SIAM J. COMPUT.
Vol. 32, No. 4, pp. 916–934

THE HIDDEN SUBGROUP PROBLEM AND QUANTUM COMPUTATION USING GROUP REPRESENTATIONS*

SEAN HALLGREN†, ALEXANDER RUSSELL‡, AND AMNON TA-SHMA§

Abstract. The hidden subgroup problem is the foundation of many quantum algorithms. An efficient solution is known for the problem over abelian groups, employed by both Simon's algorithm and Shor's factoring and discrete log algorithms. The nonabelian case, however, remains open; an efficient solution would give rise to an efficient quantum algorithm for graph isomorphism. We fully analyze a natural generalization of the algorithm for the abelian case to the nonabelian case and show that the algorithm determines the normal core of a hidden subgroup: in particular, normal subgroups can be determined. We show, however, that this immediate generalization of the abelian algorithm does not efficiently solve graph isomorphism.

Key words. quantum computation, quantum algorithms, computational complexity, representation theory, finite groups

AMS subject classifications. 81P68, 68Q17

DOI. 10.1137/S009753970139450X

1. Introduction.

Peter Shor's seminal article [27] presented efficient quantum algorithms for computing integer factorizations and discrete logarithms, problems thought to be intractable for classical computation models. A primary ingredient in these algorithms is an efficient solution to the *hidden subgroup problem* for certain abelian groups; indeed computing discrete logarithms directly reduces to the hidden subgroup problem. Formally, the hidden subgroup problem is the following.

DEFINITION 1.1. Hidden subgroup problem (HSP). *Given an efficiently computable function $f : G \to S$, from a finite group G to a set S, that is constant on (left) cosets of some subgroup H and takes distinct values on distinct cosets, determine the subgroup H.*

The general paradigm, which gives rise to efficient quantum algorithms for this problem over abelian groups, is the following.

EXPERIMENT 1.1 (experiment for the abelian HSP).

1. *Prepare the state*

$$\frac{1}{\sqrt{|G|}} \sum_{g \in G} |g, f(g)\rangle$$

and measure the second register $f(g)$. As f takes distinct values on the left cosets of

*Received by the editors August 28, 2001; accepted for publication (in revised form) December 16, 2002; published electronically June 10, 2003. A preliminary version of this article appeared in *Proceedings of the 32nd Annual ACM Symposium on Theory of Computing*, Portland, OR, 2000, pp. 627–635. The bulk of this research was completed while the authors were at the University of California, Berkeley.

http://www.siam.org/journals/sicomp/32-4/39450.html

†Computer Science Department, Caltech, MC 256-80, Pasadena, CA 91125 (hallgren@cs.caltech.edu). This author was supported in part by an NSF Mathematical Sciences Postdoctoral Fellowship, an NDSEG fellowship, a GAANN fellowship, and NSF grant CCR-9800024.

‡Department of Computer Science and Engineering, University of Connecticut, Storrs, CT 06269 (acr@cse.uconn.edu). This author was supported by NSF CAREER award CCR-0093065, NSF grant CCR-0220264, NSF grant EIA-0218443, NSF NYI grant CCR-9457799, and a David and Lucile Packard Fellowship for Science.

§Computer Science Division, Tel-Aviv University, Israel 69978 (amnon@post.tau.ac.il).

H, the resulting state is

$$\frac{1}{\sqrt{|H|}} \sum_{h \in H} |ch, f(ch)\rangle, \tag{1.1}$$

where c is an element of G selected uniformly at random.

2. *Compute the Fourier transform of the "coset" state* (1.1), *resulting in*

$$\sum_{\rho \in \hat{G}} \sqrt{\frac{1}{|G|}} \sqrt{\frac{1}{|H|}} \sum_{h \in H} \rho(ch) |\rho\rangle,$$

where $\hat{G}$ denotes the set of homomorphisms $\{\rho : G \to \mathbb{C}\}$.

3. *Measure the first register and observe a homomorphism ρ.*

A key fact about this procedure is that the resulting distribution over ρ is independent of the coset cH arising after the first stage (as the support of the first register in (1.1)). Thus, repetitions of this experiment result in the same distribution over $\hat{G}$. We note that by the principle of delayed measurement (see, e.g., [21]), measuring the second register in the first step can in fact be delayed until the end of the experiment.

It is well known that an efficient solution to the HSP for the symmetric group S_n gives, in particular, an efficient quantum algorithm for graph isomorphism. It is also known how to efficiently compute the Fourier transform over many nonabelian groups, most notably over S_n [2]. This article provides the first general understanding of the HSP over nonabelian groups: We study a natural generalization of Experiment 1.1 to nonabelian groups and explicitly describe the resulting measurement distribution. Specifically, we study the following experiment.

EXPERIMENT 1.2.

1. *Prepare the state $\sum_{g \in G} |g, f(g)\rangle$ and measure the second register $f(g)$. The resulting state is $\sum_{h \in H} |ch, f(ch)\rangle$, where c is an element of G selected uniformly at random. As above, this state is supported on a left coset cH of H.*

2. *Let $\hat{G}$ denote the set of irreducible representations of G and, for each $\rho \in \hat{G}$, fix a basis for the space on which ρ acts. Let d_ρ denote the dimension of ρ. Compute the Fourier transform of the "coset" state, resulting in*

$$\sum_{\rho \in \hat{G}} \sum_{i}^{d_\rho} \sum_{j}^{d_\rho} \sqrt{\frac{d_\rho}{|G|}} \sqrt{\frac{1}{|H|}} \left(\sum_{h \in H} \rho(ch) \right)_{i,j} |\rho, i, j\rangle.$$

3. *Measure the first register and observe a representation ρ.*

A brief discussion of the representation theory of finite groups and the associated Fourier transform appears in section 2. As before, we wish the resulting distribution to be *independent* of the actual coset cH (and so depend only on the subgroup H). This is guaranteed by measuring only the name of the representation ρ and leaving the matrix indices (the values i and j) unobserved. The question we study is whether this procedure retains enough information to determine H or, more precisely, whether $O(\log(|G|))$ samples of this distribution are enough to determine H with high probability. Our analysis of Experiment 1.2 depends on the following theorem, which describes the distribution resulting from the measurements in the above experiment.

THEOREM 1.2. *Let H be a subgroup of a group G and let ρ be an irreducible representation of G with dimension d_ρ. Let $R_H(\rho)$ denote the number of times that the trivial representation appears in ρ when decomposed as a representation of H, and*

let $I_H(\rho)$ denote the number of times that ρ appears in the permutation representation of G on the left cosets of H. Then the probability of measuring the representation ρ in Experiment 1.2 *when H is the hidden subgroup is*

$$\frac{d_\rho \cdot R_H(\rho) \cdot |H|}{|G|} = \frac{d_\rho \cdot I_H(\rho) \cdot |H|}{|G|}. \tag{1.2}$$

We first apply this to obtain the following positive result.

THEOREM 1.3. *Let H be an arbitrary subgroup of G, and let H^G be the largest subgroup of H that is normal in G. With high probability, H^G is uniquely determined by observing $m = O(\log |G|)$ independent trials of Experiment* 1.2 *when H is the hidden subgroup.*

When H is normal in G, $H = H^G$, so that this algorithm determines H with high probability. In fact, we shall see that if $\rho_1, \ldots, \rho_m$ are the representations sampled by m independent runs of Experiment 1.2 and m is sufficiently large, then $H^G = \cap_i \ker \rho_i$ with high probability. (Here $\ker \rho$ denotes the kernel of the representation ρ.)

Our reconstruction result applies to any normal subgroup H of any group G without reference to the specific way that the representations or the group elements are expressed. We proceed at this level of abstraction because there is no known canonical concise presentation for the representations (or, indeed, the elements) of a finite group G. In the same vein, there is no general method for computing the Fourier transform over an arbitrary group. Thus, while we cannot give a unified algorithm for computing the Fourier transform or a set of generators for a hidden subgroup, this does yield an algorithm for any group which admits (i) a succinct representation over which the Fourier transform can be computed, and (ii) an efficient algorithm for computing the intersection of a (polynomial-size) family of representation kernels. See section 6, where the above approach is applied to solve the HSP for Hamiltonian groups, where all subgroups are normal.

Note that it is known [8] that the HSP has polynomial (in $\log |G|$) *query* complexity for any subgroup, though the only known algorithm which achieves this uses an exponential number of quantum measurements and, hence, does not give rise to an efficient quantum algorithm for the HSP.

A corollary of Theorem 1.2 is that conjugate subgroups H_1 and H_2 (where $H_2 = gH_1g^{-1}$ for some $g \in G$) produce exactly the same distribution over ρ and hence cannot be distinguished by this process. In particular, the HSP cannot be solved by Experiment 1.2 for a group G with two distinct conjugate subgroups H_1, H_2; the symmetric group S_n is such a group.

