Trading Inverses for an Irrep in the Solovay-Kitaev Theorem

Adam Bouland[1]
Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, Berkeley, CA, USA
adam@csail.mit.edu
https://orcid.org/0000-0002-8556-8337

Māris Ozols[2]
QuSoft and University of Amsterdam, Amsterdam, Netherlands
marozols@gmail.com
https://orcid.org/0000-0002-3238-8594

Abstract

The Solovay-Kitaev theorem states that universal quantum gate sets can be exchanged with low overhead. More specifically, any gate on a fixed number of qudits can be simulated with error ϵ using merely polylog$(1/\epsilon)$ gates from any finite universal quantum gate set $\mathcal{G}$. One drawback to the theorem is that it requires the gate set $\mathcal{G}$ to be closed under inversion. Here we show that this restriction can be traded for the assumption that $\mathcal{G}$ contains an irreducible representation of any finite group G. This extends recent work of Sardharwalla *et al.* [29], and applies also to gates from the special linear group. Our work can be seen as partial progress towards the long-standing open problem of proving an inverse-free Solovay-Kitaev theorem [16, 23].

2012 ACM Subject Classification Theory of computation → Quantum computation theory

Keywords and phrases Solovay-Kitaev theorem, quantum gate sets, gate set compilation

Digital Object Identifier 10.4230/LIPIcs.TQC.2018.6

Acknowledgements We thank Aram Harrow, Vadym Kliuchnikov, and Zoltán Zimborás for helpful discussions, and Adam Sawicki for pointing us to reference [35].

1 Introduction

Quantum computing promises to solve certain problems exponentially faster than classical computers. For instance, quantum computers can factor integers [34], simulate quantum mechanics [7], and compute certain knot invariants [4] exponentially faster than the best known classical algorithms. The power of quantum computing is formalized using the notion of quantum circuits, in which polynomial number of quantum gates are applied to a standard input state, and the answer to the computational problem is obtained by measuring the final state. This results in the complexity class BQP (see [24, 20] for an introduction).

Each gate in the circuit is a unitary transformation drawn from some finite *gate set* $\mathcal{G}$; it represents elementary quantum operations that can be performed in hardware and which

[1] AB was partially supported by the NSF GRFP under Grant No. 1122374, by the Vannevar Bush Fellowship from the US Department of Defense, by ARO Grant W911NF-12-1-0541, by NSF Grant CCF-1410022, and by the NSF Waterman award under grant number 1249349.
[2] Part of this work was done while MO was at the University of Cambridge where he was supported by a Leverhulme Trust Early Career Fellowship (ECF-2015-256). MO also acknowledges hospitality of MIT where this project was initiated.

may vary between different realizations of quantum computing. Each gate can act on at most some finite number k of quantum systems at a time, where each individual system (or *qudit*) has d levels. A gate set $\mathcal{G}$ is called *universal*[3] if it is capable of approximately generating any quantum transformation on a sufficiently large number of qudits [15].

In general, the computational power of a quantum device may depend on the gate set $\mathcal{G}$ at its disposal. Clearly, if a gate set is not universal, it may have restricted computational power.[4] But *a priori*, the computational power of different universal gate sets could vary as well. This is because universality simply implies that the gates from $\mathcal{G}$ densely generate all unitaries, but it does not specify how quickly one can approximate arbitrary gates.[5]

While BQP consists of those computations that use $\mathrm{poly}(n)$ gates on an n-bit input, the degree of the polynomial for a specific algorithm could in principle depend on the actual gate set used. For example, if we are given an $O(n)$-gate algorithm over some gate set and we want to implement it using another gate set $\mathcal{G}$, we have to compile each gate to accuracy $O(1/n)$ in terms of $\mathcal{G}$. However, if our compiler uses, say, $O(1/\epsilon)$ gates to achieve accuracy $O(\epsilon)$, the total number of gates would become $O(n^2)$. This would be a strange situation for quantum computation, since the runtime of polynomial-time algorithms would be defined only up to polynomial factors. In particular, this would render Grover's speed-up useless.

Fortunately, this is not the case since the *Solovay-Kitaev theorem* [19, 20, 24, 16] (see also [14, 25]) provides a better compiler, so long as the universal gate set $\mathcal{G}$ is closed under inversion. More specifically, this theorem states that any universal gate set $\mathcal{G}$ can be used to simulate any gate U from any other universal gate set to accuracy ϵ using only $\mathrm{polylog}(1/\epsilon)$ gates from $\mathcal{G}$. Furthermore, there is an efficient algorithm, the *Solovay-Kitaev algorithm*, to perform this conversion between the gate sets.

Before formally stating the Solovay-Kitaev theorem, let us make a few remarks. First, we can assume without loss of generality that all gates in $\mathcal{G}$ are single-qudit gates in some fixed dimension d. Indeed, if $\mathcal{G}$ contains multi-qudit gates or if $\mathcal{G}$ becomes universal only on some larger number of qudits, we can simply set the new dimension to be d^k (for a sufficiently large constant k) and replace $\mathcal{G}$ by a larger gate set that consists of the original gates acting on all ordered subsets of k systems. Second, as we are now dealing with a single system, we can replace the universality of $\mathcal{G}$ by a requirement that $\mathcal{G}$ generates a *dense* subgroup of $\mathrm{SU}(d)$ [31]. Third, we can assume that $\mathcal{G}$ is itself a subset of the *special* unitary group $\mathrm{SU}(d)$ rather than $\mathrm{U}(d)$, since the global phase of a quantum gate has no physical effect. In fact, $\mathrm{U}(1)$ actually does *not* satisfy the Solovay-Kitaev theorem, hence the theorem does not hold for $\mathrm{U}(d)$ either, because in general we cannot approximate the elements of $\mathrm{U}(d)$ accurately enough due to their global phase.

With this fine print aside, we are now ready to state the theorem.

▶ **Theorem** (Solovay-Kitaev theorem [16]). *For any fixed $d \geq 2$, if $\mathcal{G} \subset \mathrm{SU}(d)$ is a finite gate set which is **closed under inverses** and densely generates $\mathrm{SU}(d)$, then there is an algorithm which outputs an ϵ-approximation to any $U \in \mathrm{SU}(d)$ using merely $O\big(\log^{3.97}(1/\epsilon)\big)$ elements from $\mathcal{G}$.*

Therefore if one wishes to change the gate set used for a BQP computation (which requires compiling each gate to $1/\,\mathrm{poly}$ accuracy), a change of gate set only incurs polylogarithmic

[3] This is also known as *physical universality*.

[4] But not always! Some gate sets which are not physically universal are nevertheless capable of universal quantum computing via an encoding; this is known as *encoded universality* [24].

