Electronic Colloquium on Computational Complexity, Report No. 107 (2019)

The Power of a Single Qubit: Two-way Quantum/Classical Finite Automata and the Word Problem for Linear Groups

Zachary Remscrim
Department of Mathematics
MIT

Abstract

The two-way quantum/classical finite automaton (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 a single qubit, can recognize, with one-sided 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.

We further demonstrate the power of 2QCFA by showing that they can recognize the word problems of a broad class of groups. In particular, we first restrict our attention to 2QCFA that: (1) have a single qubit, (2) recognize their language with one-sided bounded-error, and (3) have transition amplitudes which are algebraic numbers. We show that such 2QCFA can recognize the word problem of any finitely-generated virtually abelian group in expected polynomial time, as well as the word problem of a large class of linear groups in expected exponential time. This latter class includes all groups whose word problem is a context-free language as well as all groups whose word problem is known to be the intersection of finitely many context-free languages. As a corollary, we obtain a direct improvement on the original Ambainis and Watrous result by showing that L_{eq} can be recognized by a 2QCFA with better parameters.

We also consider those word problems which a 2QCFA can recognize with one-sided *unbounded*-error, and show that this class includes the word problem of more exotic groups such as the free product of any finite collection of finitely-generated free abelian groups. As a corollary of this result, we demonstrate that a new class of group word problems are co-stochastic languages. Lastly, we exhibit analogous results for 2QCFA with any finite number of qubits or with more general transition amplitudes, as well as results for other classic QFA models.

1 Introduction

1.1 Background

The theory of quantum computation has made amazing strides in the last several decades. Landmark results, like Shor's polynomial time quantum algorithms for integer factorization and computing the discrete logarithm [41], Grover's algorithm for unstructured search [18], and the linear system solver of Harrow, Hassadim, and Lloyd [19], have provided remarkable examples of natural problems for which quantum computers seem to have an advantage over their classical counterparts. These theoretical breakthroughs have provided strong motivation to construct quantum computers. However, while significant advancements have been made, even the most advanced experimental quantum computers that exist today are still quite limited, and are certainly not capable of implementing, on a large scale, algorithms designed for general quantum Turing machines. This naturally motivates the study of more restricted models of quantum computation.

In this paper, our goal is to understand the computational power of a small number of qubits, especially the power of a single qubit. To that end, we study the two-way quantum/classical finite

ISSN 1433-8092

automaton (2QCFA) introduced by Ambainis and Watrous [2]. Informally, a 2QCFA is a two-way deterministic finite automaton (2DFA) that has been augmented with a quantum register of constant size, i.e., a constant number of qubits. The quantum part of the machine is extremely limited; however, the model is surprisingly powerful. In particular, Ambainis and Watrous [2] showed that a 2QCFA, using only one qubit, 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. As L_{eq} and L_{pal} are both non-regular, this clearly demonstrated that 2QCFA are a more powerful model than ordinary 2DFA, which recognize precisely the regular languages [37]. Moreover, as it is known that two-way probabilistic finite automata (2PFA) can recognize L_{eq} with bounded-error in exponential time [15], but not in subexponential time [17], and cannot recognize L_{pal} with bounded-error in any time bound [14], this result also demonstrated the superiority of 2QCFA over 2PFA.

The 2QCFA model is a particular special case of the quantum finite automata (QFA) model. Many (significantly) different variants of QFA have been defined (see for example [7, 11, 21, 25, 29, 32, 35, 48], see the excellent survey [3] for a complete history), which led to many seemingly contradictory claims about the power of QFA, ranging from results that they can only recognize a certain proper subset of the regular languages, with the particular subset varying with the model, to being able to recognize precisely the regular languages, to being able to recognize wide assortments of different extremely powerful classes of languages. However, what truly sets the 2QCFA model apart is that it is realistic, in two distinct senses. Firstly, the 2QCFA constructed by Ambainis and Watrous [2] that recognize L_{eq} and L_{pal} operate under the same limitations that constrain real physical (small) quantum computers: as already noted, these 2QCFA recognize these languages with bounded-error, and have a quantum part that consists of only a single qubit; additionally, the transition amplitudes of these 2QCFA are all efficiently-computable numbers, no unreasonable assumptions are made concerning the precision of these values, and only a particularly simple type of quantum measurement is allowed. Secondly, the 2QCFA model does not impose any additional restrictions which are not physically motivated: for example, the one-way QFA defined by Kondacs and Watrous [25] permitted quantum measurements to be performed at any time; however, the result of these measurements could only be used to determine whether or not the QFA halts at any particular step, whereas a 2QCFA is free to use the result of a quantum measurement in any manner. Therefore, the 2QCFA model provides the ideal setting in which to explore the power of a small number of qubits.

We investigate the ability of a 2QCFA to recognize the word problem of a group. Informally, the word problem for a group G involves determining if the product (i.e., combination under the group operation) of a finite collection of group elements $g_1, \dots, g_k \in G$ is equal to the identity element of G. Word problems for various classes of groups have a rich and well-studied history in computational complexity theory, as there are many striking relationships between certain algebraic properties of a group G and the computational complexity of its word problem W_G. This is demonstrated by many classic results, such as the result of Anisimov [4], which showed that W_G is a regular language (REG) if and only if G is finite, or the result of Muller and Schupp [30] (see also [12]) which showed that W_G is a context-free language (CFL) if and only if W_G is a deterministic context-free language (DCFL) if an only if G is a finitely-generated virtually free group. This latter result is especially remarkable, as $\mathsf{DCFL} \subsetneq \mathsf{CFL}$, but there is no group whose word problem witnesses this separation. The landmark result of Lipton and Zalcstein [27], which showed that the word problem of any finitely-generated linear group over a field of characteristic 0 is decidable in deterministic logspace (L), has a similarly intriguing consequence. Namely, while it remains an open question whether or not CFL is contained in L, there are certainly no groups G for which $W_G \in \mathsf{CFL}$ but $W_G \notin \mathsf{L}$.

For a quantum model, such as the 2QCFA, word problems for groups are a particularly natural class of languages to study. In particular, there are several results which show that certain (generally significantly more powerful) QFA variants can recognize the word problems of particular classes of groups (see, for instance, [7, 47, 48]). Moreover, there are many other results concerning the ability of QFA to recognize certain languages that are extremely closely related to group word problems; in fact, the languages L_{eq} and L_{pal} considered by Ambainis and Watrous [2] are each closely related to the word problem of a particular group.

Fundamentally, the requirement, imposed by laws of quantum mechanics, that the quantum state of a 2QCFA must evolve unitarily forces the computation of a 2QCFA to have a certain algebraic structure. Similarly, the algebraic properties of a particular group G impose a corresponding algebraic structure on its word problem W_G. For certain classes of groups, the algebraic structure of W_G is extremely compatible with the algebraic structure of the computation of a 2QCFA; for other classes of groups, these two algebraic structures are in extreme opposition.

In this paper, we show that there is a broad class of groups for which these algebraic structures are quite compatible, which enables us to produce 2QCFA that recognize these word problems. We emphasize that, while substantially more powerful QFA variants have already been shown to recognize many of these word problems, our results hold for a significantly more limited, and physically realistic model. We discuss, in detail, the various variants of QFA, and known results concerning their ability to recognize group word problems. Additionally, as a corollary of our results concerning group word problems, we obtain a direct improvement on the Ambainis and Watrous result [2] concerning the parameters of a 2QCFA that recognizes L_{eq}. In an upcoming paper, we explore those group word problems whose algebraic structure is quite incompatible with that of a 2QCFA, as these problems are natural candidates for demonstrating an upper bound on the power of 2QCFA and other related quantum models.

1.2 Statement of the Main Results

We show that, for many groups G, the corresponding word problem W_G is recognized by a 2QCFA with “good” parameters. In order to state these results, we must make use of some terminology and notation concerning 2QCFA, the word problem of a group, and various classes of groups whose word problems are of complexity theoretic interest. A full description of the 2QCFA model can be found in Section 2.1; the definition of the word problem, as well as additional group theory background, including the definitions of the various classes of groups discussed in this section, can be found in Section 2.2. The following definition establishes some useful notation that will allow us to succinctly describe the parameters of a 2QCFA. We use $\mathbb{R}_{>0}$ to denote the positive real numbers.

Definition 1.1. We say that a language L is recognized by a $[\epsilon, \tau, d, \mathbb{T}]$-2QCFA, where $\epsilon \in \mathbb{R}_{>0}$, $\tau : \mathbb{N} \to \mathbb{N}$, $d \in \mathbb{N}$, and $\mathbb{T} \subseteq \mathbb{C}$, if there is a 2QCFA A for which the following holds.

(a) A accepts all $w \in L$ with certainty and rejects all $w \notin L$ with probability at least $1 - \epsilon$.

(b) On an input string w of length n, A runs in expected time $O(\tau(n))$.

(c) A has d quantum basis states.

(d) All transition amplitudes of A belong to $\mathbb{T}$.

The focus on the transition amplitudes of a 2QCFA warrants a bit of additional justification, as while it is standard to limit the transition amplitudes of a Turing machine in this way, it is common for finite automata to be defined without any such limitation. Firstly, this restriction is physically motivated as it is not reasonable to assume that the transition amplitudes of a physical computational device can have infinite precision; in particular, any model which relies on such infinite precision, by, for example, making use of transition amplitudes that are non-computable numbers, is not a physically realizable model. Secondly, while this constraint of physical reasonableness would apply just as well to other finite automata models, applying such a constraint would often be superfluous; for example, the class of languages recognized with bounded-error and in expected time $2^{n^{o(1)}}$ by a 2PFA with no restriction at all on its transition amplitudes is precisely the regular languages [13]. On the other hand, the power of the 2QCFA model is quite sensitive to the choice of transition amplitudes; a 2QCFA with non-computable transition amplitudes can recognize, with bounded-error and in expected polynomial time, undecidable languages [38], whereas a 2QCFA that is limited to algebraic number transition amplitudes can only recognize languages in $\mathsf{P} \cap \mathsf{L}^2$, even if permitted unbounded-error and exponential time [46]. In particular, the algebraic numbers $\overline{\mathbb{Q}}$ are arguably the "standard" choice for the permitted transition amplitudes of a quantum Turing machine (QTM). It is desirable for the definition of 2QCFA to be consistent with that of QTMs as such consistency makes it more likely that techniques developed for a 2QCFA could be applied to QTMs; therefore, $\overline{\mathbb{Q}}$ is the the natural choice for the permitted transition amplitudes of a 2QCFA. For every group G for which we can construct a 2QCFA that recognizes the word problem W_G with bounded-error, we can construct a 2QCFA, whose transition amplitudes lie in $\overline{\mathbb{Q}}$, that recognizes W_G with bounded-error. In fact, our key results remain true even if we restrict the 2QCFA to have transition amplitudes that are Gaussian rationals (i.e., numbers of the form $a + bi$, where $a, b \in \mathbb{Q}$), but for ease of exposition, and because algebraic numbers arise "naturally" in quantum computation (e.g., the entries of the 2×2 Hadamard matrix are of the form $\pm 1/\sqrt{2}$), we do not pursue this further specialization here. On the other hand, we do consider the impact of allowing transition amplitudes in the slightly broader class $\widetilde{\mathbb{C}} = \overline{\mathbb{Q}} \cup \{e^{\pi i r} | r \in (\overline{\mathbb{Q}} \cap \mathbb{R})\}$.

We begin with a simple motivating example. For a finite alphabet Σ, a letter $\sigma \in \Sigma$, and a word $w \in \Sigma^*$, let $\#(w, \sigma)$ denote the number of appearances of σ in w. Then the word problem for the group $\mathbb{Z}$ (the integers, where the group operation is addition) is the language $W_{\mathbb{Z}} = \{w \in \{a, b\}^* | \#(w, a) = \#(w, b)\}$. This language is closely related to the language $L_{eq} = \{a^m b^m | m \in \mathbb{N}\}$; in particular, $L_{eq} = (a^* b^*) \cap W_{\mathbb{Z}}$. More generally, for any positive integer k, the word problem for the group $\mathbb{Z}^k$ (the direct sum of k copies of $\mathbb{Z}$) is the language $W_{\mathbb{Z}^k} = \{w \in \{a_1, b_1, \ldots, a_k, b_k\}^* | \#(w, a_i) = \#(w, b_i), \forall i\}$. Using our terminology, Ambainis and Watrous [2] showed that, for any $\epsilon \in \mathbb{R}_{>0}$, L_{eq} is recognizable by a $[\epsilon, n^4, 2, \widetilde{\mathbb{C}}]$-2QCFA[1]; we note that the same method would easily imply the same result for $W_{\mathbb{Z}}$, and could be further adapted to produce the analogous result for $W_{\mathbb{Z}^k}$.

Our first main theorem generalizes and improves upon the above mentioned result of Ambainis and Watrous [2] in several ways. Recall that the finitely-generated virtually abelian groups are precisely those groups that have a finite-index subgroup that is isomorphic to $\mathbb{Z}^k$, for some $k \in \mathbb{N}$, where $\mathbb{Z}^0 = \{1\}$ is the trivial group (i.e., the group with one element); of course, this class of groups (properly) contains all $\mathbb{Z}^k$. We show the following.

Theorem 1.2. *There is a (universal, effectively computable) constant $C \in \mathbb{R}_{>0}$ for which the following holds. Suppose G is a finitely-generated virtually abelian group. For any $\epsilon \in \mathbb{R}_{>0}$, W_G is recognized by a $[\epsilon, n^3, 2, \widetilde{\mathbb{C}}]$-2QCFA, as well as by a $[\epsilon, n^C, 2, \overline{\mathbb{Q}}]$-2QCFA.*

By the above observation about the relationship between $W_{\mathbb{Z}}$ and L_{eq}, the following corollary is immediate.

Corollary 1.2.1. *For any $\epsilon \in \mathbb{R}_{>0}$, L_{eq} is recognized by a $[\epsilon, n^3, 2, \widetilde{\mathbb{C}}]$-2QCFA, as well as by a $[\epsilon, n^C, 2, \overline{\mathbb{Q}}]$-2QCFA, for the above constant $C \in \mathbb{R}_{>0}$.*

Note that the above corollary provides an improvement upon the result of Ambainis and Watrous [2] concerning the parameters of a 2QCFA for L_{eq} in two distinct senses. Firstly, using the same set of permissible transition amplitudes, our result has a better expected running time. Secondly, our

[1] Strictly speaking, Ambainis and Watrous [2] considered a slightly different but "equivalent" set of transition amplitudes; this equivalence will be clarified in Section 5.3

result allows for the construction of a 2QCFA for L_{eq} that is limited to having algebraic transition amplitudes, which still runs in expected polynomial time.

Let REG denote the regular languages (languages recognized by a deterministic finite automaton), CFL denote the context-free languages (languages recognized by non-deterministic pushdown automata), OCL denote the one-counter languages (languages recognized by non-deterministic pushdown automata where the stack alphabet is limited to a single symbol) and $\mathsf{poly{-}CFL}$ (resp. $\mathsf{poly{-}OCL}$) denote the intersection of finitely many context-free (resp. one-counter) languages. As $W_G \in \mathsf{poly{-}OCL}$ if and only if G is a finitely-generated virtually abelian group [23], the following corollary is also immediate.

Corollary 1.2.2. *If $W_G \in \mathsf{poly{-}OCL}$, then for any $\epsilon \in \mathbb{R}_{>0}$, W_G is recognized by a $[\epsilon, n^3, 2, \widetilde{\mathbb{C}}]$-2QCFA, as well as by a $[\epsilon, n^C, 2, \overline{\mathbb{Q}}]$-2QCFA, for the above constant $C \in \mathbb{R}_{>0}$.*

In other words, a 2QCFA, with a single qubit, can recognize any of these "multi-counter" languages in expected polynomial time, where the value of k only affects the constant hidden by the $O(\cdot)$ notation, but not the degree of the polynomial specifying the expected run-time. Moreover, as $W_G \in \mathsf{poly{-}OCL} \cap \mathsf{CFL}$ if and only if $W_G \in \mathsf{OCL}$ if and only if G is a finitely-generated virtually cyclic group [20, 23], the above corollary exhibits a wide class of non-context-free languages that are recognizable by a 2QCFA in polynomial time: the word problems W_G for any group G that is virtually rank-k free abelian, for some $k \geq 2$ (recall that $\mathbb{Z}^k$ is the rank-k free abelian group, and, for any finitely-generated virtually abelian group G, there is a unique k such that G has a finite-index subgroup isomorphic to $\mathbb{Z}^k$).

Remark. Interestingly, the limiting factor on the run-time of the $[\epsilon, n^3, 2, \widetilde{\mathbb{C}}]$-2QCFA for any of the above word problems (or L_{eq}) is not the difficulty of distinguishing strings in the language from strings not in the language, but is instead due to the apparent difficulty of using a 2QCFA to produce a Boolean random variable with a particular (rather extreme) bias. In particular, we make use of the procedure (from [2]) that allows a 2QCFA, on an input of size n, to generate a Boolean value that is 1 with probability essentially n^{-1}, in time $O(n^2)$. If, for some $\delta \in (0,1)$, it were possible for a 2QCFA to produce a Boolean variable that has value 1 with probability $n^{-\delta}$ in time $t(n)$, our technique would immediately yield a $[\epsilon, (n + t(n))n^{\delta}, 2, \widetilde{\mathbb{C}}]$-2QCFA.

We next consider groups that are built from finite-rank free groups using certain operations. First, consider the word problem W_{F_2} of the rank 2 free group F_2, which is the language over the alphabet $\Sigma = \{a, a^{-1}, b, b^{-1}\}$ defined as follows. For a word $x \in \Sigma^*$, a *matched-pair* in x is a 2-element contiguous subword of x of the form aa^{-1}, $a^{-1}a$, bb^{-1}, or $b^{-1}b$. Let $\widehat{x} \in \Sigma^*$ denote the word obtained from $x \in \Sigma^*$ by repeatedly deleting matched-pairs in x (i.e., replacing a matched-pair by the empty-string) until there are no matched-pairs remaining. Then W_{F_2} consists of precisely those words x such that $\widehat{x}$ is the empty-string. Notice that W_{F_2} is closely related to the language $L_{pal} = \{w \in \{a,b\}^* | w \text{ is a palindrome}\}$. In particular, let $\Gamma = \{a, b\} \subseteq \Sigma$ denote the alphabet over which L_{pal} is defined, and, for $w = w_1 \cdots w_n \in \Gamma^*$, where each $w_i \in \Gamma$, let $\overline{w} = w_1^{-1} \cdots w_n^{-1}$. Then, for any $w \in \Gamma^*$, $w \in L_{pal} \Leftrightarrow w\overline{w} \in W_{F_2}$. Ambainis and Watrous [2] showed that, $\forall \epsilon \in \mathbb{R}_{>0}$, there is a $D \in \mathbb{R}_{\geq 1}$ such that L_{pal} is recognized by a $[\epsilon, D^n, 2, \overline{\mathbb{Q}}]$-2QCFA, and the same method would show the same result for W_{F_2}.

More generally, we use F_k to denote the free group of rank k, for any $k \in \mathbb{N}$; in particular, F_0 is the trivial group, F_1 is the group $\mathbb{Z}$, and, for any $k \geq 2$, F_k is non-abelian. We show that the same result holds for any group built from finite-rank free groups F_k, using certain operations.

Theorem 1.3. *Suppose G is virtually a finitely-generated subgroup of a direct product of finitely many finite-rank free groups. For any $\epsilon \in \mathbb{R}_{>0}$, there is an effectively computable constant $D \in \mathbb{R}_{\geq 1}$ such that W_G is recognized by a $[\epsilon, D^n, 2, \overline{\mathbb{Q}}]$-2QCFA.*

Let $\widehat{\Pi}_2$ denote the class of groups that satisfy the hypothesis of the preceding theorem (we will explain this choice of notation shortly). Notice that all finitely-generated virtually free groups belong to $\widehat{\Pi}_2$. As $W_G \in \mathsf{CFL}$ if and only if G is a finitely-generated virtually free group [12, 30], we then immediately have the following.

Corollary 1.3.1. *If $W_G \in \mathsf{CFL}$, then for any $\epsilon \in \mathbb{R}_{>0}$, there is an effectively computable constant $D \in \mathbb{R}_{\geq 1}$ such that W_G is recognized by a $[\epsilon, D^n, 2, \overline{\mathbb{Q}}]$-2QCFA.*

Next, consider the group $H = F_2 \times F_2$, the direct product of two copies of F_2; clearly, $H \in \widehat{\Pi}_2$. However, H is not virtually free, and so $W_H \notin \mathsf{CFL}$. Moreover, H is not virtually abelian, and so $W_H \notin \mathsf{poly{-}OCL}$. This immediately implies the following corollary.

Corollary 1.3.2. *There is a group H, for which $W_H \notin \mathsf{CFL} \cup \mathsf{poly{-}OCL}$, where for any $\epsilon \in \mathbb{R}_{>0}$, there is an effectively computable constant $D \in \mathbb{R}_{\geq 1}$ such that W_H is recognized by a $[\epsilon, D^n, 2, \overline{\mathbb{Q}}]$-2QCFA.*

In fact, an even stronger version of the above corollary is true. Consider the classic example, due to Stallings [42], of a subgroup K of H which is finitely-generated, but not finitely-presented; namely, K is the kernel of the homomorphism $\pi : H = F_2 \times F_2 \to \mathbb{Z}$, where π takes each free generator of each copy of F_2 to a single generator of $\mathbb{Z}$. All groups G for which $W_G \in \mathsf{CFL} \cup \mathsf{poly{-}OCL}$ are finitely-presented, which immediately implies $W_K \notin \mathsf{CFL} \cup \mathsf{poly{-}OCL}$. Of course, we have $K \in \widehat{\Pi}_2$, which immediately implies the following corollary.

Corollary 1.3.3. *There is a group K, which is finitely-generated, but not finitely-presented (which, in particular, implies $W_K \notin \mathsf{CFL} \cup \mathsf{poly{-}OCL}$), where for any $\epsilon \in \mathbb{R}_{>0}$, there is an effectively computable constant $D \in \mathbb{R}_{\geq 1}$ such that W_K is recognized by a $[\epsilon, D^n, 2, \overline{\mathbb{Q}}]$-2QCFA.*

Remark. One could, equivalently, define $\widehat{\Pi}_2$ as the closure of the set of finite-rank free groups under the operations of (finite) direct product, passing to a finitely-generated subgroup, and passing to a finite-index overgroup. For every group $G \in \widehat{\Pi}_2$, it is known that $W_G \in \mathsf{poly{-}CFL}$ [8]. Moreover, it is conjectured that $\widehat{\Pi}_2$ is precisely the class of groups whose word problem is in $\mathsf{poly{-}CFL}$ [8] (cf. [10]).

We next consider a broader class of groups. Let $\mathbb{N}_{\geq 1}$ denote the positive natural numbers, and let $Z(H)$ denote the center of a group H. For $d \in \mathbb{N}_{\geq 1}$, let $\mathrm{U}(d, \overline{\mathbb{Q}})$ denote the group of $d \times d$ unitary matrices whose entries are algebraic numbers, where the group operation is the usual matrix multiplication; furthermore, let $\mathrm{PU}(d, \overline{\mathbb{Q}}) = \mathrm{U}(d, \overline{\mathbb{Q}})/Z(\mathrm{U}(d, \overline{\mathbb{Q}}))$ denote the d-dimensional projective unitary group with algebraic number entries. For $k \in \mathbb{N}_{\geq 1}$, let $(\mathrm{PU}(d, \overline{\mathbb{Q}}))^k$ denote the direct product of k copies of $\mathrm{PU}(d, \overline{\mathbb{Q}})$.

Theorem 1.4. *Suppose G is a finitely-generated group that has a finite-index subgroup that is isomorphic to a subgroup of $(\mathrm{PU}(d, \overline{\mathbb{Q}}))^k$, for some $d \in \mathbb{N}_{\geq 2}, k \in \mathbb{N}_{\geq 1}$. Then for any $\epsilon \in \mathbb{R}_{>0}$, there is an effectively computable constant $D \in \mathbb{R}_{\geq 1}$ such that W_G is recognized by a $[\epsilon, D^n, d, \overline{\mathbb{Q}}]$-2QCFA.*

The following corollary highlights a certain significant special case; see Section 2.4 for the notation and terminology from representation theory used in the statement of this corollary.

Corollary 1.4.1. *Suppose G is a finitely-generated group which has a faithful representation $\rho : G \to \mathrm{U}(d, \overline{\mathbb{Q}})$. Then ρ has a (unique, up to isomorphism of representations) set of irreducible subrepresentations $\{\rho_j : G \to \mathrm{U}(d_j, \overline{\mathbb{Q}})\}_{j=1}^m$ such that $\rho \cong \rho_1 \oplus \cdots \oplus \rho_m$. Let $d_{\max} = \max_j d_j$. Then for any $\epsilon \in \mathbb{R}_{>0}$, there is an effectively computable constant $D \in \mathbb{R}_{\geq 1}$ such that W_G is recognized by a $[\epsilon, D^n, d_{\max} + 1, \overline{\mathbb{Q}}]$-2QCFA.*

Remark. There is some overlap between the various classes of groups considered in each of the above theorems. For example, consider the group $\mathbb{Z}$. For any $\epsilon \in \mathbb{R}_{>0}$, Theorem 1.2 implies that $W_{\mathbb{Z}}$ is recognized by a $[\epsilon, n^C, 2, \overline{\mathbb{Q}}]$-2QCFA; as $\mathbb{Z} = F_1$, Theorem 1.3 also guarantees the existence of a 2QCFA that recognizes $W_{\mathbb{Z}}$, but with the weaker parameters $[\epsilon, D^n, 2, \overline{\mathbb{Q}}]$. Similarly, the class of groups to which Theorem 1.4 applies contains all groups to which the earlier theorems apply.

In order to state our final main result, as well as to provide appropriate context for the results listed above, we next define a certain collection of important classes of groups. We write $H_1 * H_2$ to denote the free product of groups H_1 and H_2. We define the classes of groups Σ_j and Π_j, for each $j \in \mathbb{N}$, inductively. First $\Sigma_0 = \Pi_0 = \{\mathbb{Z}, \{1\}\}$ (i.e., both of these classes consist of the two groups $\mathbb{Z}$ and the trivial group $\{1\}$). For each $j \in \mathbb{N}_{\geq 1}$, we define Σ_j as the collection of all groups G such that $\exists H_1, \ldots, H_t \in \Pi_{j-1}$, for some $t \in \mathbb{N}_{\geq 1}$, such that $G \cong H_1 * \cdots * H_t$; analogously, we define Π_j as the collection of all groups G such that $\exists H_1, \ldots, H_t \in \Sigma_{j-1}$, for some $t \in \mathbb{N}_{\geq 1}$, such that $G \cong H_1 \times \cdots \times H_t$. For example, $\Pi_1 = \{\mathbb{Z}^k | k \in \mathbb{N}\}$ is the finitely-generated free abelian groups and $\Sigma_1 = \{F_k | k \in \mathbb{N}\}$ is the finitely-generated free groups. Note that all groups in all Σ_j and Π_j are finitely-generated, and also note that the Σ_j and Π_j form a hierarchy in the obvious way. These groups form a particularly important subclass of a particularly important class of groups: the right-angled Artin groups (RAAGs). We further define $\widehat{\Sigma}_j$ (resp. $\widehat{\Pi}_j$) as the set of all finitely-generated groups that are virtually isomorphic to a (necessarily finitely-generated) subgroup of some group in Σ_j (resp. Π_j), which also form a hierarchy in the obvious way.

In particular, $\widehat{\Pi}_1$ is the class of finitely-generated virtually abelian groups, which is (precisely) the class of groups for which Theorem 1.2 demonstrates the existence of a polynomial-time 2QCFA for the corresponding word problem. Furthermore, $\widehat{\Sigma}_1$ is the class of finitely-generated virtually free groups (i.e., those groups with context-free word problem), and the class $\widehat{\Pi}_2 \supsetneq \widehat{\Sigma}_1$ is (precisely) the class of groups for which Theorem 1.3 demonstrates the existence of an exponential-time 2QCFA for the corresponding word problem. We next consider the class $\widehat{\Pi}_3$. While the relationship of this class to the class of groups to which Theorem 1.4 applies is unclear to us, we can show that the word problem of any group in this class can be recognized by a 2QCFA with one-sided *unbounded*-error. In the following we say a language L is recognized by A with *negative one-sided unbounded-error* if, $\forall w \in L$, $\Pr[A \text{ accepts } w] = 1$ and, $\forall w \notin L$, $\Pr[A \text{ rejects } w] > 0$. We then say that A is an unbounded-error $[\tau, d, \mathbb{T}]$-2QCFA for the language L if a modified version of the the conditions Definition 1.1 are satisfied, where Definition 1.1(a) is replaced by the condition that A recognizes L with negative one-sided unbounded-error and Definition 1.1(b) is replaced by the condition that A runs in time $O(\tau(n))$ (rather than *expected* time $O(\tau(n))$).

Theorem 1.5. *For any group $G \in \widehat{\Pi}_3$, the word problem W_G is recognized by an unbounded-error $[n, 2, \widetilde{\mathbb{C}}]$-2QCFA.*

Remark. Consider $G = \mathbb{Z} * \mathbb{Z}^2 \in \Sigma_2 \subsetneq \widehat{\Pi}_3$. It is conjectured that $W_G \notin \mathsf{poly{-}CFL}$ [8] and $W_G \notin \mathsf{coCFL}$ [24] (where coCFL denotes the class of languages whose complements are in CFL).

While our focus in this paper is certainly the 2QCFA model, with the further restriction to 2QCFA whose transition amplitudes are all "simple" numbers, we consider other QFA variants, beginning with 2QCFA with no restrictions on their transition amplitudes.

Theorem 1.6. *For any finitely-generated group G that has a finite-index subgroup that is isomorphic to a subgroup of $(\mathrm{PU}(d))^k$, for some $d \in \mathbb{N}_{\geq 2}, k \in \mathbb{N}_{\geq 1}$, the word problem W_G is recognized with negative one-sided unbounded-error by a 2QCFA with d quantum basis states in time $O(n)$.*

We also consider the measure-once one-way quantum finite automata (MO-1QFA) defined by Moore and Crutchfield [29]. We write $\mathcal{D}$ for the class of all groups to which any of the previously stated theorems apply.

Theorem 1.7. *For any group $G \in \mathcal{D}$, there is a MO-1QFA that recognizes W_G with negative one-sided unbounded-error.*

We say a machine M recognizes a language L with *strict cut-point* $\lambda \in \mathbb{R}$ if $\forall w \in L$, we have $\Pr[M \text{ accepts } w] > \lambda$ and $\forall w \notin L$, $\Pr[M \text{ accepts } w] \leq \lambda$. We write S to denote the stochastic languages, the class of languages L for which there is a PFA P that recognizes L for some strict cut-point; we then write coS to denote the class of languages whose complements are in S. Of course, if a MO-1QFA M recognizes a language L with negative one-sided unbounded-error, then there is a MO-1QFA $\overline{M}$ that recognizes its complement $\overline{L}$ with strict cut-point 0 (i.e., with *positive* one-sided unbounded-error), where $\overline{M}$ is obtained from M by swapping accepting and rejecting states. By [7, Theorem 3.6], any language accepted by a MO-1QFA with any strict cut-point is stochastic, which immediately implies the following corollary.

