_{m,s}, A'_{m,s} : C \to \mathrm{Den}(\mathbb{C}^Q)$ such that $Z_{m,s} \leftrightarrow (p_{m,s}, A_{m,s})$ and $Z'_{m,s} \leftrightarrow (p'_{m,s}, A'_{m,s})$. For $c \in C$, let $E_c = 1_{\mathbb{C}^Q} \otimes |c\rangle\langle c| \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. Notice that $p_{m,s}(c) = \mathrm{Tr}(E_c Z_{m,s} E_c^\dagger)$ and $p'_{m,s}(c) = \mathrm{Tr}(E_c Z'_{m,s} E_c^\dagger)$. Therefore,

$$|p_{m,s}(c) - p'_{m,s}(c)| = |\mathrm{Tr}(E_c Z_{m,s} E_c^\dagger) - \mathrm{Tr}(E_c Z'_{m,s} E_c^\dagger)| = |\mathrm{Tr}(E_c (Z_{m,s} - Z'_{m,s}) E_c^\dagger)| \le \|Z_{m,s} - Z'_{m,s}\|_1.$$

By Lemma 4.1, $\|Z_{m,s} - Z'_{m,s}\|_1 \leq \frac{s-1}{2}\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1$, $\forall s \in \mathbb{N}_{\geq 1}$, and so we conclude

$$|p_{m,s}(c) - p'_{m,s}(c)| \leq \frac{s-1}{2}\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1.$$

Notice that $p_{m,s}(c_{\mathrm{acc}})$ (resp. $p'_{m,s}(c_{\mathrm{acc}})$) is the probability that N accepts xy (resp. $x'y$) within the first s times (on a given branch of the computation) the head of N crosses the boundary between x (resp. x') and y, where any branch that attempts to perform more than m quantum measurements between consecutive boundary crossings is forced to halt and reject immediately before attempting to perform the $m+1^{\mathrm{st}}$ such quantum measurement. Let $p_N(w)$ denote the probability that N accepts an input $w \in \Sigma^*$, let $p_N(w, s)$ denote the probability that N accepts w within s steps, and let $h_N(w, s)$ denote the probability that N halts on input w within s steps.

Due to the fact that $x'y \notin L$, we must have $p_N(x'y) \leq \epsilon$. Clearly, $p'_{m,s}(c_{\mathrm{acc}}) \leq p_N(x'y)$, for any m and s, as all branches that attempt to perform more than m quantum measurements (between consecutive crossings) are considered to reject the input in the m-truncated crossing sequence. Suppose $s \leq m$. Notice that any branch that runs for a total of at most s steps before halting cannot possibly perform more than s quantum measurements (and so certainly cannot perform more than s quantum measurements between consecutive crossings between $\#_L x$ and $y\#_R$); therefore, such a branch is unaffected by m-truncation. Moreover, if a branch halts (and accepts) within s steps, it will certainly halt (and accept) within s crossings between $\#_L x$ and $y\#_R$. This implies $p_N(xy, s) \leq p_{m,s}(c_{\mathrm{acc}})$, if $s \leq m$. Therefore, if $s \leq m$, we have

$$p_N(xy,s) \leq p_{m,s}(c_{\mathrm{acc}}) \leq p'_{m,s}(c_{\mathrm{acc}}) + |p_{m,s}(c_{\mathrm{acc}}) - p'_{m,s}(c_{\mathrm{acc}})| \leq \epsilon + \frac{s-1}{2}\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1.$$

By definition, the expected running time of N on input xy is at most $T(|xy|)$; therefore, by Markov's inequality, $1 - h_N(xy, s) \leq \frac{T(|xy|)}{s}$. Due to the fact that $xy \in L$, we must have $p_N(xy) \geq 1 - \epsilon$. Therefore, for any $s, m \in \mathbb{N}_{\geq 1}$ where $s \leq m$, we have

$$1 - \epsilon \leq p_N(xy) \leq p_N(xy, s) + (1 - h_N(xy, s)) \leq \epsilon + \frac{s-1}{2}\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1 + \frac{T(|xy|)}{s}.$$

Set $s = \lceil \frac{2}{1-2\epsilon}T(n) \rceil$, and notice that $|xy| \leq n$ implies $T(|xy|) \leq T(n)$. For any $m \geq s$, we then have

$$1 - 2\epsilon \leq \frac{\lceil \frac{2}{1-2\epsilon}T(n)\rceil - 1}{2}\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1 + \frac{T(|xy|)}{\lceil \frac{2}{1-2\epsilon}T(n)\rceil} \leq \frac{T(n)}{1-2\epsilon}\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1 + \frac{1-2\epsilon}{2}.$$

Therefore,

$$T(n) \geq \frac{(1-2\epsilon)^2}{2}\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1^{-1}, \quad \forall m \geq \left\lceil \frac{2}{1-2\epsilon}T(n) \right\rceil$$

□

The following lemma shows that any "large" set of input prefixes contains a pair of input prefixes x, x' such that $N^{\hookrightarrow}_{x,m}$ and $N^{\hookrightarrow}_{x',m}$ are at "small" distance from one another, for all $m \in \mathbb{N}$.

Lemma 4.3. *Consider a 2QCFA N with quantum basis states Q, classical states C, and input alphabet Σ; let $k = |Q|$ and $d = |C|$. Further, consider any finite $X \subseteq \Sigma^*$ such that $|X| \geq 2$. Then $\forall m \in \mathbb{N}$, $\exists x, x' \in X$ such that $x \neq x'$ and $\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1 \leq 4\sqrt{2}k^4 d^3 \left(|X|^{\frac{1}{k^4 d^2}} - 1\right)^{-1}$.*

Proof. For $q, q' \in Q$ and $c, c' \in C$, let $F_{q,q',c,c'} = |q\rangle\langle q'| \otimes |c\rangle\langle c'| \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. Let $J : \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^Q \otimes \mathbb{C}^C)$ denote the *Choi isomorphism*, which is given by

$$J(\Phi) = \sum_{\substack{q,q' \in Q \\ c,c' \in C}} F_{q,q',c,c'} \otimes \Phi(F_{q,q',c,c'}), \quad \forall \Phi \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C).$$

Consider any $x \in \Sigma^*$ and $m \in \mathbb{N}$. We next observe that $J(N^{\hookrightarrow}_{x,m})$ is of a special form. We encode $J(N^{\hookrightarrow}_{x,m})$ as a $(k^2d^2) \times (k^2d^2)$ matrix $M(J(N^{\hookrightarrow}_{x,m}))$ in the natural way. The set of rows and the set of columns of $M(J(N^{\hookrightarrow}_{x,m}))$ are each indexed by $Q \times C \times Q \times C$. For a matrix M, $M[i,j]$ denotes the entry in row i and column j. For $q_1, q_2, q'_1, q'_2 \in Q$ and $c_1, c_2, c'_1, c'_2 \in C$, we define $M(J(N^{\hookrightarrow}_{x,m}))[(q_1, c_1, q_2, c_2), (q'_1, c'_1, q'_2, c'_2)] = \langle q_2c_2|N^{\hookrightarrow}_{x,m}(F_{q_1,q'_1,c_1,c'_1})|q'_2c'_2\rangle$ (where $\langle q_2c_2|$ denotes $(\langle q_2| \otimes \langle c_2|)$, etc.).

We first observe that, if $c_1 \neq c'_1$ or if $c_2 \neq c'_2$, then $M(J(N^{\hookrightarrow}_{x,m}))[(q_1, c_1, q_2, c_2), (q'_1, c'_1, q'_2, c'_2)] = 0$. To see this, recall that, by Definition 3.7(i), $N^{\hookrightarrow}_{x,m} = \mathrm{Tr}_{\mathbb{C}^{H_x}} \circ T_x \circ (K_x \circ M_x)^m \circ K_x \circ I_x$. If $c_1 \neq c'_1$, then, by inspection, $K_x(I_x(F_{q_1,q'_1,c_1,c'_1})) = 0_{\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}}$, which implies $N^{\hookrightarrow}_{x,m}(F_{q_1,q'_1,c_1,c'_1}) = 0_{\mathbb{C}^Q \otimes \mathbb{C}^C}$, which then implies $M(J(N^{\hookrightarrow}_{x,m}))[(q_1, c_1, q_2, c_2), (q'_1, c'_1, q'_2, c'_2)] = 0$. If $c_2 \neq c'_2$, then $\forall Z \in L(\mathbb{C}^Q \otimes \mathbb{C}^C)$, $\langle q_2c_2| \mathrm{Tr}_{\mathbb{C}^{H_x}}(T_x(Z))|q'_2c'_2\rangle = 0$, which implies $M(J(N^{\hookrightarrow}_{x,m}))[(q_1, c_1, q_2, c_2), (q'_1, c'_1, q'_2, c'_2)] = 0$.

By Lemma 3.8(iii), $N^{\hookrightarrow}_{x,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, which implies $J(N^{\hookrightarrow}_{x,m}) \in \mathrm{Pos}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^Q \otimes \mathbb{C}^C)$ [44, Corollary 2.27]. Therefore, for any $q_1, q_2, q'_1, q'_2 \in Q$ and for any $c_1, c_2, c'_1, c'_2 \in C$, we have $M(J(N^{\hookrightarrow}_{x,m}))[(q_1, c_1, q_2, c_2), (q'_1, c'_1, q'_2, c'_2)] = \overline{M(J(N^{\hookrightarrow}_{x,m}))[(q'_1, c'_1, q'_2, c'_2), (q_1, c_1, q_2, c_2)]}$; moreover, $M(J(N^{\hookrightarrow}_{x,m}))[(q_1, c_1, q_2, c_2), (q_1, c_1, q_2, c_2)] \in \mathbb{R}$.

