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

The Limitations of Few Qubits: One-way and Two-way Quantum Finite Automata and the Group Word Problem

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 exponential time. We prove that their result essentially cannot be improved upon: a 2QCFA (of any finite size) cannot recognize L_{pal} with bounded-error in expected time $2^{o(n)}$, on inputs of length 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. A key tool in our result is a generalization to 2QCFA of a technical lemma that was used by Dwork and Stockmeyer to prove a lower bound on the expected running time of any two-way probabilistic finite automaton that recognizes a non-regular language with bounded-error.

Furthermore, we prove strong lower bounds on the expected running time of any 2QCFA that recognizes a group word problem with bounded-error. In a recent paper, we showed that 2QCFA can recognize, with bounded-error, a broad class of group word problems in expected exponential time, and a more narrow class of group word problems in expected polynomial time. As a consequence, we can now exhibit a large family of natural languages that can be recognized with bounded-error by a 2QCFA in expected exponential time, but not in expected subexponential time. Moreover, we obtain significant progress towards a precise classification of those group word problems that can be recognized with bounded-error in expected polynomial time by a 2QCFA.

We also consider the one-way measure-once quantum finite automaton (1QFA), defined by Moore and Crutchfield, as well as a natural generalization to one-way measure-once finite automata with quantum and classical states (1QCFA). We precisely classify those groups whose word problem may be recognized with positive one-sided error (for both the bounded-error and unbounded-error cases) by a 1QFA or 1QCFA with any particular number of quantum states and any particular number of classical states; we also obtain partial results in the negative one-sided error case. As an immediate corollary, we show that allowing a 1QFA or 1QCFA to have even a single additional quantum or classical state enlarges the class of languages that may be recognized with positive one-sided error (of either type).

1 Introduction

Quantum algorithms, such as Shor's quantum polynomial time integer factorization algorithm [38], Grover's algorithm for unstructured search [18], and the linear system solver of Harrow, Hassidim, and Lloyd [19], 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

ISSN 1433-8092

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 [4]. Informally, a 2QCFA is a two-way deterministic finite automaton (2DFA) that has been augmented by a quantum register of finite size; we define the 2QCFA model formally in Section 3.1. 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 polynomial time and the language $L_{pal} = \{w \in \{a,b\}^* : w \text{ is a palindrome}\}$ in expected exponential time. In a recent paper [35], 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 [34]. A 2PFA can recognize some non-regular languages with bounded-error, given sufficient running time: in particular, a 2PFA can recognize L_{eq} with bounded-error in expected exponential time [14]. However, a 2PFA cannot recognize L_{eq} with bounded-error in expected subexponential time, by a result of Greenberg and Weiss [15]; moreover, a 2PFA cannot recognize L_{pal} with bounded-error in any time bound [13]. More generally, the landmark result of Dwork and Stockmeyer [12] showed that a 2PFA cannot recognize any non-regular language in subexponential time. In order to prove this statement, they defined a function $D_L : \mathbb{N} \to \mathbb{N}$ that captures the "hardness" of a language L; roughly speaking, $D_L(n)$ counts the number of input strings of length at most n that must be "distinguished" by any recognizer of L, in a certain sense. 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 lower bound on $T(n)$ in terms of $D_L(n)$; we will refer to this statement as the "Dwork-Stockmeyer lemma."

However, 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 any finite number of quantum basis states) can recognize with bounded-error in expected exponential time but not in expected subexponential time? These are natural questions, which, to our knowledge, were open (see, for instance, [4, 5, 44] for previous discussions of these questions).

In this paper, we answer these and other related questions. In particular, we show that 2QCFA cannot recognize L_{pal} with bounded-error in expected running time $T(n) = 2^{o(n)}$. More generally, we prove an analogue of the Dwork-Stockmeyer lemma for 2QCFA: we establish a lower bound on the expected running time $T(n)$ for any 2QCFA that recognizes any language L, where our lower bound is also in terms of $D_L(n)$. One of the key tools used in our proof is a quantum version of Hennie's [20] notion of a crossing sequence, which may be of independent interest. Crossing sequences played a key role in the aforementioned 2PFA results of Dwork and Stockmeyer [12] and of Greenberg and Weiss [15].

We also investigate which group word problems can be recognized by 2QCFA with particular resource bounds. Informally, the word problem of a 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 group G and the complexity of its word problem W_G, as has been demonstrated by many famous results, such as Anisimov's result that $W_G \in \mathsf{REG}$ (the regular languages) if and only if G is a finite group [6], the result of Muller and Schupp that $W_G \in \mathsf{CFL}$ (the context-free languages) if and only if G is a finitely-generated virtually free

group [11, 30], and the result of Lipton and Zalcstein which showed that the word problem of any finitely-generated linear group (over a field of characteristic zero) is in deterministic logspace [26]. We have recently shown that if G is a finitely-generated virtually abelian group, 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 finitely-generated linear groups of exponential growth, W_G may be recognized with bounded-error by a 2QCFA (in many cases a single-qubit 2QCFA) in expected exponential time [35].

We now show that, for any group G of exponential growth, a 2QCFA cannot recognize W_G with bounded-error in expected subexponential time, thereby providing a broad and natural class of languages that may be recognized by a 2QCFA in expected exponential time but not in expected subexponential time. We also show that, if a 2QCFA recognizes a word problem W_G with bounded-error in expected polynomial time, then G must be a finitely-generated virtually nilpotent group (i.e., G must have polynomial growth), thereby obtaining progress towards an exact classification of those group word problems recognizable by a 2QCFA in expected polynomial time.

Furthermore, we consider measure-once one-way quantum finite automata (1QFA) originally defined by Moore and Crutchfield [29] and the natural generalization to measure-once one-way finite automata with quantum and classical states (1QCFA). With *bounded-error*, a 1QFA can recognize W_G only when G is a finite group (and hence $W_G \in \mathsf{REG}$) [29]; however, the smallest such 1QFA may be considerably smaller than the smallest DFA that recognizes W_G. In particular, for the group $G = \mathbb{Z}/p\mathbb{Z}$, where p is a prime, the smallest DFA that recognizes W_G has size p, whereas there is a 1QFA with only $O(\log p)$ quantum basis states that recognizes W_G with *negative* one-sided bounded-error [2]; however, a 1QFA requires $\Omega(\frac{\log p}{\log \log p})$ quantum basis states to recognize W_G with bounded-error [1].

We show that W_G is recognized with positive one-sided *unbounded-error* by a 1QFA with at most k quantum basis states if and only if W_G is recognized with positive one-sided *bounded-error* by a 1QFA with at most k quantum basis states if and only if W_G is recognized by a DFA with at most k states if and only if $|G| \leq k$. Therefore, 1QFA with *positive* one-sided error have no advantage over DFA when recognizing word problems. Similarly, we precisely classify those word problems recognizable with positive one-sided error (in both the bounded-error and unbounded error case) by a 1QCFA with at most k quantum states and at most d classical states, for any k and d, and similarly observe that positive one-sided error provides no advantage to 1QCFA over DFA when recognizing word problems. As an immediate corollary, we show that allowing such a 1QFA or 1QCFA even a single additional quantum state or classical state enlarges the class of languages that may be recognized.

The class of groups whose word problem is recognizable by a 1QFA (or more generally a 1QCFA) with *negative* one-sided unbounded-error is considerably larger: we have recently shown that the word problem of any group with a faithful finite-dimensional projective unitary representation may be recognized by such a 1QFA [35]. This class includes, for example, all groups G for which $W_G \in \mathsf{CFL}$, as well as all groups for which it is known that $W_G \in \mathsf{poly{-}CFL}$, as well as many groups whose word problems are broadly conjectured not to be in $\mathsf{coCFL} \cup \mathsf{poly{-}CFL}$. We now obtain an exact classification of the group word problems recognizable with negative one-sided unbounded-error by a 1QFA or 1QCFA with a single-qubit (and, in the 1QCFA case, any number of classical states) as well as partial results for the general case of any finite number of quantum states.

2 Preliminaries

2.1 Quantum Computation

We briefly recall the fundamentals of quantum computation needed in this paper (see, for instance, [41] or [33] 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 $\langle A|\psi_1\rangle, |\psi_2\rangle\rangle = \langle |\psi_1\rangle, A^\dagger|\psi_2\rangle\rangle$, $\forall |\psi_1\rangle, |\psi_2\rangle \in V$. 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.

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 $\mathrm{Chan}(V) = \mathrm{Chan}(V, V)$.

Each QFA variant considered in this paper has a *quantum register* 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 QFA 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 $P_0, \ldots, P_{l-1} \in \mathrm{Proj}(\mathbb{C}^Q)$, 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 $r \in \{0, \ldots, l-1\}$ with probability $\|P_j|\psi\rangle\|^2$; if the result is r, then the quantum register collapses to the superposition $\frac{1}{\|P_j|\psi\rangle\|} P_j|\psi\rangle$. We emphasize that performing a quantum measurement changes the state of the quantum register. For $B = \{b_0, \ldots, b_{l-1}\}$, a partition of Q into l parts, we define the *quantum measurement in the computational basis* with respect to B as the von Neumann measurement where each $P_j = \sum_{q \in b_j} |q\rangle\langle q|$.

For a 2QCFA, which may perform many quantum measurements during its computation, the total state of its quantum register at any point in time is described by an *ensemble of pure states* (what some call a *mixed state*) $\{(p_i, |\psi_i\rangle)\}$, where each $p_i \in [0, 1]$ denotes the probability of the quantum register being in the superposition $|\psi_i\rangle \in \mathbb{C}^Q$, and $\sum_i p_i = 1$. Such an ensemble is described by 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 will be described by the same density operator; however, all such ensembles will "behave the same" for certain purposes. That is to say, for any ensemble described by a density operator A, applying the transformation $T \in \mathrm{U}(\mathbb{C}^Q)$ produces an ensemble described by the density operator $TAT^\dagger$. Similarly, performing the von Neumann measurement specified by $P_0, \ldots, P_{l-1} \in \mathrm{Proj}(\mathbb{C}^Q)$ on any ensemble described by a density operator A produces the result $r \in \{0, \ldots, l-1\}$ with probability $\mathrm{Tr}(P_r A P_r^\dagger)$, and if the result is r, then the ensemble collapses to an ensemble described by the density operator $\frac{1}{\mathrm{Tr}(P_r A P_r^\dagger)} P_r A P_r^\dagger$.

2.2 Group Theory and the Word Problem

In this section, we formally define the word problem of a group; for further background, see, for instance [27]. 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, [21]) 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.

3 Two-way Finite Automata with Quantum and Classical States

3.1 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 [4]. Informally, a

2QCFA is a two-way DFA that has been augmented with a quantum register of finite size; the machine may apply unitary transformations to the quantum register and perform (perhaps many) quantum measurements of its quantum register during its computation. Formally, a 2QCFA is a 10-tuple,

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

where Q is a finite set of quantum basis states, C is a finite set of classical states, Σ is a finite input alphabet, $\delta_{\text{type}}, \delta_{\text{transform}}$, and δ_{measure} collectively specify 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 specified as follows. Firstly, $\delta_{\text{type}} : \widehat{C} \times \Sigma_+ \to \{transform, measure\}$ 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, $\delta_{\text{transform}} : \delta_{\text{type}}^{-1}(transform) \to \mathrm{U}(\mathbb{C}^Q) \times C \times \{-1, 0, 1\}$ 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, $\delta_{\text{type}}(c, \sigma) = measure$, then $\delta_{\text{measure}}(c, \sigma)$ is a pair (B, f) where B is some partition of Q that specifies a quantum measurement and $f : B \to C \times \{-1, 0, 1\}$ 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 $|q_{\text{start}}\rangle$, 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, suppose $\delta_{\text{type}}(c, \sigma) = transform$ 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, $\delta_{\text{type}}(c, \sigma) = measure$, then if $\delta_{\text{measure}}(c, \sigma) = (B, f)$, N performs the quantum measurement specified by B, producing the result $b \in B$; if $f(b) = (c', d)$, 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, $\forall w \notin L$, $\Pr[N \text{ rejects } w] \geq 1 - \epsilon$. Then, for any function $T : \mathbb{N} \to \mathbb{N}$, we define $\mathsf{B2QCFA}(k, d, T, \epsilon)$ 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.

We note here that 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 [4], which required

N to halt with probability 1 on all inputs and operated under the more restrictive *negative one-sided* bounded-error recognition condition. The only other alterations that we have made to their definition of the 2QCFA model are purely done for convenience and do not affect the power of the model.

3.2 2QCFA Crossing Sequences

In this section, we develop a generalization of Hennie's [20] notion of crossing sequences to 2QCFA, in which we make use of several ideas from the 2PFA results of Dwork and Stockmeyer [12] and Greenberg and Weiss [15]. 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 (recall that we allow N to reject by entering c_{rej} or by looping). 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 an ensemble of configurations $\{(p_i, (|\psi_i\rangle, c_i, n'))\}$, with each $|\psi_i\rangle \in \Psi$ and each $c_i \in C$, where the probability of being in configuration $(|\psi_i\rangle, c_i, n')$ is given by $p_i \in [0, 1]$; we call this ensemble a *starting ensemble*. To avoid unnecessary cases later, we also allow N to start in a configuration of the form $(|\psi\rangle, c, n')$, where $c \in \{c_{\text{acc}}, c_{\text{rej}}\}$, where we adopt the convention that in such a case N immediately leaves $\#_L x$ in the configuration $(|\psi\rangle, c, n'+1)$. 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.

3.2.1 Describing Ensembles of Configurations of 2QCFA

Let $x_0 = \#_L$, let $\widehat{H}_x = \{0, \ldots, n'\}$ denote the head positions corresponding to the prefix $\#_L x$, and let $H_x = \{0, \ldots, n'+1\}$ 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 describe (non-uniquely) ensembles of configurations of N.

[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 [12] in their 2PFA result.

We first consider an ensemble of superpositions 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 of configurations of the quantum register of N 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, [33, 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 quantum measurement specified by the partition B of Q, where for each $b \in B$ we let $P_b = \sum_{q \in b} |q\rangle\langle q| \in \mathrm{Proj}(\mathbb{C}^Q)$ denote the projection operator corresponding to measurement outcome b. Then for any ensemble described by the density operator A, the probability that the result of this quantum measurement is b is given by $\mathrm{Tr}(P_b A P_b^\dagger)$, and if the result is b then the ensemble collapses to an ensemble described by the density operator $\frac{1}{\mathrm{Tr}(P_b A^\dagger)} P_b A P_b^\dagger$. 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.

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 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. In particular, 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 finite *quantum register*, that stores a superposition $|\psi\rangle \in \mathbb{C}^Q$, and a finite *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 $\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x} \cong \mathbb{C}^{Q \times C \times H_x}$, 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: n'. 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 $I_x(Z) = Z \otimes |n'\rangle\langle n'|$, $\forall Z \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. 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 $\{(p_i, (|\psi_i\rangle, c_i, n')) : i \in I\}$ 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 $n'+1$ 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 $\mathrm{Tr}_{\mathbb{C}^{H_x}}(Z) \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$.

3.2.2 Overview of 2QCFA Crossing Sequences

We begin by briefly sketching our definition of the *crossing sequence* of the 2QCFA N on the partitioned input xy. For any $m \in \mathbb{N}$, if N is run on the prefix $\#_L x$ beginning in some ensemble of configurations $\{(p_i, (|\psi_i\rangle, c_i, n')) : i \in I\}$, 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 $m + 1^{\text{st}}$ quantum measurement and instead immediately reject (we also adopt a special convention to deal with branches that are rejecting by looping, which we discuss later). We will then define the *m-truncated transfer operator* $N_{x,m} : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ 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 $N_{x,m}(Z)$. 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 $\widetilde{N}_{y,m} : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ 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 $N_{x,m}$ and $\widetilde{N}_{y,m}$ in an infinite sequence.

Crucially, we will observe that $N_{x,m}, \widetilde{N}_{y,m} \in \text{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, $\forall x, y \in \Sigma^*$, $\forall m \in \mathbb{N}$. 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. Similarly, while the m-truncated crossing operator $N_{x,m}$ completely suffices for our analysis, we also define a *non-truncated transfer operator* $N_x \in \text{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ as an accumulation point of the sequence $(N_{x,m})_{m\in\mathbb{N}}$; such an accumulation point exists due to the fact that $\text{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ is compact (see, for instance, [41, Proposition 2.28]). Using N_x and the symmetrically defined $\widetilde{N}_y$, we define the *non-truncated crossing sequence* of N on xy. The resulting analyses of these two types of crossing sequences would essentially be identical; however, we make these additional definitions as the (somewhat cleaner) non-truncated crossing sequence may be more useful in other applications.

3.2.3 Definition and Properties of 2QCFA Crossing Sequences

We now consider 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 $(2)_{\text{halt}}$ and $(2)_{\text{loop}}$ 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), $(2)_{\text{halt}}$, $(2)_{\text{loop}}$, and (3) as *key-events*. We define $\text{keyEv}_x : \widehat{C} \times \widehat{H}_x \to \{(1), (2)_{\text{halt}}, (2)_{\text{loop}}, (3)\}$ such that $\text{keyEv}_x(c, h)$ 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 $t_x : C \times H_x \to \text{U}(\mathbb{C}^Q)$, $\gamma_x : C \times H_x \to C$, 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 $\text{keyEv}_x(c, h) \in \{(1), (2)_{\text{halt}}, (3)\}$. 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 $\text{keyEv}_x(c, h) \in \{(1), (2)_{\text{halt}}\}$, let $s_{x,c,h} = \widehat{s}_{x,c,h}$, if $\text{keyEv}_x(c, h) = (3)$, 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 $(2)_{\text{halt}}$ 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 $\mathrm{keyEv}_x(c,h) = (2)_{\text{loop}}$. 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 $\mathrm{keyEv}_x(c,h) = (1)$, then N has just left $\#_L x$ for the first time; in particular, $h_x(c,h) = n' + 1$ (i.e., the head is one cell to the right of the rightmost symbol of $\#_L x$). If $\mathrm{keyEv}_x(c,h) = (2)_{\text{halt}}$, 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 $\mathrm{keyEv}_x(c,h) = (2)_{\text{loop}}$, then N is rejecting its input by looping (on this branch); in particular, $\gamma_x(c,h) = c_{\text{rej}}$. If $\mathrm{keyEv}_x(c,h) = (3)$, then $h_x(c,h) \in \widehat{H}_x$ and $\delta_{\text{type}}(\gamma_x(c,h), x_{h_x(c,h)}) = \mathit{measure}$; 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 $\mathrm{keyEv}_x(c,h) \in \{(1), (2)_{\text{halt}}, (2)_{\text{loop}}\}$, 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 $\mathrm{keyEv}_x(c,h) = (3)$, 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 $h = n'+1$), 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 $h = n' + 1$ 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 such that $c_i \in \widehat{C}$ 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 $Z \in \widetilde{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ 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 $E_{x,c,h} = t_x(c,h) \otimes |\gamma_x(c,h)\rangle\langle c| \otimes |h_x(c,h)\rangle\langle h| \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$. We then define the operator $K_x : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ such that $K_x(Z) = \sum_{c \in C, h \in H_x} E_{x,c,h} Z E_{x,c,h}^\dagger$, $\forall Z \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$.

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

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

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

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

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, [41, 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 $\sum_{c \in C, h \in H_x} E_{x,c,h}^\dagger E_{x,c,h} = 1_{\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}}$ (see, for instance, [41, Corollary 2.27]). For any $c \in C$ and $h \in H_x$, we have

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

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

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

Therefore,

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

□

We next consider the behavior of N when it perform 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 $\delta_{\text{type}}(c, x_h) = \textit{measure}$ (i.e., N will perform a quantum measurement on the next step of its computation). Define the partition $B_{x,c,h}$ of Q and the function $f_{x,c,h} : B_{x,c,h} \to C \times \{-1, 0, 1\}$ such that $\delta_{\text{measure}}(c, x_h) = (B_{x,c,h}, f_{x,c,h})$. For each $b \in B_{x,c,h}$, define $\widetilde{\gamma}_x(c, h, b) \in C$ and $d_{x,c,h,b} \in \{-1, 0, 1\}$ such that $f_{x,c,h}(b) = (\widetilde{\gamma}_x(c, h, b), d_{x,c,h,b})$ and define $\widetilde{h}_x(c, h, b) = h + d_{x,c,h,b}$. Let $P_{x,c,h,b} = \sum_{q \in b} |q\rangle\langle q| \in \text{Proj}(\mathbb{C}^Q)$ denote the projection operator corresponding to measurement outcome b. For each $b \in B_{x,c,h}$, the probability that the outcome of the quantum measurement is b is given by $\|P_{x,c,h,b}|\psi\rangle\|^2$; if the outcome is b, then the quantum register of N collapses to the superposition $\frac{1}{\|P_{x,c,h,b}|\psi\rangle\|} P_{x,c,h,b}|\psi\rangle$. Therefore, after performing the above quantum measurement, N is in an ensemble of configurations $\{(\|P_{x,c,h,b}|\psi\rangle\|^2, (\frac{1}{\|P_{x,c,h,b}|\psi\rangle\|} P_{x,c,h,b}|\psi\rangle, \widetilde{\gamma}_x(c, h, b), \widetilde{h}_x(c, h, b))) : b \in B_{x,c,h}, \|P_{x,c,h,b}|\psi\rangle\| \neq 0\}$.

We have made the above definitions of $B_{x,c,h}$, $\widetilde{\gamma}_x(c, h, b)$, 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 $B_{x,c,h} = \{Q\}$, $P_{x,c,h,Q} = 1_{\mathbb{C}^Q}$, $\widetilde{\gamma}_x(c, h, Q) = c$, and $\widetilde{h}_x(c, h, Q) = h$; 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 $b \in B_{x,c,h}$, let $\widetilde{E}_{x,c,h,b} = P_{x,c,h,b} \otimes |\widetilde{\gamma}_x(c, h, b)\rangle\langle c| \otimes |\widetilde{h}_x(c, h, b)\rangle\langle h| \in \text{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$. We then define the operator $M_x : \text{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \text{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ such that $M_x(Z) = \sum_{c \in C, h \in H_x} \sum_{b \in B_{x,c,h}} \widetilde{E}_{x,c,h,b} Z \widetilde{E}^\dagger_{x,c,h,b}$, $\forall Z \in \text{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$.