Corollary 1.7.1. *For any group $G \in \mathcal{D}$, $W_G \in \mathsf{coS}$.*

Remark. For many $G \in \mathcal{D}$, the fact that $W_G \in \mathsf{coS}$ was already known; in particular, $W_{\mathbb{Z}} \in \mathsf{coS}$ is a classic result of Rabin [36], and the fact that $W_{F_k} \in \mathsf{coS}$ was shown by Brodsky and Pippenger [7], from which one can conclude (using standard arguments from computational group theory, see for instance [30]) that $\forall G \in \widehat{\Pi}_2$, $W_G \in \mathsf{coS}$ However, for $G \in \mathcal{D} \setminus \widehat{\Pi}_2$, this result appears to be new.

1.3 Outline of the Paper

The core idea of our approach to solving the word problem of a particular group G is to construct what we have chosen to call a *distinguishing family of representations* (DFR) for G. Informally, given a group G, with identity element 1_G, a DFR for G is a "small" collection of "small" unitary representations of G that, collectively, "strongly" separate 1_G from all other $g \in G$. A (unitary) representation of a (topological) group G is a continuous homomorphism $\rho : G \to \mathrm{U}(\mathcal{H})$, where $\mathcal{H}$ is a Hilbert space, and $\mathrm{U}(\mathcal{H})$ is the group of unitary operators on $\mathcal{H}$. The Gel'fand-Raikov theorem states that the elements of any locally compact group G are separated by its unitary representations, i.e., $\forall g_1, g_2 \in G$ there is some $\mathcal{H}$ and some $\rho : G \to \mathrm{U}(\mathcal{H})$ such that $\rho(g_1) \neq \rho(g_2)$. For certain groups, stronger statements can be made; in particular, one calls a group maximally almost periodic if the previous condition still holds when $\mathcal{H}$ is restricted to be finite-dimensional. The notion of a DFR for a group G is a generalization of this idea, as it is a collection of a (constant) small number of unitary representations of G, all of which are into a Hilbert space of (constant) small dimension, such that, for any $g \in G$ other than 1_G, there is some representation ρ in the collection for which $\rho(g)$ is "far from" $\rho(1_G)$, relative to the "size" of g. This approach of recognizing W_G by computing with appropriately chosen representations of G formed the basis of the landmark result of Lipton and Zalcstein [27] which showed that $W_G \in \mathsf{L}$ when G is a finitely-generated linear group over a field of characteristic 0; however, the constraints of quantum computing will require us to make many modifications to their approach.

In Section 3, we formally define DFRs, and construct DFRs for many groups. Our constructions of DFRs crucially rely on certain results concerning Diophantine approximation, both in the traditional setting of approximation of real numbers by rational numbers, as well as in a certain non-commutative generalization, originally proposed by Gamburd, Jakobson, and Sarnak [16]; we study Diophantine approximation in Section 3.2.

In Section 4, we use a DFR for a group G to construct a 2QCFA that recognizes W_G, where the parameters of the DFR directly determine the parameters of the 2QCFA. In fact, we show that it is also possible to use a DFR for G to produce a 2QCFA that recognizes W_H, for certain groups H related to G; this observation will allow us, in certain cases, to improve the parameters of the 2QCFA.

In Section 5.1, we compare our results to existing results regarding both the classical and quantum computational complexity of the word problem. A key feature of the 2QCFA that we construct is that they operate by storing an amount of information that grows (quite quickly) with the size of the input using only a quantum register of constant size. In Section 5.2, we discuss why this is possible, and consider further implications of this extreme compression of information. In Section 5.3 we consider the various variants of QFA that have been defined.

2 Preliminaries

2.1 Quantum Computation and the 2QCFA

In this section, we briefly recall the fundamentals of quantum computation, after which we present the definition of the Ambainis and Watrous [2] two-way finite automaton with quantum and classical states (2QCFA). For additional background on quantum computation, see, for instance, [33].

The most natural way of understanding quantum computation is as a generalization of probabilistic computation. Given a probabilistic system consisting of k states, for some finite k, the particular state of that system, at some particular point in time, is given by a probability distribution over the k states. Such a probability distribution can be described by a vector $p = (p_1, \dots, p_k)$, where p_j denotes the probability that the system is in state j. As p is a probability distribution, each p_j must be a non-negative real number, and one must have $\sum_j p_j = 1$, i.e., p must be a non-negative real vector with L_1 norm 1.

Similarly, one may consider a quantum system with k basis states, where the overall state of the system at any particular time is given by a *superposition* of the k basis states. Formally, fix an orthonormal basis $|q_1\rangle, \dots, |q_k\rangle$ of $\mathbb{C}^k$, where here and throughout the paper we use the standard Bra-Ket notation. A superposition is a linear combination $\sum_j \alpha_j |q_j\rangle$, where each $\alpha_j \in \mathbb{C}$ and $\sum_j |\alpha_j|^2 = 1$. In other words, a superposition is simply an element $|\psi\rangle \in \mathbb{C}^k$ of L_2 norm 1.

Let $\mathrm{U}(k)$ denote the group of $k \times k$ unitary matrices, i.e., those matrices that preserve the norm of all vectors in $\mathbb{C}^k$. Given a quantum system currently in the superposition $|\psi\rangle$, one may apply a transformation $T \in \mathrm{U}(k)$ to the system, after which the system is in the superposition $T|\psi\rangle$. One may also perform a *quantum measurement* on a quantum system. In particular, if $B = \{B_0, \dots, B_l\}$ is a partition of $\{1, \dots, k\}$, then measuring a quantum system that is in the superposition $|\psi\rangle = \sum_j \alpha_j |q_j\rangle$ with respect to B gives the result r, with probability $p_r := \sum_{j \in B_r} |\alpha_j|^2$, for each $r \in \{0, \dots, l\}$; additionally, if the result of the measurement is r, then the state of the system *collapses* to the superposition $\frac{1}{\sqrt{p_r}} \sum_{j \in B_r} \alpha_j |q_j\rangle$. We emphasize that performing a quantum measurement on a quantum system changes the state of that system.

We now define a 2QCFA, essentially following the original definition in [2]. Informally, a 2QCFA is a two-way deterministic finite automaton that has been augmented with a finite size quantum register. Formally, a 2QCFA A is given by an 8-tuple,

$$A = \{Q, C, \Sigma, \delta, q_1, c_1, C_{acc}, C_{rej}\},$$

where $Q = \{q_1, \dots, q_k\}$ is the finite set of quantum basis states, C is the finite set of classical states, Σ is a finite alphabet, δ is the transition function, q_1 is the quantum start state, c_1 is the classical start state, and $C_{acc} \subseteq C$ and $C_{rej} \subseteq C$ (where $C_{acc} \cap C_{rej} = \emptyset$) are the accepting and rejecting states. We define the tape alphabet $\Gamma := \Sigma \cup \{\#_L, \#_R\}$ where the two distinct symbols $\#_L, \#_R \notin \Sigma$ will be used to denote, respectively, a left and right end-marker. The *quantum register* of A is the quantum part of A, i.e., the quantum system with basis states Q, which, at any point in the computation is in some superposition $|\psi\rangle = \sum_j \alpha_j |q_j\rangle$.

Each step of the computation of the 2QCFA A involves either performing a unitary transformation or a quantum measurement on its quantum register, updating the classical state, and possibly moving the tape head left or right. This behavior is encoded in the transition function δ. For each $(c, \gamma) \in (C \setminus (C_{acc} \cup C_{rej})) \times \Gamma$, $\delta(c, \gamma)$ specifies the behavior of A when it is in the classical state c and the tape head currently points to a tape alphabet symbol γ. There are two forms that $\delta(c, \gamma)$ may take, depending on whether it encodes a unitary transformation or a quantum measurement. In the first case, $\delta(c, \gamma)$ is a triple (T, c', h) where $T \in \mathrm{U}(|Q|)$ is a unitary transformation to be performed on the quantum register, $c' \in C$ is the new classical state, and $h \in \{-1, 0, 1\}$ specifies whether the tape head is to move left, stay put, or move right, respectively. In the second case, $\delta(c, \gamma)$ is a pair (B, f), where $B = \{B_0, \dots, B_l\}$ is a partition of $\{1, \dots, k\}$ (i.e., B is a family of sets specifying a quantum measurement), and $f : \{0, \dots, l\} \to C \times \{-1, 0, 1\}$ specifies the mapping from the result of that quantum measurement to the evolution of the classical part of the machine, where, if the result of the quantum measurement is r, and $f(r) = (c', h)$, then $c' \in C$ is the new classical state and $h \in \{-1, 0, 1\}$ specifies the movement of the tape head.

The computation of A on an input $w \in \Sigma^*$ is then defined as follows. If w has length n, then the tape will be of size $n + 2$ and contain the string $\#_L w \#_R$. Initially, the classical state is c_1, the quantum part of the machine is in the superposition $|q_1\rangle$, and the tape head points to the leftmost tape cell (which contains the left end-marker $\#_L$). At each step of the computation, if the classical state is currently c and the tape head is pointing to symbol γ, the machine behaves as specified by $\delta(c, \gamma)$. If, at some point in the computation, A enters an accepting state $c \in C_{acc}$ (resp. rejecting state $c \in C_{rej}$) then it immediately halts and accepts (resp. rejects) the input w. For any $w \in \Sigma^*$, we write $p_{acc}(w)$ (resp. $p_{rej}(w)$) for the probability that A will accept (resp. reject) the input w. We then say that A recognizes a language $L \subseteq \Sigma^*$ with negative one-sided bounded-error $\epsilon \in \mathbb{R}_{>0}$ if the following three conditions hold:

1. $\forall w \in \Sigma^*$, $p_{acc}(w) + p_{rej}(w) = 1$
2. $\forall w \in L$, $p_{acc}(w) = 1$
3. $\forall w \notin L$, $p_{rej}(w) \geq 1 - \epsilon$.

For a 2QCFA A, let $\mathcal{T}$ denote the set of all unitary matrices T that correspond to a unitary transformation that A may perform on its quantum register, i.e., if $A = \{Q, C, \Sigma, \delta, q_1, c_1, C_{acc}, C_{rej}\}$, $\mathcal{T}$ consists of precisely those $T \in \mathrm{U}(|Q|)$ for which $\exists (c, \gamma) \in (C \setminus (C_{acc} \cup C_{rej})) \times \Gamma$ such that $\delta(c, \gamma) = (T, \cdot, \cdot)$. The *transition amplitudes* of A are the set of numbers $\mathbb{T}$ that appear as an entry of some matrix $T \in \mathcal{T}$. Let $\widetilde{\mathbb{C}} = \overline{\mathbb{Q}} \cup \{e^{\pi i r} | r \in (\overline{\mathbb{Q}} \cap \mathbb{R})\}$. We restrict our attention to 2QCFA whose transition amplitudes $\mathbb{T} \subseteq \widetilde{\mathbb{C}}$, though we will be most interested in the special case in which $\mathbb{T} \subseteq \overline{\mathbb{Q}} \subsetneq \widetilde{\mathbb{C}}$. We note that $\widetilde{\mathbb{C}}$ is, essentially, the class of transition amplitudes used by the 2QCFA A_{AW} of Ambainis and Watrous [2] to recognize L_{eq}. In Section 5.3, we observe that A_{AW} is equivalent to a 2QCFA A'_{AW} (in the sense that A_{AW} and A'_{AW} have precisely the same probability of accepting or rejecting any input string) where the transition amplitudes of A'_{AW} are restricted to $\widetilde{\mathbb{C}}$.

2.2 Group Theory and the Word Problem of a Group

Informally, the *word problem* for a group G is the following question: given a finite sequence of elements $g_1, \dots, g_n \in G$, is $g_1 \cdots g_n$, their combination using the group operation, equal to the identity element of G? In this section, we formalize this problem.

We begin by introducing some terminology and notation from group theory; for more extensive background, see, for instance, [28]. For a group G, we write $S \subseteq G$ if the set S is a subset of G and $H \leq G$ if the group H is a subgroup of G. For $S \subseteq G$, let $\langle S \rangle$ denote the subgroup of G generated by S and let $\langle S^G \rangle$ denote the normal closure of S in G. We say that S is a *generating set* for the group G if $S \subseteq G$ and $G = \langle S \rangle$. Let $F(S)$ denote the free group on the set S. For sets S and R, where $R \subseteq F(S)$, we say G has *presentation* $\langle S|R \rangle$ if $G \cong F(S)/\langle R^{F(S)} \rangle$, in which case we write $G = \langle S|R \rangle$. If a group G has presentation $G = \langle S|R \rangle$, then S (or more precisely the image of S in G under the natural map) is a generating set for G, and if G has generating set S, then it has (many) presentations of the form $G = \langle S|R \rangle$. We say that G is *finitely-generated* if it has a generating set S that is finite, and we say that G is *finitely-presented* if it has a presentation $G = \langle S|R \rangle$ with both S and R finite. For $S \subseteq G$, let $S^{-1} = \{s^{-1} | s \in S\}$, let $\Sigma = S \cup S^{-1}$, and let Σ^* denote the free monoid on Σ. When Σ is finite (equivalently, S is finite), we say that Σ is an *alphabet*, $w \in \Sigma^*$ is a *word* over the alphabet Σ, and $L \subseteq \Sigma^*$ is a *language* over the alphabet Σ. For a group $G = \langle S|R \rangle$, we have a natural surjective homomorphism $\phi : \Sigma^* \to G$ which takes each element of Σ^* to the element of G that it represents. Lastly, we use 1_G to denote the identity element of G. We now define the *word problem* for a group.

Definition 2.1. Suppose $G = \langle S|R \rangle$, where S is finite. Then the word problem of G with respect to the presentation $\langle S|R \rangle$ is the language $W_{G=\langle S|R \rangle} = \{w \in \Sigma^* | \phi(w) = 1_G\}$ consisting of all words w over the finite alphabet $\Sigma = S \cup S^{-1}$ that represent the identity element in G. Solving the word problem for $G = \langle S|R \rangle$ means deciding membership in this language.

Note that while any group G will have infinitely many presentations, and the above definition of the word problem of a group G does depend on the particular presentation used, the particular choice of (finite) generating set or set of relators will have no bearing on the membership of this language in any of the complexity classes considered in this paper. To clarify this, let $\mathcal{L}$ denote a class of languages. We say that $\mathcal{L}$ is *closed under inverse homomorphism* if, for all pairs of finite alphabets Σ_1, Σ_2, all monoid homomorphisms $\tau : \Sigma_1^* \to \Sigma_2^*$, and every language $W \in \mathcal{L}$ over the alphabet Σ_2, we have $\tau^{-1}(W) = \{v \in \Sigma_1^* | \tau(v) \in W\} \in \mathcal{L}$. Clearly, for any class of languages $\mathcal{L}$ closed under inverse homomorphism, if $\langle S_1|R_1 \rangle$ and $\langle S_2|R_2 \rangle$, with S_1 and S_2 finite, are both presentations of the same group G, then $W_{\langle S_1|R_1 \rangle} \in \mathcal{L} \Leftrightarrow W_{\langle S_2|R_2 \rangle} \in \mathcal{L}$. As all complexity classes considered in this paper are closed under inverse homomorphism, we can reasonably speak about the complexity of the word problem for any finitely-generated group G, without reference to a particular presentation of G, and can then simply write W_G for the word problem of G.

We conclude this section with a bit of additional terminology and notation from group theory needed in later parts of the paper. We say that a group F is *free* if $F \cong F(S)$ for some set S, and we define the *rank* of F to be the cardinality of S. The rank of a free group is well-defined as $F(S) \cong F(T)$ if and only if S and T have the same cardinality. As a consequence of the same observation, there is a unique (up to isomorphism) free group of rank k, for any $k \in \mathbb{N}$, which allows us to speak about *the* free group of rank k, which we denote by $F_k := F(\{1, \dots, k\})$. We follow the convention that $F_0 = F(\emptyset) = \{1\}$, the trivial group. For a group G and a subgroup $H \leq G$, we use $[G : H]$ to denote the index of H in G; if $[G : H]$ is finite, then we say that H is a *finite index* subgroup of G. We say a group is *finite* if it is finite as a set, and *countable* if it is at most countably infinite as a set. Notice that any finitely-generated group is necessarily countable. We say a group is *cyclic* if it has a generating set consisting of a single element, *abelian* if the group operation is commutative, and *linear* if it is isomorphic to a subgroup of $\mathrm{GL}(n, k)$, where $\mathrm{GL}(n, k)$ denotes the group of $n \times n$ invertible matrices, over some field k, where the group operation is given by matrix multiplication. For any property $\mathcal{P}$ (abelian, free, etc.), we say a group is *virtually* $\mathcal{P}$ if it contains a finite-index subgroup that has $\mathcal{P}$.

2.3 Cayley Graphs

Consider a group $G = \langle S|R \rangle$, where, the generating set S is finite. As before write $\Sigma = S \cup S^{-1}$ for the union of the generators S and their inverses, Σ^* for the free monoid on Σ, and $\phi : \Sigma^* \to G$ for the natural surjection that takes $w \in \Sigma^*$ to the element of G that it represents. The (right) *Cayley graph* of G with the respect to the (symmetric) generating set Σ, which we denote $\Gamma(G, \Sigma)$, is the directed, labeled graph which has vertices G, and a directed edge from g to $g\sigma$ that is labeled σ, for each $g \in G$ and $\sigma \in \Sigma$ where $\sigma \neq 1_G$. A word $w = w_1 \cdots w_n \in \Sigma^*$, with each $w_i \in \Sigma$, specifies a path p_w in $\Gamma(G, \Sigma)$ which starts at the vertex 1_G and, on the i^{th} step, follows the edge labeled w_i. Notice that $\phi(w) = 1_G$ if and only if the path p_w terminates at the vertex 1_G.

Next, notice that the Cayley graph $\Gamma(G, \Sigma)$ depends on the particular presentation $G = \langle S|R \rangle$, insofar as different choices of the generating set S generally lead to different Σ, and, ultimately, to distinct (non-isomorphic) graphs. However, if $\langle S'|R' \rangle$ is another presentation of G, where S' is also finite, then, while it is the case that $\Gamma(G, \Sigma)$ and $\Gamma(G, \Sigma')$ will generally be non-isomorphic graphs, they will "look the same from far away."

To formalize this notion, recall that a *metric space* is a set X equipped with a map $d : X \times X \to \mathbb{R}_{\geq 0}$, where $\mathbb{R}_{\geq 0}$ denotes the non-negative real numbers, such that, $\forall x_1, x_2, x_3 \in X$, the following three properties are satisfied: $d(x_1, x_2) = 0 \Leftrightarrow x_1 = x_2$, $d(x_1, x_2) = d(x_2, x_1)$, and $d(x_1, x_3) \leq d(x_1, x_2) + d(x_2, x_3)$. Given two metric spaces (X, d) and (X', d'), we say that a function $f : X \to X'$ is a *bilipschitz equivalence* between them if f is a bijection and $\exists C \in \mathbb{R}_{>0}$ such that, $\forall x_1, x_2 \in X$,

$$\frac{1}{C} d(x_1, x_2) \leq d'(f(x_1), f(x_2)) \leq C d(x_1, x_2).$$

For G a group and S a generating set of G, the *word metric on G relative to the generating set S*, which we denote by d_S, is the usual distance metric on the Cayley graph $\Gamma(G, \Sigma)$, i.e, for any $g_1, g_2 \in G$, $d_S(g_1, g_2)$ is the length of the shortest path in $\Gamma(G, \Sigma)$ from g_1 to g_2. Equivalently, for any $g_1, g_2 \in G$, $d_S(g_1, g_2)$ is the smallest $n \in \mathbb{N}$ for which there exists a sequence $\sigma_1, \ldots, \sigma_n \in \Sigma$ such that $g_2 = g_1 \sigma_1 \cdots \sigma_n$. Notice that (G, d_S) is a metric space. It is straightforward to see that, if S and S' are two finite generating sets of G, then the identity map on G is a bilipschitz equivalence between (G, d_S) and (G, d'_S), where the constant C can be straightforwardly bounded by considering d_S and d'_S (see, for instance, [28, Proposition 5.2.4]).

When S is clear from context, we will often simply write d in place of d_S. We also define $l_S(g)$, the *length of $g \in G$ relative to the generating set S*, by $l_S(g) := d_S(1, g)$, i.e., $l_S(g)$ is the shortest length of an expression for g in the generators S and their inverses. Similarly, we write l in place of l_S, when S is clear from context.

2.4 Representation Theory Background

In this section, we state certain basic definitions and elementary results from representation theory that will be needed in the remainder of this paper. While the material in this section can be found in essentially any textbook on the (linear) representation theory of (infinite) groups, we essentially follow [26], though we deliberately avoid stating results in their full generality, to simplify the exposition as much as possible.

A *representation* of a group G over a field k is a pair (ρ, V_ρ), where V_ρ is a vector space over k, $\mathrm{GL}(V_\rho)$ denotes the group of invertible k-linear maps on V_ρ, and $\rho : G \to \mathrm{GL}(V_\rho)$ is a group homomorphism. If, furthermore, $\rho : G \to \mathrm{GL}(V_\rho)$ is injective, then we say that (ρ, V_ρ) is a *faithful* representation of G. For $v \in V_\rho$ and $g \in G$, we denote the image of v under the map $\rho(g)$ by $\rho(g)v$. This notation is used to emphasize that a representation (ρ, V_ρ) of a group G is equivalent to a linear (left) action of G on V_ρ, given by $g \cdot v = \rho(g)v$, for $g \in G$ and $v \in V_\rho$. By standard slight abuse of notation, we will often say that ρ is a representation of G, when V_ρ is clear from the context. We say that V_ρ is the *representation space* of the representation ρ. The *dimension* of a representation ρ is the (vector space) dimension of its representation space V_ρ. If ρ is a finite-dimensional representation, one may identify (non-canonically) $\mathrm{GL}(V_\rho)$ with $\mathrm{GL}(n, k)$, the group of $n \times n$ invertible matrices over the field k, by picking a particular basis of V. Such an identification allows the image of $g \in G$ under the map $\rho : G \to \mathrm{GL}(n, k)$, to be explicitly encoded in a matrix, which will be useful for computation.

In this paper, we concern ourselves, almost exclusively, with *finite-dimensional unitary representations of finitely-generated groups*, which, for such a group G, are representations of the form $\rho : G \to \mathrm{U}(n)$, for some $n \in \mathbb{N}_{\geq 1}$, where $\mathrm{U}(n)$ denotes the group of $n \times n$ unitary matrices, and for which the corresponding representation space $V_\rho = \mathbb{C}^n$. Throughout the paper, *a representation* will always mean a finite-dimensional unitary representation of a finitely-generated group, unless we explicitly note otherwise.

Remark. Generally, one defines a unitary representation of a topological group G as a representation $\rho : G \to \mathrm{U}(\mathcal{H})$, where $\mathcal{H}$ is some complex Hilbert space and $\mathrm{U}(\mathcal{H})$ denotes the group of all unitary continuous linear operators on $\mathcal{H}$, such that ρ is strongly continuous, i.e., for every $v \in \mathcal{H}$, the mapping

$G \to \mathcal{H}$ given by $g \mapsto \rho(g)v$ is continuous. However, any finitely-generated group is countable, and the natural topology for any countable group is the discrete topology, for which the continuity condition is trivially satisfied. Moreover, as previously observed, finite-dimensional representations can be concretely realized as representations into matrix groups. Therefore, this is equivalent to our simpler definition.

Consider two representations $\rho_1 : G \to \mathrm{U}(n_1)$ and $\rho_2 : G \to \mathrm{U}(n_2)$ of a group G. Let $\mathrm{Hom}_{\mathbb{C}}(n_1, n_2)$ denote the space of $\mathbb{C}$-linear maps (i.e., homomorphisms of $\mathbb{C}$ vector spaces) $\phi : \mathbb{C}^{n_1} \to \mathbb{C}^{n_2}$. A *homomorphism of representations* is a $\phi \in \mathrm{Hom}_{\mathbb{C}}(n_1, n_2)$ such that, $\forall g \in G, \forall v \in V_{\rho_1} = \mathbb{C}^{n_1}$, we have $\phi(\rho_1(g)v) = \rho_2(g)\phi(v)$. We use $\mathrm{Hom}_G(\rho_1, \rho_2)$ to denote the subspace of $\mathrm{Hom}_{\mathbb{C}}(n_1, n_2)$ consisting of all such ϕ. If there is some $\phi \in \mathrm{Hom}_G(\rho_1, \rho_2)$ that is bijective, we say that the representations ρ_1 and ρ_2 are *isomorphic*, which we denote by writing $\rho_1 \cong \rho_2$, and we call such a ϕ an *isomorphism of representations*. For an $n_1 \times n_1$ matrix A and a $n_2 \times n_2$ matrix B, we write $A \oplus B$ to denote the $(n_1 + n_2) \times (n_1 + n_2)$ block-diagonal matrix whose two diagonal blocks are given by A and B. The *direct sum of representations* ρ_1 and ρ_2 is the representation $\rho_1 \oplus \rho_2 : G \to \mathrm{U}(n_1 + n_2)$, where $(\rho_1 \oplus \rho_2)(g) = \rho_1(g) \oplus \rho_2(g)$, $\forall g \in G$.

For a representation $\rho : G \to \mathrm{U}(n)$, we say that a vector subspace V' of $V_\rho = \mathbb{C}^n$ is *stable* if $\forall g \in G, \forall v \in V'$, $\rho(g)v \in V'$. We say that the representation $\rho' : G \to \mathrm{U}(n')$ is a *subrepresentation* of ρ if there is a stable subspace V' of V_ρ, of dimension n', such that $\rho'(g)v = \rho(g)v$, $\forall g \in G, \forall v \in V'$. We say that ρ is *irreducible* if it has no non-trivial subrepresentations (i.e., the only stable subspaces of V_ρ are 0 and V_ρ itself). For any representation $\rho : G \to \mathrm{U}(n)$, there is a decomposition $\rho \cong \rho_1 \oplus \cdots \oplus \rho_m$, where the ρ_j are all irreducible subrepresentations; moreover, this decomposition is unique (up to permutation of the summands, and isomorphism of representations).

For a representation $\rho : G \to \mathrm{U}(n)$ of a group G, and a subgroup $H \leq G$, we define the *restricted representation* $\mathrm{Res}^G_H(\rho)$ to be the representation $\pi : H \to \mathrm{U}(n)$ of H, where $\pi(h) = \rho(h)$, $\forall h \in H \leq G$, i.e., this is simply the restriction of ρ to H. Next, we define a concept dual to the notion of restriction. Let $\pi : H \to \mathrm{U}(m)$ be a representation of H and let G be a finite-index overgroup of H, i.e., $H \leq G$ and $r := [G : H]$ is finite. The *induced representation* $\mathrm{Ind}^G_H(\pi)$ is the representation $\rho : G \to \mathrm{U}(mr)$, which is defined as follows. Let $T = \{g_1, \ldots, g_r\} \subseteq G$ denote a complete family of left coset representatives of H in G. Let S_r denote the symmetric group on r symbols. For each $g \in G$, let $\sigma_g \in S_r$ and $h_{g,j} \in H$ denote the (unique) elements such that, for each $j \in \{1, \ldots, r\}$, we have $gg_j = g_{\sigma_g(j)} h_{g,j}$. For each $g_j \in T$, let $g_j\mathbb{C}^m$ denote an isomorphic copy of the representation space $V_\pi = \mathbb{C}^m$. We then define V_ρ, the representation space of ρ, by $V_\rho = \bigoplus_{j=1}^r g_j\mathbb{C}^m \cong \mathbb{C}^{mr}$. To define ρ, we think of an element of V_ρ as being of the form $\sum_{j=1}^r g_j v_j$, where each $v_j \in V_\pi = \mathbb{C}^m$, and define $\rho : G \to \mathrm{U}(mr)$ such that $\forall g \in G$, $\rho(g) \sum_{j=1}^r g_j v_j = \sum_{j=1}^r g_{\sigma_g(j)} \pi(h_{g,j}) v_j$. Concretely, $\rho(g)$ is a block matrix, all of whose blocks are $m \times m$, and, in block-column j, the only non-zero block-row is $\sigma_g(j)$, and this block is given by $\pi(h_{g,j})$.

Remark. Induction and restriction, as defined above are dual in the following sense: If one lets Rep_G (resp. Rep_H) denotes, the category of representations of G (resp. H) over the field k, then Res^G_H : $\mathrm{Rep}_G \to \mathrm{Rep}_H$ and $\mathrm{Ind}^G_H : \mathrm{Rep}_H \to \mathrm{Rep}_G$ are functors and Ind^G_H is the left-adjoint of Res^G_H. We note that induction, as we have defined it, is more commonly called co-induction, and that one traditionally defines the induced representation such that induction is the right-adjoint of restriction. However, as we only consider the case when H is a finite index subgroup of G, the co-induced representation that we have defined and the induced representation that one normally defines are isomorphic. It will simply be more convenient, for our purposes, to use co-induction, though we will refer to it as induction.

Consider a representation $\rho : G \to \mathrm{U}(n)$. The *character* of ρ is the function $\chi_\rho : G \to \mathbb{C}$ given by $\chi_\rho(g) = \mathrm{Tr}(\rho(g))$, where $\mathrm{Tr}(\rho(g))$ denotes the trace of (the unitary matrix) $\rho(g)$. Let $I_d \in \mathrm{U}(d)$ denote the $d \times d$ identity matrix (i.e., the identity element of the group $\mathrm{U}(d)$), $Z(\mathrm{U}(d)) = \{e^{ir} I_d | r \in \mathbb{R}\}$ denote the center of $\mathrm{U}(d)$, $\mathrm{PU}(d) = \mathrm{U}(d)/Z(\mathrm{U}(d))$ denote the d-dimensional projective unitary group, and $\tau : \mathrm{U}(d) \to \mathrm{PU}(d)$ denote the canonical projection. Let $\mathrm{Pker}(\rho) = \{g \in G | \rho(g) \in Z(\mathrm{U}(d))\}$ denote the quasikernel of ρ; notice that $\mathrm{Pker}(\rho) = \ker(\tau \circ \rho)$, and $\ker(\rho) \leq \mathrm{Pker}(\rho) \leq G$. We say that a representation ρ of G is *projectively faithful* or simply *P-faithful* if $\mathrm{Pker}(\rho)$ is the trivial group (i.e., if only the identity element of G belongs to $\mathrm{Pker}(\rho)$). Notice that a P-faithful representation is necessarily a faithful representation. Furthermore, notice that, $\forall g \in G$, $|\chi_\rho(g)| \leq d$, and $|\chi_\rho(g)| = d \Leftrightarrow g \in \mathrm{Pker}(\rho)$. Lastly, we define a *projective unitary representation of a finitely-generated group G* to be a group homomorphism $\pi : G \to \mathrm{PU}(d)$. We will use the term *projective representation* to refer to such a representation.