Therefore, for any 2QCFA N of the assumed type, any $m \in \mathbb{N}$, and any $x \in \Sigma^*$, $M(J(N^{\hookrightarrow}_{x,m}))$ is only potentially non-zero at the k^4d^2 entries of the form $M(J(N^{\hookrightarrow}_{x,m}))[(q_1, c_1, q_2, c_2), (q'_1, c_1, q'_2, c_2)]$, and each entry below the main diagonal of $M(J(N^{\hookrightarrow}_{x,m}))$ is the complex conjugate of a particular entry above the main diagonal. We define the function $g_{N,m} : \Sigma^* \to \mathbb{R}^{k^4d^2}$ such that, $\forall x \in \Sigma^*$, $g_{N,m}(x)$ encodes all the potentially non-zero entries of $M(J(N^{\hookrightarrow}_{x,m}))$ on or above the main diagonal; to be precise, the first k^2d^2 entries are given by the entries on the main diagonal, $M(J(N^{\hookrightarrow}_{x,m}))[(q_1, c_1, q_2, c_2), (q_1, c_1, q_2, c_2)]$, for $q_1, q_2 \in Q$ and $c_1, c_2 \in C$, all of which are real numbers, and whose remaining $k^4d^2 - k^2d^2$ entries are given by encoding each of the $\frac{1}{2}(k^4d^2 - k^2d^2)$ potentially non-zero entries of $M(J(N^{\hookrightarrow}_{x,m}))$ that lie above the main diagonal as the pair of real numbers that comprise their real and imaginary parts.

Let $h = k^4d^2$. In the following, in addition to the notation for the norms of operators and of quantum channels defined in Section 2.1, we write $\|\cdot\| : \mathbb{R}^h \to \mathbb{R}_{\geq 0}$ to denote the usual Euclidean 2-norm on $\mathbb{R}^h$. It is straightforward to show that $\|\Phi\|_1 \leq \|J(\Phi)\|_1$, $\forall \Phi \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ (see, for instance [44, Section 3.4] or [29, Remark 4]). Therefore, for any $x, x' \in \Sigma^*$, we then have

$$\begin{aligned} \|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1 &\leq \|J(N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m})\|_1 = \|J(N^{\hookrightarrow}_{x,m}) - J(N^{\hookrightarrow}_{x',m})\|_1 \\ &\leq k^2d^2\|J(N^{\hookrightarrow}_{x,m}) - J(N^{\hookrightarrow}_{x',m})\|_2 \leq \sqrt{2}k^2d^2\|g_{N,m}(x) - g_{N,m}(x')\|. \end{aligned}$$

Note that $N^{\hookrightarrow}_{x,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ implies $\|N^{\hookrightarrow}_{x,m}\|_1 = 1$ [44, Corollary 3.40]. Then, $\forall q, q' \in Q, \forall c \in C$, we have $\|F_{q,q',c,c}\|_1 = 1$, which implies $\|N^{\hookrightarrow}_{x,m}(F_{q,q',c,c})\|_1 \leq 1$. Therefore, $\forall x \in \Sigma^*$, we have

$$\|g_{N,m}(x)\| \leq \|J(N^{\hookrightarrow}_{x,m})\|_2 \leq \|J(N^{\hookrightarrow}_{x,m})\|_1 \leq \sum_{q,q' \in Q, c \in C} \|N^{\hookrightarrow}_{x,m}(F_{q,q',c,c})\|_1 \leq \sum_{q,q' \in Q, c \in C} 1 = k^2d = \sqrt{h}.$$

To complete the proof, for $v_0 \in \mathbb{R}^h$ and $r \in \mathbb{R}_{>0}$, let $B(v_0, r) = \{v \in \mathbb{R}^h : \|v_0 - v\| \leq r\}$ denote the closed ball centered at v_0 of radius r in $\mathbb{R}^h$. There is a constant $c_h \in \mathbb{R}_{>0}$ such that $B(v_0, r)$ has volume $\mathrm{vol}(B(v_0, r)) = c_h r^h$. For any $x \in \Sigma^*$, $\|g_{N,m}(x)\| \leq \sqrt{h}$, which implies that, for any $\delta \in \mathbb{R}_{>0}$, $B(g_{N,m}(x), \delta)$ is properly contained in $B(0, \sqrt{h} + \delta)$. Suppose $\forall x, x' \in X$ with $x \neq x'$, we have $B(g_{N,m}(x), \delta) \cap B(g_{N,m}(x'), \delta) = \emptyset$. Then $\sqcup_{x \in X} B(g_{N,m}(x), \delta) \subseteq B(0, \sqrt{h} + \delta)$, which implies $|X| c_h \delta^h \leq c_h(\sqrt{h} + \delta)^h$. Set $\delta = \frac{2\sqrt{h}}{|X|^{1/h} - 1}$. Then $\exists x, x' \in X$, with $x \neq x'$, such that $B(g_{N,m}(x), \delta) \cap B(g_{N,m}(x'), \delta) \neq \emptyset$, which implies $\|g_{N,m}(x) - g_{N,m}(x')\| \leq 2\delta$. Therefore,

$$\|N_{x,m}^{\hookrightarrow} - N_{x',m}^{\hookrightarrow}\|_1 \leq \sqrt{2} k^2 d^2 \|g_{N,m}(x) - g_{N,m}(x')\| \leq \sqrt{2} k^2 d^2 \frac{4\sqrt{h}}{|X|^{1/h} - 1} \leq 4\sqrt{2} k^4 d^3 \left(|X|^{\frac{1}{k^4 d^2}} - 1\right)^{-1}.$$

$\square$

We now prove the main technical result of this section: a 2QCFA analogue of the Dwork and Stockmeyer lemma [9, Lemma 4.3]. That is to say, we show that, if a 2QCFA recognizes some language L with bounded error, then $T(n)$, the maximum expected running time of that 2QCFA on inputs of length at most n, is lower bounded by an appropriate function of the hardness measure $D_L(n)$.

Theorem 4.4. *If $L \in \mathsf{B2QCFA}(k, d, T(n), \epsilon)$, for some $k, d \in \mathbb{N}_{\geq 2}$, $T : \mathbb{N} \to \mathbb{N}$, and $\epsilon \in [0, \frac{1}{2})$, then $\exists N_0 \in \mathbb{N}$ such that $T(n) \geq \frac{(1-2\epsilon)^2}{16\sqrt{2}k^4 d^3} D_L(n)^{\frac{1}{k^4 d^2}}$, $\forall n \geq N_0$.*

Proof. Consider some language L over some finite alphabet Σ. By [9, Lemma 3.1], $L \in \mathsf{REG}$ if and only if $\exists b \in \mathbb{N}_{\geq 1}$ such that $D_L(n) \leq b$, $\forall n \in \mathbb{N}$. Therefore, if $L \in \mathsf{REG}$, the claim is immediate (recall that $T(n) \geq n$); for the remainder of the proof, we assume $L \notin \mathsf{REG}$.

Suppose $L \in \mathsf{B2QCFA}(k, d, T(n), \epsilon)$. For each $n \in \mathbb{N}$, we define $X_n = \{x_1, \cdots, x_{D_L(n)}\} \subseteq \Sigma^{\leq n}$ such that the x_i are pairwise (L, n)-dissimilar. As $D_L(n)$ is not bounded above by any constant, $\exists N_0 \in \mathbb{N}$ such that $D_L(N_0) \geq 2^{k^4 d^2}$. Then, $\forall n \geq N_0$, we have $|X_n| = D_L(n) \geq D_L(N_0) \geq 2^{k^4 d^2}$. Fix $n \geq N_0$ and set $m = \lceil \frac{1-2\epsilon}{2} T(n) \rceil$. By Lemma 4.3, $\exists x, x' \in X_n$ such that $x \neq x'$ and

$$\|N_{x,m}^{\hookrightarrow} - N_{x',m}^{\hookrightarrow}\|_1 \leq 4\sqrt{2} k^4 d^3 \left(|X_n|^{\frac{1}{k^4 d^2}} - 1\right)^{-1} \leq 8\sqrt{2} k^4 d^3 |X_n|^{-\frac{1}{k^4 d^2}} = 8\sqrt{2} k^4 d^3 D_L(n)^{-\frac{1}{k^4 d^2}}.$$

Fix such a pair x, x', and note that $x \not\sim_{L,n} x'$, by construction. By Lemma 4.2,

$$T(n) \geq \frac{(1 - 2\epsilon)^2}{2} \|N_{x,m}^{\hookrightarrow} - N_{x',m}^{\hookrightarrow}\|_1^{-1} \geq \frac{(1 - 2\epsilon)^2}{16\sqrt{2} k^4 d^3} D_L(n)^{\frac{1}{k^4 d^2}}.$$

$\square$

4.3 2QCFA Running Time Lower Bounds and Complexity Class Separations

Theorem 4.4 has several significant implications on the power of 2QCFA. To allow us to properly state our results, as well as to better enable us to discuss existing results, we now define a collection of complexity classes that capture the power of 2QCFA with particular resource bounds. We first define $\mathsf{B2QCFA}(T(n)) = \cup_{k,d \in \mathbb{N}_{\geq 2}, \epsilon \in [0, \frac{1}{2})} \mathsf{B2QCFA}(k, d, T(n), \epsilon)$ to be the class of languages recognizable with two-sided bounded error by a 2QCFA with any constant number of quantum basis states and any constant number of classical states, in expected time at most $T(n)$ on all inputs of length at most n. We use the standard big O, little o, Ω, etc. notation to denote the asymptotic behavior of functions. For a family $\mathcal{T}$ of functions of the form $T : \mathbb{N} \to \mathbb{N}$, let

$\mathsf{B2QCFA}(\mathcal{T}) = \cup_{T\in\mathcal{T}}\mathsf{B2QCFA}(T(n))$. We then write, for example, $\mathsf{B2QCFA}(2^{o(n)})$ to denote the union, taken over every function $T : \mathbb{N} \to \mathbb{N}$ such that $T(n) = 2^{o(n)}$, of $\mathsf{B2QCFA}(T(n))$. See Section 4.1 for the definition of D_L and related hardness measures. We immediately obtain the following corollaries of Theorem 4.4.