In light of this, one may ask whether Experiment 1.2 can distinguish between a coset cH of a nontrivial subgroup H and a coset $cH_e = \{c\}$ of the trivial subgroup $H_e = \{e\}$, as even this would be enough for solving graph isomorphism. However, even for this weaker problem we show (in section 5) the following.

THEOREM 1.4. *For the symmetric group S_n, there is a subgroup H_n so that Experiment* 1.2 *does not distinguish (even information-theoretically) the case that the hidden subgroup is the trivial subgroup from the case that the hidden subgroup is H_n. (Specifically, the distributions induced on ρ in these two cases have exponentially small distance in total variation.)*

1.1. Related work. The HSP plays a central role in most known quantum algorithms. Simon's algorithm [28] implicitly involves distinguishing the trivial subgroup from an order 2 subgroup over the group $\mathbb{Z}_2^n$. Furthermore, he has shown that a classical probabilistic oracle machine would require exponentially many oracle queries

to successfully distinguish the two cases with probability greater than 1/2. Shor [27] then gave efficient algorithms for integer factorization and the discrete log problem. In addition to solving a special case of the HSP, he also solved specific cases where the size of the underlying group is not fully known. Other generalizations have been studied by Boneh and Lipton [3], focusing on cases when a periodic function is not fixed on a coset, and Hales and Hallgren [11, 12], who generalized the results for the case when the underlying abelian group is unknown.

The efficient algorithm for the abelian HSP using the Fourier transform is well known. Other methods have been applied to this same problem by Mosca and Ekert [19]. Related problems have been studied by Kitaev [17], who gave an algorithm using eigenvalue estimation for the abelian stabilizer problem, and Hallgren [13], who gave polynomial-time quantum algorithms for Pell's equation and the principal ideal problem.

As for computing the Fourier transform, Kitaev showed how to efficiently compute the Fourier transform over any abelian group. The fastest currently known (quantum) algorithm for computing the Fourier transform over abelian groups was given by Hales and Hallgren [12]. Shallow parallel circuits for approximating the Fourier transform have been given by Cleve and Watrous in [4]. Beals [2] showed how to efficiently compute the Fourier transform over the symmetric group S_n.

For general groups, Ettinger, Høyer, and Knill [8] have shown that the HSP has polynomial query complexity, giving an algorithm that makes an exponential number of measurements. On the other hand, if one considers arbitrary functions rather than those that arise from HSPs, Aaronson [1] shows that it is not possible to distinguish a 1-1 function from a 2-1 function, even with a quantum algorithm. Several specific nonabelian groups have been studied in the context of the HSP. Ettinger and Høyer [7] give a solution for the HSP over the (nonabelian) dihedral group D_n using polynomially many measurements and exponential (classical) time. Rötteler and Beth [24] and Püschel, Rötteler, and Beth [22] have shown similar results for other specific classes of nonabelian groups. Ivanyos, Mangniez, and Santha [16] have shown how to solve certain nonabelian HSP instances using a reduction to the abelian case.

Grigni et al. [10] independently showed that measuring the representation is not enough for graph isomorphism, and they give stronger negative results. They establish the same bounds even when the row of the representation (i.e., i in Experiment 1.2 above) is measured, and similar bounds if the column (j) is measured, under the assumption that random bases are selected for each representation. They also show that the problem can be solved when the intersection of the normalizers of all subgroups of G is large. Other impossibility results have been given by Ettinger and Høyer [6, 5], determining whether *any* measurement can distinguish certain subgroup states.

2. Representation theory background. To define the Fourier transform (over a general group) we require the basic elements of representation theory, defined briefly below. For complete accounts, consult the books of Serre [26] or Harris and Fulton [15]. Throughout, we let $\mathbb{I}_d$ denote the $d \times d$ identity matrix, dropping the subscript when it can be inferred from context.

Linear representations. A representation ρ of a finite group G is a homomorphism $\rho : G \to GL(V)$, where V is a (finite-dimensional) vector space over $\mathbb{C}$ and $\mathrm{GL}(V)$ denotes the group of invertible linear operators on V. Fixing a basis for V, each $\rho(g)$ may be realized as a $d \times d$ matrix over $\mathbb{C}$, where d is the dimension of V. As ρ is a homomorphism, for any $g, h \in G$, $\rho(gh) = \rho(g)\rho(h)$ (this second product being matrix multiplication). The *dimension* d_ρ of the

representation ρ is d, the dimension of V.

A representation provides a means for investigating a group by homomorphically mapping it into a family of matrices. With this realization, the group operation is matrix multiplication, and tools from linear algebra can be applied to study the group. We shall be concerned with complex-valued functions on a group G; the representations of the group are relevant to this study, as they give rise to the natural Fourier transform in this nonabelian setting.

If $\rho : G \to \mathrm{GL}(V)$ is a representation and there is an inner product $\langle \cdot \,|\, \cdot \rangle$ defined on V, it is always possible to define a new inner product on V so that each $\rho(g)$ is unitary; we will always work under this assumption. In particular, we shall always assume an orthonormal basis for V in which the matrices corresponding to $\rho(g)$ are unitary. We let $\rho(g)_{ij}$ denote the i,jth entry of the matrix for $\rho(g)$ in this fixed basis. (See, e.g., [26] for more discussion.)

We say that two representations $\rho_1 : G \to \mathrm{GL}(V)$ and $\rho_2 : G \to \mathrm{GL}(W)$ of a group G are *isomorphic* when there is a linear isomorphism of the two vector spaces $\phi : V \to W$ so that for all $g \in G$ the diagram

$$\begin{array}{ccc} V & \xrightarrow{\rho_1(g)} & V \\ {\scriptstyle\phi}\downarrow & & {\scriptstyle\phi}\downarrow \\ W & \xrightarrow{\rho_2(g)} & W \end{array}$$

commutes. That is, for all $g \in G$, $\phi\rho_1(g) = \rho_2(g)\phi$. In this case, we write $\rho_1 \cong \rho_2$. Up to isomorphism, a finite group has a finite number of irreducible representations; we let $\hat{G}$ denote this collection (of representations).

Irreducibility. We say that a subspace $W \subset V$ is an *invariant* subspace of a representation $\rho : G \to \mathrm{GL}(V)$ if $\rho(g)W \subseteq W$ for all $g \in G$. The zero subspace and the subspace V are always invariant. If no nonzero proper subspaces are invariant, the representation is said to be *irreducible*.

Decomposition and reducibility. When a representation *does* have a nonzero proper invariant subspace $V_1 \subsetneq V$, it is always possible to find a complementary subspace V_2 (so that $V = V_1 \oplus V_2$) that is also invariant. Since V_1 is invariant, for each $g \in G$, $\rho(g)$ defines a linear map $\rho_1(g)$ from V_1 to V_1 by restriction, and it is not hard to see that $\rho_1 : G \to \mathrm{GL}(V_1)$ is in fact a representation. Similarly, define $\rho_2(g)$ to be $\rho(g)$ restricted to V_2. As $V = V_1 \oplus V_2$, the linear map $\rho(g)$ is completely determined by $\rho_1(g)$ and $\rho_2(g)$, and in this case we write $\rho = \rho_1 \oplus \rho_2$. Repeating this process, any representation ρ may be written $\rho = \rho_1 \oplus \rho_2 \oplus \cdots \oplus \rho_k$, where each ρ_i is irreducible. In particular, there is a basis in which every matrix $\rho(g)$ is block diagonal, the ith block corresponding to the ith representation in the decomposition. While this decomposition is not, in general, unique, the *number* of times a given irreducible representation appears in this decomposition (up to isomorphism) depends only on the original representation ρ.

Characters. The *character* $\chi_\rho : G \to \mathbb{C}$ of a representation ρ is defined by $\chi_\rho(g) = \mathbf{tr}\,(\rho(g))$, where $\mathbf{tr}\,(\cdot)$ denotes the trace. This function is basis independent and, as it turns out, completely determines the representation ρ. Elementary properties of trace imply that characters are in fact *class* functions, depending

only on the conjugacy class of their argument. (Specifically, for every g and h we have $\chi_\rho(hgh^{-1}) = \chi_\rho(g)$.)

