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

Zachary Remscrim
Department of Computer Science, The University of Chicago, IL, USA
remscrim@uchicago.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 $2^{O(n)}$.

We prove that their result cannot be improved upon: a 2QCFA (of any size) cannot recognize L_{pal} with bounded error in expected time $2^{o(n)}$. 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.

2012 ACM Subject Classification Theory of computation → Formal languages and automata theory; Theory of computation → Quantum computation theory

Keywords and phrases Quantum computation, Lower bounds, Finite automata

Digital Object Identifier 10.4230/LIPIcs.ITCS.2021.39

Related Version Full version of the paper `https://arxiv.org/abs/2003.09877`.

Acknowledgements 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, and to thank Professor Richard Lipton, as well as the anonymous reviewers, for several helpful comments on an earlier draft of this paper.

1 Introduction

Quantum algorithms, such as Shor's quantum polynomial time integer factorization algorithm [31], Grover's algorithm for unstructured search [16], and the linear system solver of Harrow, Hassidim, and Lloyd [17], 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. In particular, current state-of-the-art quantum computers have a very small amount of quantum memory. For example, Google's "Sycamor" quantum computer, used in their famous recent quantum supremacy experiment [5], operates on only 53 qubits.

In this paper, we study the power quantum computers that have only a small amount of memory. We begin by considering two-way finite automata 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

12th Innovations in Theoretical Computer Science Conference (ITCS 2021).
Editor: James R. Lee; Article No. 39; pp. 39:1–39:20
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

constant size. 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 [27], 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.

It is known that 2QCFA are more powerful than 2DFA and two-way probabilistic finite automata (2PFA). A 2DFA can only recognize regular languages [25]. 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 $2^{O(n)}$ [13]. However, a 2PFA cannot recognize L_{eq} with bounded error in expected time $2^{o(n)}$, by a result of Greenberg and Weiss [14]; moreover, a 2PFA cannot recognize L_{pal} with bounded error in any time bound [11]. More generally, the landmark result of Dwork and Stockmeyer [10] 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 $T(n) = \Omega\left(2^{D_L(n)^a}\right)$ [10, 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} in subexponential time, or perhaps even in 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 exponential time but not in subexponential time? These natural questions, to our knowledge, were all open (see, for instance, [2, 3, 39] for previous discussions of these questions).

In this paper, we answer these and other related questions. We first prove an analogue of the Dwork-Stockmeyer lemma for 2QCFA.

▶ **Theorem 1.** *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 $T(n) = \Omega\left(D_L(n)^a\right)$.*

This immediately implies that the result of Ambainis and Watrous [2] cannot be improved.

▶ **Corollary 2.** *2QCFA (of any size) cannot recognize L_{pal} with bounded error in time $2^{o(n)}$.*

One of the key tools used in our proof is a quantum version of Hennie's [18] 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 [10] and of Greenberg and Weiss [14]. 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.

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

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 *dequantization* 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 dequantization result was of a very different type: the class of languages recognizable by a 2QCFA, or more generally a QTM that uses $O(\log n)$ space, with algebraic number transition amplitudes (even with unbounded error and with no time bound), is contained in $\mathsf{DSPACE}(O(\log^2 n))$ [35]. This dequantization result 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 $\mathsf{DSPACE}(O(\log^2 n))$ [7].

We also investigate which group word problems can be recognized by 2QCFA or QTMs 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 [23, 9], $W_G \in \mathsf{NP} \Leftrightarrow G$ is a subgroup of a finitely presented group with polynomial Dehn function [6]. 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 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 time $2^{O(n)}$ [27].

We now show that, if G has exponential growth, then W_G cannot be recognized by a 2QCFA with bounded error in time $2^{o(n)}$, thereby providing a broad and natural class of languages that may be recognized by a 2QCFA in 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 polynomial time. Furthermore, we show analogous results for sublogarithmic-space QTMs.

2 Preliminaries

2.1 Quantum Computation

In this section, we briefly recall the fundamentals of quantum computation needed in this paper (see, for instance, [37, 24] 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 $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$. 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 $\langle |\psi\rangle, \cdot\rangle : V \to \mathbb{C}$. We write $\langle\psi|\varphi\rangle$ to denote $\langle |\psi\rangle, |\varphi\rangle\rangle$. 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 $|\rho\rangle \in V$, $|\psi\rangle\langle\varphi|(|\rho\rangle) = |\psi\rangle\langle\varphi|\rho\rangle = \langle\varphi|\rho\rangle|\psi\rangle$. Let $\mathbb{1}_V \in \mathrm{L}(V)$ denote the identity operator on V and let $\mathbb{0}_V \in \mathrm{L}(V)$ denote the zero operator on 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 $\mathrm{Herm}(V) = \{A \in \mathrm{L}(V) : A = A^\dagger\}$, $\mathrm{Pos}(V) = \{A^\dagger A : A \in \mathrm{L}(V)\}$, $\mathrm{Proj}(V) = \{A \in \mathrm{Pos}(V) : A^2 = A\}$, $\mathrm{U}(V) = \{A \in \mathrm{L}(V) : AA^\dagger = \mathbb{1}_V\}$, and $\mathrm{Den}(V) = \{A \in \mathrm{Pos}(V) : \mathrm{Tr}(A) = 1\}$ denote, respectively, the set of Hermitian, positive semi-definite, projection, unitary, and 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 $\mathbb{C}^Q \cong \mathbb{C}^k$. The quantum register stores a *superposition* $|\psi\rangle = \sum_q \alpha_q |q\rangle \in \mathbb{C}^Q$, where each $\alpha_q \in \mathbb{C}$ and $\sum_q |\alpha_q|^2 = 1$; in other words, a superposition $|\psi\rangle$ is simply an element of $\mathbb{C}^Q$ of norm 1.

Following the original definition of Ambainis and Watrous [2], 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 $P_iP_j = \mathbb{0}_{\mathbb{C}^Q}, \forall i, j$ with $i \neq j$, and $\sum_j P_j = \mathbb{1}_{\mathbb{C}^Q}$. Quantum measurement is a probabilistic process where, if the quantum register is 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 quantum measurement changes the state of the quantum register.

An *ensemble of pure states* of the quantum register is a set $\{(p_i, |\psi_i\rangle) : i \in I\}$, for some index set I, where $p_i \in [0, 1]$ denotes the probability of the quantum register being in the superposition $|\psi_i\rangle$, 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, [24, 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, when performing the von Neumann measurement specified by some $P_1, \ldots, P_l \in \mathrm{Proj}(\mathbb{C}^Q)$, the probability that the result of this measurement is r is given by $\mathrm{Tr}(P_rAP_r^\dagger)$, and if the result is r then the ensemble collapses to an ensemble described by the density operator $\frac{1}{\mathrm{Tr}(P_rAP_r^\dagger)}P_rAP_r^\dagger$.

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 of the form $\Phi : \mathrm{L}(V) \to \mathrm{L}(V')$. Define $\mathrm{T}(V) = \mathrm{T}(V, V)$ and let $\mathbb{1}_{\mathrm{L}(V)} \in \mathrm{T}(V)$ 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, $\Phi \otimes \mathbb{1}_{\mathrm{L}(W)}$ 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.* 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 $\mathrm{Chan}(V) = \mathrm{Chan}(V, V)$.

As we wish for our lower bound to be a strong as possible, we wish to consider a variant of the 2QCFA model that is as strong as possible; in particular, we will allow a 2QCFA to perform any physically realizable quantum operation on its quantum register.