Corollary 4.4.1. *If $L \in \mathsf{B2QCFA}(T(n))$, then $T(n) = D_L(n)^{\Omega(1)}$ and $T(n) = 2^{\Omega(C_L(n))}$.*

Corollary 4.4.2. *If a language L satisfies $D_L(n) = 2^{\Omega(n)}$, then $L \notin \mathsf{B2QCFA}(2^{o(n)})$.*

Notice that $D_L(n) = 2^{O(n)}$, for any language L. We next exhibit a language for which $D_L(n) = 2^{\Omega(n)}$, thereby yielding a strong lower bound on the running time of any 2QCFA that recognizes L. For $w = w_1 \cdots w_n \in \Sigma^*$, where each $w_i \in \Sigma$, let $w^{\mathrm{rev}} = w_n \cdots w_1$ denote the reversal of the string w. We consider the language $L_{pal} = \{w \in \{a,b\}^* : w = w^{\mathrm{rev}}\}$ consisting of all palindromes over the alphabet $\{a,b\}$.

Corollary 4.4.3. *$L_{pal} \notin \mathsf{B2QCFA}(2^{o(n)})$.*

Proof. For each $n \in \mathbb{N}$, let $W_n = \{w \in \{a,b\}^* : |w| = n\}$ denote all words over the alphabet $\{a,b\}$ of length n. For any $w, w' \in W_n$, with $w \neq w'$, we have $|ww^{\mathrm{rev}}| = 2n = |w'w^{\mathrm{rev}}|$, $ww^{\mathrm{rev}} \in L_{pal}$, and $w'w^{\mathrm{rev}} \notin L_{pal}$; therefore, by definition, $w \not\sim_{L_{pal},2n} w'$, $\forall w, w' \in W_n$ such that $w \neq w'$. This implies that $D_{L_{pal}}(2n) \geq |W_n| = 2^n$. Corollary 4.4.2 then implies $L_{pal} \notin \mathsf{B2QCFA}(T(n))$. $\square$

We define $\mathsf{BQE2QCFA} = \mathsf{B2QCFA}(2^{O(n)})$ to be the class of languages recognizable with two-sided bounded error in expected exponential time (with linear exponent) by a 2QCFA. Next, we say that a 2QCFA N recognizes a language L with *negative one-sided bounded error* $\epsilon \in \mathbb{R}_{>0}$ if, $\forall w \in L$, $\Pr[N \text{ accepts } w] = 1$, and, $\forall w \notin L$, $\Pr[N \text{ accepts } w] \leq \epsilon$. We define $\mathsf{coR2QCFA}(k,d,T(n),\epsilon)$ as the class of languages recognizable with negative one-sided bounded error ϵ by a 2QCFA, with at most k quantum basis states and at most d classical states, that has expected running time at most $T(n)$ on all inputs of length at most n. We define $\mathsf{coR2QCFA}(T(n))$ and $\mathsf{coRQE2QCFA}$ analogously to the two-sided bounded error case.

Ambainis and Watrous [2] showed that $L_{pal} \in \mathsf{coRQE2QCFA}$; in fact, their 2QCFA recognizer for L_{pal} has only a single-qubit (i.e., $k = 2$ quantum basis states). Clearly, $\mathsf{coR2QCFA}(T(n)) \subseteq \mathsf{B2QCFA}(T(n))$, for any T, and $\mathsf{coRQE2QCFA} \subseteq \mathsf{BQE2QCFA}$. Therefore, the class of languages recognizable by a 2QCFA with bounded error in expected *subexponential* time is properly contained in the class of languages recognizable by a 2QCFA in expected *exponential* time, as formalized by the following corollary.

Corollary 4.4.4. *We have $\mathsf{B2QCFA}(2^{o(n)}) \subsetneq \mathsf{BQE2QCFA}$ and $\mathsf{coR2QCFA}(2^{o(n)}) \subsetneq \mathsf{coRQE2QCFA}$.*

We next define $\mathsf{BQP2QCFA} = \mathsf{B2QCFA}(n^{O(1)})$ to be the class of languages recognizable with two-sided bounded error in expected polynomial time by a 2QCFA.

Corollary 4.4.5. *If $L \in \mathsf{BQP2QCFA}$, then $D_L(n) = n^{O(1)}$ and $C_L(n) = O(\log n)$. Therefore, $\mathsf{BQP2QCFA} \subseteq L/poly$.*

Proof. The first statement is a special case of Corollary 4.4.1. To see that $\mathsf{BQP2QCFA} \subseteq L/poly$, recall that, as noted in Section 4.1, $L/poly = \{L : A^{2DFA}_{L,=}(n) = n^{O(1)}\}$; clearly, for any language L and any $n \in \mathbb{N}$, $A^{2DFA}_{L,=}(n) \leq A_L(n) = D_L(n)$. $\square$

Of course, there are many languages L for which one can establish a strong lower bound on $D_L(n)$, and thereby establish a strong lower bound on the expected running time $T(n)$ of any 2QCFA that recognizes L. In Section 6, we consider the case in which L is the word problem of a group, and we show that very strong lower bounds can be established on $D_L(n)$. In the current section, we consider two especially interesting languages; the relevance of these languages was brought to our attention by Richard Lipton (personal communication). For a number $p \in \mathbb{N}$, let $\langle p \rangle_2 \in \{0,1\}^*$ denote the binary representation of p; let $L_{primes} = \{\langle p \rangle_2 : p \text{ is prime}\}$. Note that $D_{L_{primes}}(n) = 2^{\Omega(n)}$ [35], which immediately implies the following.

Corollary 4.4.6. $L_{primes} \notin \mathsf{B2QCFA}(2^{o(n)})$.

Say a string $w = w_1 \cdots w_n \in \{0,1\}^n$ has a length-3 arithmetic progression (3AP) if $\exists i, j, k \in \mathbb{N}$ such that $1 \leq i < j < k \leq n$, $j - i = k - j$, and $w_i = w_j = w_k = 1$; let $L_{3ap} = \{w \in \{0,1\}^* : w \text{ has a 3AP}\}$. It is straightforward to show the lower bound $D_{L_{3ap}}(n) = 2^{n^{1-o(1)}}$, as well as the upper bound $D_{L_{3ap}}(n) = 2^{n^{o(n)}}$. Therefore, one obtains the following lower bound on the running time of a 2QCFA that recognizes L_{3ap}, which, while still quite strong, is not as strong as that of L_{pal} or L_{primes}.

Corollary 4.4.7. $L_{3ap} \notin \mathsf{B2QCFA}\left(2^{n^{1-\Omega(1)}}\right)$.

Remark. While L_{primes} and L_{3ap} provide two more examples of natural languages for which our method yields strong lower bound on the running time of any 2QCFA recognizer, they also suggest the potential of proving a stronger lower bound for certain languages. That is to say, for L_{pal}, one has (essentially) matching lower and upper bounds on the running time of any 2QCFA recognizer; this is certainly not the case for L_{primes} and L_{3ap}. In fact, we currently do not know if either L_{primes} or L_{3ap} can be recognized by a 2QCFA with bounded error *at all* (i.e., regardless of time bound).

Lastly, we consider the issue of the *transition amplitudes* of a 2QCFA. For some 2QCFA $N = (Q, C, \Sigma, \delta_{\text{type}}, \delta_{\text{transform}}, \delta_{\text{measure}}, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$, let $V = \mathbb{C}^Q$ denote the finite-dimensional complex Hilbert space corresponding to the quantum register of N, and let $\mathcal{T} = \{t \in \mathrm{U}(V) : \delta_{\text{transform}}(c, \sigma) = (t, \cdot, \cdot), \text{ for some } (c, \sigma) \in \delta_{\text{type}}^{-1}(transform)\}$ denote the set of unitary operators that N may apply to its quantum register. For each $t \in \mathcal{T}$, there is a corresponding $|Q| \times |Q|$ complex matrix M_t that represents the linear operator $t \in \mathrm{L}(V)$ with respect to the basis $\{|q\rangle : q \in Q\}$ of V. Let $\mathcal{M} = \{M_t : t \in \mathcal{T}\}$ denote the set of all such matrices. The *transition amplitudes* of N are the set of numbers that appear as an entry of some matrix $M_t \in \mathcal{M}$.