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

(i) For any $Z \in \widehat{\text{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$, 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 $M_x \in \text{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$.

Proof. (i) Any $Z \in \widehat{\text{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ 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 $A : C \times H_x \to \text{Den}(\mathbb{C}^Q)$. For each $\widehat{c} \in C$ and each $\widehat{h} \in H_x$, let $Z_{\widehat{c},\widehat{h}} = A(\widehat{c}, \widehat{h}) \otimes |\widehat{c}\rangle\langle\widehat{c}| \otimes |\widehat{h}\rangle\langle\widehat{h}| \in \widehat{\text{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$. Let $D_{\widetilde{\gamma}_x(\widehat{c},\widehat{h},b)} = |\widetilde{\gamma}_x(\widehat{c}, \widehat{h}, b)\rangle\langle\widetilde{\gamma}_x(\widehat{c}, \widehat{h}, b)| \in \widehat{\text{Den}}(\mathbb{C}^C)$ and let $D_{\widetilde{h}_x(\widehat{c},\widehat{h},b)} = |\widetilde{h}_x(\widehat{c}, \widehat{h}, b)\rangle\langle\widetilde{h}_x(\widehat{c}, \widehat{h}, b)| \in \widehat{\text{Den}}(\mathbb{C}^{H_x})$.

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 $b \in B_{x,\widehat{c},\widehat{h}}$ is given by $\text{Tr}\left(P_{x,\widehat{c},\widehat{h},b} A(\widehat{c}, \widehat{h}) P^\dagger_{x,\widehat{c},\widehat{h},b}\right)$; if

the outcome is b, the ensemble of configurations of the quantum register will collapse to an ensemble described by $\frac{1}{\mathrm{Tr}\left(P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b}\right)}P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b}$.

Let $\widetilde{B}_{x,\widehat{c},\widehat{h},A(\widehat{c},\widehat{h})} = \left\{b \in B_{x,\widehat{c},\widehat{h}} : \mathrm{Tr}\left(P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b}\right) \neq 0\right\}$ denote those measurement outcomes that occur with non-zero probability. Note that $P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b} \in \mathrm{Den}(\mathbb{C}^Q) \subseteq \mathrm{Pos}(\mathbb{C}^Q)$, and so all eigenvalues of the operator $P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b}$ are non-negative real numbers. If we have $\mathrm{Tr}\left(P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b}\right) = 0$, then the operator $P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b}$ has only the eigenvalue 0 (with multiplicity $|Q|$), which implies that $P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b} = 0_{\mathbb{C}^Q}$ if $\mathrm{Tr}\left(P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b}\right) = 0$. 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

$$Z'_{\widehat{c},\widehat{h}} = \sum_{b \in \widetilde{B}_{x,\widehat{c},\widehat{h},A(\widehat{c},\widehat{h})}} \frac{\mathrm{Tr}\left(P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b}\right)}{\mathrm{Tr}\left(P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b}\right)} P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b} \otimes D_{\widetilde{\gamma}_x(\widehat{c},\widehat{h},b)} \otimes D_{\widetilde{h}_x(\widehat{c},\widehat{h},b)}$$

$$= \sum_{b \in \widetilde{B}_{x,\widehat{c},\widehat{h},A(\widehat{c},\widehat{h})}} P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b} \otimes D_{\widetilde{\gamma}_x(\widehat{c},\widehat{h},b)} \otimes D_{\widetilde{h}_x(\widehat{c},\widehat{h},b)}$$

$$= \sum_{b \in B_{x,\widehat{c},\widehat{h}}} P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b} \otimes D_{\widetilde{\gamma}_x(\widehat{c},\widehat{h},b)} \otimes D_{\widetilde{h}_x(\widehat{c},\widehat{h},b)}.$$

Next, suppose instead that it is *not* the case that $\widehat{c} \in \widehat{C}$, $\widehat{h} \in \widehat{H}_x$, and $\delta_{\mathrm{type}}(\widehat{c}, x_{\widehat{h}}) = \mathit{measure}$. We then define $Z'_{\widehat{c},\widehat{h}} = \sum_{b \in B_{x,\widehat{c},\widehat{h}}} P_{x,\widehat{c},\widehat{h},b}A(\widehat{c},\widehat{h})P^{\dagger}_{x,\widehat{c},\widehat{h},b} \otimes D_{\widetilde{\gamma}_x(\widehat{c},\widehat{h},b)} \otimes D_{\widetilde{h}_x(\widehat{c},\widehat{h},b)}$, 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

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

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

$$M_x(Z) = \sum_{\substack{c \in C \\ h \in H_x \\ b \in B_{x,c,h}}} \widetilde{E}_{x,c,h,b} Z \widetilde{E}^{\dagger}_{x,c,h,b}$$

$$= \sum_{\substack{c \in C \\ h \in H_x \\ b \in B_{x,c,h}}} \widetilde{E}_{x,c,h,b} \left(\sum_{\substack{\widehat{c} \in C \\ \widehat{h} \in H_x}} p(\widehat{c},\widehat{h}) A(\widehat{c},\widehat{h}) \otimes |\widehat{c}\rangle\langle\widehat{c}| \otimes |\widehat{h}\rangle\langle\widehat{h}| \right) \widetilde{E}^{\dagger}_{x,c,h,b}$$

$$= \sum_{\substack{c,\widehat{c}\in C\\ h,\widehat{h}\in H_x\\ b\in B_{x,c,h}}} F(\widehat{c},\widehat{h})\otimes|\widetilde{\gamma}_x(\widehat{c},\widehat{h},b)\rangle\langle c|\widehat{c}\rangle\langle\widehat{c}|c\rangle\langle\widetilde{\gamma}_x(\widehat{c},\widehat{h},b)|\otimes|\widetilde{h}_x(\widehat{c},\widehat{h},b)\rangle\langle h|\widehat{h}\rangle\langle\widehat{h}|h\rangle\langle\widetilde{h}_x(\widehat{c},\widehat{h},b)|$$

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

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

$$\widetilde{E}^\dagger_{x,c,h,b}\widetilde{E}_{x,c,h,b} = (P^\dagger_{x,c,h,b}\otimes|c\rangle\langle\widetilde{\gamma}_x(c,h,b)|\otimes|h\rangle\langle\widetilde{h}_x(c,h,b)|)(P_{x,c,h,b}\otimes|\widetilde{\gamma}_x(c,h,b)\rangle\langle c|\otimes|\widetilde{h}_x(c,h,b)\rangle\langle h|)$$

$$= P^\dagger_{x,c,h,b}P_{x,c,h,b}\otimes|c\rangle\langle\widetilde{\gamma}_x(c,h,b)|\widetilde{\gamma}_x(c,h,b)\rangle\langle c|\otimes|h\rangle\langle\widetilde{h}_x(c,h,b)|\widetilde{h}_x(c,h,b)\rangle\langle h|$$
$$= P_{x,c,h,b}\otimes|c\rangle\langle c|\otimes|h\rangle\langle h|.$$

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

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

By [41, 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 $N = (Q, C, \Sigma, \delta_{\mathrm{type}}, \delta_{\mathrm{transform}}, \delta_{\mathrm{measure}}, q_{\mathrm{start}}, c_{\mathrm{start}}, c_{\mathrm{acc}}, c_{\mathrm{rej}})$ 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 $\widehat{E}_{x,c,h} = 1_{\mathbb{C}^Q}\otimes|c_{\mathrm{rej}}\rangle\langle c|\otimes|h\rangle\langle h|$, 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 $T_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ [41, 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 [12] 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 $N = (Q, C, \Sigma, \delta_{\mathrm{type}}, \delta_{\mathrm{transform}}, \delta_{\mathrm{measure}}, q_{\mathrm{start}}, c_{\mathrm{start}}, c_{\mathrm{acc}}, c_{\mathrm{rej}})$.

(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 3.2.1, 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* $N_{x,m} : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ by $N_{x,m} = \mathrm{Tr}_{\mathbb{C}^{H_x}} \circ T_x \circ (K_x \circ M_x)^m \circ K_x \circ I_x$.

(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 $\{(p_i, (|\psi_i\rangle, c_i, n'+1)) : i \in I\}$ (i.e., the head position of every configuration is $n'+1 = |x|+1$, 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 $\widetilde{N}_{y,m} : \mathrm{L}(\mathbb{C}^Q\otimes\mathbb{C}^C\otimes\mathbb{C}^{H_x}) \to \mathrm{L}(\mathbb{C}^Q\otimes\mathbb{C}^C\otimes\mathbb{C}^{H_x})$ 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) $(|q_{\mathrm{start}}\rangle, c_{\mathrm{start}})$ 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 $\widetilde{N}_{y,m}$ and $N_{x,m}$. To be precise,

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

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 $n' = |x|$, the rightmost symbol of $\#_L x$), then the m-truncated stopping ensemble is described by $N_{x,m}(Z)$.

(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 $n'+1$, the leftmost symbol of $y\#_R$), then the m-truncated stopping ensemble is described by $\widetilde{N}_{y,m}(Z)$.

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

Proof. (i) For any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, let $\{(p_i, (|\psi_i\rangle, c_i, n')) : i \in I\}$ be any ensemble of configurations described by $I_x(Z)$. By Lemma 3.2(i), $\{(p_i, (t_x(c_i, n')|\psi_i\rangle, \gamma_x(c_i, n'), h_x(c_i, n'))) : i \in I\}$, 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 $\mathrm{keyEv}_x(c_i, n') \in \{(1), (2)_{\mathrm{halt}}, (2)_{\mathrm{loop}}\}$, the configuration $(t_x(c_i, n')|\psi_i\rangle, \gamma_x(c_i, n'), h_x(c_i, n')))$ 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 $\mathrm{keyEv}_x(c_i, n') = (3)$), the configuration $(t_x(c_i, n')|\psi_i\rangle, \gamma_x(c_i, n'), h_x(c_i, n')))$ 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 $\mathrm{Tr}_{\mathbb{C}^{H_x}}(T_x(K_x(I_x(Z)))) = N_{x,0}(Z)$ 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 $\{(p_i, (t_x(c_i, n')|\psi_i\rangle, \gamma_x(c_i, n'), h_x(c_i, n'))) : i \in I\}$ by performing a single quantum measurement on appropriate configurations (i.e., for each $i \in I$ such that $\gamma_x(c_i, n') \in \widehat{C}$, $h_x(c_i, n') \in \widehat{H}_x$, and $\delta_{\mathrm{type}}(\gamma_x(c_i, n'), x_{h_x(c_i, n')}) = \mathit{measure}$, we replace the configuration $(t_x(c_i, n')|\psi_i\rangle, \gamma_x(c_i, n'), h_x(c_i, n'))$ 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 $\mathrm{Tr}_{\mathbb{C}^{H_x}}(T_x(K_x(M_x(K_x(I_x(Z)))))) = N_{x,1}(Z)$ describes the 1-truncated stopping ensemble, as desired. If $m > 1$, then by continuing in this fashion, we conclude that $N_{x,m}(Z)$ 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), $N_{x,m} = \mathrm{Tr}_{\mathbb{C}^{H_x}} \circ T_x \circ (K_x \circ M_x)^m \circ K_x \circ I_x$. 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, [41, Section 2.2]).

□

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.

For the sake of completeness, we now define "non-truncated" versions of transfer operators and crossing sequences, which are, in some sense "nicer" than their m-truncated counterparts. It is straightforward to see that the argument used in Section 3.3 to establish a lower bound on the expected running time of a 2QCFA would apply, essentially identically, to non-truncated crossing sequences. We omit the details.

Definition 3.9. 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 $x \in \Sigma^*$ and any $m \in \mathbb{N}$, define $N_{x,m}$ as in Definition 3.9(i). By Lemma 3.8(iii), $N_{x,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, $\forall x \in \Sigma^*$, $\forall m \in \mathbb{N}$. We define the *non-truncated transfer operator* $N_x \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ as an accumulation point of the sequence $(N_{x,m})_{m\in\mathbb{N}}$; such an accumulation point exists due to the fact that $\mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ is compact (see, for instance, [41, Proposition 2.28]).

(ii) For any $y \in \Sigma^*$ and any $m \in \mathbb{N}$, define $\widetilde{N}_{y,m}$ as in Definition 3.9(ii). We define the "dual" *non-truncated transfer operator* $\widetilde{N}_y \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ as an accumulation point of the sequence $(\widetilde{N}_{y,m})_{m\in\mathbb{N}}$

(iii) For any $x, y \in \Sigma^*$ and any $m \in \mathbb{N}$, we then define the *non-truncated crossing sequence* of N with respect to the (partitioned) input xy to be the sequence $Z_1, Z_2, \ldots \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, where

$$Z_i = \begin{cases} |q_{\text{start}}\rangle\langle q_{\text{start}}| \otimes |c_{\text{start}}\rangle\langle c_{\text{start}}|, & i = 1 \\ \widetilde{N}_y(Z_{i-1}), & i > 1, i \text{ is even} \\ N_x(Z_{i-1}), & i > 1, i \text{ is odd.} \end{cases}$$

3.3 A 2QCFA Analogue of the Dwork-Stockmeyer Lemma

Dwork and Stockmeyer proved a lower bound [12, 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 $D_L(n)$. In this section, 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 [15] that showed that a 2PFA cannot recognize $L_{eq} = \{a^m b^m : m \in \mathbb{N}\}$ with bounded-error in subexponential time.

We begin by recalling the definition of $D_L(n)$, following [12]. 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. Consider two words $w, w' \in \Sigma^*$ such that $|w| \leq n$ and $|w'| \leq n$. We say that w and w' are (L,n)-*similar*, which we denote by writing $w \sim_{L,n} w'$, if, $\forall v \in \Sigma^*$ such that $|wv| \leq n$ and $|w'v| \leq n$, we have $wv \in L \Leftrightarrow w'v \in L$. We say that w and w' (where we continue to suppose that $|w| \leq n$ and $|w'| \leq n$) are (L,n)-*dissimilar*, which we denote by writing $w \not\sim_{L,n} w'$, if they are not (L,n)-similar; i.e., $w \not\sim_{L,n} w'$ if $\exists v \in \Sigma^*$ such that $|wv| \leq n$, $|w'v| \leq n$ and $wv \in L \Leftrightarrow w'v \notin L$. 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 $\exists w_1, \ldots, w_h \in \Sigma^*$ that are pairwise (L,n)-dissimilar (i.e., $\forall i, j$ with $i \neq j$, $w_i \not\sim_{L,n} w_j$).

In the Dwork and Stockmeyer [12] lower bound on the expected running time of any 2PFA that recognizes a non-regular 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 $N = (Q, C, \Sigma, \delta_{\text{type}}, \delta_{\text{transform}}, \delta_{\text{measure}}, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$ that recognizes L, with two-sided bounded-error $\epsilon \in \mathbb{R}_{>0}$, in expected running time at most $T(n)$ on all inputs of length n. For any $n \in \mathbb{N}$, consider $x, x' \in \Sigma^*$ such that $x \not\sim_{L,n} x'$. Fix some $y \in \Sigma^*$, such that $|xy| \leq n$, $|x'y| \leq n$, and $xy \in L \Leftrightarrow x'y \notin L$. 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 $N_{x,m}$, $N_{x',m}$, and $\widetilde{N}_{y,m}$ as in Definition 3.7. By Lemma 3.8(iii), $N_{x,m}, N_{x',m}, \widetilde{N}_{y,m} \in \text{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. We define a distance metric on $\text{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. We show that, if $D_L(n)$ is "large", then, for any m, we can find x, x' such that $x \not\sim_{L,n} x'$ and the distance between $N_{x,m}$ and $N_{x',m}$ is "small." We also show that, for m sufficiently large, if the distance between $N_{x,m}$ and $N_{x',m}$ 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 $p_N(xy) \approx p_N(x'y)$. However, as $xy \in L$, we must have $p_N(xy) \geq 1 - \epsilon$, and as $x'y \notin L$, we must have $p_N(x'y) \leq \epsilon$, 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 the above idea.

We begin by recalling the definitions of several needed norms; see, for instance [41, Section 1.1.3] for further background. We continue to use the notation established in Section 2.1. Consider a finite-dimensional complex Hilbert space V, and let $\text{L}(V)$ denote the space of $\mathbb{C}$-linear maps on V. For $p \in \mathbb{N}_{\geq 1}$, we define the *Schatten p-norm* $\|\cdot\|_p : \text{L}(V) \to \mathbb{R}_{\geq 0}$, where $\|Z\|_p = (\text{Tr}((Z^\dagger Z)^{\frac{p}{2}}))^{\frac{1}{p}}$, $\forall Z \in \text{L}(V)$. 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 $Z \in \text{L}(V)$. Similarly, we use the term *Hilbert-Schmidt norm* to refer to the Schatten 2-norm, and we note that $\|Z\|_2 = \sqrt{\sum_{i,j \in B} |\langle e_i|Z|e_j\rangle|^2}$, $\forall Z \in \text{L}(V)$, where $\{|e_i\rangle : i \in B\}$ is an orthonormal basis of V. We write $\mathbb{C}^{r \times c}$ to denote the space of $r \times c$ matrices with entries in $\mathbb{C}$. We may encode any operator $Z \in L(V)$ as an element of $\mathbb{C}^{\dim(V) \times \dim(V)}$ by choosing some basis of V; we define the above norms on the space of matrices identically. For finite-dimensional complex Hilbert spaces V and V', we again write $\text{T}(V, V')$ for the space of $\mathbb{C}$-linear maps of the form $\Phi : \text{L}(V) \to \text{L}(V')$. We define the *induced trace norm* $\|\cdot\|_1 : \text{T}(V, V') \to \mathbb{R}_{\geq 0}$, where $\|\Phi\|_1 = \sup\{\|\Phi(Z)\|_1 : Z \in \text{L}(V), \|Z\|_1 \leq 1\}$, for any $\Phi \in \text{T}(V, V')$. Observe that the induced trace norm is also a norm.

For density operators $Z, Z' \in \text{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, 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 $\|N_{x,m} - N_{x',m}\|_1$, the distance metric induced by the induced trace norm, to measure the distance

between $N_{x,m}$ and $N_{x',m}$; for m sufficiently large, this will serve as our measure of the distance between x and x'. Suppose N is run on two distinct partitioned inputs xy and $x'y$, producing two distinct m-truncated crossing sequences. We next show that if $\|N_{x,m} - N_{x',m}\|_1$ is "small", then these crossing sequences are similar.

Lemma 3.10. *Consider a 2QCFA $N = (Q, C, \Sigma, \delta_{type}, \delta_{transform}, \delta_{measure}, q_{start}, c_{start}, c_{acc}, c_{rej})$. For $x, x', y \in \Sigma^*$ and $m \in \mathbb{N}$, following Definition 3.7(iii), let $Z_1, Z_2, \ldots \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ denote the m-truncated crossing sequence when N is run on xy and let $Z'_1, Z'_2, \ldots \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ denote the m-truncated crossing sequence when N is run on $x'y$. Then $\|Z_i - Z'_i\|_1 \leq \lfloor \frac{i-1}{2} \rfloor \|N_{x,m} - N_{x',m}\|_1$, $\forall i \in \mathbb{N}_{\geq 1}$.*

Proof. Note that $\|\Phi(Z)\|_1 \leq \|Z\|_1$, $\forall Z \in \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, $\forall \Phi \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ (see, for instance, [41, 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

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

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), $N_{x,m}, N_{x',m}, \widetilde{N}_{y,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$.

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, $Z_i = \widetilde{N}_{y,m}(Z_{i-1})$ and $Z'_i = \widetilde{N}_{y,m}(Z'_{i-1})$. By the above observation concerning the contractivity of the trace norm, we then have

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

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

$$\|Z_i - Z'_i\|_1 = \|N_{x,m}(Z_{i-1}) - N_{x',m}(Z'_{i-1})\|_1 \leq \|N_{x,m}(Z_{i-1}) - N_{x,m}(Z'_{i-1})\|_1 + \|N_{x,m}(Z'_{i-1}) - N_{x',m}(Z'_{i-1})\|_1$$

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

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

Lemma 3.11. *Consider a language L over some finite alphabet Σ. Suppose $L \in \mathsf{B2QCFA}(k, d, T, \epsilon)$, for some $k, d \in \mathbb{N}$, $T : \mathbb{N} \to \mathbb{N}$, and $\epsilon \in [0, \frac{1}{2})$. Suppose further that, for some $n \in \mathbb{N}$, $\exists x, x' \in \Sigma^*$ such that $x \not\sim_{L,n} x'$. Then $T(n) \geq \frac{(1-2\epsilon)^2}{2} \|N_{x,m} - N_{x',m}\|_1^{-1}$, for any $m \geq \lceil \frac{2}{1-2\epsilon} T(n) \rceil$.*

Proof. By definition, $x \not\sim_{L,n} x'$ precisely when $\exists y \in \Sigma^*$ such that $|xy| \leq n, |x'y| \leq n$, 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 $s \in \mathbb{N}_{\geq 1}$, 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,

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

By Lemma 3.10, $\|Z_{m,s} - Z'_{m,s}\|_1 \leq \frac{s-1}{2} \|N_{x,m} - N_{x',m}\|_1$, $\forall s \in \mathbb{N}_{\geq 1}$, and so we conclude

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

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. As before, we write $p_N(w)$ to denote the probability that N accepts an input $w \in \Sigma^*$. We then define $p_N(w,s)$ to be the probability that N accepts w within s steps, and we define $h_N(w,s)$ to be 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

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

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

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

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

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

Therefore,

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

□

The following lemma shows that any "large" set of input prefixes contains a pair of input prefixes at "small" distance from one another.