[5] By a simple counting argument, generic unitaries on an n-qubit system require $\tilde{\Omega}(2^n)$ gates to implement (even approximately) irrespectively of the gate set [24, Section 4.5.4].

overhead in the input size n. In particular this implies the runtimes of quantum algorithms based on inverse-closed gate sets are well-defined up to polylog factors in n; an $O(n^c)$ algorithm using one particular universal gate set implies an $\tilde{O}(n^c)$ algorithm using any other (inversion-closed) universal gate set. It also implies that the choice of a particular universal gate set is unimportant for quantum computation; changing between gate sets incurs low overhead.

Given the central importance of the Solovay-Kitaev theorem to quantum computing, prior works have improved the theorem in various directions. For instance, a number of works (see, e.g., [21, 26, 32, 8, 9, 30, 28, 22, 27]) have decreased the overheads of the Solovay-Kitaev theorem for particular inverse-closed gate sets by improving the exponent in the logarithm from 3.97 to 1 (which is optimal) or even by improving the hidden constants in front of the logarithm. Such works are important steps towards making compilation algorithms practically efficient. Additionally prior work has shown a version of the Solovay-Kitaev algorithm for inverse-closed non-unitary matrices [2] and as well more general Lie groups [23]. Note that there is also an information-theoretic non-algorithmic version of the Solovay-Kitaev theorem with exponent 1 for generic inverse-closed gate sets [18]. This has subsequently been extended also to inverse-free gate sets [35].

In this work, rather than improving the overheads of the Solovay-Kitaev theorem, we work towards removing the assumption that the gate set contains inverses of all its gates. This is important for several reasons. First, on a theoretical level it would be surprising if the power of noiseless quantum computers could be gate set dependent. Of course, in the real world one could apply fault-tolerance [3] to allow the use of approximate inverses in place of exact inverses, but it seems strange to have to resort to such a powerful technique to deal with a seemingly minor issue which is easily stated in a noiseless setting. Furthermore, this would not answer the original mathematical question about how fast unitary gate sequences fill the space of all unitaries, since a fault-tolerant implementation corresponds to a completely positive rather than a unitary map (it implements the desired map on an encoded subspace of a larger-dimensional Hilbert space).

Second, an inverse-free Solovay-Kitaev theorem would be very helpful towards classifying the computational power of quantum gate sets. It remains open[6] to prove a classification theorem describing which gate sets $\mathcal{G}$ are capable of universal quantum computing, which are efficiently classically simulable, and which can solve difficult sampling problems like BosonSampling or IQP [1, 12]. A number of recent works have made partial progress on this problem [11, 10]. However, a common bottleneck in these proofs is that they need to invoke the Solovay-Kitaev theorem on various "postselection gadgets" to argue that one can perform hard sampling problems, and the set of these gadgets is not necessarily closed under inversion. In the above works this problem is tackled on an ad-hoc, case specific basis. An inverse-free Solovay-Kitaev theorem would simplify these proofs and expand the frontier for gate classification.

Finally, such theorem would enable further progress in quantum Hamiltonian complexity where universal gate sets are used to encode computational instructions by the interaction terms of local Hamiltonians. The ground states of such Hamiltonians have very complicated structure and computing their ground energy is typically $\mathsf{QMA}_{\mathsf{EXP}}$-complete [17], a phenomenon that can occur even when the local dimension of each individual subsystem is relatively small [5, 6]. Since low local dimension is physically more relevant, it is desirable to minimize the dimensions of these constructions even further. A significant roadblock in this is

[6] Even for the case of two-qubit gate sets [15, 11]!

the size of the universal gate set used to encode the computation. Since each gate contributes additional dimensions, one would like to have as few gates as possible. Considering how intricate and hard to optimize the known constructions [5, 6] are, getting rid of inverses would be an easy way forward.

For the reasons outlined above, we believe this longstanding open problem (noted in [16, 23]) is an important one to resolve. In this work, we make partial progress towards this goal by replacing the inverse-closedness assumption with the requirement that the gate set contains any (projective) irreducible representation (a.k.a. *irrep*) of a finite group. Roughly speaking, a *projective irrep* is a set of unitary matrices that form a group (up to a global phase) and that do not leave any non-trivial subspace invariant. A canonical example is the set of *Pauli matrices* $\{I, X, Y, Z\}$.

▶ **Theorem 1** (Solovay-Kitaev theorem with an irrep instead of inverses). *For any fixed $d \geq 2$, suppose $\mathcal{G} \subset \mathrm{SU}(d)$ is a finite gate set which densely generates $\mathrm{SU}(d)$, and furthermore $\mathcal{G}$ contains a (projective) irrep of some finite group G. Then there is an algorithm which outputs an ϵ-approximation to any $U \in \mathrm{SU}(d)$ using merely $O(\mathrm{polylog}(1/\epsilon))$ elements from $\mathcal{G}$.*

In other words, the inverses of some of the gates of $\mathcal{G}$—namely those which constitute an irrep of G—are also in $\mathcal{G}$, but the inverses of the remaining gates may not be in $\mathcal{G}$. So we are trading inverses for some other structure in the gate set $\mathcal{G}$. This extends recent work of Sardharwalla, Cubitt, Harrow and Linden [29] which proved this theorem in the special case that G is the Weyl (or generalized Pauli)[7] group. Sardharwalla *et al.*'s result has already found application in gate set classification [10]. We therefore expect that our result will likewise enable further progress on the gate set classification problem. We also extend our theorem to the non-unitary case (see Theorem 4 in Appendix C), thus generalizing the (inverse-closed) non-unitary Solovay-Kitaev theorem of [2] (this may further extend to more general Lie groups as well following [23]). We expect that this version of the theorem will be particularly useful in gate classification as postselection gadgets are often non-unitary [11].

1.1 Proof techniques

Our proof works in a similar manner to those in [16, 29]. The basic idea is to take an ϵ_0-approximation V of some gate U and improve it to an $O(\epsilon_0^2)$-approximation of U, while taking the length of the approximation from ℓ_0 to $c\ell_0$ for some constant c. Iterating this improvement step allows one to obtain a polylogarithmic overhead for compilation.[8] The key in any proof of a Solovay-Kitaev theorem is to make use of V in this construction in such a way that one does not incur $O(\epsilon_0)$ error in the resulting approximation, as one would naively have from the triangle inequality. In other words, one needs the error in the approximation of U to cancel out to lowest order in ϵ_0.

In the proof of the regular (inverse-closed) Solovay-Kitaev theorem, this is achieved using group commutators [16], which manifestly require inverses in the gate set. Sardharwalla *et al.* [29] instead achieve this by applying a group averaging function over the Weyl group. They show by direct computation that the lowest order error term in ϵ cancels out (at least in a neighborhood of the identity).

---

[7] Note the Weyl operators only form a group up to global phase, but as we only require a *projective irrep* they meet the criteria of our theorem.