Orthogonality. For two complex-valued functions f_1 and f_2 on a group G, there is a natural inner product $\langle f_1 \,|\, f_2 \rangle_G$ given by $\frac{1}{|G|}\sum_g f_1(g) f_2(g)^*$. The matrix entries of the representations of a group G are orthogonal according to this inner product: let ρ and σ be two irreducible representations of G; then

$$\langle \rho(\cdot)_{ij} \,|\, \sigma(\cdot)_{kl} \rangle_G = \begin{cases} 0 & \text{if } \rho \ncong \sigma, \\ \delta_{ik}\delta_{jl} & \text{if } \rho = \sigma. \end{cases}$$

(It is assumed here that when $\rho = \sigma$, the same basis has been selected for each.) An immediate consequence is that if χ_σ is the character of a representation σ and χ_{ρ_i} is the character of an irreducible representation ρ_i, the inner product $\langle \chi_\sigma \,|\, \chi_{\rho_i} \rangle_G$ is precisely the number of times the representation ρ_i appears in the decomposition of σ. Note that since each ρ_i is unitary, we may write

$$\langle \chi_\rho \,|\, \chi_{\rho_i} \rangle_G = \frac{1}{|G|} \sum_{g \in G} \chi_\rho(g) \chi_{\rho_i}(g^{-1}).$$

Orthogonality of the second kind. Let C be a conjugacy class of G. As mentioned before, any character χ_ρ is fixed on any conjugacy class C; we denote this value by $\chi_\rho(C)$. It holds that

$$\sum_{\rho \in \hat{G}} |\chi_\rho(C)|^2 = \frac{|G|}{|C|}. \tag{2.1}$$

This is a special case of a more general principle, and we refer the interested reader to Sagan's excellent book [25, Theorem 1.10.3].

Restriction. A representation ρ of a group G is also automatically a representation of any subgroup H. We refer to this *restricted* representation on H as $\mathrm{Res}_H \rho$. Note that even representations that are irreducible over G may be reducible when restricted to H.

Induction. There is a dual notion, that of *induction*, whereby a representation of a subgroup $H < G$ may be induced to a representation of the whole group G. We delay discussion of this to section 3.1.

We let $\ker(\rho) = \{g \in G : \rho(g) = \mathbb{I}\}$ denote the kernel of a representation ρ. As a representation ρ is a homomorphism, $\ker(\rho)$ is always a normal subgroup of G. In fact, we shall see that any normal subgroup H of G can be written $\cap_{\rho \in I} \ker(\rho)$ for some set I of irreducible representations.

As mentioned above, any representation ρ of G may be decomposed into a direct sum of irreducible representations. In fact, reiterating the comments above, if $\rho_1, \ldots, \rho_k$ are the irreducible representations of G and χ_σ is the character of the representation σ, the value

$$n_i = \langle \chi_\sigma \,|\, \chi_{\rho_i} \rangle_G$$

is precisely the number of times the irreducible representation ρ_i appears in the decomposition of the representation σ into irreducible representations. Specifically, after a unitary change of basis, the matrices $\sigma(g)$ are block diagonal, consisting of n_1

copies of $\rho_1(g)$, followed by n_2 copies of $\rho_2(g)$, etc. We denote this state of affairs by $\sigma = n_1\rho_1 \oplus \cdots \oplus n_k\rho_k$.

There are two representations which shall play a central role in our discussion:

The trivial representation. The trivial representation **1** maps every group element $g \in G$ to the 1×1 identity matrix $\mathbb{I}$. Recalling the orthogonality relations above, the function $g \mapsto 1$ is orthogonal to $\rho(\cdot)_{ij}$ for any nontrivial irreducible representation ρ; this results in the identity

$$\sum_g \rho(g) = 0 \cdot \mathbb{I}, \tag{2.2}$$

which we record in anticipation of the proof of Theorem 1.2.

The regular representation. Fix a vector space V with an orthonormal basis consisting of vectors e_g, one for every element $g \in G$. The regular representation $\mathrm{reg}_G : G \to \mathrm{GL}(V)$ is defined by $\mathrm{reg}_G(g) : e_x \mapsto e_{gx}$ for any $x \in G$. V has dimension $|G|$ and, with the basis above, $\mathrm{reg}_G(g)$ is a permutation matrix for any $g \in G$.

An interesting fact about the regular representation is that it contains every irreducible representation of G. In particular, if $\rho_1, \ldots, \rho_k$ are the irreducible representations of G with dimensions $d_{\rho_1}, \ldots, d_{\rho_k}$, then

$$\mathrm{reg}_G = d_{\rho_1}\rho_1 \oplus \cdots \oplus d_{\rho_k}\rho_k,$$

so that the regular representation contains each irreducible representation ρ exactly d_ρ times. Counting dimensions yields an important relation between the dimensions d_ρ and the order of the group:

$$|G| = \sum_{\rho \in \hat{G}} d_\rho^2. \tag{2.3}$$

The main tool in quantum polynomial-time algorithms is the Fourier transform. When G is nonabelian, this takes the form described below.

DEFINITION 2.1. *Let $f : G \to \mathbb{C}$. The* Fourier transform of f at the irreducible representation ρ, *denoted $\hat{f}(\rho)$, is the $d_\rho \times d_\rho$ matrix*

$$\hat{f}(\rho) = \sqrt{\frac{d_\rho}{|G|}} \sum_{g \in G} f(g)\rho(g).$$

We refer to the collection of matrices $\langle \hat{f}(\rho) \rangle_{\rho \in \hat{G}}$ as the *Fourier transform* of f. Thus f is mapped into $|\hat{G}|$ matrices of varying dimensions. The total number of entries in these matrices is $\sum d_\rho^2 = |G|$, by (2.3) above. The Fourier transform is linear in f; with the constants used above ($\sqrt{d_\rho / |G|}$) it is in fact unitary, taking the $|G|$ complex numbers $\langle f(g) \rangle_{g \in G}$ to $|G|$ complex numbers organized into matrices.

A familiar case in computer science is when the group is cyclic of order n. Then the linear transformation (i.e., the Fourier transform) is a Vandermonde matrix over the nth roots of unity and the representations are all one-dimensional.

In the quantum setting we identify the state $\sum_{g \in G} f_g |g\rangle$ with the function $f : G \to \mathbb{C}$ defined by $f(g) = f_g$. In this notation, $\sum_{g \in G} f(g)|g\rangle$ is mapped under the Fourier transform to

$$\sum_{\rho \in \hat{G}} \sum_{1 \le i,j \le d_\rho} \hat{f}(\rho)_{i,j} |\rho, i, j\rangle.$$

We remind the reader that $\hat{f}(\rho)_{i,j}$ is a complex number and that when the first portion of this triple is measured, we observe $\rho \in \hat{G}$ with probability

$$\sum_{1 \le i,j \le d_\rho} \left| \hat{f}(\rho)_{i,j} \right|^2 = \left\| \hat{f}(\rho) \right\|_2^2,$$

where $\|A\|_2$ is the natural norm given by $\|A\|_2^2 = \mathbf{tr}\ A^* A$.

Let f be the indicator function of a left coset of H in G; i.e., for some $c \in G$,

$$f(g) = \begin{cases} \frac{1}{\sqrt{|H|}} & \text{if } g \in cH, \\ 0 & \text{otherwise.} \end{cases}$$

Our goal is to understand the Fourier transform of f, as this determines the probability of observing ρ. Our choice to measure only the representation ρ (and not the matrix indices) depends on the following key fact about the Fourier transform, also relevant to the abelian solution.

CLAIM 2.1. *The probability of observing ρ is independent of the coset.*

Proof. We have

$$\hat{f}(\rho) = \sqrt{\frac{d_\rho}{|G|}\frac{1}{|H|}} \sum_{h \in H} \rho(ch) = \sqrt{\frac{d_\rho}{|G|}\frac{1}{|H|}} \rho(c) \sum_{h \in H} \rho(h)$$

and, since $\rho(c)$ is a unitary matrix,

$$\left\| \hat{f}(\rho) \right\|_2^2 = \frac{d_\rho}{|G|}\frac{1}{|H|} \left\| \rho(c) \sum_{h \in H} \rho(h) \right\|_2^2 = \frac{d_\rho}{|G|}\frac{1}{|H|} \left\| \sum_{h \in H} \rho(h) \right\|_2^2. \qquad \square$$

Given this we may assume, without loss of generality, that our function f is $1/\sqrt{|H|}$ on the subgroup H itself, and zero elsewhere.

3. The probability of measuring ρ. The primary question is that of the probability of observing a given ρ in Experiment 1.2. We have seen that this is determined by the linear operator $\sum_{h \in H} \rho(h)$ and begin by showing that $\frac{1}{|H|} \sum_h \rho(h)$ is a projection.