Following Watrous [35], a *selective quantum operation* $\mathcal{E}$ is specified by a set of operators $\{E_{r,j} : r \in R, j \in \{1, \ldots, l\}\} \subseteq \mathrm{L}(\mathbb{C}^Q)$, where R is a finite set and $l \in \mathbb{N}_{\geq 1}$ (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, etc.), such that $\sum_{r,j} E_{r,j}^{\dagger} E_{r,j} = \mathbb{1}_{\mathbb{C}^Q}$. For $r \in R$, we define $\Phi_r \in \mathrm{T}(\mathbb{C}^Q)$ such that, $\Phi_r(A) = \sum_j E_{r,j} A E_{r,j}^{\dagger}$, $\forall A \in \mathrm{L}(V)$. Then, if the quantum register is described by some density operator $A \in \mathrm{Den}(\mathbb{C}^Q)$, applying $\mathcal{E}$ will have result $r \in R$ with probability $\mathrm{Tr}(\Phi_r(A))$; if the result is r, then the quantum register is described by density operator $\frac{1}{\mathrm{Tr}(\Phi_r(A))} \Phi_r(A)$. Both unitary transformations and von Neumann measurements are special cases of selective quantum operations. For any $\mathcal{E}$, one may always obtain a family of operators that represent $\mathcal{E}$ with $l \leq |Q|^2$ [37, Theorem 2.22], and therefore with $l = |Q|^2$ (by defining any extraneous operators to be $\mathbb{0}_{\mathbb{C}^Q}$). Let $\mathrm{QuantOp}(\mathbb{C}^Q, R)$ denote the set of all selective quantum operations specified by some $\{E_{r,j} : r \in R, j \in \{1, \ldots, |Q|^2\}\} \subseteq \mathrm{L}(\mathbb{C}^Q)$.

2.2 Definition of the 2QCFA Model

Next, we define two-way finite automata with quantum and classical states (2QCFA), essentially following the original definition of Ambainis and Watrous [2], with a few alterations that (potentially) make the model stronger. We wish to define the 2QCFA model to be as strong as possible so that our lower bounds against this model are as general as possible.

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, $N = (Q, C, \Sigma, R, \theta, \delta, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$, where Q is a finite set of quantum basis states, C is a finite set of classical states, Σ is a finite input alphabet, R is a finite set that specifies the possible results of selective quantum operations, θ and δ are the quantum and classical parts of the transition function, $q_{\text{start}} \in Q$ is the quantum start state, $c_{\text{start}} \in C$ is the classical start state, and $c_{\text{acc}}, c_{\text{rej}} \in C$, with $c_{\text{acc}} \neq c_{\text{rej}}$, 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 $\widehat{C} = C \setminus \{c_{\text{acc}}, c_{\text{rej}}\}$ denote the non-halting classical states. The components of the transition function are as follows: $\theta : \widehat{C} \times \Sigma_+ \to \mathrm{QuantOp}(\mathbb{C}^Q, R)$ specifies the selective quantum operation that is to be performed on the quantum register and $\delta : \widehat{C} \times \Sigma_+ \times R \to C \times \{-1, 0, 1\}$ specifies how the classical state and (classical) head position evolve.

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 $|q_{\text{start}}\rangle$, 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, the selective quantum operation $\theta(c, \sigma)$ is performed on the quantum register producing some result $r \in R$. If the result was r, and $\delta(c, \sigma, r) = (c', d)$, where $c' \in C$ and $d \in \{-1, 0, 1\}$, then the classical state becomes c' and the head moves left (resp. stays put, moves right) if $d = -1$ (resp. $d = 0$, $d = 1$).

Due to the fact that applying a selective quantum operation is a probabilistic process, the computation of N on an input w is probabilistic. We say that a 2QCFA N recognizes a language L with *two-sided bounded error* ϵ if, $\forall w \in L$, $\Pr[N \text{ accepts } w] \geq 1 - \epsilon$, and, $\forall w \notin L$, $\Pr[N \text{ accepts } w] \leq \epsilon$. We then 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).

3 2QCFA Crossing Sequences

In this section, we develop a generalization of Hennie's [18] notion of crossing sequences to 2QCFA, in which we make use of several ideas from the 2PFA results of Dwork and Stockmeyer [10] and Greenberg and Weiss [14]. This notion will play a key role in our proof of a lower bound on the expected running time of a 2QCFA.

When a 2QCFA $N = (Q, C, \Sigma, R, \theta, \delta, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$ is run on an input $w = w_1 \cdots w_n \in \Sigma^*$, where each $w_i \in \Sigma$, the tape consists of $\#_L w_1 \cdots w_n \#_R$. One may describe the configuration of *a single probabilistic branch* of N at any particular point in time by a triple (A, c, h), where $A \in \text{Den}(\mathbb{C}^Q)$ describes the current state of the quantum register, $c \in C$ is the current classical state, and $h \in \{0, \ldots, n+1\}$ is the current head position. To clarify, each step of the computation of N involves applying a selective quantum operation, which is a probabilistic process that produces a particular result $r \in R$ with a certain probability (depending on the operation that is performed and the state of the quantum register); that is to say, the 2QCFA probabilistically branches, with a child for each $r \in R$.

We partition the input as $w = xy$, in some manner to be specified later. We then imagine running N beginning in the configuration $(A, c, |x|)$, where $|x|$ denotes the length of the string x (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 $\#_L x$ (by moving 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 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 $\{(p_i, (A_i, c_i, |x|)) : i \in I\}$, where the probability of being in configuration $(A_i, c_i, |x|)$ 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]. Much as it was the case that an ensemble of pure states of a quantum register can be described by a density operator, we may also describe an ensemble of configurations of a 2QCFA using density operators. This will greatly simplify our definition and analysis of the crossing sequence of a 2QCFA.

3.1 Describing Ensembles of Configurations of 2QCFA

The 2QCFA N posseses both a constant-sized *quantum register*, that is described by some density operator at any particular point in time, 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

[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 [10] in their 2PFA result.

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 $h \in H_x = \{0, \dots, |x|+1\}$ (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 $|h\rangle \in \mathbb{C}^{H_x}$. A configuration $(A, c, h) \in \mathrm{Den}(\mathbb{C}^Q) \times C \times H_x$ is then simply a state of the *combined register*, which consists of the quantum, classical, and head registers.

We then consider an *ensemble of configurations* $\{(p_i, (A_i, c_i, h_i)) : i \in I\}$, where p_i denotes the probability of being in configuration (A_i, c_i, h_i). We represent this ensemble (non-uniquely) by the density operator $Z = \sum_i \big(p_i A_i \otimes |c_i\rangle\langle c_i| \otimes |h_i\rangle\langle h_i|\big) \in \mathrm{Den}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$. Let $\widehat{i}(c,h) = \{i \in I : (c_i, h_i) = (c,h)\}$ denote the indices of those configurations in classical state c and with head position h. We then define $p : C \times H_x \to [0,1]$ such that $p(c,h) = \sum_{i \in \widehat{i}(c,h)} p_i$ is the total probability of being in classical state c and having head position h. We define $A : C \times H_x \to \mathrm{Den}(\mathbb{C}^Q)$ such that, if $p(c,h) \neq 0$, then $A(c,h) = \sum_{i \in \widehat{i}(c,h)} \frac{p_i}{p(c,h)} A_i$ is the density operator obtained by "merging" all density operators A_i that come from configurations (A_i, c_i, h_i) with classical state $c_i = c$ and head position $h_i = h$; if $p(c,h) = 0$, then we define $A(c,h)$ arbitrarily. Then $Z = \sum_{c,h} \big(p(c,h)A(c,h) \otimes |c\rangle\langle c| \otimes |h\rangle\langle h|\big)$. 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 also consider the case in which we are only interested in the states of the quantum and classical registers, but not the head position. We then analogously describe an ensemble $\{(p_i, (A_i, c_i)) : i \in I\}$ by $Z = \sum_i \big(p_i A_i \otimes |c_i\rangle\langle c_i|\big) \in \mathrm{Den}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, and we define $\widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ to be the set of all such density operators. In a starting ensemble, all configurations have the same head position: $|x|$. We define $I_x \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C, \mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ such that $I_x(Z) = Z \otimes ||x|\rangle\langle |x||$. 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 irrelevant). Let $\mathrm{Tr}_{\mathbb{C}^{H_x}} = \mathbb{1}_{\mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)} \otimes \mathrm{Tr} \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}, \mathbb{C}^Q \otimes \mathbb{C}^C)$ denote the *partial trace with respect to* $\mathbb{C}^{H_x}$.