[8] One can easily see the lengths of the gate sequences increase exponentially with each application of this operation, while the error decreases doubly exponentially, which implies the desired polylog dependence of the error.

In our proof, we also consider a group averaging function $f : \mathrm{SU}(d) \to \mathrm{SU}(d)$ based on some (projective) irrep $R : G \to \mathrm{SU}(d)$ of a finite group G:

$$f(W) := \prod_{g \in G} R(g) W R(g)^\dagger. \tag{1}$$

Our main technical contribution consists in showing that the lowest order error term cancels here, due to certain orthogonality relations obeyed by irreducible representations. We show this follows from the fact that the multiplicity of the trivial irrep in the adjoint action of any irrep is one. Therefore our proof both shows that efficient compilation can occur with a wider family of gate sets than was previously known, and also explains the mathematical reason that Sardharwalla *et al.*'s proof works as it does.

2 Proof of the main result

To aid the understanding of our main result, let us first briefly define the relevant notions from representation theory (see [33, 13] for further details).

A d-dimensional *representation* of a group G is a map $R : G \to \mathrm{U}(d)$ such that $R(g_1)R(g_2) = R(g_1 g_2)$ for all $g_2, g_2 \in G$. Similarly, R is a *projective representation* if it obeys this identity up to a global phase, i.e. $R(g_1)R(g_2) = e^{i\theta(g_1,g_2)}R(g_1 g_2)$ for some function $\theta : G \times G \to \mathbb{R}$. A representation R is *reducible* if there is a unitary map $U \in \mathrm{U}(d)$ and two other representations R_1 and R_2 of G such that $UR(g)U^\dagger = R_1(g) \oplus R_2(g)$ for all $g \in G$. If this is not the case, R is called *irreducible* (or *irrep* for short). Finally, if $A \subset B$ are two sets, we say that A is *dense* in B if for any $\varepsilon > 0$ and any $b \in B$ there exists $a \in A$ such that $\|a - b\| \leq \varepsilon$ for some suitable notion of distance $\|\cdot\|$.

▶ **Theorem 1** (Solovay-Kitaev theorem with an irrep instead of inverses). *For any fixed $d \geq 2$, suppose $\mathcal{G} \subset \mathrm{SU}(d)$ is a finite gate set which densely generates $\mathrm{SU}(d)$, and furthermore $\mathcal{G}$ contains a (projective) irrep of some finite group G. Then there is an algorithm which outputs an ϵ-approximation to any $U \in \mathrm{SU}(d)$ using merely $O(\mathrm{polylog}(1/\epsilon))$ elements from $\mathcal{G}$.*

Proof. By assumption, our gate set is of the form

$$\mathcal{G} := R(G) \cup \{U_1, U_2, \ldots, U_N\} \tag{2}$$

where $R(G) := \{R(g) : g \in G\}$ and $N \geq 0$ is some integer. Here

- $R : G \to \mathrm{SU}(d)$ is a projective irreducible representation of some finite group G,
- $U_i \in \mathrm{SU}(d)$ are some additional elements whose inverses $U_i^\dagger$ are not necessarily in $\mathcal{G}$.

For the sake of simplicity, we will assume that R is an actual irrep rather than a projective irrep (we describe how to generalize the proof to projective irreps in Appendix B). Note that by the $\mathcal{G} \subset \mathrm{SU}(d)$ assumption we implicitly require that the representation R is in $\mathrm{SU}(d)$ rather than in $\mathrm{U}(d)$. While many irreps are ruled out by this restriction, one can deal with such irreps by first converting them to projective irreps and then applying the techniques discussed in Appendix B. We divide the rest of the proof into several steps marked as below.

Original gate sequence. Given a gate $U \in \mathrm{SU}(d)$ which we wish to approximate to accuracy ϵ, we first run the usual Solovay-Kitaev algorithm (see Section 1) to obtain a sequence $S_{\epsilon/2}$ of gates whose product $\epsilon/2$-approximates U, using elements from $\mathcal{G}$ *and* their inverses. This sequence contains both elements from the set $R(G)$ (which is closed under inversion), as well as gates U_i and $U_i^\dagger$. All of these are in the gate set $\mathcal{G}$ except the $U_i^\dagger$—and there are only $O\big(\log^{3.97}(1/\epsilon)\big)$ many of these. To prove Theorem 1, it therefore suffices to give a

Solovay-Kitaev algorithm for approximating the $U_i^\dagger$ in terms of a sequence of $O(\text{polylog}(1/\epsilon))$ gates from the set $\mathcal{G}$.

More concretely, assume we show how to ϵ-approximate each $U_i^\dagger$ using $O(\log^c(1/\epsilon))$ gates from $\mathcal{G}$ for some constant $c > 0$. Then we can set $\epsilon' := \frac{\epsilon}{2}/O\big(\log^{3.97}(1/\epsilon)\big)$ and run this algorithm to ϵ'-approximate each $U_i^\dagger$ appearing in the sequence $S_{\epsilon/2}$ produced by the regular Solovay-Kitaev algorithm. If we substitute these approximations of $U_i^\dagger$ back into $S_{\epsilon/2}$, by the triangle inequality the existing error of $\epsilon/2$ in $S_{\epsilon/2}$ will be increased by another $\epsilon/2$ contributed jointly by all $U_i^\dagger$'s. These two contributions together give us the desired ϵ-approximation of U. Note that an ϵ'-approximation of $U_i^\dagger$ requires $O\big(\log^c(1/\epsilon)\big)$ gates.[9] Hence the ϵ-approximation to U in total will use $O\big(\log^{c+3.97}(1/\epsilon)\big)$ gates from $\mathcal{G}$.

Initial approximation of $U_i^\dagger$. Since $\mathcal{G}$ generates a dense subgroup of $\mathrm{SU}(d)$, there exists a finite length ℓ_0 such that length-ℓ_0 sequences of elements of $\mathcal{G}$ are ϵ_0-dense in $\mathrm{SU}(d)$, for a small fixed constant

$$\epsilon_0 := \frac{1}{6|G|(d-1)! + 2|G|^2}. \tag{3}$$

Let us pick among these sequences an initial ϵ_0-approximation of $U_i^\dagger$ and denote it by V. Then

$$\epsilon_0 \geq \|V - U_i^\dagger\| = \|VU_i - I\|, \tag{4}$$

where $\|\cdot\|$ denotes the *operator norm* which is unitarily invariant.