LEMMA 3.1. *Let ρ be an irreducible representation of G. For every subgroup $H \le G$, $\frac{1}{|H|} \sum_{h \in H} \rho(h)$ is a projection operator.*

With the right basis, then, $\frac{1}{|H|} \sum_{h \in H} \rho(h)$ is diagonal, each diagonal entry being either one or zero. The probability of observing a particular representation ρ is then proportional to the rank of $\hat{f}(\rho)$.

Proof of Lemma 3.1. Given an irreducible representation ρ of G, we are interested in the sum of the matrices $\rho(h)$ over all $h \in H$. Since we evaluate only ρ on H, we may instead consider $\mathrm{Res}_H \rho$ without changing anything. As mentioned before, though ρ is irreducible (over G), $\mathrm{Res}_H \rho$ may *not* be irreducible over H. We may, however, decompose $\mathrm{Res}_H \rho$ into irreducible representations over H, writing $\mathrm{Res}_H \rho = \sigma_1 \oplus \cdots \oplus \sigma_t$ for a sequence σ_i of (possibly repeating) irreducible representations of H. In an appropriate basis $\sum_{h \in H} \rho(h)$ is then comprised of blocks, one corresponding to

each σ_i. In particular, the matrix $\sum_{h\in H}\rho(h)$ is

$$U\begin{bmatrix} \sum_{h\in H}\sigma_1(h) & 0 & \cdots & 0 \\ 0 & \sum_{h\in H}\sigma_2(h) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{h\in H}\sigma_t(h) \end{bmatrix} U^\dagger \tag{3.1}$$

for some unitary transformation U and some irreducible representations σ_i of H (with possible repetitions). (Here † denotes conjugate transpose.) Recalling (2.2), each sum appearing on the diagonal is nonzero only when the irreducible representation is trivial, in which case it is $|H|$. □

As in the previous section, we let $f : G \to \mathbb{C}$ be the function $f(g) = \frac{1}{\sqrt{|H|}}$ for $g \in H$, and 0 otherwise. Then the probability of observing ρ in Experiment 1.2 is

$$\left\|(\hat{f}(\rho))\right\|_2^2 = \frac{d_\rho}{|G|}\frac{1}{|H|}\left\|\sum_{h\in H}\rho(h)\right\|_2^2 = \frac{d_\rho}{|G|}\frac{1}{|H|}|H|^2\left\langle \chi_\rho \,|\, \chi_{\mathbf{1}}\right\rangle_H$$
$$= \frac{|H|}{|G|}d_\rho\left\langle \chi_\rho \,|\, \chi_{\mathbf{1}}\right\rangle_H.$$

We record this result in the following theorem.

THEOREM 3.2. *For every subgroup $H \le G$, the probability of measuring ρ in Experiment* 1.2, *with hidden subgroup H, is*

$$\left\|\hat{f}(\rho)\right\|_2^2 = \frac{|H|}{|G|}d_\rho\left\langle \chi_\rho \,|\, \chi_{\mathbf{1}}\right\rangle_H.$$

Observe that this establishes the first part of Theorem 1.2, that the measurement probability equals the first expression of (1.2). Another consequence of the theorem is that the probability of observing a representation ρ depends only on the character of ρ restricted to H. As the characters are class functions, conjugate subgroups (two subgroups H_1 and H_2 are conjugate if $H_1 = gH_2g^{-1}$ for some $g \in G$) produce exactly the same distribution over ρ; this rules out using the paradigm of Experiment 1.2 with representations names alone to solve the HSP for any group containing a nonnormal subgroup.

3.1. Induced representations. We have discussed the restriction of a representation ρ of G to a subgroup H of G. There is a dual operation, *induction*, which extends to all of G a representation ρ defined on a subgroup H. We will only need to work with the representation induced from the trivial representation on H.

Let $G/H \stackrel{\text{def}}{=} \{\alpha_1, \ldots, \alpha_t\}$ be a canonical transversal for H, so that G can be written as the disjoint union $\alpha_1 H \cup \cdots \cup \alpha_t H$. Then the *induced* representation $\mathrm{Ind}_H^G\mathbf{1} : G \to \mathrm{GL}(W)$ is defined over the vector space W that has one basis vector $e_{[\alpha_i]}$ for each coset $\alpha_i H$. It is defined by linearly extending the rule

$$\mathrm{Ind}_H^G\mathbf{1}(g) : e_{[\alpha_i]} \to e_{[\alpha_j]},$$

where $\alpha_j H$ is the coset containing $g\alpha_i$. Observe that this representation is a permutation representation. As suggested by the notation, selection of a different transversal results in an isomorphic representation.

Example 3.1. $\mathrm{Ind}_{\{\mathbf{id}\}}^G\mathbf{1} \cong \mathrm{reg}_G$.

We now invoke a standard representation-theoretic result to obtain Theorem 1.2.

LEMMA 3.3 (a special case of Frobenius reciprocity; see [15, section 3.20]). *Let $H < G$ and let $\rho : G \to GL(V)$ be an irreducible representation of G. Then*

$$\langle \chi_{\mathbf{1}} \mid \chi_\rho \rangle_H = \left\langle \chi_{\mathrm{Ind}_H^G \mathbf{1}} \,\middle|\, \chi_\rho \right\rangle_G .$$

Combining this with Theorem 3.2 establishes Theorem 1.2.

Proof of Theorem 1.2. Theorem 3.2 asserts that the probability of measuring the representation ρ in Experiment 1.2 is $\frac{|H|}{|G|} d_\rho \langle \chi_{\mathbf{1}} \mid \chi_\rho \rangle_H$. By reciprocity, the number of times that the trivial representation of H appears in $\mathrm{Res}_H \rho$ is the same as the number of times that ρ appears in $\mathrm{Ind}_H^G \mathbf{1}$, that is,

$$\langle \chi_{\mathbf{1}} \mid \chi_\rho \rangle_H = \left\langle \chi_{\mathrm{Ind}_H^G \mathbf{1}} \,\middle|\, \chi_\rho \right\rangle_G .$$

Theorem 1.2 follows. □

4. A positive result: Normal subgroups and the core of H. In this section we show that $O(\log |G|)$ queries suffice to reconstruct any normal subgroup of G. In general, we show that for any subgroup H of G, the algorithm below outputs H^G, the *core of* H, which is the largest subgroup of H that is normal in G. As the product $H_1 H_2$ of two normal subgroups is again a normal subgroup of G, the core is well-defined. (In fact, the core is precisely $\bigcap_{g \in G} g H g^{-1}$.) The algorithm we study is the following.

ALGORITHM 4.1. *H is an arbitrary unknown subgroup of G; we are provided an efficiently computable function $f : G \to S$, which is constant on (left) cosets of H and takes distinct values on distinct cosets.*

1. *For $i = 1, \ldots, s = 4 \log_2 |G|$, run Experiment* 1.2 *and measure an irreducible representation, $\sigma_i \in \hat{G}$.*
2. *Let $N_i = \bigcap_{j=1}^{i} \ker \sigma_j$.*
3. *Output $N = N_s$.*

Recall that each $\ker \sigma_i$ is a normal subgroup of G, so that the resulting subgroup $N = N_s$ is normal. We will show that Algorithm 4.1 converges quickly to H^G with high probability in Theorem 4.3. We reduce the proof of this theorem to two lemmas, described in the following section. Two different sets of proofs of these lemmas are then presented in sections 4.2 and 4.3, one from the perspective of restricted representations, and one from the perspective of induced representations. The proof presented in terms of induction shows that the theorem is a consequence of the standard proof of the Mackey irreducibility criterion; for readers already acquainted with the criterion this approach may be more mnemonic than the elementary approach by restriction.

4.1. The general structure. As discussed above, Theorem 4.3 is a consequence of the following two lemmas.

LEMMA 4.1. *If the irreducible representation σ can be sampled by Experiment* 1.2 *(i.e., has nonzero probability), then $H^G \subseteq \ker(\sigma)$.*

This shows, in particular, that $H^G \trianglelefteq N_s \trianglelefteq \cdots \trianglelefteq N_1 \trianglelefteq N_0 = G$.

LEMMA 4.2. *For any subgroup $H \leq G$, if $N_i \not\subseteq H$, then* $\Pr[N_{i+1} = N_i] \leq 1/2$.

Before discussing the proofs of these lemmas, we show that together they imply Theorem 4.3. Observe that for a representation ρ of G, if $\ker \rho \subset H$, then we must have $\ker \rho \subset H^G$, as $\ker \rho$ is normal.

THEOREM 4.3. *Algorithm* 4.1 *returns H^G with probability at least* $1 - 2e^{-\log_2 |G|/8}$.

