Computational complexity and 3-manifolds and zombies

Greg Kuperberg[*] and Eric Samperton[†]

University of California, Davis

(Dated: December 30, 2017)

We show the problem of counting homomorphisms from the fundamental group of a homology 3-sphere M to a finite, non-abelian simple group G is almost parsimoniously #P-complete, when G is fixed and M is the computational input. In the reduction, we guarantee that every non-trivial homomorphism is a surjection. As a corollary, any non-trivial information about the number of non-trivial homomorphisms is computationally intractable assuming standard conjectures in computer science. In particular, deciding if there is a non-trivial homomorphism is NP-complete. Another corollary is that for any fixed integer $m \geq 5$, it is NP-complete to decide whether M admits a connected m-sheeted covering.

Given a classical reversible circuit C, we construct M so that evaluations of C with certain initialization and finalization conditions correspond to homomorphisms $\pi_1(M) \to G$. An intermediate state of C likewise corresponds to a homomorphism $\pi_1(\Sigma_g) \to G$, where Σ_g is a Heegaard surface of M of genus g. We analyze the action on these homomorphisms by the pointed mapping class group $\mathrm{MCG}_*(\Sigma_g)$ and its Torelli subgroup $\mathrm{Tor}_*(\Sigma_g)$. Using refinements of results of Dunfield-Thurston, we show that the actions of these groups are as large as possible when g is large. Our results and our construction are inspired by universality results in topological quantum computation, even though the present work is non-quantum.

1. INTRODUCTION

1.1. Statement of Results

Given a finite group G and a path-connected topological space X, let

$$H(X, G) = \{f : \pi_1(X) \to G\}$$

be the set of homomorphisms from the fundamental group of X to G. Then the number $\#H(X, G) = |H(X, G)|$ is an important topological invariant of X. For example, in the case that X is a knot complement and $G = \mathrm{Sym}(n)$ is a symmetric group, $\#H(X, G)$ was useful for compiling a table of knots with up to 15 crossings [34]. (We use both notations $\#S$ and $|S|$ to denote the cardinality of a finite set S, the former to emphasize algorithmic counting problems.)

Although these invariants can be powerful, our main result is that they are often computationally intractable, assuming that $\mathsf{P} \neq \mathsf{NP}$. We review certain considerations:

- We suppose that X is given by a finite triangulation, as a reasonable standard for computational input.

- We are interested in the case that $\#H(X, G)$ is intractable when G is fixed and X is the only computational input. We are also more interested in the case when $\#H(X, J)$ is trivial for every proper subgroup $J < G$.

- If G is abelian, then $\#H(X, G)$ is determined by the integral homology group $H_1(X) = H_1(X; \mathbb{Z})$; both of these invariants can be computed in polynomial time (Theorem 2.5). We are thus more interested in the case that $H_1(X) = 0$ and G is perfect, in particular when G is non-abelian simple.

- If X is a simplicial complex, or even an n-manifold with $n \geq 4$, then $\pi_1(X)$ can be any finitely presented group. By contrast, 3-manifold groups are highly restricted. We are more interested in the case that $X = M$ is a 3-manifold. If in addition M is closed and $H_1(M) = 0$, then M is a homology 3-sphere.

To state our main result, we pass to the related invariant $\#Q(X, G) = |Q(X, G)|$, where $Q(X, G)$ is the set of normal subgroups $\Gamma \trianglelefteq \pi_1(X)$ such that the quotient $\pi_1(X)/\Gamma$ is isomorphic to G.

Theorem 1.1. *Let G be a fixed, finite, non-abelian simple group. If M is a triangulated homology 3-sphere regarded as computational input, then the invariant $\#Q(M, G)$ is #P-complete via a parsimonious reduction. The reduction also guarantees that $\#Q(M, J) = 0$ for any non-trivial, proper subgroup $J < G$.*

Section 2.1 gives more precise definitions of the complexity theory concepts in Theorem 1.1. Briefly, a counting problem is in #P if there is a polynomial-time algorithm to verify the objects being counted; it is #P-hard if it is as hard as any counting problem in #P; and it is #P-complete if it is both in #P and #P-hard. A *parsimonious reduction* from a counting problem g to a counting problem f (to show that f is as hard as g) is a mapping h, computable in polynomial time, such that $g(x) = f(h(x))$. This standard of hardness tells us not only that $\#Q(M, G)$ is computationally intractable, but also that any partial information from it is intractable, for instance, its parity. (See Theorem 2.1.) An even stricter standard is a *Levin reduction*, which asks for a bijection between the objects being counted that is computable in polynomial time (in both directions). In fact, our proof of Theorem 1.1 yields a Levin reduction from any problem in #P to the problem $\#Q(M, G)$.

Another point of precision is that Theorem 1.1 casts $\#Q(M, G)$ as a promise problem, requiring the promise that the simplicial complex input describes a 3-manifold and more specifically a homology 3-sphere. Since this promise can be

checked in polynomial time (Proposition 2.6), this is equivalent to a non-promise problem (since an algorithm to calculate $\#Q(M,G)$ can reject input that does not satisfy the promise).

The invariants $\#H(X,G)$ and $\#Q(X,G)$ are related by the following equation:

$$|H(X,G)| = \sum_{J \leq G} |\operatorname{Aut}(J)| \cdot |Q(X,J)|. \qquad (1)$$

If $\pi_1(X)$ has no non-trivial surjections to any simple group smaller than G, as Theorem 1.1 can provide, then

$$|H(X,G)| = |\operatorname{Aut}(G)| \cdot |Q(X,G)| + 1. \qquad (2)$$

Thus we can say that $\#H(M,G)$ is *almost parsimoniously* #P-complete for homology 3-spheres. It is parsimonious except for the trivial homomorphism and up to automorphisms of G, which are both minor, unavoidable corrections. This concept appears elsewhere in complexity theory; for instance, the number of 3-colorings of a planar graph is almost parsimoniously #P-complete [6].

In particular, the fact that $\#Q(M,G)$ is parsimoniously #P-hard implies that existence is Karp NP-hardness (again see Section 2.1). Thus Theorem 1.1 has the following corollary.

Corollary 1.2. *Let G be a fixed, finite, non-abelian simple group, and let M be a triangulated homology 3-sphere regarded as computational input. Then it is Karp NP-complete to decide whether there is a non-trivial homomorphism $f : \pi_1(M) \to G$, even with the promise that every such homomorphism is surjective.*

Corollary 1.2 in turn has a corollary concerning connected covering spaces. In the proof of the corollary and later in the paper, we let $\operatorname{Sym}(n)$ be the symmetric group and $\operatorname{Alt}(n)$ be the alternating group, both acting on n letters.

Corollary 1.3. *For each fixed $n \geq 5$, it is NP-complete to decide whether a homology 3-sphere M has a connected n-sheeted cover, even with the promise that it has no connected k-sheeted cover with $1 < k < n$.*

Proof. Recall that $\operatorname{Alt}(n)$ is simple when $n \geq 5$. The n-sheeted covers $\tilde{M}$ of M are bijective with homomorphisms $f : \pi_1(M) \to \operatorname{Sym}(n)$, considered up to conjugation in $\operatorname{Sym}(n)$. If M is a homology 3-sphere, then $\pi_1(M)$ is a perfect group and we can replace $\operatorname{Sym}(n)$ by $\operatorname{Alt}(n)$. If $\tilde{M}$ is disconnected, then f does not surject onto $\operatorname{Alt}(n)$. Thus, we can apply Corollary 1.2 with $G = \operatorname{Alt}(n)$. $\square$

The idea of our proof of Theorem 1.1 is as follows. Let Σ_g be a standard oriented surface of genus g with a marked basepoint, and let G be a (not necessarily simple) finite group. Then we can interpret the set of homomorphisms, or *representation set*,

$$\hat{R}_g(G) \stackrel{\text{def}}{=} H(\Sigma_g, G) = \{f : \pi_1(\Sigma_g) \to G\}$$

as roughly the set of states of a computer memory. We can interpret a word in a fixed generating set of the pointed, oriented mapping class group $\operatorname{MCG}_*(\Sigma_g)$ as a reversible digital circuit
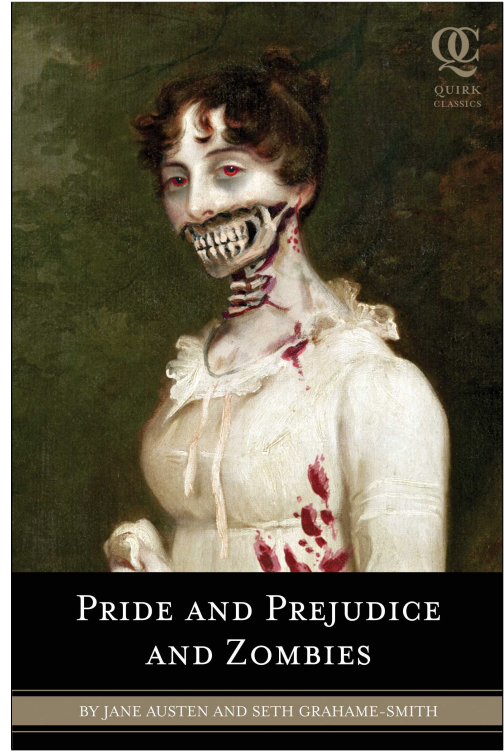

Figure 1. A classic novel, with zombies [5].

acting on $\hat{R}_g(G)$, the set of memory states. (See Section 2.2 for a discussion of complexity of circuits and reversible circuits.) Every closed, oriented 3-manifold M can be constructed as two handlebodies $(H_g)_I$ and $(H_g)_F$ that are glued together by an element $\phi \in \operatorname{MCG}_*(\Sigma_g)$. We can interpret ϕ as a reversible digital circuit in which the handlebodies partially constrain the input and output.

To understand the possible effect of ϕ, we want to decompose $\hat{R}_g(G)$ into $\operatorname{MCG}_*(\Sigma_g)$-invariant subsets. The obvious invariant of $f \in \hat{R}_g(G)$ is its image $f(\pi_1(\Sigma_g)) \leq G$; to account for it, we first restrict attention to the subset

$$R_g(G) \stackrel{\text{def}}{=} \{f : \pi_1(\Sigma_g) \to G\} \subseteq \hat{R}_g(G)$$

consisting of surjective homomorphisms.

We must also consider a less obvious invariant. Let BG be the classifying space of G, and recall that the group homology $H_*(G) = H_*(G; \mathbb{Z})$ can be defined as the topological homology $H_*(BG)$. Recall that a homomorphism $f : \pi_1(\Sigma_g) \to G$ corresponds to a map $f : \Sigma_g \to BG$ which is unique up to pointed homotopy. Every $f \in \hat{R}_g(G)$ then yields a homology class

$$\operatorname{sch}(f) \stackrel{\text{def}}{=} f_*([\Sigma_g]) \in H_2(G),$$

which we call the *Schur invariant* of f; it is $\operatorname{MCG}_*(\Sigma_g)$-invariant. Given $s \in H_2(G)$, the subset

$$R_g^s(G) \stackrel{\text{def}}{=} \{f \in R_g \mid \operatorname{sch}(f) = s\}$$

is then also $\operatorname{MCG}_*(\Sigma_g)$-invariant. Note that $\operatorname{sch}(f)$ is not always $\operatorname{Aut}(G)$-invariant because $\operatorname{Aut}(G)$ may act non-trivially

on $H_2(G)$. Fortunately, $R_g^0(G)$ is always $\text{Aut}(G)$-invariant. We summarize the relevant results of Dunfield-Thurston in the following theorem.

Theorem 1.4 (Dunfield-Thurston [14, Thms. 6.23 & 7.4])**.** *Let G be a finite group.*

1. *For every sufficiently large g (depending on G), $\text{MCG}_*(\Sigma_g)$ acts transitively on $R_g^s(G)$ for every $s \in H_2(G)$.*

2. *If G is non-abelian and simple, then for every sufficiently large g, the image of the action of $\text{MCG}_*(\Sigma_g)$ on $R_g^0(G)/\text{Aut}(G)$ is $\text{Alt}(R_g^0(G)/\text{Aut}(G))$.*

To make effective use of Theorem 1.4, we strengthen its second part in three ways to obtain Theorem 4.2. First, Theorem 1.4 holds for the pointed Torelli group $\text{Tor}_*(\Sigma_g)$. Second, we define an analogue of alternating groups for G-sets that we call *Rubik groups*, and we establish Theorem 3.10, a non-trivial structure theorem to generate a Rubik group. Theorem 4.2 gives a lift of the image of $\text{MCG}_*(\Sigma_g)$ from $\text{Alt}(R_g^0(G)/\text{Aut}(G))$ to the Rubik group $\text{Rub}_{\text{Aut}(G)}(R_g^0(G))$. Third, we still obtain the image $\text{Rub}_{\text{Aut}(G)}(R_g^0(G))$ even if we restrict to the subgroup of $\text{Tor}_*(\Sigma_g)$ that pointwise fixes $\hat{R}_g(G) \smallsetminus R_g(G)$, the set of non-surjective homomorphisms.

As a warm-up for our proof of Theorem 1.1, we can fix g, and try to interpret

$$A = R_g^0(G)/\text{Aut}(G)$$

as a computational alphabet. If g is large enough, then we can apply Theorem 1.4 to $R_{2g}^0(G)$ to obtain a universal set of reversible binary gates that act on $A^2 \subset R_{2g}^0(G)/\text{Aut}(G)^2$, implemented as mapping class elements or *gadgets*. (A gadget in computational complexity is an informal concept that refers to a combinatorial component of a complexity reduction.) The result can be related to a certain reversible circuit model $\text{RSAT}_{A,I,F}$. (See Section 2.2. The #P-hardness of RSAT, established in Theorem 2.4, is a standard result but still takes significant work.) We can convert a reversible circuit of width n to an element $\phi \in \text{MCG}_*(\Sigma_{ng})$ that acts on A^n, and then make M from ϕ. In this way, we can reduce $\#\text{RSAT}_{A,I,F}$ to $\#Q(M,G)$.

For our actual reduction, we will need to take steps to address three issues, which correspond to the three ways that Theorem 4.2 is sharper than Theorem 1.4.

- We want the larger calculation in $\hat{R}_{ng}(G)$ to avoid symbols in $\hat{R}_g(G) \smallsetminus R_g^0(G)$ that could contribute to $\#Q(M,G)$.

- We want a parsimonious reduction to $\#Q(M,G)$, which means that we must work with $R_g^0(G)$ rather than its quotient A.

- Mapping class gadgets should be elements of the Torelli group, to guarantee that M is a homology 3-sphere.

To address the first issue: We can avoid states in $R_g^s(G)$ with $s \neq 0$ because, if a surface group homomorphism $f : \pi_1(\Sigma_g) \to G$ has $\text{sch}(f) \neq 0$, then it cannot extend over a handlebody. If $f(G)$ has a non-trivial abelianization, then the fact that we will produce a homology 3-sphere will kill its participation. If f is not surjective but $f(G)$ is perfect, then we will handle this case by acting trivially on $R_g(K)$ for a simple quotient K of $f(G)$. The trivial homomorphism $z \in \hat{R}_g(G)$ is particularly problematic because it cannot be eliminated using the same techniques; we call it the *zombie symbol*. (See also Figure 1 for another source of our ideas.) We define an ad hoc reversible circuit model, ZSAT, that has zombie symbols. We reduce RSAT to ZSAT by converting the zombie symbols to warning symbols that do not finalize, unless all of the symbols are zombies. The full construction, given in Lemmas 4.1 and 4.6, is more complicated because these steps must be implemented with binary gates in $\text{MCG}_*(\Sigma_{2g})$ rather than unary gates in $\text{MCG}_*(\Sigma_g)$.