Symmetrization. Now consider the operator f on $\mathrm{SU}(d)$ defined by

$$f(W) := \prod_{g\in G} R(g) W R(g)^\dagger, \tag{5}$$

where the order of the products is taken arbitrarily, as long as the last (rightmost) element of the product corresponds to the identity element $e \in G$. We are interested in the action of f on VU_i. If we denote the difference in eq. (4) by $\mathcal{O} := VU_i - I$ and distribute the product in eq. (5) into several sums (with no $\mathcal{O}$'s, with a single copy of $\mathcal{O}$, two copies of $\mathcal{O}$, etc.), we get

$$f(VU_i) = \prod_{g\in G} R(g)(I+\mathcal{O})R(g)^\dagger \tag{6}$$

$$= I + \sum_{g\in G} R(g)\mathcal{O}R(g)^\dagger + \sum_{\substack{g,g'\in G\\ g<g'}} R(g)\mathcal{O}R(g)^\dagger R(g')\mathcal{O}R(g')^\dagger \tag{7}$$

$$+ \cdots + \prod_{g\in G} R(g)\mathcal{O}R(g)^\dagger, \tag{8}$$

where the order of terms in all products is inherited from eq. (5) and $g < g'$ refers to this order. Note that the number of terms with k copies of $\mathcal{O}$ is $\binom{|G|}{k}$.

If one were to naively apply the triangle inequality to this sum, one would obtain that

$$\|f(VU_i) - I\| \leq |G|\|\mathcal{O}\| + \binom{|G|}{2}\|\mathcal{O}\|^2 + \ldots \tag{9}$$

---

[9] One can easily see that $\log^c\big(\frac{\log^{3.97}(1/\epsilon)}{\epsilon}\big) = O\big(\log^c(1/\epsilon)\big)$ as the additional $\log^{3.97}(1/\epsilon)$ factor only adds lower order $\log\log(1/\epsilon)$ terms.

In other words, one would get that we have moved $f(VU_i)$ further from the identity than we started. To fix this, we will show that the first term of the above is actually much smaller—of order $\|\mathcal{O}\|^2$—and therefore our application of f has moved us closer to the identity. To see this, first note that using representation theory, one can show that the norm of the first-order term in eq. (8) is

$$\left\| \sum_{g\in G} R(g)\mathcal{O}R(g)^\dagger \right\| = \left\| |G| \frac{\operatorname{Tr}\mathcal{O}}{d} I \right\| \tag{10}$$

$$= |G| \frac{|\operatorname{Tr}(VU_i - I)|}{d}. \tag{11}$$

In other words, the traceless component of the first order term vanishes. This follows from certain orthogonality relations obeyed by irreps, and is proven in Claim 2 in Appendix A.

Next, we show that the trace of $\mathcal{O} = VU_i - I$ is small compared to its norm, namely

$$|\operatorname{Tr}(VU_i - I)| \leq (2^d + d!)\|VU_i - I\|^2. \tag{12}$$

This is proven in Claim 3 in Appendix A, and follows essentially because the Lie algebra of the special unitary group is traceless. Plugging this in to eq. (11), we see that

$$\left\| \sum_{g\in G} R(g)\mathcal{O}R(g)^\dagger \right\| \leq |G| \frac{2^d + d!}{d} \|VU_i - I\|^2 \tag{13}$$

$$\leq |G| \frac{2^d + d!}{d} \epsilon_0^2 \tag{14}$$

where we used Claim 3 to get the first inequality and eq. (4) to get the second.

Hence, by applying these results and then applying the triangle inequality to eq. (8) we get

$$\|f(VU_i) - I\| \leq |G| \frac{2^d + d!}{d} \epsilon_0^2 + \sum_{k=2}^{|G|} \epsilon_0^k \binom{|G|}{k} \tag{15}$$

$$\leq \left(|G| \frac{2^d + d!}{d} + \frac{|G|^2}{2} + |G|^2 \sum_{k=1}^{|G|-2} \epsilon_0^k |G|^k \right) \epsilon_0^2 \tag{16}$$

$$\leq \left(|G| \frac{2^d + d!}{d} + \frac{|G|^2}{2} + \frac{|G|^2}{2} \right) \epsilon_0^2 \tag{17}$$

Where in eq. (16) we used the fact that $\binom{|G|}{2} \leq \frac{|G|^2}{2}$ and $\binom{|G|}{k} \leq |G|^k$, and in eq. (17) we used the fact that $\epsilon_0 < \frac{1}{2|G|^2}$, so since $|G| > 2$ (as G has an irrep of dimension at least 2), we have that $\epsilon_0 |G| \leq 1/4$ so the geometric sum converges to a quantity $\leq \frac{1}{2}$.

Replacing this with a crude upper bound that $2^d \leq 2d!$ for $d > 1$, we get that

$$\|f(VU_i) - I\| \leq \left(3|G|(d-1)! + |G|^2\right)\epsilon_0^2 =: \epsilon_1 \tag{18}$$

Since we chose ϵ_0 to be $\frac{1}{2(3|G|(d-1)!+|G|^2)}$ in eq. (3), $\epsilon_1 \leq \frac{\epsilon_0}{2}$ – in other words $f(VU_i)$ is closer to the identity than VU_i.

Multiplying $f(VU_i) - I$ in eq. (18) by $U_i^\dagger$ on the right, we have that $f(VU_i)U_i^\dagger$ is an ϵ_1-approximation to $U_i^\dagger$. We chose the identity to come last in the definition of f in eq. (5), so the string of operators $f(VU_i)$ has the form

$$f(VU_i) = R(g_1)VU_iR(g_1)^\dagger R(g_2)VU_iR(g_2)^\dagger \cdots VU_i. \tag{19}$$

Since $U_i U_i^\dagger$ cancels at the end, $f(VU_i)U_i^\dagger$ is an ϵ_1-approximation to $U_i^\dagger$ using only terms from $\mathcal{G}$.

Iterative refinement. To complete the proof, we iterate this construction by considering

$$f^{(k)}(VU_i) := f(f(\cdots f(VU_i))). \tag{20}$$

Note from eq. (18) that $f^{(k)}(VU_i)U_i^\dagger$ is an ϵ_k-approximation to $U_i^\dagger$, where $\epsilon_k \leq (3|G|(d-1)! + |G|^2)\epsilon_{k-1}^2$. The length of the sequence $f^{(k)}$, denoted ℓ_k, obeys $\ell_k = |G|\ell_{k-1} + 2|G|$. Again $f^{(k)}(VU_i)U_i^\dagger$ can be expressed only in terms of elements of $\mathcal{G}$ (since the last U_i in the expansion of $f^{(k)}(VU_i)$ cancels with the rightmost $U_i^\dagger$ as before). One can easily show that these recurrence relations imply that as k grows:

- the approximation error ϵ_k shrinks doubly exponentially: $\epsilon_k \leq \frac{2\epsilon_0}{2^{2^k}}$;
- the length of the sequence ℓ_k grows exponentially: $\ell_k = O(|G|^k \ell_0)$.

