The Symmetric Group Defies Strong Fourier Sampling: Part I

Cristopher Moore
moore@cs.unm.edu
Department of Computer Science
University of New Mexico

Alexander Russell
acr@cse.uconn.edu
Department of Computer Science and Engineering
University of Connecticut

Leonard J. Schulman
schulman@cs.caltech.edu
Computer Science Department
California Institute of Technology

February 13, 2005

Abstract

We resolve the question of whether Fourier sampling can efficiently solve the hidden subgroup problem. Specifically, we show that the hidden subgroup problem over the symmetric group cannot be efficiently solved by strong Fourier sampling, even if one may perform an arbitrary POVM on the coset state. These results apply to the special case relevant to the Graph Isomorphism problem.

1 Introduction: the hidden subgroup problem

Many problems of interest in quantum computing can be reduced to an instance of the *Hidden Subgroup Problem* (HSP). We are given a group G and a function f with the promise that, for some subgroup $H \subseteq G$, f is invariant precisely under translation by H: that is, f is constant on the cosets of H and takes distinct values on distinct cosets. We then wish to determine the subgroup H by querying f.

For example, in Simon's problem [29], $G = \mathbb{Z}_2^n$ and f is an oracle such that, for some y, $f(x) = f(x+y)$ for all x; in this case $H = \{0, y\}$ and we wish to identify y. In Shor's factoring algorithm [28] G is the group $\mathbb{Z}_n^*$ where n is the number we wish to factor, $f(x) = r^x \bmod n$ for a random $r < n$, and H is the subgroup of $\mathbb{Z}_n^*$ whose index is the multiplicative order of r. Both Simon's and Shor's algorithms use the following approach, referred to as the *standard method* or *Fourier sampling* [3].

Step 1. Prepare two registers, the first in a uniform superposition over the elements of G and the second with the value zero, yielding the state

$$\psi_1 = \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g\rangle \otimes |0\rangle \quad .$$

Step 2. Query (or calculate) the function f defined on G and XOR it with the second register. This entangles the two registers and results in the state

$$\psi_2 = \frac{1}{\sqrt{|G|}} \sum_{g \in G} |g\rangle \otimes |f(g)\rangle \quad .$$

Step 3. Measure the second register. This puts the first register in a uniform superposition over one of f's level sets, i.e., one of the cosets of H, and disentangles it from the second register. If we observe the value $f(c)$, we have the state $\psi_3 \otimes |f(c)\rangle$ where

$$\psi_3 = |cH\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |ch\rangle \quad .$$

Alternately, we can view the first register as being in a mixed state with density matrix

$$\frac{1}{|G|} \sum_c |cH\rangle \langle cH|$$

where the sum includes one representative c for each of H's cosets.

Step 4. Carry out the quantum Fourier transform on ψ_3 and measure the result.

(Note that in Shor's algorithm, since $|\mathbb{Z}_n^*|$ is unknown, the Fourier transform is performed over $\mathbb{Z}_q$ for some $q = \text{poly}(n)$; see [28] or [10, 11].)

In both Simon's and Shor's algorithms, the group G is abelian; it is not hard to see that, in this abelian case, a polynomial number (i.e., polynomial in $\log |G|$) of experiments of this type determine H. In essence, each experiment yields a random element of the dual space $H^\perp$ perpendicular to H's characteristic function, and as soon as these elements span $H^\perp$ we are done.

While the *nonabelian* hidden subgroup problem appears to be much more difficult, it has very attractive applications. In particular, solving the HSP for the symmetric group S_n would provide an efficient quantum algorithm for the Graph Automorphism and Graph Isomorphism problems (see e.g. Jozsa [17] for a review). Another important motivation is the relationship between the HSP over the dihedral group with hidden shift problems [4] and cryptographically important cases of the Shortest Lattice Vector problem [23].

So far, algorithms for the HSP are only known for a few families of nonabelian groups, including wreath products $\mathbb{Z}_2^k \wr \mathbb{Z}_2$ [24]; more generally, semidirect products $K \ltimes \mathbb{Z}_2^k$ where K is of polynomial size, and groups whose commutator subgroup is of polynomial size [16]; "smoothly solvable" groups [7]; and some semidirect products of cyclic groups [14]. Ettinger and Høyer [5] provided another type of result, by showing that Fourier sampling can solve the HSP for the dihedral groups D_n in an *information-theoretic* sense. That is, a polynomial number of experiments gives enough information to reconstruct the subgroup, though it is unfortunately unknown how to determine H from this information in polynomial time.

To discuss Fourier sampling for a nonabelian group G, one needs to develop the Fourier transform over G which relies on the group's *linear representations*. For abelian groups, the Fourier basis functions are homomorphisms $\phi : G \to \mathbb{C}$ such as the familiar exponential function $\phi_k(x) = e^{2\pi i k x/n}$ for the cyclic group $\mathbb{Z}_n$. In the nonabelian case, there are not enough such homomorphisms to span the space of all $\mathbb{C}$-valued functions on G; to complete the picture, one introduces *representations* of the group, namely homomorphisms $\rho : G \to \mathsf{U}(V)$ where $\mathsf{U}(V)$ is the group of unitary matrices acting on some $\mathbb{C}$-vector space V of dimension d_ρ. It suffices to consider *irreducible* representations, namely those for which no nontrivial subspace of V is fixed by the various operators $\rho(g)$. Once a basis for each irreducible ρ is chosen, the matrix elements ρ_{ij} provide an orthogonal basis for the vector space of all $\mathbb{C}$-valued functions on G.

The quantum Fourier transform then consists of transforming (unit-length) vectors in $\mathbb{C}[G] = \{\sum_{g \in G} \alpha_g |g\rangle \mid \alpha_g \in \mathbb{C}\}$ from the basis $\{|g\rangle \mid g \in G\}$ to the basis $\{|\rho, i, j\rangle\}$ where ρ is the name of an irreducible representation and $1 \leq i, j \leq d_\rho$ index a row and column (in a chosen basis for V). Indeed, this transformation can be carried out efficiently for a wide variety of groups [2, 13, 21]. Note, however, that a nonabelian group G does not distinguish any specific basis for its irreducible representations which necessitates a rather dramatic choice on the part of the transform designer. Indeed, careful basis selection appears to be critical for obtaining efficient Fourier transforms for the groups mentioned above.

Perhaps the most fundamental question concerning the hidden subgroup problem is whether there is always a basis for the representations of G such that measuring in this basis (in Step 4, above) provides enough information to determine the subgroup H. This framework is known as *strong Fourier sampling.*

In this article, we answer this question in the negative, showing that natural subgroups of S_n cannot be determined by this process; in fact, we show this for an even more general model, where we perform an arbitrary positive operator-valued measurement (POVM) on coset states $|cH\rangle$. We emphasize that the subgroups on which we focus are among the most important special cases of the HSP, as they are those to which Graph Isomorphism naturally reduces.

Related work. The terminology "strong Fourier sampling" [9] was invented to distinguish this approach from the natural variant, called *weak Fourier sampling*, where one only measures the name of the representation ρ, and ignores the row and column information. Weak Fourier sampling is basis-independent, making it attractive from the standpoint of analysis; however, it cannot distinguish conjugate subgroups from each other, and Hallgren, Russell and Ta-Shma [12] showed that it cannot distinguish the trivial subgroup from an order-2 subgroup consisting of $n/2$ disjoint transpositions. Specifically, they used character bounds to show that the probability distribution obtained on representation names for the trivial and order-2 subgroups are exponentially close in total variation distance: it requires an exponential number of such experiments to distinguish them. Kempe and Shalev [18] have generalized this result to other conjugacy classes, and conjecture that one can do no better than classical computation with this approach.

In an effort to shed light on the power of strong Fourier sampling, Grigni, Schulman, Vazirani and Vazirani [9] showed that, for groups such as S_n, measuring in a *random* basis yields an exponentially small amount of information. This can be explained, roughly, by the fact that projecting a vector into a sufficiently high-dimensional random subspace results in tightly concentrated length. On the other hand, Moore, Rockmore, Russell and Schulman [22] showed that for the affine and q-hedral groups, measuring in a well-chosen basis can solve the HSP (at least information-theoretically) in cases where random bases cannot.

Our contribution. In this paper we show that strong Fourier sampling, in an arbitrary basis of the algorithm designer's choice, cannot solve the HSP for S_n. As in [12] we focus on order-2 subgroups of the form $\{1, m\}$ where m is an involution consisting of $n/2$ disjoint transpositions; we remark that if we fix two rigid connected graphs of size $n/2$ and consider permutations of their disjoint union, then the hidden subgroup is of this form if the graphs are isomorphic and trivial if they are not. Then we show that strong Fourier sampling—and more generally, arbitrary measurements of coset states—cannot distinguish these subgroups from each other, or from the trivial subgroup, without an exponential number of experiments.