Proof of Theorem 4.3. Let $\mathcal{D}_H$ denote the probability distribution over irreducible representations induced by Experiment 1.2. We now apply a standard martingale bound (see [20]) to prove the theorem (based on Lemmas 4.2 and 4.1). Let $\sigma_1, \ldots, \sigma_k$ be independent random variables distributed according to $\mathcal{D}_H$ with $k = 4\log_2 |G|$. Our goal is to show that

$$\Pr[N_s \neq H^G] \leq 2e^{-\log_2|G|/8}.$$

For each $i \in \{1, \ldots, k\}$, let X_i be the indicator random variable taking value 1 if $N_i \subseteq H$ or $N_{i+1} \neq N_i$, and zero otherwise. The random variables $X_1, \ldots, X_k$ are not necessarily independent, but by Lemma 4.2, $\Pr[X_i = 0 \,|\, X_1, \ldots, X_{i-1}] \leq \frac{1}{2}$, and we may apply a martingale bound. As the variables take values in the set $\{0, 1\}$, the sum $\sum_i X_i$ satisfies the Lipschitz condition (with constant 1), and we can apply Azuma's inequality to conclude that $\sum_i X_i$ is unlikely to deviate far from its expected value, which is at least $\frac{k}{2}$. In particular, we have $\Pr\left[|\sum_i X_i - \frac{k}{2}| \geq \lambda\right] \leq 2e^{-\lambda^2/2k}$, so with $\lambda = \log_2(|G|)$ we have $\Pr\left[\sum_{i=0}^{k-1} X_i \leq \log_2(|G|)\right] \leq 2e^{-\log_2(|G|)/8}$.

Therefore, with probability at least $1 - 2e^{-\log_2(|G|)/8}$ we have $N_s \subseteq H$. As N_s is normal in G, it must be the case that $N_s \subseteq H^G$. From Lemma 4.1, $H^G \subseteq N_s$; hence $N_s = H^G$, and the algorithm converges to the correct subgroup. ∎

4.2. Restricted representations. We begin by proving Lemmas 4.1 and 4.2 from the perspective of restricted representations.

By Claim 2.1, we may assume that f is distributed over H itself without loss of generality. In this case, Lemma 3.1 implies that, up to a scalar multiple, each $\hat{f}(\sigma)$ is a projection. We begin by showing that when the subgroup H is *normal*, $\hat{f}(\sigma)$ is in fact a multiple of the identity, and is nonzero precisely when H is in the kernel of σ.

LEMMA 4.4. *Let $H \trianglelefteq G$, $\rho \in \hat{G}$ have dimension d_ρ, and let $f : G \to \mathbb{C}$ be the function*

$$f(g) = \begin{cases} \frac{1}{\sqrt{H}} & \text{if } g \in H, \\ 0 & \text{otherwise.} \end{cases}$$

Then $\hat{f}(\rho) = \lambda \cdot \mathbb{I}$, where

$$\lambda = \begin{cases} \sqrt{\frac{|H|}{|G|} d_\rho} & \text{if } H \subseteq \ker \rho, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. The lemma follows from an application of Schur's lemma (see, e.g., [26]). If $H \subseteq \ker \rho$, then the lemma follows from the discussion in the previous section.

Suppose $H \not\subseteq \ker \rho$. We will show that $\hat{f}(\rho)$ must be the zero map. By Lemma 3.1 we can decompose V as $W_f \bigoplus W_a$, where $\frac{1}{|H|}\sum_{h \in H} \rho(h)$ pointwise fixes W_f and annihilates W_a. Our goal is to see that $W_f = \{0\}$. Observe that since $H \not\subseteq \ker \rho$, there is some $h_0 \in H$ for which $\rho(h_0)$ is not the identity operator and, considering that each $\rho(h)$ is unitary, their average cannot be the identity operator. (Specifically, consider a unit vector v for which $\rho(h_0)v \neq v$; note now that the average over h of $\rho(h)v$ can have unit length only if for all h_1, $\rho(h_0)v = \rho(h_1)v \neq v$.) Hence $W_f \neq V$.

Assume that $W_f \neq \{0\}$. Since ρ is irreducible over G and $W_f \neq V$, there is a vector $w'_f \in W_f$ and $g \in G$ such that $\rho(g)w'_f \notin W_f$; we may write $\rho(g)w'_f = w_f + w_a$,

with $w_a \in W_a$, $w_f \in W_f$, and $w_a \neq 0$. As $\frac{1}{|H|}\sum_{h\in H}\rho(h)$ pointwise fixes W_f we have

$$w_f + w_a = \rho(g)w'_f = \rho(g)\frac{1}{|H|}\sum_{h\in H}\rho(h)w'_f$$
$$\stackrel{*}{=} \frac{1}{|H|}\sum_{h\in H}\rho(h)\rho(g)w'_f = \frac{1}{|H|}\sum_{h\in H}\rho(h)(w_f + w_a) = w_f,$$

where equality $\stackrel{*}{=}$ follows since H is normal in G. This is a contradiction; hence $W_f = \{0\}$, as desired. □

We will now prove Lemma 4.1, which states that if the irreducible representation σ can be sampled by Experiment 1.2, then $H^G \subseteq \ker(\sigma)$.

Proof of Lemma 4.1. Let C be a set of coset representatives for H^G in H. We have

$$\sum_{h\in H}\rho(h) = \Big(\sum_{c\in C}\rho(c)\Big)\Big(\sum_{h\in H^G}\rho(h)\Big),$$

so by Lemma 4.4 we only observe ρ if $\sum_{h\in H^G}\rho(h)$ is a multiple of the identity (in which case $H^G \subseteq \ker\rho$). □

Before proving Lemma 4.2, which is for general subgroups, we will show how the statement can be proved for normal subgroups.

LEMMA 4.5. *If $H \trianglelefteq G$ and $N_i \neq H$, then* $\Pr[N_{i+1} = N_i] \leq 1/2$.

Proof. By Lemma 4.4, Theorem 3.2, and (2.3) we have

$$\Pr[N_i \subseteq \ker\rho_{i+1}] = \sum_{\rho\in\hat{G}N_i\subseteq\ker\rho}\Pr[\text{Observe }\rho]$$
$$= \sum_{\rho\in\hat{G}N_i\subseteq\ker\rho}\frac{|H|}{|G|}d_\rho^2$$
$$= \frac{|H|}{|G|}\sum_{\rho\in\widehat{G/N_i}}d_\rho^2 = \frac{|H|}{|G|}\frac{|G|}{|N_i|} \leq \frac{1}{2},$$

where changing the sum follows from the fact that representations of G that map N_i to the identity can be identified with representations of G/N_i. □

We proceed to prove Lemma 4.2, which states that for any $H \leq G$, if $N_i \not\subseteq H$, then $\Pr[N_{i+1} = N_i] \leq 1/2$.

Proof of Lemma 4.2. This proof is due to Umesh Vazirani. Let $N = \cap_{j=1}^{i}\ker\sigma_j$ be the intersection of the kernels up to step i. For an irreducible representation ρ, let r_ρ be the rank of $\hat{f}(\rho)$, i.e., the number of times the trivial representation of H appears in ρ. When $N \not\subseteq H$, we will show that the probability of N being contained in the kernel of the next representation we measure is at most $1/2$ by showing that

$$\sum_{\rho:N\subseteq\ker\rho}\frac{|H|}{|G|}d_\rho r_\rho \leq \frac{|N\cap H|}{|N|},$$

which is at most $1/2$ when $N \not\subseteq H$. Now, if the hidden subgroup had been HN, Theorem 3.2 would imply

$$\sum_{\rho:\rho\in\hat{G}}\frac{|HN|}{|G|}d_\rho r'_\rho = 1,$$

where r'_ρ is the number of times the trivial representation of HN appears in ρ. Note that $r'_\rho = r_\rho$ when $N \subseteq \ker \rho$, since

$$|H \cap N| \sum_{l \in HN} \rho(l) = \Big(\sum_{h \in H} \rho(h) \Big) \Big(\sum_{n \in N} \rho(n) \Big),$$

and $\rho(n)$ is the identity. Since $|HN| \cdot |H \cap N| = |H| \cdot |N|$, we have that

$$\sum_{\rho : N \subseteq \ker \rho} \frac{|H|}{|G|} d_\rho r_\rho = \sum_{\rho : N \subseteq \ker \rho} \frac{|H|}{|G|} d_\rho r'_\rho \leq \sum_{\rho \in \hat{G}} \frac{|H|}{|G|} d_\rho r'_\rho \leq \frac{|H \cap N|}{|N|},$$

as desired. ∎

4.3. Induced representations. We now reprove these two lemmas from the perspective of induced representations. We begin by computing the kernel of the representation $\mathrm{Ind}_H^G \mathbf{1}$.

LEMMA 4.6. $\ker(\mathrm{Ind}_H^G \mathbf{1}) = H^G$.