Lemma 3.12. *For every $k, d \in \mathbb{N}_{\geq 1}$, there is a constant $K_{k,d} \in \mathbb{R}_{>0}$, such that the following holds. Suppose N is a 2QCFA with quantum basis states Q, classical states C, and alphabet Σ, where $|Q| = k$ and $|C| = d$. Then $\forall m \in \mathbb{N}$, $\forall X \subseteq \Sigma^*$ such that X is finite and $|X| \geq 2$, $\exists x, x' \in X$ such that $x \neq x'$ and $\|N_{x,m} - N_{x',m}\|_1 \leq K_{k,d}|X|^{-\frac{1}{k^4 d^2}}$.*

Proof. We will observe that, for any 2QCFA N of the assumed type, and any $m \in \mathbb{N}$, there is a function $f_{N,m} : \Sigma^* \to [-\frac{1}{2}, \frac{1}{2}]^{k^4 d^2}$ such that $\|N_{x,m} - N_{x',m}\|_1 \leq \sqrt{2}kd\|f_{N,m}(x) - f_{N,m}(x')\|$, where here $\|f_{N,m}(x) - f_{N,m}(x')\|$ denotes the Euclidean distance between $f_{N,m}(x), f_{N,m}(x') \in [-\frac{1}{2}, \frac{1}{2}]^{k^4 d^2} \subseteq \mathbb{R}^{k^4 d^2}$. The claim will then straightforwardly follow.

For any $x \in \Sigma^*$, let $H_x = \{0, \ldots, |x|+1\}$, let $I_x : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ be as defined in Section 3.2.1, and let $K_x : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ be as defined in Definition 3.1; for any $m \in \mathbb{N}$, let $N_{x,m} : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C) \to \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x})$ be as defined in Definition 3.7(i). 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

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

By Lemma 3.8(iii), $N_{x,m} \in \mathrm{Chan}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. It is then straightforward to see that $J(N_{x,m}) \in \mathrm{Pos}(\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^Q \otimes \mathbb{C}^C)$ (see, for instance, [41, Corollary 2.27]); we next observe that $J(N_{x,m})$ is of a special form. Consider any $x \in \Sigma^*$, $m \in \mathbb{N}$, $q, q' \in Q$, and $c \in C$. We have $F_{q,q',c,c} \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$, which implies $N_{x,m}(F_{q,q',c,c}) \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. Consider any $c' \in C \setminus \{c\}$. By inspection, $K_x(I_x(F_{q,q',c,c'})) = 0_{\mathbb{C}^Q \otimes \mathbb{C}^C \otimes \mathbb{C}^{H_x}}$, which implies that $N_{x,m}(F_{q,q',c,c'}) = 0_{\mathbb{C}^Q \otimes \mathbb{C}^C}$, in this case. Therefore, $\forall x \in \Sigma^*$, $\forall m \in \mathbb{N}$, we have

$$J(N_{x,m}) = \sum_{\substack{q,q' \in Q \\ c \in C}} F_{q,q',c,c} \otimes N_{x,m}(F_{q,q',c,c}), \quad \text{with} \quad N_{x,m}(F_{q,q',c,c}) \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C),\ \forall q, q' \in Q, \forall c \in C.$$

Recall that any $Z \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ is of the form $Z = \sum_{c \in C} p(c) A(c) \otimes |c\rangle\langle c|$, for some $p : C \to [0,1]$ and $A : C \to \mathrm{Den}(\mathbb{C}^Q)$, where $\sum_{c \in C} p(c) = 1$. Furthermore, recall that $\mathrm{Den}(\mathbb{C}^Q) = \{A \in \mathrm{Pos}(\mathbb{C}^Q) : \mathrm{Tr}(A) = 1\}$ and $\mathrm{Pos}(\mathbb{C}^Q) \subseteq \mathrm{Herm}(\mathbb{C}^Q)$. We next encode $J(N_{x,m})$ as a matrix $M_{J(N_{x,m})}$ in the natural way. For a matrix M, we write $M[i,j]$ to denote the entry in row i and column j. The matrix $M_{J(N_{x,m})}$ has dimensions $(k^2d^2) \times (k^2d^2)$ and the set of rows and the set of columns of $M_{J(N_{x,m})}$ are each indexed by $Q \times C \times Q \times C$. For $q_1, q_2, q_1', q_2' \in Q$ and $c_1, c_2, c_1', c_2' \in C$, we define $M_{J(N_{x,m})}[(q_1, c_1, q_2, c_2), (q_1', c_1', q_2', c_2')] = (\langle q_2| \otimes \langle c_2|) N_{x,m}(F_{q_1,q_1',c_1,c_1'})(|q_2'\rangle \otimes |c_2'\rangle)$. By the above, if $c_1 \neq c_1'$ or $c_2 \neq c_2'$, then $M_{J(N_{x,m})}[(q_1, c_1, q_2, c_2), (q_1', c_1', q_2', c_2')] = 0$, for any $N_{x,m}$. Let $E_{Q,C} = \{((q_1, c_1, q_2, c_2), (q_1', c_1, q_2', c_2)) : q_1, q_2, q_1', q_2' \in Q, c_1, c_2 \in C\}$. Then $M_{J(N_{x,m})}$ may be non-zero only at entries specified by $E_{Q,C}$, $\forall N_{x,m}$. Notice that $|E_{Q,C}| = k^4d^2$.

For $k, d \in \mathbb{N}$, let $\mathcal{M}_{k,d}$ denote the set of all matrices $M_{J(N_{x,m})}$ for any 2QCFA N with quantum basis states Q, classical states C, and alphabet Σ, where $|Q| = k$ and $|C| = d$, any $x \in \Sigma^*$, and any $m \in \mathbb{N}$. Notice that any $M \in \mathcal{M}_{k,d}$ is positive semi-definite and has $\mathrm{Tr}(M) = 1$, which implies that $M[i,i] \in [0,1]$, $\forall i$. For any row i and column j, where $i \neq j$, consider the 2×2 induced sub-matrix $\widetilde{M}_{i,j}$ of M obtained by deleting all rows other than i and j and deleting all columns other than i and j (i.e., the first row of $\widetilde{M}_{i,j}$ is given by $M[i,i]$ and $M[i,j]$, the second row of $\widetilde{M}_{i,j}$ is given by $M[j,i]$ and $M[j,j]$). Note that the eigenvalues of $\widetilde{M}_{i,j}$ interlace those of M, which implies $0 \leq \det(\widetilde{M}_{i,j}) = M[i,i]M[j,j] - |M[i,j]|^2$. As $M[i,i], M[j,j] \in [0,1]$ and $M[i,i] + M[j,j] \leq \mathrm{Tr}(M) = 1$, we have $M[i,i]M[j,j] \leq \frac{1}{4}$. Therefore, $|M[i,j]| \leq \frac{1}{2}$, $\forall i, j$ where $i \neq j$; that is to say, the off-diagonal entries of M all lie in $\{\alpha \in \mathbb{C} : |\alpha| \leq \frac{1}{2}\}$. We define the map $R_{k,d} : \mathcal{M}_{k,d} \to [-\frac{1}{2}, \frac{1}{2}]^{k^4d^2}$ as follows. For $M \in \mathcal{M}_{k,d}$, $R_{k,d}(M)$ is the vector of real numbers whose first k^2d^2 entries are given by $\{M[((q_1, c_1, q_2, c_2), (q_1, c_1, q_2, c_2))] - \frac{1}{2} : q_1, q_2 \in Q, c_1, c_2 \in C\}$ (i.e., the diagonal entries of M, offset by $\frac{1}{2}$, for notational convenience) and whose remaining $k^4d^2 - k^2d^2$ entries are given by encoding each of the $\frac{1}{2}(k^4d^2 - k^2d^2)$ entries in $E_{Q,C}$ above the main diagonal of M as the pair of real numbers that comprise the real and imaginary part of this entry.

For any 2QCFA N of the assumed form and any $m \in \mathbb{N}$, we define the function $f_{N,m} : \Sigma^* \to [-\frac{1}{2}, \frac{1}{2}]^{k^4d^2}$ such that $f_{N,m}(x) = R_{k,d}(M_{J(N_{x,m})})$, $\forall x \in \Sigma^*$. For any $x, x' \in \Sigma^*$, we then have

$$\|N_{x,m} - N_{x',m}\|_1 \leq \|J(N_{x,m} - N_{x',m})\|_1 = \|J(N_{x,m}) - J(N_{x',m})\|_1 = \|M_{J(N_{x,m})} - M_{J(N_{x',m})}\|_1$$

$$\leq kd\|M_{J(N_{x,m})} - M_{J(N_{x',m})}\|_2 \leq \sqrt{2}kd\|f_{N,m}(x) - f_{N,m}(x')\|.$$

To complete the proof, let $h = k^4d^2$. For any $\delta \in \mathbb{R}_{>0}$, we define $B_{x,\delta} = \{v \in \mathbb{R}^h : \|f_{N,m}(x) - v\| \leq \delta\}$ to be the closed ball centered at $f_{N,m}(x)$ of radius δ in $\mathbb{R}^h$, and we also define $S_\delta = [-(\frac{1}{2}+\delta), \frac{1}{2}+\delta]^h$. The volumes of these regions are given by $\text{vol}(S_\delta) = (1+2\delta)^h$ and $\text{vol}(B_{x,\delta}) = r_h\delta^h$, where $r_h = \frac{\pi^{\frac{h}{2}}}{(\frac{h}{2})!}$ if h is even, and $r_h = \frac{2(\frac{h-1}{2}!)(4\pi)^{\frac{h-1}{2}}}{h!}$ if h is odd. Suppose $\forall x, x' \in X$ with $x \neq x'$, we have $B_{x,\delta} \cap B_{x',\delta} = \emptyset$. Then $\sqcup_{x\in X} B_{x,\delta} \subseteq S_\delta$, which implies $|X|\text{vol}(B_{x,\delta}) \leq \text{vol}(S_\delta)$. Therefore, there is a constant $l_h \in \mathbb{R}_{>0}$ (that depends only on h) such that (if, as assumed, $|X| \geq 2$) $\exists x, x' \in X$, with $x \neq x'$, such that $B_{x,\delta} \cap B_{x',\delta} \neq \emptyset$, where $\delta = l_h|X|^{-\frac{1}{h}}$. Fix such an x, x', then $B_{x,\delta} \cap B_{x',\delta} \neq \emptyset$ implies $\|f_{N,m}(x) - f_{N,m}(x')\| \leq 2\delta$, which in turn implies

$$\|N_{x,m} - N_{x',m}\|_1 \leq \sqrt{2}kd\|f_{N,m}(x) - f_{N,m}(x')\| \leq 2\sqrt{2}kd\delta = 2\sqrt{2}kdl_h|X|^{-\frac{1}{h}}.$$

Therefore, the claim holds with the constant $K_{k,d} = 2\sqrt{2}kdl_h \in \mathbb{R}_{>0}$. □

We now prove a 2QCFA analogue of the Dwork and Stockmeyer lemma [12, Lemma 4.3]; that is to say, we show that if a 2QCFA recognizes some language L with two-sided 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 their hardness measure $D_L(n)$ (defined at the beginning of this section).

Theorem 3.13. *For every $k, d \in \mathbb{N}$ and every $\epsilon \in [0, \frac{1}{2})$, there is a constant $\widehat{K}_{k,d,\epsilon} \in \mathbb{R}_{>0}$ such that, if $L \in \mathsf{B2QCFA}(k, d, T, \epsilon)$, then there is a constant $N_0 \in \mathbb{N}$ such that $T(n) \geq \widehat{K}_{k,d,\epsilon} D_L(n)^{\frac{1}{k^4d^2}}$, $\forall n \geq N_0$.*

Proof. Fix $k, d \in \mathbb{N}$ and define $K_{k,d} \in \mathbb{R}_{>0}$ as in Lemma 3.12. Fix $\epsilon \in [0, \frac{1}{2})$. We will show the claim holds with $\widehat{K}_{k,d,\epsilon} = \frac{(1-2\epsilon)^2}{2K_{k,d}}$. Consider some language L over some finite alphabet Σ. By [12, 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 $L \notin \mathsf{REG}$.

For each $n \in \mathbb{N}$, we define $X_n = \{x_1, \cdots, x_{D_L(n)}\} \subseteq \Sigma^*$ 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 $D_L(N_0) \geq 2$. Then, $\forall n \geq N_0$, we have $|X_n| = D_L(n) \geq D_L(N_0) \geq 2$. Suppose $L \in \mathsf{B2QCFA}(k, d, T, \epsilon)$. By definition, there is some 2QCFA $N = (Q, C, \Sigma, \delta_{\text{type}}, \delta_{\text{transform}}, \delta_{\text{measure}}, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$, with $|Q| = k$ and $|C| = d$, that recognizes $L \subseteq \Sigma^*$ with two-sided bounded error ϵ in expected time at most $T(n)$ on all inputs of length at most n, $\forall n \in N$. Fix $n \geq N_0$ and set $m = \lceil \frac{1-2\epsilon}{2} T(n) \rceil$. By Lemma 3.12, $\exists x, x' \in X_n$ such that $x \neq x'$ and

$$\|N_{x,m} - N_{x',m}\|_1 \leq K_{k,d}|X_n|^{-\frac{1}{k^4d^2}} = K_{k,d} D_L(n)^{-\frac{1}{k^4d^2}}.$$

Fix such a pair x, x', and note that $x \not\sim_{L,n} x'$, by construction. By Lemma 3.11,

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

□

3.4 2QCFA Time Complexity Classes

Theorem 3.13 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 $\mathsf{B2QCFA}(k,T) = \cup_{d\in\mathbb{N}_{\geq 1},\epsilon\in[0,\frac{1}{2})}\mathsf{B2QCFA}(k,d,T,\epsilon)$ to be the class of languages recognized with two-sided bounded-error by a 2QCFA with k quantum basis states and any finite number of classical states, in expected time at most $T(n)$ on all inputs of length n. We then define $\mathsf{BQE2QCFA}(k) = \cup_{c\in\mathbb{N}}\mathsf{B2QCFA}(k,2^{cn})$ (resp. $\mathsf{BQP2QCFA}(k) = \cup_{c\in\mathbb{N}}\mathsf{B2QCFA}(k,n^c)$) to be the class of languages recognized in expected exponential (resp. polynomial) time by a 2QCFA with k quantum basis states. We also define $\mathsf{B2QCFA}(T) = \cup_{k\in\mathbb{N}_{\geq 1}}\mathsf{B2QCFA}(k,T)$ to be the class of languages recognized with two-sided bounded-error by a 2QCFA with any finite number of quantum states and any finite number of classical states, in expected time at most $T(n)$ on all inputs of length n. We then define $\mathsf{BQE2QCFA} = \cup_{k\in\mathbb{N}_{\geq 1}}\mathsf{BQE2QCFA}(k)$ (resp. $\mathsf{BQP2QCFA} = \cup_{k\in\mathbb{N}_{\geq 1}}\mathsf{BQP2QCFA}(k)$) to be the class of languages recognized in expected exponential (resp. polynomial) time by a 2QCFA of any finite size.

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, $\forall w \notin L$, $\Pr[N \text{ rejects } w] \geq 1-\epsilon$. We define $\mathsf{coR2QCFA}(k,d,T,\epsilon)$ as the class of languages that are recognized 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 $\mathsf{coR2QCFA}(k,T)$, $\mathsf{coRQE2QCFA}(k)$, etc., analogously to the two-sided bounded-error case.

We use the standard big O, little o, Ω, etc. notation to denote the asymptotic behavior of functions. Notice that the hardness measure D_L defined by Dwork and Stockmeyer [12] (see the beginning of Section 3.3 for the definition of this function) satisfies $D_L(n) = O(2^{cn})$, for any language $L \subseteq \Sigma^*$, where $c = \max(\log|\Sigma|, 1)$. We immediately obtain the following corollary of Theorem 3.13.

Corollary 3.13.1. *For any $k,d \in \mathbb{N}_{\geq 1}$, any function $T:\mathbb{N}\to\mathbb{N}$, any $\epsilon \in [0,\frac{1}{2})$, and any language $L \in \mathsf{B2QCFA}(k,d,T,\epsilon)$, we have $D_L(n) = O(T(n)^{k^4d^2})$. In particular, for every language L such that $D_L = \Omega(2^{cn})$, for some $c \in \mathbb{R}_{>0}$, and for every function $T:\mathbb{N}\to\mathbb{N}$ such that $T(n) = 2^{o(n)}$, we have $L \notin \mathsf{B2QCFA}(T)$. Moreover, for every language $L \in \mathsf{BQP2QCFA}$, we have $D_L = O(n^c)$, for some $c \in \mathbb{R}_{\geq 0}$.*

For $w = w_1\cdots w_n \in \Sigma^*$, where each $w_i \in \Sigma$, let $w^{\text{rev}} = w_n\cdots w_1$ denote the reversal of the string w. We then consider the language $L_{pal} = \{w \in \{a,b\}^* : w = w^{\text{rev}}\}$ consisting of all palindromes over the alphabet $\{a,b\}$.

Corollary 3.13.2. *For every $T:\mathbb{N}\to\mathbb{N}$ such that $T(n) = 2^{o(n)}$, we have $L_{pal} \notin \mathsf{B2QCFA}(T)$.*

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 $|ww^{\text{rev}}| = 2n = |w'w^{\text{rev}}|$, $ww^{\text{rev}} \in L_{pal}$, and $w'w^{\text{rev}} \notin L_{pal}$; 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 3.13.1 then implies $L_{pal} \notin \mathsf{B2QCFA}(T)$. □

Ambainis and Watrous [4] showed that $L_{pal} \in \mathsf{coRQE2QCFA}(2)$. Clearly, $\mathsf{coRQE2QCFA}(T) \subseteq \mathsf{BQE2QCFA}(T)$, for any T, and $\mathsf{coRQE2QCFA} \subseteq \mathsf{BQE2QCFA}$. Therefore, we obtain the following.

Corollary 3.13.3. *For every function $T:\mathbb{N}\to\mathbb{N}$ such that $T(n) = 2^{o(n)}$, the following statements hold.*

(i) $\mathsf{B2QCFA}(T) \subsetneq \mathsf{BQE2QCFA}$.

(ii) $\mathsf{coR2QCFA}(T) \subsetneq \mathsf{coRQE2QCFA}$.

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 [36]. Our lower bound holds even in this setting of unrestricted transition amplitudes. For $\mathbb{F} \subseteq \mathbb{C}$, we define complexity classes $\mathsf{coR2QCFA}_{\mathbb{F}}(k, d, T, \epsilon)$, $\mathsf{coRQE2QCFA}_{\mathbb{F}}(k)$, 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 [4] showed that $L_{pal} \in \mathsf{coRQE2QCFA}_{\overline{\mathbb{Q}}}(2)$, 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 exponential expected running 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 finite size) in subexponential time, regardless of the permitted transition amplitudes.

Remark. Dwork and Stockmeyer showed that $\exists b \in \mathbb{N}_{\geq 1}$ such that $D_L(n) \leq b$, $\forall n \in \mathbb{N}$, if and only if $L \in \mathsf{REG}$ [12, Lemma 3.1]. Moreover, they showed that there is a "gap" in the sense that, if $L \notin \mathsf{REG}$, then $D_L(n) \geq \sqrt{n} - 1$, for infinitely many $n \in \mathbb{N}$ [12, Theorem 3.6]. This gap then enabled them to show that if a 2PFA recognizes a language L with bounded-error in expected subexponential time, then $L \in \mathsf{REG}$. However, it is certainly *not* the case that, if a 2QCFA recognizes a language L with bounded-error in expected subexponential time, then $L \in \mathsf{REG}$. For example, Ambainis and Watrous [4] showed that the language $L_{eq} = \{a^m b^m : m \in \mathbb{N}\} \in \mathsf{BQP2QCFA}(2)$, where, of course, $L_{eq} \notin \mathsf{REG}$.

3.5 The 2QCFA Groups

In this section, we consider the word problem W_G corresponding to a finitely-generated group G (see Section 2.2 for a full definition of the group word problem and relevant notation and terminology). 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 [35] 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. We note that the languages L_{pal} and L_{eq}, which Ambainis and Watrous [4] showed satisfy $L_{pal} \in \mathsf{coRQE2QCFA}_{\overline{\mathbb{Q}}}(2)$ and $L_{eq} \in \mathsf{BQP2QCFA}(2)$, are closely related to the word problems of the groups F_2 and $\mathbb{Z}$, respectively (see [35] for a full discussion of this correspondence).

We begin by defining the growth rate of a group (for a more thorough treatment of this material see, for instance, [27]). Consider a group $G = \langle S|R\rangle$, with S finite. Let $\Sigma = S \cup S^{-1}$, let Σ^* denote the free monoid on Σ, and let $\phi : \Sigma^* \to G$ denote the natural (monoid) homomorphism 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 3.14. *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 $D_{W_G}(2n) \geq k = \beta_{G,S}(n)$. □

Remark. In fact, one may also straightforwardly show that $D_{W_G}(2n) \leq \beta_{G,S}(n) + 1$, though we do not need this here.

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, [27, 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 3.15. 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. By making use of two very powerful results in group theory, the Tits' Alternative [40] and Gromov's theorem on groups of polynomial growth [17], 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 3.4. 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, $W_G \in \mathsf{BQP2QCFA}(2)$ to mean $W_{G=\langle S|R\rangle} \in \mathsf{BQP2QCFA}(2)$ for every presentation $G = \langle S|R\rangle$, with S finite.

Theorem 3.16. *For any finitely-generated group G, the following statements hold.*

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

(ii) *If* G *has exponential growth, then for every function* $T : \mathbb{N} \to \mathbb{N}$ *such that* $T(n) = 2^{o(n)}$, *we have* $W_G \notin \mathsf{B2QCFA}(T)$.

(iii) *If* G *is a linear group over a field of characteristic* 0, *and* G *is not virtually nilpotent, then for every function* $T : \mathbb{N} \to \mathbb{N}$ *such that* $T(n) = 2^{o(n)}$, *we have* $W_G \notin \mathsf{B2QCFA}(T)$.

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

Proof. (i) Follows immediately from Lemma 3.14 and Corollary 3.13.1.

(ii) Follows immediately from Definition 3.15(i) and part (i) of this theorem.

(iii) As a consequence of the famous Tits' Alternative [40], 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 ([40, Corollary 1],[43]). 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 [17], 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 [16], 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 3.16(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 $W_G \in \mathsf{coRQE2QCFA}_{\overline{\mathbb{Q}}}$ [35, 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, [39, Proposition 2.2]). We therefore immediately obtain the following corollary of Theorem 3.16(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 [35] for a full discussion and related results.