3.2 Definition and Properties of 2QCFA Crossing Sequences

We now formally define the notion of a crossing sequence of a 2QCFA and prove certain needed properties. We begin by establishing some notation.

▶ **Definition 3.** Consider a 2QCFA $N = (Q, C, \Sigma, R, \theta, \delta, q_{\mathrm{start}}, c_{\mathrm{start}}, c_{\mathrm{acc}}, c_{\mathrm{rej}})$. For $c \in \widehat{C} = C \setminus \{c_{\mathrm{acc}}, c_{\mathrm{rej}}\}$, $\sigma \in \Sigma_+ = \Sigma \sqcup \{\#_L, \#_R\}$, $r \in R$, and $j \in J = \{1, \dots, |Q|^2\}$, we make the following definitions.

(i) Define $E_{c,\sigma,r,j} \in \mathrm{L}(\mathbb{C}^Q)$ such that $\theta(c,\sigma) \in \mathrm{QuantOp}(\mathbb{C}^Q, R)$ is described by $\{E_{c,\sigma,r,j} : r \in R, j \in J\}$.

(ii) Define $\Phi_{c,\sigma,r} \in \mathrm{T}(\mathbb{C}^Q)$ such that $\Phi_{c,\sigma,r}(A) = \sum_j E_{c,\sigma,r,j} A E^\dagger_{c,\sigma,r,j}$, $\forall A \in \mathrm{L}(\mathbb{C}^Q)$.

(iii) Let $\gamma_{c,\sigma,r} \in C$ and $d_{c,\sigma,r} \in \{-1, 0, 1\}$ denote, respectively, the new classical state and the motion of the head, if the result of applying $\theta(c,\sigma)$ is r; i.e., $\delta(c,\sigma,r) = (\gamma_{c,\sigma,r}, d_{c,\sigma,r})$.

Consider some $x \in \Sigma^*$. Let $\widehat{H}_x = \{0, \dots, |x|\}$ denote the head positions corresponding to the prefix $\#_L x$, and let $H_x = \{0, \dots, |x|+1\}$ denote the set of possible positions the head of N may be in until it "finishes computing" on the prefix $\#_L x$. We define an operator $S_x \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ that describes a single step of the computation of N on $\#_L x$, as follows. If $(c,h) \in \widehat{C} \times \widehat{H}_x$, then $S_x(A \otimes |c\rangle\langle c| \otimes |h\rangle\langle h|)$ describes the ensemble of configurations of N after running N for a single step beginning in the configuration (A, c, h);

otherwise (i.e., if $c \in \{c_{\text{acc}}, c_{\text{rej}}\}$ or $h = |x| + 1$, which means N has "finished computing" on $\#_L x$) S_x leaves the configuration unchanged. We will observe that S_x correctly describes the behavior of N on an ensemble of configurations, and that S_x is a quantum channel.

▶ **Definition 4.** Using the notation of Definition 3, consider a 2QCFA N and a string $x \in \Sigma^*$. Let $x_h \in \Sigma$ denote the symbol of x at position h, and let $x_0 = \#_L$ denote the left end-marker.

(i) For $(c, h, r, j) \in C \times H_x \times R \times J$, define $\widetilde{E}_{x,c,h,r,j} \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ as follows.

$$\widetilde{E}_{x,c,h,r,j} = \begin{cases} E_{c,x_h,r,j} \otimes |\gamma_{c,x_h,r}\rangle \langle c| \otimes |h + d_{c,x_h,r}\rangle \langle h|, & \text{if } (c,h) \in \widehat{C} \times \widehat{H} \\ \frac{1}{\sqrt{|R||J|}} \mathbb{1}_{\mathbb{C}^Q} \otimes |c\rangle \langle c| \otimes |h\rangle \langle h|, & \text{otherwise.} \end{cases}$$

(ii) Define $S_x \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ such that

$$S_x(Z) = \sum_{(c,h,r,j) \in C \times H_x \times R \times J} \widetilde{E}_{x,c,h,r,j} Z \widetilde{E}^{\dagger}_{x,c,h,r,j}, \quad \forall Z \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}).$$

▶ **Lemma 5.** *Using the above notation, consider some $x \in \Sigma^*$ and $(A, \widehat{c}, \widehat{h}) \in \mathrm{Den}(\mathbb{C}^Q) \times \widehat{C} \times \widehat{H}_x$. Let $\widehat{Z} = A \otimes |\widehat{c}\rangle \langle \widehat{c}| \otimes \left|\widehat{h}\right\rangle \left\langle \widehat{h}\right|$. $S_x(\widehat{Z})$ describes the ensemble of configurations obtained after running N for one step, beginning in the configuration $(A, \widehat{c}, \widehat{h})$, on input prefix $\#_L x$.*

Proof. Let $\widetilde{R}_{x,\widehat{c},\widehat{h},A} = \{r \in R : \mathrm{Tr}(\Phi_{\widehat{c},x_{\widehat{h}},r}(A)) \neq 0\}$. Note that $A \in \mathrm{Den}(\mathbb{C}^Q) \subseteq \mathrm{Pos}(\mathbb{C}^Q)$, which implies $\Phi_{\widehat{c},x_{\hat{h}},r}(A) \in \mathrm{Pos}(\mathbb{C}^Q)$; therefore, we have $\mathrm{Tr}(\Phi_{\widehat{c},x_{\hat{h}},r}(A)) = 0$ precisely when $\Phi_{\widehat{c},x_{\hat{h}},r}(A) = \mathbb{0}_{\mathbb{C}^Q}$. After running N as described, it is in an ensemble of configurations

$$\left\{ \left(\mathrm{Tr}(\Phi_{\widehat{c},x_{\hat{h}},r}(A)), \left(\frac{1}{\mathrm{Tr}(\Phi_{\widehat{c},x_{\hat{h}},r}(A))} \Phi_{\widehat{c},x_{\hat{h}},r}(A), \gamma_{\widehat{c},x_{\hat{h}},r}, \widehat{h} + d_{\widehat{c},x_{\hat{h}},r} \right) \right) : r \in \widetilde{R}_{x,\widehat{c},\widehat{h},A} \right\}.$$

This ensemble of configurations is described by the density operator $\widehat{Z}'$ given by

$$\widehat{Z}' = \sum_{r \in \widetilde{R}_{x,\widehat{c},\widehat{h},A}} \left(\frac{\mathrm{Tr}(\Phi_{\widehat{c},x_{\hat{h}},r}(A))}{\mathrm{Tr}(\Phi_{\widehat{c},x_{\hat{h}},r}(A))} \Phi_{\widehat{c},x_{\hat{h}},r}(A) \otimes \left|\gamma_{\widehat{c},x_{\hat{h}},r}\right\rangle \left\langle \gamma_{\widehat{c},x_{\hat{h}},r}\right| \otimes \left|\widehat{h} + d_{\widehat{c},x_{\hat{h}},r}\right\rangle \left\langle \widehat{h} + d_{\widehat{c},x_{\hat{h}},r}\right| \right)$$

$$= \sum_{r \in R} \left(\Phi_{\widehat{c},x_{\hat{h}},r}(A) \otimes \left|\gamma_{\widehat{c},x_{\hat{h}},r}\right\rangle \left\langle \gamma_{\widehat{c},x_{\hat{h}},r}\right| \otimes \left|\widehat{h} + d_{\widehat{c},x_{\hat{h}},r}\right\rangle \left\langle \widehat{h} + d_{\widehat{c},x_{\hat{h}},r}\right| \right).$$