While other types of finite automata are often defined without any restriction on their transition amplitudes, for 2QCFA, and other types of QFA, the allowed class of transition amplitudes strongly affects the power of the model. For example, using non-computable transition amplitudes, a 2QCFA can recognize certain undecidable languages with bounded error in expected polynomial time [34]. Our lower bound holds even in this setting of unrestricted transition amplitudes. For $\mathbb{F} \subseteq \mathbb{C}$, we define complexity classes $\mathsf{coR2QCFA}_{\mathbb{F}}(k, d, T(n), \epsilon)$, $\mathsf{coRQE2QCFA}_{\mathbb{F}}$, etc., that are variants of the corresponding complexity class in which the 2QCFA are restricted to have transition amplitudes in $\mathbb{F}$. Using our terminology, Ambainis and Watrous [2] showed that $L_{pal} \in \mathsf{coRQE2QCFA}_{\overline{\mathbb{Q}}}$, where $\overline{\mathbb{Q}}$ denotes the algebraic numbers, which are, arguably, the natural choice for the permitted class of transition amplitudes of a quantum model of computation. Therefore, L_{pal} can be recognized with negative one-sided bounded error by a single-qubit 2QCFA with transition amplitudes that are all algebraic numbers in expected exponential time; however, L_{pal} cannot be recognized with two-sided bounded error (and, therefore, not with one-sided bounded error) by a 2QCFA (of any constant size) in expected subexponential time, regardless of the permitted transition amplitudes.

5 Lower Bounds on the Running Time of Small-Space QTMs

In this section, we show that our technique also yields a lower bound on the expected running time of a quantum Turing machine (QTM) that uses sublogarithmic space (i.e., $o(\log n)$ space) and recognizes a language L with bounded error. The key idea is that a QTM M that runs in expected time at most $T(n)$ and uses space at most $S(n)$ can be viewed as a sequence $(M_n)_{n\in\mathbb{N}}$ of 2QCFA, where M_n has $2^{O(S(n))}$ (classical and quantum) states and M_n simulates M on all inputs of length at most n (therefore, M_n and M have the same probability of acceptance and the same expected running time on any such input). The techniques of the previous section apply to 2QCFA with a sufficiently slowly growing number of states. See, for instance, [9, 22] for examples of this approach for classical TMs.

We begin by informally defining the *classically controlled* space-bounded QTM model that allows *intermediate measurements*, following the definitions of Ta-Shma [40], Watrous [44, Section VII.2], and (essentially, without the use of random access) van Melkebeek and Watson [26]. Such a QTM consists of three tapes: (1) a classical read-only input tape, which has a single bidirectional (classical) head, and where each cell stores a symbol from the input alphabet (with special end-markers at the left and right ends), (2) a classical one-way infinite work tape, which also has a single bidirectional (classical) head, and where each cell stores a symbol from some potentially larger (finite) alphabet, and (3) a one-way infinite quantum work tape, which has two bidirectional (classical) heads, and where each cell contains a single qubit.

The computation of the QTM is entirely *classically controlled.* Each step of the computation consists of a *quantum phase* followed by a *classical phase.* In the quantum phase, a QTM performs a quantum operation on the one or two qubits currently under the heads of the quantum work tape. Let G denote any universal set of quantum gates that each operate on at most two qubits (e.g., $G = \{\text{Hadamard, CNOT, rotation by } \frac{\pi}{8}\}$). The operation consists of either applying some particular $g \in G$ or performing a projective measurement in the computational basis; the choice of which operation to perform may depend on the current (classical) state of its finite control and the current symbols read on the input tape and classical work tape. In the classical phase, depending on the current state of its finite control, on the current symbols read on the input tape and classical work tape, and (possibly) on the (classical) result of the quantum measurement performed in the quantum phase (if a measurement was indeed performed in the previous quantum phase), the QTM updates its configuration as follows: the state of the finite control changes to a new classical state, a symbol is written on the cell of the classical work tape under the head, and the heads of all tapes move at most one cell in either direction.

A (branch of the computation of a) QTM halts and accepts/rejects its input by entering a special classical accept/reject state. As we wish to make our lower bound as strong as possible, we wish to be as generous as possible with the rejecting criteria of a QTM, and so we allow a QTM to also reject by looping (as we did with 2QCFA). Let $\mathsf{BQTISP}_\epsilon(T(n), S(n))$ denote the class of languages recognizable with two-sided bounded error $\epsilon \in [0, 1/2)$ by a QTM that runs in at most $T(n)$ expected time, and uses at most $S(n)$ space, on all inputs of length at most n; of course, only the space used on the (classical and quantum) work tapes is counted. Furthermore, let $\mathsf{BQTISP}(T(n), S(n)) = \cup_{\epsilon\in[0,1/2)}\mathsf{BQTISP}_\epsilon(T(n), S(n))$.

We note that while we only explicitly consider the above variant of QTM, the results of this section would apply to essentially any "reasonable" *classically controlled* QTM variant, such as the (other such) variant defined by Watrous [43], or the variant defined by Perdrix and Jorund [31] (where their model is modified in the usual way in order to handle sublinear space bounds). Moreover, these results would apply even to the *unreasonable* QTM variant in which the set G (of unitary transformations that the QTM could apply at any step) was allowed to be any finite set of unitary transformations (i.e., no restriction is placed on the transition amplitudes of the QTM, see the discussion at the end of Section 4.3); the analogously defined version of $\mathsf{BQTISP}(T(n), O(1))$ for such a QTM model would then be equal to $\mathsf{B2QCFA}(T(n))$.

We also note that the consideration of classically controlled QTMs is natural as the clear separation of quantum and classical parts accurately reflects the design of current and near-term experimental quantum computers, and, moreover, such a model appears far easier to analyze.

Lastly, we emphasize that the QTM model that we consider permits *intermediate measurements*. In the case of time-bounded quantum computation, allowing a QTM to perform intermediate measurements provably does not increase the power of the model. This is due to the *principle of safe storage*, which allows all measurements to be deferred until the end of a computation without affecting the running time; however, the standard technique for deferring measurements may cause a significant increase in the required space. It remains an open question whether or not allowing a space-bounded QTM to perform intermediate measurements increases the power of the model. Again, as we want our lower bound to be as strong as possible, we allow intermediate measurements. We direct the reader to [26, Section 2] for a detailed discussion of the various models of space-bounded QTMs.

As noted at the beginning of this section, we may view a QTM M that operates in space $S(n)$ as a sequence of 2QCFA with a growing number of states. This yields the following analogue of Theorem 4.4 for sublogarithmic-space QTMs.

Theorem 5.1. *Suppose* $L \in \mathsf{BQTISP}(T(n), S(n))$*, and suppose further that* $S(n) = o(\log\log D_L(n))$*. Then* $\exists b_0 \in \mathbb{R}_{>0}$ *such that,* $T(n) = \Omega\left(2^{-b_0 S(n)} D_L(n)^{2^{-b_0 S(n)}}\right)$.

Proof. By definition, there is some QTM M that recognizes L with two-sided bounded error ϵ, for some $\epsilon \in [0, 1/2)$, where M runs in expected time at most $T(n)$, and uses at most $S(n)$ space, on all inputs of length at most n. Let F denote the (finite) set of (classical) states that comprise the finite control of M, let Σ denote the (finite) input alphabet of M, and let Γ denote the (finite) alphabet of the classical work tape of M.

For each $n \in \mathbb{N}$, we define a 2QCFA M_n that correctly simulates M on any $w \in \Sigma^{\leq n}$, in the obvious way: the (only) head of the 2QCFA M_n (on its read-only input tape) directly simulates the head of the QTM M on its read-only input tape; M_n uses its classical states C_n to keep track of the state $f \in F$ of the finite control of M, the string $y \in \Gamma^{S(n)}$ that appears in the first $S(n)$ cells of the classical work tape, and the positions $h_{c-work}, h_{q-work-1}, h_{q-work-2} \in \{1, \ldots, S(n)\}$ of the heads on the (classical and quantum) work tapes; M_n uses its quantum register, which has quantum basis states Q_n, to store the first $S(n)$ qubits of the quantum work tape; the transition function of the 2QCFA M_n is defined such that, if M_n is in a classic state $c \in C_n$ which (along with the head position on the input tape) completely specifies the classical part of a configuration of M, then M_n performs the same quantum phase and classical phase that M would in this configuration. Clearly, for any $w \in \Sigma^{\leq n}$, M_n and M have the same probability of acceptance and expected running time.