We remark that our results do not preclude the existence of an efficient quantum algorithm for the HSP on S_n. Rather, they force us to consider *multi-register* algorithms, in which we prepare multiple coset states and subject them to entangled measurements. Ettinger, Høyer and Knill [6] showed that the HSP on arbitrary groups can be solved information-theoretically with a polynomial number of registers, although their algorithm takes exponential time for most groups of interest. Kuperberg [20] devised a subexponential ($2^{O(\sqrt{\log n})}$) algorithm for the HSP on the dihedral group D_n that works by performing entangled measurements on two registers at a time, and Bacon, Childs, and van Dam [1] have determined the optimal multiregister measurement for the dihedral group.

Whether a similar approach can be taken to the symmetric group is a major open question. In a companion paper, the first two authors take a step towards answering this question by showing that if we perform arbitrary entangled measurements over pairs of registers, distinguishing $H = \{1, m\}$ from the trivial group in S_n requires a superpolynomial number (specifically, $e^{\Omega(\sqrt{n}/\log n)}$) of experiments.

2 Fourier analysis over finite groups

We briefly discuss the elements of the representation theory of finite groups. Our treatment is primarily for the purposes of setting down notation; we refer the reader to [8, 27] for complete accounts.

Let G be a finite group. A *representation* ρ of G is a homomorphism $\rho : G \to \mathsf{U}(V)$, where V is a finite dimensional Hilbert space and $\mathsf{U}(V)$ is the group of unitary operators on V. The *dimension* of ρ, denoted d_ρ, is the dimension of the vector space V. By choosing a basis for V, then, each $\rho(g)$ is associated with a unitary matrix $[\rho(g)]$ so that for every $g, h \in G$, $[\rho(gh)] = [\rho(g)] \cdot [\rho(h)]$ where $\cdot$ denotes matrix multiplication.

Fixing a representation $\rho : G \to \mathsf{U}(V)$, we say that a subspace $W \subset V$ is *(left)-invariant* if $\rho(g)W \subset W$ for all $g \in G$; observe that in this case the restriction $\rho_W : G \to \mathsf{U}(W)$, given by restricting each $\rho(g)$ to W, is also a representation. When ρ has no invariant subspaces other than the trivial space $\{\mathbf{0}\}$ and V, ρ is said to be *irreducible.* When ρ is irreducible, *Schur's lemma* asserts that the centralizer of the subgroup $\mathbf{im}\,\rho \subset \mathsf{U}(V) \subset \mathsf{GL}(V)$—that is, the set of $A \in \mathsf{GL}(V)$ such that $A\rho(g) = \rho(g)A$ for all $g \in G$—consists only of the scalar matrices $\{c\mathbb{1} \mid c \in \mathbb{C}\}$. We use this fact below.

In the case when ρ is *not* irreducible, then, there is a nontrivial invariant subspace $W \subset V$ and, as the inner product $\langle \cdot, \cdot \rangle$ is invariant under the unitary maps $\rho(g)$, it is immediate that the dual subspace

$$W^\perp = \{\mathbf{u} \mid \forall \mathbf{w} \in W, \langle \mathbf{u}, \mathbf{w} \rangle = 0\}$$

is also invariant. Associated with the decomposition $V = W \oplus W^\perp$ is the natural decomposition of the operators $\rho(g) = \rho_W(g) \oplus \rho_{W^\perp}(g)$. By repeating this process, any representation $\rho : G \to \mathsf{U}(V)$ may be decomposed into a direct sum of irreducible representations: we write $\rho = \sigma_1 \oplus \cdots \oplus \sigma_k$ and, for the σ_i appearing at least once in this decomposition, $\sigma_i \prec \rho$. In general, a given σ can appear in multiply in this decomposition, in the sense that ρ may have an invariant subspace isomorphic to the direct sum of a_σ copies of σ. In this case a_σ is called the *multiplicity* of σ_i in ρ, and we write $\rho = \bigoplus_{\sigma \prec \rho} a_\sigma \sigma$ where $a_\sigma \sigma = \underbrace{\sigma \oplus \cdots \oplus \sigma}_{a_\sigma}$.

If two representations ρ and σ are the same up to a unitary change of basis, we say that they are *equivalent.* It is a fact that any finite group G has a finite number of distinct irreducible representations up to equivalence and, for a group G, we let $\widehat{G}$ denote a set of representations containing exactly one from each equivalence class. The irreducible representations of G give rise to the Fourier transform. Specifically, for a function $f : G \to \mathbb{C}$ and an element $\rho \in \widehat{G}$, define the *Fourier transform of f at ρ* to be

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

The leading coefficients are chosen to the make the transform unitary, so that it preserves inner products:

$$\langle f_1, f_2 \rangle = \sum_g f_1^*(g) f_2(g) = \sum_{\rho \in \widehat{G}} \mathbf{tr}\left(\hat{f}_1(\rho)^\dagger \cdot \hat{f}_2(\rho)\right) \quad .$$

There is a natural product operation on representations: if $\rho : G \to \mathsf{U}(V)$ and $\sigma : G \to \mathsf{U}(W)$ are representations of G, we may define a new representation $\rho \otimes \sigma : G \to \mathsf{U}(V \otimes W)$ by extending the rule $\rho \otimes \sigma(g) : \mathbf{u} \otimes \mathbf{v} \mapsto \rho(g)\mathbf{u} \otimes \sigma(g)\mathbf{v}$. In general, the representation $\rho \otimes \sigma$ is not irreducible, even when both ρ and σ are. This leads to the *Clebsch-Gordan problem*, that of decomposing $\rho \otimes \sigma$ into irreducibles.

For a representation ρ we define the *character* of ρ, denoted χ_ρ, to be the function $\chi_\rho : G \to \mathbb{C}$ given by $\chi_\rho(g) = \mathbf{tr}\,\rho(g)$. As the trace of a linear operator is invariant under conjugation, characters are constant on the conjugacy classes of G; for a conjugacy class $A = \{gag^{-1} \mid g \in G\}$, we define $\chi(A) = \chi(a)$. Characters are a powerful tool for reasoning about the decomposition of reducible representations. In particular, when $\rho = \bigoplus_i \sigma_i$ we have $\chi_\rho = \sum_i \chi_{\sigma_i}$ and, moreover, for $\rho, \sigma \in \widehat{G}$, we have the orthogonality conditions

$$\langle \chi_\rho, \chi_\sigma \rangle_G = \frac{1}{|G|} \sum_{g \in G} \chi_\rho(g) \chi_\sigma(g)^* = \begin{cases} 1 & \rho = \sigma \quad , \\ 0 & \rho \neq \sigma \quad . \end{cases}$$

Then for an irreducible representation ρ and representation σ, $\langle \chi_\rho, \chi_\sigma \rangle_G$ is equal to the multiplicity with which ρ appears in the decomposition of σ. For example, since $\chi_{\rho \otimes \sigma}(g) = \chi_\rho(g) \cdot \chi_\sigma(g)$, the multiplicity of τ in $\rho \otimes \sigma$ is $\langle \chi_\tau, \chi_\rho \chi_\sigma \rangle_G$

In general, we can consider subspaces of $\mathbb{C}[G]$ that are invariant under left multiplication, right multiplication, or both; these subspaces are called *left-*, *right-*, or *bi-invariant* respectively. Each $\rho \in \widehat{G}$ corresponds to a d_ρ^2-dimensional bi-invariant subspace of $\mathbb{C}[G]$, which can be broken up further into d_ρ d_ρ-dimensional left-invariant subspaces, or (transversely) d_ρ d_ρ-dimensional right-invariant subspaces. However, this decomposition is not unique. If ρ acts on a vector space V, then choosing an orthonormal basis for V allows us to view $\rho(g)$ as a $d_\rho \times d_\rho$ matrix. Then ρ acts on the d_ρ^2-dimensional space of such matrices by left or right multiplication, and the columns and rows correspond to left- and right-invariant spaces respectively. More generally, each left-invariant subspace isomorphic to ρ corresponds to a unit vector $\mathbf{b} \in V$.

3 The structure of the optimal measurement

In this section we show that starting with a random coset state, the optimal one-register measurement for the hidden subgroup problem is precisely an instance of strong Fourier sampling (possibly in an over-complete basis). This has been pointed out several times in the past, at varying levels of explicitness [15, 20]; we state it here for completeness. Everything we say in this section is true for the hidden subgroup problem in general. However, for simplicity we focus on the special case of the hidden subgroup problem called the *hidden conjugate problem* in [22]: there is a (non-normal) subgroup H, and we are promised that the hidden subgroup is one of its conjugates, $H^g = g^{-1}Hg$ for some $g \in G$.

We may treat the states arising after Step 3 of the procedure above as elements of the group algebra $\mathbb{C}[G]$. We use the notation $|g\rangle = 1 \cdot g \in \mathbb{C}[G]$ so that the vectors $|g\rangle$ form an orthonormal basis for $\mathbb{C}[G]$. Given a set $S \subset G$, $|S\rangle$ denotes a uniform superposition over the elements of S, $|S\rangle = (1/\sqrt{|S|}) \sum_{s \in S} |s\rangle$.