Let $B_{x,\widehat{c},\widehat{h},r} = \left|\gamma_{\widehat{c},x_{\hat{h}},r}\right\rangle \left\langle \gamma_{\widehat{c},x_{\hat{h}},r}\right| \otimes \left|\widehat{h} + d_{\widehat{c},x_{\hat{h}},r}\right\rangle \left\langle \widehat{h} + d_{\widehat{c},x_{\hat{h}},r}\right|$. If $(c,h) \in \widehat{C} \times \widehat{H}_x$, then

$$\widetilde{E}_{x,c,h,r,j} \widehat{Z} \widetilde{E}^{\dagger}_{x,c,h,r,j} = \widetilde{E}_{x,c,h,r,j} \left(A \otimes |\widehat{c}\rangle \langle \widehat{c}| \otimes \left|\widehat{h}\right\rangle \left\langle \widehat{h}\right| \right) \widetilde{E}^{\dagger}_{x,c,h,r,j}$$

$$= E_{c,x_h,r,j} A E^{\dagger}_{c,x_h,r,j} \otimes |\gamma_{c,x_h,r}\rangle \langle c|\widehat{c}\rangle \langle \widehat{c}|c\rangle \langle \gamma_{c,x_h,r}| \otimes |h + d_{c,x_h,r}\rangle \left\langle h|\widehat{h}\right\rangle \left\langle \widehat{h}|h\right\rangle \langle h + d_{c,x_h,r}|$$

$$= \begin{cases} E_{\widehat{c},x_{\hat{h}},r,j} A E^{\dagger}_{\widehat{c},x_{\hat{h}},r,j} \otimes B_{x,\widehat{c},\widehat{h},r}, & \text{if } (c,h) = (\widehat{c}, \widehat{h}) \\ \mathbb{0}_{\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}}, & \text{otherwise.} \end{cases}$$

If, instead, $(c,h) \notin \widehat{C} \times \widehat{H}_x$, then $\widetilde{E}_{x,c,h,r,j} \widehat{Z} \widetilde{E}^{\dagger}_{x,c,h,r,j} = \mathbb{0}_{\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}}$. Therefore

$$S_x(\widehat{Z}) = \sum_{(r,j) \in R \times J} \sum_{(c,h) \in C \times H_x} \widetilde{E}_{x,c,h,r,j} \widehat{Z} \widetilde{E}^{\dagger}_{x,c,h,r,j} = \sum_{(r,j) \in R \times J} \left(E_{\widehat{c},x_{\hat{h}},r,j} A E^{\dagger}_{\widehat{c},x_{\hat{h}},r,j} \otimes B_{x,\widehat{c},\widehat{h},r} \right)$$

$$= \sum_{r \in R} \left(\left(\sum_{j \in J} E_{\widehat{c},x_{\hat{h}},r,j} A E^{\dagger}_{\widehat{c},x_{\hat{h}},r,j} \right) \otimes B_{x,\widehat{c},\widehat{h},r} \right) = \sum_{r \in R} \left(\Phi_{\widehat{c},x_{\hat{h}},r}(A) \otimes B_{x,\widehat{c},\widehat{h},r} \right) = \widehat{Z}'. \qquad ◀$$

▶ **Lemma 6.** *Consider some $x \in \Sigma^*$ and $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$. If $\{(p_i, (A_i, c_i, h_i)) : i \in I\}$ is some ensemble of configurations described by Z, then $S_x(Z)$ describes the ensemble of configurations obtained by replacing each configuration with $(c_i, h_i) \in (\widehat{C} \times \widehat{H}_x)$ by the ensemble (scaled by p_i) of configurations obtained by running N for one step beginning in the configuration (A_i, c_i, h_i), and leaving each configuration with $(c_i, h_i) \notin (\widehat{C} \times \widehat{H}_x)$ unchanged.*

Proof. This follows immediately from Lemma 5 and linearity. ◀

▶ **Lemma 7.** *$S_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$, $\forall x \in \Sigma^*$.*

Proof. $\{\widetilde{E}_{x,c,h,r,j} : (c, h, r, j) \in C \times H_x \times R \times J\}$ is a *Kraus representation* of S_x; therefore, $S_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \Leftrightarrow \sum\limits_{c,h,r,j} \widetilde{E}^{\dagger}_{x,c,h,r,j} \widetilde{E}_{x,c,h,r,j} = \mathbb{1}$ [37, Corollary 2.27]. This latter statement follows from a straightforward calculation; see the full paper [26] for a proof. ◀

For $m \in \mathbb{N}$, we define the *m-truncated stopping ensemble* as the ensemble of configurations that N will be in when it "finishes computing" on $\#_L x$, as defined earlier, with the modification that if any particular branch of N runs for more than m steps, the computation of that branch will be "interrupted" immediately before it attempts to perform the $m + 1^{\text{st}}$ step and instead immediately reject. To be clear, this truncation occurs only in the *analysis* of N; we do not modify the 2QCFA. The following truncation operator T_x, which terminates all branches on which N has not yet "finished computing," will help us do this.

▶ **Definition 8.** For $(c, h) \in (C, H_x)$, let $\widehat{E}_{x,c,h} = \mathbb{1}_{\mathbb{C}^Q} \otimes |c'\rangle \langle c| \otimes |h\rangle \langle h|$, where $c' = c_{\text{rej}}$ if $(c, h) \in \widehat{C} \times \widehat{H}_x$, and $c' = c$ otherwise. We then define $T_x \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ such that $T_x(Z) = \sum_{(c,h) \in C \times H_x} \widehat{E}_{x,c,h} Z \widehat{E}^{\dagger}_{x,c,h}$.

▶ **Lemma 9.** *Using the above notation, 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 $\{(p_i, (A_i, c_i, h_i)) : i \in I\}$ is any ensemble of configurations described by Z, then $T_x(Z)$ describes the ensemble of configurations in which each configuration with $(c_i, h_i) \in \widehat{C} \times \widehat{H}_x$ is replaced by the configuration (A_i, c_{rej}, h_i) (i.e., all configurations in which N has not yet "finished computing" on $\#_L x$ become rejecting configurations) and all other configurations are left unchanged.*

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

Proof.

(i) Immediate from definitions.

(ii) As in the proof of Lemma 7, we may straightforwardly show $\sum_{c,h} \widehat{E}^{\dagger}_{x,c,h} \widehat{E}_{x,c,h} = \mathbb{1}_{\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}}$, which implies $T_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ [37, Corollary 2.27]. ◀

The following operator converts starting ensembles to m-truncated stopping ensembles.

▶ **Definition 10.** For $x \in \Sigma^*$ and $m \in \mathbb{N}$, we define the *m-truncated transfer operator* $N^{\hookrightarrow}_{x,m} = \mathrm{Tr}_{\mathbb{C}^{H_x}} \circ T_x \circ S^m_x \circ I_x \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. For $y \in \Sigma^*$, we next consider the "dual case" of running N on the suffix $y\#_R$ beginning in some ensemble of configurations $\{(p_i, (A_i, c_i, |x| + 1)) : i \in I\}$ (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 runs for more than m steps. We then define $N^{\hookleftarrow}_{y,m} \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ as the corresponding "dual" m-truncated transfer operator for y.

▶ **Lemma 11.** *Using the notation of Definition 10, the following statements hold.*

(i) *For $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, if N is run on $\#_L x$ 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) *For $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, if N is run on $y\#_R$ beginning in any ensemble of configurations described by $I_{x+1}(Z)$, 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) Immediate by Definition 10, Lemma 6, and Lemma 9(i).

(ii) Immediate by Definition 10, and analogous versions of Lemma 6, and Lemma 9(i).

(iii) By definition, $N^{\hookrightarrow}_{x,m} = \mathrm{Tr}_{\mathbb{C}^{H_x}} \circ T_x \circ S_x^m \circ I_x$. By Lemma 7 and Lemma 9(ii), we have $S_x, T_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$. 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, [37, Section 2.2]). The claim for $N^{\hookleftarrow}_{y,m}$ follows by an analogous argument. ◀

Given a 2QCFA N, we produce an equivalent N' of a certain convenient form, in much the same way that Dwork and Stockmeyer [10] converted a 2PFA to 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 start state of N'. On any input, N' will move its head to the right until it reaches $\#_R$, performing the trivial transformation to its quantum register along the way. When it reaches $\#_R$, N' will enter c'; then, N' will move its head to the left until it reaches $\#_L$, again performing the trivial transformation to its quantum register. When it reaches $\#_L$, N' will enter the original start state c_{start} and behave identically to N from this point. For the remainder of the paper, we assume all 2QCFA have this form.