To address the second issue: A direct application of Theorem 1.4 would yield a factor of $|\text{Aut}(G)|^n$ in the reduction from $\#\text{RSAT}_{A,I,F}$ to $\#H(M,G)$, when the input is a reversible circuit of width n. We want to reduce this to a single factor of $|\text{Aut}(G)|$ in order to construct a parsimonious reduction to $\#Q(M,G)$. The ZSAT model also has an action of $K = \text{Aut}(G)$ on its alphabet to model this. Lemma 4.1 addresses the problem by relying on the Rubik group refinement in Theorem 4.2, and by creating more warning symbols when symbols are misaligned relative to the group action.

To ensure that the resulting manifold is a homology 3-sphere, we implement gates in the pointed Torelli subgroup $\text{Tor}_*(\Sigma_g)$ of $\text{MCG}_*(\Sigma_g)$. This is addressed in Theorem 4.2. Recall that $\text{Tor}_*(\Sigma_g)$ is the kernel of the surjective homomorphism

$$f : \text{MCG}_*(\Sigma_g) \to H_1(\Sigma_g) \cong \text{Sp}(2g,\mathbb{Z})$$

where $H_1(\Sigma_g)$ is equipped with its integral symplectic intersection form. The proof of Theorem 4.2 uses rigidity properties of $\text{Sp}(2g,\mathbb{Z})$ combined with Goursat's lemma (Lemma 3.3).

$$\text{CSAT} \xrightarrow{\text{Sec. 2.2}} \text{RSAT} \xrightarrow{\text{Sec. 4.1}} \text{ZSAT} \xrightarrow{\text{Sec. 4.3}} \#Q(M,G)$$

Figure 2. The reductions in the proof of Theorem 1.1.

Figure 2 summarizes the main reductions in the proof of Theorem 1.1.

1.2. Related work

As far as we know, the closest prior result to our Theorem 1.1 is due to Krovi and Russell [28]. Given a link $L \subseteq S^3$, they consider a refinement $\#H(S^3 \smallsetminus L, G, C)$ of $\#H(S^3 \smallsetminus L, G)$ in which they only count the group homomorphisms that send the meridian elements of $\pi_1(S^3 \smallsetminus L)$ to a specific conjugacy

class $C \subseteq \mathrm{Alt}(m)$. They show that the exact value is #P-complete when $m \geq 5$, but they do not obtain a parsimonious reduction. Instead, they retain an exponentially small error term. In particular, they do not obtain NP-hardness for the existence problem. However, in their favor, we found it easier to prove Theorem 1.1 in the case of closed 3-manifolds than in the case of link complements, which we will address in future work [33].

	one equation		equations		homomorphisms	
finite target G	$\exists$ [21]	# [41]	$\exists$ [21]	# [41]	$\exists$	#
abelian	P	FP	P	FP	P	FP
nilpotent	P	?	NPC	#PC	?	?
solvable	?	?	NPC	#PC	?	?
non-solvable	NPC	#PC	NPC	#PC	?	?
non-ab. simple	NPC	#PC	NPC	#PC	NPC!	#PC!

Figure 3. The complexity of solving equations over or finding homomorphisms to a fixed finite target group G. Here P denotes polynomial time for a decision problem, FP denotes polynomial time for a function problem, NPC is the class of NP-complete problems, and #PC is the class of #P-complete problems. Exclamation marks indicate results in this paper.

We can also place Theorem 1.1 in the context of other counting problems involving finite groups. We summarize what is known in Figure 3. Given a finite group G, the most general analogous counting problem is the number of solutions to a system of equations that may allow constant elements of G as well as variables. Nordh and Jonsson [41] showed that this problem is #P-complete if and only if G is non-abelian, while Goldman and Russell [21] showed that the existence problem is NP-complete. If G is abelian, then any finite system of equations can be solved by the Smith normal form algorithm. These authors also considered the complexity of a single equation. In this case, the existence problem has unknown complexity if G is solvable but not nilpotent, while the counting problem has unknown complexity if G is solvable but not abelian.

If all of the constants in a system of equations over G are set to $1 \in G$, then solving the equations amounts to finding group homomorphisms $f : \Gamma \to G$ from the finitely presented group Γ given by the equations. By slight abuse of notation, we can call this counting problem $\#H(\Gamma, G)$. This is equivalent to the topological invariant $\#H(X, G)$ when X is a simplicial complex, or even a triangulated n-manifold for any fixed $n \geq 4$; in this case, given any finitely presented Γ, we can construct X with $\Gamma = \pi_1(X)$ in polynomial time. To our knowledge, Theorem 1.1 is a new result for the invariant $\#H(\Gamma, G)$, even though we specifically construct Γ to be a 3-manifold group rather than a general finitely presented group. For comparison, both the non-triviality problem and the word problem are as difficult as the halting problem for general Γ [44]. By contrast, the word problem and the isomorphism problem are both recursive for 3-manifold groups, in fact elementary recursive [4, 29].

In the other direction, if M is a closed 2-manifold, then there are well known formulas of Frobenius-Schur and Mednykh

for $\#H(M, G)$ [17, 20, 37] for any finite group G as a function of the genus and orientability of M [17, 20, 37]. Mednykh's formula was generalized by Chen [10] to the case of Seifert-fibered 3-manifolds. In Section 2.3, we give a generalization of these formulas to the class of bounded-width simplicial complexes.

Our approach to Theorem 1.1 (like that of Krovi and Russell) is inspired by quantum computation and topological quantum field theory. Every unitary modular tensor category (UMTC) $\mathscr{C}$ yields a unitary 3-dimensional topological quantum field theory [45, 46, 50]. The topological quantum field theory assigns a vector space $V(\Sigma_g)$, or *state space*, to every oriented, closed surface. It also assigns a state space $V(\Sigma_{g,n}, C)$ to every oriented, closed surface with n boundary circles, where C is an object in $\mathscr{C}$ interpreted as the "color" of each boundary circle. Each state space $V(\Sigma_{g,n}, C)$ has a projective action of the mapping class group $\mathrm{MCG}_*(\Sigma_{g,n})$. (In fact the unpointed mapping class group $\mathrm{MCG}(\Sigma_{g,n})$ acts, but we will keep the basepoint for convenience.) These mapping class group actions then extend to invariants of 3-manifolds and links in 3-manifolds.

Finally, the UMTC $\mathscr{C}$ is universal for quantum computation if the image of the mapping class group action on suitable choices of $V(\Sigma_{g,n}, C)$ is large enough to simulate quantum circuits on m qubits, with $g, n = O(m)$. If the action is only large enough to simulate classical circuits on m bits, then it is still classically universal. These universality results are important for the fault-tolerance problem in quantum computation [18, 27].

One early, important UMTC is the (truncated) category $\mathrm{rep}_q(\mathrm{SU}(2))$ of quantum representations of $\mathrm{SU}(2)$ at a principal root of unity. This category yields the Jones polynomial for a link $L \subseteq S^3$ (taking $C = V_1$, the first irreducible object) and the Jones-Witten-Reshetikhin-Turaev invariant of a closed 3-manifold. In separate papers, Freedman, Larsen, and Wang showed that $V(\Sigma_{0,n}, V_1)$ and $V(\Sigma_{g,0})$ are both quantumly universal representations of $\mathrm{MCG}_*(\Sigma_{0,n})$ and $\mathrm{MCG}_*(\Sigma_{g,0})$ [18, 19].

Universality also implies that any approximation of these invariants that could be useful for computational topology is #P-hard. The first author [32] obtained such results for the Jones polynomial (see also Aharonov-Arad [1]), while Alagic and Lo [2] obtained the analogous result for the corresponding 3-manifold invariant. Note that exact evaluation of the Jones polynomial was earlier shown to be #P-hard without quantum computation methods [24].

If G is a finite group, then the invariant $\#H(M, G)$ for a 3-manifold M also comes from a UMTC, namely the categorical double $D(\mathrm{rep}(G))$ of $\mathrm{rep}(G)$, that was treated (and generalized) by Dijkgraaf and Witten and others [12, 17, 30]. In this case, the state space $V(\Sigma_{g,0})$ is the vector space $\mathbb{C}[\hat{R}_g(G)/\mathrm{Inn}(G)]$, and the action of $\mathrm{MCG}_*(\Sigma_{g,0})$ on $V(\Sigma_{g,0})$ is induced by its action on $\hat{R}_g(G)$. Some of the objects in $D(\mathrm{rep}(G))$ are given by conjugacy classes $C \subseteq G$, and the representation of the braid group $\mathrm{MCG}_*(\Sigma_{0,n})$ with braid strands colored by a conjugacy class C yields the invariant $\#H(S^3 \smallsetminus L, G, C)$ considered by Krovi and Russell. Motivated by the fault tolerance problem, Ogburn and Preskill [42] found

that the braid group action for $G = \mathrm{Alt}(5)$ is classically universal (with C the conjugacy class of 3-cycles) and they reported that Kitaev showed the same thing for $\mathrm{Sym}(5)$. They also showed if these actions are enhanced by quantum measurements in a natural sense, then they become quantumly universal. Later Mochon [39] extended this result to any non-solvable finite group G. In particular, he proved that the action of $\mathrm{MCG}_*(\Sigma_{0,n})$ is classically universal for a suitably chosen conjugacy class C.

Mochon's result is evidence, but not proof, that $\#H(S^3 \smallsetminus L, G, C)$ is #P-complete for every fixed, non-solvable G and every suitable conjugacy class $C \subseteq G$ that satisfies his theorem. His result implies that if we constrain the associated braid group action with arbitrary initialization and finalization conditions, then counting the number of solutions to the constraints is parsimoniously #P-complete. However, if we use a braid to describe a link, for instance with a plat presentation [32], then the description yields specific initialization and finalizations conditions that must be handled algorithmically to obtain hardness results. Recall that in our proof of Theorem 1.1, the state in $\hat{R}_g(G)$ is initialized and finalized using the handlebodies $(H_g)_I$ and $(H_g)_F$. If we could choose any initialization and finalization conditions whatsoever, then it would be much easier to establish (weakly parsimonious) #P-hardness; it would take little more work than to cite Theorem 1.4.

2. COMPLEXITY AND ALGORITHMS

2.1. Complexity classes

For background on the material in this section, and some of the treatment in the next section as well, see Arora and Barak [3] and the Complexity Zoo [53].

Let A be a finite *alphabet* (a finite set with at least 2 elements) whose elements are called *symbols*, and let A^* be the set of finite words in A. We can consider three kinds of computational problems with input in A^*: decision problems d, counting problems c, and function problems f, which have the respective forms

$$d : A^* \to \{\text{yes}, \text{no}\} \qquad c : A^* \to \mathbb{N} \qquad f : A^* \to A^*. \quad (3)$$

The output set of a decision problem can also be identified with the Boolean alphabet

$$A = \mathbb{Z}/2 = \{1, 0\} \cong \{\text{true}, \text{false}\} \cong \{\text{yes}, \text{no}\}.$$

A *complexity class* C is any set of function, counting, or decision problems, which may either be defined on all of A^* or require a promise. A specific, interesting complexity class is typically defined as the set of all problems that can be computed with particular computational resources. For instance, P is the complexity class of all decision problems d such that $d(x)$ can be computed in polynomial time (in the length $|x|$ of the input x) by a deterministic Turing machine. FP is the analogous class of function problems that are also computable in polynomial time.

A *promise problem* is a function d, c, or f of the same form as (3), except whose domain can be an arbitrary subset $S \subseteq A^*$. The interpretation is that an algorithm to compute a promise problem can accept any $x \in A^*$ as input, but its output is only taken to be meaningful when it is promised that $x \in S$.

The input to a computational problem is typically a data type such as an integer, a finite graph, a simplicial complex, etc. If such a data type can be encoded in A^* in some standard way, and if different standard encodings are interconvertible in FP, then the encoding can be left unspecified. For instance, the decision problem of whether a finite graph is connected is easily seen to be in P; the specific graph encoding is not important. Similarly, there are various standard encodings of the non-negative integers $\mathbb{N}$ in A^*. Using any such encoding, we can also interpret FP as the class of counting problems that can be computed in polynomial time.

The complexity class NP is the set of all decision problems d that can be answered in polynomial time with the aid of a prover who wants to convince the algorithm (or verifier) that the answer is "yes". In other words, every $d \in \mathsf{NP}$ is given by a two-variable predicate $v \in \mathsf{P}$. Given an input x, the prover provides a witness y whose length $|y|$ is some polynomial in $|x|$. Then the verifier computes $v(x, y)$, with the conclusion that $d(x) = \text{yes}$ if and only if there exists y such that $v(x, y) = \text{yes}$. The witness y is also called a *proof* or *certificate*, and the verification v is also called a *predicate*. Likewise, a function $c(x)$ is in #P when it is given by a predicate $v(x, y)$; in this case $c(x)$ is the number of witnesses y that satisfy $v(x, y)$. For instance, whether a finite graph G (encoded as x) has a 3-coloring is in NP, while the number of 3-colorings of G is in #P. In both cases, a 3-coloring of G serves as a witness y.

A computational problem f may be NP-hard or #P-hard with the intuitive meaning that it is provably at least as difficult as any problem in NP or #P. A more rigorous treatment leads to several different standards of hardness. One quite strict standard is that any problem g in NP or #P can be reduced to the problem f by converting the input; *i.e.*, there exists $h \in \mathsf{FP}$ such that

$$g(x) = f(h(x)).$$

If $f, g \in \mathsf{NP}$, then this is called *Karp reduction*; if $f, g \in \#\mathsf{P}$, then it is called *parsimonious reduction*. Evidently, if a counting problem c is parsimoniously #P-hard, then the corresponding existence problem d is Karp NP-hard.

When a problem f is #P-hard by some more relaxed standard than parsimonious reduction, there could still be an algorithm to obtain some partial information about the value f, such as a congruence or an approximation, even if the exact value is intractable. For instance, the permanent of an integer matrix is well-known to be #P-hard [52], but its parity is the same as that of the determinant, which can be computed in polynomial time. However, when a counting problem c is parsimoniously #P-hard, then the standard conjecture that $\mathsf{NP} \not\subseteq \mathsf{BPP}$ implies that it is intractable to obtain any partial information about c. Here BPP is the set of problems solvable in randomized polynomial time with a probably correct answer.

Theorem 2.1 (Corollary of Valiant-Vazirani [51]). *Let c be a parsimoniously #P-hard problem, and let $b > a \geq 0$ be distinct, positive integers. Then distinguishing $c(x) = a$ from $c(x) = b$ is NP-hard via a Cook reduction in* BPP, *given the promise that $c(x) \in \{a, b\}$.*

When we say that an algorithm $\mathscr{A}$ obtains partial information about the value of $c(x)$, we mean that it can calculate $f(c(x))$ for some non-constant function f. Thus it can distinguish some pair of cases $c(x) = a$ and $c(x) = b$; and by Theorem 2.1, this is NP-hard. Here a *Cook reduction* is a polynomial-time algorithm $\mathscr{B}$ (in this case randomized polynomial time) that can call $\mathscr{A}$ as a subroutine.

Proof. Given a problem $d \in$ NP, Valiant and Vazirani construct a randomized algorithm $\mathscr{B}$ that calculates $d(x)$ using a collection of predicates $v_1(x, y)$ in P that usually have at most one solution in y. Thus, if an algorithm $\mathscr{A}$ can solve each problem

$$d_1(x) = \exists ? y \text{ such that } v_1(x, y) = \text{yes}$$

under the promise that at most one y exists, then $\mathscr{A}$ can be used as a subroutine to compute the original d. Such a predicate $v_1(x, y)$ may occasionally have more than one solution, but this happens rarely and still allows $\mathscr{B}$ to calculate d by the standard that its output only needs to be probably correct.

Given such a predicate $v_1(x, y)$, it is easy to construct another predicate $v_2(x, y)$ in P that has $b - a$ solutions in y for each solution to $v_1(x, y)$, and that has a other solutions in y regardless. Thus $v_2(x, y)$ has b solutions when $d_1(x) = \text{yes}$ and a solutions when $d_1(x) = \text{no}$. Thus, an algorithm $\mathscr{A}$ that can distinguish $c(x) = a$ from $c(x) = b$ can be used to calculate $d_1(x)$, and by the Valiant-Vazirani construction can be used to calculate $d(x)$. $\square$

A decision problem d which is both in NP and NP-hard is called NP-*complete*, while a counting problem which is both in #P and #P-hard is called #P-*complete*. For instance the decision problem CSAT, circuit satisfiability over an alphabet A, is Karp NP-complete, while the counting version #CSAT is parsimoniously #P-complete (Theorem 2.2). Thus, we can prove that any other problem is NP-hard by reducing CSAT to it, or #P-hard by reducing #CSAT to it.