Note that this sort of asymptotic behavior occurs simply because $\epsilon_k = O(\epsilon_{k-1}^2)$ while $\ell_k = O(\ell_{k-1})$ (though of course the value of ϵ_0 used in the recurrence may depend on the hidden constant in the big-O notation). This immediately implies that one can approximate $U_i^\dagger$ to accuracy ϵ with merely polylog overhead, as desired. More specifically, such approximation uses

$$O\big(\ell_0 \log^{\log_2 |G|}(1/\epsilon)\big) \tag{21}$$

elements of $\mathcal{G}$. By our analysis at the beginning of the proof, this gives a Solovay-Kitaev theorem with an exponent of $\log_2 |G| + 3.97$ in the polylog, completing the proof of Theorem 1. ◀

We have therefore shown that one can ϵ-approximate any $U_i^\dagger$ using only gates from our gate set $\mathcal{G}$ using merely polylog$(1/\epsilon)$ gates. The exponent of the polylog for approximating each $U_i^\dagger$ is again easily computed to be $O(\log_2 |G|)$. So putting this all together, our approximation for the overall unitary U requires

$$O\big(\log^{\log_2 |G|}(1/\epsilon)\big) \tag{22}$$

gates from $\mathcal{G}$. Note that the dependence on dimension d and order of the group G is hidden in the big-O notation, which hides a factor of ℓ_0, the length of sequences required to achieve an initial ϵ_0-net of SU(d). By a volume argument $\ell_0 = \Omega(d^2)$ [16]. In fact our choice of ϵ_0 implies that $\ell_0 = \Omega(d^3 \log d)$ in our construction.[10]

2.1 Extensions of our theorem

We have shown a Solovay-Kitaev theorem for any gate set $\mathcal{G}$ that contains an irrep of a finite group G, without requiring $\mathcal{G}$ to be inverse-closed. Our result can be easily generalized in two directions.

First, our proof also works if instead of an irrep we have a *projective* irrep. That is, a map $R : G \to \mathrm{SU}(d)$ such that, for any $g_1, g_2 \in G$,

$$R(g_1)R(g_2) = e^{i\theta(g_1,g_2)}R(g_1 g_2) \tag{23}$$

[10] Since an ϵ_0-ball occupies $\Theta(\epsilon_0^{d^2})$ volume in SU(d), $\ell_0 = \Omega\big(d^2 \log(1/\epsilon_0)\big)$ [16]. Since we set $\epsilon_0 = (2|G|^2 + 6|G|(d-1)!)^{-1}$ in eq. (3), we have that $\ell_0 = \Omega\big(d^3 \log d\big)$ since $\log d!$ scales as $O(d \log d)$ by Stirling's formula.

for some collection of phases[11] $\theta(g_1, g_2) \in [0, 2\pi)$. In such case one still has a Solovay-Kitaev theorem for any universal gate set that includes $R(G)$. For instance, the Pauli matrices $\{I, X, Y, Z\}$ form a projective irrep, but not an irrep (though the matrices $\{\pm 1, \pm i\} \cdot \{I, X, Y, Z\}$ do form an irrep). Since the exponent of the logarithm of our version of the Solovay-Kitaev theorem contains $\log_2 |G|$, this generalization improves the exponent (e.g. using the four Pauli matrices instead of the eight-element Pauli group improves the exponent by an additive 2). We give details on why projective irreps suffice in Appendix B.

Second, we note that our proof can be extended to the *special linear group* $\mathrm{SL}(d, \mathbb{C})$ as well. That is, one can also efficiently compile non-singular matrices, so long as a (projective) irrep is present in a gate set that is universal for $\mathrm{SL}(d, \mathbb{C})$. A Solovay-Kitaev Theorem (with inverses) for the special linear group was first shown by Aharnov, Arad, Eban and Landau [2], who used it to prove that additive approximations to the Tutte polynomial are BQP-hard in many regimes. It was also applied by [11] to the problem of classifying quantum gate sets, where it arose naturally because the "postselection gadgets" used in their proof are non-unitary. For a formal description of the non-unitary version of this theorem, please see Appendix C. Since postselection gadgets are often non-unitary [11], we likewise expect this version of the theorem will be more useful for gate classification problems.

3 Open problems

The main unresolved problem left by our work is to prove a generic inverse-free Solovay-Kitaev theorem, which has been a longstanding open problem [16, 23].

▶ **Conjecture** (Inverse-free Solovay-Kitaev theorem). *For any fixed $d \geq 2$, if $\mathcal{G} \subset \mathrm{SU}(d)$ is a finite gate set which densely generates $\mathrm{SU}(d)$, then there is an algorithm which outputs an ϵ-approximation to any $U \in \mathrm{SU}(d)$ using merely $O(\mathrm{polylog}(1/\epsilon))$ elements from $\mathcal{G}$.*

One can easily see that for any universal gate set (possibly without inverses), one can ϵ-approximate arbitrary unitaries with $O(1/\epsilon)$ overhead. This follows from simply running the Solovay-Kitaev theorem with inverses, and then approximating each inverse $W^\dagger$ with W^k for some integer k (which one can do with $O(1/\epsilon)$ overhead as this is simply composing irrational rotations about a single axis). However current approaches seem to be unable to improve this compilation algorithm from $O(1/\epsilon)$ to $\mathrm{polylog}(1/\epsilon)$. As discussed in Section 1.1, current proofs of the Solovay-Kitaev theorem require a special cancellation of error terms in order to convert an ϵ-approximation of some operator into an $O(\epsilon^2)$-approximation. This cancellation of error terms can be achieved by taking group commutators [16] or, as in this work and [29], it can be achieved by averaging over irreps and using the orthogonality of irreps. However, there is no known technique for achieving this sort of error cancellation without having some structure in the gate set.[12]

Additionally, a natural question is whether the value of ϵ_0 can be improved. This would improve the scaling of our result with dimension. In our result (and in the inverse-closed Solovay-Kitaev Theorem) the big-O notation hides a factor of ℓ_0—the length of the initial sequences required to achieve an ϵ_0-net. In our result ϵ_0 scales as $1/d!$, and hence a volume argument implies $\ell_0 = \Omega(d^3 \log d)$. In contrast the (inverse-closed) Solovay-Kitaev theorem

[11] The quantity $e^{i\theta(g_1,g_2)}$ is also known as a *Schur multiplier* of G.

[12] For example, Zhiyenbayev, Akulin, and Mandilara [36] have recently studied an alternative setting where instead of inverses a certain "isotropic" property of the gates is assumed.

merely requires $\epsilon_0 = \Theta(1)$ resulting in $\ell_0 = \Omega(d^2)$ [16]. It is a natural question if one can improve the value of ϵ_0 and therefore improve dimension dependence of our construction.