Finally, we define the *m-truncated crossing sequence*.

▶ **Definition 12.** For $x, y \in \Sigma^*$ and $m \in \mathbb{N}$, the *m-truncated crossing sequence* of N with respect to the (partitioned) input xy is 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,

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

▶ Remark. 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 $y\#_R$.

4 Lower Bounds on the Running Time of 2QCFA

Dwork and Stockmeyer proved a lower bound [10, Lemma 4.3] on the expected running time $T(n)$ of any 2PFA that recognizes any language L with bounded error, 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 [14], which showed that 2PFA cannot recognize L_{eq} in subexponential time.

4.1 Nonregularity

For a language L, Dwork and Stockmeyer [10] defined a particular "hardness measure" $D_L : \mathbb{N} \to \mathbb{N}$, which they called the *nonregularity* of L, as follows. Let Σ be a finite alphabet, $L \subseteq \Sigma^*$ a language, and $n \in \mathbb{N}$. 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 $\exists y \in \Sigma^{\leq n-\max(|x|,|x'|}$, 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 $D_L(n)$ to be 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., $x_i \not\sim_{L,n} x_j$, $\forall i,j$ with $i \neq j$).

In fact, D_L has been defined by many authors, both before and after Dwork and Stockmeyer, who gave many different names to this quantity and who (repeatedly) rediscovered certain basic facts about it; we refer the reader to the excellent paper of Shallit and Breitbart [30] for a detailed history of the study of D_L and related hardness measures.

4.2 A 2QCFA Analogue of the Dwork-Stockmeyer Lemma

We now prove that an analogue of the Dwork-Stockmeyer lemma holds for 2QCFA. The main idea is as follows. Suppose the 2QCFA N recognizes $L \subseteq \Sigma^*$, with two-sided bounded error ϵ, in expected time at most $T(n)$. We show that, if $D_L(n)$ is "large," then, for any $m \in \mathbb{N}$, we can find $x, x' \in \Sigma^{\leq n}$ such that $x \not\sim_{L,n} x'$ and the distance between the corresponding m-truncated transfer operators $N^{\hookrightarrow}_{x,m}$ and $N^{\hookrightarrow}_{x',m}$ is "small." By definition, $\exists y \in \Sigma^{\leq n-\max(|x|,|x'|)}$, 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 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 partitioned inputs xy and $x'y$ will be similar; in particular, if $T(n)$ is "small," 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.

Recall that the *trace norm* $\|\cdot\|_1 : \mathrm{L}(V) \to \mathbb{R}_{\geq 0}$ is given by $\|Z\|_1 = \mathrm{Tr}(\sqrt{Z^\dagger Z})$, $\forall Z \in \mathrm{L}(V)$, and the *induced trace norm* $\|\cdot\|_1 : \mathrm{T}(V,V') \to \mathbb{R}_{\geq 0}$, is given $\|\Phi\|_1 = \sup\{\|\Phi(Z)\|_1 : Z \in \mathrm{L}(V), \|Z\|_1 \leq 1\}$, $\forall \Phi \in \mathrm{T}(V,V')$. Suppose N is run on two distinct partitioned inputs xy and $x'y$, producing two distinct m-truncated crossing sequences, following Definition 12. We first show that if $\|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1$ is "small", then these crossing sequences are similar.

▶ **Lemma 13.** *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 $\|Z_i - Z'_i\|_1 \leq \lfloor \frac{i-1}{2} \rfloor \|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1$, $\forall i \in \mathbb{N}_{\geq 1}$.*

Proof. 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 $\|Z_1 - Z'_1\|_1 = 0$. Note that $\|\Phi(Z)\|_1 \leq \|Z\|_1$, $\forall Z \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, $\forall \Phi \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ [37, 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

$\|\Phi(Z)-\Phi(Z')\|_1 = \|\Phi(Z-Z')\|_1 \leq \|Z-Z'\|_1$. By Lemma 11(iii), $N^{\hookrightarrow}_{x,m}, N^{\hookrightarrow}_{x',m}, N^{\hookleftarrow}_{y,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. For i even, $Z_i = N^{\hookleftarrow}_{y,m}(Z_{i-1})$ and $Z'_i = N^{\hookleftarrow}_{y,m}(Z'_{i-1})$. We then have

$$\|Z_i - Z'_i\|_1 = \|N^{\hookleftarrow}_{y,m}(Z_{i-1}) - N^{\hookleftarrow}_{y,m}(Z'_{i-1})\|_1 \leq \|Z_{i-1} - Z'_{i-1}\|_1.$$

For odd $i > 1$, $Z_i = N^{\hookrightarrow}_{x,m}(Z_{i-1})$ and $Z'_i = N^{\hookrightarrow}_{x',m}(Z'_{i-1})$. We have $\|Z\|_1 = 1$, $\forall Z \in \mathrm{Den}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, which implies $\|\Phi(Z)\|_1 \leq \|\Phi\|_1$, $\forall \Phi \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. Therefore,

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

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

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

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

▶ **Lemma 14.** *Consider a language $L \subseteq \Sigma^*$. Suppose $L \in \mathsf{B2QCFA}(k, d, T(n), \epsilon)$, for some $k, d \in \mathbb{N}_{\geq 2}$, $T : \mathbb{N} \to \mathbb{N}$, and $\epsilon \in [0, \frac{1}{2})$. If, for some $n \in \mathbb{N}$, $\exists x, x' \in \Sigma^{\leq n}$ such that $x \not\sim_{L,n} x'$, then $T(n) \geq \frac{(1-2\epsilon)^2}{2} \|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1^{-1}$, $\forall m \geq \lceil \frac{2}{1-2\epsilon} T(n) \rceil$.*

Proof. By definition, $x \not\sim_{L,n} x'$ precisely when $\exists y \in \Sigma^*$ such that $xy, x'y \in \Sigma^{\leq n}$, 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 $c \in C$, let $E_c = \mathbb{1}_{\mathbb{C}^Q} \otimes |c\rangle\langle c| \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. For $s \in \mathbb{N}_{\geq 1}$, define $p_{m,s}, p'_{m,s} : C \to [0,1]$ such that $p_{m,s}(c) = \mathrm{Tr}(E_c Z_{m,s} E_c^\dagger)$ and $p'_{m,s}(c) = \mathrm{Tr}(E_c Z'_{m,s} E_c^\dagger)$. Then, for any $c \in C$, Lemma 13 implies

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

$$\leq \|Z_{m,s} - Z'_{m,s}\|_1 \leq \frac{s-1}{2} \|N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}\|_1.$$

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 runs for more than m steps between consecutive boundary crossings is forced to halt and reject immediately before attempting to perform the $m+1^{\mathrm{st}}$ such step. 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.