We mention three variations of parsimonious reduction. A counting function c is *weakly parsimoniously* #P-hard if for every $b \in$ #P, there are $f, g \in$ FP such that

$$b(x) = f(c(g(x)), x).$$

The function c is *almost parsimoniously* #P-hard if f does not depend on x, only on $c(g(x))$. In either case, we can also ask for $f(c, x)$ to be 1-to-1 on the set of values of c with $f^{-1} \in$ FP, linear or affine linear in c, etc. So, Theorem 1.1 says that $\#H(M, G)$ is almost parsimoniously #P-complete.

Finally, suppose that $c(x)$ counts the number of solutions to $v(x, y)$ and $b(x)$ counts the number of solutions to $u(x, y)$. Then a *Levin reduction* is a map $h \in$ FP and a bijection f with $f, f^{-1} \in$ FP such that

$$u(x, y) = v(h(x), f(y)).$$

Obviously Levin reduction implies parsimonious reduction.

2.2. Circuits

Given an alphabet A, a *gate* is a function $\alpha : A^k \to A^\ell$. A *gate set* Γ is a finite set of gates, possibly with varying sizes of domain and target, and a *circuit* over Γ is a composition of gates in Γ in the pattern of a directed, acyclic graph. A gate set Γ is *universal* if every function $f : A^n \to A^m$ has a circuit. For example, if $A = \mathbb{Z}/2$, then the gate set

$$\Gamma = \{\text{AND}, \text{OR}, \text{NOT}, \text{COPY}\}$$

is universal, where AND, OR, and NOT are the standard Boolean operations and the COPY gate is the diagonal embedding $a \mapsto (a, a)$.

Let A be an alphabet with a universal gate set Γ, and suppose that A has a distinguished symbol yes $\in A$. Choose a standard algorithm to convert an input string $x \in A^*$ to a circuit C_x with one output. Then the *circuit satisfiability problem* $\text{CSAT}_{A,\Gamma}(x)$ asks whether the circuit C_x has an input y such that $C_x(y) = \text{yes}$. It is not hard to construct a Levin reduction of $\text{CSAT}_{A,\Gamma}$ from any one alphabet and gate set to any other, so we can just call any such problem CSAT. CSAT also has an obvious counting version #CSAT.

Theorem 2.2 (Cook-Levin-Karp). CSAT *is Karp* NP-*complete and* #CSAT *is parsimoniously* #P-*complete.*

(See Arora-Barak [3, Sec. 6.1.2 & Thm. 17.10] for a proof of Theorem 2.2.)

We will need two variations of the circuit model that still satisfy Theorem 2.2: Reversible circuits and planar circuits.

A *reversible circuit* [16] is a circuit C in which every gate $\alpha : A^k \to A^k$ in the gate set Γ is a bijection; thus the evaluation of C is also a bijection. We say that Γ is *reversibly universal* if for any sufficiently large n, the gates of Γ in different positions generate either $\text{Alt}(A^n)$ or $\text{Sym}(A^n)$. (If $|A|$ is even, then we cannot generate any odd permutations when n is larger than the size of any one gate in Γ.)
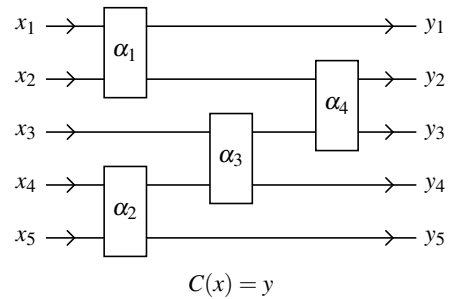

$$C(x) = y$$

Figure 4. A planar, reversible circuit.

A circuit C is *planar* if its graph is a planar graph placed in a rectangle in the plane, with the inputs on one edge and the output on an opposite edge. The definition of a universal gate for general circuits carries over to planar circuits; likewise the definition for reversible circuits carries over to reversible planar circuits. (See Figure 4.) We can make a circuit

or a reversible circuit planar using reversible SWAP gates that take (a,b) to (b,a). Likewise, any universal gate set becomes planar-universal by adding the SWAP gate. Thus, the planar circuit model is equivalent to the general circuit model.

The reduction from general circuits to reversible circuits is more complicated.

Lemma 2.3. *Let A be an alphabet for reversible circuits.*

1. *If $|A| \geq 3$, then $\Gamma = \text{Alt}(A^2)$ is a universal set of binary gates.*

2. *If $|A| = 2$, then $\Gamma = \text{Alt}(A^3)$ is a universal set of ternary gates.*

3. *If $|A|$ is even, then $\text{Sym}(A^n) \subseteq \text{Alt}(A^{n+1})$.*

Different versions of Lemma 2.3 are standard in the reversible circuit literature. For instance, when $A = \mathbb{Z}/2$, the foundational paper [16] defines the Fredkin gate and the Toffoli gate, each of which is universal together with the NOT gate. Nonetheless, we did not find a proof for all values of $|A|$, so we give one.

Proof. Case 3 of the lemma is elementary, so we concentrate on cases 1 and 2. We will show by induction on n that Γ generates $\text{Alt}(A^n)$. The hypothesis hands us the base of induction $n = 3$ when $|A| = 2$ and $n = 2$ when $|A| \geq 3$. So, we assume a larger value of n and we assume by induction that the case $n - 1$ is already proven.

We consider the two subgroups in $\text{Alt}(A^n)$ that are given by Γ-circuits that act respectively on the left $n-1$ symbols or the right $n-1$ symbols. By induction, both subgroups are isomorphic to $\text{Alt}(A^{n-1})$, and we call them $\text{Alt}(A^{n-1})_L$ $\text{Alt}(A^{n-1})_R$. They in turn have subgroups $\text{Alt}(A^{n-2})_L^{|A|}$ and $\text{Alt}(A^{n-2})_R^{|A|}$ which are each isomorphic to $\text{Alt}(A^{n-2})^{|A|}$ and each act on the middle $n-2$ symbols; but in one case the choice of permutation $\alpha \in \text{Alt}(A^{n-2})$ depends on the leftmost symbol, while in the other case it depends on the rightmost symbol. By taking commutators between these two subgroups, we obtain all permutations in $\text{Alt}(\{a\} \times A^{n-2} \times \{b\})$ for every pair of symbols (a,b). Moreover, we can repeat this construction for every subset of $n-2$ symbols. Since $n \geq 3$, and since $n \geq 4$ when $|A| = 2$, we know that $|A^{n-2}| \geq 3$. We can thus apply Lemma 3.1 in the next section to the alternating subgroups that we have obtained. $\square$

Lemma 2.3 motivates the definition of a canonical reversible gate set Γ for each alphabet A. (Canonical in the sense that it is both universal and constructed intrinsically from the finite set A.) If $|A|$ is odd, then we let $\Gamma = \text{Alt}(A^2)$. If $|A| \geq 4$ is even, then we let $\Gamma = \text{Sym}(A^2)$. Finally, if $|A| = 2$, then we let $\Gamma = \text{Sym}(A^3)$. By Lemma 2.3, each of these gate sets is universal. Moreover, each of these gate sets can be generated by any universal gate set, possibly with the aid of an ancilla in the even case.

In one version of reversible circuit satisfiability, we choose two subsets $I, F \subseteq A$, interpreted as initialization and finalization alphabets. We define the problem $\text{RSAT}_{A,I,F}$ as follows: The input x represents a reversible circuit C_x of some width

n over the alphabet A, with gates taken from some universal gate set Γ. Then C_x is said to be satisfied if there is a circuit input $y \in I^n \subseteq A^n$ such that $C_x(y) \in F^n \subseteq A^n$. The satisfiability problem $\text{RSAT}_{A,I,F}(x)$ asks whether such a witness y exists, while as usual the counting problem $\#\text{RSAT}_{A,I,F}(x)$ asks for the number of witnesses y. Note that if either $F = A$ or $|I| = 1$, then $\text{RSAT}_{A,I,F}$ is trivial. Since C_x is a reversible circuit, it is just as easy to construct its inverse C_x^{-1}, so likewise $\text{RSAT}_{A,I,F}$ is also trivial if either $I = A$ or $|F| = 1$.

Theorem 2.4. *Consider A, I, F, and Γ with Γ universal and $2 \leq |I|, |F| < |A|$. Then $\text{RSAT}_{A,I,F}$ is Karp NP-hard and $\#\text{RSAT}_{A,I,F}$ is parsimoniously #P-hard.*

Theorem 2.4 is also a standard result in reversible circuit theory, but we again give a proof because we did not find one.

Proof. We consider a sequence RSAT_i of versions of the reversible circuit problem. We describe the satisfiability version for each one, and implicitly define the counting version $\#\text{RSAT}_i$ using the same predicate.

- RSAT_1 uses the binary alphabet $A = \mathbb{Z}/2$ and does not have I or F. Instead, some of the input bits are set to 0 while others are variable, and the decision output of a circuit is simply the value of the first bit.

- RSAT_2 also has $A = \mathbb{Z}/2$ with an even number of input and output bits. Half of the input bits and output bits are set to 0, while the others are variable. A circuit C is satisfied by finding an input/output pair x and $C(x)$ that satisfy the constraints.

- RSAT_3 is $\text{RSAT}_{A,I,F}$ with I and F disjoint and $|A \smallsetminus (I \cup F)| \geq 2$.

- RSAT_4 is $\text{RSAT}_{A,I,F}$ with the stated hypotheses of the theorem.

We claim parsimonious reductions from $\#\text{CSAT}$ to $\#\text{RSAT}_1$, and from $\#\text{RSAT}_i$ to $\#\text{RSAT}_{i+1}$ for each i.

Step 1: We can reduce CSAT to RSAT_1 through the method of *gate dilation* and ancillas. Here an *ancilla* is any fixed input to the circuit that is used for scratch space; the definition of RSAT_1 includes ancillas. To define gate dilation, we can let A be any alphabet with the structure of an abelian group. If $\alpha : A^k \to A$ is a gate, then we can replace it with the reversible gate

$$\beta : A^{k+1} \to A^{k+1} \qquad \beta(x,a) = (x, \alpha(x) + a),$$

where $x \in A^k$ is the input to α and $a \in A$ is an ancilla which is set to $a = 0$ when β replaces α. The gate β is called a *reversible dilation* of α. We can similarly replace every irreversible COPY gate with the reversible gate

$$\text{COPY} : A^2 \to A^2 \qquad \text{COPY}(x,a) = (x, x + a),$$

where again a is an ancilla set to $a = 0$. Dilations also leave extra output symbols, but under the satisfiability rule of RSAT_1, we can ignore them.

In the Boolean case $A = \mathbb{Z}/2$, the reversible COPY gate is denoted CNOT (controlled NOT), while the dilation of AND is denoted CCNOT (doubly controlled NOT) and is called the Toffoli gate. We can add to this the uncontrolled NOT gate

$$\text{NOT}(x) = x + 1.$$

These three gates are clearly enough to dilate irreversible Boolean circuits. (They are also a universal gate set for reversible computation.)
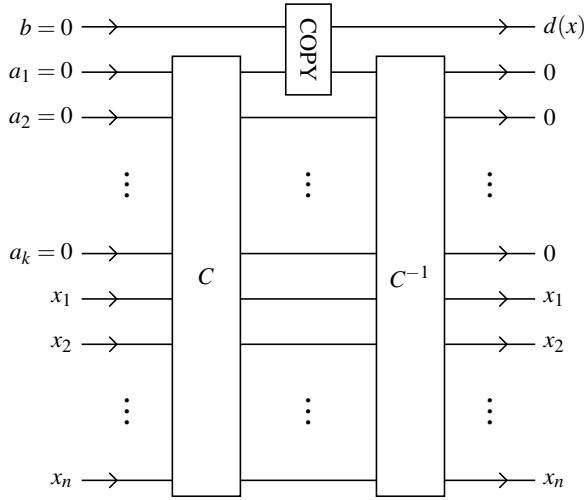

Figure 5. Using uncomputation to reset ancilla values.

Step 2: We can reduce RSAT_1 to RSAT_2 using the method of uncomputation. Suppose that a circuit C in the RSAT_1 problem has an n-bit variable input register $x = (x_1, x_2, \ldots, x_n)$ and a k-bit ancilla register $a = (a_1, a_2, \ldots, a_k)$. Suppose that C calculates decision output $d(x)$ in the a_1 position (when $a_1 = 0$ since it is an ancilla). Then we can make a new circuit C_1 with the same x and a and one additional ancilla bit b, defined by applying C, then copying the output to b and negating b, then applying C^{-1}, as in Figure 5. If $n = k + 1$, then C_1 is a reduction of C from RSAT_1 to RSAT_2. If $n > k + 1$, then we can pad C_1 with $n - k - 1$ more ancillas and do nothing with them to produce a padded circuit C_2. If $n < k + 1$, then we can pad C_1 with $k + 1 - n$ junk input bits, and at the end of C_1, copy of these junk inputs to $k + 1 - n$ of the first k ancillas; again to produce C_2. (Note that $k + 1 - n \leq k$ since we can assume that C has at least one variable input bit.) In either of these cases, C_2 is a reduction of C from RSAT_1 to RSAT_2.

Step 3: We can reduce RSAT_2 to RSAT_3 by grouping symbols and permuting alphabets. As a first step, let $A_1 = \mathbb{Z}/2 \times \mathbb{Z}/2$ with $I_1 = F_1 = \{(0,0),(1,0)\}$. Then we can reduce RSAT_2 to $\text{RSAT}_{A_1,I_1,F_1}$ by pairing each of input or output bit with an ancilla; we can express each ternary gate over $\mathbb{Z}/2$ in terms of binary gates over A_1. Now let A_2 be any alphabet with disjoint I_2 and F_2, and with at least two symbols not in I_2 or F_2. Then we can embed (A_1, I_1, F_1) into (A_2, I_2, F_2) arbitrarily, and extend any gate $\alpha : A_1^k \to A_1^k$ (with $k \in \{1,2\}$, say) arbitrarily to a gate $\beta : A_2^k \to A_2^k$ which is specifically

an even permutation. This reduces $\text{RSAT}_2 = \text{RSAT}_{A_1,I_1,F_1}$ to $\text{RSAT}_3 = \text{RSAT}_{A_2,I_2,F_2}$.

Step 4: Finally, (A_3, I_3, F_3) is an alphabet that is not of our choosing, and we wish to reduce $\text{RSAT}_3 = \text{RSAT}_{A_2,I_2,F_2}$ to $\text{RSAT}_4 = \text{RSAT}_{A_3,I_3,F_3}$. We choose k such that

$$|A_3|^k \geq |I_3|^k + |F_3|^k + 2.$$

We then let $A_2 = A_3^k$ and $I_2 = I_3^k$, and we choose $F_2 \subseteq A_2 \smallsetminus I_2$ with $|F_2| = |F_3|^k$. A circuit in $\text{RSAT}_{A_2,I_2,F_2}$ can now be reduced to a circuit in $\text{RSAT}_{A_3,I_3,F_3}$ by grouping together k symbols in A_3 to make a symbol in A_2. Since $I_2 = I_3^k$, the initialization is the same. At the end of the circuit, we convert finalization in F_2 to finalization in F_3^k with some unary permutation of the symbols in A_2. $\square$

2.3.   Standard algorithms

In this section we will review a few standard algorithms that supplement Theorem 1.1. Instead of hardness results, they are all easiness results. (Note that Theorem 2.7 produces a conditional type of easiness, namely predicates that can be evaluated in polynomial time.)