Let $k_n = |Q_n| = 2^{S(n)}$ denote the number of quantum basis states of M_n and let $d_n = |C_n| = |F||\Gamma|^{S(n)} S(n)^3$ denote the number of classical states of M_n. Then, $\exists b_0 \in \mathbb{R}_{>0}, \exists \widehat{N}_0 \in \mathbb{N}$ such that, $\forall n \geq \widehat{N}_0$, we have $k_n^4 d_n^3 \leq 2^{b_0 S(n)}$. Moreover, as $S(n) = o(\log\log D_L(n))$, $\exists \widetilde{N}_0 \in \mathbb{N}$ such that, $\forall n \geq \widetilde{N}_0$, $D_L(n)^{2^{-b_0 S(n)}} \geq 2$. Set $N_0 = \max(\widehat{N}_0, \widetilde{N}_0)$. For any $n \geq N_0$, we may then construct $X_n \subseteq \Sigma^{\leq n}$ such that $|X_n| = D_L(n) \geq 2$ and the elements of X_n are pairwise (L, n)-dissimilar. By Lemma 4.3, $\exists x, x' \in X_n$ such that $x \neq x'$ and

$$\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1 \leq 4\sqrt{2} k_n^4 d_n^3 \left(D_L(n)^{\frac{1}{k_n^4 d_n^2}} - 1\right)^{-1} \leq (4\sqrt{2}) 2^{b_0 S(n)} \left(D_L(n)^{2^{-b_0 S(n)}} - 1\right)^{-1}.$$

Let $a_\epsilon = \frac{(1-2\epsilon)^2}{2} \in \mathbb{R}_{>0}$. By Lemma 4.2,

$$T(n) \geq a_\epsilon \|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1^{-1} \geq \frac{a_\epsilon}{4\sqrt{2}} 2^{-b_0 S(n)} \left(D_L(n)^{2^{-b_0 S(n)}} - 1 \right) \geq \frac{a_\epsilon}{8\sqrt{2}} 2^{-b_0 S(n)} D_L(n)^{2^{-b_0 S(n)}}$$

□

Remark. Recall that, for any language L, $D_L(n) = 2^{O(n)}$; therefore, the supposition of the above theorem that $S(n) = o(\log\log D_L(n))$ implies $S(n) = o(\log n)$, and so this theorem only applies to QTMs that use sublogarithmic space. Moreover, this requirement also implies that $D_L(n) = \omega(1)$, and hence $L \notin \mathsf{REG}$ [9, Lemma 3.1]; of course, for any $L \in \mathsf{REG}$, we trivially have $L \in \mathsf{BQTISP}(n, O(1))$.

Note that, if $S(n) = o(\log n)$, then for any constants $b_1, b_2 \in \mathbb{R}_{>0}$, $2^{-b_1 S(n)} \geq n^{-b_2}$, for all sufficiently large n. We therefore obtain the following corollary.

Corollary 5.1.1. *If a language L satisfies $D_L(n) = 2^{\Omega(n)}$, then $L \notin \mathsf{BQTISP}\left(2^{n^{1-\Omega(1)}}, o(\log n)\right)$. In particular, $L_{pal} \notin \mathsf{BQTISP}\left(2^{n^{1-\Omega(1)}}, o(\log n)\right)$.*

Remark. Of course, L_{pal} can be recognized by a *deterministic* Turing machine in $O(\log n)$ space (and, trivially, polynomial time). Therefore, the previous corollary exhibits a natural problem for which polynomial time *quantum* Turing machines cannot (asymptotically) outperform polynomial time *deterministic* Turing machines in terms of the amount of space used.

6 The Word Problem of a Group

We begin by formally defining the word problem of a group; for further background, see, for instance [25]. For a set S, let $F(S)$ denote the free group on S. For sets S, R such that $R \subseteq F(S)$, let N denote the normal closure of R in $F(S)$; for a group G, if $G \cong F(S)/N$, then we say that G *has presentation* $\langle S|R \rangle$, which we denote by writing $G = \langle S|R \rangle$.

Suppose $G = \langle S|R \rangle$, with S finite; we now define $W_{G=\langle S|R \rangle}$, *the word problem of G with respect to the presentation $\langle S|R \rangle$*. We define the set of formal inverses S^{-1}, such that, for each $s \in S$, there is a unique corresponding $s^{-1} \in S^{-1}$, and $S \cap S^{-1} = \emptyset$. Let $\Sigma = S \sqcup S^{-1}$, let Σ^* denote the free monoid over Σ, and let $\phi : \Sigma^* \to G$ be the natural (monoid) homomorphism that takes each string in Σ^* to the element of G that it represents. We use 1_G to denote the identity element of G. Then $W_{G=\langle S|R \rangle} = \phi^{-1}(1_G)$.

We say that G is *finitely generated* if it has a presentation $\langle S|R \rangle$ where S is finite. Note that the word problem of G is only defined when G is finitely generated and that the definition of the word problem does depend on the particular presentation. However, it is well-known (see, for instance, [18]) that if $\mathcal{L}$ is any complexity class that is closed under inverse homomorphism, then if $\langle S|R \rangle$ and $\langle S'|R' \rangle$ are both presentations of some group G, and S and S' are both finite, then $W_{G=\langle S|R \rangle} \in \mathcal{L} \Leftrightarrow W_{G=\langle S'|R' \rangle} \in \mathcal{L}$. As all complexity classes considered in this paper are easily seen to be closed under inverse homomorphism, we will simply write $W_G \in \mathcal{L}$ to mean that $W_{G=\langle S|R \rangle} \in \mathcal{L}$, for every presentation $G = \langle S|R \rangle$, with S finite.

We note that the languages L_{pal} and L_{eq}, which Ambainis and Watrous [2] showed satisfy $L_{pal} \in \mathsf{coRQE2QCFA}_{\overline{\mathbb{Q}}}$ and $L_{eq} \in \mathsf{BQP2QCFA}$, are closely related to the word problems of the groups F_2 and $\mathbb{Z}$, respectively (see [33] for a full discussion of this correspondence).

In this section, we consider the (quantum) computational complexity of the word problem W_G corresponding to a finitely generated group G. We will show that there is a close correspondence between D_{W_G} and the *growth rate* of the group G, which will enable us to exhibit a strong lower bound on the expected running time of a 2QCFA that recognizes a word problem from a particular class of groups. By combining these lower bounds with a recent result of ours [33] that showed that 2QCFA can recognize certain wide classes of group word problems within particular time bounds, we obtain a natural class of languages that 2QCFA can recognize with bounded error in expected exponential time, but not in expected subexponential time, as well as strong statements about the class of group word problems that a 2QCFA can recognize with bounded error in expected polynomial time.

6.1 The Growth Rate of a Group and Nonregularity

Consider a group $G = \langle S|R\rangle$, with S finite. As above, let $\Sigma = S \sqcup S^{-1}$, and let $\phi : \Sigma^* \to G$ denote the natural map that takes each string in Σ^* to the element of G that it represents. For $g \in G$, we define the *length of g with respect to S*, which we denote by $l_S(g)$, as the smallest $m \in \mathbb{N}$ such that $\exists \sigma_1, \ldots, \sigma_m \in \Sigma$ such that $g = \phi(\sigma_1 \cdots \sigma_m)$. For $n \in \mathbb{N}$, we define $B_{G,S}(n) = \{g \in G : l_S(g) \leq n\}$ and we further define $\beta_{G,S}(n) = |B_{G,S}(n)|$, which we call the *growth rate of G with respect to S*. The following straightforward lemma demonstrates an important relationship between $\beta_{G,S}$ and $D_{W_{G=\langle S|R\rangle}}$.

Lemma 6.1. *Suppose $G = \langle S|R\rangle$ with S finite. Using the notation established above, let $W_G := W_{G=\langle S|R\rangle} = \phi^{-1}(1_G)$ denote the word problem of G with respect to this presentation. Then, $\forall n \in \mathbb{N}$, $D_{W_G}(2n) \geq \beta_{G,S}(n)$.*

Proof. Fix $n \in \mathbb{N}$, let $k = \beta_{G,S}(n)$, and let $B_{G,S}(n) = \{g_1, \ldots, g_k\}$. For a string $x = x_1 \cdots x_m \in \Sigma^*$, where each $x_j \in \Sigma$, let $|x| = m$ denote the (string) length of x and define $x^{-1} = x_m^{-1} \cdots x_1^{-1}$. Note that, $\forall g \in G$, $l_S(g) = \min_{w \in \phi^{-1}(g)} |w|$. Therefore, for each $i \in \{1, \ldots, k\}$ we may define $w_i \in \phi^{-1}(g_i)$ such that $|w_i| = l_S(g_i)$. Observe that $w_i w_i^{-1} \in W_G$ and $|w_i w_i^{-1}| = 2|w_i| = 2l_S(g_i) \leq 2n$; moreover, for each $j \neq i$, we have $w_j w_i^{-1} \notin W_G$ and $|w_j w_i^{-1}| = |w_j| + |w_i| = l_S(g_j) + l_S(g_i) \leq 2n$. Therefore, $w_1, \ldots, w_k$ are pairwise $(W_G, 2n)$-dissimilar, which implies $D_{W_G}(2n) \geq k = \beta_{G,S}(n)$. $\square$

Remark. In fact, one may also easily show that $D_{W_G}(2n) \leq \beta_{G,S}(n)+1$, though we do not need this here. Essentially, $\beta_{G,S}(n)$ is (another) equivalent characterization of the nonregularity $D_{W_G}(2n)$ (see Section 4.1 for a discussion of the many such characterizations of nonregularity).