Note that $x'y \notin L$ implies $p_N(x'y) \leq \epsilon$. 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 steps (between consecutive crossings) are considered to reject the input in the m-truncated crossing sequence. Suppose $s \leq m$. Any branch that runs for a total of at most s steps before halting is unaffected by m-truncation. Moreover, if a branch accepts within s steps, it will certainly accept within s crossings between $\#_L x$ and $y\#_R$. This implies $p_N(xy, s) \leq p_{m,s}(c_{\mathrm{acc}})$. Therefore, if $s \leq m$,

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

The expected running time of N on input xy is at most $T(|xy|)$. By Markov's inequality, $1 - h_N(xy, s) \leq \frac{T(|xy|)}{s}$. Note that $xy \in L$ implies $p_N(xy) \geq 1 - \epsilon$. Thus, for any $m \geq s \geq 1$,

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

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 $m \geq s$,

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

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

▶ **Lemma 15.** *Consider a 2QCFA $N = (Q, C, \Sigma, R, \theta, \delta, q_{start}, c_{start}, c_{acc}, c_{rej})$. Let $k = |Q|$ and $d = |C|$. 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 $\| N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m} \|_1 \leq 4\sqrt{2} k^4 d^2 \left(|X|^{\frac{1}{k^4 d^2}} - 1 \right)^{-1}$.*

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 $J(\Phi) = \sum_{(q,q',c,c') \in Q^2 \times C^2} F_{q,q',c,c'} \otimes \Phi(F_{q,q',c,c'}), \forall \Phi \in \mathrm{T}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. Consider any $x \in \Sigma^*$ and $m \in \mathbb{N}$. We first show that, if $(c_1, c_2) \neq (c_1', c_2')$, then $\langle q_2 c_2 | \, N^{\hookrightarrow}_{x,m}(F_{q_1,q_1',c_1,c_1'}) \, | q_2' c_2' \rangle = 0$. To see this, recall that, by Definition 10, $N^{\hookrightarrow}_{x,m} = \mathrm{Tr}_{\mathbb{C}^{H_x}} \circ T_x \circ S_x^m \circ I_x$. If $c_1 \neq c_1'$, then $N^{\hookrightarrow}_{x,m}(F_{q_1,q_1',c_1,c_1'}) = \mathbb{0}_{\mathbb{C}^Q \otimes \mathbb{C}^C}$, which implies $\langle q_2 c_2 | \, N^{\hookrightarrow}_{x,m}(F_{q_1,q_1',c_1,c_1'}) \, | q_2' c_2' \rangle = 0$. If $c_2 \neq c_2'$, then $\forall Z \in L(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$, $\langle q_2 c_2 | \, \mathrm{Tr}_{\mathbb{C}^{H_x}}(T_x(Z)) \, | q_2' c_2' \rangle = 0$, which implies $\langle q_2 c_2 | \, N^{\hookrightarrow}_{x,m}(F_{q_1,q_1',c_1,c_1'}) \, | q_2' c_2' \rangle = 0$.

Therefore, $\langle q_2 c_2 | \, N^{\hookrightarrow}_{x,m}(F_{q_1,q_1',c_1,c_1'}) \, | q_2' c_2' \rangle$ is only potentially non-zero at the $k^4 d^2$ elements where $(c_1, c_2) = (c_1', c_2')$. By Lemma 11(iii), $N^{\hookrightarrow}_{x,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, which implies $J(N^{\hookrightarrow}_{x,m}) \in \mathrm{Pos}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^Q \otimes \mathbb{C}^C)$ [37, Corollary 2.27]. Therefore, the elements where $(q_1, q_2) \neq (q_1', q_2')$ come in conjugate pairs, and the elements with $(q_1, q_2) \neq (q_1', q_2')$ are real. We define the function $g_{N,m} : \Sigma^* \to \mathbb{R}^{k^4 d^2}$ such that $g_{N,m}(x)$ encodes all the potentially non-zero $\langle q_2 c_2 | \, N^{\hookrightarrow}_{x,m}(F_{q_1,q_1',c_1,c_1'}) \, | q_2' c_2' \rangle$, without redundancy (only encoding one element of a conjugate pair). To be precise, the first $k^2 d^2$ entries of $g_{N,m}(x)$ are given by $\{ \langle q_2 c_2 | \, N^{\hookrightarrow}_{x,m}(F_{q_1,q_1,c_1,c_1}) \, | q_2 c_2 \rangle : q_1, q_2 \in Q, c_1, c_2 \in C \} \subseteq \mathbb{R}$. Establish some total order $\geq$ on Q, and let $\widehat{Q^4} = \{ (q_1, q_1', q_2, q_2') \in Q^4 : q_1' > q_1 \text{ or } (q_1' = q_1 \text{ and } q_2' > q_2) \}$. The remaining $k^4 d^2 - k^2 d^2$ entries are given by encoding each of the $\frac{1}{2}(k^4 d^2 - k^2 d^2)$ potentially non-zero entries $\{ \langle q_2 c_2 | \, N^{\hookrightarrow}_{x,m}(F_{q_1,q_1',c_1,c_1}) \, | q_2' c_2 \rangle : (q_1, q_1', q_2, q_2') \in \widehat{Q^4}, c_1, c_2 \in C \} \subseteq \mathbb{C}$ as the pair of real numbers that comprise their real and imaginary parts.

Let $h = k^4 d^2$. Let $\| \cdot \| : \mathbb{R}^h \to \mathbb{R}_{\geq 0}$ denote the Euclidean 2-norm and $\| \cdot \|_2 : \mathrm{L}(V) \to \mathbb{R}_{\geq 0}$ denote the Schatten 2-norm. Note that $\| \Phi \|_1 \leq \| J(\Phi) \|_1, \forall \Phi$ [37, Section 3.4]. We have,

$$\| N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m} \|_1 \leq \| J(N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}) \|_1 \leq \sqrt{\mathrm{rank}(J(N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}))} \| J(N^{\hookrightarrow}_{x,m} - N^{\hookrightarrow}_{x',m}) \|_2$$

$$\leq \sqrt{h} \| J(N^{\hookrightarrow}_{x,m}) - J(N^{\hookrightarrow}_{x',m}) \|_2 \leq \sqrt{2h} \| g_{N,m}(x) - g_{N,m}(x') \|.$$

Note that $N^{\hookrightarrow}_{x,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, which implies $\| N^{\hookrightarrow}_{x,m} \|_1 = 1$ [37, 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,

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

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 $\mathbb{R}^h$, which has volume $\mathrm{vol}(B(v_0, r)) = c_h r^h$, for some constant $c_h \in \mathbb{R}_{>0}$. By the above, $\| g_{N,m}(x) \| \leq \sqrt{h}$, which implies that $B(g_{N,m}(x), \delta) \subseteq B(0, \sqrt{h} + \delta)$, $\forall \delta \in \mathbb{R}_{>0}$. 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,

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

We now prove a 2QCFA analogue of the Dwork-Stockmeyer lemma.

▶ **Theorem 16.** *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 $T(n) \geq \frac{(1-2\epsilon)^2}{16\sqrt{2}k^4d^2}D_L(n)^{\frac{1}{k^4d^2}}$, $\forall n \geq N_0$.*

Proof. Consider some $L \subseteq \Sigma^*$. By [10, Lemma 3.1], $L \in \mathsf{REG} \Leftrightarrow \exists b \in \mathbb{N}_{\geq 1}$ such that $D_L(n) \leq b$, $\forall n \in \mathbb{N}$. Thus, if $L \in \mathsf{REG}$, the claim is immediate (recall that $T(n) \geq n$). Next, suppose $L \notin \mathsf{REG}$. For $n \in \mathbb{N}$, 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 $D_L(N_0) \geq 2^{k^4d^2}$. Then, $\forall n \geq N_0$, we have $|X_n| = D_L(n) \geq D_L(N_0) \geq 2^{k^4d^2}$. Fix $n \geq N_0$ and set $m = \lceil\frac{1-2\epsilon}{2}T(n)\rceil$. By Lemma 15, $\exists x,x' \in X_n$ such that $x \neq x'$ and

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

Fix such a pair x,x', and note that $x \not\sim_{L,n} x'$, by construction. By Lemma 14,

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

4.3 2QCFA Running Time Lower Bounds and Complexity Class Separations

Let $\mathsf{B2QCFA}(T(n)) = \cup_{k,d\in\mathbb{N}_{\geq 2},\epsilon\in[0,\frac{1}{2})}\mathsf{B2QCFA}(k,d,T(n),\epsilon)$ denote the class of languages recognizable with two-sided bounded error by a 2QCFA with any constant number of quantum and classical states, in expected time at most $T(n)$. 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 $\mathsf{B2QCFA}(T(n))$. Let $C_L : \mathbb{N} \to \mathbb{N}$ denote the *one-way deterministic communication complexity* of testing membership in L; note that $C_L(n) = \log D_L(n), \forall n$ [8]. We immediately obtain the following corollaries of Theorem 16.