3 Distinguishing Family of Representations

Our primary tool for constructing a 2QCFA for the word problem for a group G is a *distinguishing family of representations* (DFR) for the group G. Informally, a DFR for a group G is a "small" family of "small" unitary representations of G such that, for each $g \in G$ where $g \neq 1_G$, the family contains at least one representation which "strongly" separates g from 1_G. The following definition formalizes this, by introducing parameters to quantify the above fuzzy notions. In this definition, and in the remainder of the paper, we write $G_{\neq 1_G}$ for the subset $G \setminus \{1_G\}$ of G. Recall that $\widetilde{\mathbb{C}} = \overline{\mathbb{Q}} \cup \{e^{\pi i r} | r \in (\overline{\mathbb{Q}} \cap \mathbb{R})\}$ is the set of transition amplitudes that we permit for 2QCFA, in this paper. Let $M_d(\widetilde{\mathbb{C}})$ denote the set of $d \times d$ matrices with entries in $\widetilde{\mathbb{C}}$.

Definition 3.1. Consider a group $G = \langle S | R \rangle$, with S finite. Let $\Sigma = S \cup S^{-1}$ denote the corresponding symmetric generating set of G and let $l(g)$ denote the length of any $g \in G$ relative to this symmetric generating set (i.e., $l(g)$ is the distance from 1_G to g in the Cayley graph $\Gamma(G, \Sigma)$). Then for $k, d \in \mathbb{N}$, with $k \geq 1, d \geq 2$, and $\tau : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ a monotone non-increasing function, we define a $[k, d, \tau]$-*distinguishing family of representations*(DFR) for G, to be a set $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ where the following conditions hold.

(i) $\forall j \in \{1, \ldots, k\}$, $\rho_j : G \to \mathrm{U}(d)$ is a representation of G.

(ii) $\forall g \in G_{\neq 1_G}$, $\exists j \in \{1, \ldots, k\}$ such that $|\chi_{\rho_j}(g)| \leq d - \tau(l(g))$.

(iii) $\forall s \in S, \forall j \in \{1, \ldots, k\}$, $\exists Y_1, \ldots, Y_t \in \mathrm{U}(d) \cap M_d(\widetilde{\mathbb{C}})$, for some finite t, such that $\rho_j(s) = \prod_i Y_i$.

Remark. As was the case for the definition of the word problem, the definition of a DFR for a group G does depend on the particular choice of presentation for G. However, much as it was the case that any reasonable choice of presentation did not affect the computational complexity of the word problem, for the complexity classes considered in this paper, it is also the case that any reasonable choice of presentation will not affect the existence of a DFR, for any of the groups considered in this paper. Moreover, the requirements that all representations must have the same dimension, and that this dimension $d \geq 2$, are simply done for convenience and ease of notation; it will have no effect on the parameters of the 2QCFAs that are ultimately constructed.

3.1 Application to the Word Problem

We now discuss how a DFR for G will be used to solve the word problem for G. We begin by clarifying the sense in which a DFR strongly separates each $g \in G_{\neq 1_G}$ from 1_G. Suppose $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ is a $[k, d, \tau]$-DFR for G. As before, we write $I_d = 1_{\mathrm{U}(d)} \in \mathrm{U}(d)$ for the $d \times d$ identity matrix, $\ker(\rho_j) = \{g \in G | \rho_j(g) = I_d\}$ for the kernel of ρ_j, $Z(\mathrm{U}(d)) = \{e^{ir} I_d | r \in \mathbb{R}\}$ for the center of $\mathrm{U}(d)$ and $\mathrm{Pker}(\rho_j) = \{g \in G | \rho_j(g) = Z(\mathrm{U}(d))\}$ for the quasikernel of ρ_j. Clearly, $1_G \in \cap_j \mathrm{Pker}(\rho_j)$, as ρ_j is a group homomorphism and so must take 1_G to $1_{\mathrm{U}(d)} = I_d \in Z(\mathrm{U}(d))$. Notice ρ_j is not assumed to be

P-faithful or even faithful, and, in fact, it will often be desirable to have the ρ_j be non-faithful as it will lead to more efficient 2QCFAs. Therefore, there may be $g \in G_{\neq 1_G}$ for which, for certain j, we have $g \in \operatorname{Pker}(\rho_j)$. As previously observed, $g \in \operatorname{Pker}(\rho_j)$ exactly when $|\chi_{\rho_j}(g)| = d$, and so we may have $g \in G_{\neq 1_G}$ where for some (perhaps many) j, $|\chi_{\rho_j}(g)| = d$. However, the second defining property of a DFR guarantees not only that $\cap_j \operatorname{Pker}(\rho_j) = \{1_G\}$, but, much more strongly, that all $g \in G_{\neq 1}$ are "far from" being in $\cap_j \operatorname{Pker}(\rho_j)$ in that, for each $g \in G_{\neq 1}$, there is some j such that $|\chi_{\rho_j}(g)|$ is at distance at least $\tau(l(g))$ from having value d. The fundamental approach to solving the word problem for g is to test if $g \in \cap_j \operatorname{Pker}(\rho_j)$, where this can be done as any g is either in $\cap_j \operatorname{Pker}(\rho_j)$ or far from being in $\cap_j \operatorname{Pker}(\rho_j)$. The following proposition is then immediate, but we explicitly state it as it is a central notion in our quantum approach to the word problem.

Proposition 3.2. *Consider a group $G = \langle S|R\rangle$ with a $[k, d, \tau]$-DFR $\{\rho_1, \dots, \rho_k\}$. Then, $\forall g \in G$,*

$$g = 1_G \Leftrightarrow g \in \cap_j \operatorname{Pker}(\rho_j) \Leftrightarrow |\chi_{\rho_1}(g)| = \cdots = |\chi_{\rho_k}(g)| = d.$$

Furthermore,

$$g \in G_{\neq 1_G} \Leftrightarrow \exists j \in \{1, \dots, k\} \text{ such that } |\chi_{\rho_j}(g)| \leq d - \tau(l(g)).$$

In particular, $\rho_1 \oplus \cdots \oplus \rho_k : G \to \mathrm{U}(kd)$ is a faithful representation of G.

In the following definition, we establish some terminology that will better allow us to describe particular types of DFR. We first need a bit of notation. Let $\mathrm{U}(d, \overline{\mathbb{Q}})$ denote the subgroup of $\mathrm{U}(d)$ consisting of matrices whose entries are algebraic numbers, let $\mathbb{C}^*$ denote the multiplicative group of the field $\mathbb{C}$, and let $S_1 = \{e^{ir} | r \in \mathbb{R}\} \leq \mathbb{C}^*$ denote the circle group. For any $H \leq \mathbb{C}^*$, let $S_1(H) = S_1 \cap H$, and let $\mathrm{T}(d, H)$ denote the group of all diagonal matrices D where each diagonal entry $D_{jj} \in S_1(H)$. Notice that, $\forall H \leq \mathbb{C}^*$, $\mathrm{T}(d, H) \leq \mathrm{U}(d)$; moreover, $\mathrm{T}(d, \overline{\mathbb{Q}}) \leq \mathrm{U}(d, \overline{\mathbb{Q}})$. We also define $\mathrm{T}(d) = \mathrm{T}(d, \mathbb{C}^*)$. For a homomorphism $\rho_j : G \to \mathrm{U}(d)$, let $\rho_j(G)$ denote the image of G under ρ_j, and notice that $\rho_j(G) \leq \mathrm{U}(d)$.

Definition 3.3. Suppose $\mathcal{F} = \{\rho_1, \dots, \rho_k\}$ is a $[k, d, \tau]$-DFR for a group G.

(i) If, $\forall j \in \{1, \dots, k\}$, $\rho_j(G) \leq \mathrm{U}(d, \overline{\mathbb{Q}}) \leq \mathrm{U}(d)$, we say $\mathcal{F}$ is an *algebraic* DFR.

(ii) If, for some $H \leq \mathbb{C}^*$, we have that, $\forall j \in \{1, \dots, k\}$, $\rho_j(G) \leq \mathrm{T}(d, H) \leq \mathrm{U}(d)$, we say $\mathcal{F}$ is a *H-diagonal* DFR.

Notice that $\overline{\mathbb{Q}} \leq \mathbb{C}^*$, and, moreover, that a $\overline{\mathbb{Q}}$-diagonal DFR is also an algebraic DFR; we call such a DFR an *algebraic diagonal* DFR. Let $\widetilde{E} = \{e^{\pi i r} | r \in (\overline{\mathbb{Q}} \cap \mathbb{R})\}$, and note that $\widetilde{\mathbb{C}} = \overline{\mathbb{Q}} \cup \widetilde{E}$ and $\widetilde{E} \leq \mathbb{C}^*$. In the following, we will restrict our attention to algebraic DFRs and $\widetilde{E}$-diagonal DFRs; and so from this point forward, we will only use the term "DFR" to refer to one of these special cases. Using a DFR for a group G, it will be possible to construct a 2QCFA for W_G, the word problem of G, where the parameters of the DFR will strongly impact the parameters of the resulting 2QCFA. In particular, as will be shown in Section 4, a $[k, d, \tau]$-DFR for G can be used to produce a 2QCFA for W_G which requires only d quantum states, $k + c$ classical states (for a universal constant $c > 0$), and has expected running time approximately $O(\tau(n))^{-1})$. Moreover, if the DFR is algebraic, all transition amplitudes of the resulting 2QCFA will be algebraic numbers; whereas a $\widetilde{E}$-diagonal DFR will yield a 2QCFA whose transition amplitudes are in $\widetilde{\mathbb{C}}$. For all groups for which we can construct DFRs, we can in fact construct algebraic DFRs; however, in a single important special case we can construct a non-algebraic DFR for a group that has better parameters than our best algebraic DFR for that group. As we wish to explore the trade-off between the permitted complexity of the transition amplitudes of a 2QCFA and the performance of that 2QCFA, we consider both algebraic DFRs and not-necessarily-algebraic DFRs.

Notice that if a group G has a diagonal DFR, then G must be a finitely-generated abelian group. To see this, suppose $\mathcal{F} = \{\rho_1, \dots, \rho_k\}$ is a H-diagonal $[k, d, \tau]$-DFR of G, for some $H \leq \mathbb{C}^*$. As noted in Proposition 3.2, $\rho_1 \oplus \cdots \oplus \rho_k : G \to \mathrm{T}(kd, H) \leq \mathrm{T}(kd)$ is a faithful representation of G, and so we have $G \leq \mathrm{T}(kd)$. Of course, $\mathrm{T}(kd)$ is an abelian group, and so any subgroup of $\mathrm{T}(kd)$ is also abelian; in particular, G is abelian. Moreover, the definition of a DFR requires that G have presentation $\langle S|R\rangle$, with S finite, and so G is finitely-generated. In the other direction, we will show that any finitely-generated abelian group G has a diagonal algebraic $[1, 2, C_1 n^{-C_2}]$-DFR, and a $\widetilde{E}$-diagonal $[1, 2, C_3 n^{-\delta}]$-DFR, for any $\delta \in \mathbb{R}_{>0}$ and certain constants $C_1, C_2, C_3 \in \mathbb{R}_{>0}$.

Remark. Additionally, notice that, for any $N \in \mathrm{U}(d) \cap M_d(\widetilde{\mathbb{C}})$, there is a d-dimensional permutation matrix P_N such that $P_N N P_N^{-1} \in \mathrm{U}(d_1, \overline{\mathbb{Q}}) \times \mathrm{T}(d_2, \widetilde{E})$, for some $d_1, d_2 \in \mathbb{N}_{\geq 1}$, where $d_1 + d_2 = d$. Here, we are thinking of elements $M \in \mathrm{U}(d_1, \overline{\mathbb{Q}}) \times \mathrm{T}(d_2, \widetilde{E})$ concretely as $d \times d$ block-diagonal matrices, with an upper-left $d_1 \times d_1$ diagonal block given by some $M' \in \mathrm{U}(d_1, \overline{\mathbb{Q}})$ and a lower-right $d_2 \times d_2$ diagonal block given by some $M'' \in \mathrm{T}(d_2, \widetilde{E})$, i.e., $M = M' \oplus M''$, where here $\oplus$ denotes the direct sum of matrices; therefore, we have $M = \widehat{M'}\widehat{M''}$, where $\widehat{M'} = M' \oplus I_{d_2} \in \mathrm{U}(d, \overline{\mathbb{Q}})$ and $\widehat{M''} = I_{d_1} \oplus M'' \in \mathrm{T}(d, \widetilde{E})$. Applying this to $M_N := P_N N P_N^{-1} \in \mathrm{U}(d_1, \overline{\mathbb{Q}}) \times \mathrm{T}(d_2, \widetilde{E})$, we can write $M_N = \widehat{M'_N}\widehat{M''_N}$ and therefore $N = P_N^{-1}\widehat{M'_N}\widehat{M''_N}P_N$, where $P_N, P_N^{-1}, \widehat{M'_N} \in \mathrm{U}(d, \overline{\mathbb{Q}})$ and $\widehat{M''_N} \in \mathrm{T}(d, \widetilde{E})$. Therefore, the condition expressed in Definition 3.1(iii) for a DFR $\mathcal{F} = \{\rho_1, \dots, \rho_k\}$ of a group $G = \langle S|R\rangle$, is equivalent to the statement that $\forall s \in S, \forall j \in \{1, \dots, k\}$, $\exists Y_1, \dots, Y_t \in \mathrm{U}(d, \overline{\mathbb{Q}}) \cup \mathrm{T}(d, \widetilde{E})$, such that $\rho_j(s) = \prod_i Y_i$; moreover, as S is a generating set for G, the preceding statement holds $\forall s \in S$ precisely when it holds $\forall g \in G$.

It will also be shown, in Section 4, that it is possible to solve the word problem for G using a DFR for a finite-index subgroup $H \leq G$; the resulting 2QCFA for W_G will have the same parameters as the 2QCFA for W_H, except for an increase in the number of classical states. For many groups, this will allow the construction of a 2QCFA for the word problem with smaller quantum part, though larger classical part. As quantum states are, arguable, more "expensive" than classical states, this is a desirable trade-off, which motivates the following generalization of a DFR. Recall that, for some property $\mathcal{P}$ (e.g., free, abelian, etc.) a group G is said to be *virtually-$\mathcal{P}$* if it contains a finite-index subgroup H that has $\mathcal{P}$.

Definition 3.4. We say that a finitely-generated group G *virtually* has a $[k, d, \tau]$-DFR, if there is some $H \leq G$, with $[G : H]$ finite, such that H has a $[k, d, \tau]$-DFR.

The goal is then to show that a wide collection of groups virtually have $[k, d, \tau]$-DFRs with good parameters, with a preference for algebraic DFRs.

3.2 Diophantine Approximation

Our constructions of DFRs rely crucially on certain results concerning Diophantine approximation. We begin by establishing some notation that will be used throughout this section, as well as in the remainder of the paper. We write $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$, and $\overline{\mathbb{Q}}$ to denote, respectively, the integers, rational numbers, real numbers, complex numbers, and algebraic numbers. We also write $\mathbb{R}_{>0}$ to denote the positive real numbers, $\mathbb{R}_{\geq 1}$ to denote the real numbers that are at least 1, $\mathbb{Z}_{\neq 0}$ to denote the non-zero integers, etc. For $\alpha \in \mathbb{C}$, we denote the magnitude of α by $|\alpha|$. For $\alpha \in \mathbb{R}$, $\|\alpha\|$ denotes the distance between α and its nearest integer, i.e., $\|\alpha\| = \min_{m \in \mathbb{Z}}|\alpha - m|$. If the value of a particular constant C depends on numbers α, β, γ, we write $C = C(\alpha, \beta, \gamma)$.

Most fundamentally, the Diophantine approximation question asks how well a particular real number α can be approximated by rational numbers. Of course, as $\mathbb{Q}$ is dense in $\mathbb{R}$, one can choose $\frac{p}{q} \in \mathbb{Q}$ so as to make the quantity $|\alpha - \frac{p}{q}|$ arbitrarily small; for this reason, one considers $\frac{p}{q}$ to be a "good" approximation to α only when $|\alpha - \frac{p}{q}|$ is small compared to a suitable function of q. One then considers

α to be poorly approximated by rationals if, for some "small" constant $d \in \mathbb{R}_{\geq 2}$, there is a constant $C = C(\alpha, d) \in \mathbb{R}_{>0}$ such that, $\forall (p, q) \in \mathbb{Z} \times \mathbb{Z}_{\neq 0}$, we have $|\alpha - \frac{p}{q}| \geq C|q|^{-d}$, where the smallness of d determines just how poorly approximable α is. Notice that

$$\left|\alpha - \frac{p}{q}\right| = |q|^{-1}|q\alpha - p| \geq |q|^{-1} \min_{m\in\mathbb{Z}}|q\alpha - m| = |q|^{-1}\|q\alpha\|,$$

which implies

$$\left|\alpha - \frac{p}{q}\right| \geq C|q|^{-d},\ \forall (p, q) \in \mathbb{Z} \times \mathbb{Z}_{\neq 0} \Leftrightarrow \|q\alpha\| \geq C|q|^{-(d-1)},\ \forall q \in \mathbb{Z}_{\neq 0}.$$

Of particular relevance to us is the following result, due to Schmidt [39], that real, irrational algebraic numbers are poorly approximated by rationals, in two dual senses.

Proposition 3.5. *[39] Let $\alpha_1, \ldots, \alpha_k \in (\mathbb{R} \cap \overline{\mathbb{Q}})$ such that $1, \alpha_1, \ldots, \alpha_k$ are linearly independent over $\mathbb{Q}$. For any $\epsilon \in \mathbb{R}_{>0}$, $\exists C = C(\alpha_1, \ldots, \alpha_k, \epsilon) \in \mathbb{R}_{>0}$ such that the following hold.*

(i) $\forall q \in \mathbb{Z}_{\neq 0}$, $\exists j \in \{1, \ldots, k\}$ such that $\|q\alpha_j\| \geq C|q|^{-(\frac{1}{k}+\epsilon)}$.

(ii) $\forall (q_1, \ldots, q_k) \in \mathbb{Z}^k$, where $q_{max} := \max_j |q_j| > 0$, we have $\|q_1\alpha_1 + \ldots + q_k\alpha_k\| \geq Cq_{max}^{-(k+\epsilon)}$.

We also require the following result concerning the Diophantine properties of linear forms in logarithms of algebraic numbers, due to Baker [6].

Proposition 3.6. *[6] Let $L = \{\beta \in \mathbb{C}_{\neq 0} | e^\beta \in \overline{\mathbb{Q}}\}$. For any $\beta_1, \ldots, \beta_k \in L$ that are linearly independent over $\mathbb{Q}$, there is an effectively computable constant $C = C(\beta_1, \ldots, \beta_k) \in \mathbb{R}_{>0}$ such that, $\forall (q_1, \ldots, q_k) \in \mathbb{Z}^k$ where $q_{max} := \max_j |q_j| > 0$, we have $|q_1\beta_1 + \cdots + q_k\beta_k| \geq (eq_{max})^{-C}$.*

Additionally, we require the following result of Gamburd, Jakobson, and Sarnak [16], concerning the Diophantine properties of $\mathrm{SU}(2, \overline{\mathbb{Q}})$, the group of 2×2 unitary matrices of determinant 1 whose entries are algebraic numbers, as well as a particular generalization to $\mathrm{U}(d, \overline{\mathbb{Q}})$. We first need a bit of notation. For a group G, and a finite collection of elements $h_1, \ldots, h_k \in G$, let $H = \langle h_1, \ldots, h_k\rangle$ denote the subgroup of G generated by $h_1, \ldots, h_k$ and let $\Sigma = \{h_1, \ldots, h_k, h_1^{-1}, \ldots, h_k^{-1}\}$ denote the corresponding symmetric generating set. For any $h \in H$, let $l(h)$ denote the length of H with respect to Σ. Let $M_d(S)$ denote the set of $d \times d$ matrices with entries in some set S. For $M \in M_d(\mathbb{C})$ let $\|M\|_{\mathrm{HS}}$ denote the Hilbert-Schmidt norm (i.e., $\|M\|_{\mathrm{HS}}^2 = \sum_{i,j}|M_{ij}|^2$), and note that for any $g \in \mathrm{SU}(2)$, $\|g \pm I_d\|_{\mathrm{HS}}^2 = 2|\mathrm{Tr}(g) \mp 2|$.

Proposition 3.7. *[16] For any $h_1, \ldots, h_k \in \mathrm{SU}(2, \overline{\mathbb{Q}})$, there is an effectively computable constant $C = C(h_1, \ldots, h_k) \in \mathbb{R}_{\geq 1}$, such that $\forall h \in H = \langle h_1, \ldots, h_k\rangle$ for which $h \neq \pm I_d$, we have $\|h \pm I_d\|_{HS} \geq C^{-l(h)}$.*

We now prove a straightforward generalization of the preceding result of Gamburd, Jakobson, and Sarnak [16]. Recall that the center of $\mathrm{U}(d, \overline{\mathbb{Q}})$ is given by $Z(\mathrm{U}(d, \overline{\mathbb{Q}})) = \{e^{ir}I_d | r \in \mathbb{R}, e^{ir} \in \overline{\mathbb{Q}}\}$.

Lemma 3.8. *For any $h_1, \ldots, h_k \in \mathrm{U}(d, \overline{\mathbb{Q}})$, there is an effectively computable constant $C = C(h_1, \ldots, h_k) \in \mathbb{R}_{\geq 1}$, such that $\forall h \in H = \langle h_1, \ldots, h_k\rangle$, if $h \notin Z(\mathrm{U}(d, \overline{\mathbb{Q}}))$, then $|\mathrm{Tr}(h)| \leq d - C^{-l(h)}$.*

Proof. Notice that $Z(\mathrm{U}(1, \overline{\mathbb{Q}})) = \mathrm{U}(1, \overline{\mathbb{Q}})$, and so the conclusion is vacuously true when $d = 1$; we assume for the remainder of the proof that $d \geq 2$.

We begin by following, essentially, the proof in [16]. As $\{h_1, \ldots, h_k\}$ is a finite subset of $M_d(\overline{\mathbb{Q}})$, there is some finite degree extension K of $\mathbb{Q}$ such that $\{h_1, \ldots, h_k\} \subseteq M_d(K)$. Let $\mathcal{O}_K$ denote the ring of integers of K and set $N \in \mathbb{Z}_{>0}$ sufficiently large such that $Nh_i \in M_d(\mathcal{O}_K)$, $\forall i$. Let s denote the degree of K over $\mathbb{Q}$, and let $\sigma_1, \dots, \sigma_s$ denote the s distinct embeddings of K in $\mathbb{C}$, where σ_1 is the identity map. Each $\sigma_j : K \to \mathbb{C}$ induces a map $M_d(K) \to M_d(\mathbb{C})$ in the obvious way, which we also denote by σ_j. For brevity, we write $\|\cdot\|$ in place of $\|\cdot\|_{\text{HS}}$ throughout this proof. Let $B = \max_{i,j}\|\sigma_j(h_i)\|$, and notice that $B \geq \sqrt{d}$ as $h_j \in \mathrm{U}(d)$ implies $\|\sigma_1(h_j)\| = \|h_j\| = \sqrt{d}$.

Fix $h \notin Z(\mathrm{U}(d, \overline{\mathbb{Q}}))$. In particular, $h \neq I_d = 1_H$, and so $l(h) \geq 1$. As $\|\cdot\|$ is submultiplicative, we then have $\|\sigma_j(h)\| \leq B^{l(h)}$, $\forall j$. For $r, c \in \{1, \dots, d\}$, and W a $d \times d$ matrix, we write $W[r, c]$ to denote the entry of W in row r and column c.

There are two cases. First, suppose there is some r such that $h[r, r] \neq h[1, 1]$. Fix such an r. Let y denote the $d \times d$ matrix given by $y = h - h[1, 1]I_d$ and notice that $y[r, r] = h[r, r] - h[1, 1] \neq 0$. For every j, we have

$$|\sigma_j(y[r, r])| = |\sigma_j(h[r, r]) - \sigma_j(h[1, 1])| \leq |\sigma_j(h[r, r])| + |\sigma_j(h[1, 1])| \leq 2\|\sigma_j(h)\| \leq 2B^{l(h)}.$$

By construction, $N^{l(h)}h \in M_d(\mathcal{O}_K)$, $\forall h \in H = \langle h_1, \dots, h_k\rangle$, which immediately implies $N^{l(h)}y = N^{l(h)}(h - h[1, 1]I_d) \in M_d(\mathcal{O}_K)$. Therefore, $N^{l(h)}y[r, r]$ is some non-zero element of $\mathcal{O}_K$, which implies $\prod_j \sigma_j(N^{l(h)}y[r, r]) \in \mathbb{Z}_{\neq 0}$. By the above, $|\sigma_j(N^{l(h)}y[r, r])| \leq 2(BN)^{l(h)} \leq (2BN)^{l(h)}$, $\forall j$. Therefore,

$$|y[r, r]| = |\sigma_1(y[r, r])| = N^{-l(h)}|\sigma_1(N^{l(h)}y[r, r])| \geq N^{-l(h)}\frac{1}{\prod_{j>1}|\sigma_j(N^{l(h)}y[r, r])|} \geq ((2B)^{d-1}N^d)^{-l(h)}.$$

Notice that

$$|h[r, r] + h[1, 1]|^2 + |h[r, r] - h[1, 1]|^2 = 2|h[r, r]|^2 + 2|h[1, 1]|^2 \leq 4.$$

Therefore,

$$|h[r, r] + h[1, 1]| \leq \sqrt{4 - |h[r, r] - h[1, 1]|^2} \leq 2 - \frac{1}{4}|h[r, r] - h[1, 1]|^2 = 2 - \frac{1}{4}|y[r, r]|^2 \leq 2 - C^{-l(h)},$$

where $C = ((2BN)^{2d}) \geq 1$ (notice $l(h) \geq 1$, $B \geq \sqrt{d} \geq 1$, and $N \geq 1$). Therefore,

$$|\mathrm{Tr}(h)| = \left|\sum_i h[i, i]\right| \leq |h[r, r] + h[1, 1]| + \left|\sum_{i \notin \{1, r\}} h[i, i]\right| \leq 2 - C^{-l(h)} + (d - 2) = d - C^{-l(h)}.$$

Next, suppose instead $h[r, r] = h[1, 1]$, $\forall r$. As $h \notin Z(\mathrm{U}(d, \overline{\mathbb{Q}}))$, there must then be some $r, c \in \{1, \dots, d\}$, $r \neq c$, such that $h[r, c] \neq 0$ (if there were no such r, c, then $h = h[1, 1]I_d \in Z(\mathrm{U}(d, \overline{\mathbb{Q}}))$). Fix such a pair r, c. For every j, we have

$$|\sigma_j(h[r, c])| \leq \|\sigma_j(h)\| \leq B^{l(h)}.$$

Furthermore, $N^{l(h)}h[r, c]$ is some non-zero element of $\mathcal{O}_K$, and so

$$|h[r, c]| = N^{-l(h)}|\sigma_1(N^{l(h)}h[r, c])| \geq N^{-l(h)}\frac{1}{\prod_{j>1}|\sigma_j(N^{l(h)}h[r, c])|} \geq (B^{d-1}N^d)^{-l(h)}.$$

As $|h[r, r]|^2 + |h[r, c]|^2 \leq 1$, we have

$$|h[r, r]| \leq \sqrt{1 - |h[r, c]|^2} \leq 1 - \frac{1}{2}|h[r, c]|^2 \leq 1 - C^{-l(h)}.$$

Therefore,

$$|\mathrm{Tr}(h)| = \left|\sum_i h[i, i]\right| \leq |h[r, r]| + \left|\sum_{i \neq r} h[i, i]\right| \leq 1 - C^{-l(h)} + (d - 1) = d - C^{-l(h)}.$$

□

By expressing the above condition in the language of representation theory, we then immediately have the following.

Corollary 3.8.1. *Suppose the group $G = \langle S|R\rangle$, with S finite, has a representation $\rho : G \to \mathrm{U}(d, \overline{\mathbb{Q}})$. Then there is an effectively computable constant $C = C(G, S, \rho) \in \mathbb{R}_{\geq 1}$, such that, if $g \notin \mathrm{Pker}(\rho)$, then $|\chi_\rho(g)| \leq d - C^{-l(g)}$, where $l(g)$ denotes the length of g with respect to the generating set S.*

3.3 Constructions of Distinguishing Families of Representations

We now show that a wide collection of groups virtually have $[k, d, \tau]$-DFRs with good parameters. We accomplish this by first constructing DFRs for only a small family of special groups. We then present several constructions in which a DFR for a group, or more generally a family of DFRs for a family of groups, is used to produce a DFR for a related group. This will allow us to construct DFRs with good parameters for a wide class of groups, and, ultimately, show that an even wider class of groups virtually have $[k, d, \tau]$-DFRs with good parameters. We begin with a straightforward lemma expressing a useful character bound. In this lemma, and throughout this section, we continue to write group operations multiplicatively, and so, for $g \in G$ and $h \in \mathbb{Z}$, if $h > 0$ then g^h denotes the element of G obtained by combining h copies of g with the group operation, if $h < 0$ then then g^h denotes the element obtained by combining h copies of g^{-1}, and if $h = 0$ then g^h is 1_G, by the usual convention on an empty product. For any $d \in \mathbb{N}_{\geq 1}$, let $\mathbf{1}_d : G \to \mathrm{U}(d)$ denote the trivial representation, i.e., $\mathbf{1}_d(g) = I_d$, $\forall g \in G$. As before, $S_1 = \{e^{ir}|r \in \mathbb{R}\}$ denotes the circle group.