Corollary 3.16.1. *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* $W_G \notin \mathsf{B2QCFA}(T)$.

We have also recently shown that $W_G \in \mathsf{coRQP2QCFA}_{\overline{\mathbb{Q}}}(2) \subseteq \mathsf{BQP2QCFA}$, $\forall G \in \mathcal{G}_{\mathbf{vAb}}$ [35, Theorem 1.2] (i.e., the word problem of any finitely-generated virtually abelian group can be recognized with negative one-sided bounded-error by a single-qubit 2QCFA in expected polynomial time). By Theorem 3.16(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 [35] 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. 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, [24, Theorem 12] for these facts, as well as for their application towards understanding the computational complexity of the group word problem). Note 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, [24, Lemma 2]). This implies that, if $W_G \in \mathsf{BQP2QCFA}$, for some $G \in \mathcal{G}_{\mathbf{vNilp}} \setminus \mathcal{G}_{\mathbf{vAb}}$, then $W_H \in \mathsf{BQP2QCFA}$. We have therefore proven the following proposition.

Proposition 3.17. *If $W_H \notin \mathsf{BQP2QCFA}$, where H is the Heisenberg group, then for any finitely-generated group G, $W_G \in \mathsf{BQP2QCFA} \Leftrightarrow W_G \in \mathsf{coRQP2QCFA}_{\overline{\mathbb{Q}}}(2) \Leftrightarrow G \in \mathcal{G}_{\boldsymbol{vAb}}$.*

In Section 6.1, we further explore the relationship between our results concerning the class of group word problems recognizable by 2QCFA with particular resource bounds, and known results concerning the class of group word problems recognizable by various classical models of computation.

4 One-way Measure-once QFA

4.1 Definition of the 1QFA Model

We now define the measure-once one-way quantum finite automata (1QFA) model, following the original definition given by Moore and Crutchfield [29]. In particular, we follow the convention that the quantum register of a 1QFA is described by some $\langle\psi|$ (rather than $|\psi\rangle$ as in the case of 2QCFA), which reverses many other conventions. We do this both to be consistent with the notation used by Moore and Crutchfield as well as to allow 1QFA to apply operators in the "natural" left-to-right order.

Informally, the 1QFA model can be thought of as a modification of the one-way probabilistic finite automata model in which the probabilistic states are replaced by quantum states. Formally, a 1QFA is a 5-tuple $M = (V, \Sigma, \delta, \langle\psi_{\text{start}}|, V_{\text{acc}})$, where V is a finite-dimensional complex Hilbert space, Σ is a finite alphabet, $\delta : \Sigma \to \mathrm{U}(V)$ is the transition function, $\langle\psi_{\text{start}}| \in V$, with $\|\langle\psi_{\text{start}}|\| = 1$, is the initial superposition of the quantum register, and $V_{\text{acc}} \subseteq V$ is the accepting subspace of V. We define $V_{\text{rej}} = V_{\text{acc}}^{\perp}$, the orthogonal complement of V_{acc} in V, and we write P_{acc} (resp. P_{rej}) to denote the projection operator onto V_{acc} (resp. V_{rej}).

On an input string $w \in \Sigma^*$, the 1QFA M operates as follows. The quantum register of M is initially in the superposition $\langle\psi_{\text{start}}|$. Then, M reads the string w from left to right, and, when reading the symbol $\sigma \in \Sigma$, performs the transformation $\delta(\sigma)$ to its quantum register (where, of course, transformations are now applied on the right). After reading the entire string w, M performs

the quantum measurement specified by $\{P_{\text{acc}}, P_{\text{rej}}\}$. If the result of that measurement is P_{acc}, then M accepts w; otherwise, M rejects w.

For $w \in \Sigma^*$, let $\langle\psi_M(w)|$ denote the state of the quantum register of M immediately after reading the entire string w, before performing the quantum measurement. By slight abuse of notation, let $\delta : \Sigma^* \to \mathrm{U}(V)$ denote the unique monoid homomorphism induced by $\delta : \Sigma \to \mathrm{U}(V)$. We then have $\langle\psi_M(w)| = \langle\psi_{\text{start}}|\delta(w)$. We use $p_M(w)$ to denote the probability that M accepts w, where we have $p_M(w) = \|\langle\psi_M(w)|P_{\text{acc}}\|^2 = \|\langle\psi_{\text{start}}|\delta(w)P_{\text{acc}}\|^2$.

While the above form of the definition of a 1QFA is most convenient for our purposes, we note that a 1QFA $M = (V, \Sigma, \delta, \langle\psi_{\text{start}}|, V_{\text{acc}})$, where $k = \dim(V)$, could also be specified by a 5-tuple $(Q, \Sigma, \delta, \langle\psi_{\text{start}}|, V_{\text{acc}})$, where $Q = \{q_1, \ldots, q_k\}$ is a finite set of quantum basis states that corresponds to an orthonormal basis $\{\langle q_1|, \ldots, \langle q_k|\}$ of $V \cong \mathbb{C}^k$. For this reason, we will refer to a 1QFA as having k quantum basis states if the dimension of its underlying Hilbert space is k.

4.2 The 1QFA Groups

We wish to investigate the class of groups G whose word problem W_G can be recognized by a 1QFA with any particular number of quantum basis states and with any particular type of error. We first must establish some notation.

We say that a 1QFA M recognizes a language $L \subseteq \Sigma^*$ with *zero-error* if, $\forall w \in L$, $p_M(w) = 1$, and $\forall w \notin L$, $p_M(w) = 0$. We say that M recognizes L with *positive one-sided unbounded-error* if, $\forall w \in L$, $p_M(w) > 0$, and $\forall w \notin L$, $p_M(w) = 0$. We say that M recognizes L with positive one-sided *bounded-error* if $\exists \lambda \in \mathbb{R}_{>0}$ such that $\forall w \in L$, $p_M(w) > \lambda$, and $\forall w \notin L$, $p_M(w) = 0$. Analogously, we say that M recognizes L with *negative* one-sided unbounded-error (resp. bounded-error) if, $\forall w \in L$, $p_M(w) = 1$, and $\forall w \notin L$, $p_M(w) < 1$ (resp. $\exists \lambda \in \mathbb{R}_{>0}$ such that $\forall w \notin L$, $p_M(w) < 1 - \lambda$).

Then, for each $k \in \mathbb{N}_{\geq 1}$, we define the complexity class $\mathsf{E1QFA}(k)$ (resp. $\mathsf{N1QFA}(k)$, $\mathsf{R1QFA}(k)$, $\mathsf{coN1QFA}(k)$, $\mathsf{coR1QFA}(k)$, $\mathsf{B1QFA}(k)$) to be the set of languages recognized with zero-error (resp. positive one-sided unbounded-error, positive one-sided bounded-error, negative one-sided unbounded-error, negative one-sided bounded-error, two-sided bounded-error) by some 1QFA with at most k quantum basis states. We further define the complexity classes $\mathsf{E1QFA} = \cup_{k \in \mathbb{N}_{\geq 1}} \mathsf{E1QFA}(k)$, $\mathsf{N1QFA} = \cup_{k \in \mathbb{N}_{\geq 1}} \mathsf{N1QFA}(k)$, etc., to be those languages recognized in such a manner by a 1QFA with any finite number of quantum basis states.

A deterministic finite automaton (DFA) is a 5-tuple $D = (C, \Sigma, \gamma, c_{\text{start}}, F)$, where C is the finite set of (classical) states, Σ is the finite input alphabet, $\gamma : C \times \Sigma \to C$ is the transition function, $c_{\text{start}} \in C$ is the start state, and $F \subseteq C$ is the set of accepting states. We say that D is a *permutation*-DFA (or group finite automaton) if $\forall \sigma \in \Sigma$ and $\forall c \in C$, there is a unique $c' \in C$ such that $\gamma(c', \sigma) = c$. Let $\mathsf{DFA}(k)$ (resp. $\mathsf{permDFA}(k)$) denote the class of languages recognized by a DFA (resp. permutation-DFA) with at most k states. Let REG (resp. pREG) denote the regular languages (resp. group languages): the class of languages recognized by a DFA (resp. permutation-DFA) with any finite number of states.

Note that $\mathsf{E1QFA} = \mathsf{R1QFA} = \mathsf{coR1QFA} = \mathsf{B1QFA} = \mathsf{pREG} \subsetneq \mathsf{REG}$ [8]. However, there are languages L for which the smallest 1QFA that recognizes L with bounded-error is much smaller than the smallest DFA that recognizes L. For example, for a prime p, consider the group $\mathbb{Z}/p\mathbb{Z}$ (the integers modulo p, with the group operations being addition), and let $W_{\mathbb{Z}/p\mathbb{Z}}$ denote its word problem. It is straightforward to see that $W_{\mathbb{Z}/p\mathbb{Z}} \in \mathsf{DFA}(p)$, but that $W_{\mathbb{Z}/p\mathbb{Z}} \notin \mathsf{DFA}(k)$, for any $k < p$. However, there is a constant C such that $W_G \in \mathsf{coR1QFA}(C \log(2p))$ [2]. Therefore, with bounded-error, 1QFA can recognize only a proper subset of the regular languages, but, for certain languages, 1QFA have a "succinctness" advantage over DFA. Furthermore, note that 1QFA with unbounded-error are incomparable to DFA (i.e., there is a language L such that $L \in \mathsf{N1QFA}$ but

$L \notin \mathsf{REG}$ and a language L' such that $L' \notin \mathsf{N1QFA}$ but $L' \in \mathsf{REG}$ (see, for instance, [8]), which then implies the analogous result holds for $\mathsf{coN1QFA}$).

We show that the class of group word problems recognized by a 1QFA with k states with *positive* one-sided error (bounded-error or unbounded-error) is precisely the same as the class of group word problems recognized by a DFA with k states. However, we also show that the class of group word problems recognized with *negative* one-sided unbounded-error by a 1QFA with only 2 states vastly exceeds the class of group word problems recognized by a DFA with any number of states. This situation is precisely analogous to that of pushdown automata (PDA), as the class of group word problems recognized by deterministic PDA and by PDA with "positive" non-determinism are identical [30], and the class of group word problems recognized by PDA with "negative" non-determinism is much larger [24]. We discuss this similarity more fully in Section 6.1.

4.3 Classification of the N1QFA, R1QFA, and E1QFA Groups

In this section, we precisely classify the $\mathsf{N1QFA}(k)$, $\mathsf{R1QFA}(k)$, and $\mathsf{E1QFA}(k)$ groups, for any $k \in \mathbb{N}_{\geq 1}$.

Theorem 4.1. *For any finitely-generated group G, and any $k \in \mathbb{N}_{\geq 1}$, the following are equivalent.*

(i) $W_G \in \mathsf{N1QFA}(k)$

(ii) $W_G \in \mathsf{R1QFA}(k)$

(iii) $W_G \in \mathsf{E1QFA}(k)$

(iv) $W_G \in \mathsf{DFA}(k)$

(v) $W_G \in \mathsf{permDFA}(k)$

(vi) $|G| \leq k$

Proof. (i) $\Rightarrow$ (vi): Suppose $G = \langle S|R\rangle$, with S finite. Let $\Sigma = S \cup S^{-1}$, let $\phi : \Sigma^* \to G$ denote the natural map, and let $W_{G=\langle S|R\rangle} = \phi^{-1}(1_G)$ denote the word problem of G with respect to this presentation. If $W_G \in \mathsf{N1QFA}(k)$, then, by definition, there is a 1QFA $M = (V, \Sigma, \delta, \langle\psi_{\text{start}}|, V_{\text{acc}})$ that recognizes $W_{G=\langle S|R\rangle}$ with positive one-sided unbounded-error, where $\dim(V) = k$. Let $V_{\text{rej}} = V_{\text{acc}}^{\perp}$. Note that, for any $w \in \Sigma^*$, we have $\langle\psi_M(w)| \in V_{\text{rej}} \Leftrightarrow p_M(w) = 0 \Leftrightarrow \phi(w) \neq 1_G$. Assume, for contradiction, $|G| > k$. Let $g_1, \ldots, g_{k+1} \in G$ denote $k+1$ distinct elements of G, and, for each j, fix $x_j \in \phi^{-1}(g_j) \subseteq \Sigma^*$. Let $\widehat{\Psi} = \{\langle\psi_M(x_1)|, \ldots, \langle\psi_M(x_{k+1})|\} \subseteq V$ and notice that $|\widehat{\Psi}| = k+1 > k = \dim(V)$. This immediate implies that there must be some r such that $\langle\psi_M(x_r)| \in \text{span}\{\langle\psi_M(x_j)| : j \neq r\}$; fix such an r. Then, $\forall j \neq r$, the fact that $g_j \neq g_r$ implies $1_G \neq g_j g_r^{-1} = \phi(x_j)\phi(x_r)^{-1} = \phi(x_j x_r^{-1})$, which then implies $\langle\psi_M(x_j x_r^{-1})| \in V_{\text{rej}}$. Notice that, $\forall y, z \in \Sigma^*$, $\langle\psi_M(yz)| = \langle\psi_{\text{start}}|\delta(yz) = \langle\psi_{\text{start}}|\delta(y)\delta(z) = \langle\psi_M(y)|\delta(z)$. In particular, $\forall j \neq r$, we have $\langle\psi_M(x_j)|\delta(x_r^{-1}) = \langle\psi_M(x_j x_r^{-1})| \in V_{\text{rej}}$. As $\langle\psi_M(x_r)| \in \text{span}\{\langle\psi_M(x_j)| : j \neq r\}$, this then implies that $\langle\psi_M(x_r)|\delta(x_r^{-1}) \in V_{\text{rej}}$. However, $\langle\psi_M(x_r)|\delta(x_r^{-1}) = \langle\psi_M(x_r x_r^{-1})| \notin V_{\text{rej}}$ as $\phi(x_r x_r^{-1}) = 1_G$. This contradiction allows us to conclude that, if $W_G \in \mathsf{N1QFA}(k)$, then $|G| \leq k$.

The remaining needed implications all immediately follow from existing results or are obvious.
(vi) $\Leftrightarrow$ (v) $\Leftrightarrow$ (iv): See, for instance, [30, Lemma 1] or [6].
(v) $\Rightarrow$ (iii): $\mathsf{permDFA}(k) \subseteq \mathsf{E1QFA}(k)$ (see, for instance, [8, Theorem 3.3]).
(iii) $\Rightarrow$ (ii): Clearly, $\mathsf{E1QFA}(k) \subseteq \mathsf{R1QFA}(k)$.
(ii) $\Rightarrow$ (i): Clearly, $\mathsf{R1QFA}(k) \subseteq \mathsf{N1QFA}(k)$. □

For any $k \in \mathbb{N}_{\geq 1}$, there is a group of size k (e.g., $\mathbb{Z}/k\mathbb{Z}$). We immediately obtain the following corollaries.

Corollary 4.1.1. *For any $k \in \mathbb{N}_{\geq 1}$, the following statements hold.*

(i) $\mathsf{N1QFA}(k) \subsetneq \mathsf{N1QFA}(k+1)$.

(ii) $\mathsf{R1QFA}(k) \subsetneq \mathsf{R1QFA}(k+1)$.

(iii) $\mathsf{E1QFA}(k) \subsetneq \mathsf{E1QFA}(k+1)$.

(iv) There is a language $L \in \mathsf{permDFA}(k+1)$ such that $L \notin \mathsf{N1QFA}(k)$.

Corollary 4.1.2. *For a finitely-generated group G, $W_G \in \mathsf{N1QFA} \Leftrightarrow W_G \in \mathsf{E1QFA} \Leftrightarrow W_G \in \mathsf{R1QFA} \Leftrightarrow W_G \in \mathsf{REG} \Leftrightarrow G$ is finite.*

4.4 The coN1QFA Groups

4.4.1 Normal Form 1QFA

We next consider the classification of the $\mathsf{coN1QFA}(k)$ groups. We begin by defining the notion of a "normal form" for a 1QFA that recognizes some word problem W_G with negative one-sided unbounded-error, which will be more convenient to analyze. We will then show that any W_G recognized by a 1QFA in this fashion is always recognized by a normal form 1QFA.

Definition 4.2. Consider a group $G = \langle S|R\rangle$, with S finite. Suppose that the 1QFA $M = (V, \Sigma, \delta, \langle\psi_{\text{start}}|, V_{\text{acc}})$ recognizes $W_G := W_{G=\langle S|R\rangle}$ with negative one-sided unbounded-error. We say that M is of *normal form* if the following conditions hold.

(i) $\delta(w^{-1}) = \delta(w)^{-1}$, $\forall w \in \Sigma^*$.

(ii) $V_{\text{acc}} = \text{span}\{\langle\psi_M(w)| : w \in W_G\}$

(iii) $V = \text{span}\{\langle\psi_M(w)| : w \in \Sigma^*\}$.

Lemma 4.3. *Consider a group $G = \langle S|R\rangle$, with S finite, and define $W_G := W_{G=\langle S|R\rangle}$. Suppose that the 1QFA $M = (V, \Sigma, \delta, \langle\psi_{start}|, V_{acc})$ recognizes W_G with negative one-sided unbounded-error. Let $\widehat{V}_{acc} = \text{span}\{\langle\psi_M(w)| : w \in W_G\} \subseteq V_{acc}$ and let $B \leq \text{U}(V)$ denote the subgroup of the unitary group on V consisting of those elements under which $\widehat{V}_{acc}$ is stable (i.e., $B = \{b \in \text{U}(V) : v \in \widehat{V}_{acc} \Rightarrow vb \in \widehat{V}_{acc}\}$). Then the following statements hold.*

(i) $w \in W_G \Leftrightarrow \delta(w) \in B$, $\forall w \in \Sigma^*$.

(ii) $\delta(w)^{-1}\delta(x)\delta(w) \in B$, $\forall w \in \Sigma^*, \forall x \in W_G$.

Proof. (i) First, suppose $\delta(w) \in B$. Note that the empty string $\epsilon \in W_G$, which implies $\langle\psi_{\text{start}}| = \langle\psi_{\text{start}}|\delta(\epsilon) \in \widehat{V}_{\text{acc}}$. Then $\langle\psi_M(w)| = \langle\psi_{\text{start}}|\delta(w) \in \widehat{V}_{\text{acc}} \subseteq V_{\text{acc}}$, which implies $p_M(w) = 1$, and so $w \in W_G$.

Next, suppose, $w \in W_G$. Then $\exists x_1, \ldots, x_s \in W_G$ such that $\{\langle\psi_M(x_1)|, \ldots, \langle\psi_M(x_s)|\}$ is a basis of $\widehat{V}_{\text{acc}}$. For each j, we have $x_j w \in W_G$, which implies $\langle\psi_M(x_j)|\delta(w) = \langle\psi_M(x_j w)| \in \widehat{V}_{\text{acc}}$. This immediately implies that, $\forall v \in \widehat{V}_{\text{acc}}$, we have $v\delta(w) \in \widehat{V}_{\text{acc}}$.

(ii) Begin by noting that $\delta(w^{-1}w) = \delta(w^{-1})\delta(w)$, which implies $\delta(w)^{-1} = \delta(w^{-1}w)^{-1}\delta(w^{-1})$. Clearly, $w^{-1}w \in W_G$ and $w^{-1}xw \in W_G$, and so the first part of this lemma implies $\delta(w^{-1}w) \in B$ and $\delta(w^{-1}xw) \in B$. Therefore,

$$\delta(w)^{-1}\delta(x)\delta(w) = \delta(w^{-1}w)^{-1}\delta(w^{-1})\delta(x)\delta(w) = \delta(w^{-1}w)^{-1}\delta(w^{-1}xw) \in B.$$

□

Lemma 4.4. *Consider a group $G = \langle S|R\rangle$, with S finite, and define $W_G := W_{G=\langle S|R\rangle}$. If $W_G \in$ coN1QFA(k), then there is a normal form 1QFA with at most k quantum basis states that recognizes W_G with negative one-sided unbounded-error.*

Proof. By definition, there is a 1QFA $M = (V, \Sigma, \delta, \langle\psi_{\text{start}}|, V_{\text{acc}})$ that recognizes W_G with negative one-sided unbounded-error, where $\dim(V) \leq k$. We define the 1QFA $\widehat{M} = (V, \Sigma, \widehat{\delta}, \langle\psi_{\text{start}}|, V_{\text{acc}})$ to be identical to M with the exception of its transition function $\widehat{\delta} : \Sigma \to \mathrm{U}(V)$, which we now specify. For each $\sigma \in \Sigma$, if $\sigma \in S$, then $\widehat{\delta}(\sigma) = \delta(\sigma)$; if, instead, $\sigma \in S^{-1}$, then $\widehat{\delta}(\sigma) = \delta(\sigma^{-1})^{-1}$. Then $\widehat{\delta}(\sigma^{-1}) = \widehat{\delta}(\sigma)^{-1}$, $\forall\sigma \in \Sigma$, which immediately implies that $\widehat{\delta}(w^{-1}) = \widehat{\delta}(w)^{-1}$, $\forall w \in \Sigma^*$. Therefore, $\widehat{M}$ satisfies Definition 4.2(i). Fix $w = w_1 \cdots w_n \in \Sigma^*$, where each $w_j \in \Sigma$. Define $\widehat{V}_{\text{acc}} = \text{span}\{\langle\psi_M(w)| : w \in W_G\} \subseteq V_{\text{acc}}$ and $B = \{b \in \mathrm{U}(V) : v \in \widehat{V}_{\text{acc}} \Rightarrow vb \in \widehat{V}_{\text{acc}}\}$ as in Lemma 4.3.