▶ **Corollary 17.** *If $L \in \mathsf{B2QCFA}(T(n))$, then $D_L(n) = T(n)^{O(1)}$ and $C_L(n) = O(\log T(n))$.*

▶ **Corollary 18.** *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 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 $w = w_1\cdots w_n \in \Sigma^*$, let $w^{\mathrm{rev}} = w_n\cdots w_1$ denote the reversal of the string w. Let $L_{pal} = \{w \in \{a,b\}^* : w = w^{\mathrm{rev}}\}$ consist of all palindromes over the alphabet $\{a,b\}$.

▶ **Corollary 19.** *$L_{pal} \notin \mathsf{B2QCFA}(2^{o(n)})$.*

Proof. For $n \in \mathbb{N}$, 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, $w \not\sim_{L_{pal},2n} w'$, $\forall w,w' \in W_n$ such that $w \neq w'$. This implies that $D_{L_{pal}}(2n) \geq |W_n| = 2^n$. Corollary 18 then implies $L_{pal} \notin \mathsf{B2QCFA}(T(n))$. ◀

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. 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 *subexponential* time is properly contained in the class of languages recognizable by a 2QCFA in *exponential* time.

▶ **Corollary 20.** $\mathsf{B2QCFA}(2^{o(n)}) \subsetneq \mathsf{BQE2QCFA}$ *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. See the full version [26] for a proof of the following corollary.

▶ **Corollary 21.** *If* $L \in \mathsf{BQP2QCFA}$*, then* $D_L(n) = n^{O(1)}$*. Therefore,* $\mathsf{BQP2QCFA} \subseteq L/poly$.

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 $p \in \mathbb{N}$, let $\langle p \rangle_2 \in \{0,1\}^*$ denote its binary representation; let $L_{primes} = \{\langle p \rangle_2 : p \text{ is prime}\}$. Note that $D_{L_{primes}}(n) = 2^{\Omega(n)}$ [29], which immediately implies the following.

▶ **Corollary 22.** $L_{primes} \notin \mathsf{B2QCFA}(2^{o(n)})$.

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

▶ 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).

4.4 Transition Amplitudes of 2QCFA

As in Definition 3, for some 2QCFA $N = (Q, C, \Sigma, R, \theta, \delta, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$, let $\{E_{c,\sigma,r,j} : r \in R, j \in J\} \subseteq \mathrm{L}(\mathbb{C}^Q)$ denote the set of operators that describe the selective quantum operation $\theta(c, \sigma) \in \mathrm{QuantOp}(\mathbb{C}^Q, R)$ that is applied to the quantum register when the classical state of N is $c \in \widehat{C}$ and the head of N is over the symbol $\sigma \in \Sigma_+$. The *transition amplitudes* of N are the set of numbers $\{\langle q| E_{c,\sigma,r,j} |q'\rangle : c \in \widehat{C}, \sigma \in \Sigma_+, r \in R, j \in J, q, q' \in Q\} \subseteq \mathbb{C}$.

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 [28]. 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 subexponential time, regardless of the permitted transition amplitudes.

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

We next 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). The key idea is that a QTM M that uses $S(n)$ space 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.

We consider the *classically controlled* space-bounded QTM model that allows *intermediate measurements*, following the definition of Watrous [35]. While several such QTM models have been defined, we focus on this model as we wish to prove our lower bound in the greatest generality possible. We note that the definitions of such QTM models by, for instance, Ta-Shma [32], Watrous [36, Section VII.2], and (essentially, without the use of random access) van Melkebeek and Watson[22] are special cases of the QTM model that we consider. In the case of time-bounded quantum computation, it is well-known that allowing a QTM to perform intermediate measurements provably does not increase the power of the model; very recently, this fact has also been shown to hold in the simultaneously time-bounded and space-bounded setting [12]. Let $\mathsf{BQTISP}(T(n), S(n))$ denote the class of languages recognizable with bounded error by a QTM in time $T(n)$ and space $S(n)$. See the full version [26] for a complete definition of the QTM model and a proof of the following theorem.

▶ **Theorem 24.** *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, $T(n) = \Omega\big(2^{-b_0 S(n)} D_L(n)^{2^{-b_0 S(n)}}\big)$.*

▶ **Corollary 25.** *If $D_L(n) = 2^{\Omega(n)}$, 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* TM in $O(\log n)$ space (and, trivially, polynomial time). Therefore, the previous corollary exhibits a natural problem for which polynomial time *quantum* TM cannot outperform polynomial time *deterministic* TM 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 [21]. 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 $W_{G=\langle S|R\rangle} = \phi^{-1}(1_G)$. Note that the definition of the word problem does depend on the choice presentation. However, 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 $W_{G=\langle S|R\rangle} \in \mathcal{L} \Leftrightarrow W_{G=\langle S'|R'\rangle} \in \mathcal{L}$ [19]. 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.

6.1 The Growth Rate of a Group and Nonregularity

Consider a group $G = \langle S|R\rangle$, with S finite. Define Σ and ϕ as in the previous section. For $g \in G$, let $l_S(g)$ denote the smallest $m \in \mathbb{N}$ such that $\exists \sigma_1, \dots, \sigma_m \in \Sigma$ 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 26.** *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 $B_{G,S}(n) = \{g_1, \dots, g_k\}$. 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 $i \in \{1, \dots, k\}$ 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, $w_1, \dots, w_k$ are pairwise $(W_G, 2n)$-dissimilar, which implies $D_{W_G}(2n) \geq k = \beta_{G,S}(n)$. ◀

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 $\forall r \in \mathbb{R}_{\geq 0}$, $f_1(r) \leq C_1 f_2(C_1 r + C_2) + C_2$; we write $f_1 \sim f_2$ if both $f_1 \prec f_2$ and $f_2 \prec f_1$. Suppose $\langle S|R\rangle$ and $\langle S'|R'\rangle$ are both presentations of G, with S and S' finite. It is straightforward to show that $\beta_{G,S}$ and $\beta_{G,S'}$ are non-decreasing, and that $\beta_{G,S} \sim \beta_{G,S'}$ [21, Proposition 6.2.4]. For this reason, we will simply write β_G to denote the growth rate of G.

▶ **Definition 27.** Suppose G is a finitely generated group. If $\beta_G \sim (n \mapsto e^n)$, we say *$G$ has exponential growth.* If $\exists c \in \mathbb{R}_{\geq 0}$ such that $\beta_G \prec (n \mapsto n^c)$, we say *$G$ has polynomial growth.* Otherwise, 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 [34] and Gromov's theorem on groups of polynomial growth [15], 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. We obtain an analogous result for sublogarithmic-space QTMs; see the full paper [26] for details.

▶ **Theorem 28.** *For any finitely generated group G, the following statements hold.*

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

(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 26 and Corollary 17.

(ii) Follows immediately from Definition 27 and part (i) of this theorem.

(iii) As a consequence of the famous Tits' Alternative [34], 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 ([34, Corollary 1],[38]). 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, $\beta_G \prec (n \mapsto n^{ck^4 d^2})$, which implies G has polynomial growth. By Gromov's theorem on groups of polynomial growth [15], a finitely generated group has polynomial growth precisely when it is virtually nilpotent. ◀

▶ Remark. All *known* G of intermediate growth have $\beta_G \sim (n \mapsto e^{n^c})$, for some $c \in (1/2, 1)$. Therefore, a strong lower bound may be established on the running time of any 2QCFA that recognizes W_G, for any known group of intermediate growth.