Lemma 3.9. *Consider the cyclic group $G = \langle a|R_G\rangle$. Fix $r \in \mathbb{R}$ and define the representation $\phi : G \to S_1 \cong \mathrm{U}(1)$ such that $a \mapsto e^{2\pi ir}$; further define the representation $\rho : G \to \mathrm{T}(2)$ by $\rho = \phi \oplus \mathbf{1}_1$. Suppose that $h \in \mathbb{Z}$ and $\epsilon \in \mathbb{R}_{>0}$ satisfy $\|hr\| \geq \epsilon$. Then $\chi_\rho(a^h) \leq 2 - \frac{19\pi^2}{24}\epsilon^2$.*

Proof. We have

$$\chi_\rho(a^h) = e^{2\pi ihr} + 1 = e^{\pi ihr}\left(e^{\pi ihr} + e^{-\pi ihr}\right) = 2e^{\pi ihr}\cos(\pi hr).$$

As we must necessarily have $\epsilon \leq \frac{1}{2}$, it immediately follows that

$$|\chi_\rho(a^h)| = 2|\cos(\pi hr)| \leq 2\cos(\pi\epsilon) \leq 2\left(1 - \frac{(\pi\epsilon)^2}{2} + \frac{(\pi\epsilon)^4}{24}\right) \leq 2 - (\pi\epsilon)^2 + \frac{\pi^2(\pi\epsilon)^2}{48} \leq 2 - \frac{19\pi^2}{24}\epsilon^2.$$

□

We next construct DFRs for a very narrow class of special groups: (i) for any $m \in \mathbb{N}_{\geq 2}$, $\mathbb{Z}_m = \langle a|a^m\rangle$, the integers modulo m, where the group operation is addition (modulo m), (ii) $\mathbb{Z} = \langle a|\rangle$, the integers, where the group operations is addition, and (iii) $F_2 = \langle a, b|\rangle$ the free (non-abelian) group of rank 2. Note that $\mathbb{Z} = F_1$ is the free abelian group of rank 1.

Lemma 3.10. *For any $m \in \mathbb{N}_{\geq 2}$, $\mathbb{Z}_m = \langle a|a^m\rangle$ has a diagonal algebraic $\left[1, 2, \frac{19\pi^2}{24m^2}\right]$-DFR.*

Proof. Fix $m \in \mathbb{N}_{\geq 2}$, define the representation $\phi : \mathbb{Z}_m = \langle a|a^m\rangle \to S_1(\overline{\mathbb{Q}})$ such that $a \mapsto e^{\frac{2\pi i}{m}}$, and define the representation $\rho : \mathbb{Z}_m \to \mathrm{T}(2, \overline{\mathbb{Q}})$ where $\rho = \phi \oplus \mathbf{1}_1$. Then $\{\rho\}$ is a diagonal algebraic DFR for $\mathbb{Z}_m$, with the desired parameters. To see this, consider any $q \in \mathbb{Z}_m$, where $q \neq 1_{\mathbb{Z}_m}$. Then q can be expressed as $q = a^h$, for $h \in \mathbb{Z}$, $h \not\equiv 0 \mod m$. Let $r = \epsilon = \frac{1}{m}$. As we clearly have $\|hr\| \geq \epsilon$, Lemma 3.9 immediately implies $\chi_\rho(q) \leq 2 - \frac{19\pi^2}{24m^2}$. □

Lemma 3.11. *Let $\mathbb{Z} = \langle a|\rangle$.*

(i) $\forall \delta \in \mathbb{R}_{>0}$, $\exists C = C(\delta) \in \mathbb{R}_{>0}$ *such that* $\mathbb{Z} = \langle a | \rangle$ *has a* $\widetilde{E}$*-diagonal* $[1 + \lfloor \frac{2}{\delta} \rfloor, 2, Cn^{-\delta}]$*-DFR.*

(ii) Let R *denote the set of* $r \in (\mathbb{R} \backslash \mathbb{Q}) \cap (0,1)$ *for which* $e^{2\pi i r} \in \overline{\mathbb{Q}}$ *(e.g.,* $\widehat{r} = \frac{1}{2\pi} \cos^{-1}\left(\frac{3}{5}\right)$ *is irrational and has* $e^{2\pi i \widehat{r}} = \frac{3+4i}{5}$*, and so* $\widehat{r} \in R$*). Fix* $r \in R$*, define the representation* $\phi : \mathbb{Z} = \langle a | \rangle \to S_1(\overline{\mathbb{Q}})$ *such that* $a \mapsto e^{2\pi i r}$*, and define the representation* $\rho : \mathbb{Z} \to T(2, \overline{\mathbb{Q}})$ *as* $\rho = \phi \oplus \mathbf{1}_1$*. Then there are (effectively computable) constants* $C_1 = C_1(r), C_2 = C_2(r) \in \mathbb{R}_{>0}$ *for which* $\{\rho\}$ *is a diagonal algebraic* $[1, 2, C_2 n^{-C_1}]$*-DFR for* $\mathbb{Z}$*.*

Proof. (i) Let $k = 1 + \lfloor \frac{2}{\delta} \rfloor$ and $\eta = \frac{\delta}{2} - \frac{1}{k} > 0$. Fix $\alpha_1, \ldots, \alpha_k \in (\overline{\mathbb{Q}} \cap \mathbb{R})$ such that $1, \alpha_1, \ldots, \alpha_k$ are linearly independent over $\mathbb{Q}$. For each $j \in \{1, \ldots, k\}$ define the representation $\phi_j : \mathbb{Z} = \langle a | \rangle \to S_1(\widetilde{E})$ such that $a \mapsto e^{2\pi i \alpha_j}$, and let the representation $\rho_j : \mathbb{Z} \to \mathrm{T}(2, \widetilde{E})$ be given by $\rho_j = \phi_j \oplus \mathbf{1}_1$. Then for an appropriately chosen $C \in \mathbb{R}_{>0}$, $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ is a $[1 + \lfloor \frac{2}{\delta} \rfloor, 2, Cn^{-\delta}]$-DFR for $\mathbb{Z}$. To see this, notice that, by Proposition 3.5(i), $\exists D \in \mathbb{R}_{>0}$, such that $\forall q \in \mathbb{Z}_{\neq 0}$ (i.e., $\forall q \in \mathbb{Z}$ where $q \neq 1_{\mathbb{Z}} = 0$), $\exists j$ such that

$$\|q\alpha_j\| \geq D|q|^{-(\frac{1}{k} + \eta)} = D|q|^{-\frac{\delta}{2}}.$$

Therefore, for any $q \in \mathbb{Z}_{\neq 0}$, if we take j as above, then by Lemma 3.9, (with $r = \alpha_j$, $\epsilon = D|q|^{-\frac{\delta}{2}}$, and $h = q$) we have

$$|\chi_{\rho_j}(q)| \leq 2 - \frac{19\pi^2}{24} D^2 |q|^{-\delta}.$$

Therefore, $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ is a diagonal $[1 + \lfloor \frac{2}{\delta} \rfloor, 2, Cn^{-\delta}]$-DFR for $\mathbb{Z}$, with $C = \frac{19\pi^2}{24} D^2 > 0$.

(ii) As in Proposition 3.6, let $L = \{\beta \in \mathbb{C}_{\neq 0} | e^{\beta} \in \overline{\mathbb{Q}}\}$. Notice that $\pi i \in L$, as $e^{\pi i} = -1 \in \overline{\mathbb{Q}}$. By definition, $2\pi i r \in L$, which immediately implies $\pi i r \in L$. Also by definition, r is irrational, which implies $\pi i r$ and πi are linearly independent over $\mathbb{Q}$. Therefore, by Proposition 3.6, $\exists D \in \mathbb{R}_{>0}$ such that $\forall (q, m) \in \mathbb{Z}^2$ where $q_{max} := \max(|q|, |m|) > 0$, we have

$$|q\pi i r - m\pi i| \geq (eq_{max})^{-D}.$$

Consider any $q \in \mathbb{Z}_{\neq 0}$. For fixed q and varying $m \in \mathbb{Z}$, $|q\pi i r - m\pi i|$ attains its minimum when m is the closest integer to qr, which we denote by $\mathrm{round}(qr)$. Notice that $|\mathrm{round}(qr)| \leq |q|$, as $r \in (0,1)$ and $q \in \mathbb{Z}$. Therefore, for any $q \in \mathbb{Z}_{\neq 0}$, we have

$$\|qr\| = \min_{m \in \mathbb{Z}} |qr - m| = \frac{1}{\pi} \min_{m \in \mathbb{Z}} |q\pi i r - m\pi i| = \frac{1}{\pi} |q\pi i r - \mathrm{round}(qr)\pi i| \geq \frac{1}{\pi} |eq|^{-D}.$$

Applying Lemma 3.9, we conclude

$$\chi_{\rho}(q) \leq 2 - \frac{19}{24} |eq|^{-2D}.$$

Therefore, $\{\rho\}$ is a $[1, 2, C_2 n^{-C_1}]$-DFR for $\mathbb{Z}$, with $C_1 = 2D$ and $C_2 = \frac{19}{24} e^{-2D}$.

□

Remark. We note that the above constructions of diagonal DFRs for $\mathbb{Z}^k$ are quite similar to the technique used by Ambainis and Watrous [2] to produce a 2QCFA A_{AW} that recognizes L_{eq} (cf. [7, 36]). To clarify this similarity, let SO(2) denote the dimension-2 special orthogonal group, i.e., the group of all rotation matrices

$$R_{\theta} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \quad \theta \in \mathbb{R}.$$

Let $\mathbb{A} = \{\pi r | r \in (\overline{\mathbb{Q}} \cap \mathbb{R})\}$ (and so $\widetilde{E} = e^{i\mathbb{A}}$), notice that $\mathbb{A}$ is a subgroup of the group (under addition) $\mathbb{R}$, and let $\mathrm{SO}(2, \cos(\mathbb{A}))$ denote the subgroup of SO(2) consisting of those matrices whose entries lie in

$\cos(\mathbb{A})$ (i.e., the subgroup consisting of those R_θ with $\theta \in \mathbb{A}$). Then, essentially, A_{AW} recognizes L_{eq} by making use of, in our language, a DFR $\{\pi\}$ for $\mathbb{Z}$, for a representation $\pi : \mathbb{Z} \to \mathrm{SO}(2, \cos(\mathbb{A})) \leq \mathrm{U}(2)$. As $\mathrm{SO}(2, \cos(\mathbb{A})) \cong \mathrm{T}(1, \widetilde{E})$, π induces a representation $\widehat{\pi} : \mathbb{Z} \to \mathrm{T}(1, \widetilde{E})$, and $\widehat{\pi} \oplus \mathbf{1}_1 : \mathbb{Z} \to \mathrm{T}(2, \widetilde{E})$ is then a representation of the type we have produced above. Moreover, we note that the technique used by Ambainis and Watrous [2] relied on the fact that the number $\sqrt{2} \in \overline{\mathbb{Q}}$ is poorly approximated by rationals; our constructions above make use of more general Diophantine approximation results.

Lemma 3.12. *There is an effectively computable constant $C \in \mathbb{R}_{\geq 1}$, such that $F_2 = \langle a, b | \rangle$ has an algebraic $[1, 2, C^{-n}]$-DFR.*

Proof. First, define the representation $\pi : F_2 \to SO(3, \mathbb{Q})$ by

$$a \mapsto \frac{1}{5}\begin{pmatrix} 3 & -4 & 0 \\ 4 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix} \text{ and } b \mapsto \frac{1}{5}\begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & -4 \\ 0 & 4 & 3 \end{pmatrix}.$$

This is the "standard" faithful representation of F_2 into SO(3) used in many treatments of the Banach-Tarski paradox. Recall that SU(2) is the double cover of SO(3), i.e., $\mathrm{SU}(2)/Z(\mathrm{SU}(2)) \cong \mathrm{SO}(3)$. Then π induces a homomorphism $\widehat{\pi} : F_2 \to \mathrm{SU}(2)/Z(\mathrm{SU}(2))$ in the obvious way, which, by the universal property of the free group, can be lifted to the representation $\rho : F_2 \to \mathrm{SU}(2, \overline{\mathbb{Q}})$ given by

$$a \mapsto \frac{1}{\sqrt{5}}\begin{pmatrix} 2+i & 0 \\ 0 & 2-i \end{pmatrix} \text{ and } b \mapsto \frac{1}{\sqrt{5}}\begin{pmatrix} 2 & i \\ i & 2 \end{pmatrix}.$$

Then for any $g \in F_2$, where $g \neq 1_{F_2}$, $\rho(g) \notin Z(\mathrm{SU}(2, \overline{\mathbb{Q}})) = \{\pm I_2\}$, and, therefore, $\rho(g) \notin Z(\mathrm{U}(2, \overline{\mathbb{Q}}))$ (as $\rho(g) \in \mathrm{SU}(2)$, and $Z(\mathrm{SU}(2, \overline{\mathbb{Q}})) = \mathrm{SU}(2) \cap Z(\mathrm{U}(2, \overline{\mathbb{Q}}))$). Therefore, by Corollary 3.8.1, $\{\rho\}$ is an algebraic $[1, 2, C^{-n}]$-DFR for F_2. □

Remark. Similarly, we note that the method used in the proof of the preceding lemma to produce a DFR for F_2 is, fundamentally, the same construction used by Ambainis and Watrous [2] to produce a 2QCFA for L_{pal}. However, the algebraic structure of F_2 allows a substantially simpler argument to be used.

We now consider several constructions of new DFRs from existing DFRs. We emphasize that all results in the following lemmas are constructive in the sense that, given the supposed DFR or collection of DFRs, each corresponding proof provides an explicit construction of the new DFR. We begin by considering conversions of a DFR of a group G to a DFR with different parameters of the same group G.

Lemma 3.13. *Suppose $\mathcal{F}_G = \{\rho_1, \ldots, \rho_k\}$ is a $[k, d, \tau]$-DFR for $G = \langle S | R \rangle$. Then the following statements hold.*

(i) G has a $[1, kd, \tau]$-DFR.

(ii) If $d' \in \mathbb{N}$ and $d' > d$, then G has a $[k, d', \tau]$-DFR.

(iii) Suppose $S' \subseteq G$ is a finite generating set for G. Then there is an effectively computable constant $C = C(S, S') \in \mathbb{R}_{>0}$ such that $\mathcal{F}_G$ is also a $[k, d, \tau \circ \eta_C]$-DFR for $G = \langle S' | R' \rangle$, where $\eta_C : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is given by $\eta_C(n) = Cn$.

If, moreover, $\mathcal{F}_G$ is an algebraic (resp. diagonal) DFR, then each newly constructed DFR is also algebraic (resp. diagonal).

Proof. (i) Consider the representation $\rho : G \to \mathrm{U}(kd)$ of G given by $\rho = \rho_1 \oplus \cdots \oplus \rho_k$. As $\mathcal{F}_G = \{\rho_1, \ldots, \rho_k\}$ is a $[k, d, \tau]$-DFR for G, it satisfies the property Definition 3.1(ii); for each $g \in G_{\neq 1_G}$, set j_g to be the corresponding value of $j \in \{1, \ldots, k\}$ provided by the property. Therefore, for each $g \in G_{\neq 1_G}$, we have,

$$|\chi_\rho(g)| = \left|\sum_j \chi_{\rho_j}(g)\right| \leq |\chi_{\rho_{j_g}}(g)| + \left|\sum_{j \neq j_g} \chi_{\rho_i}(g)\right| \leq d - \tau(l(g)) + (k-1)d \leq kd - \tau(l(g)).$$

(ii) For each j, define the representation $\widehat{\rho}_j = \rho_j \oplus \mathbf{1}_{d'-d}$. Then $\{\widehat{\rho}_1, \ldots, \widehat{\rho}_k\}$ is a $[k, d', \tau]$-DFR, by an argument analogous to the above proof of (i).

(iii) Let $\Gamma(G, \Sigma)$ (resp. $\Gamma(G, \Sigma')$) denote the Cayley graph of G with (symmetric) generating sets $\Sigma = S \cup S^{-1}$ (resp. $\Sigma' = S' \cup S'^{-1}$). Let d_S and $d_{S'}$ denote the corresponding word metrics. Then $id_G : G \to G$, the identity map on G, is a bilipschitz equivalence between (G, d'_S) and (G, d_S) (see, for instance, [28, Proposition 5.2.4]), and so $\exists C = C(S, S') \in \mathbb{R}_{>0}$, which is straightforwardly computable, such that, $\forall g_1, g_2 \in G$, $\frac{1}{C} d_{S'}(g_1, g_2) \leq d_S(g_1, g_2) \leq C d_{S'}(g_1, g_2)$. We then write $l_S(g)$ and $l_{S'}(g)$ for the length of $g \in G$ with respect to each of the generating sets S and S', i.e., $l_S(g) = d_S(g, 1_G)$ and $l_{S'}(g) = d_{S'}(g, 1_G)$. By the above, $l_S(g) \leq C l_{S'}(g)$. As $\mathcal{F}_G$ is a $[k, d, \tau]$-DFR for G, we have that for each $g \in G_{\neq 1_G}$, $\exists j_g \in \{1, \ldots, k\}$ such that $|\chi_{\rho_{j_g}}(g)| \leq d - \tau(l_S(g))$. As $l_S(g) \leq C l_{S'}(g)$, and τ is monotone non-increasing, we then have $\tau(l_S(g)) \geq \tau(C l_{S'}(g))$, which immediately implies $|\chi_{\rho_{j_g}}(g)| \leq d - \tau(C l_{S'}(g))$, as desired.

□

Next, we show that a DFR of G and a DFR of H can be used to produce a DFR of $G \times H$, the direct product of G and H. In the following, for a group Q the commutator of elements $q_1, q_2 \in Q$, is denoted by $[q_1, q_2] = q_1^{-1} q_2^{-1} q_1 q_2$. For functions $\tau, \tau' : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, we define the function $\tau^{\min}_{\tau,\tau'} : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ by $\tau^{\min}_{\tau,\tau'}(m) := \min(\tau(m), \tau'(m))$, $\forall m \in \mathbb{R}_{>0}$.

Lemma 3.14. *Consider groups $G = \langle S_G | R_G \rangle$ and $H = \langle S_H | R_H \rangle$, with $S_G \cap S_H = \emptyset$. Let $R_{com} = \{[g, h] | g \in S_G, h \in S_H\}$. If G has a $[k, d, \tau]$-DFR and H has a $[k', d', \tau']$-DFR, then $G \times H = \langle S_G \sqcup S_H | R_G \cup R_H \cup R_{com} \rangle$ has a $[k + k', \max(d, d'), \tau^{\min}_{\tau,\tau'}]$-DFR. Moreover, if G and H have algebraic (resp. diagonal) DFRs with the above parameters, then $G \times H$ has an algebraic (resp. diagonal) DFR with the above parameters.*

Proof. By Lemma 3.13(ii), we may assume, without loss of generality, that $d' = d$ (i.e., we increase the smaller of d, d' to $\max(d, d')$). Let $\mathcal{F}_G = \{\rho_1, \ldots, \rho_k\}$ be a $[k, d, \tau]$-DFR for G and $\mathcal{F}_H = \{\pi_1, \ldots, \pi_{k'}\}$ a $[k', d, \tau']$-DFR for H. For each $j \in \{1, \ldots, k\}$, define a representation $\widehat{\rho}_j : G \times H \to \mathrm{U}(d)$ such that, $\forall (g, h) \in G \times H$, $\widehat{\rho}_j(g, h) = \rho_j(g)$. Analogously, for each $j \in \{1, \ldots, k'\}$, we define a representation $\widehat{\pi}_j : G \times H \to \mathrm{U}(d)$ such that, $\forall (g, h) \in G \times H$, $\widehat{\pi}_j(g, h) = \pi_j(h)$. Then $\mathcal{F}_{G \times H} = \{\widehat{\rho}_1, \ldots, \widehat{\rho}_k, \widehat{\pi}_1, \ldots, \widehat{\pi}_{k'}\}$ is the desired DFR. To see this, first notice that, $\forall (g, h) \in G \times H$, $l(g, h) = l(g) + l(h)$, where we write $l(g, h)$ in place of $l((g, h))$, to avoid cumbersome notation. By definition, τ and τ' are monotone non-increasing, and so, $\forall (g, h) \in G \times H$, we have $\tau(l(g, h)) \leq \tau(l(g))$ and $\tau'(l(g, h)) \leq \tau'(l(h))$. As $\mathcal{F}_G$ is a $[k, d, \tau]$-DFR for G, we have that for each $g \in G_{\neq 1_G}$, $\exists j_g \in \{1, \ldots, k\}$ such that $|\chi_{\rho_{j_g}}(g)| \leq d - \tau(l(g))$. Analogously, for each $h \in H_{\neq 1_H}$, $\exists j_h \in \{1, \ldots, k'\}$ such that $|\chi_{\pi_{j_h}}(h)| \leq d - \tau(l(h))$.

Consider $(g, h) \in G \times H$, where $(g, h) \neq 1_{G \times H} = (1_G, 1_H)$. Then we must have $g \neq 1_G$ or $h \neq 1_H$. If $g \neq 1_G$, then, by the above $\exists j_g$ such that

$$|\chi_{\widehat{\rho}_{j_g}}(g, h)| = |\chi_{\rho_{j_g}}(g)| \leq d - \tau(l(g)) \leq d - \tau(l(g, h)).$$

If, $h \neq 1_H$, then, analogously, $\exists j_h$ such that

$$|\chi_{\widehat{\pi}_{j_h}}(g, h)| = |\chi_{\pi_{j_h}}(h)| \leq d - \tau'(l(h)) \leq d - \tau'(l(g, h)).$$

Therefore, for any $(g, h) \in (G \times H)_{\neq 1_{G\times H}}$, there is some representation $\beta \in \mathcal{F}_{G\times H}$ for which

$$|\chi_\beta(g, h)| \leq \max(d - \tau(l(g, h)), d - \tau'(l(g, h))) = d - \min(\tau(l(g, h)), \tau'(l(g, h))) = d - \tau^{\min}_{\tau,\tau'}(l(g, h)).$$

□

Now, we show that a DFR of a group G can be used to produce a DFR of a finitely-generated subgroup of G, or of a finite-index overgroup of G.

Lemma 3.15. *Suppose $\mathcal{F}_G = \{\rho_1, \ldots, \rho_k\}$ is a $[k, d, \tau]$-DFR for $G = \langle S_G|R_G\rangle$. For $C \in \mathbb{R}_{>0}$, let $\eta_C : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ be given by $\eta_C(n) = Cn$. Then the following statements hold.*

(i) Suppose $H \leq G$, where $H = \langle S_H|R_H\rangle$, with S_H finite. Then there is an effectively computable constant $C \in \mathbb{R}_{>0}$ such that H has a $[k, d, \tau \circ \eta_C]$-DFR. If, moreover, $\mathcal{F}_G$ is an algebraic (resp. diagonal) DFR, then H will also have an algebraic (resp. diagonal) DFR with the claimed parameters.

(ii) Suppose $G \leq Q$, where $Q = \langle S_Q|R_Q\rangle$, with S_Q finite, $S_G \subseteq S_Q$, and $[Q : G] := r$ finite. Then there is an effectively computable constant $C \in \mathbb{R}_{>0}$ such that Q has a $[k, dr, \tau \circ \eta_C]$-DFR. If, moreover, $\mathcal{F}_G$ is an algebraic DFR, then Q will also have an algebraic DFR with the claimed parameters.

Proof. (i) As $H \leq G$, G admits a presentation $\langle S'_G|R'_G\rangle$ such that S'_G is finite and $S_H \subseteq S'_G$. Writing $l_{S_H}(h)$ for the length of $h \in H$ relative to the generating set S_H and $l_{S'_G}(g)$ for the length of $g \in G$ relative to the generating set S'_G, we immediately have that $l_{S_H}(h) \geq l_{S'_G}(h)$, $\forall h \in H \leq G$. By Lemma 3.13(iii), $\exists C \in \mathbb{R}_{>0}$ such that $\mathcal{F}_G$ is a $[k, d, \tau\circ\eta_C]$-DFR of $G = \langle S'_G|R'_G\rangle$. Let $\tau' = \tau \circ \eta_C$ and let $\mathcal{F}_H = \{\pi_1, \ldots, \pi_k\}$, where $\pi_j = \mathrm{Res}^G_H(\rho_j)$. As $\mathcal{F}_G$ is a $[k, d, \tau']$-DFR for G, we have that for each $h \in H \leq G$, where $h \neq 1_H = 1_G$, $\exists j_h \in \{1, \ldots, k\}$ such that $|\chi_{\rho_{j_h}}(h)| \leq d - \tau'(l_{S'_G}(h))$. Notice that $\chi_{\pi_j}(h) = \chi_{\rho_j}(h)$, $\forall h \in H, \forall j \in \{1, \ldots, k\}$. As τ' is monotone non-increasing, $\tau'(l_{S_H}(h)) \leq \tau'(l_{S'_G}(h))$. Therefore, $\forall h \in H_{\neq 1_H}$, $\exists j_h$ such that

$$|\chi_{\pi_{j_h}}(h)| = |\chi_{\rho_{j_h}}(h)| \leq d - \tau'(l_{S'_G}(h)) \leq d - \tau'(l_{S_H}(h)).$$

Therefore, $\mathcal{F}_H$ is a $[k, d, \tau \circ \eta]$-DFR for H.

(ii) For each $j \in \{1, \ldots, k\}$, let $\pi_j = \mathrm{Ind}^Q_G(\rho_j) : Q \to \mathrm{U}(kr)$. Then $\mathcal{F}_Q = \{\pi_1, \ldots, \pi_k\}$ is the desired DFR. To see this, let $T \subseteq Q$ be a complete family of left coset representatives of G in Q, where $1_Q \in T$. Notice that $|T| = [Q : G] = r$, with r finite. Then, for any $q \in Q$, we have (see, for instance, [26, Proposition 2.7.35])

$$\chi_{\pi_j}(q) = \sum_{\substack{t\in T\\ t^{-1}qt\in G}} \chi_{\rho_j}(t^{-1}qt).$$

Let $l_Q(q)$ denote the length of $q \in Q$ relative to S_Q and $l_G(g)$ denote the length of $g \in G \leq Q$ relative to S_G. Then $\exists C \in \mathbb{R}_{>0}$ such that $l_G(g) \leq Cl_Q(g)$, $\forall g \in G$, as $[Q : G]$ is finite. As τ is monotone non-increasing, $\tau(l_G(g)) \geq \tau(Cl_Q(g))$, $\forall g \in G$. Additionally, $\tau(l(g)) \leq d$, $\forall g \in G_{\neq 1_G}$. Therefore, if $g \in G_{\neq 1_G}$, then $d \geq \tau(l_G(g)) \geq \tau(Cl_Q(q))$.

Fix $q \in Q_{\neq 1_Q}$. First, suppose $q \in G$. As $\mathcal{F}_G = \{\rho_1, \ldots, \rho_k\}$ is a $[k, d, \tau]$-DFR for G, we conclude that there is some j such that $|\chi_{\rho_j}(q)| \leq d - \tau(l_G(q)) \leq d - \tau(Cl_Q(q))$. This immediately implies

$$|\chi_{\pi_j}(q)| = \left|\sum_{\substack{t\in T\\ t^{-1}qt\in G}} \chi_{\rho_j}(t^{-1}qt)\right| \leq |\chi_{\rho_j}(1_Q^{-1}q1_Q)| + \left|\sum_{\substack{t\in T\setminus 1_Q\\ t^{-1}qt\in G}} \chi_{\rho_j}(t^{-1}qt)\right| \leq d - \tau(Cl_Q(q)) + (r-1)d.$$

Therefore, there is some j such that $|\chi_{\pi_j}(q)| \leq dr - \tau(Cl_Q(q))$, if $q \in G$. Next, suppose instead $q \notin G$ and let $m = |\{t \in T | t^{-1}qt \in G\}|$. As $q \notin G$, $1_Q^{-1} q 1_Q = q \notin G$, and so $m \leq |T| - 1 = r - 1$. Therefore, $\forall j$, we have

$$|\chi_{\pi_j}(q)| = \left| \sum_{\substack{t \in T \\ t^{-1}qt \in G}} \chi_{\rho_j}(t^{-1}qt) \right| \leq dm \leq dr - d \leq dr - \tau(Cl_Q(q)).$$

Therefore, $\forall q \in Q_{\neq 1_Q}$, $\exists j$ such that $|\chi_{\pi_j}(q)| \leq dr - \tau(Cl_Q(q))$, as desired.

□

Remark. Notice that an immediate consequence of the preceding Lemma is that any group G that virtually has a DFR also has a DFR, but with worse parameters. In particular, if G virtually has a $[k, d, \tau]$-DFR, it has some subgroup $H \leq G$, $[G : H] := r$ finite, where H has a $[k, d, \tau]$-DFR. Then Lemma 3.15(ii) guarantees that $\exists C \in \mathbb{R}_{>0}$ such that G has a $[k, dr, \tau \circ \eta_C]$-DFR, where $\eta_C : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ is given by $\eta_C(n) = Cn$. However, the increase in the dimension of the representation space from d to dr has a corresponding increase in the size of the quantum register 2QCFA for the word problem of G, which is undesirable. As discussed earlier, it will be possible to solve the word problem for G using DFR for its subgroup H, thereby avoiding this issue.

We now present the main technical result of this section: the construction of DFRs, with good parameters, for a wide class of groups. We begin with the finitely-generated (virtually) abelian groups. Recall that any finitely-generated abelian group G admits a unique decomposition $G \cong \mathbb{Z}^r \times \mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_l}$, where m_i divides m_{i+1}, $\forall i \in \{1, \ldots, l-1\}$ (it is more standard to express the previous decomposition using the direct sum, as such a decomposition is most naturally thought of as coproduct, which, in the category of abelian groups, is the direct sum; however, as direct products and direct sums of a finite number of groups are canonically isomorphic, we express the above decomposition using the direct product, for ease of notation). Clearly, $\mathbb{Z}^r \times \mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_l}$ has presentation $\langle a_1, \ldots, a_{r+l} | R(r, m_1, \ldots, m_l) \rangle$ where $R(r, m_1, \ldots, m_l) = \{a_i^{m_i} | i \in \{1, \ldots, l\}\} \cup \{[a_i, a_j] | i, j \in \{1, \ldots, r+l\}\}$ with $[a_i, a_j] = a_i^{-1} a_j^{-1} a_i a_j$ denoting the commutator of a_i and a_j.

Theorem 3.16. *There is an (effectively computable) constant $C_1 \in \mathbb{R}_{>0}$, such that, for any finitely-generated abelian group $G = \mathbb{Z}^r \times \mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_l} = \langle a_1, \ldots, a_{r+l} | R(r, m_1, \ldots, m_l) \rangle$, the following statements hold.*

(i) Suppose $r = 0$. In the trivial case in which $l = 0$, i.e., G is the trivial group, G has a diagonal algebraic $[1, 2, 2]$-DFR. Otherwise, G has a diagonal algebraic $\left[l, 2, \frac{19\pi^2}{24m_l^2}\right]$-DFR.

(ii) If $r \neq 0$, $\exists C_2 \in \mathbb{R}_{>0}$, with C_2 effectively computable, such that G has a diagonal algebraic $\left[r + l, 2, C_2 n^{-C_1}\right]$-DFR.

(iii) If $r \neq 0$, then $\forall \delta \in \mathbb{R}_{>0}$, $\exists C_3 \in \mathbb{R}_{>0}$ such that G has a $\widetilde{E}$-diagonal $\left[r(1 + \lfloor \frac{2}{\delta} \rfloor) + l, 2, C_3 n^{-\delta}\right]$-DFR.