3.1 The optimal POVM consists of strong Fourier sampling

The most general type of measurement allowed in quantum mechanics is a *positive operator-valued measurement* (POVM). A POVM with a set of possible outcomes J consists of a set of positive operators $\{M_j \mid j \in J\}$ subject to the completeness condition,

$$\sum_j M_j = \mathbb{1} \ . \tag{3.1}$$

Since positive operators are self-adjoint, they can be orthogonally diagonalized, and since their eigenvalues are positive, they may be written as a positive linear combination of projection operators (see e.g. [26, §10]). Any POVM may thus be refined so that each $M_j = a_j \mu_j$ where μ_j is a projection operator and a_j is positive and real.

The result of this measurement on the state $|\psi\rangle$ is a random variable, taking values in J, that is equal to $j \in J$ with probability $P_j = a_j \langle\psi| \mu_j |\psi\rangle$. Note that outcomes j need not correspond to subgroups directly; the algorithm designer is free to carry out a polynomial number t of experiments, observing outcomes $j_1, \ldots, j_t$, and then apply some additional computation to find the most likely subgroup given these observations.

If g is chosen from G uniformly so that the hidden subgroup is a uniformly random conjugate of H, we wish to find a POVM that maximizes the probability of correctly identifying g from the coset state $|H^g\rangle$. (Of course, to identify a conjugate H^g, we only need to specify g up to an element of the normalizer of H.) Since a random left coset of H^g can be written $cgH^g = cHg$ for a random $c \in G$, the probability we observe outcome j is

$$P_j = a_j \frac{1}{|G|} \sum_{c \in G} \langle cHg| \mu_j |cHg\rangle \ . \tag{3.2}$$

Ip [15] observed that in the special case that each outcome j corresponds to a subgroup, maximizing the probability that j is correct subject to the constraint (3.1) gives a semi-definite program. Since such programs are convex, the optimum is unique and is a fixed point of any symmetries possessed by the problem.

However, our proof relies on an elementary "symmetrization" argument. Given a group element $x \in G$, let $L_x |g\rangle = |xg\rangle$ denote the unitary matrix corresponding to left group multiplication by x. In particular, applying L_x maps one left coset onto another: $|cHg\rangle = L_c |Hg\rangle$. Writing

$$P_j = a_j \frac{1}{|G|} \sum_{c \in G} \langle cHg| \mu_j |cHg\rangle = a_j \left\langle Hg \left| \frac{1}{|G|} \sum_{c \in G} L_c^\dagger \mu_j L_c \right| Hg \right\rangle \ ,$$

we conclude that replacing μ_j for each j with the symmetrization

$$\mu'_j = \frac{1}{|G|} \sum_{g \in G} L_g^\dagger \mu_j L_g$$

does not change the resulting probability distribution P_j. Since μ'_j commutes with L_x for every $x \in G$ and provides exactly the same information as the original μ_j, we may assume without loss of generality that the optimal POVM commutes with L_x for every $x \in G$.

It is easy to see that any projection operator that commutes with left multiplication projects onto a left-invariant subspace of $\mathbb{C}[G]$, and we can further refine the POVM so that each μ_j projects onto an *irreducible* left-invariant subspace. Each such space is contained in the bi-invariant subspace corresponding to some irreducible representation ρ, in which case we write $\mathbf{im}\,\mu_j \subseteq \rho$. As discussed in Section 2, a given irreducible left-invariant subspace corresponds to some unit vector $\mathbf{b}$ in the vector space V on which ρ acts. Thus we can write

$$\mu_j = |\mathbf{b}_j\rangle \langle \mathbf{b}_j| \otimes \frac{1}{d_\rho} \mathbb{1}_{d_\rho}$$

where $\mathbb{1}_{d_\rho}$ acts within the left-invariant subspace. Let $B = \{\mathbf{b}_j \mid \mathbf{im}\,\mu_j \in \rho\}$; then (3.1) implies a completeness condition for each $\rho \in \widehat{G}$,

$$\sum_{\mathbf{b}_j \in B} a_j |\mathbf{b}_j\rangle \langle \mathbf{b}_j| = \mathbb{1}_{d_\rho} \tag{3.3}$$

and so B is a (possibly over-complete) basis for V. In other words, the optimal POVM consists of first measuring the representation name ρ, and then performing a POVM on the vector space V with possible outcomes B. Another way to see this is to regard the choice of coset as a mixed state; then its density matrix is block-diagonal in the Fourier basis, and so as Kuperberg puts it [20] measuring the representation name "sacrifices no entropy."

We note that in the special case that this POVM is a von Neumann measurement—that is, when the B is an orthonormal basis for V—then it corresponds to measuring the column of ρ in that basis, which is how strong Fourier sampling is usually defined. (As pointed out in [9], nothing is gained by measuring the row, since we have a random left coset cHg and left-multiplying by a random element c in an irreducible representation completely mixes the probability across the rows in each column. Here this is reflected by the fact that μ_j is a scalar in each left-invariant subspace.) However, in general the optimal measurement might consist of an over-complete basis, or *frame*, in each ρ, consisting of the vectors $\{\mathbf{b}_j\}$ and the weights a_j.

Now that we know μ_j takes this form, let us change notation. Given $\rho \in \widehat{G}$ acting on a vector space V and a unit vector $\mathbf{b} \in V$, let $\Pi^\rho_{\mathbf{b}} = |\mathbf{b}\rangle \langle \mathbf{b}| \otimes \mathbb{1}_{d_\rho}$ denote the projection operator onto the left-invariant subspace corresponding to $\mathbf{b}$. Then $\mu_j = \Pi^\rho_{\mathbf{b}_j}$, and (3.2) becomes

$$P_j = a_j \frac{1}{|G|} \sum_{c \in G} \left\| \Pi^\rho_{\mathbf{b}_j} |cHg\rangle \right\|^2 = a_j \left\| \Pi^\rho_{\mathbf{b}_j} |Hg\rangle \right\|^2 \quad . \tag{3.4}$$

We can write this as the product of the probability $P(\rho)$ that we observe ρ, times the conditional probability $P(\rho, \mathbf{b}_j)$ that we observe $\mathbf{b}_j$. Note that by (3.3),

$$\Pi^\rho = \sum_{\mathbf{b}_j \in B} a_j \Pi^\rho_{\mathbf{b}_j}$$

is the projection operator onto the bi-invariant subspace corresponding to ρ. Then

$$P_j = P(\rho) P(\rho, \mathbf{b}_j)$$

where

$$\begin{aligned} P(\rho) &= \left\| \Pi^\rho |H\rangle \right\|^2 && (3.5) \\ P(\rho, \mathbf{b}_j) &= a_j \left\| \Pi^\rho_{\mathbf{b}_j} |Hg\rangle \right\|^2 \Big/ P(\rho) \quad . && (3.6) \end{aligned}$$

Note that $P(\rho, \mathbf{b}_j)$ depends on g but $P(\rho)$ does not, which is why weak sampling is incapable of distinguishing conjugate subgroups.

3.2 The probability distribution for a conjugate subgroup

Now let us use the fact that $|H\rangle$ is a superposition over a subgroup, and calculate $P(\rho)$ and $P(\rho, \mathbf{b}_j)$ as defined in (3.5) and (3.6). This will set the stage for asking whether we can distinguish different conjugates of H from from the trivial subgroup and from each other.

Fix an irreducible representation ρ that acts on a vector space V. Then Fourier transforming the state

$$|H\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |h\rangle$$

yields the coefficient

$$\widehat{H}(\rho) = \sqrt{\frac{d_\rho}{|H||G|}} \sum_{h \in H} \rho(h) = \sqrt{\frac{d_\rho |H|}{|G|}}\, \Pi_H$$

where $\Pi_H = (1/|H|) \sum_{h \in H} \rho(h)$ is a projection operator onto a subspace of V. The probability that we observe ρ is then the norm squared of this coefficient,

$$P(\rho) = \left\| \widehat{H}(\rho) \right\|^2 = \frac{d_\rho |H|}{|G|}\, \mathbf{rk}\, \Pi_H \tag{3.7}$$

and, as stated above, this is the same for all conjugates H^g. The conditional probability that we observe the vector $\mathbf{b}_j$, given that we observe ρ, is then

$$P(\rho, \mathbf{b}_j) = a_j \frac{\left\| \Pi^\rho_{\mathbf{b}_j} |H\rangle \right\|^2}{P(\rho)} = a_j \frac{\left\| \widehat{H}(\rho) \mathbf{b}_j \right\|^2}{P(\rho)} = a_j \frac{\|\Pi_H \mathbf{b}_j\|^2}{\mathbf{rk}\, \Pi_H} \quad . \tag{3.8}$$

In the case where H is the trivial subgroup, $\Pi_H = \mathbb{1}_{d_\rho}$ and $P(\rho, \mathbf{b}_j)$ is given by

$$P(\rho, \mathbf{b}_j) = \frac{a_j}{d_\rho} \quad . \tag{3.9}$$

We call this the *natural distribution* on the frame $B = \{\mathbf{b}_j\}$. In the case that B is an orthonormal basis, $a_j = 1$ and this is simply the uniform distribution.