Proof. We begin with the forward inclusion. Indeed, if $x \in \ker(\mathrm{Ind}_H^G \mathbf{1})$, then $\mathrm{Ind}_H^G \mathbf{1}(x)$ is the identity mapping, i.e., for every $g \in G$, $\mathrm{Ind}_H^G \mathbf{1}(x) : [gH] \to [gH]$, or equivalently, $[xgH] = [gH]$. In particular, for $g = e$ we get $xH = H$, and therefore $x \in H$. Now, as $\ker(\mathrm{Ind}_H^G \mathbf{1})$ is normal and is contained in H we must have $\ker(\mathrm{Ind}_H^G \mathbf{1}) \subseteq H^G$.

For the reverse inclusion, suppose that $x \in H^G$. Then for any $g \in G$, there is some $x' \in H^G \subseteq H$ such that $xg = gx'$. Therefore, $\mathrm{Ind}_H^G \mathbf{1}(x)[gH] = [xgH] = [gx'H] = [gH]$, and we see that $\mathrm{Ind}_H^G \mathbf{1}(x)$ is the identity mapping. Hence, $x \in \ker(\mathrm{Ind}_H^G \mathbf{1})$. ∎

Now, by Theorem 1.2, any σ that can be sampled by Experiment 1.2 appears in $\mathrm{Ind}_H^G \mathbf{1}$, and we therefore conclude that $H^G \subseteq \ker(\mathrm{Ind}_H^G \mathbf{1}) \subseteq \ker(\sigma)$; Lemma 4.1 follows. This also gives a simple decomposition of $\mathrm{Ind}_H^G \mathbf{1}$ when H is normal.

LEMMA 4.7. *Let $N \trianglelefteq G$. Then* $\mathrm{Ind}_N^G \mathbf{1} = \bigoplus_{\rho \in \hat{G}, N \subseteq \ker(\rho)} d_\rho \rho$.

Proof. Suppose $\mathrm{Ind}_N^G \mathbf{1} = \bigoplus_{\rho \in \hat{G}} n_\rho \rho$. We have

$$n_\rho = \Big\langle \chi_{\mathrm{Ind}_N^G \mathbf{1}} \Big| \chi_\rho \Big\rangle_G \;=\; \langle \chi_{\mathbf{1}} \,|\, \chi_\rho \rangle_N \;=\; \frac{1}{|N|} \sum_{x \in N} \chi_\rho(x) \;=\; d_\rho,$$

where the second equality is by Frobenius reciprocity, and the last one is because $N \subseteq \ker \rho$. Note that $n_\rho = 0$ if $N \not\subset \ker \rho$. ∎

We now prove Lemma 4.2.

Proof of Lemma 4.2. Denote $N = N_i$. For $\rho \in \hat{G}$, let $m_\rho = \langle \chi_\rho \,|\, \chi_{\mathbf{1}} \rangle_H$. We know that

$$\Pr_{\sigma \in \mathcal{D}_H} [N_i \subseteq \ker \sigma] = \frac{|H|}{|G|} \sum_{\substack{\rho \in \hat{G} \\ N_i \subseteq \ker(\rho)}} m_\rho d_\rho,$$

$$\mathrm{Ind}_H^G \mathbf{1} = \bigoplus_{\rho \in \hat{G}} m_\rho \rho,$$

$$\mathrm{Ind}_N^G \mathbf{1} = \bigoplus_{\substack{\rho \in \hat{G} \\ N \subseteq \ker(\rho)}} d_\rho \rho,$$

where the first equation is by Theorem 3.2, the second because $m_\rho = \langle \chi_\rho \,|\, \chi_{\mathbf{1}} \rangle_H = \langle \chi_{\mathrm{Ind}_H^G \mathbf{1}} | \chi_\rho \rangle_G$ by Frobenius reciprocity (Lemma 3.3), and the last one by Lemma 4.7. We observe that

$$\left\langle \chi_{\mathrm{Ind}_H^G \mathbf{1}} \,\middle|\, \chi_{\mathrm{Ind}_N^G \mathbf{1}} \right\rangle_G = \sum_{\substack{\rho \in \hat{G} \\ N \subseteq \ker(\rho)}} d_\rho m_\rho \langle \chi_\rho \,|\, \chi_\rho \rangle_G = \sum_{\substack{\rho \in \hat{G} \\ N \subseteq \ker(\rho)}} d_\rho m_\rho$$

and thus is proportional to the probability that $N_i \subseteq \ker(\sigma_i)$. We complete the proof with an argument similar to that given in Serre [26] for the proof of Mackey's irreducibility criterion. By Frobenius reciprocity,

$$\left\langle \chi_{\mathrm{Ind}_H^G \mathbf{1}} \,\middle|\, \chi_{\mathrm{Ind}_N^G \mathbf{1}} \right\rangle_G = \left\langle \chi_{\mathbf{1}} \,\middle|\, \chi_{\mathrm{Res}_H \ \mathrm{Ind}_N^G \mathbf{1}} \right\rangle_H .$$

Decomposing the restricted induction (see [26, section 7.3]) we have

$$\left\langle \chi_{\mathrm{Ind}_H^G \mathbf{1}} \,\middle|\, \chi_{\mathrm{Ind}_N^G \mathbf{1}} \right\rangle_G = \bigoplus_{g \in H \backslash G / N} \left\langle \chi_{\mathbf{1}} \,\middle|\, \chi_{\mathrm{Ind}_{N_g}^H \mathbf{1}} \right\rangle_H ,$$

where $N_g = H \cap gNg^{-1}$ is a subgroup of H, and g runs over all representatives of the double cosets $H \setminus g/N$ of G. Using Frobenius reciprocity again we see that

$$\begin{aligned} \left\langle \chi_{\mathrm{Ind}_H^G \mathbf{1}} \,\middle|\, \chi_{\mathrm{Ind}_N^G \mathbf{1}} \right\rangle_G &= \bigoplus_{g \in H \backslash G / N} \langle \chi_{\mathbf{1}} \,|\, \chi_{\mathbf{1}} \rangle_{N_g} \\ &= |H \setminus G / N|. \end{aligned}$$

However, N is normal in G. Hence for any $g \in G$, $H \backslash g/N = HNg$. Furthermore, as H is a group and N is normal in G, HN is also a group. Hence, $|H \setminus G/N| = |G|/|HN|$. Thus,

$$\Pr_{\sigma \in \mathcal{D}_H} [N_i \subseteq \ker \sigma] = \frac{|H|}{|G|} \sum_{\substack{\rho \in \hat{G} \\ N_i \subseteq \ker(\rho)}} m_\rho d_\rho = \frac{|H|}{|G|} \frac{|G|}{|HN|} = \frac{|H|}{|HN|},$$

which is at most $1/2$ when $N \not\subseteq H$. ∎

5. A negative result: Determining triviality in S_n. In this section we show that a well-known reduction of graph isomorphism for finding a hidden subgroup over S_n cannot work using Experiment 1.2.

Graph automorphism is the problem of determining whether a graph G has a nontrivial automorphism and is easier than graph isomorphism [18]. A natural special case occurs when the graph G consists of two disjoint connected rigid graphs G_1, G_2 (i.e., $\mathrm{Aut}(G_1) = \mathrm{Aut}(G_2) = \{e\}$). In this case there are two possibilities for the automorphism group of G.

CLAIM 5.1. *Let the graph G be the disjoint union of the two connected rigid graphs G_1 and G_2 and let n denote the number of vertices of G. Then*

1. *if $G_1 \not\approx G_2$, then $\mathrm{Aut}(G) = \{e\}$, and*
2. *if $G_1 \approx G_2$, then $\mathrm{Aut}(G) = \{e, \sigma\}$, where $\sigma \in S_n$ is a permutation with $n/2$ disjoint 2-cycles.*

Proof. For the first part notice that any automorphism maps a connected component onto a connect component. In our case we have two connected components

G_1 and G_2. However, G_1 and G_2 are not isomorphic and have no nontrivial automorphisms.

For the second part, let σ reflect an automorphism between G_1 and G_2. Now, suppose there was another nontrivial automorphism τ. Then $\sigma\tau$ is also an automorphism, and $\sigma\tau$ maps the connected component of G_1 onto G_1, and G_2 onto G_2. As G_1 and G_2 have no nontrivial automorphisms it follows that $\sigma\tau = 1$, $\tau = \sigma^{-1} = \sigma$. □

Thus, if one knows how to solve the HSP for S_n, or if one knows how to distinguish between cosets of a trivial subgroup and the cosets of a nontrivial subgroup, one can give an efficient quantum algorithm for graph automorphism. In particular, one might attempt to reconstruct $H = \mathrm{Aut}(G)$ based on the result of the following experiment.