Let $\mathcal{G}_{\mathbf{vAb}}$ (resp. $\mathcal{G}_{\mathbf{vNilp}}$) denote the collection of all finitely generated virtually abelian (resp. nilpotent) groups. Let $\mathrm{U}(k, \overline{\mathbb{Q}})$ denote the group of $k \times k$ unitary matrices with algebraic number entries, and let $\mathcal{U}$ consist of all finitely generated subgroups of any $\mathrm{U}(k, \overline{\mathbb{Q}})$. We have recently shown that if $G \in \mathcal{U}$, then $W_G \in \mathsf{coRQE2QCFA}_{\overline{\mathbb{Q}}}$ [27, 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, $\mathcal{U} \cap \mathcal{G}_{\mathbf{vNilp}} = \mathcal{G}_{\mathbf{vAb}}$ [33, Proposition 2.2]. We, therefore, obtain the following corollary of Theorem 28, which exhibits a broad and natural class of languages that a 2QCFA can recognize in exponential time, but not in subexponential time.

▶ **Corollary 29.** $\forall G \in \mathcal{U} \setminus \mathcal{G}_{\mathbf{vAb}}$*, we have* $W_G \in \mathsf{coRQE2QCFA}_{\overline{\mathbb{Q}}}$ *but* $W_G \notin \mathsf{B2QCFA}(2^{o(n)})$.

We have also recently shown that $W_G \in \mathsf{coRQP2QCFA}_{\overline{\mathbb{Q}}}(2) \subseteq \mathsf{BQP2QCFA}$, $\forall G \in \mathcal{G}_{\mathbf{vAb}}$ [27, Theorem 1.2]. By Theorem 28, 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 $W_G \in \mathsf{BQP2QCFA}$. Consider the (three-dimensional discrete) Heisenberg group $H = \langle x, y, z | z = [x, y], [x, z] = [y, z] = 1 \rangle$. W_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 [27] for further discussion). We next show that if $W_H \notin \mathsf{BQP2QCFA}$, then we have a complete classification of those word problems recognizable by 2QCFA in polynomial time.

▶ **Proposition 30.** *If* $W_H \notin \mathsf{BQP2QCFA}$*, then* $W_G \in \mathsf{BQP2QCFA} \Leftrightarrow G \in \mathcal{G}_{\mathbf{vAb}}$.

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 $\forall G \in \mathcal{G}_{\mathbf{vNilp}} \setminus \mathcal{G}_{\mathbf{vAb}}$, G has a subgroup isomorphic to H [20, Theorem 12]. It is straightforward to see that $\mathsf{BQP2QCFA}$ is closed under inverse homomorphism and intersection with regular languages. Therefore, if $W_G \in \mathsf{BQP2QCFA}$, then $W_H \in \mathsf{BQP2QCFA}$ [20, Lemma 2]. ◀

References

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

2 Andris Ambainis and John Watrous. Two-way finite automata with quantum and classical states. *Theoretical Computer Science*, 287(1):299–311, 2002.

3 Andris Ambainis and Abuzer Yakaryılmaz. Automata and quantum computing. *arXiv preprint*, 2015. `arXiv:1507.01988`.

4 Ao V Anisimov. Group languages. *Cybernetics and Systems Analysis*, 7(4):594–601, 1971.

5 Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando GSL Brandao, David A Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.

6 J-C Birget, A Yu Ol'shanskii, Eliyahu Rips, and Mark V Sapir. Isoperimetric functions of groups and computational complexity of the word problem. *Annals of Mathematics*, pages 467–518, 2002.

7 Allan Borodin, Stephen Cook, and Nicholas Pippenger. Parallel computation for well-endowed rings and space-bounded probabilistic machines. *Information and Control*, 58(1-3), 1983.

8 Anne Condon, Lisa Hellerstein, Samuel Pottle, and Avi Wigderson. On the power of finite automata with both nondeterministic and probabilistic states. *SIAM Journal on Computing*, 27(3):739–762, 1998.

9 Martin J Dunwoody. The accessibility of finitely presented groups. *Inventiones mathematicae*, 81(3):449–457, 1985.

10 Cynthia Dwork and Larry Stockmeyer. A time complexity gap for two-way probabilistic finite-state automata. *SIAM Journal on Computing*, 19(6):1011–1023, 1990.

11 Cynthia Dwork and Larry Stockmeyer. Finite state verifiers I: The power of interaction. *Journal of the ACM (JACM)*, 39(4):800–828, 1992.

12 Bill Fefferman and Zachary Remscrim. Eliminating intermediate measurements in space-bounded quantum computation, 2020. `arXiv:2006.03530`.

13 Rūsiņš Freivalds. Probabilistic two-way machines. In *International Symposium on Mathematical Foundations of Computer Science*, pages 33–45. Springer, 1981.

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

15 Michael Gromov. Groups of polynomial growth and expanding maps (with an appendix by Jacques Tits). *Publications Mathématiques de l'IHÉS*, 53:53–78, 1981.

16 Lov K Grover. A fast quantum mechanical algorithm for database search. *Proceedings of the Twenty-Eighth Annual ACM Symposium of Theory of Computing*, pages 212–219, 1996.

17 Aram W Harrow, Avinatan Hassidim, and Seth Lloyd. Quantum algorithm for linear systems of equations. *Physical review letters*, 103(15):150502, 2009.

18 Fred C Hennie. One-tape, off-line Turing machine computations. *Information and Control*, 8(6):553–578, 1965.

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

20 Derek F Holt, Sarah Rees, Claas E Röver, and Richard M Thomas. Groups with context-free co-word problem. *Journal of the London Mathematical Society*, 71(3):643–657, 2005.

21 Clara Löh. *Geometric group theory*. Springer, 2017.

22 Dieter van Melkebeek and Thomas Watson. Time-space efficient simulations of quantum computations. *Theory of Computing*, 8(1):1–51, 2012.

23 David E Muller and Paul E Schupp. Groups, the theory of ends, and context-free languages. *Journal of Computer and System Sciences*, 26(3):295–310, 1983.

24 Michael A Nielsen and Isaac Chuang. Quantum computation and quantum information, 2002.

25 Michael O Rabin and Dana Scott. Finite automata and their decision problems. *IBM journal of research and development*, 3(2):114–125, 1959.

26 Zachary Remscrim. Lower bounds on the running time of two-way quantum finite automata and sublogarithmic-space quantum turing machines, 2020. `arXiv:2003.09877`.

27 Zachary Remscrim. The Power of a Single Qubit: Two-Way Quantum Finite Automata and the Word Problem. In *47th International Colloquium on Automata, Languages, and Programming (ICALP 2020)*, volume 168 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 139:1–139:18, 2020.

28 AC Say and Abuzer Yakaryilmaz. Magic coins are useful for small-space quantum machines. *Quantum Information & Computation*, 17(11-12):1027–1043, 2017.

29 Jeffrey Shallit. Automaticity IV: sequences, sets, and diversity. *Journal de théorie des nombres de Bordeaux*, 8(2):347–367, 1996.

30 Jeffrey Shallit and Yuri Breitbart. Automaticity I: Properties of a measure of descriptional complexity. *Journal of Computer and System Sciences*, 53(1):10–25, 1996.

31 Peter W Shor. Algorithms for quantum computation: Discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*. Ieee, 1994.

32 Amnon Ta-Shma. Inverting well conditioned matrices in quantum logspace. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 881–890, 2013.

33 Andreas Thom. Convergent sequences in discrete groups. *Canadian Mathematical Bulletin*, 56(2):424–433, 2013.

34 Jacques Tits. Free subgroups in linear groups. *Journal of Algebra*, 20(2):250–270, 1972.

35 John Watrous. On the complexity of simulating space-bounded quantum computations. *Computational Complexity*, 12(1-2):48–84, 2003.

36 John Watrous. Encyclopedia of complexity and system science, chapter quantum computational complexity, 2009. `arXiv:0804.3401`.

37 John Watrous. *The theory of quantum information*. Cambridge University Press, 2018.

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

39 Abuzer Yakaryilmaz and AC Cem Say. Succinctness of two-way probabilistic and quantum finite automata. *Discrete Mathematics and Theoretical Computer Science*, 12(4):19–40, 2010.