This probability distribution over B changes for a conjugate H^g in the following way. Again ignoring left multiplication since the columns are left-invariant, the Fourier transform becomes

$$\widehat{Hg}(\rho) = \sqrt{\frac{d_\rho |H|}{|G|}}\, \Pi_H \rho(g)$$

and we have

$$P(\rho, \mathbf{b}_j) = a_j \frac{\|\Pi_H g \mathbf{b}_j\|^2}{\mathbf{rk}\, \Pi_H}$$

where we write $g\mathbf{b}$ for $\rho(g)\mathbf{b}$. It is not hard to show that, for any $\mathbf{b} \in V$, the *expected* value of $\|\Pi_H(g\mathbf{b})\|^2$, over the choice of $g \in G$, is $\mathbf{rk}\ \Pi_H / d_\rho$. Our primary technical contribution is a method for establishing concentration results for this random variable.

4 The variance of projection through a random involution

In this section we focus on the case where $H = \{1, m\}$ for an element m chosen uniformly at random from a conjugacy class $[m]$ of involutions. (Observe that order is preserved under conjugation so that if m is an involution, then so are all elements of $[m]$.) Given an irreducible representation $\rho : G \to \mathsf{U}(V)$ and a vector $\mathbf{b} \in V$, we bound the variance, over the choice of m, of the probability $P(\rho, \mathbf{b})$ that $\mathbf{b}$ is observed given that we observed ρ. Our key insight is that this variance depends on how the tensor product representation $\rho \otimes \rho^*$ decomposes into irreducible representations σ, and how the vector $\mathbf{b} \otimes \mathbf{b}^*$ projects into the constituent orthogonal irreducibles.

Recall that, if a representation ρ is reducible, it can be written as an orthogonal direct sum of irreducibles $\rho = \bigoplus_{\sigma \prec \rho} a_\sigma \sigma$ where a_σ is the multiplicity of σ. We let Π^ρ_σ denote the projection operator whose image is $a_\sigma \sigma$, that is, the span of all the irreducible subspaces isomorphic to σ.

Lemma 1. *Let ρ be a representation of a group G acting on a space V and let $\mathbf{b} \in V$. Let m be an element chosen uniformly from a conjugacy class $[m]$ of involutions. If ρ is irreducible, then*

$$\mathrm{Exp}_m \langle \mathbf{b}, m\mathbf{b} \rangle = \frac{\chi_\rho([m])}{\dim \rho} \|\mathbf{b}\|^2 \quad .$$

If ρ is reducible, then

$$\mathrm{Exp}_m \langle \mathbf{b}, m\mathbf{b} \rangle = \sum_{\sigma \prec \rho} \frac{\chi_\sigma([m])}{\dim \sigma} \|\Pi^\rho_\sigma \mathbf{b}\|^2 \quad .$$

Proof. Fix a particular element $\mu \in [m]$. Let $\rho([m])$ denote the average of $\rho(m)$ over $[m]$; this is

$$\rho([m]) = \frac{1}{|[m]|} \sum_{m \in [m]} \rho(m) = \frac{1}{|G|} \sum_{g \in G} \rho(g^{-1} \mu g) = \frac{1}{|G|} \sum_{g \in G} \rho(g)^\dagger \rho(\mu)\, \rho(g) \quad .$$

Observe that $\rho([m])$ commutes with $\rho(g)$ for all $g \in G$ and hence, by Schur's lemma, its action on any irreducible subspace is multiplication by a scalar. Note that for the scalar linear operator $A = \alpha \mathbb{1}_d$ acting on a space of dimension d, we have $\alpha = \mathbf{tr}\, A/d$. In particular, if ρ is irreducible then

$$\rho([m]) = \frac{\chi_\rho([m])}{\dim \rho} \mathbb{1}_{d_\rho}$$

and so

$$\mathrm{Exp}_m \langle \mathbf{b}, m\mathbf{b} \rangle = \langle \mathbf{b}, \rho([m])\mathbf{b} \rangle = \frac{\chi_\rho([m])}{\dim \rho} \|\mathbf{b}\|^2 \quad .$$

If ρ is reducible, these same considerations apply to each irreducible subspace:

$$\rho([m]) = \sum_{\sigma \prec \rho} \frac{\chi_\sigma([m])}{\dim \sigma} \Pi^\rho_\sigma$$

and so

$$\mathrm{Exp}_m \langle \mathbf{b}, m\mathbf{b} \rangle = \langle \mathbf{b}, \rho([m])\mathbf{b} \rangle = \sum_{\sigma \prec \rho} \frac{\chi_\sigma([m])}{\dim \sigma} \langle \mathbf{b}, \Pi^\rho_\sigma \mathbf{b} \rangle = \sum_{\sigma \prec \rho} \frac{\chi_\sigma([m])}{\dim \sigma} \|\Pi^\rho_\sigma \mathbf{b}\|^2 \quad .$$

□

Turning now to the second moment of $\langle \mathbf{b}, m\mathbf{b} \rangle$, we observe that

$$|\langle \mathbf{b}, m\mathbf{b} \rangle|^2 = \langle \mathbf{b}, m\mathbf{b} \rangle \langle \mathbf{b}, m\mathbf{b} \rangle^* = \langle \mathbf{b} \otimes \mathbf{b}^*, m\mathbf{b} \otimes m\mathbf{b}^* \rangle = \langle \mathbf{b} \otimes \mathbf{b}^*, m(\mathbf{b} \otimes \mathbf{b}^*) \rangle,$$

where the action of m on the vector $\mathbf{b} \otimes \mathbf{b}^*$ is precisely given by the action of G in the representations $\rho \otimes \rho^*$. This will allow us to express the second moment of the inner product $\langle \mathbf{b}, m\mathbf{b} \rangle$ in terms of the projections of $\mathbf{b} \otimes \mathbf{b}^*$ into the irreducible constituents of the tensor product representation $\rho \otimes \rho^*$.

Lemma 2. *Let ρ be a representation of a group G acting on a space V and let $\mathbf{b} \in V$. Let m be an element chosen uniformly at random from a conjugacy class $[m]$ of involutions. Then*

$$\mathrm{Exp}_m \left|\langle \mathbf{b}, m\mathbf{b}\rangle\right|^2 = \sum_{\sigma \prec \rho\otimes\rho^*} \frac{\chi_\sigma([m])}{\dim \sigma} \left\|\Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2 \quad .$$

Proof. We write the second moment as a first moment over the product representation $\rho \otimes \rho^*$: as above, $\left|\langle \mathbf{b}, m\mathbf{b}\rangle\right|^2 = \langle \mathbf{b}\otimes\mathbf{b}^*, m(\mathbf{b}\otimes\mathbf{b}^*)\rangle$, so that

$$\mathrm{Exp}_m \left|\langle \mathbf{b}, m\mathbf{b}\rangle\right|^2 = \mathrm{Exp}_m \langle \mathbf{b}\otimes\mathbf{b}^*, m(\mathbf{b}\otimes\mathbf{b}^*)\rangle$$

and applying Lemma 1 completes the proof. □

Now let $\Pi_m = \Pi_H$ denote the projection operator given by

$$\Pi_m \mathbf{v} = \frac{\mathbf{v} + m\mathbf{v}}{2} \quad .$$

For a given vector $\mathbf{b} \in B$, we will focus on the expectation and variance of $\|\Pi_m \mathbf{b}\|^2$. These are given by the following lemma.

Lemma 3. *Let ρ be an irreducible representation acting on a space V and let $\mathbf{b} \in V$. Let m be an element chosen uniformly at random from a conjugacy class $[m]$ of involutions. Then*

$$\begin{aligned} \mathrm{Exp}_m \|\Pi_m \mathbf{b}\|^2 &= \frac{1}{2}\|\mathbf{b}\|^2 \left(1 + \frac{\chi_\rho([m])}{\dim\rho}\right) && (4.1)\\ \mathrm{Var}_m \|\Pi_m \mathbf{b}\|^2 &\le \frac{1}{4} \sum_{\sigma \prec \rho\otimes\rho^*} \frac{\chi_\sigma([m])}{\dim \sigma} \left\|\Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2 \quad . && (4.2)\end{aligned}$$

Proof. For the expectation,

$$\begin{aligned} \mathrm{Exp}_m \|\Pi_m \mathbf{b}\|^2 &= \mathrm{Exp}_m \langle \mathbf{b}, \Pi_m \mathbf{b}\rangle \\ &= \frac{1}{2}\mathrm{Exp}_m \left(\langle \mathbf{b}, \mathbf{b}\rangle + \langle \mathbf{b}, m\mathbf{b}\rangle\right) \\ &= \frac{1}{2}\|\mathbf{b}\|^2 \left(1 + \frac{\chi_\rho([m])}{\dim\rho}\right) \end{aligned}$$

where the last equality follows from Lemma 1.