Proof. Apply Lemma 3.11(ii), for an arbitrary x that satisfies the hypothesis of the lemma, e.g., $x = \frac{1}{2\pi} \cos^{-1}\left(\frac{3}{5}\right)$. Then there are effectively computable constants $D_1, D_2 \in \mathbb{R}_{>0}$ such that $\mathbb{Z} = \langle a | \rangle$ has a diagonal algebraic $[1, 2, D_2 n^{-D_1}]$-DFR, which we call $\mathcal{F}$. We set $C_1 = D_1$. Now, consider the finitely-generated abelian group $G = \mathbb{Z}^r \times \mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_l}$.

(i) When $l = 0$, the claim immediately follows by considering the representation $\rho : \{1\} \to \mathrm{U}(2)$, for which $\rho(1) = I_2$. Suppose $l > 0$. By Lemma 3.10, each factor $\mathbb{Z}_{m_i} = \langle a | a^{m_i} \rangle$ has a diagonal algebraic $\left[1, 2, \frac{19\pi^2}{24m_i^2}\right]$-DFR. Notice that $m_1 \leq \cdots \leq m_l$, as each m_i divides m_{i+1}. The existence of the desired DFR is then an immediate consequence of Lemma 3.14.

(ii) Using the DFR $\mathcal{F}$ of $\mathbb{Z}$, Lemma 3.14 implies $H_1 := \mathbb{Z}^r$ has a diagonal algebraic $[r, 2, D_2 n^{-C_1}]$-DFR $\mathcal{H}_1$. If $l = 0$, then $G = H_1$; therefore, $\mathcal{H}_1$ is the desired DFR for G, with $C_2 = D_2$, and we are done. If $l > 0$, part (i) of this lemma shows $H_2 := \mathbb{Z}_{m_1} \times \cdots \times \mathbb{Z}_{m_l}$ has a diagonal algebraic $\left[l, 2, \frac{19\pi^2}{24m_l^2}\right]$-DFR $\mathcal{H}_2$. Set $C_2 = \min(D_2, \frac{19\pi^2}{24m_l^2})$. By Lemma 3.14, we conclude $G = H_1 \times H_2$ has a DFR with the claimed parameters.

(iii) By Lemma 3.11(i), $\exists D \in \mathbb{R}_{>0}$ such that $\mathbb{Z} = \langle a | \rangle$ has a $\widetilde{E}$-diagonal $\left[1 + \lfloor \frac{2}{\delta} \rfloor, 2, Dn^{-\delta}\right]$-DFR, $\mathcal{F}'$. The remainder of the proof is precisely analogous to that of part(ii), using $\mathcal{F}'$ in place of $\mathcal{F}$.

□

Recall that we say a group is virtually abelian if it has a finite-index subgroup that is abelian. Notice that all finite groups are virtually abelian as any finite group contains the (trivial) abelian group $\{1\}$ as a (necessarily) finite-index subgroup. As defined in Section 1.2, we write $\widehat{\Pi}_1$ to denote the class of all finitely-generated virtually abelian groups. Recall that, for any $G \in \widehat{\Pi}_1$, there is a unique $r \in \mathbb{N}$ such that G has a finite-index subgroup isomorphic to $\mathbb{Z}^r$. The following is immediate.

Corollary 3.16.1. *There is an effectively computable constant $C \in \mathbb{R}_{>0}$, such that, for any $G \in \widehat{\Pi}_1$, the following holds.*

(i) If G is finite, there are (trivially) effectively computable $D \in \mathbb{R}_{>0}$ and $K \in \mathbb{N}_{>0}$, such that G virtually has a diagonal algebraic $[K, 2, D]$-DFR.

(ii) There are effectively computable $D \in \mathbb{R}_{>0}$ and $K \in \mathbb{N}_{>0}$, such that G virtually has a diagonal algebraic $[K, 2, Dn^{-C}]$-DFR.

(iii) $\forall \delta \in \mathbb{R}_{>0}$, $\exists D \in \mathbb{R}_{>0}$ and $K \in \mathbb{N}_{>0}$, such that G virtually has a $\widetilde{E}$-diagonal $\left[K, 2, Dn^{-\delta}\right]$-DFR.

Consider the group $\mathbb{Z}^r$, for $r \in \mathbb{N}_{\geq 1}$. By Theorem 3.16, we see that $\mathbb{Z}^r$ has a diagonal algebraic $\left[r, 2, C_2 n^{-C_1}\right]$-DFR, as well as a $\widetilde{E}$-diagonal $\left[r(1 + \lfloor \frac{2}{\delta} \rfloor), 2, C_3 n^{-\delta}\right]$-DFR, for any $\delta \in \mathbb{R}_{>0}$, for particular constants C_1, C_2, C_3. While these DFRs suffice for establishing all of our results concerning the recognizability of the word problem for $\mathbb{Z}^r$, we next exhibit a different construction of a DFR for $\mathbb{Z}^r$, which we will require in Section 3.5. In the following, for a commutative (unital) ring R, let $\mathrm{SO}(2, R)$ denote the group of 2×2 orthogonal matrices of determinant 1 whose entries lie in R. For a set of prime numbers $\mathcal{P} = \{p_1, \ldots, p_m\}$, let $\mathbb{Z}[\frac{1}{p_1}, \ldots, \frac{1}{p_m}]$ denote the ring obtained by adjoining $\frac{1}{p_1}, \ldots, \frac{1}{p_m}$ to the ring $\mathbb{Z}$, i.e., $\mathbb{Z}[\frac{1}{p_1}, \ldots, \frac{1}{p_m}]$ is the localization of $\mathbb{Z}$ away from $\mathcal{P}$. Notice that $\mathrm{SO}(2, \mathbb{Z}[\frac{1}{p_1}, \ldots, \frac{1}{p_m}]) \leq \mathrm{SO}(2, \mathbb{Q}) \leq \mathrm{SU}(2, \mathbb{Q}) \leq \mathrm{SU}(2, \overline{\mathbb{Q}})$.

Lemma 3.17. *Consider the group $\mathbb{Z}^r = \langle S_r | R_r \rangle$, where $S_r = \{a_1, \ldots, a_r\}$ and $R_r = \{[a_i, a_j] | i, j \in \{1, \ldots, r\}\}$. There is a representation $\rho : \mathbb{Z}^r \to \mathrm{SO}(2, \mathbb{Z}[\frac{1}{p_1}, \ldots, \frac{1}{p_r}])$ such that, for some effectively computable constants $D_1, D_2 \in \mathbb{R}_{>0}$, $\{\rho\}$ is a $[1, 2, D_2 n^{-D_1}]$-algebraic DFR for $\mathbb{Z}^r$.*

Proof. Fundamentally, we follow the construction of Tan [43] of the rational points on the unit circle. Let p_j denote the j^{th} prime number that is congruent to 1 modulo 4, and let $m_j, n_j \in \mathbb{N}$ denote the (unique) values which satisfy $p_j = m_j^2 + n_j^2$ and $m_j > n_j > 0$. Define the representation $\rho : \mathbb{Z}^r \to \mathrm{SO}(2, \mathbb{Z}[\frac{1}{p_1}, \ldots, \frac{1}{p_r}])$ such that

$$a_j \mapsto \frac{1}{p_j} \begin{pmatrix} m_j^2 - n_j^2 & 2m_j n_j \\ -2m_j n_j & m_j^2 - n_j^2 \end{pmatrix}, \quad \forall j \in \{1, \ldots, r\}.$$

Notice that $\rho(a_j)$ has eigenvalues $p_j^{-1}(m_j^2 - n_j^2 \pm 2m_jn_ji)$. As $\mathrm{SO}(2,\mathbb{Z}[\frac{1}{p_1},\dots,\frac{1}{p_r}])$ is abelian, the $\rho(a_j)$ are simultaneously diagonalizable. Define $Y \in \mathrm{U}(2)$ such that, $\forall j$, $Y\rho(a_j)Y^{-1} = D_j$, where D_j is a 2×2 diagonal matrix whose diagonal entries are the eigenvalues $p_j^{-1}(m_j^2 - n_j^2 \pm 2m_jn_ji)$. Define $\alpha_j \in (\mathbb{R}\cap(-\pi,\pi))$ such that $D_j = \mathrm{diag}[e^{i\alpha_j}, e^{-i\alpha_j}]$.

For some $(q_1,\dots,q_r) \in \mathbb{Z}^r$, consider the element $g = a_1^{q_1}\cdots a_r^{q_r} \in \mathbb{Z}^r$. Then

$$\chi_\rho(g) = \mathrm{Tr}\left(\prod_{j=1}^r \rho(a_j)^{q_j}\right) = \mathrm{Tr}\left(\prod_{j=1}^r \left(Y\rho(a_j)Y^{-1}\right)^{q_j}\right) = e^{i\sum_j q_j\alpha_j} + e^{-i\sum_j q_j\alpha_j} = 2\cos\left(\sum_j q_j\alpha_j\right).$$

Let $L = \{\beta \in \mathbb{C}_{\neq 0} | e^\beta \in \overline{\mathbb{Q}}\}$. Let $\beta_0 = i\pi$ and, for $j \in \{1,\dots,r\}$, let $\beta_j = i\alpha_j$. Then $\beta_0,\dots,\beta_r \in L$. By [43, Theorem 1], ρ is P-faithful, which immediately implies $\beta_0,\dots,\beta_r$ are linearly independent over $\mathbb{Q}$. By Proposition 3.6, there is an effectively computable constant $C \in \mathbb{R}_{>0}$ such that, $\forall(q_0,\dots,q_r) \in \mathbb{Z}^{r+1}$, where $q_{\max} := \max_j|q_j| > 0$, we have $|\sum_j q_j\beta_j| \geq (eq_{\max})^{-C}$.

Consider any $g = a_1^{q_1}\cdots a_r^{q_r} \in \mathbb{Z}^r_{\neq 1_{\mathbb{Z}^r}}$ (i.e., not all $q_i = 0$). Let $q_0 = \mathrm{round}(\frac{1}{\pi}\sum_{j=1}^r q_j\alpha_j)$ and observe that, by construction $|\alpha_j| \leq \pi$, $\forall j$, and so $|q_0| \leq \sum_{j=1}^r |q_j| = l(g)$. Therefore, $q_{\max} := \max_{j\in\{0,\dots,r\}} q_j \leq l(g)$, which implies

$$\min_{m\in\mathbb{Z}}\left|m\pi + \sum_{j=1}^r q_j\alpha_j\right| = \left|q_0\pi + \sum_{j=1}^r q_j\alpha_j\right| = \left|q_0\beta_0 + \sum_{j=1}^r q_j\beta_j\right| \geq (el(g))^{-C}.$$

Therefore,

$$|\chi_\rho(g)| = 2\left|\cos\left(\sum_j q_j\alpha_j\right)\right| \leq 2 - C'\min_{m\in\mathbb{Z}}\left|m\pi + \sum_{j=1}^r q_j\alpha_j\right|^2 \leq 2 - C'(el(g))^{-2C},$$

for a constant $C' \in \mathbb{R}_{>0}$. We then conclude that $\{\rho\}$ is a $[1,2,D_2n^{-D_1}]$-algebraic DFR for $\mathbb{Z}^r$, where $D_1 = 2C$ and $D_2 = C'e^{-2C}$. □

Next, we consider groups that can be built, in certain ways, from finitely-generated free groups.

Theorem 3.18. *Suppose $G = \langle S|R\rangle$, with S finite, such that $G \leq F_{r_1}\times\cdots\times F_{r_t}$, for some $r_1,\dots,r_t \in \mathbb{N}$. Then there is an effectively computable constant $C \in \mathbb{R}_{\geq 1}$ such that G has an algebraic $[t,2,C^{-n}]$-DFR. In particular, for any $r \in \mathbb{N}$, $F_r = \langle a_1,\dots,a_r|\rangle$ has an algebraic $[1,2,C^{-n}]$-DFR.*

Proof. We first show that, for any $r \in \mathbb{N}$, there is an (effectively computable) constant $C \in \mathbb{R}_{\geq 1}$ such that $F_r = \langle a_1,\dots,a_r|\rangle$ has an algebraic $[1,2,C^{-n}]$-DFR. As $F_0 = \{1\}$ and $F_1 = \mathbb{Z}$, Theorem 3.16 immediately implies the claim when $r \in \{0,1\}$. Next, consider the case in which $r = 2$. By Lemma 3.12, $\exists C \in \mathbb{R}_{\geq 1}$ such that the free group of rank 2, $F_2 = \langle a_1,a_2|\rangle$, has an algebraic $[1,2,C^{-n}]$-DFR. If $r > 2$, then fix r, and note that, by the Nielsen-Schreier theorem, F_2 has a finite-index subgroup isomorphic to F_r. The result immediately follows from Lemma 3.15(i).

Next, suppose $G = \langle S|R\rangle$, with S finite, such that $G \leq F_{r_1}\times\cdots\times F_{r_t}$, for some $r_1,\dots,r_t \in \mathbb{N}$. By the previous paragraph, each F_{r_i} has an algebraic $[1,2,C_i^{-n}]$-DFR, for some $C_i \in \mathbb{R}_{\geq 1}$. Lemma 3.14 implies that $F_{r_1}\times\cdots\times F_{r_t}$ has an algebraic $[t,2,C^{-n}]$-DFR, where $C = \max_i C_i$, and Lemma 3.15(i) then implies G has a DFR with the claimed parameters. □

As defined in Section 1.2, we write $\Sigma_1 = \{F_k | k \in \mathbb{N}\}$ to denote the finitely-generated (i.e., finite-rank) free groups, Π_2 to denote the class of groups isomorphic to a direct product of some finite set of groups in Σ_1, and $\widehat{\Pi}_2$ to denote the class of finitely-generated groups that are virtually a (necessarily finitely-generated) subgroup of some group in Π_2. The following corollary is immediate.

Corollary 3.18.1. *Suppose $G \in \widehat{\Pi}_2$. Then G has a finite-index subgroup H such that $H \cong F_{r_1} \times \cdots \times F_{r_t}$, for some $r_1, \ldots, r_t \in \mathbb{N}$. Then there is an effectively computable constant $C \in \mathbb{R}_{\geq 1}$ such that G virtually has an algebraic $[t, 2, C^{-n}]$-DFR.*

Remark. Note that the above theorem (and its corollary) would also hold if we considered finitely-generated subgroups of $F_{r_1} \times \cdots \times F_{r_t} \times A$ (and their finite-index overgroups), where A is a finite abelian group. This, of course, does not extend the class of groups for which we can construct a DFR, but it does allow for a slightly better bound on the number of classical states of the 2QCFA for certain word problems (by "moving" the finite abelian group into the DFR).

We conclude with a "generic" construction, that, in a certain sense, covers all groups that have algebraic DFRs. We remark that while this does partially subsume all other results in this section, it does not do so completely, as the earlier constructions of DFRs for certain particular groups will, in several important special cases, have parameters that are better than those guaranteed by this construction.

Theorem 3.19. *Consider a group $G = \langle S|R \rangle$, with S finite, where G is not the trivial group. Suppose G has a faithful representation $\pi : G \to \mathrm{U}(l, \overline{\mathbb{Q}})$. Then π has a (unique, up to isomorphism of representations) set of irreducible subrepresentations $\{\pi_j : G \to \mathrm{U}(d_j, \overline{\mathbb{Q}})\}_{j=1}^m$ such that $\pi \cong \pi_1 \oplus \cdots \oplus \pi_m$. Let $d_{\max} = \max_j d_j$. Define the value d as follows: if $\cap_j \operatorname{Pker}(\pi_j) = \{1_G\}$, let $d = d_{\max}$, otherwise, let $d = d_{\max} + 1$. Partition the non-trivial π_j into isomorphism classes (i.e., only consider those π_j which are not the trivial representation; π_{j_1} and π_{j_2} belong to the same isomorphism class precisely when $\pi_{j_1} \cong \pi_{j_2}$) and let k denote the number of isomorphism classes that appear. Then there is an effectively computable constant $C \in \mathbb{R}_{\geq 1}$ such that G has an algebraic $[k, d, C^{-n}]$-DFR.*

Proof. Notice that, as G is not the trivial group, we must have $d \geq 2$. Assume, for notational convenience, that the π_j are ordered such that $\pi_1, \ldots, \pi_k$ are representatives of the k distinct isomorphism classes of the non-trivial representations that appear among the π_j. For each $j \in \{1, \ldots, k\}$, define the representation $\rho_j : G \to \mathrm{U}(d, \overline{\mathbb{Q}})$ as $\rho_j = \pi_j \oplus \mathbf{1}_{d-d_j}$. By Corollary 3.8.1, we have that for each $j \in \{1, \ldots, k\}$, there is an effectively computable constant $C_j = C_j(G, S, \rho_j) \in \mathbb{R}_{\geq 1}$ such that, $\forall g \notin \operatorname{Pker}(\rho_j)$, $|\chi_{\rho_j}(g)| \leq d - C_j^{-l(g)}$. Set $C = \max_j C_j$.

Next, notice that $\cap_j \operatorname{Pker}(\rho_j) = \{1_G\}$. If $\cap_j \operatorname{Pker}(\pi_j) = \{1_G\}$, then this is obvious. Suppose $\cap_j \operatorname{Pker}(\pi_j) \neq \{1_G\}$. Then $d = d_{\max} + 1 > d_j$, $\forall j$, which implies $\rho_j = \pi_j \oplus \mathbf{1}_{t_j}$, where $t_j := d - d_j \geq 1$. Therefore, for each j, $\rho_j(G) \cap Z(\mathrm{U}(d, \overline{\mathbb{Q}})) = I_d$, and so, by definition, $\operatorname{Pker}(\rho_j) = \ker(\rho_j)$. As π is faithful,

$$\{1_G\} = \bigcap_{j=1}^{m} \ker(\pi_j) = \bigcap_{j=1}^{k} \ker(\rho_j) = \bigcap_{j=1}^{k} \operatorname{Pker}(\rho_j).$$

This immediate implies that, $\forall g \in G_{\neq 1_G}$, $\exists j_g$ such that $g \notin \operatorname{Pker}(\rho_{j_g})$, which implies

$$|\chi_{\rho_{j_g}}(g)| \leq d - C_{j_g}^{-l(g)} \leq d - C^{-l(g)}.$$

Therefore, $\{\rho_1, \ldots, \rho_k\}$ is an algebraic $[k, d, C^{-n}]$-DFR for G. □

Corollary 3.19.1. *Suppose G is a finitely-generated group. Suppose further that G has a finite-index subgroup H such that H is isomorphic to a subgroup $\widetilde{H}$ of $(\mathrm{U}(d, \overline{\mathbb{Q}}))^k$ (the direct product of k copies of $(\mathrm{U}(d, \overline{\mathbb{Q}}))$), for some $d, k \in \mathbb{N}_{\geq 1}$. Then there is an effectively computable constant $C \in \mathbb{R}_{\geq 1}$ such that G virtually has an algebraic $[k, d+1, C^{-n}]$-DFR. Moreover, if $\widetilde{H}$ is not the trivial group and $\widetilde{H} \cap Z((\mathrm{U}(d, \overline{\mathbb{Q}}))^k)$ is the trivial group (in particular, this implies $d \geq 2$), then G virtually has an algebraic $[k, d, C^{-n}]$-DFR.*

Remark. While we demonstrated that many groups have a DFR, it is of course not the case that all groups have a DFR. Let $\mathcal{U}$ (resp. $\mathcal{L}$) denote the set of finitely-generated groups that have a faithful finite-dimensional unitary (resp. $\mathbb{C}$-linear) representation. As noted in Proposition 3.2, the existence of a DFR for a finitely-generated group G immediately implies $G \in \mathcal{U}$. Of course, $\mathcal{U} \subseteq \mathcal{L}$ and, in fact, $\mathcal{U} \subsetneq \mathcal{L}$. To see this, note that, by the Tits' alternative [45], $\mathcal{L}$ is divided into two classes: those groups that are virtually solvable, and those groups that have a subgroup isomorphic to F_2. By [44, Proposition 2.2], the only virtually solvable groups in $\mathcal{U}$ are virtually abelian, and so a DFR could certainly not exist for any virtually solvable $G \in \mathcal{L}$ that is not virtually abelian. On the other hand, we have demonstrated the existence of a DFR for every finitely-generated virtually abelian group.

3.4 Projective DFRs

Thus far, we have considered DFRs that consist of ordinary representations; that is to say, a DFR $\mathcal{F} = \{\rho_1, \ldots, \rho_j\}$ of a group G is a collection of representations (i.e., group homomorphisms) $\rho_j : G \to \mathrm{U}(d)$. We next consider a slight generalization to projective representations. As before, we write $S_1 = \{e^{ir} | r \in \mathbb{R}\}$ for the circle group, $Z(\mathrm{U}(d)) = S_1 I_d$ for the center of $\mathrm{U}(d)$, and $\mathrm{PU}(d) = \mathrm{U}(d)/Z(\mathrm{U}(d))$ for the d-dimensional projective unitary group. Recall that a projective representation of G is a group homomorphism $\rho : G \to \mathrm{PU}(d)$. For any such ρ, we may define a function (not necessarily a group homomorphism) $\widehat{\rho} : G \to \mathrm{U}(d)$, such that, for each $g \in G$, $\widehat{\rho}(g) \in \mathrm{U}(d)$ is some lift of $\rho(g) \in \mathrm{PU}(d)$, i.e., $\gamma \circ \widehat{\rho} = \rho$, where $\gamma : \mathrm{U}(d) \to \mathrm{PU}(d)$ is the canonical projection. For a particular projective representation ρ of G, it may or may not be possible to choose some $\widehat{\rho}$ that is a unitary representation of G (i.e., such that $\widehat{\rho} : G \to \mathrm{U}(d)$ is a group homomorphism).

A 2-*cocycle* is a mapping $\zeta : G \times G \to S_1$ such that $\zeta(g_1, g_2)\zeta(g_1 g_2, g_3) = \zeta(g_1, g_2 g_3)\zeta(g_2, g_3)$, $\forall g_1, g_2, g_3 \in G$. For any $\widehat{\rho}$, there is a unique ζ such that, $\forall g_1, g_2 \in G$, we have $\widehat{\rho}(g_1)\widehat{\rho}(g_2) = \zeta(g_1, g_2)\widehat{\rho}(g_1, g_2)$. Then, for any two lifts $\widehat{\rho}_1$ and $\widehat{\rho}_2$, we have $|\chi_{\widehat{\rho}_1}(g)| = |\chi_{\widehat{\rho}_2}(g)|$. Therefore, the function $|\chi_\rho(\cdot)| : G \to \mathbb{C}$ given by $|\chi_\rho(g)| = |\chi_{\widehat{\rho}}(g)|$, $\forall g \in G$, is well-defined. We then define a $[k, d, \tau]$-PDFR as a set of projective representations $\mathcal{F} = \{\rho_1, \ldots, \rho_j\}$ that satisfies Definition 3.1 where "representation" is replaced by "projective representation" in that definition. As we will observe in the following section, the same process that allows a DFR for a group G to be used to produce a 2QCFA for the word problem W_G, can also be applied to a PDFR. Clearly, any DFR is a PDFR, with the same parameters. However, as a projective representation of a group cannot always be lifted to an ordinary representation of that group, this is a generalization.

Let $\mathrm{PU}(d, \overline{\mathbb{Q}}) = \mathrm{U}(d, \overline{\mathbb{Q}})/Z(\mathrm{U}(d, \overline{\mathbb{Q}}))$ denote the d-dimensional projective unitary group with algebraic number entries. If a PDFR consists entirely of representations into $\mathrm{PU}(d, \overline{\mathbb{Q}})$, we say it is an *algebraic* PDFR. The following variant of Theorem 3.19 follows by a precisely analogous proof.

Theorem 3.20. *Suppose the group $G = \langle S, R \rangle$, with S finite, has a family $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ of projective representations $\rho_j : G \to \mathrm{PU}(d, \overline{\mathbb{Q}})$, such that $\cap_j \ker(\rho_j) = \{1_G\}$. Then there is an effectively computable constant $C \in \mathbb{R}_{\geq 1}$ such that $\mathcal{F}$ is an algebraic $[k, d, C^{-n}]$-PDFR for G.*

3.5 Unbounded DFRs

As noted in Proposition 3.2, if $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ is a DFR for a group G, then $\cap_j \mathrm{Pker}(\rho_j) = \{1_G\}$. However, a crucial element in the definition of a DFR is the requirement that, much more strongly, all $g \in G_{\neq 1_G}$ are "far" from being in $\cap_j \mathrm{Pker}(\rho_j)$; in particular, if $\mathcal{F}$ is a $[k, d, \tau]$-DFR, then for each $g \in G_{\neq 1_G}$ there is some j such that $|\chi_{\rho_j}(g)| \leq d - \tau(l(g))$. This requirement is essential in order for the construction of a 2QCFA for W_G using a DFR for G, which we present in Section 4, to operate with one-sided bounded-error. We next consider a generalization of a DFR, in which this requirement is removed; as we will then see in Section 4, these less constrained DFRs will still yield a 2QCFA that recognizes the corresponding group word problem with one-sided *unbounded-error*.

Definition 3.21. We say $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ is an *unbounded-error* $[k, d]$-DFR for a group $G = \langle S|R \rangle$ if the conditions of Definition 3.1 hold, where Definition 3.1(ii) is replaced by:

Definition 3.1(ii)': $\forall g \in G_{\neq 1_G}$, $\exists j \in \{1, \ldots, k\}$ such that $|\chi_{\rho_j}(g)| < d$. This condition is equivalent to $\cap_j \operatorname{Pker}(\rho_j) = \{1_G\}$.

Similarly, if the ρ_j are projective representations, then we say $\mathcal{F}$ is an unbounded-error $[k, d]$-PDFR. We also define algebraic, diagonal, and virtual unbounded-error DFRs in the obvious way.

Of course, any $[k, d, \tau]$-DFR for G is also an unbounded-error $[k, d]$-DFR for G. In the other direction, any algebraic unbounded-error $[k, d]$-DFR is also an algebraic $[k, d, C^{-n}]$-DFR, for some constant $C \in \mathbb{R}_{\geq 1}$, by Corollary 3.8.1; furthermore, by the same argument that follows the discussion of Definition 3.3, only a finitely-generated abelian group could have a diagonal unbounded-error $[k, d]$-DFR, and all finitely-generated abelian groups were shown to have DFRs in Theorem 3.16. Therefore, in order to obtain something new, we must consider unbounded-error DFRs that are neither algebraic nor diagonal.

Recall that for groups $G = \langle S_G|R_G \rangle$ and $H = \langle S_H|R_H \rangle$, the *free product* of G and H, which we denote $G * H$, is the group $G * H = \langle S_G \sqcup S_H | R_G \sqcup R_H \rangle$ (here, for notational convenience, we assume that S_G and S_H are disjoint; of course, if G and H are both subgroups of some group, there could be non-trivial $g \in G$ and $h \in H$ such that $g = h$, but we view these as two distinct copies of a single element). We consider the possibility of using DFRs for G and H to produce a DFR for $G * H$. First, suppose both G and H have *linear* representations (we emphasize here that these are the more general sort of representations discussed at the beginning of Section 2.4, rather than the special case of unitary representations) $\rho : G \to \mathrm{GL}(n, \mathbb{F})$ and $\pi : H \to \mathrm{GL}(n, \mathbb{F})$, for some $n \in \mathbb{N}_{\geq 1}$ and field $\mathbb{F}$. It is a classic result of Nisnevič [34] that if ρ and π are both P-faithful representations, then $G * H$ has a P-faithful representation $\gamma : G * H \to \mathrm{GL}(n, \mathbb{F}')$ where $\mathbb{F}'$ is some field of the same characteristic as $\mathbb{F}$. While the technique used to show Nisnevič's result [34] do not, immediately, seem to carry over to the case of *unitary* representations, the technique used by Shalen [40] to prove a certain generalization of the preceding result does directly apply to the unitary case. However, unfortunately, $\mathbb{F}'$ may be a substantially more "complex" field than $\mathbb{F}$; for example, if G and H have P-faithful representations over the field $\overline{\mathbb{Q}}$, these constructions will produce a P-faithful representation of $G * H$ over some transcendental extension of $\overline{\mathbb{Q}}$. While that latter issue of the complexity of the field is an obstacle to using DFRs of G and H to produce a DFR of $G * H$, we are still able to construct an unbounded-error DFR of $G * H$, in a certain interesting case.

Lemma 3.22. *For any $r \in \mathbb{N}_{\geq 1}$, $\mathbb{Z} * \mathbb{Z}^r$ has an unbounded-error $[1, 2]$-DFR.*

Proof. Fix r. Let $S_r = \{x_1, \ldots, x_r\}$ and let $R_r = \{[x_i, x_j] | i, j \in \{1, \ldots, r\}\}$. By Lemma 3.17, the group $A := \mathbb{Z}^r = \langle S_r | R_r \rangle$ has a P-faithful representation $\rho : A \to \mathrm{SU}(2, \mathbb{Q})$, and the group $B := \mathbb{Z} = \langle \{y\} | \rangle$ has a P-faithful representation $\pi : B \to \mathrm{SU}(2, \mathbb{Q})$. Notice that, $\forall a \in A_{\neq 1_A}$ both off-diagonal entries of the matrix $\rho(a)$ are nonzero. To see this, consider some $a \in A_{\neq 1_A}$. As $\rho(a) \in \mathrm{SU}(2)$, its two off-diagonal entries are equal in magnitude, and so they are both zero or both nonzero. If they are both zero, then $\rho(a)$ is diagonal; however, the only diagonal matrices in $\mathrm{SU}(2, \mathbb{Q})$ are $\{\pm I_2\}$, which would then imply $\rho(a) \in \{\pm I_2\} = Z(\mathrm{SU}(2))$, which contradicts the fact that ρ is P-faithful. By a symmetric argument, $\forall b \in B_{\neq 1_B}$, both off-diagonal entries of the matrix $\pi(b)$ are nonzero.