Theorem 2.5. *The integer homology $H_*(X)$ of a finite simplicial complex X can be computed in polynomial time.*

Briefly, Theorem 2.5 reduces to computing the Smith normal form of an integer matrix and a corresponding matrix factorization [13]. Kannan and Bachem [26] showed that a Smith factorization can be computed in polynomial using a refinement of the standard Smith normal form algorithm based on row and column operations.

Proposition 2.6. *If X is a finite simplicial complex given as computational input, then it can be confirmed in polynomial time whether X is a closed 3-manifold M, and whether M is a homology 3-sphere.*

Proof. To be concrete, X is described by a set of vertices and a set of subsets of those vertices representing simplices. We can then trivially check the first two properties:

1. That the every maximal simplex is 3-dimensional.

2. That the link of every edge is a polygon.

It follows that the link $\text{lk}(v)$ of every vertex v is a surface; to check that that M is a closed 3-manifold, we want to know that every $\text{lk}(v)$ is a 2-sphere. We can confirm this for instance by computing $H_*(\text{lk}(v))$ using Theorem 2.5. Then to confirm that M is a homology 3-sphere (including that it is orientable), we can again use Theorem 2.5 to calculate $H_*(M)$. $\square$

Theorem 2.7. *If G is a fixed finite group and X is a finite, connected simplicial complex regarded as the computational input, then $\#H(X,G)$ and $\#Q(X,G)$ are both in $\#\text{P}$.*

Proof. By choosing a spanning tree for the 1-skeleton of X, we can convert its 2-skeleton to a finite presentation P of $\pi_1(X)$. Then we can describe a homomorphism $f : \pi_1(X) \to G$ by the list of its values on the generators in P. This serves as a certificate; the verifier should then check whether the values satisfy the relations in P. This shows that $\#H(X, G)$ is in #P.

The case of $\#Q(X, G)$ is similar but slightly more complicated. The map f is surjective if and only if its values on the generators in P generate G; the verifier can check this. The verifier can also calculate the $\mathrm{Aut}(G)$-orbit of f. Given an ordering of the generators and an ordering of the elements of G, the verifier can accept f only when it is alphabetically first in its orbit. Since only surjections are counted and each orbit is only counted once, we obtain that $\#Q(X, G)$ certificates are accepted. □

In the input to the third algorithm, we decorate a finite simplicial complex X with a complete ordering of its simplices (of all dimensions) that refines the partial ordering of simplices given by inclusion. If there are n simplices total, then for each $0 \leq k \leq n$, we let X_k be subcomplex formed by the first k simplices, so that $X_0 = \emptyset$ and $X_n = X$. Each X_k has a relative boundary $\mathrm{bd}(X_k)$ in X. (Here we mean boundary in the set of general topology rather than manifold theory, *i.e.*, closure minus interior.) We define the *width* of X with its ordering to be the maximum number of simplices in any $\mathrm{bd}(X_k)$.

Theorem 2.8. *If G is a fixed finite group and X is a finite, connected simplicial complex with a bounded-width ordering, then $\#H(X, G)$ and $\#Q(X, G)$ can be computed in polynomial time (non-uniformly in the wdith).*

It is easy to make triangulations for all closed surfaces with uniformly bounded width. For instance, we can make such a triangulation of an orientable surface Σ_g from a Morse function chosen so that each regular level is either one or two circles. With more effort, we can make a bounded-width triangulation of a Seifert-fibered 3-manifold M using a bounded-width triangulation of its orbifold base. Thus Theorem 2.8 generalizes the formulas of Mednykh [37] and Chen [10] in principle, although in practice their formulas are more explicit and use better decompositions than triangulations. Theorem 2.8 also applies to 3-manifolds with bounded Heegaard genus, or more generally bounded Morse width.

Proof. We can calculate $|H(X, G)|$ using the formalism of non-abelian simplicial cohomology theory with coefficients in G [43]. In this theory, we orient the edges of X and we mark a vertex $x_0 \in X$ as a basepoint. A 1-cocycle is then a function from the edges of X to G that satisfies a natural coboundary condition on each triangle, while a 0-cochain is a function from the vertices to G that takes the value 1 at x_0. The 1-cocycle set $Z^1(X; G)$ has no natural group structure when G is non-commutative, while the relative 0-cochain set $C^0(X, x_0; G)$ is a group that acts freely on $Z^1(X; G)$. Then the set of orbits

$$H^1(X, x_0; G) \stackrel{\text{def}}{=} Z^1(X; G)/C^0(X, x_0; G).$$

can be identified with the representation set $H(X, G)$, while if X has v vertices, then $C^0(X, x_0; G) \cong G^{v_1}$. Thus

$$|H(X, G)| = |Z^1(X; G)|/|G|^{v-1}.$$

Our approach is to compute $|Z^1(X; G)|$ and divide. We can then also obtain $|Q(X, G)|$ from $|H(X, G)|$ by applying Möbius inversion to equation (1).

The algorithm is an example of *dynamical programming* in computer science. Working by induction for each k from 0 to n, it maintains a vector v_k of non-negative integers that consists of the number of ways to extend each 1-cocycle on $\mathrm{bd}(X_k)$ to a 1-cocycle of X_K. The dimension of v_k may be exponential in the number of edges of $\mathrm{bd}(X_k)$, but since that is bounded, the dimension of v_k is also bounded. It is straightforward to compute v_{k+1} from v_k when we pass from X_k to X_{k+1}. If $X_{k+1} \smallsetminus X_k$ is an edge, then v_{k+1} consists of $|G|$ copies of v_k. If $X_{k+1} \smallsetminus X_k$ is a triangle and $\mathrm{bd}(X_{k+1})$ has the same edges as $\mathrm{bd}(X_k)$, then v_{k+1} is a subvector of v_k. If $\mathrm{bd}(X_{k+1})$ has fewer edges than $\mathrm{bd}(X_k)$, then v_{k+1} is obtained from v_k by taking local sums of entries. □

3. GROUP THEORY

In this section we collect some group theory results. We do not consider any of these results to be especially new, although we found it challenging to prove Theorem 3.10.

3.1. Generating alternating groups

Lemma 3.1 (*Cf.* [11, Lem. 7]). *Let S be a finite set and let $T_1, T_2, \ldots, T_n \subseteq S$ be a collection of subsets with at least 3 elements each, whose union is S, and that form a connected graph under pairwise intersection. Then the permutation groups $\mathrm{Alt}(T_i)$ together generate $\mathrm{Alt}(S)$.*

Proof. We argue by induction on $|S \smallsetminus T_1|$. If $T_1 = S$, then there is nothing to prove. Otherwise, we can assume (possibly after renumbering the sets) that there is an element $a \in T_1 \cap T_2$ and an element $b \in T_2 \smallsetminus T_1$. Let $\alpha \in \mathrm{Alt}(T_2)$ be a 3-cycle such that $\alpha(a) = b$. Then the 3-cycles in $\mathrm{Alt}(T_1)$, and their conjugates by α, and α itself if it lies in $\mathrm{Alt}(T_1 \cup \{b\})$, include all 3-cycles in $\mathrm{Alt}(T_1 \cup \{b\})$. Thus we generate $\mathrm{Alt}(T_1 \cup \{b\})$ and we can replace T_1 by $T_1 \cup \{b\}$. □

3.2. Joint surjectivity

Recall the existence half of the Chinese remainder theorem: If $d_1, d_2, \ldots, d_n$ are pairwise relatively prime integers, then the canonical homomorphism

$$f : \mathbb{Z} \to \mathbb{Z}/d_1 \times \mathbb{Z}/d_2 \times \cdots \times \mathbb{Z}/d_n$$

is (jointly) surjective. The main hypothesis is "local" in the sense that it is a condition on each pair of divisors d_i and d_j, namely $\gcd(d_i, d_j) = 1$. For various purposes, we will need

non-commutative joint surjectivity results that resemble the classic Chinese remainder theorem. (But we will not strictly generalize the Chinese remainder theorem, although such generalizations exist.) Each version assumes a group homomorphism

$$f : K \to G_1 \times G_2 \times \cdots \times G_n$$

that surjects onto each factor G_i, and assumes certain other local hypotheses, and concludes that f is jointly surjective. Dunfield-Thurston [14, Lem. 3.7] and the first author [31, Lem. 3.5] both have results of this type and call them "Hall's lemma", but Hall [23, Sec. 1.6] only stated without proof a special case of Dunfield and Thurston's lemma. Ribet [47, Lem. 3.3] also has such a lemma with the proof there attributed to Serre. In this paper, we will start with a generalization of Ribet's lemma.

We define a group homomorphism

$$f : K \to G_1 \times G_2 \times \cdots \times G_n$$

to be *k-locally surjective* for some integer $1 \le k \le n$ if it surjects onto every direct product of k factors. Recall also that if G is a group, then $G' = [G, G]$ is a notation for its commutator subgroup.

Lemma 3.2 (After Ribet-Serre [47, Lem. 3.3]). *Let*

$$f : K \to G_1 \times G_2 \times \cdots \times G_n$$

be a 2-locally surjective group homomorphism, such that also its abelianization

$$f_{\mathrm{ab}} : K \to (G_1)_{\mathrm{ab}} \times (G_2)_{\mathrm{ab}} \times \cdots \times (G_n)_{\mathrm{ab}}$$

is $\lceil (n+1)/2 \rceil$-locally surjective. Then

$$f(K) \ge G_1' \times G_2' \times \cdots \times G_n'.$$

Proof. We argue by induction on n. If $n = 2$, then there is nothing to do. Otherwise let $t = \lceil (n+1)/2 \rceil$ and note that $n > t > n/2$. Let

$$\pi : G_1 \times G_2 \times \cdots \times G_n \to G_1 \times G_2 \times \cdots \times G_t$$

be the projection onto the first t factors. Then $\pi \circ f$ satisfies the hypotheses, so

$$\pi(f(K)) \ge G_1' \times G_2' \times \cdots \times G_t'.$$

Morever, $(\pi \circ f)_{\mathrm{ab}}$ is still t-locally surjective, which is to say that

$$\pi(f(K))_{\mathrm{ab}} = (G_1)_{\mathrm{ab}} \times (G_2)_{\mathrm{ab}} \times \cdots \times (G_t)_{\mathrm{ab}}.$$

Putting these two facts together, we obtain

$$\pi(f(K)) = G_1 \times G_2 \times \cdots \times G_t.$$

Repeating this for any t factors, we conclude that f is t-locally surjective.

Given any two elements $g_t, h_t \in G_t$, we can use t-local surjectivity to find two elements

$$(g_1, g_2, \ldots, g_{t-1}, g_t, 1, 1, \ldots, 1),$$
$$(1, 1, \ldots, 1, h_t, h_{t+1}, \ldots, h_n) \in f(K).$$

Their commutator then is $[g_t, h_t] \in G_t \cap f(K)$. Since g_t and h_t are arbitrary, we thus learn that $G_t' \le f(K)$, and since this construction can be repeated for any factor, we learn that

$$f(K) \ge G_1' \times G_2' \times \cdots \times G_n',$$

as desired. $\qquad\square$

We will also use a complementary result, Goursat's lemma, which can be used to establish 2-local surjectivity. (Indeed, it is traditional in some papers to describe joint surjectivity results as applications of Goursat's lemma.)

Lemma 3.3 (Goursat [8, 22]). *Let G_1 and G_2 be groups and let $H \le G_1 \times G_2$ be a subgroup that surjects onto each factor G_i. Then there exist normal subgroups $N_i \trianglelefteq G_i$ such that $N_1 \times N_2 \le H$ and $H/(N_1 \times N_2)$ is the graph of an isomorphism $G_1/N_1 \cong G_2/N_2$.*

For instance, if G_1 is a simple group, then either $H = G_1 \times G_2$ or H is the graph of an isomorphism $G_1 \cong G_2$. In other words, given a joint homomorphism

$$f = f_1 \times f_2 : K \to G_1 \times G_2$$

which surjects onto each factor, either f is surjective or f_1 and f_2 are equivalent by an isomorphism $G_1 \cong G_2$. We can combine this with the perfect special case of Lemma 3.2 to obtain exactly Dunfield and Thurston's version.

Lemma 3.4 ([14, Lem. 3.7]). *If*

$$f : K \to G_1 \times G_2 \times \cdots \times G_n$$

is a group homomorphism to a direct product of non-abelian simple groups, and if no two factor homomorphism $f_i : K \to G_i$ and $f_j : K \to G_j$ are equivalent by an isomorphism $G_i \cong G_j$, then f is surjective.

Corollary 3.5. *Let K be a group and let*

$$N_1, N_2, \ldots, N_n \triangleleft K$$

be distinct maximal normal subgroups with non-abelian simple quotients $G_i = K/N_i$. Then

$$G_1 \cong (N_2 \cap N_3 \cap \cdots \cap N_n)/(N_1 \cap N_2 \cap \cdots \cap N_n).$$

Proof. We can take the product of the quotient maps to obtain a homomorphism

$$f : K \to G_1 \times G_2 \times \cdots \times G_n$$

that satisfies Lemma 3.4. Thus we can restrict f to

$$f^{-1}(G_1) = N_2 \cap N_3 \cap \cdots \cap N_n$$

to obtain a surjection

$$f : N_2 \cap N_3 \cap \cdots \cap N_n \twoheadrightarrow G_1.$$

This surjection yields the desired isomorphism. $\qquad\square$

We will use a more direct corollary of Lemma 3.3. We say that a group G is *normally Zornian* if every normal subgroup of G is contained in a maximal normal subgroup. Clearly every finite group is normally Zornian, and so is every simple group. A more interesting result implied by Neumann [40, Thm. 5] is that every finitely generated group is normally Zornian. (Neumann's stated result is that every subgroup is contained in a maximal subgroup, but the proof works just as well for normal subgroups. He also avoided the axiom of choice for this result, despite our reference to Zorn's lemma.) Recall also the standard concept that a group H is *involved* in another group G if H is a quotient of a subgroup of G.

Lemma 3.6. *Suppose that*

$$f : K \to G_1 \times G_2$$

is a group homomorphism that surjects onto the first factor G_1, and that G_1 is normally Zornian. Then:

  1. *If no simple quotient of G_1 is involved in G_2, then $f(K)$ contains G_1.*

  2. *If f surjects onto G_2, and no simple quotient of G_1 is a quotient of G_2, then f is surjective.*

Proof. Case 1 reduces to case 2, since we can replace G_2 by the projection of $f(K)$ in G_2. In case 2, Lemma 3.3 yields isomorphic quotients $G_1/N_1 \cong G_2/N_2$. Since G_1 is normally Zornian, we may further quotient G_1/N_1 to produce a simple quotient Q, and we can quotient G_2/N_2 correspondingly. $\square$

Finally, we have a lemma to calculate the simple quotients of a direct product of groups.

Lemma 3.7. *If*

$$f : G_1 \times G_2 \times \cdots \times G_n \twoheadrightarrow Q$$

is a group homomorphism from a direct product to a non-abelian simple quotient, then it factors through a quotient map $f_i : G_i \to Q$ for a single value of i.

Proof. The lemma clearly reduces to the case $n = 2$ by induction. If

$$f : G_1 \times G_2 \twoheadrightarrow Q$$

is a simple quotient, then $f(G_1)$ and $f(G_2)$ commute with each other, so they are normal subgroups of the group that they generate, which by hypothesis is Q. So each of $f(G_1)$ and $f(G_2)$ is either trivial or equals Q. Since Q is non-commutative, then $f(G_1)$ and $f(G_2)$ cannot both be Q, again because they commute with each other. Thus one of G_1 and G_2 is in the kernel of f, and f factors through a quotient of the other one. $\square$

3.3. Integer symplectic groups

Recall that for any integer $g \geq 1$ and any commutative ring A, there is an integer symplectic group $\mathrm{Sp}(2g, A)$, by definition the set of automorphisms of the free A-module A^{2g} that preserves a symplectic inner product. Likewise the projective symplectic group $\mathrm{PSp}(2g, A)$ is the quotient of $\mathrm{Sp}(2g, A)$ by its center (which is trivial in characteristic 2 and consists of $\pm I$ otherwise). For each prime p and each $g \geq 1$, the group $\mathrm{PSp}(2g, \mathbb{Z}/p)$ is a finite simple group, except for $\mathrm{PSp}(2, \mathbb{Z}/2)$, $\mathrm{PSp}(2, \mathbb{Z}/3)$, and $\mathrm{PSp}(4, \mathbb{Z}/2)$ [9, Thm. 11.1.2]. Moreover, $\mathrm{PSp}(2g, \mathbb{Z}/p)$ is never isomorphic to an alternating group when $g \geq 2$ (because it has the wrong cardinality).