For the variance, we first calculate the second moment,

$$\begin{aligned} \mathrm{Exp}_m \|\Pi_m \mathbf{b}\|^4 &= \mathrm{Exp}_m \left|\langle \mathbf{b}, \Pi_m \mathbf{b}\rangle\right|^2 \\ &= \frac{1}{4}\mathrm{Exp}_m \left|\langle \mathbf{b}, \mathbf{b}\rangle + \langle \mathbf{b}, m\mathbf{b}\rangle\right|^2 \\ &= \frac{1}{4}\mathrm{Exp}_m \left(\left|\langle \mathbf{b}, \mathbf{b}\rangle\right|^2 + 2\Re\langle \mathbf{b}, \mathbf{b}\rangle\langle \mathbf{b}, m\mathbf{b}\rangle + \left|\langle \mathbf{b}, m\mathbf{b}\rangle\right|^2\right) \\ &= \frac{1}{4}\left(\|\mathbf{b}\|^4 + 2\|\mathbf{b}\|^4 \frac{\chi_\rho([m])}{\dim\rho} + \sum_{\sigma \prec \rho\otimes\rho^*} \frac{\chi_\sigma([m])}{\dim \sigma} \left\|\Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2\right) \end{aligned}$$

where in the last line we applied Lemmas 1 and 2 and the fact that any character evaluated at an involution is real. Then

$$\begin{aligned} \mathrm{Var}_m \|\Pi_m \mathbf{b}\|^2 &= \mathrm{Exp}_m \|\Pi_m \mathbf{b}\|^4 - \left(\mathrm{Exp}_m \|\Pi_m \mathbf{b}\|^2\right)^2 \\ &= \frac{1}{4}\left[\sum_{\sigma \prec \rho\otimes\rho^*} \frac{\chi_\rho([m])}{\dim \rho} \left\|\Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2 - \|\mathbf{b}\|^4 \left(\frac{\chi_\rho([m])}{\dim\rho}\right)^2\right] \quad . && (4.3)\end{aligned}$$

Ignoring the second term, which is negative, gives the stated result. □

Finally, we point out that since

$$\mathrm{Exp}_m \|\Pi_m \mathbf{b}\|^2 = \|\mathbf{b}\|^2 \frac{\mathbf{rk}\,\Pi_m}{\dim \rho}$$

we have

$$\frac{\mathbf{rk}\,\Pi_m}{\dim \rho} = \frac{1}{2}\left(1 + \frac{\chi_\rho([m])}{\dim \rho}\right) \quad , \tag{4.4}$$

a fact which we will use below.

5 The representation theory of the symmetric group

In this section we record the particular properties of S_n and its representation theory applied in the proofs of the main results. The irreducible representations of S_n are labeled by Young diagrams, or equivalently by integer partitions of n,

$$\lambda = (\lambda_1, \ldots, \lambda_t)$$

where $\sum_i \lambda_i = n$ and $\lambda_i \geq \lambda_{i+1}$ for all i. The *conjugate* Young diagram λ' is obtained by flipping λ about the diagonal: $\lambda' = (\lambda'_1, \ldots, \lambda'_{\lambda_1})$ where $\lambda'_j = |\{i \mid \lambda_i \geq j\}|$. In particular, $\lambda'_1 = t$. The number of such diagrams, equal to the number of conjugacy classes in S_n, is the partition number $p(n)$, which obeys

$$p(n) = (1 + o(1)) \frac{1}{4\sqrt{3} \cdot n}\, e^{\pi\sqrt{2n/3}} = e^{\Theta(\sqrt{n})} \quad .$$

We denote these irreducibles S^λ, their characters χ^λ, and their dimensions d^λ. In particular, S^λ is the trivial or parity representation if λ is a single row (n) or a single column $(1, \ldots, 1)$ respectively, and $S^{\lambda'}$ is the (tensor) product of S^λ with the parity representation.

The dimensions of the representations of the symmetric group are given by the remarkable *hook length formula*:

$$d^\lambda = \frac{n!}{\prod_c \mathrm{hook}(c)} \quad ,$$

where this product runs over all cells of the Young diagram associated with λ and hook(c) is the number of cells appearing in either the same column or row as c, excluding those that are above or to the left of c.

For example, the partition $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (6, 5, 3, 2)$ is associated with the diagram shown in Figure 1 below. The hook associated with the cell $(2, 2)$ in this diagram appears in Figure 2; it has length 6.

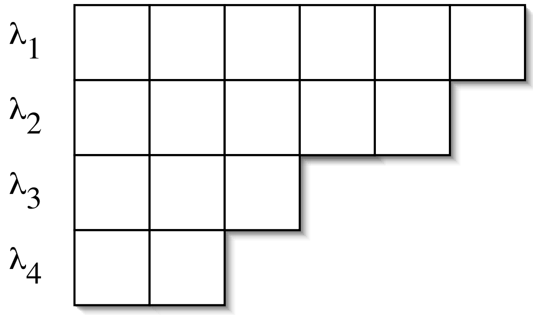

Figure 1: The Young diagram for $\lambda = (6, 5, 3, 2)$.

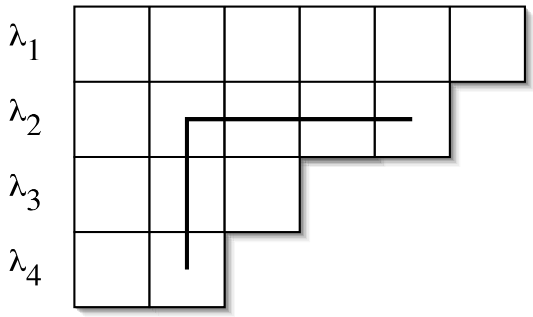

Figure 2: A hook of length 6.

The symmetric groups have the property that every representation S^λ possesses a basis in which its matrix elements are real, and so all its characters are real. However, in a given basis S^λ might be complex, so we will refer below to the complex conjugate representation $(S^\lambda)^*$ (not to be confused with $S^{\lambda'}$).

The study of the asymptotic properties of the representations of S_n typically focuses on the *Plancherel* distribution. (See, e.g., Kerov's monograph [19] for further discussion.) For a general group G, this is the probability distribution obtained on $\widehat{G}$ by assigning ρ the probability density $d_\rho^2/|G|$. One advantage of this distribution is that the density at ρ is proportional to its contribution, dimensionwise, to the group algebra $\mathbb{C}[G]$; this will allow us to reduce structural questions about representations chosen according to this distribution to questions about the regular and conjugation representation. Note that in the context of the hidden subgroup problem, the Plancherel distribution is exactly the one obtained by performing weak Fourier sampling on the trivial hidden subgroup.

In the symmetric groups a fair amount is known about representations chosen according to the Plancherel distribution. In particular, Vershik and Kerov have shown that with high probability they have dimension equal to $e^{\Theta(\sqrt{n})}\sqrt{n!}$.

Theorem 4 ([30]). *Let S^λ be chosen from $\widehat{S_n}$ according to the Plancherel distribution. Then there exist positive constants c_1 and c_2 for which*

$$\lim_{n\to\infty} \Pr\left[e^{-c_1\sqrt{n}}\sqrt{n!} \le d^\lambda \le e^{-c_2\sqrt{n}}\sqrt{n!}\right] = 1 \ .$$

Vershik and Kerov have also obtained estimates for the *maximum* dimension of a representation in $\widehat{S_n}$:

Theorem 5 ([30]). *There exist positive constants $\check{c}$ and $\hat{c}$ such that for all $n \ge 1$,*

$$e^{-\check{c}\sqrt{n}}\sqrt{n!} \le \max_{S^\lambda\in\widehat{S_n}} d^\lambda \le e^{-\hat{c}\sqrt{n}}\sqrt{n!} \ .$$

Along with these estimates, we shall require some (one-sided) large-deviation versions of Theorem 4, recorded below.