We now fundamentally follow (the proof of) Shalen [40, Proposition 1.3] to produce a P-faithful representation of $A * B \cong \mathbb{Z} * \mathbb{Z}^r$. Fix $\alpha \in ((\mathbb{R} \cap \overline{\mathbb{Q}}) \setminus \mathbb{Q})$, let $\lambda = e^{\pi i \alpha}$, and notice that, by the Gel'fond-Schneider theorem, $\lambda \notin \overline{\mathbb{Q}}$. Let $\Lambda = \operatorname{diag}[\lambda, \lambda^2]$, the 2×2 diagonal matrix with diagonal entries λ and λ^2, and observe that $\Lambda \in \mathrm{T}(2, \widetilde{E})$. Define the representation $\widehat{\rho} : A \to \mathrm{SU}(2)$ by $\widehat{\rho}(a) = \Lambda \rho(a) \Lambda^{-1}$, $\forall a \in A$. Define the representation $\gamma : A * B \to \mathrm{SU}(2)$ such that $\gamma(a) = \widehat{\rho}(a)$, $\forall a \in A$ and $\gamma(b) = \pi(b)$, $\forall b \in B$ (where γ is uniquely defined by the universal property of the free product). By Shalen [40, Proposition 1.3], γ is a P-faithful representation. Moreover, $\pi(y) \in \mathrm{SU}(2, \mathbb{Q}) \leq \mathrm{U}(2, \overline{\mathbb{Q}})$, and for each

$x_j \in S_r$, $\widehat{\rho}(x_j) = \Lambda\rho(x_j)\Lambda^{-1}$, and so $\widehat{\rho}(x_j)$ is the product of three matrices in $\mathrm{U}(2, \overline{\mathbb{Q}}) \cup \mathrm{T}(2, \widetilde{E})$. As $\{y\} \sqcup S_r$ is a generating set for $A * B$, this implies that the image of each such generator under γ is expressible as the product of at most three matrices in $\mathrm{U}(2, \overline{\mathbb{Q}}) \cup \mathrm{T}(2, \widetilde{E})$. Therefore, $\{\gamma\}$ is an unbounded-error $[1, 2]$-DFR for $A * B \cong \mathbb{Z} * \mathbb{Z}^r$. □

As defined in Section 1.2, $\Pi_1 = \{\mathbb{Z}^k | k \in \mathbb{N}\}$ denotes the finitely-generated free abelian groups, Σ_2 denotes the set of groups isomorphic to a free product of some finite set of groups in Π_1, Π_3 denotes the set of groups isomorphic to a direct product of some finite set of groups in Σ_2, and $\widehat{\Pi}_3$ denotes the set of finitely-generated groups that are virtually a (necessarily finitely-generated) subgroup of some group in Π_3.

Theorem 3.23. *Suppose $G \in \widehat{\Pi}_3$, then G virtually has an unbounded-error $[k, 2]$-DFR, for some $k \in \mathbb{N}$.*

Proof. Consider a group $H \in \Sigma_2$. Such an H is of the form $H \cong \mathbb{Z}^{r_1} * \cdots * \mathbb{Z}^{r_m}$, for some $r_1, \ldots, r_m \in \mathbb{N}$. Let $r = \max_j r_j$. Then, by a straightforward application of the Kurosh subgroup theorem, H embeds in $\mathbb{Z} * \mathbb{Z}^r$, which implies H has an unbounded-error $[1, 2]$-DFR, by Lemma 3.22. Next, consider a group $L \in \Pi_3$; such a group is of the form $L \cong H_1 \times \cdots \times H_k$, for some $H_1, \ldots, H_k \in \Sigma_2$. As all such H_j have unbounded-error $[1, 2]$-DFRs, we conclude, by an argument identical to that of Lemma 3.14, that L has an unbounded-error $[k, 2]$-DFR. Finally, for any $G \in \widehat{\Pi}_3$, G has a finitely-index subgroup K such that K is isomorphic to a finitely-generated subgroup of some $L \in \Pi_3$. As just observed, any such L has an unbounded-error $[k, 2]$-DFR, for some k, and so, by the same argument as in Lemma 3.15(i), K has an unbounded-error $[k, 2]$-DFR. We then conclude G virtually has an unbounded-error $[k, 2]$-DFR, as desired. □

4 2QCFA for the Word Problem

In this section, we use a DFR for a group G to construct a 2QCFA for the word problem of G, as well as for certain other groups related to G. Consider the group $G = \langle S|R\rangle$, with S finite. As before, we write $\Sigma = S \cup S^{-1}$ for the finite symmetric generating set of S which serves as our alphabet for the word problem of G, Σ^* for the free monoid on Σ, $\phi : \Sigma^* \to G$ for the natural map that takes each word in Σ^* to the element of G that it represents, and $l(g)$ for the length of $g \in G$ with respect to Σ, i.e., the distance from 1_G to g in the Cayley graph $\Gamma(G, \Sigma)$. Recall that the word problem of G is $W_G := W_{G=\langle S|R\rangle} = \{w \in \Sigma^* | \phi(w) = 1_G\}$. Suppose $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ is a $[k, d, \tau]$-DFR for G. As noted in Proposition 3.2, if $w \in W_G$, then $|\chi_{\rho_j}(g)| = d$, $\forall j$, and if $w \notin W_G$, then there is some j where it is not merely the case that $|\chi_{\rho_j}(g)| < d$, but, much more strongly, $|\chi_{\rho_j}(g)| \leq d - \tau(l(g))$. We will show that a 2QCFA can solve the word problem for G by checking the above condition. We will also observe that a PDFR or unbounded-error DFR for G can also be used to produce a 2QCFA for W_G, where essentially all aspects of the construction are the same, modulo some small notational details.

4.1 Computing with DFRs

We first consider a simple way in which a 2QCFA could directly compute $\rho_j(\phi(w))$, where, here, "compute $\rho_j(\phi(w))$" means placing the quantum part of the 2QCFA in a superposition that encodes $\rho_j(\phi(w))$. This initial approach will be somewhat inefficient, in terms of the size of the quantum register, and so, after describing this approach, we will then improve it. We do this both so as to present certain features of the algorithm as simply as possible, as well as to allow certain generalizations in which having a full encoding of $\rho_j(\phi(w))$ is useful (in particular, we use this encoding when considering the MO-1QFA [29]). We also note that this approach has a certain formal similarity to the "standard" way of using an "ordinary" quantum computer that is limited to pure quantum states and unitary transformations to simulate a quantum computer that is augmented with the ability to use mixed quantum states and superoperators (see, for instance, [3]).

Let $|q_1\rangle, \dots, |q_{d^2}\rangle$ denote the standard basis of $\mathbb{C}^{d^2}$. Let $\mathbb{C}^{d\times d}$ denote the vector space of $d \times d$ complex matrices (under addition); clearly, $\mathbb{C}^{d\times d} \cong \mathbb{C}^{d^2}$, as vector spaces. This vector space isomorphism is witnessed by the usual vectorization homomorphism $\mathrm{vec} : \mathbb{C}^{d\times d} \to \mathbb{C}^{d^2}$, where, for any $M \in \mathbb{C}^{d\times d}$, $\mathrm{vec}(M)$ is the vector where the first d elements are given by the first column of M, the next d elements of $\mathrm{vec}(M)$ are given by the second column of M, and so on. By interpreting an element M of (the multiplicative group) $\mathrm{U}(d)$ as an element of the vector space $\mathbb{C}^{d\times d}$ in the obvious way, we can encode M in a quantum superposition by normalizing $\mathrm{vec}(M)$. In particular, for any $M \in \mathrm{U}(d)$ we define $|M\rangle = \frac{1}{\sqrt{d}}\mathrm{vec}(M)$. Recall that, for any $M, L, R \in \mathbb{C}^{d\times d}$, $\mathrm{vec}(LMR) = (R^\top \otimes L)\mathrm{vec}(M)$, where here $\otimes$ denotes the Kronecker product. Therefore, for $M, L, R \in \mathrm{U}(d)$, $|LMR\rangle = (R^\top \otimes L)|M\rangle$ and $R^\top \otimes L \in \mathrm{U}(d^2)$. This allows us to multiply M by matrices, both on the left and on the right, by applying a corresponding unitary transformation to $|M\rangle$; though, in the following we will only require the special case in which we multiply on the left. Define the homomorphism $\gamma : \mathrm{U}(d) \to \mathrm{U}(d^2)$ such that, for any $L \in \mathrm{U}(d)$, $\gamma(L) = I_d \otimes L$, and so $\gamma(L)|M\rangle = |LM\rangle$. For $w \in \Sigma^*$, write $w = w_1, \dots, w_n$ where each $w_i \in \Sigma$. Of course, as each $w_i \in \Sigma$, $\phi(w_i) = w_i$, in that ϕ takes w_i as a formal symbol in the input alphabet Σ to the corresponding group element $w_i \in \Sigma \subseteq G$. For notational convenience, we will ignore this distinction and write w_i in place of $\phi(w_i)$, for $w_i \in \Sigma$.

Throughout this section, we will always read the string w "backwards" by initially positioning the tape head at the rightmost symbol and scanning the tape from right to left. We do this so as to respect both the convention that representations are left actions and the convention that the string w is written from left to right along the tape. As a 2QCFA can move its head in either direction, this is perfectly permissible. However, we emphasize that all results in this paper would apply equally well if one is only allowed to read the string w in the "forwards" direction, as would be the case for the MO-1QFA [29] or one-way QFA with reset [48], for example. This follows as $1_G = \phi(w) = \phi(w_1) \cdots \phi(w_n)$ precisely when $1_G = \phi(w)^{-1} = \phi(w_n)^{-1} \cdots \phi(w_1)^{-1}$. This convention motivates the following definition of a certain restricted type of subroutine of a 2QCFA A in which A is only permitted to scan the tape once, from right to left, exclusively applying unitary transformations to its quantum register along the way.

Definition 4.1. Consider a group $G = \langle S|R\rangle$, with S finite. Suppose A is a 2QCFA with quantum basis states $Q = \{q_1, \dots, q_{|Q|}\}$, $|Q| \geq 2$, quantum start state $q_1 \in Q$, and alphabet $\Sigma = S \cup S^{-1}$. Recall that, on input $w \in \Sigma^*$, the tape of A is the string $\#_L w \#_R$, where $\#_L$ and $\#_R$ are the left and right end-markers, respectively.

(i) Suppose $|\psi_1\rangle = \sum_h \alpha_h |q_h\rangle$ and $|\psi_2\rangle = \sum_h \beta_h |q_h\rangle$ are two, not necessarily distinct, superpositions of the basis states Q, where all $\alpha_h, \beta_h \in \overline{\mathbb{Q}}$. There are (many) $T \in \mathrm{U}(|Q|, \overline{\mathbb{Q}})$ such that $T|\psi_1\rangle = |\psi_2\rangle$. Let $T_{|\psi_1\rangle \to |\psi_2\rangle}$ denote an arbitrary such T.

(ii) Suppose $\pi : G \to \mathrm{U}(|Q|)$ is a group homomorphism, and $|\psi\rangle = \sum_h \alpha_h |q_h\rangle$ a superposition of the basis states Q, where $\alpha_h \in \overline{\mathbb{Q}}$, $\forall h$. Then the *unitary round* $\mathcal{U}(\pi, |\psi\rangle)$ is a particular subcomputation of A on w, defined as follows. The round begins with the quantum register in the superposition $|q_1\rangle$ and the tape head at the right end of the tape. On reading $\#_R$, A performs the unitary transformation $T_{|q_1\rangle \to |\psi\rangle}$ to its quantum register, and moves its head to the left. On reading a symbol $\sigma \in \Sigma$ (which is interpreted as the element $\phi(\sigma) = \sigma \in \Sigma \subseteq G$), A performs the unitary transformation $\pi(\sigma)$ to the quantum register and moves its head left. When the tape head first reaches the left end of the tape (i.e., the first time the symbol $\#_L$ is read), A performs the identity transformation to its quantum register, and does not move its head, at which point the round ends. As $\phi : \Sigma^* \to G$ is a (monoid) homomorphism and $\pi : G \to \mathrm{U}(|Q|)$ is a (group) homomorphism, we immediately conclude that, at the end of the round, the quantum register is in the superposition $\pi(\phi(w))|\psi\rangle$.

We now observe that a 2QCFA can compute $\rho_j(\phi(w))$ in a single unitary round, in the sense that, when the round is over, the 2QCFA will have $\rho_j(\phi(w))$ encoded in its quantum register. We consider a 2QCFA A with quantum basis states $q_1, \ldots, q_{d^2}$ and quantum start state q_1. Then, after the round $\mathcal{U}(\gamma \circ \rho_j, |I_d\rangle)$, the quantum register will be in the superposition $\gamma(\rho_j(\phi(w)))|I_d\rangle = |\rho_j(\phi(w))\rangle$. Moreover, $|\chi_{\rho_j}(\phi(w_1, \ldots, w_n))| = |\mathrm{Tr}(\rho_j(\phi(w_1, \ldots, w_n)))|$ is also easily obtainable from this encoding as there is a $T \in \mathrm{U}(d^2)$ for which the first entry of $T|M\rangle$ is $\frac{1}{d}\mathrm{Tr}(M)$, for any $M \in \mathrm{U}(d)$. We leave the straightforward details of completing this procedure to the reader, as we will now consider a more efficient encoding.

In particular, the above procedure is rather wasteful, in that it explicitly encodes $\rho_j(\phi(w))$ as the quantum superposition $|\rho_j(\phi(w))\rangle$, which requires d^2 quantum states. Of course, we only wish to determine if $|\chi_{\rho_j}(\phi(w))| = d$, which does not require explicitly computing $\rho_j(\phi(w))$. Fundamentally, the idea is that, as ρ_j is a d-dimensional unitary representation, we can obtain the needed information about $\rho_j(\phi(w))$ by simply applying it to appropriately chosen $|\psi\rangle \in \mathbb{C}^d$, and such $|\psi\rangle$ can be stored using only d quantum states. Define the subset $Y_j \subseteq G$ to be the elements of G strongly separated from 1_G by ρ_j, i.e., $Y_j = \{g \in G \,||\chi_{\rho_j}(g)| \leq d - \tau(l(g))\}$. By definition, $\cup_j Y_j = G_{\neq 1_G}$, though the Y_j are not necessarily disjoint. In particular, $\forall j$, we have $1_G \notin Y_j$. For $n \in \mathbb{N}$, let Σ^n denote all words in Σ^* of (string) length exactly n (i.e., all sequences $\sigma_1, \ldots, \sigma_n \in \Sigma$). Notice that $l(\phi(w)) \leq n$, and so, as τ is monotone non-increasing, $\tau(l(\phi(w))) \geq \tau(n)$. We will show that, given some $w \in \Sigma^n$ as input, a 2QCFA can make a constant number (i.e., independent of n) right-to-left passes over the input, and perform a constant number of quantum measurements such that, if $\phi(w) \in Y_j$, then with non-negligible probability (related to $\tau(n)$), the results of those measurements will allow the machine to conclude, with certainty, that $\phi(w) \neq 1_G$. This motivates the following extension of Definition 4.1(ii), in which we allow the 2QCFA to perform quantum measurements, in a certain restricted way.

Definition 4.2. Using the notation of Definition 4.1, let $B = \{B_0, B_1\}$ be the partition of $\{1, \ldots, |Q|\}$ given by $B_0 = \{2, \ldots, |Q|\}$ and $B_1 = \{1\}$.

(i) For $M \in \mathrm{U}(|Q|)$, a *measurement round* $\mathcal{M}(\pi, |\psi\rangle, M)$ is a sub-computation of A that begins with the unitary round $\mathcal{U}(\pi, |\psi\rangle)$. Then A performs the unitary transformation M, and does not move its head. After which A performs the quantum measurement specified by B, producing the result $r \in \{0, 1\}$, A records r in its classical state, and does not move its head, at which point the round is over. The *result* of the measurement round is the result r of the quantum measurement. If $r = 1$, then the quantum register is in the superposition $|q_1\rangle$ and if $r = 0$ then the quantum register is in some superposition of the form $\sum_{h>1} \alpha_h |q_h\rangle$.

(ii) A *reset* consists of A moving its head directly to the right end of the tape, without altering its quantum register. That is to say, when reading $\#L$ or any $\sigma \in \Sigma$, A must perform the identity transformation on its quantum register and move its head one step to the right. When $\#R$ is encountered for the first time, A must again perform the identity transformation on its quantum register and A must not move its head, after which the reset is complete.

(iii) For $p \in \mathbb{N}_{\geq 1}$, a $[\leq p]$-pass measurement round of A on input w consists of A performing at most p measurement rounds, where the overall result is the AND of the results of individual measurement rounds, and which stops as soon as any result of 0 is obtained. Formally, we define a $[\leq p]$-pass measurement round $\mathcal{M}\left[(\pi_1, |\psi_1\rangle, M_1), \ldots, (\pi_p, |\psi_p\rangle, M_p)\right]$ as follows. Initialize a counter $j = 1$ (A keeps track of j using its classical states). A repeatedly does the following: A performs the measurement round $\mathcal{M}(\pi_j, |\psi_j\rangle)$ producing the result r_j, if $r_j = 0$ or $j = p$, we are done and the result is r_j, otherwise (in particular, notice this requires $r_j = 1$ and so the quantum register is $|q_1\rangle$) A increments the counter to $j+1$, performs a reset, and continues (and of course does *not* continue to remember r_j).

We now show that a 2QCFA can distinguish w for which $\phi(w) \in Y_j$ from w for which $\phi(w) = 1_G$.

Lemma 4.3. *Suppose $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ is a $[k, d, \tau]$-DFR (or algebraic PDFR) for a group $G = \langle S|R\rangle$, with S finite. Fix any j, consider the representation ρ_j, and define $Y_j = \{g \in G \big| |\chi_{\rho_j}(g)| \leq d - \tau(l(g))\}$, as above.*

(a) If $\mathcal{F}$ is a diagonal DFR, then there is a 2QCFA A, with only d quantum basis states, such that, $\forall w \in \Sigma^n$, the result $r \in \{0,1\}$ of a single measurement round of A on w satisfies the following.

(i) (Perfect Completeness) If $\phi(w) = 1_G$, then $\Pr[r=1] = 1$.

(ii) (Soundness) If $\phi(w) \in Y_j \subseteq G_{\neq 1_G}$, then $\Pr[r=0] \geq \frac{\tau(n)}{d}$.

(b) Otherwise, there is a 2QCFA B, with only d quantum basis states, such that, $\forall w \in \Sigma^n$, the result $r \in \{0,1\}$ of a $[\leq (d+1)]$-pass measurement round of B on w satisfies the following.

(i) (Perfect Completeness) If $\phi(w) = 1_G$, then $\Pr[r=1] = 1$.

(ii) (Soundness) If $\phi(w) \in Y_j \subseteq G_{\neq 1_G}$, then $\Pr[r=0] \geq \left(\frac{\tau(n)}{2d(d-1)}\right)^2$.

Moreover, in either case, all transition amplitudes of the 2QCFA will belong to $\overline{\mathbb{Q}} \cup E$, where E is the collection of entries of the family of matrices $\rho_j(s)$, as s varies over S.

Proof. First, suppose $\mathcal{F}$ is a DFR. Recall that the definition of a DFR requires $d \geq 2$. Fix an orthonormal basis $|q_1\rangle, \ldots, |q_d\rangle$ of $\mathbb{C}^d$, and let $|1\rangle = \frac{1}{\sqrt{d}}\sum_j |q_j\rangle$. Fix any $F \in \mathrm{U}(d, \overline{\mathbb{Q}})$ such that all entries in the first row of F are equal to $\frac{1}{\sqrt{d}}$. For concreteness, we take F as the usual (unitary) $d \times d$ DFT matrix, i.e., set $\omega = e^{-\frac{2\pi i}{d}}$, then, for any $u, v \in \{1, \ldots, d\}$, the (u,v) entry of F is given by $F_{u,v} = \frac{1}{\sqrt{d}}\omega^{(u-1)(v-1)}$. Notice that, for any $M \in \mathrm{U}(d)$, if $|\psi\rangle := FM|1\rangle$, then $|\psi\rangle = \left(\frac{1}{d}\sum_{u,v} M_{u,v}\right)|q_1\rangle + \sum_{h>1}\alpha_h|q_h\rangle$, for some $\alpha_2, \ldots, \alpha_d \in \mathbb{C}$. If, moreover, M is a diagonal matrix, then $|\psi\rangle = \frac{1}{d}\operatorname{Tr}(M)|q_1\rangle + \sum_{h>1}\alpha_h|q_h\rangle$.

First, we consider the case in which $\mathcal{F}$ is a diagonal DFR. The 2QCFA A will have the d quantum basis states $Q = \{q_1, \ldots, q_d\}$, and q_1 will be its quantum start state. A performs the measurement round $\mathcal{M}(\rho_j, |1\rangle, F)$, producing the result r, which we now show has the claimed properties. Immediately before performing the quantum measurement, the quantum register is in the superposition

$$F\rho_j(\phi(w))|1\rangle = \frac{1}{d}\chi_{\rho_j}(\phi(w))|q_1\rangle + \sum_{h>1}\alpha_h|q_h\rangle,$$

for some $\alpha_2, \ldots, \alpha_d \in \mathbb{C}$. By definition, $\Pr[r=1] = |\frac{1}{d}\chi_{\rho_j}(\phi(w))|^2$. If $\phi(w) = 1_G$, then $\chi_{\rho_j}(\phi(w)) = d$, and so $\Pr[r=1] = 1$, as desired. If $\phi(w) \in Y_j$, then

$$|\chi_{\rho_j}(\phi(w))| \leq d - \tau(l(\phi(w))) \leq d - \tau(n).$$

This immediately implies,

$$\Pr[r=0] \geq 1 - \frac{1}{d^2}(d - \tau(n))^2 = 2\frac{\tau(n)}{d} - \left(\frac{\tau(n)}{d}\right)^2 \geq \frac{\tau(n)}{d},$$

where the last inequality follows from the fact that $\tau(n) \leq d$, which shows that the result of the measurement round performed by A has the claimed parameters.

Next, we consider the case in which $\mathcal{F}$ is not diagonal. The 2QCFA B will also have the d quantum basis states $Q = \{q_1, \ldots, q_d\}$, with q_1 its quantum start state. For each, $v \in \{1, \ldots, d\}$, let $P_v \in \mathrm{U}(d, \overline{\mathbb{Q}})$ denote an arbitrary permutation matrix with a 1 in entry $(1, v)$. B performs the $[\leq (d+1)]$-pass measurement round $\mathcal{M}\left[(\rho_j, |1\rangle, F), (\rho_j, |q_1\rangle, P_1), (\rho_j, |q_2\rangle, P_2), \ldots, (\rho_j, |q_d\rangle, P_d)\right]$, producing the result r. To see that r has the claimed properties, let $M = \rho_j(\phi(w))$. By the above,

$$FM|1\rangle = \frac{1}{d}\sum_{u,v} M_{u,v}|q_1\rangle + \sum_{h>1} \alpha'_h |q_h\rangle,$$

for some $\alpha'_2, \ldots, \alpha'_d \in \mathbb{C}$. Fix w such that $\phi(w) \in Y_j$, and define δ to be the maximum, taken over u, v such that $u \neq v$, of $|M_{u,v}|$. We wish to bound $|\alpha'_1| = |\frac{1}{d}\sum_{u,v} M_{u,v}|$. We have,

$$\left|\frac{1}{d}\sum_{u,v} M_{u,v}\right| = \frac{1}{d}\left|\mathrm{Tr}(M) + \sum_{\substack{u,v \\ u\neq v}} M_{u,v}\right| \leq \frac{1}{d}\left(|\mathrm{Tr}(M)| + \sum_{\substack{u,v \\ u\neq v}} |M_{u,v}|\right) \leq \frac{1}{d}\left(d - \tau(n) + \delta d(d-1)\right).$$

In particular, if $\delta \leq \frac{\tau(n)}{2d(d-1)}$, then, letting r_1 denote the result of the first quantum measurement performed by B, we have

$$\Pr[r = 0] \geq \Pr[r_1 = 0] = 1 - \left|\frac{1}{d}\sum_{u,v} M_{u,v}\right|^2 \geq 1 - \left(1 - \frac{\tau(n)}{2d}\right)^2 \geq \frac{\tau(n)}{2d}.$$

As this proves the claimed bound, assume for the remainder of the proof that $\delta \geq \frac{\tau(n)}{2d(d-1)}$, and so there is some u', v', $u' \neq v'$ such that $|M_{u',v'}| \geq \frac{\tau(n)}{2d(d-1)}$. Notice that $P_{v'} M |q_{v'}\rangle$ is of the form $\sum_h \beta_h |q_h\rangle$ where the β_h are a permutation of the entries in column v' of M. In particular, $\beta_1 = M_{v',v'}$, so there is some $h > 1$ such that $\beta_h = M_{u',v'}$. Let p_{v+1} denote the probability that B performs the $(v+1)^{\text{th}}$ quantum measurement (recall that a multiple pass measurement round will stop as soon as a result of 0 is obtained) and let r_{v+1} denote the result of that measurement, assuming that it is performed. Then

$$\Pr[r = 0] \geq (1 - p_{v+1}) + \Pr[r_{v+1} = 0]p_{v+1} \geq \Pr[r_{v+1} = 0] = \sum_{h>1} |\beta_h|^2 \geq |M_{u',v'}|^2 \geq \left(\frac{\tau(n)}{2d(d-1)}\right)^2.$$

Therefore, we have shown that, $\forall w$ such that $\phi(w) \in Y_j$, the result r produced by B satisfies the claimed lower bound on $\Pr[r = 0]$. All that remains is to observe that, if $\phi(w) = 1_G$, then each of the measurements possibly performed by B will always have value 1. This follows from the fact that $\rho_j(\phi(w)) = I_d$, and so, when each of the above quantum measurements are performed, the quantum register of B is in the superposition $|q_1\rangle$. Therefore, the multiple pass measurement round performed by B has the claimed parameters.

The claim concerning the transition amplitudes immediately follows from the fact that the only unitary transformations performed by either A or B fall in two classes: (1) those of the type $\rho_j(\sigma)$, for $\sigma \in \Sigma$ (note that, for $s \in S$, $\rho_j(s^{-1}) = \rho_j(s)^{-1} = \rho_j(s)^\dagger$, where $\dagger$ denotes conjugate-transpose), whose transition amplitudes belong to E by definition, or (2) a transformation $T_{|q_1\rangle \to 1}$, $T_{|q_1\rangle \to |q_v\rangle}$, F, I_d, or P_v, all of whose transition amplitudes are clearly algebraic numbers.

If, instead, $\mathcal{F}$ is an algebraic PDFR, simply take arbitrary lifts $\widehat{\rho}_j : G \to \mathrm{U}(d, \overline{\mathbb{Q}})$ (functions, not necessarily homomorphisms, such that the composition of the natural projection $\mathrm{U}(d) \to \mathrm{PU}(d)$ with each $\widehat{\rho}_j$ yields ρ_j), and apply the above argument to the $\widehat{\rho}_j$. □

Remark. The particular parameters used to define "soundness" above can straightforwardly be improved. However, as we view d as a small constant (in particular, we are most interested in the case when $d = 2$), such improvements would not affect the asymptotics of any final results, so we do not pursue them here.

Similarly, in the unbounded-error case, we have the following, by a proof precisely analogous to that of Lemma 4.3(b) above.

Lemma 4.4. *Consider a group $G = \langle S|R \rangle$, with S finite and suppose $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ is a set of representations $\rho_j : G \to \mathrm{U}(d)$. Then there is a 2QCFA B, with only d quantum basis states, such that, $\forall w \in \Sigma^*$, the result $r \in \{0, 1\}$ of a $[\leq (d + 1)]$-pass measurement round of B on w satisfies the following.*

(i) (Perfect Completeness) If $\phi(w) = 1_G$, then $\Pr[r = 1] = 1$.

(ii) (Soundness) If $\phi(w) \notin \mathrm{Pker}(\rho_j)$, then $\Pr[r = 0] > 0$.

Moreover all transition amplitudes of the 2QCFA will belong to $\overline{\mathbb{Q}} \cup E$, where E is the collection of entries of an appropriate finite factorization of each matrix the family of matrices $\rho_j(s)$, as s varies over S. Furthermore, the above claims all also hold if instead $\mathcal{F}$ is a set of projective representations.

4.2 Constructions of 2QCFA for Word Problems

Now, by combining the results of the previous section, the constructions of DFRs from Section 3.3, and standard techniques from computational group theory, we show that 2QCFA can solve the word problem for a wide class of groups. We first show that a DFR (with appropriate parameters) for a group G can be used to produce a 2QCFA for W_G, where for ease of exposition we split this into two cases according to the parameters of the DFR. We then show that, if H is a finite-index subgroup of G, a 2QCFA for W_H can be used to produce a 2QCFA for W_G. Finally, we show, for many groups, there is a 2QCFA that recognizes its word problem. We use the notation of Definition 1.1 when stating the parameters of the 2QCFA.

Lemma 4.5. *Suppose $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ is a E-diagonal $[k, d, C_1 n^{-C_2}]$-DFR for a group $G = \langle S|R \rangle$, with S finite, and $C_1, C_2 \in \mathbb{R}_{>0}$. Then $\forall \epsilon \in \mathbb{R}_{>0}$, there is a $[\epsilon, n^{\lceil C_2 \rceil + 2}, d, \overline{\mathbb{Q}} \cup E]$-2QCFA A that recognizes $W_G := W_{G=\langle S|R \rangle}$.*

Proof. Define the subsets $Y_j \subseteq G_{\neq 1_G}$ as in Lemma 4.3, and recall that $G_{\neq 1_G} = \cup_j Y_j$. The 2QCFA A will recognize W_G by running the subroutine of Lemma 4.3(a), for each j. If $\phi(w) \neq 1_G$, then, for at least some j, this subroutine will, with sufficient probability, produce a result that allows one to conclude with certainty, that $\phi(w) \neq 1_G$, at which point A will immediately reject. To assure that w for which $\phi(w) = 1_G$ are accepted, A will periodically run a subroutine that accepts with some small probability and continues otherwise, using the technique from Ambainis and Watrous [2]. In particular, for $m, y \in \mathbb{N}$, let $\mathcal{R}(m, y)$ denote the subroutine that, on an input of length $n \in \mathbb{N}$ produces a result $b \in \{0, 1\}$, where $\Pr[b = 1] = (n + 1)^{-m} 2^{-y}$, within expected running time $O(n^2)$ (see [2] for details; in brief, if the 2QCFA starts with its head over the first symbol to the right of $\#L$ and performs an unbiased one-dimensional random walk along the tape until either of the end-markers are encountered, then the probability that $\#R$ is the first end-marker encountered is $(n + 1)^{-1}$; by repeating this procedure m times, and generating unbiased random bits y times, the desired b can be produced).

We now fill in the details. A has the quantum basis states $|q_1\rangle, \ldots, |q_d\rangle$, where q_1 is the quantum start state. A performs the following procedure.