While $\beta_{G,S}$ does depend on the particular choice of the generating set S, the dependence is quite minor, in a sense that we now clarify. For a pair of non-decreasing functions $f_1, f_2 : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, we write $f_1 \prec f_2$ if $\exists C_1, C_2 \in \mathbb{R}_{>0}$ such that $\forall r \in \mathbb{R}_{\geq 0}$, $f_1(r) \leq C_1 f_2(C_1 r + C_2) + C_2$; if both $f_1 \prec f_2$ and $f_2 \prec f_1$, then we say that f_1 is *quasi-equivalent* to f_2, which we denote by $f_1 \sim f_2$. We extend a growth function $\beta_{G,S} : \mathbb{N} \to \mathbb{N}$ to $\beta_{G,S} : \mathbb{R}_{\geq 0} \to \mathbb{N}$ by defining $\beta_{G,S}(r) = \beta_{G,S}(\lceil r \rceil)$, $\forall r \in \mathbb{R}_{\geq 0}$. Suppose $G = \langle S'|R'\rangle$, where S' is finite. It is straightforward to show that $\beta_{G,S}$ and $\beta_{G,S'}$ are non-decreasing, and that $\beta_{G,S} \sim \beta_{G,S'}$ (see, for instance, [25, Proposition 6.2.4]). For this reason, we will often omit S and simply write β_G to denote the growth rate of G, when we only care about the growth rate up to quasi-equivalence. We then make the following definition.

Definition 6.2. Suppose G is a finitely generated group.

(i) If $\beta_G \sim (n \mapsto e^n)$, we say G *has exponential growth.*

(ii) If $\exists c \in \mathbb{R}_{\geq 0}$ such that $\beta_G \prec (n \mapsto n^c)$, we say G *has polynomial growth.*

(iii) If G has neither polynomial growth nor exponential growth, we say G *has intermediate growth.*

Note that, for any finitely generated group G, we have $\beta_G \prec (n \mapsto e^n)$, and so the term "intermediate" growth is justified.

6.2 Word Problems Recognizable by 2QCFA and Small-Space QTMs

By making use of two very powerful results in group theory, the Tits' Alternative [42] and Gromov's theorem on groups of polynomial growth [14], we exhibit useful lower bounds on D_{W_G}, which in turn allows us to show a strong lower bound on the expected running time of a 2QCFA that recognizes W_G. In the following, we use the notation for complexity classes established in Section 4.3. As previously noted, the membership of W_G in any of the complexity classes in question does not depend on the particular choice of presentation, and so we write, for example, $W_G \in \mathsf{BQP2QCFA}$ to mean $W_{G=\langle S|R\rangle} \in \mathsf{BQP2QCFA}$ for some (equivalently every) presentation $G = \langle S|R\rangle$, with S finite.

Theorem 6.3. *For any finitely generated group G, the following statements hold.*

(i) If $W_G \in \mathsf{B2QCFA}(k, d, T(n), \epsilon)$, then $\beta_G \prec (n \mapsto T(n)^{k^4d^2})$.

(ii) If G has exponential growth, then $W_G \notin \mathsf{B2QCFA}(2^{o(n)})$.

(iii) If G is a linear group over a field of characteristic 0, and G is not virtually nilpotent, then $W_G \notin \mathsf{B2QCFA}(2^{o(n)})$.

(iv) If $W_G \in \mathsf{BQP2QCFA}$, then G is virtually nilpotent.

Proof. (i) Follows immediately from Lemma 6.1 and Corollary 4.4.1.

(ii) Follows immediately from Definition 6.2(i) and part (i) of this theorem.

(iii) As a consequence of the famous Tits' Alternative [42], every finitely generated linear group over a field of characteristic 0 either has polynomial growth or exponential growth, and has polynomial growth precisely when it is virtually nilpotent ([42, Corollary 1], [45]). The claim then follows by part (ii) of this theorem.

(iv) If $W_G \in \mathsf{BQP2QCFA}$, then $W_G \in \mathsf{B2QCFA}(k, d, n^c, \epsilon)$ for some $k, d, c \in \mathbb{N}_{\geq 1}, \epsilon \in [0, \frac{1}{2})$. By part (i) of this theorem, $\beta_G \prec (n \mapsto n^{ck^4d^2})$, which implies G has polynomial growth. By Gromov's theorem on groups of polynomial growth [14], a finitely generated group has polynomial growth precisely when it is virtually nilpotent.

□

Remark. We note that, while finitely generated groups of intermediate growth provably exist [13], all known groups of intermediate growth have growth rate quasi-equivalent to $(n \mapsto e^{n^c})$, for some $c \in (1/2, 1)$. Therefore, if W_G is the word problem for one of these known groups of intermediate growth, a strong lower bound may be established on D_{W_G}, which in turn allows a strong lower bound to be established on the running time of any 2QCFA that recognizes W_G for one of these known groups of intermediate growth. We also note that one may show that the conclusion of Theorem 6.3(iv) still holds even if W_G is only assumed to be recognized in slightly super-polynomial time. In particular, by a quantitative version of Gromov's theorem due to Shalom and Tal [37, Corollary 1.10], $\exists c \in \mathbb{R}_{>0}$ such that if $\beta_{G,S}(n) \leq n^{c(\log\log n)^c}$, for some $n > 1/c$, then G is virtually nilpotent.

Let $\mathcal{G}_{\mathbf{vAb}}$ (resp. $\mathcal{G}_{\mathbf{vNilp}}$) denote the collection of all finitely generated virtually abelian (resp. nilpotent) groups. Let $\overline{\mathbb{Q}}$ denote the algebraic numbers and let $\mathrm{U}(k, \overline{\mathbb{Q}})$ denote the group of $k \times k$ unitary matrices with entries in $\overline{\mathbb{Q}}$, and let $\mathcal{U}$ denote the family of finitely generated groups G such that G is isomorphic to a subgroup of $\mathrm{U}(k, \overline{\mathbb{Q}})$, for some k. We have recently shown that if $G \in \mathcal{U}$, then $W_G \in \mathsf{coRQE2QCFA}_{\overline{\mathbb{Q}}}$ [33, Corollary 1.4.1]. Observe that $\mathcal{G}_{\mathbf{vAb}} \subseteq \mathcal{U}$ and that all groups in $\mathcal{U}$

are finitely generated linear groups over a field of characteristic zero. Moreover, $\mathcal{U} \cap \mathcal{G}_{\mathbf{vNilp}} = \mathcal{G}_{\mathbf{vAb}}$ (see, for instance, [41, Proposition 2.2]). We therefore immediately obtain the following corollary of Theorem 6.3(iii), which exhibits a broad and natural class of languages that a 2QCFA can recognize with bounded error in expected exponential time, but not in expected subexponential time. We note that $\mathcal{U} \setminus \mathcal{G}_{\mathbf{vAb}}$ is a rather wide class of groups, see [33] for a full discussion and related results.

Corollary 6.3.1. *For any $G \in \mathcal{U} \setminus \mathcal{G}_{\boldsymbol{vAb}}$ and for any $T : \mathbb{N} \to \mathbb{N}$ such that $T(n) = 2^{o(n)}$, we have $W_G \in \mathsf{coRQE2QCFA}_{\overline{\mathbb{Q}}}$ but $W_G \notin \mathsf{B2QCFA}(T(n))$.*

Let $\mathsf{coRQP2QCFA}_{\overline{\mathbb{Q}}}(2)$ denote the class of languages recognizable with negative one-side bounded error by a 2QCFA, with a single-qubit quantum register and algebraic number transition amplitudes, in expected polynomial time. We have also recently shown that $W_G \in \mathsf{coRQP2QCFA}_{\overline{\mathbb{Q}}}(2) \subseteq \mathsf{BQP2QCFA}$, $\forall G \in \mathcal{G}_{\mathbf{vAb}}$ [33, Theorem 1.2]. By Theorem 6.3(iv), if $W_G \in \mathsf{BQP2QCFA}$, then $G \in \mathcal{G}_{\mathbf{vNilp}}$. This naturally raises the question of whether or not there is some $G \in \mathcal{G}_{\mathbf{vNilp}} \setminus \mathcal{G}_{\mathbf{vAb}}$ such that $W_G \in \mathsf{BQP2QCFA}$. In particular, consider the (three-dimensional discrete) Heisenberg group $H = \langle x, y, z | z = [x, y], [x, z] = [y, z] = 1 \rangle$ (where $[x, y] = x^{-1}y^{-1}xy$ denotes the commutator of x and y and we have expressed the relators as equations, rather than words in $F(x, y, z)$, for convenience). The word problem W_H of the Heisenberg group H is a natural choice for a potential "hard" word problem for 2QCFA, due to the lack of faithful finite-dimensional unitary representations of H (see [33] for further discussion). In fact, it is possible, and perhaps plausible, that W_H cannot be recognized with bounded error by a 2QCFA in any time bound. We next show that if $W_H \notin \mathsf{BQP2QCFA}$, then we have a complete classification of those word problems recognizable by 2QCFA in expected polynomial time.