Lemma 6. *Let S^λ be chosen according to the Plancherel distribution on $\widehat{S_n}$.*

1. *Let $\delta = \pi\sqrt{2/3}$. Then for sufficiently large n, $\Pr\left[d^\lambda \le e^{-\delta\sqrt{n}}\sqrt{n!}\right] < e^{-\delta\sqrt{n}}$.*

2. *Let $0 < c < 1/2$. Then $\Pr[d^\lambda \le n^{cn}] = n^{-\Omega(n)}$.*

Proof. For the first bound, setting $d = e^{-\delta\sqrt{n}}\sqrt{n!}$ and using $p(n) < e^{\delta\sqrt{n}}$, we have

$$\sum_{S^\lambda : d^\lambda \le d} \frac{(d^\lambda)^2}{n!} \le p(n)\frac{d^2}{n!} < e^{-\delta\sqrt{n}} \ .$$

For the second bound, recalling Sterling's approximation $n! \sim \sqrt{2\pi n}(n/e)^n$, we have

$$\frac{1}{n!}\sum_{\lambda : d^\lambda \le n^{cn}} (d^\lambda)^2 \le \frac{p(n)n^{2cn}}{n!} = n^{-(1-2c)n}e^{O(n)} = n^{-\Omega(n)} \ .$$

□

Finally, we will also apply Roichman's [25] estimates for the characters of the symmetric group:

Definition 1. *For a permutation $\pi \in S_n$, define the* support *of π, denoted* $\operatorname{supp}(\pi)$*, to be the cardinality of the set $\{k \in [n] \mid \pi(k) \ne k\}$.*

Theorem 7 ([25]). *There exist constants $b > 0$ and $0 < q < 1$ so that for $n > 4$, for every conjugacy class C of S_n, and every irreducible representation S^λ of S_n,*

$$\left|\frac{\chi^\lambda(C)}{d^\lambda}\right| \le \left(\max\left(q, \frac{\lambda_1}{n}, \frac{\lambda_1'}{n}\right)\right)^{b\cdot\operatorname{supp}(C)} \ ,$$

where $\operatorname{supp}(C) = \operatorname{supp}(\pi)$ *for any $\pi \in C$.*

In our application, we take n to be even and consider involutions m in the conjugacy class of elements consisting of $n/2$ disjoint transpositions, $M = M_n = \{\sigma\left((12)(34)\cdots(n-1\ n)\right)\sigma^{-1} \mid \sigma \in S_n\}$. Note that each $m \in M_n$ is associated with one of the $(n-1)!!$ perfect matchings of n things, and that $\operatorname{supp}(m) = n$.

6 Strong Fourier sampling over S_n

We consider the hidden subgroup $H = \{1, m\}$, where m is chosen uniformly from $M = M_n \subset S_n$, the conjugacy class

$$\{\pi^{-1}((1\ 2)(3\ 4)\cdots(n-1\ n))\pi \mid \pi \in S_n\} \ ;$$

we assume throughout that n is even. We start by measuring the name of an irreducible representation, yielding S^λ for a diagram λ. We remark that Hallgren, Russell, and Ta-Shma [12] established that the probability distribution on λ is exponentially close to the Plancherel distribution in total variation. We allow the algorithm designer to choose an arbitrary POVM, with a frame $B = \{\mathbf{b}_j\}$ and weights $\{a_j\}$ obeying the completeness condition (3.3). We will show that with high probability (over m and λ), the conditional distribution induced on the vectors B is exponentially close to the natural distribution (3.9) on B. It will follow by the triangle inequality that it requires an exponential number of single-register experiments to distinguish two involutions from each other or, in fact, distinguish H from the trivial subgroup.

For simplicity, and to illustrate our techniques, we first prove this for a von Neumann measurement, i.e., where B is an orthonormal basis for S^λ. In this case, we show that the probability distribution on B (or equivalently, on the columns of S^λ) is exponentially close to the uniform distribution.

6.1 von Neumann measurements

Theorem 8. *Let $B = \{\mathbf{b}\}$ be an orthonormal basis for an irreducible representation S^λ. Given the hidden subgroup $H = \{1, m\}$ where m is chosen uniformly at random from M, let $P_m(\mathbf{b})$ be the probability that we observe the vector $\mathbf{b}$ conditioned on having observed the representation name S^λ, and let U be the uniform distribution on B. Then there is a constant $\delta > 0$ such that for sufficiently large n, with probability at least $1 - e^{-\delta n}$ in m and λ, we have*

$$\|P_m - U\|_1 < e^{-\delta n} \ .$$

Proof. First, recall from (3.8) in Section 3 that the conditional distribution on B is given by (since $a_j = 1$)

$$P_m(\mathbf{b}) = P(S^\lambda, \mathbf{b}) = \frac{\|\Pi_m \mathbf{b}\|^2}{\mathbf{rk}\,\Pi_m} \ . \tag{6.1}$$

Our strategy will be to bound $\mathrm{Var}_m \|\Pi_m \mathbf{b}\|^2$ using Lemma 3, and apply Chebyshev's inequality to conclude that it is almost certainly close to its expectation. Recall, however, that our bounds on the variance of $\|\Pi_m \mathbf{b}\|^2$ depend on the decomposition of $S^\lambda \otimes (S^\lambda)^*$ is into irreducibles and, furthermore, on the projection of $\mathbf{b} \otimes \mathbf{b}^*$ into these irreducible subspaces. Matters are somewhat complicated by the fact that certain S^μ appearing in $S^\lambda \otimes (S^\lambda)^*$ may contribute more to the variance than others. While Theorem 7 allows us to bound the contribution of those constituent irreducible representations S^μ for which μ_1 and μ_1' are much smaller than n, those which violate this condition could conceivably contribute large terms to the variance estimates. Fortunately, we will see that the total fraction of the space $S^\lambda \otimes (S^\lambda)^*$, dimensionwise, consisting of such S^μ is small with overwhelming probability. Despite this, we cannot preclude the possibility that for a *specific* vector $\mathbf{b}$, the quantity $\mathrm{Var}\,\|\Pi_m \mathbf{b}\|^2$ is large, as $\mathbf{b}$ may project solely into spaces of the type described above. On the other hand, as these troublesome spaces amount to a small fraction of $S^\lambda \otimes (S^\lambda)^*$, only a few $\mathbf{b}$ can have this property, and we will see that this suffices to control the distance in total variation to the uniform distribution.

Specifically, let $0 < c < 1/4$ be a constant, and let $\Lambda = \Lambda_c$ denote the collection of Young diagrams μ with the property that either $\mu_1 \geq (1-c)n$ or $\mu_1' \geq (1-c)n$. We have the following upper bounds on the cardinality of Λ and the dimension of any S^μ with $\mu \in \Lambda$:

Lemma 9. *Let $p(n)$ denote the number of integer partitions of n. Then $|\Lambda| \leq 2cnp(cn)$, and $d^\mu < n^{cn}$ for any $\mu \in \Lambda$.*

Proof. For the first statement, note that removing the top row of a Young diagram μ with $\mu_1 \geq (1-c)n$ gives a Young diagram of size $n - \mu_1 \leq cn$. The number of these is at most $p(cn)$, and summing over all such μ_1 gives $cnp(cn)$. The case $\mu'_1 \geq (1-c)n$ is similar, and summing the two gives $|\Lambda| \leq 2cnp(cn)$.

Now let $\mu \in \Lambda$ with $\mu_1 \geq (1-c)n$. By the hook-length formula, since the ith cell from the right in the top row has $\text{hook}(c) \geq i$, $d^\mu < n!/\mu_1! \leq n!/((1-c)n)! \leq n^{cn}$. The case $\mu'_1 \geq (1-c)n$ is similar. □

As a result, the representations associated with diagrams in Λ constitute a negligible fraction of $\widehat{S_n}$; specifically, from Lemma 6, part 2, the probability that a λ drawn according to the Plancherel distribution falls into Λ is $n^{-\Omega(n)}$. The following lemma shows that this is also true for the distribution $P(\rho)$ induced on $\widehat{S_n}$ by weak Fourier sampling the coset state $|H\rangle$.

Lemma 10. *Let $d < 1/2$ be a constant and let n be sufficiently large. Then there is a constant $\gamma > 0$ such that we observe a representation S^λ with $d^\lambda \geq n^{dn}$ with probability at least $1 - n^{-\gamma n}$.*

Proof. The proof is nearly identical to that of Lemma 6, and follows an argument from [9]. Recall from (3.7) in Section 3 that we observe an irreducible ρ with probability $P(\rho) = (d_\rho |H|/|G|)\mathbf{rk}\ \pi_H$. Given $|H| = 2$, $|G| = n!$, and $\mathbf{rk}\ \pi_H \leq d_\rho$, we have $P(\rho) \leq 2d_\rho^2/n!$. Since the total number of irreducibles of S_n is $p(n)$ and $n! > n^n e^{-n}$, we have

$$\sum_{S^\lambda : d^\lambda < n^{dn}} P(S^\lambda) < 2p(n)n^{2dn}/n! < 2p(n)e^n n^{(2d-1)n}$$

and setting $\gamma < 1 - 2d$ completes the proof. □

On the other hand, for a representation S^μ with $\mu \notin \Lambda$, Theorem 7 implies that

$$\left| \frac{\chi^\mu(M)}{d^\mu} \right| \leq \big(\max(q, 1-c)\big)^{bn} \leq e^{-\alpha n} \tag{6.2}$$

for a constant $\alpha \geq bc > 0$. Thus the contribution of such an irreducible to the variance estimate of Lemma 3 is exponentially small. In addition, let E_0 be the event that $d^\lambda \geq n^{dn}$ as in Lemma 10; then conditioning on E_0 and setting d such that $c < 1/4 < d$, Lemma 9 implies that $\lambda \notin \Lambda$, and Equations (4.4) and (6.2) imply

$$\frac{d^\lambda}{2}\left(1 - e^{-\alpha n}\right) \leq \mathbf{rk}\,\Pi_m \leq \frac{d^\lambda}{2}\left(1 + e^{-\alpha n}\right) \ . \tag{6.3}$$