Use the classical states to store a counter $j \in \{1, \ldots, k\}$, initialized to 1

```
Repeat indefinitely:
    Move the head to the right end of the tape, leaving the quantum register unchanged
    Run the subroutine of Lemma 4.3(a) with $\rho_j$ producing the result $r$
    If $r = 0$ then reject
    Add 1 to $j$, where the addition is performed modulo $k$
    If $j = k$ then
        Run the subroutine $\mathcal{R}(\lceil C_2 \rceil, \lceil \log(\frac{\epsilon C_1}{d}) \rceil)$, giving the result $b$
        If $b = 1$ then accept
```

We now show that A has the claimed parameters. Clearly, A has d basis states and the transition amplitudes of A belong to $\overline{\mathbb{Q}} \cup E$. To see the remaining claims, fix a string w and let n denote its (string) length. Consider a subcomputation of the above computation of A that begins when the counter $j = 1$ and A is at the beginning of the "Repeat indefinitely" loop, and ends as soon as A accepts or rejects, or after k complete iterations of the "Repeat indefinitely" loop. Let p_{acc} and p_{rej} denote, respectively, the probability that such a subcomputation ends with A accepting or rejecting. Let E_j denote the event that such a subcomputation actually runs the subroutine of Lemma 4.3(a) with ρ_j (note that the only way this does not happen is if A has already rejected for some $\widetilde{j} < j$), let p_j denote the probability that E_j occurs, and let r_j denote the result produced by this subroutine, if E_j occurs. Notice that

$$\Pr[b = 1 | E_k] = 2^{-\left\lceil \log(\frac{\epsilon C_1}{d}) \right\rceil}(n+1)^{-\lceil C_2 \rceil} > 0.$$

First, suppose $w \notin W_G$. There is at least one j' such that $\phi(w) \in Y_{j'}$. Therefore, when the counter $j = j'$, Lemma 4.3(a)(ii) guarantees that $\Pr[r_{j'} = 0 | E_{j'}] \geq \frac{C_1}{d} n^{-C_2}$. Notice that the event that A rejects in such a subcomputation is the (disjoint) union of the event A rejects before step j' (i.e., $E_{j'}$ does not occur) and the event A rejects at step j' or later. Therefore,

$$p_{rej} = (1 - p_{j'})1 + \sum_{j \geq j'} p_j \Pr[r_j = 0 | E_j] \geq (1 - p_{j'}) + p_{j'} \Pr[r_{j'} = 0 | E_{j'}] \geq \Pr[r_{j'} = 0 | E_{j'}] \geq \frac{C_1}{d} n^{-C_2}.$$

We also have

$$p_{acc} = p_k \Pr[b = 1 | E_k] < \Pr[b = 1 | E_k] = 2^{-\left\lceil \log(\frac{\epsilon C_1}{d}) \right\rceil}(n+1)^{-\lceil C_2 \rceil} \leq \epsilon \frac{C_1}{d}(n+1)^{-\lceil C_2 \rceil} \leq \epsilon p_{rej}.$$

As we repeat such subcomputations until A either accepts or rejects, we have

$$\Pr[A \text{ rejects } w | w \notin W_G] = \frac{p_{rej}}{p_{acc} + p_{rej}} \geq \frac{p_{rej}}{\epsilon p_{rej} + p_{rej}} = \frac{1}{1 + \epsilon} \geq 1 - \epsilon.$$

Next, instead suppose $w \in W_G$. Then Lemma 4.3(a)(i) guarantees that every use of the subroutine of Lemma 4.3(a) will produce $r = 1$. This implies $p_{rej} = 0$, $p_k = 1$, and

$$p_{acc} = p_k \Pr[b = 1 | E_k] = 2^{-\left\lceil \log(\frac{\epsilon C_1}{d}) \right\rceil}(n+1)^{-\lceil C_2 \rceil} \geq \epsilon \frac{C_1}{2d}(n+1)^{-\lceil C_2 \rceil} > 0.$$

As we repeat such subcomputations until A either accepts or rejects, we have

$$\Pr[A \text{ accepts } w | w \in W_G] = \frac{p_{acc}}{p_{acc} + p_{rej}} = 1.$$

This completes the proof of the claim that A recognizes W_G with one-sided error ϵ. Lastly, to see that A has the claimed expected running time, let p_{halt} denote the probability that any given subcomputation of the above form ends with A halting (i.e., accepting or rejecting). When $w \in W_G$,

$$p_{halt} = p_{acc} + p_{rej} \geq \epsilon \frac{C_1}{2d}(n+1)^{-\lceil C_2 \rceil}.$$

When $w \notin W_G$,

$$p_{halt} = p_{acc} + p_{rej} \geq p_{rej} \geq \frac{C_1}{2d} n^{-C_2} \geq \epsilon \frac{C_1}{2d}(n+1)^{-\lceil C_2 \rceil}.$$

Therefore the expected number of executions of such subcomputations is $O(n^{\lceil C_2 \rceil})$. Each subcomputation of the above form consists of at most k passes through the "Repeat indefinitely" loop. Each pass involves a single use of the subroutine of Lemma 4.3(a), which runs in time $O(n)$; additionally, the pass in which the counter $j = k$ also involves a single use of the subroutine $\mathcal{R}$, which runs in time $O(n^2)$. Therefore, A runs in expected time $O(n^{\lceil C_2 \rceil + 2})$, as desired. □

Lemma 4.6. *Suppose $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ is an algebraic $[k, d, C^{-n}]$-DFR (or PDFR) for a group $G = \langle S|R \rangle$, with S finite, and $C \in \mathbb{R}_{\geq 1}$. Then $\forall \epsilon \in \mathbb{R}_{>0}$, there is a $[\epsilon, K^n, d, \overline{\mathbb{Q}}]$-2QCFA A that recognizes $W_G := W_{G=\langle S|R \rangle}$, for some constant $K = K(C) \in \mathbb{R}_{\geq 1}$.*

Proof. We proceed almost exactly as in the proof of Lemma 4.5, with the only modification arising from the fact that the substantially weaker bound on the parameter τ of the DFR has a corresponding decrease in the probability that the subroutine of Lemma 4.3 can distinguish w with $|\chi_{\rho_j}(\phi(w))| = d$ from w with $|\chi_{\rho_j}(\phi(w))| \neq d$. As before, A will periodically run a subroutine that accepts with some small probability, though the above issue requires that this is done with a substantially smaller probability than in the proof of Lemma 4.5.

A has the quantum basis states $|q_1\rangle, \ldots, |q_d\rangle$, where q_1 is the quantum start state. For $p \in \overline{\mathbb{Q}} \cap [0, 1]$, let $\mathcal{B}(p)$ denote the subroutine that produces a biased random Boolean value x, such that $\Pr[x = 1] = p$, which operates as follows. We start with the quantum register in the superposition $|q_1\rangle$. Let $|\psi\rangle = \sqrt{p}|q_1\rangle + \sqrt{1-p}|q_2\rangle$. We then perform the unitary transformation $T_{|q_1\rangle \to |\psi\rangle}$, followed by the quantum measurement with respect to the partition $B_0 = \{2 \ldots, d\}, B_1 = \{1\}$. The result 1 occurs with probability p. If the result is 0, we then perform the unitary transformation $T_{|q_2\rangle \to |q_1\rangle}$ to return the quantum register to the superposition $|q_1\rangle$. The head of the 2QCFA does not move during this subroutine.

For $p \in \overline{\mathbb{Q}} \cap [0, 1]$, $y \in \mathbb{N}$, let $\mathcal{R}'(p, y)$ denote the subroutine that, on an input of length $n \in \mathbb{N}$ produces a result $b \in \{0, 1\}$, where $\Pr[b = 1] = p^n 2^{-y}$, and has running time $O(n)$. $\mathcal{R}'(p, y)$ operates by scanning the tape once, from left to right. On symbols other than the end-markers, $\mathcal{B}(p)$ is run; if the result is 0, the subroutine immediately halts with the result of 0, otherwise it continues reading the next symbol. When the right end-marker $\#R$ is encountered, the subroutine generates up to y unbiased bits, one after the other. If any of these bits are 0, the subroutine immediately halts with the result of 0; if all y bits are 1, the subroutine halts with the result of 1. Notice that the transition amplitudes needed to implement $\mathcal{R}'$ are all algebraic numbers.

A performs the following procedure.

```
Use the classical states to store a counter $j \in \{1, \ldots, k\}$, initialized to 1
Repeat indefinitely:
    Move the head to the right end of the tape, leaving quantum register unchanged
    Run the subroutine of Lemma 4.3(b) with $\rho_j$ producing the result $r$
    If $r = 0$ then reject
    Add 1 to $j$, where the addition is performed modulo $k$
    If $j = k$ then
        Run the subroutine $\mathcal{R}'(\frac{1}{\lceil C^2 \rceil}, \lceil \log(\frac{\epsilon}{4d^4}) \rceil)$, giving the result $b$
        If $b = 1$ then accept
```

All remaining parts of the proof are identical to that of Lemma 4.5, and so we omit the details. □

We now show that a 2QCFA for W_G can be constructed from a 2QCFA for W_H, if H is a finite-index normal subgroup of G.

Lemma 4.7. *Consider a group $H = \langle S_H | R_H \rangle$, with S_H finite, and suppose that A_H is a 2QCFA that recognizes W_H, which operates in the manner of Lemma 4.5 or Lemma 4.6. Further suppose G is a group such that $H \leq G$ and $[G : H]$ is finite. Then G admits a presentation $G = \langle S_G | R_G \rangle$, with S_G finite, such that there is a 2QCFA A_G that recognizes W_G. Moreover, A_G has the same acceptance criteria, asymptotic expected running time, number of quantum basis states, and class of transition amplitudes as A_H.*

Proof. Following (essentially) [30] (with the exception that we do not assume H is a *normal* subgroup of G), we now construct a convenient presentation for G. We begin by establishing some notation. Let $l = [G : H]$, and let $g_1, \ldots, g_l$ denote a complete family of left coset representatives of H in G, where $g_1 = 1_G$. We assume for notational convenience that $S_H \cap S_H^{-1} = \emptyset$ (and so, in particular, $1_H \notin S_H$). Let $\Sigma_H = S_H \sqcup S_H^{-1}$, $S_G = S_H \sqcup (g_2, \ldots, g_l)$, and $\Sigma_G = S_G \cup S_G^{-1}$. Let $\phi_H : \Sigma_H^* \to H$ and $\phi_G : \Sigma_G^* \to G$ be the natural maps. Let $T_l = \{1, \ldots, l\}$.

As the g_i are a complete family of left coset representatives of H in G, every element $g \in G$ can be expressed uniquely as some $g_i h$, where $i \in T_l$ and $h \in H$. In particular, for any $\sigma \in \Sigma_G$ and $j \in T_l$, consider the element $\sigma g_j \in G$; there is unique $i \in T_l$ and $h \in H$ such that $\sigma g_j = g_i h$. Therefore, we can define functions $\alpha : \Sigma_G \times T_l \to T_l$ and $\beta : \Sigma_G \times T_l \to H$, such that

$$\sigma g_j = g_{\alpha(\sigma,j)} \beta(\sigma, j), \ \forall \sigma \in \Sigma_G, \forall j \in T_l.$$

Let $\tau : H \to F(S_H)$ be the function that takes each $h \in H$ to some element in the free group on S_H such that $h = \tau(h)$, as elements of H. Then G has presentation $\langle S_G | R_G \rangle$, where S_G is as defined above and

$$R_G = R_H \cup \left\{ g_{\alpha(\sigma,j)} \tau(\beta(\sigma, j)) g_j^{-1} \sigma^{-1} | \sigma \in \Sigma_G, j \in T_l \right\}.$$

We now construct a 2QCFA A_G that recognizes $W_G := W_{G = \langle S_G | R_G \rangle}$. Consider an input $w \in \Sigma_G^*$, and let $|w|$ denote the (string) length of w, i.e., $w = w_1 \cdots w_{|w|}$, where each $w_i \in \Sigma_G$. For any $p \in \{0, \ldots, |w|\}$, let $w^p = w_{|w|-p+1} \cdots w_{|w|}$ denote the suffix of w of length p; in particular, w^0 is the empty string. A_G must determine if $\phi_G(w) = 1_G = g_1 1_H$. The key idea is that A_G will make many right-to-left passes over its input, such that, after A_G has read the suffix w^p, if $\phi_G(w^p) = g_m h$, then A_G will have the values $m \in T_l$ and $h \in H$ "stored" in its internal state, in an appropriate sense. Namely, A_G will keep track of $m \in T_l$ using its classical states, and A_G will keep track of h by simulating A_H.

We now fill in the details. A_G has the same quantum basis states as A_H, which we will denote $|q_1\rangle, \ldots, |q_d\rangle$, and quantum start state q_1. A_G begins by moving its head to the far right end of the tape, leaving its quantum register in the superposition $|q_1\rangle$. A_G will store a value $t \in T_l$ using its classical states, where t is initialized to 1. A_G then repeatedly scans its input in the manner prescribed by A_H, i.e., A_G makes many right-to-left passes reading the input word w, and A_G also performs the simulated coin flipping via random walks of A_H. During each right-to-left pass, A_G will maintain the property that after reading the suffix w^p, if $\phi_G(w^p) = g_m h$, then the stored value $t = m$ and A_N will have been simulated on a string $\widehat{w^p} \in \Sigma_H^*$ (read "backwards"), where $\phi_H(\widehat{w^p}) = h$.

A_G accomplishes this as follows. Suppose A_G has already read the particular suffix w_p and $\phi_G(w^p) = g_m h$, and is now about to read the next symbol, $\sigma := w_{|w|-p}$. After reading σ, we want A_G to update its internal state (both classical and quantum) to correspond to the word $w^{p+1} = \sigma \circ w^p$. By construction, $\sigma g_m = g_{\alpha(\sigma,m)} \beta(\sigma, m)$, and so

$$\phi_G(w^{p+1}) = \phi_G(\sigma \circ w^p) = \phi_G(\sigma)\phi_G(w^p) = \sigma g_m h = g_{\alpha(\sigma,m)} \beta(\sigma, m) h.$$

Define the function $\widehat{\beta} : \Sigma_G \times T_l \to \Sigma_H^*$ such that $\widehat{\beta}(\kappa, j)$ is any word in Σ_H^* of minimum (string) length such that $\phi_N(\widehat{\beta}(\kappa, j)) = \beta(\kappa, j)$, $\forall \kappa \in \Sigma_G, \forall j \in T_l$. A_G then updates its stored value $t \in T_l$ from m

to $\alpha(\sigma, m)$ and simulates A_H on $\widehat{\beta}(\sigma, m)$. That is to say, at this point A_H has been simulated on the string $\widehat{w^p}$, where $\phi_H(\widehat{w^p}) = h$; A_G then feeds the string $\widehat{\beta}(\sigma, m)$ to A_H (from right-to-left), after which A_H will have been simulated on $\widehat{\beta}(\sigma, m) \circ \widehat{w^p}$, as desired. During this process of feeding the string $\widehat{\beta}(\sigma, m)$ to A_H, A_G does not move its head.

All that remains is to define the acceptance criteria of A_G. Suppose A_G has just made a complete pass over the input, simulating A_H along the way, and then possibly also performed a simulated coin-flipping procedure, if A_H so demanded. A_G also has the value m in its internal state, such that $\phi_G(w) = g_m h$. At this point (the simulation of) A_H may or may not have halted. A_G behaves as follows. If $m \neq 1$, A_G immediately rejects. If $m = 1$, then if A_H has halted (accepting or rejecting the input), then A_G halts, accepting if A_H accepted and rejecting if A_H rejected. If $m = 1$ and A_H has not halted, A_G continues. It immediately follows from the above argument that A_G recognizes W_G and that A_G has all the claimed properties. □

Using the above Lemma, and the constructions of DFR and PDFR from Section 3, the main theorems stated in the introduction concerning the bounded-error recognizability of the word problem straightforwardly follow.

Proof of Theorem 1.2. By Corollary 3.16.1, there is an effectively computable constant $C \in \mathbb{R}_{>0}$ such that, for any finitely-generated virtually abelian group G, claims Corollary 3.16(ii) and Corollary 3.16(iii) hold. Fix such a group G. By Corollary 3.16(ii), G virtually has a diagonal algebraic $[K_1, 2, D_1 n^{-C}]$-DFR. By Corollary 3.16(iii), with $\delta = 0.9$, G also virtually has a $\widetilde{E}$-diagonal $[K_2, 2, D_2 n^{-0.9}]$-DFR. By Lemma 4.7, we conclude that, for any $\epsilon \in \mathbb{R}_{>0}$, W_G is recognized by a $[\epsilon, n^C, 2, \overline{\mathbb{Q}}]$-2QCFA, as well as by a $[\epsilon, n^3, 2, \widetilde{\mathbb{C}}]$-2QCFA. □

Remark. We note that the limiting factor on the expecting running time of the $[\epsilon, n^3, 2, \widetilde{\mathbb{C}}]$-2QCFA for W_G is not the difficulty of distinguishing strings in W_G from strings not in W_G, but is instead the difficulty of producing an appropriately biased Boolean random variable. In particular, by Corollary 3.16(iii), any such G virtually has a $\widetilde{E}$-diagonal $[K_2, 2, D_2 n^{-\delta}]$-DFR, for arbitrarily small $\delta > 0$. However, while we can use the random walk technique of Ambainis and Watrous [2] to produce a Boolean random variable b such that $\Pr[b = 1] = n^{-1}$, we do not know how to produce a b' such that $\Pr[b = 1] = n^{-\gamma}$, for some $\gamma \in (0, 1)$, which prevents us from fully exploiting the improved parameters of the DFR.

Proof of Theorem 1.3. Direct consequence of Corollary 3.18.1 and Lemma 4.7. □

Proof of Theorem 1.4. Let H denote a finite-index subgroup of G such that $H \cong (\mathrm{PU}(d, \overline{\mathbb{Q}}))^k$. By definition, we have an injective group homomorphism $\rho : H \to (\mathrm{PU}(d, \overline{\mathbb{Q}}))^k$. For $j \in \{1, \ldots, k\}$, let $\tau_j : (\mathrm{PU}(d, \overline{\mathbb{Q}}))^k \to \mathrm{PU}(d, \overline{\mathbb{Q}})$ denote the canonical projection onto the j^{th} factor and define the representation $\rho_j : H \to \mathrm{PU}(d, \overline{\mathbb{Q}})$ by $\rho_j = \tau_j \circ \rho$. As $\ker \rho = \{1_H\}$, we then have $\bigcap_j \ker \rho_j = \{1_H\}$. The result is then a direct consequence of applying Theorem 3.20 to $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ and then applying Lemma 4.7. □

Proof of Corollary 1.4.1. Follows immediately from Theorem 1.4. Alternatively, this is a direct consequence of Corollary 3.19.1 and Lemma 4.7. □

We next consider the unbounded-error case. We first show the following lemma, which is the unbounded-error version of a combination of Lemma 4.6 and Lemma 4.7.

Lemma 4.8. *Consider a group $H = \langle S_H | R_H \rangle$, with S_H finite, and suppose that $\mathcal{F} = \{\rho_1, \ldots, \rho_k\}$ is a set of representations $\rho_j : H \to \mathrm{U}(d)$ such that $\bigcap_j \mathrm{Pker}(\rho_j) = \{1_H\}$. Further suppose G is a group such that $H \leq G$ and $[G : H]$ is finite. Then G admits a presentation $G = \langle S_G | R_G \rangle$, with S_G finite, such* *that there is a 2QCFA A_G, with d quantum basis states, that recognizes W_G with negative one-sided unbounded-error. Moreover, the above claim also holds if the ρ_j are projective representations.*

Proof. First, by an argument similar to that of Lemma 4.6, we construct a 2QCFA A_H for W_H. By Lemma 4.4, for each $j \in \{1, \dots, k\}$, there is a 2QCFA subroutine B_j that produces a result $r_j \in \{0, 1\}$ such that, on any input $w \in \Sigma^*$, if $\phi(w) = 1_H$, then $\Pr[r_j = 1] = 1$, and if $\phi(w) \notin \mathrm{Pker}(\rho_j)$, then $\Pr[r_j = 0] > 0$. A_H operates as follows. For each j, run B_j producing r_j. If $r_j = 0$, then immediately reject; otherwise continue with the next j. Then (i.e., if A_H completes the above procedure for every $j \in \{1, \dots, k\}$ without rejecting) immediately accept. The correctness of this procedure is immediate, as $\cap_j \mathrm{Pker}(\rho_j) = \{1_H\}$. Using A_H we then produce the 2QCFA for A_G precisely as in the proof of Lemma 4.7. □

We now prove the remaining theorems stated in Section 1.2.

Proof of Theorem 1.5. By Theorem 3.23, any $G = \langle S|R\rangle \in \widehat{\Pi}_3$ has an unbounded-error $[k, 2]$-DFR $\mathcal{F} = \{\rho_1, \dots, \rho_k\}$, for some constant k. By applying Lemma 4.8, we produce the desired 2QCFA A_G for W_G (note that, by definition, $\forall j \in \{1, \dots, k\}, \forall s \in S$, $\rho_j(s)$ is expressible as the product of a finite number of matrices in $\mathrm{U}(d) \cap M_d(\widetilde{\mathbb{C}})$, which then implies that the transition amplitudes of A_G all lie in $\widetilde{\mathbb{C}}$). □

Proof of Theorem 1.6. Let H denote a finite-index subgroup of G such that $H \cong (\mathrm{PU}(d))^k$. As in the proof of Theorem 1.4, we obtain a set $\mathcal{F} = \{\rho_1, \dots, \rho_k\}$ of representations $\rho_j : H \to \mathrm{PU}(d)$ where $\cap_j \ker \rho_j = \{1_H\}$. The result follows by Lemma 4.8 and Lemma 4.7. □

Proof of Theorem 1.7. Recall that $\mathcal{D}$ is defined to be the class of all groups G for which a 2QCFA for W_G is produced by any of the preceding theorems. Then $\mathcal{D}$ consists precisely of those finitely-generated groups G that have a P-faithful projective representation $\rho : G \to \mathrm{PU}(d)$, for some finite d. Let $\tau : \mathrm{U}(d) \to \mathrm{PU}(d)$ denote the canonical projection and let $\widehat{\rho} : G \to \mathrm{U}(d)$ denote a function (not necessarily a homomorphism) such that $\tau \circ \widehat{\rho} = \rho$. A Moore-Crutchfield MO-1QFA (see [29] for complete definition) recognizes W_G by using the "full" encoding defined at the beginning of Section 4.1, i.e., after reading the partial input word w the quantum register stores the entire $d \times d$ matrix $\widehat{\rho}(\phi(w))$; after reading the entire word, a single quantum measurement is performed to determine if the magnitude of the trace of this matrix is d. □

5 Discussion

5.1 Computational Complexity of the Word Problem

Word problems for finitely-generated groups have long been studied by complexity theorists, and it has been observed that there are many deep connections between the algebraic structure of a particular group and the computational complexity of its word problem. In this section, we compare the results that we have obtained concerning the ability of a 2QCFA to recognize certain group word problems with existing results for "simple" classical and quantum models.

We use the following notation for complexity classes: REG denotes the regular languages (languages recognized by deterministic finite automata), CFL (resp. DCFL) denotes the context-free (resp. deterministic context-free) languages (languages recognized by non-deterministic (resp. deterministic) pushdown automata), OCL (resp. DOCL) denotes the one-counter (resp. deterministic one-counter) languages (languages recognized by non-deterministic (resp. deterministic) pushdown automata where the stack alphabet is limited to a single symbol), poly−CFL (resp. poly−DCFL, poly−OCL, poly−DOCL) denotes the intersection of finitely many context-free (resp. deterministic context-free, one-counter, deterministic one-counter) languages, and L denotes deterministic logspace (languages recognized by deterministic Turing machines with read-only input tape and read/write work tape of size logarithmic in the input).

Using the notation of Section 1.2, we write $\widehat{\Pi}_0$ (resp. $\widehat{\Pi}_1$, $\widehat{\Sigma}_1$, $\widehat{\Pi}_2$) 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). 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. The following proposition, 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 5.1. *([4, 5, 8, 12, 20, 23, 27, 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) $G \in \widehat{\Pi}_2 \Rightarrow W_G \in \mathsf{poly{-}DCFL} \subsetneq \mathsf{poly{-}CFL}$.

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

Proof. Statements $(i), (ii), (iii), (v),$ and (vi) were shown, respectively, in [4],[20],[23], [8], and [27]. In [30], it was shown that G is free if and only if $W_G \in \mathsf{CFL}$ and G is accessible, in [12], it was shown that all finitely-presented groups are accessible, and in [5] 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]. □

Our results have a close correspondence to the above mentioned results. By Theorem 1.2 (resp. Theorem 1.3), for all groups $G \in \widehat{\Pi}_1 \supsetneq \widehat{\Pi}_0 \supsetneq \widehat{\{1\}}$ (resp. $G \in \widehat{\Pi}_2 \supsetneq \widehat{\Pi}_1 \cup \widehat{\Sigma}_1 \supsetneq \widehat{\Pi}_0 \supsetneq \widehat{\{1\}}$), W_G is recognized, with one-sided bounded-error, in expected polynomial (resp. exponential) time, by a 2QCFA with a single qubit and algebraic number transition amplitudes. Of course, as our fundamental approach to solving the group word problem is to construct a DFR for a group G, and as any such DFR yields a faithful finite-dimensional unitary representation of G, any such $G \in \mathcal{L}$, as a faithful finite-dimensional *unitary* representation is (a special case of) a faithful finite-dimensional *linear* representation over $\mathbb{C}$.

It is particularly interesting that, while there are strict inclusions $\mathsf{DCFL} \subsetneq \mathsf{CFL}$, $\mathsf{DOCL} \subsetneq \mathsf{OCL}$, and $\mathsf{poly{-}DOCL} \subsetneq \mathsf{poly{-}OCL}$, there are no groups whose word problem witnesses any of these separations. That is to say, the deterministic and non-deterministic versions of each of these models (pushdown automata, one-counter automata, etc.) can recognize word problems for precisely the same class of groups. Let Σ_2 denote the finitely-generated groups that are a free product of finitely many finite-rank free abelian groups, and $\widehat{\Pi}_3$ denote the finitely-generated groups that are virtually a subgroup of a direct product of finitely many groups in Σ_2. By Theorem 1.5, for all $G \in \widehat{\Pi}_3 \supsetneq \widehat{\Pi}_2$, the word problem W_G is recognized with one-sided *unbounded-error* in expected exponential time by a 2QCFA with a single qubit and transition amplitudes in $\widetilde{\mathbb{C}}$. Of course, language recognition with one-sided unbounded-error is naturally a non-deterministic analogue of language recognition with one-sided bounded-error, which raises the question of whether or not adding non-determinism to the 2QCFA model allows the recognition of a larger class of group word problems. In particular, consider the group $G = \mathbb{Z} * \mathbb{Z}^2$, and notice that $G \in \Sigma_2 \subsetneq \widehat{\Pi}_3$ (and so W_G is recognized by an unbounded-error 2QCFA), but $G \notin \widehat{\Pi}_2$. The complexity of W_G has been considered by many authors and it is conjectured that $W_G \notin \mathsf{poly{-}CFL}$ [8](cf. [10]) and that $W_G \notin \mathsf{coCFL}$ [24]. We ask the following question.

Open Problem 5.2. *Is W_G, the word problem of the group $G = \mathbb{Z} * \mathbb{Z}^2$, recognizable with one-sided bounded-error by a 2QCFA with algebraic number transition amplitudes? More generally, is the word problem of every group of the form $\mathbb{Z} * \mathbb{Z}^r$, $r \in \mathbb{N}$ recognizable by such a 2QCFA?*

In particular, by Lemma 3.17, for any distinct primes $p_1, p_2 \equiv 1 \mod 4$, and any prime $q \equiv 1 \mod 4$ (where q is not necessarily distinct from p_1, p_2), we have P-faithful representations $\rho : \mathbb{Z}^2 = \langle x_1, x_2 | [x_1, x_2] \rangle \to \mathrm{SO}(2, \mathbb{Z}[\frac{1}{p_1}, \frac{1}{p_2}]) \leq \mathrm{SO}(2, \mathbb{Q})$ and $\pi : \mathbb{Z} = \langle y | \rangle \to \mathrm{SO}(2, \mathbb{Z}[\frac{1}{q}]) \leq \mathrm{SO}(2, \mathbb{Q})$. By use of Shalen's method [40, Proposition 1.3], we can produce a P-faithful representation $\gamma : G \to \mathrm{SU}(2)$ of $G = \mathbb{Z} * \mathbb{Z}^2$ (see the proof of Theorem 3.23). While $\gamma(y) = \pi(y) \in \mathrm{SO}(2, \mathbb{Q})$ (here we identify y with its image under the natural inclusion $\mathbb{Z} \to \mathbb{Z} * \mathbb{Z}^2$), we have, for $j \in \{1, 2\}$, that $\gamma(x_j) = \Lambda \rho(x_j) \Lambda^{-1}$, where $\rho(x_j) \in \mathrm{SO}(2, \mathbb{Q})$, but $\Lambda \in \mathrm{T}(2, \widetilde{E})$. This added complexity of the numbers that appear in these matrices produces two issues. Firstly, our construction of a 2QCFA that makes use of this γ to recognize W_G has transition amplitudes that lie in $\widetilde{\mathbb{C}} = \overline{\mathbb{Q}} \cup \widetilde{E}$; whereas, we would prefer to construct a 2QCFA whose transition amplitudes are limited to $\overline{\mathbb{Q}}$. Secondly, it is then unclear if a useful bound can be obtained on the distance $2 - |\chi_\gamma(g)|$, for $g \in G_{\neq 1_G}$ in terms of $l(g)$, as each such $\chi_\gamma(g)$ is a polynomial $f \in R[z]$ evaluated at some transcendental number λ, where $R = \mathbb{Z}[\frac{1}{p_1}, \frac{1}{p_2}, \frac{1}{q}]$ and $R[z]$ denotes the polynomial ring over the ring R in a single indeterminate z. We naturally ask if this construction, or a construction like this (for example, by the same construction, one also has a P-faithful representation of G into $\mathrm{SU}(2, R[z, z^{-1}])$, where $R[z, z^{-1}]$ denotes the ring of Laurent polynomials in a single indeterminate over R), could be adapted to yield an algebraic DFR of G, or similar groups.

Open Problem 5.3. *Does the group $\mathbb{Z} * \mathbb{Z}^2$ have an algebraic $[k, d, C^{-n}]$-DFR, for some $k \in \mathbb{N}_{\geq 1}$, $d \in \mathbb{N}_{\geq 2}$, $C \in \mathbb{R}_{\geq 1}$. More generally, does every group $\mathbb{Z} * \mathbb{Z}^r$, $r \in \mathbb{N}$ have such a DFR? Even more, generally, is the class of groups which have DFRs of this type closed under free product?*

Remark. Of course, such a DFR would immediately yield a 2QCFA of the desired type for the corresponding word problem. Moreover, recall that Σ_2 consists of all groups of the form $\mathbb{Z}^{r_1} * \cdots * \mathbb{Z}^{r_m}$, for some $m, r_1, \ldots, r_m \in \mathbb{N}$, and that any such groups embeds in (i.e., is a subgroup of) $\mathbb{Z} * \mathbb{Z}^r$, where $r = \max_j r_j$. By Lemma 3.15(i), if $\mathbb{Z} * \mathbb{Z}^r$ has a DFR then $\mathbb{Z}^{r_1} * \cdots * \mathbb{Z}^{r_m}$ has a DFR with essentially the same parameters. Therefore, if all such $\mathbb{Z} * \mathbb{Z}^r$ have DFRs of the desired type, then so do all groups in Σ_2, which would then imply all groups in $\widehat{\Pi}_3$ virtually have such a DFR, by an application of Lemma 3.14 and Lemma 3.15(i).