We want to apply Lemma 3.6 to the symplectic group $\mathrm{Sp}(2g, \mathbb{Z})$, since it is the quotient of the mapping class group $\mathrm{MCG}_*(\Sigma_g)$ by the Torelli group $\mathrm{Tor}_*(\Sigma_g)$. To this end, we can describe its simple quotients when $g \geq 3$.

Lemma 3.8. *If $g \geq 3$, then the simple quotients of $\mathrm{Sp}(2g, \mathbb{Z})$ are all of the form $\mathrm{PSp}(2g, \mathbb{Z}/p)$, where p is prime and the quotient map is induced by the ring homomorphism from $\mathbb{Z}$ to $\mathbb{Z}/p$.*

As the proof will indicate, Lemma 3.8 is a mutual corollary of two important results due to others: the congruence subgroup property of Mennicke and Bass-Lazard-Serre, and the Margulis normal subgroup theorem.

Note that the finite simple quotients of $\mathrm{Sp}(4, \mathbb{Z})$ are only slightly different. The best way to repair the result in this case is to replace both $\mathrm{Sp}(4, \mathbb{Z})$ and $\mathrm{Sp}(4, \mathbb{Z}/2)$ by their commutator subgroups of index 2. Meanwhile given the well-known fact that $\mathrm{PSp}(2, \mathbb{Z}) \cong C_2 * C_3$, any simple group generated by an involution and an element of order 3 is a simple quotient of $\mathrm{Sp}(2, \mathbb{Z})$, and this is a very weak restriction. However, we only need Lemma 3.8 for large g.

Proof. We note first that $\mathrm{Sp}(2g, \mathbb{Z})$ is a perfect group when $g \geq 3$, so every possible simple quotient is non-abelian, and every such quotient is also a quotient of $\mathrm{PSp}(2g, \mathbb{Z})$. It is a special case of the Margulis normal subgroup theorem [36] that $\mathrm{PSp}(2g, \mathbb{Z})$ is *just infinite* for $g \geq 2$, meaning that all quotient groups are finite. Meanwhile, a theorem of Mennicke and Bass-Lazard-Serre [7, 38] says that $\mathrm{Sp}(2g, \mathbb{Z})$ has the *congruence subgroup property*, meaning that all finite quotients factor through $\mathrm{Sp}(2g, \mathbb{Z}/n)$ for some integer $n > 1$. Every finite quotient of $\mathrm{PSp}(2g, \mathbb{Z})$ likewise factors through $\mathrm{PSp}(2g, \mathbb{Z}/n)$, so we only have to find the simple quotients of $\mathrm{PSp}(2g, \mathbb{Z}/n)$.

Clearly if a prime p divides n, then $\mathrm{PSp}(2g, \mathbb{Z}/p)$ is a simple quotient of $\mathrm{PSp}(2g, \mathbb{Z}/n)$. We claim that there are no others. Let N be the kernel of the joint homomorphism

$$f : \mathrm{PSp}(2g, \mathbb{Z}/n) \to \prod_{p | n \text{ prime}} \mathrm{PSp}(2g, \mathbb{Z}/p).$$

If $\mathrm{PSp}(2g, \mathbb{Z}/n)$ had another simple quotient, necessarily non-abelian, then by Corollary 3.5, it would also be a simple quotient of N. It is easy to check that N is nilpotent, so all of its simple quotients are abelian. $\square$

3.4. Rubik groups

Recall the notation that $G' = [G, G]$ is the commutator subgroup of a group G.

If G is a group and X is a G-set, then we define the G-set symmetric group $\text{Sym}_G(X)$ to be the group of permutations of X that commute with the action of G. (Equivalently, $\text{Sym}_G(X)$ is the group of automorphisms of X as a G-set.) In the case that there are only finitely many orbits, we define the *Rubik group* $\text{Rub}_G(X)$ to be the commutator subgroup $\text{Sym}_G(X)'$. (For instance, the actual Rubik's Cube group has a subgroup of index two of the form $\text{Rub}_G(X)$, where $G = C_6$ acts on a set X with 12 orbits of order 2 and 8 orbits of order 3.)

If every G-orbit of X is free and X/G has n elements, then we can recognize $\text{Sym}_G(X)$ as the restricted wreath product

$$\text{Sym}_G(X) \cong G \text{wr}_{X/G} \text{Sym}(X/G) \cong G \text{wr}_n \text{Sym}(n).$$

We introduce the more explicit notation

$$\text{Sym}(n, G) \overset{\text{def}}{=} G \text{wr}_n \text{Sym}(n)$$

$$\text{Alt}(n, G) \overset{\text{def}}{=} G \text{wr}_n \text{Alt}(n)$$

$$\text{Rub}(n, G) \overset{\text{def}}{=} \text{Sym}(n, G)'.$$

We can describe $\text{Rub}(n, G)$ as follows. Let G_{ab} be the abelianization of G, and define a map $\sigma : G^n \to G_{\text{ab}}$ by first abelianizing G^n and then multiplying the n components in any order. Let $\text{AD}(n, G) \leq G^n$ (AD as in "anti-diagonal") be the kernel of σ. Then:

Proposition 3.9. *For any integer $n > 1$ and any group G, the commutator subgroup of* $\text{Sym}(n, G)$ *is given by*

$$\text{Rub}(n, G) = \text{AD}(n, G) \rtimes \text{Alt}(n).$$

Proof. It is easy to check that $(\ker \sigma) \rtimes \text{Alt}(n)$ is a normal subgroup of $\text{Sym}(n, G)$ and that the quotient is the abelian group $G_{\text{ab}} \times C_2$. This shows that

$$\text{AD}(n, G) \rtimes \text{Alt}(n) \supseteq \text{Rub}(n, G).$$

To check the opposite inclusion, note that $\text{AD}(n, G) \rtimes \text{Alt}(n)$ is generated by the union of $(G')^n$, $\text{Alt}(n) = \text{Sym}(n)'$, and all permutations of elements of the form

$$(g, g^{-1}, 1, \ldots, 1) \in G^n.$$

Clearly $\text{Rub}(n, G)$ contains the former two subsets. Since

$$(g, g^{-1}, 1, \ldots, 1) = [(g, 1, 1, \ldots, 1), (1\ 2)]$$

(and similarly for other permutations), we see

$$\text{AD}(n, G) \rtimes \text{Alt}(n) \leq \text{Rub}(n, G).$$

We conclude with the desired equality. $\qquad\square$

The main result of this section is a condition on a group homomorphism to $\text{Rub}(n, G)$ that guarantees that it is surjective. We say that a group homomorphism

$$f : K \to \text{Sym}(n, G)$$

is *G-set k-transitive* if it acts transitively on ordered lists of k elements that all lie in distinct G-orbits.

Theorem 3.10. *Let G be a group and let $n \geq 7$ be an integer such that $\text{Alt}(n - 2)$ is not a quotient of G. Suppose that a homomorphism*

$$f : K \to \text{Rub}(n, G)$$

is G-set 2-transitive and that its composition with the projection $\text{Rub}(n, G) \to \text{Alt}(n)$ is surjective. Then f is surjective.

Proof of Theorem 3.10. In the proof we will mix Cartesian product notation for elements of G^n with cycle notation for permutations. The proof is divided into three steps.

Step 1: We let $H = f(K)$, and we consider its normal subgroup

$$D \overset{\text{def}}{=} H \cap G^n.$$

We claim that D is 2-locally surjective. To this end, we look at the subgroup $\text{Alt}(n - 2) \leq \text{Alt}(n)$ that fixes the last two letters (say). Then there is a projection

$$\pi : G^n \rtimes \text{Alt}(n - 2) \to G^2 \times \text{Alt}(n - 2)$$

given by retaining only the last two coordinates of $g \in G^n$. We let

$$J = \pi(H \cap (G^n \rtimes \text{Alt}(n - 2))).$$

Since H is G-set 2-transitive, the group J surjects onto G^2; since H surjects onto $\text{Alt}(n)$, J surjects onto $\text{Alt}(n - 2)$. Thus we can apply Lemma 3.6 to the inclusion

$$J \leq G^2 \times \text{Alt}(n - 2).$$

Since $\text{Alt}(n - 2)$ is not a quotient of G and therefore not G^2 either (by Lemma 3.7), we learn that

$$J = G^2 \times \text{Alt}(n - 2)$$

and that

$$G^2 \leq H \cap (G^n \rtimes \text{Alt}(n - 2)).$$

So the group $D = H \cap G^n$ surjects onto the last two coordinates of G^n. Since we can repeat the argument for any two coordinates, D is 2-locally surjective.

Step 2: Suppose that G is abelian. Then D is a subgroup of G^n which is 2-locally surjective. Since G^n is abelian, conjugation of elements of D by elements of H that surject onto $\text{Alt}(n)$ coincides with conjugation by $\text{Alt}(n)$; thus D is $\text{Alt}(n)$-invariant. By step 1, for each $g_1 \in G$, there exists an element

$$d_1 \overset{\text{def}}{=} (g_1, 1, g_3, g_4, \ldots, g_n) \in D.$$

We now form a commutator with elements in $\text{Alt}(n)$ to obtain

$$d_2 \overset{\text{def}}{=} [d_1, (1\ 2)(3\ 4)] = (g_1, g_1^{-1}, g_3 g_4^{-1}, g_3^{-1} g_4, 1, \ldots, 1) \in D.$$

$$d_3 \overset{\text{def}}{=} [d_2, (1\ 2\ 5)(3\ 4)(6\ 7)] = (g, 1, 1, 1, g^{-1}, 1, \ldots, 1) \in D.$$

The Alt(n)-orbit of d_3 generates AD(n,G), thus $D =$ AD(n,G).

Step 3: In the general case, step 2 tells us that $D_{ab} =$ AD(n,G_{ab}) is $(n-1)$-locally surjective. This together with step 1 tells us that $D \le G^n$ satisfies the hypothesis of Lemma 3.2, which tells us that $D = $ AD(n,G). It remains only to show that Alt$(n) \le H$. It suffices to show that H/D contains (indeed is) Alt(n) in the quotient group

$$\text{Alt}(n,G)/D \cong (G^n/D) \rtimes \text{Alt}(n) \cong (G^n/D) \times \text{Alt}(n).$$

Now let $D_0 = (G^n \cap H)/D$, so that H surjects onto $D_0 \times$ Alt(n). Since Alt(n) is not a quotient of D_0 (for one reason, because G^n/D is abelian), we can thus apply Lemma 3.6 to conclude that H/D contains Alt(n). $\square$

Lemma 3.11. *If G is a group and $n \ge 5$, then* Rub$(n,G)/$AD$(n,G) \cong $ Alt(n) *is the unique simple quotient of* Rub(n,G).

Proof. We first claim that Rub(n,G) is a perfect group. For any two elements $g,h \in G$, we can take commutators such as

$$[(g_1,g_1^{-1},1,1,\ldots,1),(g_2,1,g_2^{-1},1,\ldots,1)]$$
$$= ([g_1,g_2],1,1,\ldots,1) \in \text{AD}(n,G)',$$

to conclude that

$$(G')^n = \text{AD}(n,G)' \le \text{Rub}(n,G)'.$$

We can thus quotient Rub(n,G) by $(G')^n$ and replace G by G_{ab}, or equivalently assume that G is abelian. In this case, we can take commutators such as

$$[(g,1,g^{-1},1,1,\ldots,1),(1\,2)(4\,5)]$$
$$= (g,g^{-1},1,1,\ldots,1) \in \text{Rub}(n,G)'$$

to conclude that AD$(n,G) \le $ Rub$(n,G)'$. Meanwhile Alt$(n) \le$ Rub$(n,G)'$ because it is a perfect subgroup. Thus Rub(n,G) is perfect.

Suppose that

$$f : \text{Rub}(n,G) \twoheadrightarrow Q$$

is a second simple quotient map, necessarily non-abelian. Then Corollary 3.5 tells us that f is also surjective when restricted to AD(n,G). If G is abelian, then so is AD(n,G) and this is immediately impossible. Otherwise we obtain that the restriction of f to AD$(n,G)' = (G')^n$ is again surjective, and we can apply Lemma 3.7 to conclude that $f|_{(G')^n}$ factors through a quotient $h : G' \to Q$ on a single factor. But then $(\ker f) \cap (G')^n$ would not be invariant under conjugation by Alt(n) even though it is the intersection of two normal subgroups of Rub(n,G), a contradiction. $\square$

4. PROOF OF THEOREM 1.1

In this section, we will prove Theorem 1.1 in three stages. In Section 4.1, we define an ad hoc circuit model called ZSAT

in which the alphabet has a group action and also has an unwanted zombie state z. Despite its contrived features, RSAT reduces to ZSAT, which is thus #P-complete. In Section 4.2, we refine Theorem 1.4 of Dunfield and Thurston in several ways for our purposes. Finally in Section 4.3, we build a homology 3-sphere M from a ZSAT circuit that satisfies the requirements of Theorem 1.1.

4.1. Zombies

Let G be a non-trivial finite group and let A be an alphabet which is a G-set with a single fixed point z, the *zombie symbol*, and otherwise with free orbits. We choose two G-invariant alphabets $I, F \subsetneq A \smallsetminus \{z\}$, and we assume that

$$|I|, |F| \ge 2|G| \qquad I \ne F \qquad |A| \ge 2|I \cup F| + 3|G| + 1. \quad (4)$$

(The second and third conditions are for convenience rather than necessity.) With these parameters, we define a planar circuit model that we denote ZSAT$_{G,A,I,F}$ that is the same as RSAT$_{A,I\cup\{z\},F\cup\{z\}}$ as defined in Section 2.2, *except* that the gate set is Rub$_G(A^2)$. This gate set is not universal in the sense of RSAT because every gate and thus every circuit is G-equivariant. (One can show that it is universal for G-equivariant circuits by establishing an analogue of Lemma 2.3 with the aid of Theorem 3.10, but we will not need this.) More explicitly, in the ZSAT model we consider G-equivariant planar circuits C that are composed of binary gates in Rub$_G(A^2)$, and satisfiability is defined by the equation $C(x) = y$ with $x \in (I \cup \{z\})^n$ and $y \in (F \cup \{z\})^n$.

Lemma 4.1. #ZSAT$_{G,A,I,F}$ *is almost parsimoniously #P-complete. More precisely, if $c \in$ #P, then there is an $f \in$ FP such that*

$$\#\text{ZSAT}_{G,A,I,F}(f(x)) = |G|c(x) + 1. \quad (5)$$

Equation (5) has the same form as equation (2), and for an equivalent reason: The input $(z,z,\ldots,z)$ trivially satisfies any ZSAT circuit (necessarily at both ends), while G acts freely on the set of other circuit solutions.

Proof. We take the convention that A is a left G-set. We choose a subset $A_0 \subsetneq A$ that has one representative from each free G-orbit of A. (In other words, A_0 is a section of the free orbits.)

We say that a data state $(a_1,a_2,\ldots,a_n)$ of a ZSAT circuit of width n is *aligned* if it has no zombie symbols and if there is a single element $g \in G$ such that $ga_i \in A_0$ for all i. The idea of the proof is to keep zombie symbols unchanged (which is why they are called zombies) and preserve alignment in the main reduction, and then add a postcomputation that converts zombies and misaligned symbols into warning symbols in a separate warning alphabet. The postcomputation cannot work if all symbols are zombies, but it can work in all other cases.