We turn now to the problem of bounding the multiplicities with which representations S^μ, for $\mu \in \Lambda$, can appear in $S^\lambda \otimes (S^\lambda)^*$. While no explicit decomposition is known for $S^\lambda \otimes (S^\lambda)^*$, the *endomorphism representations* of S_n, we record a coarse bound below which will suffice for our purposes. Recall that character of $S^\lambda \otimes (S^\lambda)^*$ is $\chi^\lambda \cdot (\chi^\lambda)^* = (\chi^\lambda)^2$ as characters of S_n are real. The multiplicity of the representation S^μ in $S^\lambda \otimes (S^\lambda)^*$ is $\langle \chi^\mu, (\chi^\lambda)^2 \rangle_G$. However, this is equal to $\langle \chi^\mu \chi^\lambda, \chi^\lambda \rangle_G$, the multiplicity of S^λ in the representation $S^\mu \otimes S^\lambda$. Counting dimensions, this is clearly no more than $\dim(S^\mu \otimes S^\lambda)/\dim S^\lambda = d^\mu$. Hence the multiplicity of S^μ in $S^\lambda \otimes (S^\lambda)^*$ is never more than d^μ; we have

$$\langle \chi^\mu, (\chi^\lambda)^2 \rangle_G \leq d^\mu \ . \tag{6.4}$$

Let $L \subset S^\lambda \otimes (S^\lambda)^*$ be the subspace consisting of copies of representations S^μ with $\mu \in \Lambda$, and let Π_L be the projection operator onto this subspace. By Lemma 9, we have

$$\dim L \leq \sum_{\mu \in \Lambda} (d^\mu)^2 \leq 2cnp(cn)n^{2cn} = e^{O(\sqrt{n})} n^{2cn} \ .$$

Note by Lemma 10 with $d > 2c$, $\dim L$ is a vanishingly small fraction of d^λ.

As B is an orthonormal basis for S^λ, the vectors $\{\mathbf{b} \otimes \mathbf{b}^* \mid \mathbf{b} \in B\}$ are mutually orthogonal in $S^\lambda \otimes (S^\lambda)^*$. Therefore,

$$\sum_{\mathbf{b} \in B} \|\Pi_L(\mathbf{b} \otimes \mathbf{b}^*)\|^2 \leq \dim L \ .$$

In particular, if $B_L \subset B$ denotes the set of $\mathbf{b}$ with the property that

$$\|\Pi_L(\mathbf{b} \otimes \mathbf{b}^*)\|^2 \geq e^{-\alpha n} \ ,$$

then, conditioning on E_0, $|B_L|$ cannot be larger than $e^{\alpha n} \dim L = e^{O(n)} n^{2cn} = n^{-\Omega(n)} d^\lambda$. For a basis vector $\mathbf{b} \notin B_L$, we apply Lemma 3 to bound the variance as follows: assuming pessimistically that $\chi^\mu(M)/d^\mu = 1$ for all $\mu \in \Lambda$, and applying (6.2), gives

$$\begin{aligned} \mathrm{Var}_m \|\Pi_m \mathbf{b}\|^2 &\leq \frac{1}{4}\left[\sum_{S^\mu : \mu \in \Lambda} \left\|\Pi_\mu^\lambda(\mathbf{b} \otimes \mathbf{b}^*)\right\|^2 + \sum_{S^\mu : \mu \notin \Lambda} \frac{\chi^\mu(M)}{d^\mu} \left\|\Pi_\mu^\lambda(\mathbf{b} \otimes \mathbf{b}^*)\right\|^2\right] \\ &\leq \frac{1}{4}\left[e^{-\alpha n} + e^{-\alpha n} \sum_{S^\mu : \mu \notin \Lambda} \left\|\Pi_\mu^\lambda(\mathbf{b} \otimes \mathbf{b}^*)\right\|^2\right] \leq \frac{1}{2} e^{-\alpha n} \ . \end{aligned}$$

In this case, by Chebyshev's inequality,

$$\Pr\left[\left|\|\Pi_m \mathbf{b}\|^2 - \mathrm{Exp}_m \|\Pi_m \mathbf{b}\|^2\right| \geq e^{-\alpha n/3}\right] \leq e^{-\alpha n/3} \ . \tag{6.5}$$

We say that a basis vector is *bad* if this bound is violated, i.e.,

$$\left|\|\Pi_m \mathbf{b}\|^2 - \mathrm{Exp}_m \|\Pi_m \mathbf{b}\|^2\right| \geq e^{-\alpha n/3} \ .$$

Let B_{bad} denote the subset of B consisting of bad basis vectors (observe that while B_L depends only on the choice of λ, B_{bad} depends also on m). Then (6.5) implies

$$\mathrm{Exp}_m |B_{\text{bad}}| \leq e^{-\alpha n/3} d^\lambda.$$

Now let E_1 be the event that

$$|B_{\text{bad}}| < e^{-\alpha n/6} d^\lambda \ ;$$

then by Markov's inequality, E_1 occurs with probability at least $1 - e^{-\alpha n/6}$.

Now let us separate $\|P_m - U\|_1$ into contributions from basis vectors outside and inside $B_L \cup B_{\text{bad}}$:

$$\|P_m - U\|_1 = \sum_{\mathbf{b} \notin B_L \cup B_{\text{bad}}} |P_m(\mathbf{b}) - U| + \sum_{\mathbf{b} \in B_L \cup B_{\text{bad}}} |P_m(\mathbf{b}) - U| \ . \tag{6.6}$$

The first sum is taken only over vectors $\mathbf{b}$ for which

$$\left|\|\Pi_m \mathbf{b}\|^2 - \mathrm{Exp}_m \|\Pi_m \mathbf{b}\|^2\right| < e^{-\alpha n/3} \ .$$

Conditioning on E_0 and recalling that $P_m(\mathbf{b}) = \|\Pi_m \mathbf{b}\|^2 / \mathbf{rk}\,\Pi_m$, the rank estimate of (6.3) gives

$$\sum_{\mathbf{b} \notin B_L \cup B_{\text{bad}}} |P_m(\mathbf{b}) - U| \leq \frac{e^{-\alpha n/3}}{\mathbf{rk}\,\Pi_m} \cdot d^\lambda \leq \frac{2e^{-\alpha n/3}}{1 - e^{-\alpha n}} < 4e^{-\alpha n/3} \ . \tag{6.7}$$

It follows that $P_m(B_L \cup B_{\text{bad}})$ is at most $|B_L \cup B_{\text{bad}}|/d^\lambda + 4e^{-\alpha n/3}$. Therefore, since conditioning on E_0 and E_1 we have $|B_L \cup B_{\text{bad}}| \leq (n^{-\Omega(n)} + e^{-\alpha n/6}) d^\lambda < 2e^{-\alpha n/6} d^\lambda$, the second sum in (6.6) is at most

$$\sum_{\mathbf{b} \in B_L \cup B_{\text{bad}}} \left|P_m(\mathbf{b}) - \frac{1}{d^\lambda}\right| \leq P_m(B_L \cup B_{\text{bad}}) + \frac{|B_L \cup B_{\text{bad}}|}{d^\lambda} < 4e^{-\alpha n/6} + 4e^{-\alpha n/3} < 5e^{-\alpha n/6} \ . \tag{6.8}$$

Then combining (6.6), (6.7) and (6.8),

$$\|P_m - U\|_1 < 4e^{-\alpha n/3} + 5e^{-\alpha n/6} < 6e^{-\alpha n/6}$$

with probability at least $\Pr[E_0 \wedge E_1] \geq 1 - n^{-\gamma n} - e^{-\alpha n/6} \geq 1 - 2e^{-\alpha n/6}$. We complete the proof by setting $\delta < \alpha/6$. □

6.2 Arbitrary POVMs

We now generalize the proof of Theorem 8 to the case where the algorithm designer is allowed to choose an arbitrary finite frame $B = \{\mathbf{b}\}$ of unit length vectors in S^λ, with a family of positive real weights $a_\mathbf{b}$, that satisfy the completeness condition

$$\sum_{\mathbf{b}} a_\mathbf{b} \left|\mathbf{b}\right\rangle \left\langle\mathbf{b}\right| = \mathbb{1} \ . \tag{6.9}$$

(Note that this is simply (3.3) where we have written $\mathbf{b}$ and $a_\mathbf{b}$ instead of $\mathbf{b}_j$ and a_j.)