We next show that $\exists b_0, \ldots, b_n \in B$ such that, $\forall j \in \{0, \ldots, n\}$, $\widehat{\delta}(w_{n-j+1} \cdots w_n) = \delta(w_{n-j+1} \cdots w_n)b_j$, by induction on j. For $j = 0$, writing ϵ for the empty-string, we have $\delta(\epsilon) = 1_{\mathrm{U}(V)} = \widehat{\delta}(\epsilon)$, and so the claim holds with $b_0 = 1_{\mathrm{U}(V)}$. For $j \geq 1$, let $\sigma = w_{n-j+1} \in \Sigma$, $y = w_{n-j+2} \cdots w_n \in \Sigma^*$ and $z = w_{n-j+1} \cdots w_n = \sigma y \in \Sigma^*$. We have

$$\widehat{\delta}(z) = \widehat{\delta}(\sigma)\widehat{\delta}(y) = \widehat{\delta}(\sigma)\delta(y)b_{j-1}.$$

If $\sigma \in S$, then $\widehat{\delta}(\sigma) = \delta(\sigma)$, and the claim holds with $b_j = b_{j-1}$. Suppose instead $\sigma \in S^{-1}$. Then $\widehat{\delta}(\sigma) = \delta(\sigma^{-1})^{-1}$, which implies

$$\widehat{\delta}(z) = \delta(\sigma^{-1})^{-1}\delta(y)b_{j-1} = \delta(z)\delta(z)^{-1}\delta(\sigma^{-1})^{-1}\delta(y)b_{j-1} = \delta(z)\left(\delta(y)^{-1}\delta(\sigma^{-1})\delta(z)\right)^{-1}b_{j-1}.$$

By Lemma 4.3(ii), $\delta(y)^{-1}\delta(\sigma^{-1}\sigma)\delta(y) \in B$. This implies

$$\left(\delta(y)^{-1}\delta(\sigma^{-1})\delta(z)\right)^{-1}b_{j-1} = \left(\delta(y)^{-1}\delta(\sigma^{-1})\delta(\sigma y)\right)^{-1}b_{j-1} = \left(\delta(y)^{-1}\delta(\sigma^{-1}\sigma)\delta(y)\right)^{-1}b_{j-1} \in B.$$

Therefore, $\widehat{\delta}(z) = \delta(z)b_j$, where $b_j = \left(\delta(y)^{-1}\delta(\sigma^{-1})\delta(z)\right)^{-1}b_{j-1} \in B$, as desired.

Therefore, by the above, $\delta(w)^{-1}\widehat{\delta}(w) \in B$, $\forall w \in \Sigma^*$. Note that $\|vbP_{\text{acc}}\| = \|vP_{\text{acc}}\|$, $\forall v \in V, \forall b \in B$. Therefore, for every $w \in \Sigma^*$, we have

$$p_{\widehat{M}}(w) = \|\langle\psi_{\text{start}}|\widehat{\delta}(w)P_{\text{acc}}\|^2 = \|\langle\psi_{\text{start}}|\delta(w)\delta(w)^{-1}\widehat{\delta}(w)P_{\text{acc}}\|^2 = \|\langle\psi_{\text{start}}|\delta(w)P_{\text{acc}}\|^2 = p_M(w).$$

Therefore, $\widehat{M}$ recognizes W_G with negative one-sided unbounded-error. To complete the proof, we define the 1QFA $M' = (V', \Sigma, \widehat{\delta}, \langle\psi_{\text{start}}|, V'_{\text{acc}})$ to have Hilbert space $V' = \text{span}\{\langle\psi_{\widehat{M}}(w)| : w \in \Sigma^*\}$, accepting subspace $V'_{\text{acc}} = \text{span}\{\langle\psi_{\widehat{M}}(w)| : w \in W_G\}$, and to otherwise be identical to $\widehat{M}$. Notice that $\langle\psi_{M'}(w)| = \langle\psi_{\text{start}}|\widehat{\delta}(w) = \langle\psi_{\widehat{M}}(w)|$, $\forall w \in \Sigma^*$, and so M' satisfies satisfies Definition 4.2(ii) and Definition 4.2(iii). As $\widehat{\delta}$ is unchanged, M' also satisfies Definition 4.2(i). We have $V'_{\text{acc}} \subseteq V_{\text{acc}}$, which implies that, for any $w \notin W_G$, we have $p_{M'}(w) \leq p_{\widehat{M}}(w) < 1$. Clearly, for any $w \in W_G$, we have $p_{M'}(w) = 1$. Therefore, M' recognizes W_G with negative one-sided unbounded-error, and M' is of normal form. □

Consider a group $G = \langle S|R\rangle$, with S finite, and suppose that the normal form 1QFA $M = (V, \Sigma, \delta, \langle\psi_{\text{start}}|, V_{\text{acc}})$ recognizes $W_G := W_{G=\langle S|R\rangle}$ with negative one-sided unbounded-error. Let $T = \{\delta(w) : w \in \Sigma^*\}$ and $N = \{\delta(w) : w \in W_G\}$. For groups H, K we write $H \leq K$ if H is a subgroup of K and $H \triangleleft K$ if H is a normal subgroup of K. Notice that $N \triangleleft T \leq \mathrm{U}(V)$. Let $B = \{b \in \mathrm{U}(V) : v \in \widehat{V}_{\text{acc}} \Rightarrow vb \in V_{\text{acc}}\}$. By Lemma 4.3(i), we have $N \leq B$ and, for any $x, y \in \Sigma^*$, we have $xy^{-1} \in W_G \Leftrightarrow \delta(xy^{-1}) \in N$. Due to the fact that $xy^{-1} \in W_G \Leftrightarrow \phi(x) = \phi(y)$ and $\delta(x)\delta(y)^{-1} = \delta(xy^{-1})$, we conclude that $\delta(x)\delta(y)^{-1} \in N \Leftrightarrow \phi(x) = \phi(y)$. Let $\eta : T \to G$ denote the group homomorphism defined such that, for any $t \in T$, we have $\eta(t) = \phi(w)$, where $w \in \delta^{-1}(t)$. Note that η is well-defined and has $\ker \eta = N$, by the above observation. Therefore, $G \cong T/N$. In other words, T is an extension of G by N; i.e., we have the short exact sequence $1 \to N \to T \xrightarrow{\eta} G \to 1$.

4.4.2 Single-Accept 1QFA

We now consider finite automata that may only have a single accepting state. We define a single-accept-1QFA (resp. single-accept-DFA, single-accept-NFA, etc.) to be a 1QFA (resp. DFA, NFA, etc.) that has only a single accepting state; in particular, a single-accept-1QFA is a 1QFA where $\dim(V_{\text{acc}}) = 1$. As before, REG denotes the regular languages (i.e., those languages recognizable by a DFA, or equivalently by an NFA). Note that, for any word problem W_G, if W_G is recognizable by a DFA (resp. NFA, PDA), then W_G is recognizable by a single-accept-DFA (resp. single-accept-NFA, resp. single-accept-PDA). For NFA, this follows immediately from the fact that, $\forall L \in \mathsf{REG}$, L is recognizable by a single-accept-NFA. However, $\exists L \in \mathsf{REG}$ such that L is not recognizable by a single-accept-DFA. In other words, the restriction of having a single accepting state reduces the power of DFA, but it does not shrink the class of group word problems that may be recognized by a DFA, nor does it reduce the power of NFA at all. For 1QFA, it is not immediately clear if single-accept-1QFA are as powerful as (general) 1QFA. By Theorem 4.1, if there is a 1QFA with k basis states that recognizes some W_G, with either zero-error, positive one-sided bounded-error, or positive one-sided unbounded-error, then there is a single-accept-1QFA of the same size that recognizes W_G with the same error type; that is to say, the restriction of having a single accepting state does not affect the class of groups recognizable with any of these error types by a 1QFA of any particular size. We define $\mathsf{single{-}accept{-}coN1QFA}(k)$ to be the class of languages recognizable with negative one-sided unbounded-error by a 1QFA with k basis states, and we define $\mathsf{single{-}accept{-}coN1QFA} = \cup_{k\in\mathbb{N}_{\geq 1}} \mathsf{single{-}accept{-}coN1QFA}(k)$ to be the class of languages recognizable with negative one-sided unbounded-error by a 1QFA of any finite size. We will show that, if k is sufficiently small, then $W_G \in \mathsf{coN1QFA}(k) \Rightarrow W_G \in \mathsf{single{-}accept{-}coN1QFA}(k)$; we are not presently able to show the claim for all k. For any k, if $W_G \in \mathsf{single{-}accept{-}coN1QFA}(k)$, then, by Lemma 4.4, W_G is recognizable (with negative one-sided unbounded-error) by a normal form single-accept-1QFA with k basis states. This will allow us to strongly constrain those group word problems in such a class.

Lemma 4.5. *Consider a group $G = \langle S|R\rangle$, with S finite. Suppose that the normal form 1QFA $M = (V, \Sigma, \delta, \langle\psi_{start}|, V_{acc})$ recognizes $W_G := W_{G=\langle S|R\rangle}$ with negative one-sided unbounded-error and $\dim(V_{acc}) = 1$. Let $N = \{\delta(w) : w \in W_G\}$. The following statements hold.*

(i) $\forall n \in N, \forall w \in \Sigma^*$, n has (left) eigenvector $\langle\psi_M(w)|$.

(ii) N is an abelian group.

Proof. (i) Note that $\dim(V_{\text{acc}}) = 1$ implies $P_{\text{acc}} = |\psi_{\text{start}}\rangle\langle\psi_{\text{start}}|$ (recall that the empty string

$\epsilon \in W_G$). Let $t = \delta(w)$ and fix $x \in \delta^{-1}(n) \subseteq W_G$. Then $wxw^{-1} \in W_G$ immediately implies

$$1 = p_M(wxw^{-1}) = \|\langle\psi_{\text{start}}|\delta(wxw^{-1})P_{\text{acc}}\|^2 = |\langle\psi_{\text{start}}|tnt^{-1}|\psi_{\text{start}}\rangle|^2.$$

Due to the fact that $\|\langle\psi_{\text{start}}|\| = 1$ and $t, n \in \mathrm{U}(V)$, we conclude that $\langle\psi_{\text{start}}|tn = \lambda\langle\psi_{\text{start}}|t$, for some $\lambda \in \mathbb{C}$, with $|\lambda| = 1$. Therefore, n has (left) eigenvector $\langle\psi_{\text{start}}|t = \langle\psi_{\text{start}}|\delta(w) = \langle\psi_M(w)|$, $\forall n \in N, \forall w \in \Sigma^*$.

(ii) By Definition 4.2(iii), we have $V = \text{span}\{\langle\psi_M(w)| : w \in \Sigma^*\}$. By the first part of this lemma, this implies that there is a common eigenbasis shared by all elements of N, and so all elements of N are simultaneously diagonalizable; therefore, $N \leq \mathrm{U}(V)$ is abelian.

□

Lemma 4.6. *Consider a group $G = \langle S|R\rangle$, with S finite. Suppose $W_G \in \mathsf{coN1QFA}(k)$, where $k \leq 3$. Then $W_G \in \mathsf{single{-}accept{-}coN1QFA}(k)$.*

Proof. If $G = \{1\}$ (the trivial group), then $W_G = \Sigma^* \in \mathsf{single{-}accept{-}coN1QFA}(1)$; therefore, we assume for the remainder of the proof that G is not the trivial group. If $W_G \in \mathsf{coN1QFA}(k)$, then, by definition, there is a 1QFA $M = (V, \Sigma, \delta, \langle\psi_{\text{start}}|, V_{\text{acc}})$ that recognizes $W_{G=\langle S|R\rangle}$ with negative one-sided unbounded-error where $\dim(V) \leq k$. We cannot have $\dim(V_{\text{acc}}) = 0$, as that would imply that the language of M is the empty language, nor can we have $\dim(V_{\text{acc}}) = k$, as that would imply $V_{\text{acc}} = V$, which in turn would imply that the language of M is Σ^*.

If $k \leq 2$, then the claim has been proven. If $k = 3$, then by the above, we must have $\dim(V_{\text{acc}}) \in \{1, 2\}$. If $\dim(V_{\text{acc}}) = 1$, then M itself is the desired single-accept-1QFA. Suppose instead that $\dim(V_{\text{acc}}) = 2$. By Lemma 4.3(i), V_{acc} is stable under the action of N, which then implies $V_{\text{rej}} = V_{\text{acc}}^{\perp}$ is also stable under the action of $N \leq \mathrm{U}(V)$. For any $w \notin W_G$, and any $v \in V_{\text{rej}} \backslash \{0\}$, if $v\delta(w) \in V_{\text{rej}}$, then $V_{\text{rej}}\delta(w) = V_{\text{rej}}$ (as $\dim(V_{\text{rej}}) = 1$), which implies $V_{\text{acc}}\delta(w) = V_{\text{acc}}$, which is impossible by Lemma 4.3(i); therefore, we must have $v\delta(w) \notin V_{\text{rej}}$, $\forall w \notin W_G$, $\forall v \in V_{\text{rej}} \setminus \{0\}$. Thus, the 1QFA $M' = (V, \Sigma, \delta, \langle\psi'_{\text{start}}|, V'_{\text{acc}})$ with $\langle\psi_{\text{start}}| \in V_{\text{rej}}$ (of unit norm) and $V'_{\text{acc}} = V_{rej}$ also recognizes $W_{G=\langle S|R\rangle}$ with negative one-sided unbounded-error, but has $\dim(V_{\text{acc}}) = 1$. □

For groups H, N, Q, we say that H is *an extension of Q by N* if we have a short exact sequence $1 \to N \to H \to Q \to 1$ (i.e., $N \triangleleft H$ and $Q \cong H/N$). If N has some property $\mathcal{P}$ (e.g., finite, abelian, nilpotent, etc.) and Q has some property $\mathcal{R}$, then we say that H is $\mathcal{P}$-by-$\mathcal{R}$ (note the order here). We say that a group K is *virtually* $\mathcal{P}$ if K has a finite-index subgroup that has property $\mathcal{P}$. We say that a property $\mathcal{P}$ is *subgroup-closed* if, for any group K that has property $\mathcal{P}$, any subgroup $S \leq K$ also has property $\mathcal{P}$. For any subgroup-closed $\mathcal{P}$, a group K is virtually $\mathcal{P}$ precisely when it is $\mathcal{P}$-by-finite. Note that the properties of being abelian, nilpotent, or solvable are all subgroup-closed. Let $\mathcal{G}_{\mathbf{vAb}}$ (resp. $\mathcal{G}_{\mathbf{vNilp}}$, $\mathcal{G}_{\mathbf{vSolv}}$) denote the collection of all finitely-generated virtually abelian (resp. nilpotent, solvable) groups, where $\mathcal{G}_{\mathbf{vAb}} \subseteq \mathcal{G}_{\mathbf{vNilp}} \subseteq \mathcal{G}_{\mathbf{vSolv}}$. Before considering the general case (in the following section), we first classify those groups $G \in \mathcal{G}_{\mathbf{vSolv}}$ for which $W_G \in \mathsf{single{-}accept{-}coN1QFA}$.

Theorem 4.7. *For any $G \in \mathcal{G}_{\boldsymbol{vSolv}}$, we have $W_G \in \mathsf{single{-}accept{-}coN1QFA} \Leftrightarrow G \in \mathcal{G}_{\boldsymbol{vAb}}$.*

Proof. Firstly, $W_G \in \mathsf{single{-}accept{-}coN1QFA}$, $\forall G \in \mathcal{G}_{\mathbf{vAb}}$ [35, Theorem 1.7]. In the other direction, suppose $G \in \mathcal{G}_{\mathbf{vSolv}}$ and $W_G \in \mathsf{single{-}accept{-}coN1QFA}$. Let $G = \langle S|R\rangle$, with S finite. Then by Lemma 4.4, there is a normal form single-accept-1QFA $M = (V, \Sigma, \delta, \langle\psi_{\text{start}}|, V_{\text{acc}})$ that recognizes $W_{G=\langle S|R\rangle}$ with negative one-sided unbounded-error, where $\dim(V) \leq k$, for some finite k (it is easy to see that the 1QFA produced by the proof of this lemma preserves the property that

$\dim(V_{\text{acc}}) = 1$). Let $N = \{\delta(w) : w \in W_G\}$ and $T = \{\delta(w) : w \in \Sigma^*\}$. As observed at the end of Section 4.4.1, T is an extension of G by N. By Lemma 4.5(ii), N is abelian. Therefore, T is abelian-by-(solvable-by-finite), which implies that T is (abelian-by-solvable)-by-finite, which then implies that T is solvable-by-finite. As $T \leq \mathrm{U}(V)$, where $\dim(V) \leq k$ for k finite, we then conclude that T is abelian-by-finite (see, for instance, [39, Proposition 2.2]). As $G \cong T/N$, this then implies that G is abelian-by-finite; therefore, $G \in \mathcal{G}_{\mathbf{vAb}}$. □

4.4.3 Representations of coN1QFA Groups

We will show that the existence of a 1QFA, of a particular type, that recognizes W_G implies the existence of a faithful finite-dimensional projective unitary representation of G.

We briefly recall the needed notation and terminology from representation theory; we refer the reader to [25] or [42] for more thorough background, as well as our recent paper [35] for the particular application of representation theory to the recognizability of group word problem by QFA. A *finite-dimensional unitary* (resp. *projective unitary*) *representation* of a group H is a pair (ρ, V) where V is a finite-dimensional complex Hilbert space and $\rho : H \to \mathrm{U}(V)$ (resp. $\rho : H \to \mathrm{PU}(V)$) is a group homomorphism (as we restrict our attention to groups that are countable, our definition agrees with the standard notion of a unitary represntation, which requires that ρ is a strongly continuous homomorphism of topological groups, as this property is trivially satisfied when H is given the (natural) discrete topology). Throughout this section, we use the term *representation* to refer exclusively to representations of these types. By standard slight abuse of notation, we refer to ρ as being a representation of H, when V is clear from context, and we also refer to V as being a representation of H, when ρ is clear from context. We say a representation ρ is *faithful* if $\ker \rho$ is trivial.

A representation $\rho : H \to \mathrm{U}(V)$ gives a (right) action of H on V where, for $h \in H$ and $v \in V$, $v \cdot h = v\rho(h)$. If a subspace $V' \subseteq V$ is stable under this action (i.e., $\forall h \in H, \forall v \in V'$ we have $v\rho(h) \in V'$), then the restriction of $\rho(h)$ to V' yields a representation $\rho' : H \to \mathrm{U}(V')$ of H; we say (ρ', V') is a *subrepresentation* of (ρ, V). We say that the representation (ρ, V) is *irreducible* if $\dim(V) \neq 0$ and (ρ, V) has no non-trivial subrepresentations (i.e., the only stable subspaces of V under the action of H are V and $\{0\}$). Suppose $\pi : H \to \mathrm{U}(\widehat{V})$ is also a representation of H. A *homomorphism of representations* from ρ to π is a $\mathbb{C}$-linear map $\Phi : V \to \widehat{V}$ such that $\Phi(v\rho(h)) = \Phi(v)\pi(h)$, $\forall h \in H, \forall v \in V$; if, moreover, Φ is a bijection, we say that Φ is an *isomorphism of representations*, and that (ρ, V) and $(\pi, \widehat{V})$ are *isomorphic*, which we denote by writing $\rho \cong \pi$.

For representations $\rho_1 : H \to \mathrm{U}(V_1)$ and $\rho_2 : H \to \mathrm{U}(V_2)$ of H, their *direct sum* is the representation $\rho_1 \oplus \rho_2 : H \to \mathrm{U}(V_1 \oplus V_2)$ where $(v_1 + v_2)((\rho_1 \oplus \rho_2)(h)) = v_1\rho_1(h) + v_2\rho_2(h)$, $\forall h \in H$, $\forall v_1 \in V_1$, $\forall v_2 \in V_2$. Any representation $\rho : H \to \mathrm{U}(V)$ is *semisimple*: there is a set $\{\rho_i : H \to \mathrm{U}(V_i) : i \in I\}$ of irreducible subrepresentations of ρ such that $\rho \cong \oplus_i \rho_i$. Note that such a decomposition is not unique, in a manner which we now clarify. Suppose $\pi : H \to \mathrm{U}(V')$ is some irreducible representation of H. Then for any decomposition $\rho \cong \oplus_i \rho_i$, where $\{\rho_i : H \to \mathrm{U}(V_i) : i \in I\}$ are irreducible subrepresentations of ρ, the subspace $\bigoplus_{\rho_i \cong \pi} V_i \subseteq V$ is the same (see, for instance, [25, Proposition 2.7.7]); let $M_V(\pi) = \bigoplus_{\rho_i \cong \pi} V_i$ denote the *π-isotypic component* of V. Then V admits a decomposition into its isotypic components, $V = \oplus_j M_V(\pi_j)$, for some set $\{\pi_j : H \to \mathrm{U}(V_j') : j \in J\}$ of irreducible representations of H, which is unique (up to reordering of the isotypic components).

We now show that the existence of a 1QFA, of the appropriate type, that recognizes W_G, implies the existence of a faithful finite-dimensional projective unitary representation of G.

Lemma 4.8. *Consider a group $G = \langle S|R\rangle$, with S finite. Suppose that the normal form 1QFA $M = (V, \Sigma, \delta, \langle\psi_{start}|, V_{acc})$ recognizes $W_G := W_{G=\langle S|R\rangle}$ with negative one-sided unbounded-error. Let $N = \{\delta(w) : w \in W_G\}$. Then (id, V), where $\mathrm{id} : \mathrm{U}(V) \to \mathrm{U}(V)$ is the identity map, is (trivially) a faithful finite-dimensional unitary representation of N. Let $V = \oplus_{i=1}^{r} V_i$ denote the decomposition of V (as a representation of N) into its isotypic components. Suppose that N is an abelian group and that $\langle\psi_{start}| \in V_1$. The following statements hold.*

(i) $\dim(V_i) = \dim(V_j)$, $\forall i, j$.

(ii) $\exists H \leq G$, *with* $[G : H] = r$, *such that* H *has a faithful representation* $\rho : H \to \mathrm{PU}(V_1)$.

(iii) G *has a faithful representation* $\pi : G \to \mathrm{PU}(V)$.