More precisely, we let $W \subseteq A \smallsetminus (I \cup F \cup \{z\})$ be a G-invariant subalphabet of size $|I \cup F| + 2|G|$ which we call the *warning alphabet*, and we distinguish two symbols $z_1, z_2 \in W$

in distinct orbits. Using Theorem 2.4 as a lemma, we will reduce a circuit C in the planar, reversible circuit model $\mathsf{RSAT}_{(I \cup F)/G, I/G, F/G}$ with binary gates to a circuit D in $\mathsf{ZSAT}_{G,A,I,F}$. To describe the reduction, we identify each element of $(I \cup F)/G$ with its lift in A_0.

We let D have the same width n as C. To make D, we convert each binary gate γ of the circuit C in $\mathsf{RSAT}_{(I \cup F)/G, I/G, F/G}$ to a gate δ in $\mathsf{ZSAT}_{G,A,I,F}$ in sequence. After all of these gates, D will also have a postcomputation stage. Given γ, we define δ as follows:

1. Of necessity,

$$\delta(z,z) = (z,z).$$

2. If $a \in I \cup F$, then

$$\delta(z,a) = (z,a) \qquad \delta(a,z) = (a,z).$$

3. If $a_1, a_2 \in (I \cup F) \cap A_0$, $g_1, g_2 \in G$, and

$$\gamma(a_1, a_2) = (b_1, b_2),$$

then

$$\delta(g_1 a_1, g_2 a_2) = (g_1 b_1, g_2 b_2).$$

4. We extend δ to the rest of A^2 so that $\delta \in \mathrm{Rub}_G(A^2)$.

By cases 1 and 2, zombie symbols stay unchanged. Cases 1, 2, and 3 together keep the computation within the subalphabet $I \cup F \cup \{z\}$, while case 3 preserves alignments, as well as misalignments.

The postcomputation uses a gate $\alpha : A^2 \to A^2$ such that:

1. Of necessity,

$$\alpha(z,z) = (z,z).$$

2. If $a \in I \cap A_0$, then

$$\alpha(z,a) = (z_1, a) \qquad \alpha(a,z) = (z_2, a)$$

3. If $a_1, a_2 \in I \cup F$ are misaligned, then

$$\alpha(a_1, a_2) = (\beta(a_1), a_2)$$

for some G-equivariant bijection

$$\beta : I \cup F \to W \smallsetminus (G z_1 \cup G z_2).$$

4. If $a_1, a_2 \in I \cup F$ are aligned, then

$$\alpha(a_1, a_2) = (a_1, a_2).$$

5. We extend α to the rest of A^2 so that $\alpha \in \mathrm{Rub}_G(A^2)$.

We apply this gate α to each adjacent pair of symbols (a_i, a_{i+1}) for i ranging from 1 to $n-1$ in order. The final effect is that, if some (but not all) of the input symbols are zombies, or if any two symbols are misaligned, then the postcomputation in D creates symbols in the warning alphabet W.

Any input to D with either zombies or misaligned symbols cannot finalize, since the main computation preserves these syndromes and the postcomputation then produces warning symbols that do not finalize. The only spurious input that finalizes is the all-zombies state $(z, z, \ldots, z)$, and otherwise each input that C accepts yields a single aligned G-orbit. Thus we obtain the relation

$$\#D = |G| \#C + 1$$

between the number of inputs that satisfy C and the number that satisfy D, as desired. $\qquad\square$

4.2. Theorem 1.4 refined

In this subsection and the next one, we let G be a fixed finite simple group, and we use "eventually" to mean "when the genus g is sufficiently large".

Recall from Section 1 that we consider several sets of homomorphisms of the fundamental group of the surface Σ_g to G:

$$\hat{R}_g(G) \stackrel{\text{def}}{=} \{ f : \pi_1(\Sigma_g) \to G \}$$
$$R_g(G) \stackrel{\text{def}}{=} \{ f : \pi_1(\Sigma_g) \twoheadrightarrow G \} \subseteq \hat{R}_g(G)$$
$$R_g^s(G) \stackrel{\text{def}}{=} \{ f \in R_g \mid \mathrm{sch}(f) = s \}.$$

For convenience we will write $R_g = R_g(G)$, etc., and only give the argument of the representation set when the target is some group other than G.

The set $\hat{R}_g$ has an action of $K = \mathrm{Aut}(G)$ and a commuting action of $\mathrm{MCG}_*(\Sigma_g)$, so we obtain a representation map

$$\rho : \mathrm{MCG}_*(\Sigma_g) \to \mathrm{Sym}_K(\hat{R}_g).$$

Since $\mathrm{MCG}_*(\Sigma_g)$ is perfect for $g \geq 3$ [15, Thm. 5.2] (and we are excluding small values of g), we can let the target be $\mathrm{Rub}_K(\hat{R}_g)$ instead. Now R_g and R_g^0 are both invariant subsets under both actions; in particular the representation map projects to maps to $\mathrm{Sym}_K(\hat{R}_g \smallsetminus R_g)$ and $\mathrm{Sym}_K(R_g^0)$. At the same time, $\mathrm{MCG}_*(\Sigma_g)$ acts on $H_1(\Sigma_g) \cong \mathbb{Z}^{2g}$, and we get a surjective representation map

$$\tau : \mathrm{MCG}_*(\Sigma_g) \to \mathrm{Sp}(2g, \mathbb{Z}),$$

whose kernel is by definition the Torelli group $\mathrm{Tor}_*(\Sigma_g)$.

The goal of this subsection is the following theorem.

Theorem 4.2. *The image of the joint homomorphism*

$$\rho_{R_g^0} \times \rho_{\hat{R}_g \smallsetminus R_g} \times \tau : \mathrm{MCG}_*(\Sigma_g)$$
$$\to \mathrm{Rub}_K(R_g^0) \times \mathrm{Rub}_K(\hat{R}_g \smallsetminus R_g) \times \mathrm{Sp}(2g, \mathbb{Z})$$

eventually contains $\mathrm{Rub}_K(R_g^0)$.

Comparing Theorem 4.2 to the second part of Theorem 1.4, it says that Theorem 1.4 still holds for the smaller Torelli group $\mathrm{Tor}_*(\Sigma_g)$, and after that the action homomorphism is still surjective if we lift from $\mathrm{Alt}(R_g^0/K)$ to $\mathrm{Rub}_K(R_g^0)$. Its third implication is that we can restrict yet further to the subgroup of $\mathrm{Tor}_*(\Sigma_g)$ that acts trivially on $\hat{R}_g \smallsetminus R_g$, the set of non-surjective homomorphisms to G.

The proof uses a lemma on relative sizes of representation sets.

Lemma 4.3. *Eventually*

$$|R_g^0/K| > |\hat{R}_g \smallsetminus R_g|.$$

Proof. Informally, if g is large and we choose a homomorphism $f \in \hat{R}_g$ at random, then it is a surjection with very high probability; if it is a surjection, then its Schur invariant $\mathrm{sch}(f)$ is approximately equidistributed. In detail, Dunfield-Thurston [14, Lems. 6.10 & 6.13] show that

$$\lim_{g \to \infty} \frac{|R_g|}{|\hat{R}_g|} = 1 \qquad \lim_{g \to \infty} \frac{|R_g^0|}{|R_g|} = \frac{1}{|H_2(G)|}.$$

Thus

$$\lim_{g \to \infty} \frac{|\hat{R}_g \smallsetminus R_g|}{|R_g^0/K|} = |H_2(G)| \cdot |K| \cdot \left(\lim_{g \to \infty} \frac{|\hat{R}_g|}{|R_g|} - 1 \right) = 0. \quad \square$$

Proof of Theorem 4.2. We first claim that $\rho_{R_g^0}$ by itself is eventually surjective. Note that the action of K on R_g^0 is free; thus we can apply Theorem 3.10 if $\rho_{R_g^0}$ satisfies suitable conditions. By part 2 of Theorem 1.4, $\rho_{R_g^0}$ is eventually surjective when composed with the quotient $\mathrm{Rub}_K(R_g^0) \to \mathrm{Sym}(R_g^0/K)$. Meanwhile, part 1 of Theorem 1.4 says that $\mathrm{MCG}_*(\Sigma_g)$ eventually acts transitively on $R_g^0(G^2)$. Since G is simple, Lemma 3.4 tells us that the homomorphisms $f \in R_g^0(G^2)$ correspond to pairs of surjections

$$f_1, f_2 : \Sigma_g \twoheadrightarrow G$$

that are inequivalent under $K = \mathrm{Aut}(G)$. This eventuality is thus the condition that the action of $\mathrm{MCG}_*(\Sigma_g)$ is K-set 2-transitive in its action on R_g^0. (*Cf.* Lemma 7.2 in [14].) Thus $\rho_{R_g^0}$ eventually satisfies the hypotheses of Theorem 3.10 and is surjective.

The map τ also surjects $\mathrm{MCG}_*(\Sigma_g)$ onto $\mathrm{Sp}(2g, \mathbb{Z})$. Lemmas 3.8 and 3.11 thus imply that $\mathrm{Rub}_K(R_g^0)$ and $\mathrm{Sp}(2g, \mathbb{Z})$ do not share any simple quotients. By Lemma 3.6, $\mathrm{MCG}_*(\Sigma_g)$ surjects onto $\mathrm{Rub}_K(R_g^0) \times \mathrm{Sp}(2g, \mathbb{Z})$. Equivalently, $\ker \tau = \mathrm{Tor}_*(\Sigma_g)$ surjects onto $\mathrm{Rub}_K(R_g^0)$.

Finally we consider

$$\rho_{R_g^0} \times \rho_{\hat{R}_g \smallsetminus R_g} : \mathrm{Tor}_*(\Sigma_g) \to \mathrm{Rub}_K(R_g^0) \times \mathrm{Rub}_K(\hat{R}_g \smallsetminus R_g),$$

which we have shown surjects onto the first factor. The unique simple quotient $\mathrm{Alt}(R_g^0/K)$ of $\mathrm{Rub}_K(R_g^0)$ is eventually not involved in $\mathrm{Rub}_K(\hat{R}_g \smallsetminus R_g)$ because it is too large. More precisely, Lemma 4.3 implies that eventually

$$|\mathrm{Alt}(R_g^0/K)| > |\mathrm{Alt}(\hat{R}_g \smallsetminus R_g)| > |\mathrm{Rub}_K(\hat{R}_g \smallsetminus R_g)|.$$

Thus we can apply Lemma 3.6 to conclude that the image of $\mathrm{Tor}_*(\Sigma_g)$ contains $\mathrm{Rub}_K(R_g^0)$, which is equivalent to the conclusion. $\square$

4.3. Mapping class gadgets

In this subsection and the next one, we will finish the proof of Theorem 1.1. We want to convert a suitable ZSAT circuit C of width n to a homology 3-sphere M. To this end, we choose some sufficiently large g that depends only on the group G, and we let Σ_{ng} be a Heegaard surface of M. This Heegaard surface will be decorated in various ways that we summarize in Figure 6. We use the additional notation that $\Sigma_{g,k}$ is a surface of genus g with k boundary circles, with a basepoint on one of its circles. We give $\Sigma_{g,k}$ the representation set

$$\hat{R}_{g,k} \stackrel{\mathrm{def}}{=} \{ f : \pi_1(\Sigma_{g,k}) \to G \}.$$

We let $\mathrm{MCG}_*(\Sigma_{g,k})$ be the relative mapping class group (that fixes $\partial \Sigma_{g,k}$); it naturally acts on $\hat{R}_{g,k}$.

We attach two handlebodies $(H_{ng})_I$ and $(H_{ng})_F$ to Σ_{ng} so that

$$(H_{ng})_I \cup (H_{ng})_F \cong S^3.$$

Although an actual sphere S^3 is not an interesting homology sphere for our purposes, our goal is to construct a homeomorphism $\phi \in \mathrm{MCG}_*(\Sigma_{ng})$ so that

$$M \stackrel{\mathrm{def}}{=} (H_{ng})_I \sqcup_\phi (H_{ng})_F$$

is the 3-manifold that we will produce to prove Theorem 1.1. (We could let ϕ be an element of the unpointed mapping class group here, but it is convenient to keep the basepoint.)

We identify n disjoint subsurfaces

$$(\Sigma_{g,1})_1, (\Sigma_{g,1})_2, \ldots, (\Sigma_{g,1})_n \subseteq \Sigma_{ng}$$

which are each homeomorphic to $\Sigma_{g,1}$. The handlebodies $(H_{ng})_I$ and $(H_{ng})_F$ likewise have sub-handlebodies $(H_g)_{I,i}$ and $(H_g)_{F,i}$ of genus g associated with $(\Sigma_{g,1})_i$ and positioned so that

$$(H_g)_{I,i} \cup (H_g)_{F,i} \cong B^3.$$

We also choose another set of subsurfaces

$$(\Sigma_{2g,1})_{1,2}, (\Sigma_{2g,1})_{2,3}, \ldots, (\Sigma_{2g,1})_{n-1,n} \subseteq \Sigma_{ng}$$

such that

$$(\Sigma_{g,1})_i, (\Sigma_{g,1})_{i+1} \subseteq (\Sigma_{2g,1})_{i,i+1}.$$

Finally we mark basepoints for each subsurface $(\Sigma_{g,1})_i$ and $(\Sigma_{2g,1})_{i,i+1}$, and one more basepoint $p_0 \in \Sigma_{ng}$, and we mark a set of connecting paths as indicated in Figure 6.

The circuit conversion is summarized in Figure 7. We will use the computational alphabet

$$A \stackrel{\mathrm{def}}{=} R_g^0 \cup \{z\} \subseteq \hat{R}_g \subseteq \hat{R}_{g,1},$$
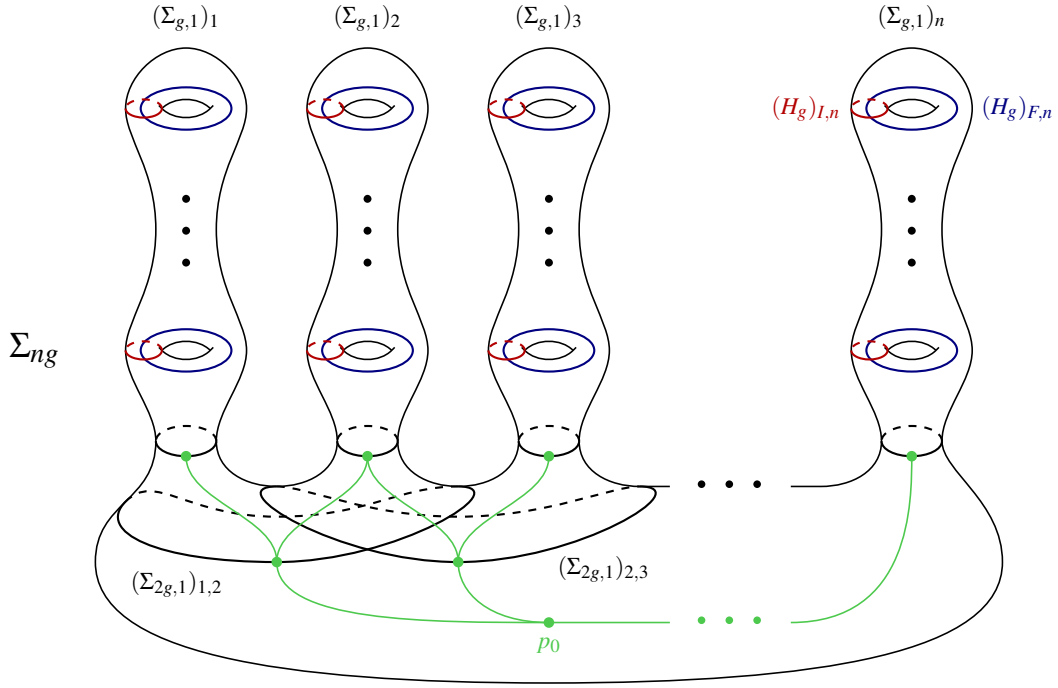
Figure 6. The Heegaard surface Σ_{ng} with disjoint subsurfaces $(\Sigma_{g,1})_i$. Circles that contract in $(H_g)_{I,i} \subset (H_{ng})_I$ are in red, while circles that contract in $(H_g)_{F,i} \subset (H_{ng})_F$ are in blue. The subsurfaces $(\Sigma_{2g,1})_{i,i+1}$ are also indicated. The system of basepoints and connecting paths is in green.