EXPERIMENT 5.1. *Let G be a graph such that either $\mathrm{Aut}(G) = \{e\}$ or $\mathrm{Aut}(G) = \{e, \sigma\}$.*

1. *Compute $\Sigma_{\pi\in S_n}|\pi, \pi(G)\rangle$ and measure the second register $\pi(G)$. The resulting state is $\Sigma_{h\in H}|ch, f(ch)\rangle$ for some coset cH of H. Furthermore, c is uniformly distributed over G.*
2. *Compute the Fourier transform of the coset state, which is*

$$\sum_{\rho\in\hat{G}} \sqrt{\frac{d_\rho}{|G|}} \sqrt{\frac{1}{|H|}} \Bigg(\sum_{h\in H} \rho(ch) \Bigg)_{i,j} |\rho, i, j\rangle.$$

3. *Measure the first register and observe a representation ρ.*

We show that even for this particular case of graph isomorphism (and graph automorphism) the experiment fails to distinguish nonisomorphic pairs of graphs from isomorphic pairs of graphs; Theorem 1.4 follows.

THEOREM 5.1. *Let G_1 and G_2 be two rigid, connected graphs with n vertices. Let $\mathcal{D}_N(\rho)$ be the probability of sampling ρ in Experiment 5.1 when $G_1 \not\approx G_2$, and let $\mathcal{D}_I(\rho)$ be the probability when $G_1 \approx G_2$. Then $|\mathcal{D}_N - \mathcal{D}_I|_1 \le 2^{-\Omega(n)}$.*

Proof. We present the proof from [10], which simplifies the proof of [14]. When $G_1 \not\approx G_2$, $H = \{e\}$, so $\mathcal{D}_N(\rho) = d_\rho^2/n!$ by Theorem 3.2. When $G_1 \approx G_2$, and G_1 and G_2 are both connected and rigid, $H = \{e, \tau\}$. By Theorem 3.2,

$$\mathcal{D}_I(\rho) \;=\; \frac{|H|}{|G|} d_\rho \left\langle \chi_{\mathbf{1}} \,|\, \chi_\rho \right\rangle_H .$$

The subgroup H has only two elements, e and τ; hence

$$\left\langle \chi_{\mathbf{1}} \,|\, \chi_\rho \right\rangle_H = \frac{1}{2}(\chi_\rho(e) + \chi_\rho(\tau)) = \frac{1}{2}(d_\rho + \chi_\rho(\tau)).$$

That is, $\mathcal{D}_I(\rho) = \frac{d_\rho}{n!}(d_\rho + \chi_\rho(\tau))$, and so

$$\begin{aligned} \sum_\rho |\mathcal{D}_I(\rho) - \mathcal{D}_N(\rho)| &= \frac{1}{n!} \sum_\rho d_\rho |\chi_\rho(\tau)| \\ &\le \frac{1}{n!} \sqrt{\sum_\rho d_\rho^2} \sqrt{\sum_\rho |\chi_\rho(\tau)|^2} = \frac{1}{\sqrt{n!}} \sqrt{\sum_\rho |\chi_\rho(\tau)|^2} \end{aligned}$$

by the Cauchy–Schwarz inequality and (2.3).

By (2.1), $\sum_{\rho\in\hat{G}} |\chi_\rho(\tau)|^2 = |G|/|\{\tau\}|$, where $\{\tau\}$ is the conjugacy class of τ. However, two permutations share the same conjugacy class if and only if they have

the same cycle decomposition. In our case τ has cycle decomposition into $n/2$ pairs. Thus

$$|\{\tau\}| = \frac{\binom{n}{n/2}(\frac{n}{2})!}{2^{n/2}},$$

where $\binom{n}{n/2}$ is the number of possibilities for choosing the first element in each of the $n/2$ pairs, $(\frac{n}{2})!$ is the number of possibilities for arranging the remaining $n/2$ elements in the pairs, and each ordering is counted exactly $2^{n/2}$ times.

Altogether,

$$\sum_{\rho} |\mathcal{D}_I(\rho) - \mathcal{D}_N(\rho)| \leq \frac{1}{\sqrt{n!}}\sqrt{\frac{n!}{|\{\tau\}|}} \;=\; \sqrt{\frac{2^{(n/2)}(n/2)!}{n!}} \;\leq\; 2^{-\Omega(n)},$$

as desired. □

6. Finding hidden subgroups in Hamiltonian groups. A group G is Hamiltonian if all subgroups are normal. In light of Theorem 3.2, a hidden subgroup of a Hamiltonian group G is determined with high probability by $O(\log |G|)$ samples of the distribution induced by Experiment 1.2. In this section we show that for Hamiltonian groups, generators for the hidden subgroup can be computed efficiently from these samples. As the Fourier transform over such groups can be efficiently computed, this gives an efficient quantum algorithm for the HSP over Hamiltonian groups.

All abelian groups are Hamiltonian; the only nonabelian Hamiltonian groups are of the form

$$G \cong \mathbb{Z}_2^k \times B \times Q,$$

where $Q = \{\pm 1, \pm i, \pm j, \pm k\}$ is the quaternion group and B is an abelian group with exponent b coprime with 2. For a detailed description of such groups, see Rotman's excellent book [23].

We begin by briefly discussing the case when G is abelian. If G is simply the cyclic group $\mathbb{Z}_n$, the representations are the functions $\rho_s : z \mapsto \exp(2\pi isz/n)$, and the reconstruction algorithm, when it succeeds, yields a collection $\{\rho_s | s \in S\}$ with the property that $H = \cap_{s\in S} \ker \rho_s$. Observe that $\rho_s(h)\rho_t(h) = \rho_{s+t \bmod n}(h)$ and that $\rho_s(h) = 1$ implies $\rho_{st \bmod n}(h) = 1$ for all $t \in \mathbb{Z}$. Hence $\cap_s \ker \rho_s = \ker \rho_d$, where d is the greatest common divisor of n and the elements in S. Then H is the cyclic subgroup of $\mathbb{Z}_n$ generated by n/d.

In general, an abelian group G is isomorphic to a direct sum $\mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_k}$, and we assume that this decomposition is known. The irreducible representations are the functions $\rho_{s_1,\ldots,s_k}(z_1, \ldots, z_k) = \prod_{j=1}^{k} \exp(2\pi i s_j z_j/n_j)$, and, as above, we begin with a collection $\{\rho_{\vec{s}} \mid \vec{s} = (s_1, \ldots, s_k) \subset S\}$ so that $H = \cap_S \ker \rho_{\vec{s}}$. Then

$$\begin{aligned} (h_1, \ldots, h_k) \in H \quad &\Leftrightarrow \quad \text{for all } \vec{s} \in S,\ \exp\left(\sum_j \frac{2\pi i s_j h_j}{n_j}\right) = 1 \\ &\Leftrightarrow \quad \text{for all } \vec{s} \in S,\ \sum_j s_j q_j h_j \equiv 0 \bmod N, \end{aligned} \tag{6.1}$$

where N is the least common multiple of the n_j, and $q_j = N/n_j$. For convenience, we treat the family of equalities appearing in line (6.1) as a system of equations over

the ring $\mathbb{Z}_N$; then a solution $\vec{h}$ of this system corresponds to the element ($h_1 \bmod n_1, \ldots, h_k \bmod n_k$) of H. Collect these equations together into a matrix R. Though $\mathbb{Z}_N$ may not be a field, it is easy to check that a matrix over $\mathbb{Z}_N$ may be diagonalized in polynomial time with the following two operations:

- for some pair $i \neq j$, swap row (column) i with row (column) j;
- for some pair $i \neq j$, add a multiple of row (column) i to row (column) j.

This results in a system $D \cdot F \cdot \vec{h} = \vec{0}$, where D is diagonal and F is invertible. Any vector $\vec{h}'$ for which $D\vec{h}' = \vec{0}$ may then be transformed into a solution $\vec{h}$ of the original equation and, moreover, if $\vec{h}'$ is selected at random in the null space of D, then the resulting $\vec{h}'$ will give rise to a random element of H. Selection of $O(\log |G|)$ random elements in this way yields a generating set for H with high probability.

Finally, consider a Hamiltonian group of the form $G = \mathbb{Z}_2^k \times B \times Q$. An irreducible representation ρ of G is a tensor product $\zeta \otimes \beta \otimes \kappa$, where $\zeta \in \widehat{\mathbb{Z}_2^k}$, $\beta \in \hat{B}$, and $\kappa \in \hat{Q}$. (As $\mathbb{Z}_2^k$ and B are abelian, in this case the tensor product may be replaced with the regular product in $\mathbb{C}$.)