Proof. (i) Let $T = \{\delta(w) : w \in \Sigma^*\}$. As $N \triangleleft T$, it is easy to see that, $\forall i \in \{1, \ldots, r\}$, $\forall t \in T$, we have $V_i t \subseteq V_{j_{i,t}}$, for some $j_{i,t} \in \{1, \ldots, r\}$ (see, for instance, [25, Proposition 2.7.7(3)]). By Definition 4.2(iii), $\langle\psi_{\mathrm{start}}|$ is a cyclic vector for T, which implies $\exists t_2, \ldots, t_r \in T$ such that $V_1 t_j \subseteq V_j$, $\forall j \in \{2, \ldots, r\}$. As $t_j t_j^{-1} \in N$, we must also have $V_j t_j^{-1} \subseteq V_1$, which implies $V_1 t_j = V_j$ and $V_j t_j^{-1} = V_1$, $\forall j \in \{2, \ldots, r\}$. Therefore, $\dim(V_i) = \dim(V_j)$, $\forall i, j$.

(ii) By the above, the group T has a (right) action on $\{V_i : i \in \{1, \ldots, r\}\}$ given by $V_i t = V_{j_{i,t}}$, $\forall t \in T$, $\forall i \in \{1, \ldots, r\}$. Let $T_{V_1} = \{t \in T : V_1 t = V_1\}$ denote the stabilizer subgroup of T with respect to V_1. Let $t_1 = 1_T \in T$, and let $t_2, \ldots, t_r \in T$ be as defined above. For any $t \in T$, we have $V_1 t t_{j_{1,t}}^{-1} = V_{j_{1,t}} t_{j_{1,t}}^{-1} = V_1$, which implies $t t_{j_{1,t}}^{-1} \in T_{V_1}$. Therefore, $T = \sqcup_j T_{V_1} t_j$; that is to say, $\{t_j : j \in \{1, \ldots, r\}\}$ is a complete family of right coset representatives of T_{V_1} in T, which implies $[T : T_{V_1}] = r$. As $N \triangleleft T$ and $N \leq T_{V_1} \leq T$, we have $N \triangleleft T_{V_1}$. We then have $[T/N : T_{V_1}/N] = [T : T_{V_1}] = r$, and as $G \cong T/N$, we conclude that $\exists H \leq G$, with $[G : H] = r$, such that $H \cong T_{V_1}/N$.

By definition, V_1 is stable under the action of $T_{V_1} \leq \mathrm{U}(V)$, and, as N is abelian, N acts as a scalar on V_1; this immediately implies the existence of a faithful representation $\zeta : T_{V_1}/N \to \mathrm{PU}(V_1)$ of T_{V_1}/N. This, in turn, implies the existence of a faithful representation $\rho : H \to \mathrm{PU}(V_1)$, due to the fact that $H \cong T_{V_1}/N$.

(iii) The representation ρ of H induces the representation $\mathrm{Ind}_H^G(\rho)$ of its finite-index overgroup G, where $\mathrm{Ind}_H^G(\rho) : G \to \mathrm{PU}(V_1 \otimes \mathbb{C}^r)$ is easily seen to be faithful (see [28] for a precise definition of the induced representation). By the first part of this lemma, $r \dim(V_1) = \dim(V)$, which implies $V_1 \otimes \mathbb{C}^r \cong V$. Therefore, G has a faithful representation $\pi : G \to \mathrm{PU}(V)$.

□

Theorem 4.9. *Consider a group $G = \langle S|R\rangle$, with S finite. The following statements hold.*

(i) For any $k \in \mathbb{N}_{\geq 1}$, suppose G has a faithful representation $\pi : G \to \mathrm{PU}(\mathbb{C}^k)$. If G is abelian or $k \leq 2$, then $W_G \in \mathsf{coN1QFA}(k)$; in general, $W_G \in \mathsf{coN1QFA}(k^2)$.

(ii) If $W_G \in \mathsf{single{-}accept{-}coN1QFA}(k)$, then G has a faithful representation $\pi : G \to \mathrm{PU}(\mathbb{C}^k)$.

(iii) $W_G \in \mathsf{coN1QFA}(1)$ if and only if G is the trivial group.

(iv) $W_G \in \mathsf{coN1QFA}(2)$ if and only if G has a faithful representation $\pi : G \to \mathrm{PU}(\mathbb{C}^2)$.

(v) If $W_G \in \mathsf{coN1QFA}(3)$, then G has a faithful representation $\pi : G \to \mathrm{PU}(\mathbb{C}^3)$.

Proof. (i) [35, Theorem 1.7].

(ii) Suppose $W_G \in \mathsf{single{-}accept{-}coN1QFA}(k)$. By Lemma 4.4, there is a normal form single-accept-1QFA $M = (V, \Sigma, \delta, \langle\psi_{\text{start}}|, V_{\text{acc}})$ that recognizes $W_{G=\langle S|R\rangle}$ with negative one-sided unbounded-error where $\dim(V) \leq k$ (it is easy to see that the 1QFA produced by the proof of this lemma preserves the property that $\dim(V_{\text{acc}}) = 1$). Let $N = \{\delta(w) : w \in W_G\}$. By Lemma 4.5, N is abelian and $\langle\psi_{\text{start}}|$ is an eigenvector of every $n \in N$. Therefore, $\langle\psi_{\text{start}}| \in M_V(\theta)$, where $\theta : N \to \mathrm{U}(V_{\text{acc}})$ is the (necessarily irreducible, as $\dim(V_{\text{acc}}) = 1$) representation of N given by restricting the operators $n \in N \leq \mathrm{U}(V)$ to the (stable) subspace $V_{\text{acc}} \subseteq V$. In particular, M meets the hypothesis of Lemma 4.8. By Lemma 4.8(iii), G has a faithful representation $\pi : G \to \mathrm{PU}(\mathbb{C}^k)$.

(iii) Immediate from definitions.

(iv) If $W_G \in \mathsf{coN1QFA}(2)$, then Lemma 4.6 implies $W_G \in \mathsf{single{-}accept{-}coN1QFA}(2)$. By part (ii) of this theorem, G has a faithful representation $\pi : G \to \mathrm{PU}(\mathbb{C}^k)$. In the other direction, if G has a faithful representation $\pi : G \to \mathrm{PU}(\mathbb{C}^k)$, then by part (i) of this theorem, $W_G \in \mathsf{coN1QFA}(2)$.

(v) Analogous to the proof of part (iv).

□

Corollary 4.9.1. *We have* $\mathsf{coN1QFA}(1) \subsetneq \mathsf{coN1QFA}(2) \subsetneq \mathsf{coN1QFA}(3) \subsetneq \mathsf{coN1QFA}(4)$.

5 One-way Measure-once QCFA

5.1 Definition of the 1QCFA Model

We next consider measure-once one-way quantum finite automata with quantum and classical states (1QCFA). Informally, a 1QCFA can be thought of as a Moore and Crutchfield [29] type measure-once 1QFA that has been augmented with a DFA-based control, or, equivalently, as a DFA that has been augmented with a quantum register of finite size that is measured once after reading the entire input (cf. the measure-many one-way QFA with quantum and classical states of Zheng et al. [45]). Formally, a 1QCFA is a 9-tuple $N = (V, C, \Sigma, \delta, \gamma, \langle\psi_{\text{start}}|, c_{\text{start}}, V_0, F_{\text{acc}})$, where V is a finite-dimensional complex Hilbert space, C is a finite set of classical states, Σ is a finite alphabet, $\delta : C \times \Sigma \to \mathrm{U}(V)$ is the quantum part of the transition function, $\gamma : C \times \Sigma \to C$ is the classic part of the transition function, $\langle\psi_{\text{start}}| \in V$, with $\|\langle\psi_{\text{start}}|\| = 1$, is the initial superposition of the quantum register, $c_{\text{start}} \in C$ is the initial classical state, $V_0 \subseteq V$ specifies the single quantum measurement to be performed, and $F_{\text{acc}} \subseteq C \times \{0, 1\}$ specifies the conditions under which the machine accepts or rejects its input. Let $V_1 = V_0^{\perp} \subseteq V$, and let P_0 (resp. P_1) denote the orthogonal projection operator onto V_0 (resp. V_1). Note that we follow the convention that the quantum register of a 1QCFA is described by some $\langle\psi|$ (and so operators are applied on the right), as was the case for 1QFA.

On an input string $w \in \Sigma^*$, the 1QCFA N operates as follows. Initially, N is in the classic state c_{start} and its quantum register is in the configuration $\langle\psi_{\text{start}}|$. Then, N reads the string w from left to right, one symbol at a time; when reading the symbol $\sigma \in \Sigma$, if the classic state is currently $c \in C$, then N performs the transformation $\delta(c, \sigma)$ to its quantum register and changes its classic state to $\gamma(c, \sigma)$. After reading the entire string w, N performs the quantum measurement specified by $\{P_0, P_1\}$ producing the result $b \in \{0, 1\}$. Then N accepts w precisely when $(c, b) \in F_{\text{acc}}$, where c is the classic state after reading the entire string.

We next extend the functions γ and δ to $\gamma : C \times \Sigma^* \to C$ and $\delta : C \times \Sigma^* \to \mathrm{U}(V)$ in the natural way. Namely, for $c \in C$ and $x \in \Sigma^*$, suppose that the classic state of N is initially c and that N reads the input substring $x \in \Sigma^*$; then $\gamma(c, x)$ denotes the classic state of N after reading all of x and $\delta(c, x)$ denotes the total unitary transformation applied to the quantum register of N. In particular, $\gamma(c, \epsilon) = c$ and $\delta(c, \epsilon) = 1_{\mathrm{U}(V)}$, where ϵ denotes the empty-string. If $x = \sigma y$, for some $\sigma \in \Sigma$ and $y \in \Sigma^*$, then $\gamma(c, x) = \gamma(\gamma(c, \sigma), y)$ and $\delta(c, x) = \delta(c, \sigma)\delta(\gamma(c, \sigma), y)$. For an input string $w \in \Sigma^*$, let $\langle \psi_N(w)| = \langle \psi_{\mathrm{start}}|\delta(c_{\mathrm{start}}, w)$ denote the configuration of the quantum register of N immediately after reading the entire string w, before performing the quantum measurement. Furthermore, for $w \in \Sigma^*$, let $p_N(w)$ denote the probability that N accepts w. Define $R_w \subseteq \{0, 1\}$ such that $F_{\mathrm{acc}} \cap (\gamma(c_{\mathrm{start}}, w) \times \{0, 1\}) = \gamma(c_{\mathrm{start}}, w) \times R_w$. Then $p_N(w) = \sum_{r \in R_w} \| \langle \psi_N(w)|P_r \|^2 = \sum_{r \in R_w} \| \langle \psi_{\mathrm{start}}|\delta(c_{\mathrm{start}}, w)P_r \|^2$. In particular, if $R_w = \{0, 1\}$, then $p_N(w) = 1$; if $R_w = \emptyset$, then $p_N(w) = 0$; if $R_w = \{r\}$, then $p_N(w) = \| \langle \psi_{\mathrm{start}}|\delta(c_{\mathrm{start}}, w)P_r \|^2$.

Much as was the case for 1QFA, a 1QCFA could be specified using a finite set of quantum basis states Q in place of V, where $|Q| = \dim(V)$; for this reason, we say that a 1QCFA has k quantum basis states if the dimension of its underlying Hilbert space is k. Then, for $k, d \in \mathbb{N}_{\geq 1}$, we define the complexity class $\mathsf{E1QCFA}(k, d)$ (resp. $\mathsf{N1QCFA}(k, d)$, $\mathsf{R1QCFA}(k, d)$, $\mathsf{coN1QCFA}(k, d)$, $\mathsf{coR1QCFA}(k, d)$) to be the set of languages recognized with zero-error (resp. positive one-sided unbounded-error, positive one-sided bounded-error, negative one-sided unbounded-error, negative one-sided bounded-error) by some 1QCFA with at most k quantum basis states and at most d classical states (where each error-type is as defined in Section 4.1). We further define the complexity classes $\mathsf{E1QCFA} = \cup_{k,d \in \mathbb{N}_{\geq 1}} \mathsf{E1QCFA}(k, d)$, $\mathsf{N1QCFA} = \cup_{k,d \in \mathbb{N}_{\geq 1}} \mathsf{N1QCFA}(k, d)$, etc., to be those languages recognized in such a manner by a 1QCFA with any finite number of quantum basis states and classical states.

Note that the 1QFA model can be thought of as a special case of the 1QCFA model with only one classical state. That is to say, given a 1QCFA $N = (V, C, \Sigma, \delta, \gamma, \langle \psi_{\mathrm{start}}|, c_{\mathrm{start}}, V_0, F_{\mathrm{acc}})$, where $|C| = 1$ (and so $C = \{c_{\mathrm{start}}\}$), there is an equivalent 1QFA $M = (V, \Sigma, \delta', \langle \psi_{\mathrm{start}}|, V_{\mathrm{acc}})$, where M is equivalent to N in the sense that, $\forall w \in \Sigma^*$, $p_M(w) = p_N(w)$. In particular, the remaining pieces of M are defined as follows. First, $\forall \sigma \in \Sigma$, $\delta'(\sigma) = \delta(c_{\mathrm{start}}, \sigma)$. Then, defining $T \subseteq \{0, 1\}$ such that $F_{\mathrm{acc}} = \{c_{\mathrm{start}}\} \times T$, we set $V_{\mathrm{acc}} = \cup_{t \in T} V_t$. It is straightforward to see that M is equivalent to N. In the other direction, any 1QFA M immediately yields an equivalent 1QCFA N with a single classical state. Therefore, $\forall k \in \mathbb{N}_{\geq 1}$, we have $\mathsf{E1QFA}(k) = \mathsf{E1QCFA}(k, 1)$, $\mathsf{N1QFA}(k) = \mathsf{N1QCFA}(k, 1)$, etc.

5.2 The Relationship Between 1QCFA and 1QFA Groups

In this section, we establish the precise relationship between the class of groups whose word problem is recognized (with any particular error-type) by a 1QCFA with at most k quantum basis states and at most d classical states, and the class of groups whose word problem is recognized (with the same error-type) by a 1QFA with at most k quantum basis states. To avoid making equivalent statements for each of $\mathsf{E1QCFA}(k, d)$, $\mathsf{N1QCFA}(k, d)$, etc., we first establish a bit of additional notation. Let $\mathcal{T}$ denote an error-type; that is to say, $\mathcal{T}$ is one of E, coN, etc., and $\mathcal{T}\mathsf{1QCFA}(k, d)$ denotes the corresponding class $\mathsf{E1QCFA}(k, d)$, $\mathsf{coN1QCFA}(k, d)$, etc. We define $\mathcal{T}\mathsf{1QFA}(k)$ analogously. We show that, for any $k, d \in \mathbb{N}_{\geq 1}$, we have $W_G \in \mathcal{T}\mathsf{1QCFA}(k, d)$ if and only if $\exists H \leq G$ such that $[G : H] \leq d$ and $W_H \in \mathcal{T}\mathsf{N1QFA}(k)$. As an immediately corollary, we establish a precise classification of those group word problems in $\mathsf{E1QCFA}(k, d)$, $\mathsf{R1QCFA}(k, d)$, and $\mathsf{N1QCFA}(k, d)$, for any $k, d \in \mathbb{N}_{\geq 1}$. We require several lemmas.

Lemma 5.1. *Consider a group $G = \langle S|R \rangle$, with S finite, and define $W_G := W_{G=\langle S|R \rangle}$. Suppose the 1QCFA $N = (V, C, \Sigma, \delta, \gamma, \langle \psi_{start}|, c_{start}, V_0, F_{acc})$ recognizes W_G (with any particular error-type). Furthermore, suppose $\exists \widehat{c} \in C, \exists \widehat{x} \in W_G$ such that $\gamma(\widehat{c}, \widehat{x}) \neq \widehat{c}$. Then there is a 1QCFA*

$N' = (V, C', \Sigma, \delta', \gamma', \langle\psi'_{start}|, c'_{start}, V_0, F'_{acc})$ *that recognizes* W_G *with the same error-type, where* $C' = C \setminus \{\widehat{c}\}$.

Proof. Let $\widehat{c}' = \gamma(\widehat{c}, \widehat{x}) \neq \widehat{c}$. Fundamentally, N' is obtained from N by removing the single classic state $\widehat{c}$ and modifying both the classic and quantum part of the transition function such that whenever N reads a symbol $\sigma \in \Sigma$ that would cause it to transition into the classic state $\widehat{c}$, N' instead simulates N on the string $\sigma\widehat{x}$, transitioning into the classic state $\widehat{c}'$ and performing the appropriate transformation on its quantum register. More precisely, for each $c \in C'$ and $\sigma \in \Sigma$, if $\gamma(c, \sigma) = \widehat{c}$, then

$$\gamma'(c, \sigma) = \gamma(c, \sigma\widehat{x}) = \gamma(\gamma(c, \sigma), \widehat{x}) = \gamma(\widehat{c}, \widehat{x}) = \widehat{c}'$$

and

$$\delta'(c, \sigma) = \delta(c, \sigma\widehat{x}) = \delta(c, \sigma)\delta(\gamma(c, \sigma), \widehat{x}) = \delta(c, \sigma)\delta(\widehat{c}, \widehat{x}).$$

If $\gamma(c, \sigma) \neq \widehat{c}$, then $\gamma'(c, \sigma) = \gamma(c, \sigma)$ and $\delta'(c, \sigma) = \delta(c, \sigma)$. If $\widehat{c} \neq c_{\text{start}}$, then $c'_{\text{start}} = c_{\text{start}}$ and $\langle\psi'_{\text{start}}| = \langle\psi_{\text{start}}|$; if $\widehat{c} = c_{\text{start}}$, then $c'_{\text{start}} = \gamma(c_{\text{start}}, \widehat{x}) = \widehat{c}'$ and $\langle\psi'_{\text{start}}| = \langle\psi_{\text{start}}|\delta(c_{\text{start}}, \widehat{x})$. Lastly, $F'_{\text{acc}} = F_{\text{acc}} \cap (C' \times \{0, 1\})$.

To see that N' recognizes W_G with the same type of error as N, consider an input string $w' = w'_1 \cdots w'_n \in \Sigma^*$, where each $w_j \in \Sigma$. Let $J = \{j \in \{0, \ldots, n\} : \gamma(c_{\text{start}}, w'_1 \cdots w'_j) = \widehat{c}\}$ and $r = |J|$. Then, when reading the input w', N visits $\widehat{c}$ precisely r times. If $r = 0$, then we are done, as the computation of N' on w' is identical to that of N on w'. If $r > 0$, then let $J = \{j_1, \ldots, j_r\}$, where $j_1 < \ldots < j_r$. Let $y_1 = w'_1 \cdots w'_{j_1}$ and, for each $i \in \{2, \ldots, r\}$, let $y_i = w'_{j_{i-1}+1} \cdots w'_{j_i}$. Then $w' = y_1 \cdots y_r z$, for some $z \in \Sigma^*$. Let ϵ denote the empty-string. Then $y_1 = \epsilon$ precisely when $\widehat{c} = c_{\text{start}}$, and $z = \epsilon$ precisely when $\widehat{c} = \gamma(c_{\text{start}}, w')$. Clearly, $y_2, \ldots, y_r \neq \epsilon$. Define $\widetilde{w} = y_1\widehat{x}y_2\widehat{x} \cdots y_r\widehat{x}z$, and observe that $\widehat{x} \in W_G$ implies $\phi(\widetilde{w}) = \phi(w')$. We next show that N' on input w' behaves identically to N on input $\widetilde{w}$. We extend γ' and δ' to $\gamma' : C' \times \Sigma^* \to C'$ and $\delta' : C' \times \Sigma^* \to C'$ as defined in Section 5.1.

Observe that $\gamma'(c'_{\text{start}}, y_1) = \widehat{c}'$. To see this, note that if $\widehat{c} = c_{\text{start}}$, then $y_1 = \epsilon$ and $c'_{\text{start}} = \gamma(c_{\text{start}}, \widehat{x})$ which implies $\gamma'(c'_{\text{start}}, y_1) = \gamma'(\gamma(c_{\text{start}}, \widehat{x}), \epsilon) = \gamma(c_{\text{start}}, \widehat{x}) = \widehat{c}'$. If, instead, $\widehat{c} \neq c_{\text{start}}$, then $c'_{\text{start}} = c_{\text{start}}$ and $y_1 \neq \epsilon$; we may then write $y_1 = t\sigma$, for some $t \in \Sigma^*$, $\sigma \in \Sigma$, and then define $d = \gamma(c_{\text{start}}, t)$. As y_1 is, by definition, the shortest prefix of w' such that $\gamma(c_{\text{start}}, y_1) = \widehat{c}$, we then have $\gamma'(c'_{\text{start}}, t) = \gamma(c_{\text{start}}, t) = d \in C' \subseteq C$, as N' behaves identically to N when reading t; furthermore, $\gamma(d, \sigma) = \widehat{c}$, which then implies $\gamma'(d, \sigma) = \widehat{c}'$. Therefore,

$$\gamma'(c'_{\text{start}}, y_1) = \gamma'(c'_{\text{start}}, t\sigma) = \gamma'(\gamma'(c'_{\text{start}}, t), \sigma) = \gamma'(d, \sigma) = \widehat{c}'.$$

For each $i \in \{2, \ldots, r\}$, we have $y_i \neq \epsilon$, and so we may write $y_i = t_i\sigma_i$ for some $t_i \in \Sigma^*$, $\sigma_i \in \Sigma$. By the same reasoning as above, we then have, $\gamma'(\widehat{c}', t_i) = \gamma(\widehat{c}', t_i) = d_i$, for some $d_i \in C' \subseteq C$ such that $\gamma(d_i, \sigma_i) = \widehat{c}$; this implies

$$\gamma'(\widehat{c}', y_i) = \gamma'(\widehat{c}', t_i\sigma_i) = \gamma'(\gamma'(\widehat{c}', t_i), \sigma_i) = \gamma'(d_i, \sigma_i) = \widehat{c}'.$$

Therefore,

$$\gamma'(c'_{\text{start}}, w') = \gamma'(c'_{\text{start}}, y_1 \cdots y_r z) = \gamma'(\cdots\gamma'(\gamma'(c'_{\text{start}}, y_1), y_2)\cdots, z) = \gamma'(\cdots\gamma'(\gamma'(\widehat{c}', y_2), y_3)\cdots, z)$$

$$= \gamma'(\cdots\gamma'(\gamma'(\widehat{c}', y_3), y_4)\cdots, z) = \cdots = \gamma'(\widehat{c}', z) = \gamma(\widehat{c}', z) = \gamma(\gamma(c_{\text{start}}, y_1\widehat{x} \cdots y_r\widehat{x}), z) = \gamma(c_{\text{start}}, \widetilde{w}).$$

By an analogous argument, we also conclude $\langle\psi'_{\text{start}}|\delta'(c'_{\text{start}}, w') = \langle\psi_{\text{start}}|\delta(c_{\text{start}}, \widetilde{w})$, which then implies $p_{N'}(w') = p_N(\widetilde{w})$. Finally, $\phi(\widetilde{w}) = \phi(w')$ implies that N' recognizes W_G with the same error-type as N. □

Lemma 5.2. *For any finitely-generated group G and error-type $\mathcal{T}$, if $W_G \in \mathcal{T}\mathsf{1QCFA}(k,d)$, then $\exists H \leq G$, with $[G:H] \leq d$, such that $W_H \in \mathcal{T}\mathsf{1QFA}(k)$.*

Proof. Let d' denote the minimal value such that $W_G \in \mathcal{T}\mathsf{1QCFA}(k,d')$; then $1 \leq d' \leq d$. Let $G = \langle S_G | R_G \rangle$ with S_G finite. By definition, there is a 1QCFA $N = (V, C, \Sigma, \delta, \gamma, \langle \psi_{\text{start}}|, c_{\text{start}}, V_0, F_{\text{acc}})$ that recognizes $W_G := W_{G=\langle S_G|R_G\rangle}$ with error-type $\mathcal{T}$, where $\dim(V) \leq k$ and $|C| = d'$.