$\mathrm{ZSAT}_{K,A,I,F}$	$H(M,G)$
n-symbol memory	Heegaard surface Σ_{ng}
1-symbol memory	computational subsurface $\Sigma_{g,1}$
binary gate	element of $\mathrm{MCG}_*(\Sigma_{2g})$
circuit C	mapping class $\phi \in \mathrm{MCG}_*(\Sigma_{ng})$
alphabet A	homomorphisms $\pi_1(\Sigma_g^1) \to G$
zombie symbol: $z \in A$	trivial map $z : \pi_1(\Sigma_g^1) \to G$
memory state: $x \in A^n$	homomorphism $f : \pi_1(\Sigma_{ng}) \to G$
initialization: $x \in (I \cup \{z\})^n$	f extends to $\pi_1((H_{ng})_I)$
finalization: $y \in (F \cup \{z\})^n$	f extends to $\pi_1((H_{ng})_F)$
solution: $C(x) = y$	homomorphism $f : \pi_1(M) \to G$

Figure 7. A correspondence between ZSAT and $H(M,G)$.

where $z : \pi_1(\Sigma_g) \to G$ is (as first mentioned in Section 1.1) the trivial homomorphism and the zombie symbol, and the inclusion $\hat{R}_g \subseteq \hat{R}_{g,1}$ comes from the inclusion of surfaces $\Sigma_{g,1} \subseteq \Sigma_g$. Note that $\hat{R}_g^0$ is precisely the subset of $\hat{R}_{g,1}$ consisting of homomorphisms

$$f : \pi_1(\Sigma_{g,1}) \to G$$

that are trivial on the peripheral subgroup $\pi_1(\partial\Sigma_{g,1})$.

Each subsurface $(\Sigma_{g,1})_i$ is interpreted as the "memory unit" of a single symbol $x_i \in A$. Using the connecting paths in Σ_{ng} between the basepoints of its subsurfaces, and since each x_i is

trivial on $\pi_1(\partial\Sigma_{g,1})$, a data register

$$x = (x_1, x_2, \ldots, x_n) \in A^n$$

combines to form a homomorphism

$$f : \pi_1(\Sigma_{ng}) \to G.$$

In particular, if $x \neq (z, z, \ldots, z)$, then $f \in R_{ng}$. In other words, f is surjective in this circumstance because one of its components x_i is already surjective. (Note that the converse is not true: we can easily make a surjective f whose restriction to each $(\Sigma_{g,1})_i$ is far from surjective.)

For every subgroup $K \leq G$, we define $I(K)$ to be the set of surjections

$$x : \pi_1(\Sigma_{g,1}) \twoheadrightarrow K$$

that come from a homomorphism

$$x : \pi_1((H_g)_I) \twoheadrightarrow K.$$

We define $F(K)$ in the same way using $(H_g)_F$. A priori we know that $I(K), F(K) \subseteq R_{g,1}(K)$. This inclusion can be sharpened in two significant respects.

Lemma 4.4. *The sets $I(K)$ and $F(K)$ are subsets of $R_g^0(K)$. If K is non-trivial, then they are disjoint.*

Proof. First, since $\partial\Sigma_{g,1}$ bounds a disk in $(H_g)_I$, we obtain that $I(K), F(K) \subseteq R_g(K)$. Second, since any x in $I(K)$ or

$F(K)$ extends to a handlebody, the cycle $x_*([\Sigma_g])$ is null-homologous in BG and therefore $\mathrm{sch}(x) = 0$. Third, since $(H_g)_I \cup (H_g)_F \cong B^3$ is simply connected, a surjective homomorphism $x \in R_g(K)$ cannot extend to both handlebodies if K is non-trivial. Therefore $I(K)$ and $F(K)$ are disjoint in this case. □

The gadgets that serve as binary gates are mapping class elements $\alpha \in \mathrm{MCG}_*((\Sigma_{2g,1})_{i,i+1})$ that act on two adjacent memory units $(\Sigma_{g,1})_i$ and $(\Sigma_{g,1})_{i+1}$. We summarize the effect of the local subsurface inclusions on representation sets. In order to state it conveniently, if X and Y are two pointed spaces, we define a modified wedge $X \vee_\lambda Y$, where λ is a connecting path between the basepoint of X and the basepoint of Y. Figure 8 shows a surjection from Σ_{2g} to $\Sigma_g \vee_\lambda \Sigma_g$, while Figure 6 has copies of $\Sigma_{g,1} \vee_\lambda \Sigma_{g,1}$ (which has a similar surjection from $\Sigma_{2g,1}$.

Lemma 4.5. *The inclusions and surjections*

$$
\begin{array}{ccc}
\Sigma_{2g,1} & \subseteq & \Sigma_{2g} \\
\downarrow & & \downarrow \\
\Sigma_{g,1} \vee_\lambda \Sigma_{g,1} & \subseteq & \Sigma_g \vee_\lambda \Sigma_g
\end{array}
$$

yield the inclusions

$$
\begin{array}{ccccccc}
\hat{R}_{2g,1} & \supseteq & \hat{R}_{2g} & \supseteq & R_{2g} & \supseteq & R_{2g}^0 \\
\| & & \cup | & & \cup | & & \cup | \\
\hat{R}_{g,1} \times \hat{R}_{g,1} & \supseteq & \hat{R}_g \times \hat{R}_g & \supseteq & R_g \times R_g & \supseteq & R_g^0 \times R_g^0
\end{array} . \quad (6)
$$

For every pair of subgroups $K_1, K_2 \le G$ that generate $K \le G$, they also yield

$$
R_g^0(K_1) \times R_g^0(K_2) \subseteq R_{2g}^0(K). \quad (7)
$$

Finally, they yield

$$
A \times A \subseteq R_{2g}^0 \cup \{z_{2g}\}, \quad (8)
$$

where $z_g \in R_g$ is the trivial map in genus g and $z_{2g} = (z_g, z_g)$.
Proof. The horizontal inclusions are all addressed above; the real issue is the vertical inclusions and equalities. We consider the vertical inclusions from left to right in diagram (6). The surjection

$$
\sigma_1 : \Sigma_{2g,1} \twoheadrightarrow \Sigma_{g,1} \vee_\lambda \Sigma_{g,1}
$$

is an isomorphism of π_1, while the surjection

$$
\sigma_0 : \Sigma_{2g} \twoheadrightarrow \Sigma_g \vee_\lambda \Sigma_g
$$

is a surjection in π_1. This implies the first two vertical relations. Then, if two homomorphisms

$$
f_1, f_2 : \pi_1(\Sigma_g) \twoheadrightarrow G
$$

are each surjective, then they are certainly jointly surjective; this implies the third relation. Finally, the surjection σ_0 yields the formula

$$
\mathrm{sch}((f_1, f_2)) = \mathrm{sch}(f_1) + \mathrm{sch}(f_2). \quad (9)
$$

The reason is that the image $\sigma_0([\Sigma_{2g}])$ of the fundamental class of Σ_{2g} is the sum of the fundamental classes of the two Σ_g components. This yields the fourth, leftmost inclusion because equation (9) then reduces to $0 = 0 + 0$.

To treat (7), we claim that if $\mathrm{sch}_{K_i}(f_i) = 0$, then $\mathrm{sch}_K(f_i) = 0$. This follows from the fact that each map from Σ_g to the classifying space BK_i and BK forms a commutative triangle with the map $BK_i \to BK$. With this remark, inclusion (7) can be argued in the same way as the inclusion $R_g^0 \times R_g^0 \subseteq R_{2g}^0$.

Finally for inclusion (8), recall that $A = R_g^0 \cup \{z_g\}$, and that $z_{2g} = (z_g, z_g)$ since in each case z is the trivial homomorphism. The inclusions

$$
R_g^0 \times \{z_g\}, \{z_g\} \times R_g^0 \subseteq R_{2g}^0
$$

can be argued the same way as before: Given the two homomorphisms f_1, f_2, even if one of them is the trivial homomorphism z_g, the surjectivity of the other one gives us joint surjectivity. Moreover, the trivial homomorphism z_g has a vanishing Schur invariant $\mathrm{sch}_G(z_g) = 0$ relative to the target group G. □

4.4. End of the proof

We combine Theorem 4.2 with Lemmas 4.5 and 4.1 to convert a circuit C in $\mathrm{ZSAT}_{K,A,I,F}$ to a mapping class $\phi \in \mathrm{MCG}_*(\Sigma_{ng})$ using mapping class gadgets. To apply Lemma 4.1, we need to verify the conditions in (4). These follow easily from asymptotic estimates on the cardinality of A and I [14, Lems. 6.10 & 6.11].

For each $\gamma \in \mathrm{Rub}_K(A \times A)$, we choose an $\alpha \in \mathrm{Tor}_*(\Sigma_{2g,1})$ such that:

1. α acts by γ on $A \times A$.

2. α acts by an element of $\mathrm{Rub}_K(R_{2g}^0)$ that fixes $R_{2g}^0 \smallsetminus (A \times A)$.

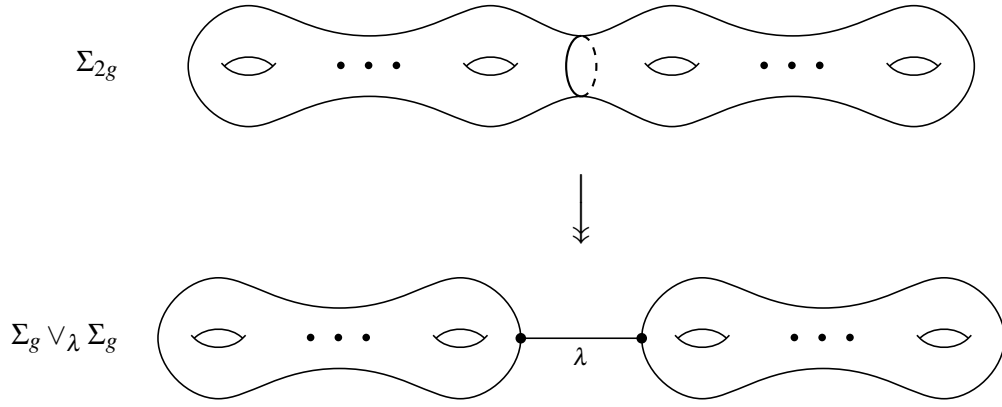3. α fixes $\hat{R}_{2g} \smallsetminus R_{2g}$.

Given a circuit C in $\mathrm{ZSAT}_{K,A,I,F}$, we can replace each gate $\gamma \in \mathrm{Rub}_K(A \times A)$ that acts on symbols i and $i+1$ by the corresponding local mapping class $\alpha \in \mathrm{Tor}_*((\Sigma_{2g,1})_{(i,i+1)})$. Then we let ϕ be the composition of the gadgets α.

Lemma 4.6. *Let*

$$
M \stackrel{\text{def}}{=} (H_{ng})_I \sqcup_\phi (H_{ng})_F.
$$

Then

1. *M is a homology 3-sphere.*

2. *If $1 \lneq K \lneq G$ is a non-trivial, proper subgroup of G, then $Q(M, K) = \emptyset$.*

3. *$\#H(M, G) = \#C$.*

Figure 8. From Σ_{2g} to $\Sigma_g \vee_\lambda \Sigma_g$

Proof. Point 1 holds because by construction, $\phi \in \mathrm{Tor}(\Sigma_{2g})$.

To address points 2 and 3, we decompose ϕ as a composition of local gadgets,

$$\phi = \alpha_m \circ \alpha_{m-1} \circ \cdots \circ \alpha_2 \circ \alpha_1, \qquad (10)$$

and we insert parallel copies $(\Sigma_{ng})_j$ of the Heegaard surface with $0 \le j \le m$, so the ith gadget α_j yields a map

$$\alpha_j : (\Sigma_{ng})_{j-1} \to (\Sigma_{ng})_j$$

from the $(j-1)$-st to the j-th surface. Each α_j is a non-trivial homeomorphism

$$\alpha_j : (\Sigma_g)_{j-1,(i,i+1)} \to (\Sigma_{ng})_{j,(i,i+1)}$$

for some i, and is the identity elsewhere. We use this decomposition to analyze the possibilities for a group homomorphism

$$f : \pi_1(M) \to G.$$

The map f restricts to a homomorphism

$$f_j : \pi_1((\Sigma_{ng})_j) \to G,$$

and then further restricts to a homomorphism

$$f_{j,i} : \pi_1((\Sigma_{g,1})_{j,i}) \to G$$

for the ith memory unit for each i. It is convenient to interpret $\hat{R}_{g,1} \supseteq A$ as the superalphabet of all possible symbols that could in principle arise as the state of a memory unit.

By construction, each initial symbol $f_{0,i}$ extends to the handlebody $(H_g)_{I,i}$. Thus $f_{0,i} \in I(K)$ for some subgroup $1 \le K \le G$, and all cases are disjoint from A other than $K=1$ and $K=G$. Likewise at the end, each $f_{m,i} \in F(K)$ for some K. By construction, each α_j fixes both $R^0_{2g} \smallsetminus (A \times A)$ and $\hat{R}_{2g} \smallsetminus R_{2g}$. This fixed set includes all cases $R^0(K_1) \times R^0(K_2)$, and therefore all cases $I(K_1) \times I(K_2)$, other than $K_1, K_2 \in \{1, G\}$. Thus every initial symbol $f_{0,i} \in I(K) \not\subseteq A$ is preserved by every gadget α_j, and then can't finalize because $I(K) \cap F(K) = \emptyset$. Among other things, this establishes point 2 of the lemma.

This derivation also restricts the initial state f_0 to A^n. In this case, each α_j acts in the same way on A^n as the corresponding gate γ_j; consequently it leaves the set A^n invariant. Considering both the circuit action and initialization and finalization, these states exactly match the behavior of the circuit C under the rules of $\mathrm{ZSAT}_{K,A,I,F}$. $\qquad\square$

To complete the proof of Theorem 1.1, we only need to efficiently triangulate the 3-manifold $M \stackrel{\mathrm{def}}{=} (H_{ng})_I \sqcup_\phi (H_{ng})_F$. The first step is to refine the decoration of Σ_{ng} shown in Figure 6 to a triangulation. It is easy to do this with polynomial complexity in n (or in ng, but recall that g is fixed). We can also give each subsurface $(\Sigma_{g,1})_i$ thesame triangulation for all i, as well as each subsurface $(\Sigma_{2g,1})_{i,i+1}$. It is also routine to extend any such triangulation to either $(H_{ng})_I$ or $(H_{ng})_F$ with polynomial (indeed linear) overhead: Since by construction the triangulation of each $(\Sigma_{g,1})_i$ is the same, we pick some extension to $(H_g)_I$ and $(H_g)_F$ and use it for each $(H_g)_{I,i}$ and each $(H_g)_{F,i}$. The remainder of $(H_{ng})_I$ and $(H_{ng})_F$ is a 3-ball whose boundary has now been triangulated; any triangulation of the boundary of a 3-ball can be extended to the interior algorithmically and with polynomial complexity.