We briefly review the representation theory of the quaternion group. Q has five irreducible representations: four one-dimensional and one two-dimensional. The one-dimensional representations arise as the irreducible representations of the abelian quotient $Q/\{\pm 1\} \cong \mathbb{Z}_2 \oplus \mathbb{Z}_2$. The two-dimensional representation τ realizes Q as a subgroup of $\mathbf{SU}_2$, where

$$\tau(1) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \tau(i) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$
$$\tau(j) = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \qquad \tau(k) = \begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix},$$

and $\tau(-q) = -\tau(q)$ for each $q \in Q$.

As above, we assume that we have a set of samples S so that

$$H = \bigcap_{\zeta \otimes \beta \otimes \kappa \in S} \ker(\zeta \otimes \beta \otimes \kappa).$$

It is sufficient to show that for a given element $q \in Q$, one can generate a collection of random elements of $H \cap \mathbb{Z}_2^k \times B \times \{q\}$, for if these collections are large enough, then their union yields a set of generators for H with high probability.

Fixing an element $q \in Q$, consider a specific sample $\zeta \otimes \beta \otimes \kappa$. There are two cases to consider:

- If κ is one-dimensional, the condition $\zeta(z) \otimes \beta(b) \otimes \kappa(q) = 1$ may be interpreted as an equation over $\mathbb{Z}_N$, where $N = b2^{k+1}$, as in the abelian case above. (Note that $\kappa(q) = \pm 1$ contributes a constant to the equation; as $2 \mid N$, this constant can be suitably represented as $\exp(2\pi i t/N)$ for $t = N/2$.)
- If κ is two-dimensional, the condition

  $$\zeta(z) \otimes \beta(b) \otimes \kappa(q) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

  cannot be satisfied unless $q = \pm 1$. When $q = \pm 1$, this may be interpreted as a pair of equations over $\mathbb{Z}_N$, where $N = 2^{k+1}b$, each equation corresponding to a diagonal entry of the matrix. (Note that $\kappa(q)$ contributes a constant ± 1 to each equation; as $2 \mid N$, these constants can be suitably represented as $\exp(2i\pi t/N)$ for $t = N/2$.)

Now the solution proceeds as in the abelian case. For each $q \in Q$, the above procedure is used to compute $c \log |G|$ random elements of $H \cap \mathbb{Z}_2^n \times B \times \{q\}$ (unless this intersection is empty). If c is chosen appropriately, the union of these sets generates H with high probability.

Acknowledgment. The authors thank Umesh Vazirani for many helpful discussions and the simplification of several of the proofs.

REFERENCES

[1] S. Aaronson, *Quantum lower bounds for the collision problem*, in Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing, Montreal, Quebec, Canada, 2002, Association for Computing Machinery, New York, 2002, pp. 635–642.

[2] R. Beals, *Quantum computation of Fourier transforms over symmetric groups*, in Proceedings of the Twenty-Ninth Annual ACM Symposium on Theory of Computing, El Paso, TX, 1997, Association for Computing Machinery, New York, 1997, pp. 48–53.

[3] D. Boneh and R. Lipton, *Quantum cryptanalysis of hidden linear functions (extended abstract)*, in Advances in Cryptology—CRYPTO '95, D. Coppersmith, ed., Lecture Notes in Comput. Sci. 963, Springer-Verlag, Berlin, 1995, pp. 424–437.

[4] R. Cleve and J. Watrous, *Fast parallel circuits for the quantum Fourier transform*, in Proceedings of the 41st Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press, Los Alamitos, CA, 2000, pp. 526–536.

[5] M. Ettinger and P. Høyer, *A Quantum Observable for the Graph Isomorphism Problem*, Los Alamos preprint, quant-ph/9901029, 1999; available online from http://lanl.arxiv.org/abs/quant-ph/9901029.

[6] M. Ettinger and P. Høyer, *Quantum State Detection via Elimination*, Los Alamos preprint, quant-ph/9905099, 1999; available online from http://lanl.arxiv.org/abs/quant-ph/9905099.

[7] M. Ettinger and P. Høyer, *On quantum algorithms for noncommutative hidden subgroups*, Adv. Appl. Math., (2000), pp. 239–251.

[8] M. Ettinger, P. Høyer, and E. Knill, *Hidden Subgroup States Are Almost Orthogonal,* Los Alamos preprint, quant-ph/9901034, 1999; available online from http://lanl.arxiv.org/abs/quant-ph/9901034.

[9] M. Goldmann and A. Russell, *The complexity of solving equations over finite groups*, Inform. Comput., 178 (2002), pp. 253–262.

[10] M. Grigni, L. Schulman, M. Vazirani, and U. Vazirani, *Quantum mechanical algorithms for the nonabelian hidden subgroup problem*, in Proceedings of the Thirty-Third Annual ACM Symposium on Theory of Computing, Crete, Greece, 2001, Association for Computing Machinery, New York, 2001, pp. 68–74.

[11] L. Hales and S. Hallgren, *Quantum fourier sampling simplified*, in Proceedings of the Thirty-First Annual ACM Symposium on Theory of Computing, Atlanta, Georgia, 1999, Association for Computing Machinery, New York, 1999, pp. 330–338.

[12] L. Hales and S. Hallgren, *An improved quantum fourier transform algorithm and applications*, in Proceedings of the 41st Annual Symposium on Foundations of Computer Science, IEEE Computer Society Press, Los Alamitos, CA, 2000, pp. 515–525.

[13] S. Hallgren, *Polynomial-time quantum algorithms for Pell's equation and the principal ideal problem*, in Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing, Montreal, Quebec, Canada, 2002, Association for Computing Machinery, New York, 2002, pp. 653–658.

[14] S. Hallgren, A. Russell, and A. Ta-Shma, *Normal subgroup reconstruction and quantum computation using group representations*, in Proceedings of the Thirty-Second Annual ACM Symposium on Theory of Computing, Portland, Oregon, 2000, Association for Computing Machinery, New York, 2000.

[15] J. Harris and W. Fulton, *Representation Theory*, Grad. Texts in Math. 129, Springer-Verlag, New York, 1991.

[16] G. Ivanyos, F. Magniez, and M. Santha, *Efficient quantum algorithms for some instances of the nonabelian hidden subgroup problem*, in Proceedings of the Thirteenth Annual ACM Symposium on Parallel Algorithms and Architectures, Heraklion, Crete Island, Greece, 2001, Association for Computing Machinery, New York, 2001, pp. 263–270.

[17] A. Kitaev, *Quantum computations: Algorithms and error correction*, Russian Math. Surveys, 52 (1997), pp. 1191–1249.
[18] J. Köbler, U. Schöning, and J. Torán, *The Graph Isomorphism Problem: Its Structural Complexity*, Progr. Theoret. Comput. Sci., Birkhäuser Boston, Boston, 1993.
[19] M. Mosca and A. Ekert, *The hidden subgroup problem and eigenvalue estimation on a quantum computer*, in Proceedings of the 1st NASA International Conference on Quantum Computing and Quantum Communications, C. Williams, ed., Lecture Notes in Comput. Sci. 1509, Springer-Verlag, Berlin, 1999, pp. 174–188.
[20] R. Motwani and P. Raghavan, *Randomized Algorithms*, Cambridge University Press, Cambridge, UK, 1995.
[21] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, Cambridge, UK, 2000.
[22] M. Püschel, M. Rötteler, and T. Beth, *Fast quantum fourier transforms for a class of non-abelian groups*, in Proceedings of Applied Algebra Algebraic Algorithms, and Error-Correcting Codes (AAECC-13), Lecture Notes in Comput. Sci. 1719, Springer-Verlag, Berlin, 1999, pp. 148–159.
[23] J. Rotman, *An Introduction to the Theory of Groups*, 4th ed., Grad. Texts in Math. 149, Springer-Verlag, Berlin, 1995.
[24] M. Rötteler and T. Beth, *Polynomial-Time Solution to the Hidden Subgroup Problem for a Class of Non-Abelian Groups,* Los Alamos preprint, quant-ph/9812070, 1998; available online from http://lanl.arxiv.org/abs/quant-ph/9812070.
[25] B. E. Sagan, *The Symmetric Group*, Wadsworth and Brooks/Cole, Pacific Grove, CA, 1991.
[26] J.-P. Serre, *Linear Representations of Finite Groups*, Springer-Verlag, New York, 1977.
[27] P. W. Shor, *Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer*, SIAM J. Comput., 26 (1997), pp. 1484–1509.
[28] D. R. Simon, *On the power of quantum computation*, SIAM J. Comput., 26 (1997), pp. 1474–1483.