Observe that, $\forall c \in C, \forall x \in W_G$, we must have $\gamma(c,x) = c$. To see this, notice that, if $\exists c \in C, \exists x \in W_G$ such that $\gamma(c,x) \neq c$, then Lemma 5.1 would imply the existence of a 1QCFA N' with $d'-1$ classical states that also recognizes W_G with error-type $\mathcal{T}$; this would then imply $W_G \in \mathcal{T}\mathsf{1QCFA}(k, d'-1)$, which contradicts the minimality of d'.

Next, notice that, $\forall c \in C$, $\forall y, z \in \Sigma^*$, if $\phi(y) = \phi(z)$, then $\gamma(c,y) = \gamma(c,z)$. This follows from the fact that, if $\phi(y) = \phi(z)$, then $zy^{-1} \in W_G$, which implies $\gamma(c, zy^{-1}) = c$; we then have

$$\gamma(c,y) = \gamma(\gamma(c, zy^{-1}), y) = \gamma(c, zy^{-1}y) = \gamma(\gamma(c,z), y^{-1}y) = \gamma(c,z).$$

We then define the function $\eta : G \to C$ such that, for any $g \in G$, $\eta(g) = \gamma(c_{\text{start}}, y)$, for some $y \in \phi^{-1}(g)$. The function η is well-defined as for any other $z \in \phi^{-1}(g)$, the above implies $\gamma(c_{\text{start}}, y) = \gamma(c_{\text{start}}, z)$. We then define $H = \eta^{-1}(c_{\text{start}})$.

To see that H is a group, and, therefore, that $H \leq G$, notice first that H is non-empty as $\gamma(c_{\text{start}}, x) = c_{\text{start}}$, $\forall x \in W_G = \phi^{-1}(1_G)$, which implies $1_H = 1_G \in H$. Next, consider any $h_1, h_2 \in H$, and fix any $y \in \phi^{-1}(h_1)$ and $z \in \phi^{-1}(h_2)$. Then

$$\eta(h_1h_2) = \gamma(c_{\text{start}}, yz) = \gamma(\gamma(c_{\text{start}}, y), z) = \gamma(c_{\text{start}}, z) = c_{\text{start}} \Rightarrow h_1h_2 \in H,$$

and

$$\eta(h_1^{-1}) = \gamma(c_{\text{start}}, y^{-1}) = \gamma(\gamma(c_{\text{start}}, y), y^{-1}) = \gamma(c_{\text{start}}, yy^{-1}) = c_{\text{start}} \Rightarrow h_1^{-1} \in H.$$

Let $C = \{c_1, \ldots, c_{d'}\}$ and fix $y_1, \ldots, y_{d'} \in \Sigma^*$ such that $\gamma(c_{\text{start}}, y_i) = c_i$ (if such a y_i did not exist, then the state c_i is never entered when reading any possible input string; then c_i could be deleted, which would yield a 1QCFA with $d'-1$ classic states that recognizes W_G with the same error-type as N, thereby contradicting the minimality of d'). For each $i \in \{1, \ldots, d'\}$, let $t_i = \phi(y_i) \in G$. Then $\eta(t_i) = \gamma(c_{\text{start}}, y_i) = c_i$. Consider $g \in G$ and fix $z \in \phi^{-1}(g)$. We have $\gamma(c_{\text{start}}, z) = \eta(g) = c_i = \eta(t_i)$, for some $i \in \{1, \ldots, d'\}$; we then have

$$\eta(gt_i^{-1}) = \gamma(c_{\text{start}}, zy_i^{-1}) = \gamma(\gamma(c_{\text{start}}, z), y_i^{-1}) = \gamma(\eta(t_i), y_i^{-1}) = \gamma(\gamma(c_{\text{start}}, y_i), y_i^{-1}) = c_{\text{start}}.$$

Therefore, if $\eta(g) = \eta(t_i)$, then $gt_i^{-1} \in H$, which implies $g \in Ht_i$. We then conclude that $\{Ht_i : i \in \{1, \ldots, d'\}\}$ is a complete set of right cosets of H in G; in particular $[G:H] = d' \leq d$, as desired.

All that remains is to show $W_H \in \mathcal{T}\mathsf{1QFA}(k)$. We first establish a convenient presentation of H (as discussed earlier, membership of W_H in $\mathcal{T}\mathsf{1QFA}(k)$ does not depend on the particular choice of presentation beyond the requirement that the generating set is finite). Let $S_G = \{g_1, \ldots, g_{|S_G|}\}$. For each $l \in \{1, \ldots, d'\}$ and each $j \in \{1, \ldots, |S_G|\}$, we have $t_l g_j \in Ht_i$, for a unique $i \in \{1, \ldots, d'\}$; let $h_{l,j} \in H$ and $r(l,j) \in \{1, \ldots, d'\}$ denote the unique values such that $t_l g_j = h_{l,j} t_{r(l,j)}$. Let $S_H = \{h_{l,j} : l \in \{1, \ldots, d'\}, j \in \{1, \ldots, |S_G|\}\}$. It is straightforward to verify that S_H is a (finite) generating set for H. Define R_H such that $H = \langle S_H | R_H \rangle$. Let $\Sigma_H = S_H \sqcup S_H^{-1}$ (we may assume, without loss of generality, that $S_H \cap S_H^{-1} = \emptyset$) and let $\phi_H : \Sigma_H^* \to H$ denote the natural map.

We exhibit a 1QFA $M = (V, \Sigma_H, \delta', \langle \psi_{\text{start}}|, V'_{\text{acc}})$ that recognizes $W_H := W_{H=\langle S_H|R_H\rangle} = \phi_H^{-1}(1_H) = \phi_H^{-1}(1_G)$ with the appropriate error-type. The 1QFA M has the same underlying Hilbert space V and quantum start configuration $\langle \psi_{\text{start}}|$ as the original 1QCFA N. To define the

transition function $\delta' : \Sigma_H \to \mathrm{U}(V)$, we first define $z_1, \ldots, z_{|S_G|} \in \Sigma^*$ such that $\phi(z_j) = g_j$. Then, for $l \in \{1, \ldots, d'\}$, $j \in \{1, \ldots, |S_G|\}$, we have

$$h_{l,j} = t_l g_j t_{r(l,j)}^{-1} = \phi(y_l)\phi(z_j)\phi(y_{r(l,j)})^{-1} = \phi(y_l z_j y_{r(l,j)}^{-1}).$$

We then define $\delta'(h_{l,j}) = \delta(c_{\mathrm{start}}, y_l z_j y_{r(l,j)}^{-1})$ and $\delta'(h_{l,j}^{-1}) = \delta(c_{\mathrm{start}}, y_{r(l,j)} z_j^{-1} y_l^{-1})$. We define $F \subseteq \{0,1\}$ such that $F_{\mathrm{acc}} \cap (c_{\mathrm{start}} \times \{0,1\}) = (c_{\mathrm{start}} \times F)$. Finally, we define $V'_{\mathrm{acc}} = \cup_{f \in F} V_f$ (recall that $V_1 = V_0^{\perp} \subseteq V$).

To see that M recognizes W_H with the appropriate error-type, consider an input string $w = w_1 \cdots w_n \in \Sigma_H^*$, where each $w_i \in \Sigma_H$. For $i \in \{1, \ldots, n\}$, define $b_i \in \Sigma^*$ such that, if $w_i = h_{l,j}$, then $b_i = y_l z_j y_{r(l,j)}^{-1}$, and if instead $w_i = h_{l,j}^{-1}$, then $b_i = y_{r(l,j)} z_j^{-1} y_l^{-1}$. Notice that, by construction $\phi(b_i) = \phi_H(w_i)$, $\delta'(w_i) = \delta(c_{\mathrm{start}}, b_i)$, and $\gamma(c_{\mathrm{start}}, b_i) = \eta(\phi(b_i)) = c_{\mathrm{start}}$. Let $\widetilde{w} = b_1 \cdots b_n \in \Sigma^*$. Then $\phi_H(w) = \phi_H(w_1) \cdots \phi_H(w_n) = \phi(b_1) \cdots \phi(b_n) = \phi(\widetilde{w})$. Moreover, $\gamma(c_{\mathrm{start}}, b_i) = c_{\mathrm{start}}$ implies

$$\delta(c_{\mathrm{start}}, \widetilde{w}) = \delta(c_{\mathrm{start}}, b_1 \cdots b_n) = \delta(c_{\mathrm{start}}, b_1)\delta(\gamma(c_{\mathrm{start}}, b_1), b_2 \cdots b_n) = \delta(c_{\mathrm{start}}, b_1)\delta(c_{\mathrm{start}}, b_2 \cdots b_n)$$

$$= \delta(c_{\mathrm{start}}, b_1)\delta(c_{\mathrm{start}}, b_2)\delta(\gamma(c_{\mathrm{start}}, b_2), b_3 \cdots b_n) = \cdots = \delta(c_{\mathrm{start}}, b_1)\delta(c_{\mathrm{start}}, b_2) \cdots \delta(c_{\mathrm{start}}, b_n).$$

Therefore,

$$\delta'(w) = \delta'(w_1) \cdots \delta'(w_n) = \delta(c_{\mathrm{start}}, b_1) \cdots \delta(c_{\mathrm{start}}, b_n) = \delta(\widetilde{w}).$$

This immediately implies $p_M(w) = p_N(\widetilde{w})$; as noted above, $\phi_H(w) = \phi(\widetilde{w})$, which then implies M recognizes W_H with the same error-type with which N recognizes W_G. □

Lemma 5.3. *Consider finitely-generated groups G, H, with $H \leq G$ and $[G : H] = d$. If $W_H \in \mathcal{T}\mathsf{1QFA}(k)$, for some error-type $\mathcal{T}$, then $W_G \in \mathcal{T}\mathsf{1QCFA}(k, d)$.*

Proof. We make use of the standard "coset automaton" construction (see, for instance [30, Lemma 3] or [35, Lemma 4.7]) to show that a 1QFA for W_H can be used to produce a 1QCFA for W_G. Let $H = \langle S_H | R_H \rangle$, for some finite set S_H. We begin by constructing a convenient presentation for G. Let $C = \{c_1, \ldots, c_d\} \subseteq G$, with $c_1 = 1_G = 1_H$, denote a complete set of right coset representatives of H in G (i.e., for each $g \in G$, there is a unique $c_i \in C$ such that $g \in Hc_i$). We then define $S_G = S_H \sqcup \{c_2, \ldots, c_d\}$ and $\Sigma_G = S_G \cup S_G^{-1}$. For $\sigma \in \Sigma_G \subseteq G$ and $c_j \in C \subseteq G$, consider the element $c_j\sigma \in G$; there is a unique $h \in H$ and $c_i \in C$ such that $c_j\sigma = hc_i$. We then define functions $\beta : C \times \Sigma_G \to H$ and $\gamma : C \times \Sigma_G \to C$ such that $c_j\sigma = \beta(c_j, \sigma)\gamma(c_j, \sigma)$, $\forall \sigma \in \Sigma_G, \forall c_j \in C$. Then $G = \langle S_G | R_G \rangle$, where $R_G = R_H \cup \{\sigma^{-1} c_j^{-1} \beta(c_j, \sigma)\gamma(c_j, \sigma) : \sigma \in \Sigma_G, c_j \in C\}$.

Let $\Sigma_H = S_H \cup S_H^{-1}$, let $\phi_H : \Sigma_H^* \to H$ and $\phi_G : \Sigma_G^* \to G$ denote the natural maps, and let $W_H := W_{H=\langle S_H | R_H \rangle} = \phi_H^{-1}(1_H)$ and $W_G := W_{G=\langle S_G | R_G \rangle} = \phi_G^{-1}(1_G)$ denote the word problems of the groups H and G, respectively. By definition, there is a 1QFA $M = (V, \Sigma_H, \delta, \langle \psi_{\mathrm{start}}|, V_{\mathrm{acc}})$, with $\dim(V) \leq k$, that recognizes W_H with error-type $\mathcal{T}$.

We next exhibit a 1QCFA $N = (V, C, \Sigma_G, \delta', \gamma, \langle \psi_{\mathrm{start}}|, c_{\mathrm{start}}, V_0, F_{\mathrm{acc}})$ that recognizes W_G with error-type $\mathcal{T}$. The main idea is that, for an input $w = w_1 \cdots w_n \in \Sigma_G^*$, if $\phi_G(w) = \widehat{h}\widehat{c}$, for some $\widehat{h} \in H$ and $\widehat{c} \in C$, then $w \in W_G \Leftrightarrow \phi_G(w) = 1_G \Leftrightarrow (\widehat{h} = 1_H$ and $\widehat{c} = 1_G)$. The 1QCFA N will operate such that, after reading the prefix $w_1 \cdots w_l$ of any length $l \in \{0, \ldots, n\}$, if $\phi_G(w_1 \cdots w_l) = hc_i$, then the classic state of N will be c_i and N will have used its quantum states to simulate M on a string $x \in \Sigma_H^*$ such that $\phi_H(x) = h$. In particular, after reading the entire input w, the classic state of N will be $\widehat{c}$ and M will have been simulated on a string $\widehat{x}$ such that $\phi_H(\widehat{c}) = \widehat{h}$, which will allow N to determine if $w \in W_G$.

We now fill in the details of the definition of N. The 1QCFA N has the same underlying Hilbert space V and initial quantum configuration $\langle \psi_{\mathrm{start}}|$ as the 1QFA M, it has alphabet Σ_G, and its

set of classic states C and classic transition function $\gamma : C \times \Sigma_G \to C$ are as defined during the above construction of the desired presentation of G. To define the remaining parts of N, we first define $\widehat{\beta} : C \times \Sigma_G \to \Sigma_H^*$ such that $\phi_H(\widehat{\beta}(c_j, \sigma)) = \beta(c_j, \sigma)$, $\forall c_j \in C, \forall \sigma \in \Sigma_G$. We then define $\delta' : C \times \Sigma_G \to \mathrm{U}(V)$ such that $\delta'(c_j, \sigma) = \delta(\widehat{\beta}(c_j, \sigma))$. Lastly, we define $c_{\text{start}} = c_1 = 1_G$, $V_0 = V_{\text{acc}}$, and $F_{\text{acc}} = \{(c_{\text{start}}, 0)\}$.

All that remains is to show that N recognizes W_G with the appropriate type of error. We first extend γ and δ' to $\gamma : C \times \Sigma_G^* \to C$ and $\delta' : C \times \Sigma_G^* \to \mathrm{U}(V)$ as specified in Section 5.1 and extend δ to $\delta : \Sigma_H^* \to \mathrm{U}(V)$ as specified in Section 4.1. For any $w = w_1 \cdots w_n \in \Sigma_G^*$, where each $w_i \in \Sigma_G$, and for any $l \in \{0, \ldots, n\}$, let $c(w, l) = \gamma(c_{\text{start}}, w_1 \cdots w_l)$ denote the classic state of N after reading the prefix $w_1 \cdots w_l$ of w of length l (in particular $c(w, 0) = \gamma(c_{\text{start}}, \epsilon) = c_{\text{start}}$, where ϵ denotes the empty string) and let $c(w) = c(w, n) = \gamma(c_{\text{start}}, w)$ denote the classic state of N after reading the entire string w. For any $l \in \{1, \ldots, n\}$, we define $y(w, l) = \widehat{\beta}(c(w, l-1), w_l) \in \Sigma_H^*$, and we also define $y(w, 0) = \epsilon$. Lastly, we define $x(w, l) = y(w, 1) \cdots y(w, l)$ and $x(w) = x(w, n)$.

We next show that $\phi_G(w_1 \cdots w_l) = \phi_H(x(w, l))c(w, l)$, $\forall w = w_1 \cdots w_n \in \Sigma_G^*$, $\forall l \in \{0, \ldots, n\}$ (note that $c(w, l) \in C \subseteq G$). This claim follows straightforwardly by induction on l. For $l = 0$, the claim is immediate. For $l > 0$, by definition, $c(w, l)\phi_G(w_{l+1}) = \beta(c(w, l), w_{l+1})\gamma(c(w, l), w_{l+1})$, which then implies

$$\phi_G(w_1 \cdots w_{l+1}) = \phi_G(w_1 \cdots w_l)\phi_G(w_{l+1}) = \phi_H(x(w, l))c(w, l)\phi_G(w_{l+1})$$

$$= \phi_H(x(w, l))\beta(c(w, l), w_{l+1})\gamma(c(w, l), w_{l+1}) = \phi_H(x(w, l))\phi_H(\widehat{\beta}(c(w, l), w_{l+1}))c(w, l+1)$$

$$= \phi_H(x(w, l))\phi_H(y(w, l+1))c(w, l+1) = \phi_H(x(w, l+1))c(w, l+1).$$

Therefore, $w \in W_G \Leftrightarrow \phi_G(w) = 1_G \Leftrightarrow (\phi_H(x(w)) = 1_H$ and $c(w) = 1_G = c_{\text{start}})$. Moreover,

$$\delta'(c_{\text{start}}, w) = \prod_{l=1}^{n} \delta'(c(w, l-1), w_l) = \prod_{l=1}^{n} \delta(\widehat{\beta}(c(w, l-1), w_l)) = \prod_{l=1}^{n} \delta(y(w, l)) = \delta\left(\prod_{l=1}^{n} y(w, l)\right) = \delta(x(w)).$$

As $F_{\text{acc}} = \{(c_{\text{start}}, 0)\}$, we conclude that if $c(w) \neq c_{\text{start}}$, then $p_N(w) = 0$, if instead $c(w) = c_{\text{start}}$, then

$$p_N(w) = \|\langle \psi_{\text{start}} | \delta'(c_{\text{start}}, w) P_0 \|^2 = \|\langle \psi_{\text{start}} | \delta(x(w)) P_{\text{acc}} \|^2 = p_M(x(w)).$$

Therefore, $\forall w \in W_G$, $c(w) = c_{\text{start}}$, and $p_N(w) = p_M(x(w))$, for some $x(w) \in W_H$; $\forall w \notin W_G$, either $c(w) \neq c_{\text{start}}$, in which case $p_N(w) = 0$, or $c(w) = c_{\text{start}}$, in which case $p_N(w) = p_M(x(w))$, for some $x(w) \notin W_H$. Therefore, N recognizes W_G with error-type $\mathcal{T}$. □

By combining Lemma 5.2 and Lemma 5.3, we obtain the relationship claimed at the start of this section between 1QCFA groups and 1QFA groups.

Theorem 5.4. *For any finitely-generated group G, any $k, d \in \mathbb{N}_{\geq 1}$, and any error-type $\mathcal{T}$, we have $W_G \in \mathcal{T}\mathsf{1QCFA}(k, d)$ if and only if $\exists H \leq G$ such that $[G : H] \leq d$ and $W_H \in \mathcal{T}\mathsf{1QFA}(k)$.*

5.3 Classification of the N1QCFA, R1QCFA, and E1QCFA Groups

We then have the following classification of the $\mathsf{N1QCFA}(k, d)$, $\mathsf{R1QCFA}(k, d)$, and $\mathsf{E1QCFA}(k, d)$ groups.

Theorem 5.5. *For any finitely-generated group G, and any $k, d \in \mathbb{N}_{\geq 1}$, the following are equivalent.*

(i) $W_G \in \mathsf{N1QCFA}(k, d)$

(ii) $W_G \in \mathsf{R1QCFA}(k,d)$

(iii) $W_G \in \mathsf{E1QCFA}(k,d)$

(iv) $\exists H \leq G$ *such that* $[G:H] \leq d$ *and* $|H| \leq k$.

Proof. Follows immediately from Theorem 4.1 and Theorem 5.4. □

Corollary 5.5.1. *For any finitely-generated group* G*, and any* $k, d \in \mathbb{N}_{\geq 1}$*, if* $W_G \in \mathsf{N1QCFA}(k,d)$*, then* $W_G \in \mathsf{permDFA}(kd)$.

Corollary 5.5.2. *For any* $k, d \in \mathbb{N}_{\geq 1}$*, the following statements hold.*

(i) $\mathsf{N1QCFA}(k,d) \subsetneq \mathsf{N1QCFA}(k+1,d)$.

(ii) $\mathsf{N1QCFA}(k,d) \subsetneq \mathsf{N1QCFA}(k,d+1)$.

(iii) $\mathsf{R1QCFA}(k,d) \subsetneq \mathsf{R1QCFA}(k+1,d)$.

(iv) $\mathsf{R1QCFA}(k,d) \subsetneq \mathsf{R1QCFA}(k,d+1)$.

(v) $\mathsf{E1QCFA}(k,d) \subsetneq \mathsf{E1QCFA}(k+1,d)$.