A somewhat simpler open problem is whether our theorem can be improved by considering particular orders of the group elements in eq. (5). The function $f(U)$ which we iterate when proving Theorem 1 is defined by averaging over the irrep of G in an arbitrary order; our theorem essentially works because if U is ϵ-close to the identity then $f(U)$ is $O(\epsilon^2)$-close to the identity. However, we have found by direct calculation that for the 2-dimensional irrep of S_3, considering particular orders of the group can lead to the $O(\epsilon^2)$ terms cancelling out as well, leaving only $O(\epsilon^3)$ terms. It is an interesting open problem if these additional cancellations can be generalized to other groups. If so, they would improve the $\log_2 |G|$ in the exponent of the logarithm of our result to $\log_k |G|$, where k is the lowest order remaining error term.

Finally, we note one may be able to extend our results to compilation over more general Lie groups, just as Kuperberg extended the inverse-closed Solovay-Kitaev theorem to arbitrary connected Lie groups whose Lie algebra is perfect [23]. We leave this as an open problem.

A Auxiliary claims

▶ **Claim 2.** *If R is a d-dimensional (projective) irrep of some finite group G and M is any $d \times d$ complex matrix then*

$$\sum_{g \in G} R(g) M R(g)^\dagger = |G| \frac{\operatorname{Tr} M}{d} I. \tag{24}$$

Proof. If R and R' are any two irreps of a finite group G, with dimensions d_R and $d_{R'}$ respectively, their matrix entries obey the following orthogonality relations [33]:

$$\frac{d_R}{|G|} \sum_{g \in G} R(g)_{ij} \overline{R'(g)_{kl}} = \delta_{RR'} \delta_{ik} \delta_{jl}, \qquad \forall i, j \in \{1, \ldots, d_R\}, \quad \forall k, l \in \{1, \ldots, d_{R'}\}. \tag{25}$$

In particular, if $R = R'$ and we write the matrix entries as $R(g)_{ij} = \langle i|R(g)|j\rangle$ then

$$\frac{d}{|G|} \sum_{g \in G} \langle i|R(g)|j\rangle \langle l|R(g)^\dagger|k\rangle = \delta_{ik} \delta_{jl}, \qquad \forall i, j, k, l \in \{1, \ldots, d\} \tag{26}$$

where $d := d_R = d_{R'}$. If we multiply both sides by $|i\rangle\langle k|$ and then sum over i and k, we get

$$\frac{d}{|G|} \sum_{g \in G} R(g)|j\rangle \langle l|R(g)^\dagger = I \delta_{jl}, \qquad \forall j, l \in \{1, \ldots, d\}. \tag{27}$$

If $M = \sum_{j,l=1}^{d} m_{jl} |j\rangle\langle l|$ then by linearity,

$$\frac{d}{|G|} \sum_{g \in G} R(g) M R(g)^\dagger = I \sum_{j,l=1}^{d} m_{jl} \delta_{jl} = I \operatorname{Tr} M, \tag{28}$$

which completes the proof. ◀

Another way to see this result is by noticing that the adjoint action of R decomposes as a direct sum of the trivial representation (acting on the 1-dimensional space spanned by the identity matrix) and a $(d^2 - 1)$-dimensional representation without any trivial component. This follows from Schur's first lemma. The result then follows by the orthogonality relations obeyed by the irrep decomposition of the adjoint action.

▶ **Claim 3.** *If* $M \in \mathrm{SL}(d,\mathbb{C})$ *then* $|\mathrm{Tr}\, M - d| \leq (2^d + d!)\|M - I\|^2$.

Proof. Let $A := M - I$ and denote the entries of A by a_{ij} where $i, j = 1, \ldots, d$. We know that $1 = \det M = \det(A + I)$, so expanding in terms of the a_{ij}'s, we have that

$$1 = \sum_{\sigma \in \mathrm{S}_d} \mathrm{sgn}(\sigma) \prod_{i=1}^{d} (a_{i\sigma(i)} + \delta_{i\sigma(i)}). \tag{29}$$

Now let us simply take out the term with $\sigma = \varepsilon$, the identity permutation:

$$1 = \prod_{i=1}^{d} (a_{ii} + 1) + \sum_{\sigma \in \mathrm{S}_d \setminus \{\varepsilon\}} \mathrm{sgn}(\sigma) \prod_{i=1}^{d} (a_{i\sigma(i)} + \delta_{i\sigma(i)}). \tag{30}$$

And now expanding the first term we see

$$1 = 1 + \sum_{i=1}^{d} a_{ii} + \sum_{i \neq j} a_{ii} a_{jj} + \cdots + a_{11} a_{22} \cdots a_{dd} + \sum_{\sigma \in \mathrm{S}_d \setminus \{\varepsilon\}} \mathrm{sgn}(\sigma) \prod_{i=1}^{d} (a_{i\sigma(i)} + \delta_{i\sigma(i)}), \tag{31}$$

which implies

$$-\mathrm{Tr}\, A = \sum_{i \neq j} a_{ii} a_{jj} + \cdots + a_{11} a_{22} \cdots a_{dd} + \sum_{\sigma \in \mathrm{S}_d \setminus \{\varepsilon\}} \mathrm{sgn}(\sigma) \prod_{i=1}^{d} (a_{i\sigma(i)} + \delta_{i\sigma(i)}). \tag{32}$$

Now observe that each of the terms on the right hand side is quadratic in the a_{ij}'s—this is because any non-identity permutation displaces at least two items. Let $c \leq 2^d + d!$ denote the number of the terms present, which is constant in any fixed dimension d. Hence we have that

$$|\mathrm{Tr}\, M - d| = |\mathrm{Tr}\, A| \leq c \max_{i,j} |a_{ij}|^2 \leq c\|A\|^2 = c\|M - I\|^2 \tag{33}$$

where we used $|a_{ij}| \leq \|A\|$ in the last inequality (this follows by choosing the j-th standard basis vector in the definition of the operator norm). ◀

Note that this claim, i.e. that elements ϵ-close to the identity have trace substantially smaller than ϵ, is a reflection of the fact that the Lie algebra of the special linear group is traceless.

B Representations vs projective representations

Throughout our proof of Theorem 1, we assumed that R is an irrep of the group G. Here we show that the same construction works also for a projective irrep of G. In other words, even if $R(g_1)R(g_2) = e^{i\theta(g_1,g_2)} R(g_1 g_2)$ for some phase $\theta(g_1, g_2) \in [0, 2\pi)$, our version of the Solovay-Kitaev theorem still holds. As the Weyl operators merely form a projective representation, this allows our result to strictly generalize that of [29]. Intuitively, such generalization is to be expected since global phases are non-physical in quantum theory. We make this precise below.