We next consider known results concerning those group word problems recognizable by particular QFA variants. Ambainis and Watrous, in the paper in which the 2QCFA model was first defined [2], demonstrated the power of this new model by showing that it is capable of recognizing certain languages that probabilistic two-way finite automata cannot. In particular, they considered the languages $L_{eq} = \{a^m b^m | m \in \mathbb{N}\}$ and $L_{pal} = \{w \in \{a, b\}^* | w \text{ is a palindrome}\}$. They showed that a 2QCFA, with only two quantum basis states (i.e., a single qubit quantum register), can recognize L_{eq} (resp. L_{pal}) with one-sided bounded-error in expected polynomial (resp. exponential) time. As noted in the introduction, while neither L_{eq} nor L_{pal} are group word problems, they are closely related to word problems for $\mathbb{Z}$ and F_2, respectively. This observation allows us to reinterpret the above results of Ambainis and Watrous [2] in terms of the word problem.

In addition to results of the above form, which, implicitly, study the quantum computational complexity of the word problem for certain groups, some authors have explicitly considered this question. We now briefly recall certain results that are especially relevant to this paper (see the survey [3] for a complete history). In the following we write MO-1QFA for the measure-once one-way QFA (defined in [29]), MM-1QFA for the measure-many one-way QFA (defined in [25]) and 1QFA$\circlearrowleft$ for the one-way QFA with restart (defined in [48]). Let $\mathsf{S}^{=}_{\mathbb{Q}}$ denote the class of languages L for which there is a PFA (probabilistic finite automaton) P, all of whose transition amplitudes are rational numbers, such that, $\forall w \in L$, the probability that P accepts w is exactly $\frac{1}{2}$, and, $\forall w \notin L$, the probability that P accepts w differs from $\frac{1}{2}$.

Brodsky and Pippenger [7] showed that the languages W_{F_k}, $k \in \mathbb{N}$ (in particular, recall $F_1 = \mathbb{Z}$) can be recognized, with negative one-sided *unbounded-error*, by a MO-1QFA. Yakaryilmaz and Say [48] showed that, any language $L \in \mathsf{S}^{=}_{\mathbb{Q}}$ can be recognized by a MM-1QFA, with negative one-sided *unbounded-error*, and by a 1QFA$\circlearrowleft$ or 2QCFA, with negative one-sided *bounded-error*, in expected *exponential time.* As L_{eq} and L_{pal} both belong to $\mathsf{S}^{=}_{\mathbb{Q}}$, this result, partially, subsumes the original result from Ambainis and Watrous [2]. However, in addition to the difference in expected running time in the case of L_{eq}, we also note that there is a significant difference between the sizes of the quantum registers of the machines produced in these two results. In particular, the result of Yakaryilmaz and Say [48] was obtained by using the technique of Yakaryilmaz and Say [47] to directly simulate a PFA with a MM-1QFA; this leads to a MM-1QFA with a number of quantum basis states given by a particular constant plus the (potentially non-trivially large) constant number of states of the PFA: a similar statement holds for the constructions of 1QFA$\circlearrowleft$ and 2QCFA. For example, the 1QFA$\circlearrowleft$ and 2QCFA constructed by Yakaryilmaz and Say [48] to recognize L_{pal} have 15 quantum basis states (and therefore require 4 qubits), as opposed to the 2 quantum basis states (i.e., 1 qubit) of Ambainis and Watrous [2]. Similarly, as $W_{F_k} \in \mathsf{S}^{=}_{\mathbb{Q}}$, $\forall k \in \mathbb{N}$, the result of Yakaryilmaz and Say [48] shows that the word problems of these groups can be recognized by a 2QCFA of our type; though, a direct application of their construction would yield a 2QCFA with larger quantum part than that of our construction, or that of Ambainis and Watrous [2]. Of course, our results also apply to the 1QFA$\circlearrowleft$ model.

5.2 Information Compression, Unstable Stacks, and Quantum Pointers and Counters

The 2QCFA constructed by Ambainis and Watrous [2] that recognize L_{eq} and L_{pal} do so using only a single qubit; as they noted, this demonstrates that quantum computational models can perform a particularly interesting sort of extreme information compression. We next observe that the same phenomenon occurs in our constructions of 2QCFA. Consider a group $G = \langle S|R \rangle$, with S finite, where we assume for notational convenience, that $1_G \notin S$. Let $\Sigma = S \cup S^{-1}$ denote the corresponding symmetric generating set that forms the alphabet for the word problem W_G and let $\phi : \Sigma^* \to G$ denote the natural map that takes each word in Σ^* to the group element it represents. The core idea of our 2QCFA A for the word problem W_G is to scan the input word $w = w_1 \cdots w_n \in \Sigma^*$ (where each $w_j \in \Sigma$) and, after the partial word $w_1 \cdots w_t$ has been read[2] the quantum register of A stores the group element $g_t := \phi(w_1 \cdots w_t) \in G$. Interestingly, for a wide collection of groups, A is able to store g_t using only a single qubit. Therefore, in a certain sense, such an A is storing an unbounded amount of information in a single qubit.

To clarify the information stored in this single qubit, let $\Gamma(G, \Sigma)$ denote the Cayley graph of G with respect to Σ. Then A operates by following the path p_w specified by w in $\Gamma(G, \Sigma)$, using its quantum register to store the current vertex in this graph (i.e., after t steps along the path p_w, the element $g_t \in G$ is stored). Let $B_{G,\Sigma}(n) = \{g \in G | l_S(g) \leq n\}$ denote the closed ball in $\Gamma(G, \Sigma)$ of radius n centered at 1_G, and let $f_{G,\Sigma}(n) = |B_{G,\Sigma}(n)|$ denote the *growth rate* of G. On an input w, A scans w and, after having read the first t symbols of w, A has the element $g_t \in G$ corresponding to that prefix of w stored in its single qubit. On inputs of string length n, g_t may vary over the entirety of $B_{G,\Sigma}(n)$.

[2] Strictly speaking, our algorithm reads the input word "backwards" (i.e., from w_n to w_1), though in this section we will describe the algorithm as though it operated in the "forwards" direction, for the sake of clarity. As discussed in Section 4.1, the algorithm could be adapted to operate in the "forwards" direction.

In order to store an arbitrary element of $B_{G,\Sigma}(n)$ such that it is (information theoretically) possible to perfectly discern the identity of that element, one requires $\log(f_{G,\Sigma}(n))$ (classical) bits. Moreover, by Holevo's theorem [22], this same task requires $\log f_{G,\Sigma}(n)$ qubits. For the remainder of this section, we ignore the uninteresting case in which G is a finite group (as then W_G is a regular language), and consider only finitely-generated (necessarily countably) infinite groups, and so $f_{G,\Sigma}$ is necessarily a growing function of n. Therefore, we must first make clear why our approach, which encodes such an element using only a single qubit, does not violate Holevo's theorem. The key observation is that, while all $\log f_G(n)$ bits of information are truly stored in the single qubit, one is extremely limited in the manner in which that information may be accessed. In particular, this information may only be accessed by performing a quantum measurement, which only (probabilistically) indicates whether or not the currently stored value g_t is equal to the identity element 1_G; moreover, performing this quantum measurement completely destroys all information stored in this qubit. This extremely severe restriction on the manner in which the information content of a qubit may be accessed prevents one from reconstructing information stored within the qubit in a manner inconsistent with Holevo's theorem. On the other hand, this restriction is perfectly consistent with the manner in which A operates when solving the word problem of G, and so it provides no impediment to using a single qubit to store information in a radically compressed way.

We next quantify the extent to which our constructions of 2QCFA compress information. For two monotone non-decreasing functions $f_1, f_2 : \mathbb{R}_{\geq 1} \to \mathbb{R}_{\geq 0}$, we write $f_1 \prec f_2$ if there are constants $C_1, C_2 \in \mathbb{R}_{>0}$ such that, $f_1(x) \leq C_1 f_2(C_1 x + C_2) + C_2$, $\forall x \in \mathbb{R}_{\geq 0}$, and we write $f_1 \sim f_2$ if both $f_1 \prec f_2$ and $f_2 \prec f_1$. Note that while the exact value of $f_{G,\Sigma}(n)$ does depend on the choice of symmetric generating set Σ, the asymptotic behavior does not, in that $f_{G,\Sigma} \sim f_{G,\Sigma'}$, for any other finite symmetric generating set Σ' (see, for instance, [28, Proposition 6.2.4]); therefore, we will often simply write f_G in place of $f_{G,\Sigma}$ when only the asymptotic behavior is relevant. We say G is of *polynomial growth* if $f_G \sim n^C$, for some $C \in \mathbb{R}_{\geq 0}$, and of *exponential growth* if $f_G \sim C^n$, for some $C \in \mathbb{R}_{>0}$. In this paper we have restricted our attention to finitely-generated linear groups over a field of characteristic 0; we denote this class of groups by $\mathcal{L}$. By the famous Tits' alternative [45], every $G \in \mathcal{L}$ is either of polynomial or exponential growth; in particular, every $G \in \mathcal{L}$ that is not virtually nilpotent has exponential growth (more explicitly, any $G \in \mathcal{L}$ that is virtually solvable but not virtually nilpotent, or which has a subgroup isomorphic to F_2, has exponential growth). For example, any finitely-generated virtually abelian group G has polynomial growth; more precisely, $f_G \sim n^r$, where $r \in \mathbb{N}$ is the unique value such that G has a finite-index subgroup isomorphic to $\mathbb{Z}^r$. By the discussion of the previous paragraph, one requires $\log f_G(n) \sim \log(n)$ classical bits to unambiguously store an element of $B_{G,\Sigma}(n)$. By Theorem 1.2, for any such G, there is a 2QCFA A, which has only a single-qubit and algebraic number transition amplitudes, that recognizes W_G, with negative one-sided bounded-error, in expected polynomial time. In particular, A stores this arbitrary element of $B_{G,\Sigma}(n)$ using only a single qubit. More dramatically, by Theorem 1.3, for any finitely-generated virtually free group G, there is a 2QCFA A, which has only a single-qubit and algebraic number transition amplitudes, that recognizes W_G, with negative one-sided bounded-error, in expected exponential time. Any such G which is not virtually cyclic (i.e., any such G that is neither finite nor virtually $\mathbb{Z}$, or equivalently any such G that has a subgroup isomorphic to F_2) has exponential growth, which means that one requires $\log f_G(n) \sim n^C$ classical bits to unambiguously store an element of $B_{G,\Sigma}(n)$, for some constant $C \in \mathbb{R}_{>0}$. Yet, A still stores an arbitrary element of $B_{G,\Sigma}(n)$ using only one qubit.

The above examples, and more generally all of the 2QCFA that we have constructed for various word problems, demonstrate the extreme sort of information compression that a 2QCFA is capable of performing. On the other hand, this extreme compression does not come without a cost, as it directly impacts the running time of our 2QCFA. We note that information compression of this form is by no means a new idea in quantum computation, as techniques like quantum fingerprinting [9] and dense quantum coding [1] explicitly involve such compression, and, moreover, many quantum algorithms, including Shor's quantum factoring algorithm [41], crucially rely on this sort of compression to achieve their apparent speedup relative to their classical counterparts. Nevertheless, both the original Ambainis and Watrous 2QCFA result [2] and our approach push this idea down to the much weaker computational model of 2QCFA, and introduce techniques that might also be useful for more powerful quantum models.

Next, we consider other consequences and interpretations of this extreme compression of information. We first observe that the 2QCFA Q_{eq} constructed by Ambainis and Watrous [2] to recognize $L_{eq} = \{a^m b^m | m \in \mathbb{N}\}$ operates in much the same way as the natural deterministic pushdown automaton (DPDA) for this language. Namely there is DPDA P_{eq} (which is in fact a deterministic one-counter automaton) that recognizes L_{eq} by scanning the input, pushing a symbol 1 onto the stack when reading each a and popping a 1 off the stack when reading each b; P_{eq} rejects if the input is not of the form a^*b^* or if it ever reads the symbol b when the stack is empty, and accepts otherwise. Similarly, Q_{eq} recognizes L_{eq} by scanning the input, applying a unitary transformation T when reading each a and T^{-1} when reading each b. Therefore, after reading the prefix $w_1 \cdots w_t$ of the input $w_1 \cdots w_n$, the single qubit register of Q_{eq} has the value $d_t := \#(w_1 \cdots w_t, a) - \#(w_1 \cdots w_t, b)$ stored (where $\#(x, \sigma)$ denotes the number of appearances of the letter σ in the word x). Then Q_{eq} rejects if the input is not of the form a^*b^* or if $d_n \neq 0$, and accepts otherwise. There is a similar relationship between the 2QCFA Q_{pal} constructed by Ambainis and Watrous [2] to recognize $L_{pal} = \{w \in \{a,b\}^* | w \text{ is a palindrome}\}$ and the natural two-pass DPDA P_{pal} for this language. Namely, in P_{pal}'s first pass over the input, it pushes each symbol that is read onto the stack, and in its second pass over the input, it pops one symbol at a time from the stack and compares with the input; P_{pal} rejects if a mismatch is found, and accepts otherwise. In Q_{pal}'s first pass over the input, it performs the unitary transformation A when a is read and B when b is read; in its second pass over the input, it performs the unitary transformation A^{-1} when a is read and B^{-1} when b is read. A and B are chosen so they satisfy no non-trivial relations (i.e., they generate a free group), and so the input is a palindrome precisely when the overall transformation performed is the identity, which is the criterion that Q_{pal} checks to determine whether to accept or reject the input.

Similarly, there is also a close correspondence between the 2QCFA that we have constructed for various word problems and the natural multi-pass DPDA for those word problems. In a certain sense, both our constructions, and those of Ambainis and Watrous [2] discussed above use a single qubit to simulate a stack, albeit an "unstable" stack. Of course, given an (actual) stack, one can, for example, check if the stack is empty, or determine the value of the symbol on top of the stack, without altering the state of the stack; on the other hand, given this unstable simulation of a stack with a single qubit, one can only obtain information by performing a quantum measurement, which destroys all stored information. Moreover, on an (actual) stack, symbols are stored discretely, whereas this unstable simulation mashes together all symbols into the state of a single qubit. Nevertheless, this technique yields 2QCFA, with only one qubit, that can recognize any word problem $W_G \in \mathsf{CFL}$, and, more generally, any W_G known to be in $\mathsf{poly}{-}\mathsf{DCFL}$ or even $\mathsf{poly}{-}\mathsf{CFL}$ (see Theorem 1.3 and the discussion that follows it). Therefore, it appears that these deficiencies in the unstable single qubit simulation of a stack pose no problem regarding the recognition of word problems. The situation is far less clear if one considers recognition of general languages, rather than simply group word problems; that is to say, the relationship between the class of languages recognizable by 2QCFA and the complexity class CFL (and, more generally, $\mathsf{poly}{-}\mathsf{DOCL}$ and $\mathsf{poly}{-}\mathsf{DCFL}$) is quite unclear. We note that this is not an unusual phenomenon as, for example, while it remains an open question whether or not $\mathsf{CFL} \subseteq \mathsf{L}$, it is known that for any word problem W_G, $W_G \in \mathsf{CFL} \Rightarrow W_G \in \mathsf{L}$ [27, 30]. On the other hand, perhaps this unstable simulation of a stack has an advantage over an actual stack, as we have shown that a 2QCFA, with a single qubit and transition amplitudes in $\widetilde{\mathbb{C}}$, can recognize the word problem W_H of the group $H = \mathbb{Z} * \mathbb{Z}^2$ with negative one-sided *unbounded-error*; note that $W_H \notin \mathsf{CFL}$ and it is conjectured that $W_H \notin \mathsf{poly{-}CFL}$ [8](cf. [10]) and that $W_H \notin \mathsf{coCFL}$ [24]. In particular, the mashing together of all symbols on the stack into the state of a single qubit might actually be an advantage when solving the word problem.

The core technical fact enabling the proof of Theorem 1.2, which guarantees the existence of a single qubit 2QCFA that recognizes W_G, for any $G \in \widehat{\Pi}_1$ (the finitely-generated virtually abelian groups), is that the group $\mathbb{Z}$ has a faithful one-dimensional representation, as this fact allows us to use a single qubit to store a counter in a single qubit. More generally, $\mathbb{Z}^r$ has a faithful one-dimensional representation, for any $r \in \mathbb{N}$, which in fact allows us to use a single qubit to store r counters; while we did not use such a representation in the construction of our 2QCFA for W_G when $G \in \widehat{\Pi}_1$, as this produces worse parameters than using the ability of a 2QCFA to make multiple passes over its input, this representation did play a crucial role in our construction of a 2QCFA for W_G with $G \in \widehat{\Pi}_3$ (see Theorem 3.23 and the discussion surrounding it). More generally, all of our constructions of 2QCFA with a single qubit for some W_G relied on the existence of a faithful two-dimensional representation of G, or of some group closely related to G. Such a representation of a group G allows a single qubit to be used to store an arbitrary element $g \in G$; moreover, crucially, it also allows one to perform limited computation on the stored element. In particular, one can replace the stored element g with the element $g\sigma$, for any $\sigma \in \Sigma$ (where Σ is a finite symmetric generating set of G); being able to perform this computational step is essential to our approach to the word problem. One can naturally interpret these constructions as making use of a single qubit to store a *quantum pointer* into the Cayley graph $\Gamma(G, \Sigma)$, where the quantum pointer points to a single vertex in $\Gamma(G, \Sigma)$ (i.e., a single $g \in G$), and one can, in a single computational step, update the pointer so that it points to a neighbor of the current vertex. In the special case when $G = \mathbb{Z}^r$, we can view this quantum pointer more simply as an *r-wide quantum counter*, which stores r integers. Therefore, much as a deterministic logspace Turing machine can use its $O(\log(n))$ space work tape to store a pointer into its length n input string, or to store a finite number of counters that can each store a $O(\log(n))$-bit integer, these constructions of 2QCFA use a single qubit to store a pointer or counter that would require as many as poly(n) classical bits to directly record. Of course, our implementations of quantum pointers and counters are quite limited, as in order to obtain any information, one must perform a quantum measurement; the result of this measurement only indicates (probabilistically) if the value pointed to is 1_G (in the case of an r-wide quantum counter, this corresponds to all counters having value 0), and performing the measurement destroys all stored information. Nevertheless, it is natural to ask what other sort of computational tasks one might be able to perform using quantum pointers and counters. In particular, we note here that, as discussion in Section 4.1, one can store the entire $d \times d$ matrix corresponding to the value of a d-dimensional unitary representation using d^2 quantum basis states.

5.3 Quantum Finite Automata Variants

As noted in the introduction, many distinct variants of quantum finite automata (QFA) have been defined (see for example [7, 11, 21, 25, 29, 32, 35, 48], see the excellent survey [3] for a complete history). These numerous types of QFA differ significantly in terms of their language recognition power: at one extreme, the measure-once one-way quantum finite automata (MO-1QFA) defined by Moore and Crutchfield [29] can recognize, with bounded-error, precisely the group languages [7], a particularly limited subset of the regular languages; at the other extreme, many QFA variants can recognize undecidable languages, in some cases even with bounded-error, (see, for instance, [38, 47]). Fundamentally, this massive discrepancy in the capabilities of early definitions of QFAs was a result of two significant issues. Firstly, as real, physical implementations of quantum computers were then, and in many ways still are now, in the early experimental phase, it was not entirely clear what features the theoretical QFA model should possess. Secondly, while the core idea of quantum computation is to be a model of computation that benefits from the unique features of quantum mechanics, those same quantum mechanical principles also provide sharp restrictions.

For example, measure-many one-way quantum finite automata (MM-1QFA), defined in [25], can naturally be thought of as a modification of one-way probabilistic finite automata (1PFA) in which the probabilistic states are replaced by quantum states. When reading each symbol of the input, a MM-1QFA performs a single unitary transformation, rather than a stochastic transformation, on its state space, followed by a single quantum measurement of its state space, where the result of that measurement determines whether the machine accepts its input, rejects its input, or continues reading the input. The class of languages recognized, with bounded-error, by MM-1QFA is a proper subset of the regular languages, which might seem to suggest a setting where quantum computers are actually *weaker* than their classical counterparts. However, this weakness is somewhat artificial. In particular, the laws of quantum mechanics tell us that the state of a quantum system evolves unitarily, and so the requirement that the MM-1QFA may only perform unitary transformations on its state space is physically justified; on the other hand, the requirement that the MM-1QFA may only perform a single quantum measurement when reading each input symbol, and that the result of this measurement may only be used to determine whether or not the machine halts at this stage of the computation, has no physical motivation. That is to say, if a machine is capable of performing a quantum measurement at a particular point in time, then it would be quite reasonable to allow the machine to freely make use of the result of that quantum measurement. If one removes this artificial restriction on the use of quantum measurement then one may obtain a QFA model for which the class of languages recognized, with bounded-error, is precisely the regular languages (e.g., the variant of the MM-1QFA defined in [21], which, after reading each input symbol, may freely perform a sequence of operations, consisting of both unitary transformations and quantum measurements, where the particular operation performed at any stage of the sequence may depend on the results of quantum measurements performed earlier in the sequence).

More generally, one may consider any of the other "standard" one-way QFA variants which are "physically realistic" (i.e., whose definition is consistent with the limitations imposed on real, physical quantum computers by the laws of quantum mechanics; see the survey [3] for a thorough discussion of these models). The class of languages recognized, with bounded-error, by each such model is some subset of the regular languages (more precisely, many of these models recognize exactly the regular languages, others recognize various, often distinct, proper subsets of the regular languages). The requirement that these models have a quantum part that is both small and whose operation is consistent with the laws of quantum mechanics, as well as the requirement that a language must be accepted with bounded-error, make these models quite realistic. However, given that these models can only recognize, at most, the regular languages, they are not very powerful. On the other hand, many versions of QFA have been shown to be able to recognize L_{eq}, L_{pal}, or other related languages, both before and after the 2QCFA result. For example, in [25], it was shown that a two-way quantum finite automaton with a quantum head can also recognize L_{eq} with bounded-error in polynomial time; however, on an input of size n, one would require $\log(n)$ qubits to implement the quantum head. There are also many results which demonstrate the power of variants of the QFA model where it is allowed to recognize a language with unbounded-error, or for which the transition amplitudes of the QFA are non-computable numbers, or for which the model is augmented with various not necessarily physically realizable quantum phenomena (where we again direct the reader to the survey [3]). These models are powerful, but they are not very realistic. On the other hand, the 2QCFA model operates under all the constraints demanded by physical realism, yet it is surprisingly powerful. As demonstrated both in the original Ambainis and Watrous results [2], as well as in this paper, a 2QCFA, with a single qubit and with one-sided bounded error, recognize a very broad class of languages.

The issue of the allowable class of transition amplitudes of a 2QCFA merits further discussion. As in Section 2.1, for a 2QCFA A, let $\mathcal{T}$ denote the set of all unitary matrices T that correspond to a unitary transformation that A may perform on its quantum register, and let $\mathbb{T}$ denote the transition amplitudes of A, which are the set of numbers that appear as an entry of some matrix $T \in \mathcal{T}$. In this paper, we have restricted our attention to 2QCFA whose transition amplitudes $\mathbb{T}$ satisfy $\mathbb{T} \subseteq \overline{\mathbb{Q}}$ or, more generally, $\mathbb{T} \subseteq \widetilde{\mathbb{C}}$. While the stronger restriction to $\overline{\mathbb{Q}}$ is quite natural in quantum computation, the weaker restriction to $\widetilde{\mathbb{C}}$ is motivated by the fact that this is, essentially, the class of transition amplitudes used by the 2QCFA A_{AW} of Ambainis and Watrous [2] that recognizes L_{eq}, which we now clarify. As before, let $\mathbb{A} = \{\pi r | r \in \overline{\mathbb{Q}} \cap \mathbb{R}\}$, $\widetilde{E} = \{e^{ia} | a \in \mathbb{A}\}$, $\widetilde{\mathbb{C}} = \overline{\mathbb{Q}} \cup \widetilde{E}$, and $\mathrm{SO}(2, \cos(\mathbb{A}))$ denote the group of 2×2 special orthogonal matrices with entries in $\cos(\mathbb{A})$. Then A_{AW} has transition matrices $\mathcal{T}_{\mathrm{AW}} \subseteq \mathrm{SO}(2, \cos(\mathbb{A})) \cup \mathrm{U}(2, \overline{\mathbb{Q}})$. As $\widetilde{E} = e^{i\mathbb{A}}$, we are justified in saying that $\widetilde{\mathbb{C}}$ is "essentially" the class of transition amplitudes needed by A_{AW}; moreover, there is a 2QCFA A'_{AW} that is equivalent to A_{AW}, in the sense that A'_{AW} and A_{AW} have precisely the same probability of accepting any input string, but the transition amplitudes of A'_{AW} are restricted to $\widetilde{\mathbb{C}}$. To see this, notice that $\mathrm{SO}(2, \cos(\mathbb{A}))$ is conjugate to a subgroup of $\mathrm{T}(2, \widetilde{E})$ by some element of $\mathrm{U}(2, \overline{\mathbb{Q}})$; i.e., $\exists M \in \mathrm{U}(2, \overline{\mathbb{Q}})$ such that, $\forall N \in \mathrm{SO}(2, \cos(\mathbb{A}))$, $MNM^{-1} \in \mathrm{T}(2, \widetilde{E})$. Produce the 2QCFA A'_{AW} from A_{AW} by replacing each unitary transition $N \in \mathrm{SO}(2, \cos(\mathbb{A}))$ by the sequence of transitions M, MNM^{-1}, M^{-1}. More generally, this sort of decomposition holds for any unitary matrices with entries in $\widetilde{\mathbb{C}}$, as noted in the remark following Definition 3.3: each such matrix $N \in \mathrm{U}(d) \cap M_d(\widetilde{\mathbb{C}})$ can be expressed as $N = M_1 T M_2$, for $M_1, M_2 \in \mathrm{U}(d, \overline{\mathbb{Q}})$ and $T \in \mathrm{T}(d, \widetilde{E})$.

As discussed in Section 5.2, our approach to the word problem fundamentally depends on being able to store an unbounded amount of information in a single qubit, which therefore requires extremely high precision in the state of that qubit. However, this does not imply any unreasonably stringent constraints on the precision of the 2QCFA. That is to say, a qubit is implemented by some physical quantum system, such as a spin-$\frac{1}{2}$ particle, and its precision is guaranteed by the rules of quantum mechanics; the remainder of a 2QCFA, where here we are imagining a concrete physical realization of a 2QCFA as a computational device, interacts with the physical qubit system to perform computation. A 2QCFA can interact with a qubit in only two ways: either by performing a unitary transformation or a quantum measurement. It is only the precision of these interactions that one must consider. While a qubit may store a tremendous amount of information, a quantum measurement of a qubit will only return a single bit of information, and collapse the state of that qubit, destroying all information stored within; in particular, the fact that a qubit stores many bits of information in no way increases the difficulty of performing an accurate quantum measurement. Moreover, as unitary transformations do not compound errors, the needed precision of the unitary transformations is relatively mild; furthermore, our key constructions of DFR only rely on using numbers that are hard to approximate by rational numbers, and this property is held by almost all numbers, a random but static error in the corresponding unitary transformation would cause no issue.

Acknowledgments

The author would like to express his sincere gratitude to Professor Michael Sipser for many years of mentorship and support, without which this work would not have been possible, as well as to thank Professor David Vogan for a very helpful conversation.

References

[1] 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.

[2] 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.

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

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

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

[6] A. Baker, *Transcendental number theory.* Cambridge university press, 1990.

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

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

[9] H. Buhrman, R. Cleve, J. Watrous, and R. De Wolf, "Quantum fingerprinting," *Physical Review Letters*, vol. 87, no. 16, p. 167902, 2001.

[10] T. Ceccherini-Silberstein, M. Coornaert, F. Fiorenzi, P. E. Schupp, and N. W. Touikan, "Multipass automata and group word problems," *Theoretical Computer Science*, vol. 600, pp. 19–33, 2015.

[11] M. P. Ciamarra, "Quantum reversibility and a new model of quantum automaton," in *International Symposium on Fundamentals of Computation Theory.* Springer, 2001, pp. 376–379.

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

[13] 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.

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

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

[16] A. Gamburd, D. Jakobson, and P. Sarnak, "Spectra of elements in the group ring of su (2)," *Journal of the European Mathematical Society*, vol. 1, no. 1, pp. 51–85, 1999.

[17] 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.

[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] 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.

[21] M. Hirvensalo, "Quantum automata with open time evolution," *International Journal of Natural Computing Research (IJNCR)*, vol. 1, no. 1, pp. 70–85, 2010.

[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] A. Kondacs and J. Watrous, "On the power of quantum finite state automata," in *Proceedings 38th Annual Symposium on Foundations of Computer Science.* IEEE, 1997, pp. 66–75.

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

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

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

[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] V. Nisnewitsch, "Uber gruppen, die durch matrizen uber einem kommutativen feld isomorph darstellbar sind," *Matematicheskii Sbornik*, vol. 50, no. 3, pp. 395–403, 1940.

[35] K. Paschen, *Quantum finite automata using ancilla qubits.* Univ., Fak. für Informatik, Bibliothek, 2000.

[36] M. O. Rabin, "Probabilistic automata," *Information and control*, vol. 6, no. 3, pp. 230–245, 1963.

[37] 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.

[38] 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.

[39] W. M. Schmidt, "Simultaneous approximation to algebraic numbers by rationals," *Acta Mathematica*, vol. 125, no. 1, pp. 189–201, 1970.

[40] P. B. Shalen, "Linear representations of certain amalgamated products," *Journal of Pure and Applied Algebra*, vol. 15, no. 2, pp. 187–197, 1979.

[41] 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.

[42] J. Stallings, "A finitely presented group whose 3-dimensional integral homology is not finitely generated," *American Journal of Mathematics*, vol. 85, no. 4, pp. 541–543, 1963.

[43] L. Tan, "The group of rational points on the unit circle," *Mathematics Magazine*, vol. 69, no. 3, pp. 163–171, 1996.

[44] A. Thom, "Convergent sequences in discrete groups," *Canadian Mathematical Bulletin*, vol. 56, no. 2, pp. 424–433, 2013.

[45] J. Tits, "Free subgroups in linear groups," *Journal of Algebra*, vol. 20, no. 2, pp. 250–270, 1972.

[46] J. Watrous, "On the complexity of simulating space-bounded quantum computations," *Computational Complexity*, vol. 12, no. 1-2, pp. 48–84, 2003.

[47] A. Yakaryilmaz and A. C. Say, "Languages recognized by nondeterministic quantum finite automata," *Quantum Information & Computation*, vol. 10, no. 9, pp. 747–770, 2010.

[48] ——, "Succinctness of two-way probabilistic and quantum finite automata," *Discrete Mathematics and Theoretical Computer Science*, vol. 12, no. 4, pp. 19–40, 2010.