(vi) $\mathsf{E1QCFA}(k,d) \subsetneq \mathsf{E1QCFA}(k,d+1)$.

(vii) There is a language $L \in \mathsf{permDFA}(kd+1)$ *such that* $L \notin \mathsf{N1QCFA}(k,d)$.

Proof. For any $k, d \in \mathbb{N}_{\geq 1}$, let $G_{k,d} = (\mathbb{Z}/k\mathbb{Z}) \times (\mathbb{Z}/d\mathbb{Z})$. Note that $\mathbb{Z}/k\mathbb{Z} \leq G_{k,d}$, $[G_{k,d} : \mathbb{Z}/k\mathbb{Z}] = d$, and $|\mathbb{Z}/k\mathbb{Z}| = k$. Therefore, the preceding theorem implies $W_{G_{k,d}} \in \mathsf{E1QCFA}(k,d) \subseteq \mathsf{R1QCFA}(k,d) \subseteq \mathsf{N1QCFA}(k,d)$. Moreover, if, for some $k', d' \in \mathbb{N}_{\geq 1}$, $\exists H \leq G_{k,d}$ such that $[G_{k,d} : H] \leq d'$ and $|H| \leq k'$ then $kd = |G_{k,d}| = [G_{k,d} : H]|H| \leq k'd'$. Therefore, if $k'd' < kd$, the preceding theorem also implies $W_{G_{k,d}} \notin \mathsf{N1QCFA}(k',d') \supseteq \mathsf{R1QCFA}(k',d') \supseteq \mathsf{E1QCFA}(k',d')$. Lastly, $W_{\mathbb{Z}/(kd+1)\mathbb{Z}} \in \mathsf{permDFA}(kd+1)$ by [30, Lemma 1], and $W_{\mathbb{Z}/(kd+1)\mathbb{Z}} \notin \mathsf{N1QCFA}(k,d)$ by the above. □

Corollary 5.5.3. *For a finitely-generated group* G*,* $W_G \in \mathsf{N1QCFA} \Leftrightarrow W_G \in \mathsf{R1QCFA} \Leftrightarrow W_G \in \mathsf{E1QCFA} \Leftrightarrow W_G \in \mathsf{N1QFA} \Leftrightarrow W_G \in \mathsf{R1QFA} \Leftrightarrow W_G \in \mathsf{E1QFA} \Leftrightarrow W_G \in \mathsf{REG} \Leftrightarrow G$ *is a finite group.*

6 Discussion

6.1 The Computational Complexity of Group Word Problems

Let CFL (resp. DCFL) denote the context-free languages (resp. deterministic context-free languages), the class of languages recognizable by a non-deterministic (resp. deterministic) pushdown automaton. Let OCL (resp. DOCL) denote the one-counter languages (resp. deterministic one-counter languages), the class of languages recognizable by a non-deterministic (resp. deterministic) pushdown automaton whose stack alphabet is limited to a single symbol. Let $\mathsf{poly{-}CFL}$ (resp. $\mathsf{poly{-}DCFL}$, $\mathsf{poly{-}OCL}$, $\mathsf{poly{-}DOCL}$) denote the class of languages expressible as the intersection of finitely many context-free (resp. deterministic context-free, one-counter, deterministic one-counter) languages. Let L denote the class of languages recognizable by a deterministic logspace Turing machine.

In order to make clear the relationship between a class of groups and the computational complexity of the corresponding word problems, we write $\widehat{\Pi}_0$ (resp. $\widehat{\Pi}_1$, $\widehat{\Sigma}_1$, $\widehat{\Pi}_2$, $\widehat{\Sigma}_2$) for the finitely-generated groups that are virtually cyclic (resp. abelian, free, a subgroup of a direct product of finitely many finite-rank free groups, a subgroup of a free product of finitely many finite-rank free abelian groups); see [35] for a more thorough explanation of this choice of notation. We also write $\widehat{\{1\}}$ for the finite groups, and $\mathcal{L}$ for the set of all finitely-generated groups G that are linear groups over some field of characteristic 0. As before, let $\mathcal{G}_{\mathbf{vNilp}}$ (resp. $\mathcal{G}_{\mathbf{vSolv}}$) denote the collection of all finitely-generated virtually nilpotent (resp. solvable) groups. Let $\overline{\mathbb{Q}}$ denote the algebraic numbers, 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.

The following proposition (which appeared in our recent paper [35]), which collects the results of many authors, demonstrates the extremely strong relationship between the computational complexity of W_G and certain algebraic properties of G.

Proposition 6.1. *([6, 7, 9, 11, 21, 23, 26, 30, 31]) Let G be a finitely-generated group, with word problem W_G.*

(i) $G \in \widehat{\{1\}} \Leftrightarrow W_G \in \mathsf{REG}$.

(ii) $G \in \widehat{\Pi}_0 \Leftrightarrow W_G \in \mathsf{OCL} \Leftrightarrow W_G \in \mathsf{DOCL}$.

(iii) $G \in \widehat{\Pi}_1 \Leftrightarrow W_G \in \mathsf{poly{-}OCL} \Leftrightarrow W_G \in \mathsf{poly{-}DOCL}$.

(iv) $G \in \widehat{\Sigma}_1 \Leftrightarrow W_G \in CFL \Leftrightarrow W_G \in \mathsf{DCFL}$.

(v) $G \in \widehat{\Pi}_2 \Rightarrow W_G \in \mathsf{poly{-}DCFL} \cap \mathsf{coCFL}$.

(vi) $G \in \mathcal{L} \Rightarrow W_G \in \mathsf{L}$.

Proof. Statements $(i), (ii), (iii), (v),$ and (vi) were shown, respectively, in [6],[21],[23], [9], and [26]. In [30], it was shown that G is free if and only if $W_G \in \mathsf{CFL}$ and G is accessible, in [11], it was shown that all finitely-presented groups are accessible, and in [7] it was shown that all context-free groups are finitely-presented, which implies the first equivalence in (iv). The second equivalence in (iv) was shown in [31]. □

We have shown that, if $G \in \widehat{\Pi}_1 = \mathcal{G}_{\mathbf{vAb}}$, then $W_G \in \mathsf{coRQP2QCFA}_{\overline{\mathbb{Q}}}(2) \subseteq \mathsf{BQP2QCFA}$ [35, Theorem 1.2]; moreover, if $W_G \in \mathsf{BQP2QCFA}$, then $G \in \mathcal{G}_{\mathbf{vNilp}}$ (Theorem 3.16(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 3.17). This naturally raises the following question.

Open Problem 6.2. *Is there a group $G \in \mathcal{G}_{\boldsymbol{vNilp}} \backslash \mathcal{G}_{\boldsymbol{vAb}}$ such that $W_G \in \mathsf{BQP2QCFA}$? In particular, is $W_H \in \mathsf{BQP2QCFA}$, where H is the Heisenberg group?*

We have shown that, if $G \in \mathcal{U}$, then $W_G \in \mathsf{coRQE2QCFA}_{\overline{\mathbb{Q}}} \subseteq \mathsf{BQE2QCFA}$ [35, Corollary 1.4.1]. Note that $\widehat{\Pi}_2 \subseteq \mathcal{U} \subseteq \mathcal{L}$, $\mathcal{G}_{\mathbf{vNilp}} \subseteq \mathcal{G}_{\mathbf{vSolv}} \subseteq \mathcal{L}$, and $\mathcal{U} \cap \mathcal{G}_{\mathbf{vSolv}} = \mathcal{G}_{\mathbf{vAb}} = \widehat{\Pi}_1$. We have also shown that, if G has exponential growth, then $W_G \notin \mathsf{B2QCFA}(T)$, for any $T : \mathbb{N} \to \mathbb{N}$ such that $T(n) = 2^{o(n)}$ (Theorem 3.16(ii)). This naturally raises several questions.

Open Problem 6.3. *(i) Is there a group $G \in \mathcal{L}$ such that $W_G \notin \mathsf{BQE2QCFA}$?*

(ii) Is there a group $G \in (\mathcal{G}_{\boldsymbol{vSolv}} \setminus \mathcal{G}_{\boldsymbol{vAb}}) \subseteq \mathcal{L}$ such that $W_G \in \mathsf{BQE2QCFA}$?

(iii) Is there an infinite (finitely-generated) Kazhdan group G such that $W_G \in \mathsf{BQE2QCFA}$?

(iv) Is there a group G of intermediate growth such that $W_G \in \mathsf{BQE2QCFA}$?

(v) Is there a group G such that $W_G \in \mathsf{BQE2QCFA}$, but $W_G \notin \mathsf{L}$?

We have shown that any W_G that a 1QCFA (and as a special case a 1QFA) can recognize with *positive* one-sided unbounded-error can also be recognized by an equivalently sized DFA (Corollary 5.5.1). This is precisely analogous to the situation for PDA: the class of group word problems recognizable by a deterministic PDA is identical to the class of group word problems recognizable by a PDA with *positive* ($\exists$) non-determinism [30]. We have also shown that 1QFA (and, therefore, 1QCFA) can recognize with *negative* one-sided unbounded-error a very broad class of groups. This is again precisely analogous to the situation for PDA: there are many groups G for which $W_G \notin \mathsf{DCFL}$ but $W_G \in \mathsf{coCFL}$ (i.e., W_G is recognizable by a PDA with *negative* ($\forall$) non-determinism). For any $G \in \widehat{\Pi}_2$, we have $W_G \in \mathsf{coN1QFA}$; moreover, for the group $\mathbb{Z} * \mathbb{Z}^2 \notin \widehat{\Pi}_2$ (but which does satisfy $\mathbb{Z} * \mathbb{Z}^2 \in \widehat{\Sigma}_2$), we have $W_{\mathbb{Z}*\mathbb{Z}^2} \in \mathsf{coN1QFA}(2)$ [35, Theorem 1.7]. The complexity of $W_{\mathbb{Z}*\mathbb{Z}^2}$ has been considered by many authors and it is conjectured that $W_G \notin \mathsf{poly{-}CFL}$ [9](cf. [10]) and that $W_{\mathbb{Z}*\mathbb{Z}^2} \notin \mathsf{coCFL}$ [24].

Open Problem 6.4. *Is there a group G such that $W_G \in \mathsf{coN1QFA}$, but $W_G \notin (\mathsf{poly{-}CFL \cup coCFL})$?*

6.2 The Quantum Register of a 2QCFA and Collapse-2QCFA

Consider a 2QCFA $N = (Q, C, \Sigma, \delta_{\text{type}}, \delta_{\text{transform}}, \delta_{\text{measure}}, q_{\text{start}}, c_{\text{start}}, c_{\text{acc}}, c_{\text{rej}})$. Suppose that $Q = \{q_0, \ldots, q_{|Q|-1}\}$. We consider the special case in which N only performs quantum measurements defined by the partition $B = \{\{q_0\}, \ldots, \{q_{|Q|-1}\}\}$ of Q into singletons; that is to say N only performs *complete* measurements of its quantum register with respect to the computational basis, as opposed to the more general case in which N is allowed to perform *partial* measurements of its quantum register. If the quantum register of N is in the superposition $|\psi\rangle \in \Psi$ when performing this quantum measurement, then the probability that the result is $q_r \in Q$ (where here, and throughout this section, we denote the result $\{q_r\} \in B$ simply by $q_r \in Q$, for brevity) is $|\langle q_r|\psi\rangle|^2$; if the result is q_r, then the state of the quantum register collapses to $|q_r\rangle$. In particular, after performing a single quantum measurement, all information in the quantum register is destroyed, which greatly simplifies the description of the behavior of N. We call such a 2QCFA a *collapse*-2QCFA.

As collapse-2QCFA are a special case of 2QCFA, Theorem 3.13 (which gives a lower bound on the expected running time of any 2QCFA that recognizes a language L in terms of the hardness measure D_L) of course also applies to collapse-2QCFA, though we note that a more elementary analysis could yield this result for collapse-2QCFA. Importantly, if N has only a single qubit (i.e., $|Q| = 2$), then the only non-trivial quantum measurement that N can perform is the measurement defined by the partition $B = \{\{q_0\}, \{q_1\}\}$, and so every single-qubit 2QCFA is a collapse-2QCFA. In particular, the result of Ambainis and Watrous [4], which showed that a single-qubit 2QCFA can recognize L_{pal} and L_{eq}, implies these languages are recognizable by collapse-2QCFA. We also note that our result [35], which showed that a 2QCFA can recognize many group word problems, always produced 2QCFA that were, in fact, collapse-2QCFA. This naturally raises the following question.

Open Problem 6.5. *Are collapse-2QCFA equivalent in power to (general) 2QCFA? That is to say, if some language L is recognized with bounded-error by a 2QCFA (with some finite number of classical and quantum states), is L also recognized with bounded-error by some collapse-2QCFA (with some, possibly larger, finite number of classical and quantum states)?*

Crucially, all of the collapse-2QCFA constructed to prove these results operate by using the finite-size quantum register of a 2QCFA to store an amount of classical information that grows (often quite quickly) with the length of the input (see [35] for a full discussion of this phenomenon). As Ambainis and Watrous noted, by Holevo's theorem [22], it is impossible to store more than b classical bits of information using a b-qubit quantum register, if one wishes to be able to perfectly reconstruct all stored information by performing (destructive) quantum measurements on the quantum register. These collapse-2QCFA do not violate Holevo's theorem, nor any of the other bounds on the manner in which the information stored in a quantum register may be accessed (see, for instance, [3, 32]), as any complete quantum measurement of a quantum register with k basis states (i.e., $\log k$ qubits), yields at most $\log k$ classical bits of information about the state of the quantum register before the measurement was performed, and after the measurement has been performed, the state of the quantum register has collapsed to a single pure state $|q\rangle$ (i.e., all other information has been destroyed).

As we have observed, the m-truncated transfer operators $N_{x,m}$ (and their non-truncated counterpart N_x), which describe the behavior of a (general) 2QCFA N when computing on a prefix $\#_L x$, are quantum channels. Recall that the distance metric on $\mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C)$ induced by the trace norm $\|\cdot\| : \mathrm{L}(\mathbb{C}^Q \otimes \mathbb{C}^C) \to \mathbb{R}_{\geq 0}$ is contractive under the application of any quantum channel. In particular, $\|N_{x,m}(Z) - N_{x,m}(Z')\|_1 \leq \|Z - Z'\|_1$, $\forall x \in \Sigma^*$, $\forall m \in \mathbb{N}$, $\forall Z, Z' \in \widehat{\mathrm{Den}}(\mathbb{C}^Q \otimes \mathbb{C}^C)$. This implies another restriction on the manner in which a (general) 2QCFA may access the information stored in its quantum register, as it shows that the classically controlled computation of a 2QCFA cannot perform quantum measurements and use the results of those quantum measurements to guide the later steps of its computation in such a way so as to increase the distance between a pair of starting density operators. We emphasize that this is both true for collapse-2QCFA (for a somewhat more obvious reason) as well as general 2QCFA.

References

[1] F. Ablayev and A. Gainutdinova, "On the lower bounds for one-way quantum automata," in *International Symposium on Mathematical Foundations of Computer Science*. Springer, 2000, pp. 132–140.

[2] A. Ambainis and N. Nahimovs, "Improved constructions of quantum automata," *Theoretical Computer Science*, vol. 410, no. 20, pp. 1916–1922, 2009.

[3] A. Ambainis, A. Nayak, A. Ta-Shma, and U. Vazirani, "Dense quantum coding and quantum finite automata," *Journal of the ACM (JACM)*, vol. 49, no. 4, pp. 496–511, 2002.

[4] A. Ambainis and J. Watrous, "Two-way finite automata with quantum and classical states," *Theoretical Computer Science*, vol. 287, no. 1, pp. 299–311, 2002.

[5] A. Ambainis and A. Yakaryılmaz, "Automata and quantum computing," *arXiv preprint arXiv:1507.01988*, 2015.

[6] A. V. Anisimov, "Group languages," *Cybernetics and Systems Analysis*, vol. 7, no. 4, pp. 594–601, 1971.

[7] ——, "Some algorithmic problems for groups and context-free languages," *Cybernetics and Systems Analysis*, vol. 8, no. 2, pp. 174–182, 1972.

[8] A. Brodsky and N. Pippenger, "Characterizations of 1-way quantum finite automata," *SIAM Journal on Computing*, vol. 31, no. 5, pp. 1456–1478, 2002.

[9] T. Brough, "Groups with poly-context-free word problem," *Groups Complexity Cryptology*, vol. 6, no. 1, pp. 9–29, 2014.

[10] T. Ceccherini-Silberstein, M. Coornaert, F. Fiorenzi, P. E. Schupp, and N. W. Touikan, "Mul-

tipass automata and group word problems," *Theoretical Computer Science*, vol. 600, pp. 19–33, 2015.

[11] M. J. Dunwoody, "The accessibility of finitely presented groups," *Inventiones mathematicae*, vol. 81, no. 3, pp. 449–457, 1985.

[12] C. Dwork and L. Stockmeyer, "A time complexity gap for two-way probabilistic finite-state automata," *SIAM Journal on Computing*, vol. 19, no. 6, pp. 1011–1023, 1990.

[13] ——, "Finite state verifiers i: The power of interaction," *Journal of the ACM (JACM)*, vol. 39, no. 4, pp. 800–828, 1992.

[14] R. Freivalds, "Probabilistic two-way machines," in *International Symposium on Mathematical Foundations of Computer Science*. Springer, 1981, pp. 33–45.

[15] A. G. Greenberg and A. Weiss, "A lower bound for probabilistic algorithms for finite state machines," *Journal of Computer and System Sciences*, vol. 33, no. 1, pp. 88–105, 1986.

[16] R. I. Grigorchuk, "Degrees of growth of finitely generated groups, and the theory of invariant means," *Mathematics of the USSR-Izvestiya*, vol. 25, no. 2, p. 259, 1985.

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

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

[19] A. W. Harrow, A. Hassidim, and S. Lloyd, "Quantum algorithm for linear systems of equations," *Physical review letters*, vol. 103, no. 15, p. 150502, 2009.

[20] F. C. Hennie, "One-tape, off-line turing machine computations," *Information and Control*, vol. 8, no. 6, pp. 553–578, 1965.

[21] T. Herbst, "On a subclass of context-free groups," *RAIRO-Theoretical Informatics and Applications-Informatique Théorique et Applications*, vol. 25, no. 3, pp. 255–272, 1991.

[22] A. S. Holevo, "Bounds for the quantity of information transmitted by a quantum communication channel," *Problemy Peredachi Informatsii*, vol. 9, no. 3, pp. 3–11, 1973.

[23] D. F. Holt, M. D. Owens, and R. M. Thomas, "Groups and semigroups with a one-counter word problem," *Journal of the Australian Mathematical Society*, vol. 85, no. 2, pp. 197–209, 2008.

[24] D. F. Holt, S. Rees, C. E. Röver, and R. M. Thomas, "Groups with context-free co-word problem," *Journal of the London Mathematical Society*, vol. 71, no. 3, pp. 643–657, 2005.

[25] E. Kowalski, *An introduction to the representation theory of groups*. American Mathematical Society, 2014, vol. 155.

[26] R. J. Lipton and Y. Zalcstein, "Word problems solvable in logspace," *Journal of the ACM (JACM)*, vol. 24, no. 3, pp. 522–526, 1977.

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

[28] G. W. Mackey *et al.*, "Unitary representations of group extensions. i," *Acta Mathematica*, vol. 99, pp. 265–311, 1958.

[29] C. Moore and J. P. Crutchfield, "Quantum automata and quantum grammars," *Theoretical Computer Science*, vol. 237, no. 1-2, pp. 275–306, 2000.

[30] D. E. Muller and P. E. Schupp, "Groups, the theory of ends, and context-free languages," *Journal of Computer and System Sciences*, vol. 26, no. 3, pp. 295–310, 1983.

[31] ——, "The theory of ends, pushdown automata, and second-order logic," *Theoretical Computer Science*, vol. 37, pp. 51–75, 1985.

[32] A. Nayak, "Optimal lower bounds for quantum automata and random access codes," in *40th Annual Symposium on Foundations of Computer Science (Cat. No. 99CB37039)*. IEEE, 1999, pp. 369–376.

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

[34] M. O. Rabin and D. Scott, "Finite automata and their decision problems," *IBM journal of research and development*, vol. 3, no. 2, pp. 114–125, 1959.
[35] Z. Remscrim, "The power of a single qubit: Two-way quantum/classical finite automata and the word problem for linear groups," in *Electronic Colloquium on Computational Complexity (ECCC)*, no. 107, 2019, pp. 1–50.
[36] A. Say and A. Yakaryilmaz, "Magic coins are useful for small-space quantum machines," *Quantum Information & Computation*, vol. 17, no. 11-12, pp. 1027–1043, 2017.
[37] Y. Shalom and T. Tao, "A finitary version of gromovs polynomial growth theorem," *Geometric and Functional Analysis*, vol. 20, no. 6, pp. 1502–1547, 2010.
[38] P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *Proceedings 35th annual symposium on foundations of computer science*. Ieee, 1994, pp. 124–134.
[39] A. Thom, "Convergent sequences in discrete groups," *Canadian Mathematical Bulletin*, vol. 56, no. 2, pp. 424–433, 2013.
[40] J. Tits, "Free subgroups in linear groups," *Journal of Algebra*, vol. 20, no. 2, pp. 250–270, 1972.
[41] J. Watrous, *The theory of quantum information*. Cambridge University Press, 2018.
[42] B. A. Wehrfritz, "Infinite linear groups: an account of the group-theoretic properties of infinite groups of matrices," 1973.
[43] J. A. Wolf *et al.*, "Growth of finitely generated solvable groups and curvature of riemannian manifolds," *Journal of differential Geometry*, vol. 2, no. 4, pp. 421–446, 1968.
[44] A. Yakaryilmaz and A. C. Say, "Succinctness of two-way probabilistic and quantum finite automata," *Discrete Mathematics and Theoretical Computer Science*, vol. 12, no. 4, pp. 19–40, 2010.
[45] S. Zheng, D. Qiu, L. Li, and J. Gruska, "One-way finite automata with quantum and classical states," in *Languages Alive*. Springer, 2012, pp. 273–290.

ECCC ISSN 1433-8092
https://eccc.weizmann.ac.il