We insert more triangulated structure in between $(H_{ng})_I$ and $(H_{ng})_F$ to realize the homeomorphism ϕ. Recalling equation (10) in the proof of Lemma 4.6, ϕ decomposes into local mapping class gadgets α_j. Only finitely many $\alpha \in \mathrm{MCG}_*(\Sigma_{g,1})$ are needed, since we only need one representative for each $\gamma \in \mathrm{Rub}_K(A \times A)$. At this point it is convenient to use a blister construction. We make a 3-manifold W_α whose boundary is two copies of $\Sigma_{2g,1}$ (with its standard triangulation) that meet at their boundary circle, and so that W_α is a relative mapping cylinder for the homeomorphism α. If α_j acts on $(\Sigma_{2g,1})_{i,i+1}$, then we can have $(\Sigma_{ng})_{j-1}$ and $(\Sigma_{ng})_j$ coincide outside of $(\Sigma_{2g,1})_{j-1,(i,i+1)}$ and $(\Sigma_{2g,1})_{j,(i,i+1)}$, so that their union $(\Sigma_{ng})_{j-1} \cup (\Sigma_{ng})_j$ is a branched surface. We insert W_α and its triangulation in the blister within $(\Sigma_{ng})_{j-1} \cup (\Sigma_{ng})_j$; see Figure 9.
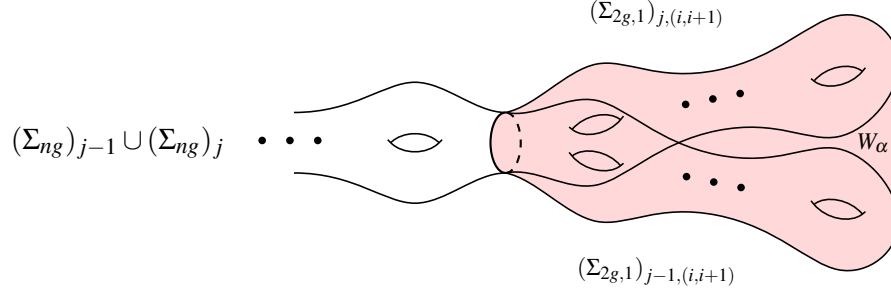
Figure 9. The blister W_α between $(\Sigma_{2g,1})_{j-1,(i,i+1)}$ and $(\Sigma_{2g,1})_{j,(i,i+1)}$.

5. FINAL REMARKS AND QUESTIONS

5.1. Sharper hardness

Even though the proof of Theorem 1.1 is a polynomially efficient reduction, for any fixed, suitable target group G, it is not otherwise particularly efficient. Various steps of the proof require the genus g (which is used to define the symbol alphabet Y_g^0) to be sufficiently large. In fact, the crucial Theorem 1.4 does not even provide a constructive lower bound on g. Dunfield and Thurston [14] discuss possibilities to improve this bound, and they conjecture that $g \geq 3$ suffices in Theorem 1.4 for many or possibly all choices of G. We likewise believe that there is some universal genus g_0 such that Theorem 4.2 holds for all $g \geq g_0$.

In any case, the chain of reduction summarized in Figure 2 is not very efficient either. What we really believe is that the random 3-manifold model of Dunfield and Thurston also yields computational hardness. More precisely, Johnson showed that the Torelli group $\text{Tor}(\Sigma_g)$ is finitely generated for $g \geq 3$ [25]. This yields a model for generating a random homology 3-sphere: We choose $\phi \in \text{Tor}(\Sigma_g)$ by evaluating a word of length ℓ in the Johnson generators, and then we let

$$M \stackrel{\text{def}}{=} (H_g)_I \sqcup_\phi (H_g)_F.$$

Our Theorem 4.2 implies that [14, Thm. 7.1] holds in this model, *i.e.*, that the distribution of $\#Q(M,G)$ converges to Poisson with mean $|H_2(G)|/|\text{Out}(G)|$ if we first send $\ell \to \infty$ and then send $g \to \infty$. We also conjecture that $\#Q(M,G)$ is hard on average in the sense of average-case computational complexity [3, Ch. 18] if ℓ grows faster than g.

Speaking non-rigorously, we conjecture that it is practical to randomly generate triangulated homology 3-spheres M in such a way that no one will ever know the value of $\#Q(M,G)$, say for $G = A_5$. Hence, no one will ever know whether such an M has a connected 5-sheeted cover.

5.2. Other spaces

Maher [35] showed that the probability that a randomly chosen M in the Dunfield-Thurston model is hyperbolic con-

verges to 1 as $\ell \to \infty$, for any fixed $g \geq 2$. Maher notes that the same result holds if M is a homology 3-sphere made using the Torelli group, for any $g \geq 3$. Thus our conjectures in Section 5.1 would imply that $\#Q(M,G)$ is computationally intractable when M is a hyperbolic homology 3-sphere.

We conjecture that a version of Theorem 1.1 holds when M fibers over a circle. In this case M cannot be a homology 3-sphere, but it can be a homology $S^2 \times S^1$. If M fibers over a circle, then the invariant $H(M,G)$ is obviously analogous (indeed a special case of) counting solutions to $C(x) = x$ when C is a reversible circuit. However, the reduction from C to M would require new techniques to avoid spurious solutions.

In forthcoming work [33], we will prove an analogue of Theorem 1.1 when M is a knot complement. We will use a theorem of Roberts and Venkatesh [48] which is itself an analogue of Theorem 1.4 for braid group actions.

5.3. Non-simple groups

We consider the invariant $\#H(M,G)$ for a general finite group G.

Recall that the *perfect core* G_{per} of a group G is its unique largest perfect subgroup; if G is finite, then it is also the limit of its derived series. If M is a homology sphere, then its fundamental group is perfect and $H(M,G) = H(M,G_{\text{per}})$. We conjecture then that a version of Theorem 1.1 holds for any finite, perfect group G. More precisely, we conjecture that Theorem 1.1 holds for $Q(M,G)$ when G is finite and perfect, and that the rest of $H(M,G)$ is explained by non-surjective homomorphisms $f : G \to G$. Mochon's analysis [39] in the case when G is non-solvable can be viewed as a partial result towards this conjecture.

If G is finite and G_{per} is trivial, then this exactly the case that G is solvable. In the case when M is a link complement, Ogburn and Preskill [42] non-rigorously conjecture that $H(M,G)$ is not "universal" for classical computation. It is very believable that the relevant actions of braid groups and mapping class groups are too rigid for any analogue of the second half of Theorem 1.4 to hold. Rowell [49] more precisely conjectured that $\#H(M,G)$ can be computed in polynomial time for any link complement M and any finite, solvable G. We are much less confident that this more precise conjecture is true.

[1] Dorit Aharonov and Itai Arad, *The* BQP-*hardness of approximating the Jones polynomial*, New J. Phys. **13** (2011), 035019, arXiv:quant-ph/0605181. 1.2

[2] Gorjan Alagic and Catharine Lo, *Quantum invariants of 3-manifolds and NP vs #P*, Quantum Inf. Comput. **17** (2017), no. 1-2, 125–146, arXiv:1411.6049. 1.2

[3] Sanjeev Arora and Boaz Barak, *Computational complexity: a modern approach*, Cambridge University Press, Cambridge, 2009. 2.1, 2.2, 5.1

[4] Matthias Aschenbrenner, Stefan Friedl, and Henry Wilton, *Decision problems for 3-manifolds and their fundamental groups*, Geom. Topol. Monogr. **19** (2015), 201–236, arXiv:1405.6274. 1.2

[5] Jane Austen and Seth Grahame-Smith, *Pride and prejudice and zombies*, Quirk Books, Philadelphia, 2009. 1

[6] Régis Barbanchon, *On unique graph 3-colorability and parsimonious reductions in the plane*, Theoret. Comput. Sci. **319** (2004), no. 1-3, 455–482. 1.1

[7] H. Bass, M. Lazard, and J.-P. Serre, *Sous-groupes d'indice fini dans* SL(n, **Z**), Bull. Amer. Math. Soc. **70** (1964), 385–392. 3.3

[8] Kristine Bauer, Debasis Sen, and Peter Zvengrowski, *A generalized Goursat lemma*, Tatra Mt. Math. Publ. **64** (2015), 1–19, arXiv:1109.0024. 3.3

[9] Roger W. Carter, *Simple groups of Lie type*, Wiley Classics Library, John Wiley & Sons, Inc., New York, 1989, Reprint of the 1972 original. 3.3

[10] Haimiao Chen, *Applying TQFT to count regular coverings of Seifert 3-manifolds*, J. Geom. Phys. **62** (2012), no. 6, 1347–1357, arXiv:1110.0186. 1.2, 2.3

[11] Persi Diaconis, R. L. Graham, and William M. Kantor, *The mathematics of perfect shuffles*, Adv. in Appl. Math. **4** (1983), no. 2, 175–196. 3.1

[12] Robbert Dijkgraaf and Edward Witten, *Topological gauge theories and group cohomology*, Comm. Math. Phys. **129** (1990), no. 2, 393–429. 1.2

[13] Bruce Randall Donald and David Renpan Chang, *On the complexity of computing the homology type of a triangulation*, 32nd Annual Symposium on Foundations of Computer Science, IEEE, 1991, pp. 650–661. 2.3

[14] Nathan M. Dunfield and William P. Thurston, *Finite covers of random 3-manifolds*, Invent. Math. **166** (2006), no. 3, 457–521, arXiv:math/0502567. 1.4, 3.2, 3.4, 4.2, 4.4, 5.1

[15] Benson Farb and Dan Margalit, *A primer on mapping class groups*, Princeton Mathematical Series, vol. 49, Princeton University Press, Princeton, NJ, 2012. 4.2

[16] Edward Fredkin and Tommaso Toffoli, *Conservative logic*, Internat. J. Theoret. Phys. **21** (1981/82), no. 3-4, 219–253. 2.2, 2.2

[17] Daniel S. Freed and Frank Quinn, *Chern-Simons theory with finite gauge group*, Comm. Math. Phys. **156** (1993), no. 3, 435–472, arXiv:hep-th/9111004. 1.2

[18] Michael H. Freedman, Michael Larsen, and Zhenghan Wang, *A modular functor which is universal for quantum computation*, Comm. Math. Phys. **227** (2002), no. 3, 605–622, arXiv:quant-ph/0001108. 1.2

[19] Michael H. Freedman, Michael J. Larsen, and Zhenghan Wang, *The two-eigenvalue problem and density of Jones representation of braid groups*, Comm. Math. Phys. **228** (2002), no. 1, 177–199, arXiv:math/0103200. 1.2

[20] Georg Frobenius and Issai Schur, *Über die reellen Darstellungen der endlichen Gruppen*, Sitzungsber. Königlich Preussis-chen Akad. Wiss. **8** (1906), 186–208. 1.2

[21] Mikael Goldmann and Alexander Russell, *The complexity of solving equations over finite groups*, Inform. and Comput. **178** (2002), no. 1, 253–262. 1.2, 1.2

[22] Édouard Goursat, *Sur les substitutions orthogonales et les divisions régulières de l'espace*, Ann. Sci. École Norm. Sup. **6** (1889), 9–102. 3.3

[23] Philip Hall, *The Eulerian functions of a group*, Quart. J. Math. **7** (1936), 134–151. 3.2

[24] François Jaeger, Dirk L. Vertigan, and Dominic Welsh, *On the computational complexity of the Jones and Tutte polynomials*, Math. Proc. Cambridge Philos. Soc. **108** (1990), no. 1, 35–53. 1.2

[25] Dennis Johnson, *The structure of the Torelli group. I. A finite set of generators for $\mathscr{I}$*, Ann. of Math. (2) **118** (1983), no. 3, 423–442. 5.1

[26] Ravindran Kannan and Achim Bachem, *Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix*, SIAM J. Comput. **8** (1979), no. 4, 499–507. 2.3

[27] Robert König, Greg Kuperberg, and Ben Reichardt, *Quantum computation with Turaev-Viro codes*, Ann. Phys. **325** (2010), 2707–2749, arXiv:1002.2816. 1.2

[28] Hari Krovi and Alexander Russell, *Quantum Fourier transforms and the complexity of link invariants for quantum doubles of finite groups*, Comm. Math. Phys. **334** (2015), no. 2, 743–777, arXiv:1210.1550. 1.2

[29] Greg Kuperberg, *Algorithmic homeomorphism of 3-manifolds as a corollary of geometrization*, arXiv:1508.06720. 1.2

[30] ———, *Involutory Hopf algebras and 3-manifold invariants*, Internat. J. Math. **2** (1991), no. 1, 41–66, arXiv:math/9201301. 1.2

[31] ———, *Denseness and Zariski denseness of Jones braid representations*, Geom. Topol. **15** (2011), 11–39, arXiv:0909.1881. 3.2

[32] ———, *How hard is it to approximate the Jones polynomial?*, Theory Comput. **11** (2015), 183–219, arXiv:0908.0512. 1.2

[33] Greg Kuperberg and Eric Samperton, *Coloring invariants of knots and links are often intractable*, in preparation. 1.2, 5.2

[34] W. B. Raymond Lickorish, *An introduction to knot theory*, Graduate Texts in Mathematics, vol. 175, Springer-Verlag, New York, 1997. 1.1

[35] Joseph Maher, *Random Heegaard splittings*, J. Topol. **3** (2010), no. 4, 997–1025, arXiv:0809.4881. 5.2

[36] Grigoriĭ A. Margulis, *Discrete subgroups of semisimple Lie groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 17, Springer-Verlag, 1991. 3.3

[37] Alexander Mednykh, *Determination of the number of nonequivalent coverings over a compact Riemann surface*, Dokl. Akad. Nauk SSSR **239** (1978), no. 2, 269–271. 1.2, 2.3

[38] Jens Mennicke, *Zur Theorie der Siegelschen Modulgruppe*, Math. Ann. **159** (1965), 115–129. 3.3

[39] Carlos Mochon, *Anyons from nonsolvable finite groups are sufficient for universal quantum computation*, Phys. Rev. A **67** (2003), 022315, arXiv:quant-ph/0206128. 1.2, 5.3

[40] Bernhard H. Neumann, *Some remarks on infinite groups*, J. London Math. Soc. **12** (1937), 120–127. 3.2

[41] Gustav Nordh and Peter Jonsson, *The complexity of counting solutions to systems of equations over finite semigroups*, Computing and combinatorics, Lecture Notes in Comput. Sci., vol. 3106, Springer, Berlin, 2004, pp. 370–379. 1.2, 1.2

[42] R. Walter Ogburn and John Preskill, *Topological quantum computation*, Quantum computing and quantum communications (Palm Springs, CA, 1998), Lecture Notes in Comput. Sci., vol. 1509, Springer, Berlin, 1999, pp. 341–356. 1.2, 5.3

[43] Paul Olum, *Non-abelian cohomology and van Kampen's theorem*, Ann. of Math. (2) **68** (1958), 658–668. 2.3

[44] Bjorn Poonen, *Undecidable problems: a sampler*, ch. 10, pp. 211–241, Cambridge University Press, 2014, arXiv:1204.0299. 1.2

[45] Nicolai Yu. Reshetikhin and Vladimir G. Turaev, *Ribbon graphs and their invariants derived from quantum groups*, Comm. Math. Phys. **127** (1990), no. 1, 1–26. 1.2

[46] ———, *Invariants of 3-manifolds via link polynomials and quantum groups*, Invent. Math. **103** (1991), no. 3, 547–597. 1.2

[47] Kenneth A. Ribet, *On l-adic representations attached to modular forms*, Invent. Math. **28** (1975), 245–275. 3.2, 3.2

[48] David P. Roberts and Akshay Venkatesh, *Hurwitz monodromy and full number fields*, Algebra Number Theory **9** (2015), no. 3, 511–545, arXiv:1401.7379. 5.2

[49] Eric C. Rowell, *Two paradigms for topological quantum computation*, Advances in quantum computation, Contemp. Math., vol. 482, Amer. Math. Soc., Providence, RI, 2009, arXiv:0803.1258, pp. 165–177. 5.3

[50] Vladimir G. Turaev, *Quantum invariants of knots and 3-manifolds*, W. de Gruyter, 1994. 1.2

[51] L. G. Valiant and V. V. Vazirani, *NP is as easy as detecting unique solutions*, Theoret. Comput. Sci. **47** (1986), no. 1, 85–93. 2.1

[52] Leslie G. Valiant, *The complexity of computing the permanent*, Theoret. Comput. Sci. **8** (1979), no. 2, 189–201. 2.1

[53] *The Complexity Zoo*, http://www.complexityzoo.com/. 2.1