Proposition 6.4. *If $W_H \notin \mathsf{BQP2QCFA}$, where H is the Heisenberg group, then for any finitely generated group G, $W_G \in \mathsf{BQP2QCFA} \Leftrightarrow W_G \in \mathsf{coRQP2QCFA}_{\overline{\mathbb{Q}}}(2) \Leftrightarrow G \in \mathcal{G}_{\boldsymbol{vAb}}$.*

Proof. By the above discussion, it suffices to show the following claim: if $W_G \in \mathsf{BQP2QCFA}$, for some $G \in \mathcal{G}_{\mathbf{vNilp}} \setminus \mathcal{G}_{\mathbf{vAb}}$, then $W_H \in \mathsf{BQP2QCFA}$. Begin by noting that $\mathsf{BQP2QCFA}$ is easily seen to be closed under inverse homomorphism and intersection with regular languages. Suppose G and G' are finitely generated groups such that G' is (isomorphic to) a subgroup of G, if $W_G \in \mathsf{BQP2QCFA}$, then $W_{G'} \in \mathsf{BQP2QCFA}$ (see, for instance, [19, Lemma 2]). It is well-known that $H \in \mathcal{G}_{\mathbf{vNilp}} \setminus \mathcal{G}_{\mathbf{vAb}}$ and, $\forall G \in \mathcal{G}_{\mathbf{vNilp}} \setminus \mathcal{G}_{\mathbf{vAb}}$, G has a subgroup isomorphic to H (see, for instance, [19, Theorem 12] for these facts, as well as for their application towards understanding the computational complexity of the group word problem). ☐

We next obtain the following analogue of Theorem 6.3 for small-space QTMs.

Theorem 6.5. *For any finitely generated group G, the following statements hold.*

(i) If G has exponential growth, then $W_G \notin \mathsf{BQTISP}(2^{n^{1-\Omega(1)}}, o(\log n))$.

(ii) If G is a linear group over a field of characteristic 0, and G is not virtually nilpotent, then $W_G \notin \mathsf{BQTISP}(2^{n^{1-\Omega(1)}}, o(\log n))$.

(iii) If $W_G \in \mathsf{BQTISP}(n^{O(1)}, o(\log \log \log n))$, then G is virtually nilpotent.

Proof. (i) Follows immediately from Definition 6.2(i), Corollary 5.1.1, and Lemma 6.1.

(ii) The claim follows from the Tits' Alternative [42] and part (i) of this theorem.

(iii) If $W_G \in \mathsf{BQTISP}(n^{O(1)}, o(\log \log \log n))$, then $\forall c \in \mathbb{R}_{>0}$ and for all sufficiently large n we have, by Theorem 5.1, $D_L(n) \leq n^{c(\log \log n)^c}$. By Lemma 6.1 and the quantitative version of Gromov's theorem due to Shalom and Tal [37, Corollary 1.10], G is virtually nilpotent.

☐

7 Discussion

In this paper, we established strong lower bounds on the expected running time of 2QCFA, or sublogarithmic-space QTMs, that recognize particular languages with bounded error. In particular, the language L_{pal} had been shown by Ambainis and Watrous [2] to be recognizable with bounded error by a single-qubit 2QCFA in expected time $2^{O(n)}$. We have given a matching lower bound: no 2QCFA (of any size) can recognize L_{pal} with bounded error in expected time $2^{o(n)}$. Moreover, we have shown that no QTM, that runs in expected time $2^{n^{1-\Omega(1)}}$ and uses space $o(\log n)$, can recognize L_{pal} with bounded error. This latter results is especially interesting, as a *deterministic* TM can recognize L_{pal} using space $O(\log n)$ (and, of course, polynomial time); therefore, polynomial time *quantum* TMs have no (asymptotic) advantage over polynomial time *deterministic* TMs in terms of the amount of space needed to recognize L_{pal}.

Our main technical result, Theorem 4.4, showed that, if a language L is recognized with bounded error by a 2QCFA in expected time $T(n)$, then $\exists a \in \mathbb{R}_{>0}$ (that depends only on the number of states of the 2QCFA) such that $T(n) = \Omega(D_L(n)^a)$, where D_L is the Dwork-Stockmeyer *nonregularity* of L. This result is extremely (qualitatively) similar to the landmark result of Dwork and Stockmeyer [9, Lemma 4.3], which showed that, if a language L is recognized with bounded error by a 2PFA in expected time $T(n)$, then $\exists a \in \mathbb{R}_{>0}$ (that depends only on the number of states of the 2PFA) such that $T(n) = \Omega(2^{D_L(n)^a})$. We again note that both of these lower bounds are tight.

We conclude by stating a few interesting open problems. While our lower bound on the expected running time $T(n)$, of a 2QCFA that recognizes a language L, in terms of $D_L(n)$ cannot be improved, it is natural to ask if one could establish a lower bound on $T(n)$ in terms of a different hardness measure of L that would be stronger for certain languages. Generalizing the definitions made in Section 4.1, let $\mathcal{F}$ denote a class of finite automata (e.g., DFA, NFA, 2DFA, etc.), let L be a language over some alphabet Σ, and let $A^{\mathcal{F}}_{L,\leq}(n) = \min\{|M| : M \in \mathcal{F} \text{ and } L(M) \cap \Sigma^{\leq n} = L \cap \Sigma^{\leq n}\}$ denote the smallest number of states of an automaton of type $\mathcal{F}$ that agrees with L on all strings of length at most n. As discussed earlier, $A^{DFA}_{L,\leq}(n) = D_L(n)$, for any language L and for any $n \in \mathbb{N}$. Recall that DFA and 2DFA both recognize precisely the regular languages [32], but for some $\widehat{L} \in \mathsf{REG}$, the smallest 2DFA that recognizes $\widehat{L}$ might have many fewer states than the smallest DFA that recognizes $\widehat{L}$. In fact, there is a sequence of regular languages $(L_k)_{k\in\mathbb{N}}$ such that L_k can be recognized by a $5k+5$-state 2DFA, but any DFA that recognizes L_k requires at least k^k states [27]; however, this is (essentially) the largest succinctness advantage possible, as any language recognizable by a d-state 2DFA is recognizable by a $(d+2)^{d+1}$-state DFA [38]. Of course, for any language L, we have $A^{2DFA}_{L,\leq}(n) \leq A^{DFA}_{L,\leq}(n)$, $\forall n$. For certain languages L, we have $A^{2DFA}_{L,\leq}(n) \ll A^{DFA}_{L,\leq}(n)$, $\forall n$; most significantly, this holds for the languages L_{pal} and L_{eq} shown by Ambainis and Watrous [2] to be recognizable with bounded error by 2QCFA in, respectively, expected exponential time and expected polynomial time. In particular, it is easy to show that $A^{DFA}_{L_{pal},\leq}(n) = 2^{\Theta(n)}$, $A^{DFA}_{L_{eq},\leq}(n) = \Theta(n)$, and $A^{2DFA}_{L_{pal},\leq}(n) = n^{\Theta(1)}$; moreover, $A^{2DFA}_{L_{eq},\leq}(n) = \log^{\Theta(1)}(n)$ [20, Theorem 3 and Corollary 4]. In fact, this same phenomenon occurs for all the group word problems that we can show [33] are recognized by 2QCFA. Might this be true for all languages recognizable by 2QCFA?

Open Problem 7.1. *If a language L is recognizable with bounded error by a 2QCFA in expected time $T(n)$, does a stronger lower bound than $T(n) = (A^{2DFA}_{L,\leq}(n))^{\Omega(1)}$ hold?*

We have shown that the class of languages recognizable with bounded error by a 2QCFA in expected polynomial time is contained in $\mathsf{L/poly}$. This type of *dequantumization* result, which shows that the class of languages recognizable by a particular quantum model is contained in the class of languages recognizable by a particular classical model, is analogous to the Adleman-type [1] *derandomization* result $\mathsf{BPL} \subseteq \mathsf{L/poly}$. It is natural to ask if our dequantumization result might be extended, either to 2QCFA that run in a larger time bound, or to small-space QTM. Note that $\mathsf{L/poly} = \{L : A^{2DFA}_{L,=}(n) = n^{O(1)}\} = \{L : A^{2DFA}_{L,\leq}(n) = n^{O(1)}\} \supsetneq \{L : A^{DFA}_{L,\leq}(n) = n^{O(1)}\}$. This further demonstrates the value of the preceding open problem, as any improvement in the lower bound on $T(n)$ in terms of $A^{2DFA}_{L,\leq}(n)$ would directly translate into an improved dequantumization result.

The seminal paper of Lipton and Zalcstein [24] showed that, if a finitely generated group G has a faithful finite-dimensional (linear) representation over a field of characteristic 0, then $W_G \in \mathsf{L}$ (deterministic logspace). We [33] recently adapted their technique to show that 2QCFA can recognize the word problem W_G of any group G that belongs to a certain (proper) subset of the set of groups to which their result applies: any group G that has a faithful finite-dimensional *unitary* representation of a certain special type. The requirement, imposed by the laws of quantum mechanics, that the state of the quantum register of a 2QCFA must evolve unitarily, prevents a 2QCFA from (directly) implementing the Lipton-Zalcstein algorithm for any other groups; on the other hand, for those groups G that do have such a representation, these same laws allow a 2QCFA to recognize W_G using only a constant amount of space. The word problem W_G of any group G that lacks such a representation (for example, all $G \in \mathcal{G}_{\mathbf{vNilp}} \setminus \mathcal{G}_{\mathbf{vAb}}$, or any infinite Kazhdan group, or any group of intermediate growth) seems to be a plausible candidate for a hard problem for 2QCFA (see [33] for further discussion).