Suppose that we have a projective representation R of a finite group G. It is convenient to think of $R(G)$ as a subset of the *projective unitary group* $\mathrm{PU}(d)$ that consists of equivalence classes of elements of $\mathrm{U}(d)$ that differ only by global phase. Note that $\mathrm{PU}(d) = \mathrm{PSU}(d)$, the *special* projective unitary group, since $\overline{\det(U)} U \in \mathrm{SU}(d)$ for any $U \in \mathrm{U}(d)$. Now, consider

extending the projective representation R in $\mathrm{PSU}(d)$ into a representation[13] R' in $\mathrm{SU}(d)$. Since

$$\mathrm{PSU}(d) = \mathrm{SU}(d)/\mathbb{Z}_d, \tag{34}$$

i.e. the only difference between projective and non-projective representations are factors of $e^{2\pi i/d} I$, this merely increases the size of the group by an integer multiple k which is a divisor of d. Let us denote this larger group by G'.

Now consider applying our proof of Theorem 1 to R' and G'. The corresponding averaging operator is

$$f'(W) := \prod_{g \in G'} R'(g) W R'(g)^\dagger. \tag{35}$$

Our proof essentially uses two facts:

1. The trace of W is small relative to its distance from the identity (Claim 3).
2. The traceless component of W vanishes to lowest order because from Claim 2 we have that for any tracless $\mathcal{O}$,

$$\sum_{g \in G'} R'(g) \mathcal{O} R'(g)^\dagger = 0. \tag{36}$$

Note that if $g, h \in G'$ are such that $R'(g) = e^{i\theta} R'(h)$ for some $\theta \in \mathbb{R}$, then they contribute identical terms in the above sum, since the global phase factors commute through and cancel out. Since the any projectively equivalent group elements g, h contribute the same quantity to the sum, and G' is simply a (projective) k-fold cover of G, this means that we can rewrite eq. (36) as

$$k \sum_{g \in G} R'(g) \mathcal{O} R'(g)^\dagger = 0, \tag{37}$$

where we have simply summed over one representative from each set of projectively equivalent representatives.

Therefore, if we had instead considered averaging over the projective representation only using the original averaging operator (which involves a factor k fewer products),

$$f(W) := \prod_{g \in G} R(g) W R(g)^\dagger, \tag{38}$$

the corresponding sum in eq. (36) (which is the above sum divided by k) would be 0 as well. Therefore, the cancellation of lowest-order terms for the traceless component of the error—i.e. the second fact listed above—still holds. Furthermore, the first fact is true independent of the group G considered, and is simply a fact about matrices of determinant 1 which are close to the identity. Therefore, the proof of Theorem 1 works exactly as before if R is merely a projective representation.

C Extension to the special linear group

In this appendix we describe how to extend our proof of Theorem 1 to the non-unitary case. Namely, we want to approximate some matrix $M \in \mathrm{SL}(d, \mathbb{C})$, our gate set $\mathcal{G} \subset \mathrm{SL}(d, \mathbb{C})$ is

[13] This is known as a *central extension* of the representation.

dense in $\mathrm{SL}(d,\mathbb{C})$, and it contains a (possibly non-unitary) irrep of a finite group G as well as some additional gates $U_i \in \mathrm{SL}(d,\mathbb{C})$.

Let us argue that an ϵ-approximation of M can be obtained using the same algorithm as in the proof of Theorem 1, but with one minor change. Namely, in the first step of the algorithm one must apply the non-unitary Solovay-Kitaev Theorem (with inverses) of Aharonov, Arad, Eban, and Landau [2] rather than the usual unitary Solovay-Kitaev Theorem (with inverses). As before, the problem therefore reduces to finding an expression for the elements U_i^{-1} in terms of $\mathcal{G}$. Note that no other step of our proof requires any matrices to be unitary! Recall that the heart of the proof was in showing that if VU_i is ϵ-close to I then $f(VU_i)$ is $O(\epsilon^2)$-close to I, where V denotes the initial ϵ_0-approximation of U_i^{-1}. The key facts that we used to show this are:

- Claim 2, which states that the traceless component of VU_i vanishes to first order under the application of f due to the orthogonality of irreps.
- Claim 3, which states that matrices of determinant 1 which are ϵ-close to the identity have trace $O(\epsilon^2)$.

Neither of these depends on the matrices involved being unitary—indeed the Schur orthogonality relations between irreps in eq. (25) also hold for non-unitary irreps. Therefore, our proof implies the following:

▶ **Theorem 4.** *For any fixed $d \geq 2$, suppose $\mathcal{G} \subset \mathrm{SL}(d,\mathbb{C})$ is a finite gate set that contains a (projective) irrep of some finite group G. Let $r > 0$ be any fixed radius, let B_r be the ball of radius r about the identity in $\mathrm{SL}(d,\mathbb{C})$, and suppose that $\mathcal{G}$ densely generates all transformations in B_r. Then there is an algorithm which outputs an ϵ-approximation to any $M \in B_r$ using merely $O(\mathrm{polylog}(1/\epsilon))$ elements from $\mathcal{G}$.*

Other than the replacement of $\mathrm{SU}(d)$ with $\mathrm{SL}(d,\mathbb{C})$, the only thing that differs between this theorem and Theorem 1 is the additional restriction that the matrix M we are approximating is a finite distance from the identity (as is present in the non-unitary Solovay-Kitaev theorem of [2] as well). This restriction arises simply because $\mathrm{SL}(d,\mathbb{C})$ is not compact, and approximating elements very far from the identity requires longer sequences of gates. For instance, it requires more applications of the gate $\left(\begin{smallmatrix} 2 & 0 \\ 0 & 1/2 \end{smallmatrix}\right)$ to reach $\left(\begin{smallmatrix} 2^{1000} & 0 \\ 0 & 2^{-1000} \end{smallmatrix}\right)$ than it requires to reach $\left(\begin{smallmatrix} 2^2 & 0 \\ 0 & 2^{-2} \end{smallmatrix}\right)$. Since points arbitrarily far from the identity require arbitrarily long gate sequences to approximate, one cannot upper bound the length of sequences required to ϵ-approximate arbitrary $M \in \mathrm{SL}(d,\mathbb{C})$ as a function of ϵ only—rather the length would depend on the distance of M to the identity as well. Restricting M's distance to the identity allows one to upper bound the length of the approximating sequence in terms of ϵ only.

References

1. Scott Aaronson and Alex Arkhipov. The computational complexity of linear optics. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 333–342. ACM, 2011. doi:10.1145/1993636.1993682.
2. Dorit Aharonov, Itai Arad, Elad Eban, and Zeph Landau. Polynomial quantum algorithms for additive approximations of the Potts model and other points of the Tutte plane, 2007. arXiv:quant-ph/0702008.
3. Dorit Aharonov and Michael Ben-Or. Fault-tolerant quantum computation with constant error rate. *SIAM Journal on Computing*, 38(4):1207–1282, 2008. doi:10.1137/S0097539799359385.

4 Dorit Aharonov, Vaughan Jones, and Zeph Landau. A polynomial quantum algorithm for approximating the Jones polynomial. *Algorithmica*, 55(3):395–421, Nov 2009. doi:10.1007/s00453-008-9168-0.

5 Johannes Bausch, Toby Cubitt, and Maris Ozols. The complexity of translationally invariant spin chains with low local dimension. *Annales Henri Poincaré*, 18(11):3449–3513, Nov 2017. doi:10.1007/s00023-017-0609-7.

6 Johannes Bausch and Stephen Piddock. The complexity of translationally invariant low-dimensional spin lattices 3D. *Journal of Mathematical Physics*, 58(11):111901, 2017. doi:10.1063/1.5011338.

7 Dominic W. Berry, Andrew M. Childs, and Robin Kothari. Hamiltonian simulation with nearly optimal dependence on all parameters. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 792–809. IEEE, Oct 2015. doi:10.1109/FOCS.2015.54.

8 Alex Bocharov, Martin Roetteler, and Krysta M. Svore. Efficient synthesis of probabilistic quantum circuits with fallback. *Phys. Rev. A*, 91(5):052317, May 2015. doi:10.1103/PhysRevA.91.052317.

9 Alex Bocharov, Martin Roetteler, and Krysta M. Svore. Efficient synthesis of universal repeat-until-success quantum circuits. *Phys. Rev. Lett.*, 114(8):080502, Feb 2015. doi:10.1103/PhysRevLett.114.080502.

10 Adam Bouland, Joseph F. Fitzsimons, and Dax E. Koh. Complexity classification of conjugated Clifford circuits. *Proc. 33rd Computational Complexity Conference (CCC)*, 2018. arXiv:1709.01805.

11 Adam Bouland, Laura Mančinska, and Xue Zhang. Complexity classification of two-qubit commuting Hamiltonians. In Ran Raz, editor, *31st Conference on Computational Complexity (CCC 2016)*, volume 50 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 28:1–28:33, Dagstuhl, Germany, 2016. Schloss Dagstuhl–Leibniz-Zentrum für Informatik. doi:10.4230/LIPIcs.CCC.2016.28.

12 Michael J. Bremner, Richard Jozsa, and Dan J. Shepherd. Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, 2010. doi:10.1098/rspa.2010.0301.

13 Andrew M. Childs. Fourier analysis in nonabelian groups. Lecture notes at University of Waterloo, 2013. URL: http://www.cs.umd.edu/~amchilds/teaching/w13/l06.pdf.

14 Andrew M. Childs. Lecture notes on quantum algorithms. Lecture notes at University of Maryland, 2017. URL: http://www.cs.umd.edu/~amchilds/qa/qa.pdf.

15 Andrew M. Childs, Debbie Leung, Laura Mančinska, and Maris Ozols. Characterization of universal two-qubit Hamiltonians. *Quantum Information & Computation*, 11(1&2):19–39, Jan 2011. arXiv:1004.1645.

16 Christopher M. Dawson and Michael A. Nielsen. The Solovay-Kitaev algorithm. *Quantum Information & Computation*, 6(1):81–95, Jan 2006. arXiv:quant-ph/0505030.

17 Daniel Gottesman and Sandy Irani. The quantum and classical complexity of translationally invariant tiling and Hamiltonian problems. *Theory of Computing*, 9(2):31–116, 2013. doi:10.4086/toc.2013.v009a002.

18 Aram W. Harrow, Benjamin Recht, and Isaac L. Chuang. Efficient discrete approximations of quantum gates. *Journal of Mathematical Physics*, 43(9):4445–4451, 2002. doi:10.1063/1.1495899.

19 Alexei Yu. Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191–1249, 1997. doi:10.1070/RM1997v052n06ABEH002155.

20 Alexei Yu. Kitaev, Alexander Shen, and Mikhail N. Vyalyi. *Classical and Quantum Computation*, volume 47 of *Graduate Studies in Mathematics*. American Mathematical Society, 2002. URL: https://books.google.com/books?id=qYHTvHPvmG8C.

21 Vadym Kliuchnikov. Synthesis of unitaries with Clifford+T circuits, 2013. arXiv:1306.3200.

22 Vadym Kliuchnikov, Dmitri Maslov, and Michele Mosca. Practical approximation of single-qubit unitaries by single-qubit quantum Clifford and T circuits. *IEEE Transactions on Computers*, 65(1):161–172, Jan 2016. doi:10.1109/TC.2015.2409842.

23 Greg Kuperberg. How hard is it to approximate the Jones polynomial? *Theory of Computing*, 11(6):183–219, 2015. doi:10.4086/toc.2015.v011a006.

24 Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010. URL: https://books.google.com/books?id=-s4DEy7o-a0C.

25 Māris Ozols. The Solovay-Kitaev theorem. Essay at University of Waterloo, 2009. URL: http://home.lu.lv/~sd20008/papers/essays/Solovay-Kitaev.pdf.

26 Adam Paetznick and Krysta M. Svore. Repeat-until-success: Non-deterministic decomposition of single-qubit unitaries. *Quantum Information & Computation*, 14(15&16):1277–1301, 2014. arXiv:1311.1074.

27 Ori Parzanchevski and Peter Sarnak. Super-Golden-Gates for PU(2). *Advances in Mathematics*, 327:869–901, 2018. Special volume honoring David Kazhdan. doi:10.1016/j.aim.2017.06.022.

28 Neil J. Ross and Peter Selinger. Optimal ancilla-free Clifford+T approximation of Z-rotations. *Quantum Information & Computation*, 16(11&12):0901–0953, 2016. doi:10.26421/QIC16.11-12.

29 Imdad S.B. Sardharwalla, Toby S. Cubitt, Aram W. Harrow, and Noah Linden. Universal refocusing of systematic quantum noise, 2016. arXiv:1602.07963.

30 Peter Sarnak. Letter to Aaronson and Pollington on the Solvay-Kitaev Theorem and Golden Gates, 2015. URL: http://publications.ias.edu/sarnak/paper/2637.

31 Adam Sawicki and Katarzyna Karnas. Universality of single-qudit gates. *Annales Henri Poincaré*, Aug 2017. doi:10.1007/s00023-017-0604-z.

32 Peter Selinger. Efficient Clifford+T approximation of single-qubit operators. *Quantum Information & Computation*, 15(1&2):159–180, 2015. arXiv:1212.6253.

33 Jean-Pierre Serre. *Linear Representations of Finite Groups*. Springer, 2012. URL: https://books.google.com/books?id=9mT1BwAAQBAJ.

34 Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997. doi:10.1137/S0097539795293172.

35 Péter Pál Varjú. Random walks in compact groups. *Documenta Mathematica*, 18:1137–1175, 2013. arXiv:1209.1745.

36 Y. Zhiyenbayev, V. M. Akulin, and A. Mandilara. Quantum compiling with diffusive sets of gates, 2017. arXiv:1708.08909.