Open Problem 7.2. *Is there a finitely generated group G that does not have a faithful finite-dimensional projective unitary representation for which $W_G \in \mathsf{BQE2QCFA}$?*

Concerning those groups with word problem recognizable by a 2QCFA in expected polynomial time, we have shown that, if $G \in \mathcal{G}_{\mathbf{vAb}}$, then $W_G \in \mathsf{coRQP2QCFA}_{\overline{\mathbb{Q}}}(2) \subseteq \mathsf{BQP2QCFA}$ [33, Theorem 1.2]; moreover, if $W_G \in \mathsf{BQP2QCFA}$, then $G \in \mathcal{G}_{\mathbf{vNilp}}$ (Theorem 6.3(iv)). We have also shown, if $W_H \notin \mathsf{BQP2QCFA}$, where $H \in \mathcal{G}_{\mathbf{vNilp}}$ is the (three-dimensional discrete) Heisenberg group, then the classification of those groups whose word problem is recognizable by a 2QCFA in expected polynomial time would be complete; in particular, we would have $W_G \in \mathsf{BQP2QCFA} \Leftrightarrow G \in \mathcal{G}_{\mathbf{vAb}}$ (Proposition 6.4). This naturally raises the following question.

Open Problem 7.3. *Is there a group $G \in \mathcal{G}_{\mathbf{vNilp}} \setminus \mathcal{G}_{\mathbf{vAb}}$ such that $W_G \in \mathsf{BQP2QCFA}$? In particular, is $W_H \in \mathsf{BQP2QCFA}$, where H is the Heisenberg group?*

Acknowledgments

The author would like to express his sincere gratitude to Professor Michael Sipser for many years of mentorship and support, without which this work would not have been possible, as well as to thank Professor Richard Lipton for several helpful comments on an earlier draft of this paper.

References

[1] Adleman, L. Two theorems on random polynomial time. In *19th Annual Symposium on Foundations of Computer Science (sfcs 1978)* (1978), IEEE, pp. 75–83.

[2] Ambainis, A., and Watrous, J. Two-way finite automata with quantum and classical states. *Theoretical Computer Science 287*, 1 (2002), 299–311.

[3] Ambainis, A., and Yakaryilmaz, A. Automata and quantum computing. *arXiv preprint arXiv:1507.01988* (2015).

[4] Anisimov, A. V. Group languages. *Cybernetics and Systems Analysis 7*, 4 (1971), 594–601.

[5] Birget, J.-C., Ol'shanskii, A. Y., Rips, E., and Sapir, M. V. Isoperimetric functions of groups and computational complexity of the word problem. *Annals of Mathematics* (2002), 467–518.

[6] Borodin, A., Cook, S., and Pippenger, N. Parallel computation for well-endowed rings and space-bounded probabilistic machines. *Information and Control 58*, 1-3 (1983), 113–136.

[7] Condon, A., Hellerstein, L., Pottle, S., and Wigderson, A. On the power of finite automata with both nondeterministic and probabilistic states. *SIAM Journal on Computing 27*, 3 (1998), 739–762.

[8] Dunwoody, M. J. The accessibility of finitely presented groups. *Inventiones mathematicae 81*, 3 (1985), 449–457.

[9] Dwork, C., and Stockmeyer, L. A time complexity gap for two-way probabilistic finite-state automata. *SIAM Journal on Computing 19*, 6 (1990), 1011–1023.

[10] Dwork, C., and Stockmeyer, L. Finite state verifiers i: The power of interaction. *Journal of the ACM (JACM) 39*, 4 (1992), 800–828.

[11] Freivalds, R. Probabilistic two-way machines. In *International Symposium on Mathematical Foundations of Computer Science* (1981), Springer, pp. 33–45.

[12] Greenberg, A. G., and Weiss, A. A lower bound for probabilistic algorithms for finite state machines. *Journal of Computer and System Sciences 33*, 1 (1986), 88–105.

[13] Grigorchuk, R. I. Degrees of growth of finitely generated groups, and the theory of invariant means. *Mathematics of the USSR-Izvestiya 25*, 2 (1985), 259.

[14] Gromov, M. Groups of polynomial growth and expanding maps (with an appendix by jacques tits). *Publications Mathématiques de l'IHÉS 53* (1981), 53–78.

[15] Grover, L. K. A fast quantum mechanical algorithm for database search. *Proceedings of the Twenty-Eighth Annual ACM Symposium of Theory of Computing* (1996), 212–219.

[16] Harrow, A. W., Hassidim, A., and Lloyd, S. Quantum algorithm for linear systems of equations. *Physical review letters 103*, 15 (2009), 150502.

[17] Hennie, F. C. One-tape, off-line turing machine computations. *Information and Control 8*, 6 (1965), 553–578.

[18] Herbst, T. On a subclass of context-free groups. *RAIRO-Theoretical Informatics and Applications-Informatique Théorique et Applications 25*, 3 (1991), 255–272.

[19] Holt, D. F., Rees, S., Röver, C. E., and Thomas, R. M. Groups with context-free co-word problem. *Journal of the London Mathematical Society 71*, 3 (2005), 643–657.

[20] Ibarra, O. H., and Ravikumar, B. Sublogarithmic-space turing machines, nonuniform space complexity, and closure properties. *Mathematical systems theory 21*, 1 (1988), 1–17.

[21] Kaņeps, J., and Freivalds, R. Minimal nontrivial space complexity of probabilistic one-way turing machines. In *International Symposium on Mathematical Foundations of Computer Science* (1990), Springer, pp. 355–361.

[22] Karp, R. M. Some bounds on the storage requirements of sequential machines and turing machines. *Journal of the ACM (JACM) 14*, 3 (1967), 478–489.

[23] Kitaev, A. Y. Quantum computations: algorithms and error correction. *Uspekhi Matematicheskikh Nauk 52*, 6 (1997), 53–112.

[24] Lipton, R. J., and Zalcstein, Y. Word problems solvable in logspace. *Journal of the ACM (JACM) 24*, 3 (1977), 522–526.

[25] Löh, C. *Geometric group theory.* Springer, 2017.

[26] Melkebeek, D. v., and Watson, T. Time-space efficient simulations of quantum computations. *Theory of Computing 8*, 1 (2012), 1–51.

[27] Meyer, A. R., and Fischer, M. J. Economy of description by automata, grammars, and formal systems. In *12th Annual Symposium on Switching and Automata Theory (swat 1971)* (1971), IEEE, pp. 188–191.

[28] Muller, D. E., and Schupp, P. E. Groups, the theory of ends, and context-free languages. *Journal of Computer and System Sciences 26*, 3 (1983), 295–310.

[29] Nechita, I., Puchala, Z., Pawela, L., and Zyczkowski, K. Almost all quantum channels are equidistant. *Journal of Mathematical Physics 59*, 5 (2018), 052201.

[30] Nielsen, M. A., and Chuang, I. Quantum computation and quantum information, 2002.

[31] Perdrix, S., and Jorrand, P. Classically controlled quantum computation. *Mathematical Structures in Computer Science 16*, 4 (2006), 601–620.

[32] Rabin, M. O., and Scott, D. Finite automata and their decision problems. *IBM journal of research and development 3*, 2 (1959), 114–125.

[33] Remscrim, Z. The power of a single qubit: Two-way quantum/classical finite automata and the word problem for linear groups. *Electronic Colloquium on Computational Complexity (ECCC) 26* (2019), 107.

[34] Say, A., and Yakaryilmaz, A. Magic coins are useful for small-space quantum machines. *Quantum Information & Computation 17*, 11-12 (2017), 1027–1043.

[35] Shallit, J. Automaticity iv: sequences, sets, and diversity. *Journal de théorie des nombres de Bordeaux 8*, 2 (1996), 347–367.

[36] Shallit, J., and Breitbart, Y. Automaticity i: Properties of a measure of descriptional complexity. *Journal of Computer and System Sciences 53*, 1 (1996), 10–25.

[37] Shalom, Y., and Tao, T. A finitary version of gromov's polynomial growth theorem. *Geometric and Functional Analysis 20*, 6 (2010), 1502–1547.

[38] Shepherdson, J. C. The reduction of two-way automata to one-way automata. *IBM Journal of Research and Development 3*, 2 (1959), 198–200.

[39] Shor, P. W. Algorithms for quantum computation: Discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science* (1994), Ieee, pp. 124–134.

[40] Ta-Shma, A. Inverting well conditioned matrices in quantum logspace. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing* (2013), pp. 881–890.

[41] Thom, A. Convergent sequences in discrete groups. *Canadian Mathematical Bulletin 56*, 2 (2013), 424–433.

[42] Tits, J. Free subgroups in linear groups. *Journal of Algebra 20*, 2 (1972), 250–270.

[43] Watrous, J. On the complexity of simulating space-bounded quantum computations. *Computational Complexity 12*, 1-2 (2003), 48–84.

[44] Watrous, J. *The theory of quantum information.* Cambridge University Press, 2018.

[45] Wolf, J. A., et al. Growth of finitely generated solvable groups and curvature of riemannian manifolds. *Journal of differential Geometry 2*, 4 (1968), 421–446.

[46] Yakaryilmaz, A., and Say, A. C. Succinctness of two-way probabilistic and quantum finite automata. *Discrete Mathematics and Theoretical Computer Science 12*, 4 (2010), 19–40.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il