Theorem 11. *Let $B = \{\mathbf{b}\}$ be a frame with weights $\{a_\mathbf{b}\}$ satisfying the completeness condition* (6.9) *for an irreducible representation S^λ. Given the hidden subgroup $H = \{1, m\}$ where m is chosen uniformly at random from M, let $P_m(\mathbf{b})$ be the probability that we observe the vector $\mathbf{b}$ conditioned on having observed the representation name S^λ, and let N be the natural distribution* (3.9) *on B. Then there is a constant $\delta > 0$ such that for sufficiently large n, with probability at least $1 - e^{-\delta n}$ in m and λ, we have*

$$\|P_m - N\|_1 < e^{-\delta n} \ .$$

Proof. Recall from (3.8) in Section 3 that the conditional distribution on B is given by

$$P_m(\mathbf{b}) = P(S^\lambda, \mathbf{b}) = a_\mathbf{b} \frac{\|\Pi_m \mathbf{b}\|^2}{\mathbf{rk}\, \Pi_m}$$

and the natural distribution (3.9) is given by $N(\mathbf{b}) = a_\mathbf{b}/d^\lambda$.

The proof of Theorem 8 goes through with a few modifications. First, let us change some semantics: given a subset $A \subseteq B$, we let $|A|$ denote the weighted size of A,

$$|A| = \sum_{\mathbf{b} \in A} a_\mathbf{b} \ .$$

With this definition, the total probability that falls in A under the natural distribution is $N(A) = |A|/d^\lambda$. Then we will use the following lemma:

Lemma 12. *Let L be a subspace of $S^\lambda \otimes (S^\lambda)^*$ and let Π_L be the projection operator onto L. Then*

$$\sum_{\mathbf{b} \in B} a_\mathbf{b} \left\|\Pi_L(\mathbf{b} \otimes \mathbf{b}^*)\right\|^2 \le \dim L \ . \tag{6.10}$$

Proof. First note that a vector $\mathbf{e} \in S^\lambda \otimes (S^\lambda)^*$ has entries $\mathbf{e}_{j,k}$ for $1 \le j, k \le d^\lambda$. There is a unique linear operator E on S^λ whose matrix entries are $E_{j,k} = \mathbf{e}_{j,k}$, and the inner product $\langle \mathbf{b} \otimes \mathbf{b}^*, \mathbf{e} \rangle$ in $S^\lambda \otimes (S^\lambda)^*$ can then be written as the bilinear form $\langle \mathbf{b}, E\mathbf{b} \rangle$ in S^λ. The Frobenius norm of E is $\|E\|^2 = \mathbf{tr}\, E^\dagger E = \|\mathbf{e}\|^2$.

Now let $\{\mathbf{e}_i\}$ be an orthonormal basis for L and let E_i be the operator corresponding to $\mathbf{e}_i$. Then

$$\begin{aligned}
\sum_{\mathbf{b} \in B} a_\mathbf{b} \left|\langle \mathbf{b} \otimes \mathbf{b}^*, \mathbf{e}_i \rangle\right|^2 &= \sum_{\mathbf{b} \in B} a_\mathbf{b} \left|\langle \mathbf{b}, E_i \mathbf{b} \rangle\right|^2 \le \sum_{\mathbf{b} \in B} a_\mathbf{b} \|\mathbf{b}\|^2 \|E_i \mathbf{b}\|^2 = \sum_{\mathbf{b} \in B} a_\mathbf{b} \|E_i \mathbf{b}\|^2 \\
&= \sum_{\mathbf{b} \in B} a_\mathbf{b}\, \mathbf{tr} \left(E_i^\dagger \left|\mathbf{b}\right\rangle \left\langle\mathbf{b}\right| E_i \right) = \mathbf{tr} \left[E_i^\dagger \left(\sum_{\mathbf{b} \in B} a_\mathbf{b} \left|\mathbf{b}\right\rangle \left\langle\mathbf{b}\right| \right) E_i \right] = \mathbf{tr}\, E_i^\dagger E_i = \|\mathbf{e}_i\|^2 = 1
\end{aligned}$$

where we used the Cauchy-Schwartz inequality in the second line and completeness in the second. Summing over the $\dim L$ basis vectors $\mathbf{e}_i$ then gives (6.10). □

We define Λ and E_0 as before, and Lemmas 9 and 10 still apply. As before, let $L \subset S^\lambda \otimes (S^\lambda)^*$ be the subspace consisting of copies of representations S^μ with $\mu \in \Lambda$, and let $B_L \subset B$ denote the set of $\mathbf{b}$ with the property that

$$\left\|\Pi_L(\mathbf{b} \otimes \mathbf{b}^*)\right\|^2 \ge e^{-\alpha n} \ .$$

Then Lemma 12 implies that

$$|B_L| \leq e^{\alpha n} \dim L$$

where $|B_L|$ is defined as above. We again define B_{bad} as the set of $\mathbf{b} \in B \setminus B_L$ such that

$$\left| \|\Pi_m \mathbf{b}\|^2 - \text{Exp}_m \|\Pi_m \mathbf{b}\|^2 \right| \geq e^{-\alpha n/3}$$

and Chebyshev's and Markov's inequalities imply that the event E_1, namely

$$|B_{\text{bad}}| < e^{-\alpha n/6} d^\lambda \ ,$$

occurs with probability at least $1 - e^{-\alpha n/6}$.

We separate $\|P_m - N\|_1$ as we did $\|P_m - U\|_1$ before:

$$\|P_m - N\|_1 = \sum_{\mathbf{b} \notin B_L \cup B_{\text{bad}}} |P_m(\mathbf{b}) - N(\mathbf{b})| + \sum_{\mathbf{b} \in B_L \cup B_{\text{bad}}} |P_m(\mathbf{b}) - N(\mathbf{b})| \ . \tag{6.11}$$

Since $\left| \|\Pi_m \mathbf{b}\|^2 - \text{Exp}_m \|\Pi_m \mathbf{b}\|^2 \right| < e^{-\alpha n/3}$ for all $\mathbf{b} \notin B_L \cup B_{\text{bad}}$, and since $\sum_{\mathbf{b}} a_{\mathbf{b}} = d^\lambda$, conditioning on E_0 and using (6.3) bounds the first sum as follows,

$$\sum_{\mathbf{b} \notin B_L \cup B_{\text{bad}}} |P_m(\mathbf{b}) - N(\mathbf{b})| \leq \frac{e^{-\alpha n/3}}{\mathbf{rk}\, \Pi_m} \cdot d^\lambda \leq \frac{2e^{-\alpha n/3}}{1 - e^{-\alpha n}} < 4e^{-\alpha n/3} \ . \tag{6.12}$$

It follows that $P_m(B_L \cup B_{\text{bad}})$ is at most $|B_L \cup B_{\text{bad}}|/d^\lambda + 4e^{-\alpha n/3}$. Conditioning on E_0 and E_1, we have $|B_L \cup B_{\text{bad}}| \leq (n^{-\Omega(n)} + e^{-\alpha n/6}) d^\lambda < 2e^{-\alpha n/6} d^\lambda$. Thus the second sum in (6.11) is at most

$$\sum_{\mathbf{b} \in B_L \cup B_{\text{bad}}} |P_m(\mathbf{b}) - N(\mathbf{b})| \leq P_m(B_L \cup B_{\text{bad}}) + N(B_L \cup B_{\text{bad}}) < 4e^{-\alpha n/6} + 4e^{-\alpha n/3} < 5e^{-\alpha n/6} \ . \tag{6.13}$$

Then combining (6.11), (6.12) and (6.13),

$$\|P_m - N\|_1 < 4e^{-\alpha n/3} + 5e^{-\alpha n/6} < 6e^{-\alpha n/6}$$

with probability at least $\Pr[E_0 \wedge E_1] \geq 1 - n^{-\gamma n} - e^{-\alpha n/6} \geq 1 - 2e^{-\alpha n/6}$. We complete the proof by setting $\delta < \alpha/6$ as before. □

Acknowledgments.

This work was supported by NSF grants CCR-0093065, PHY-0200909, EIA-0218443, EIA-0218563, CCR-0220070, and CCR-0220264. We are grateful to Denis Thérien, McGill University, and Bellairs Research Institute for organizing a workshop at which this work began; to Dorit Aharonov, Daniel Rockmore, and Umesh Vazirani for helpful conversations; Chris Lomont for pointing out several typos; and to Tracy Conrad and Sally Milius for their support and tolerance. C.M. also thanks She Who Is Not Yet Named for her impending arrival, and for providing a larger perspective.

References

[1] David Bacon, Andrew Childs, and Wim van Dam. Optimal measurements for the dihedral hidden subgroup problem. Preprint, quant-ph/0501044 (2005).

[2] Robert Beals. Quantum computation of Fourier transforms over symmetric groups. *Proc. 29th Annual ACM Symposium on the Theory of Computing*, pages 48–53, 1997.

[3] Ethan Bernstein and Umesh Vazirani. Quantum complexity theory (preliminary abstract). *Proc. 25th Annual ACM Symposium on the Theory of Computing*, pages 11–20, 1993.

[4] Wim van Dam, Sean Hallgren, and Lawrence Ip. Quantum algorithms for some hidden shift problems. *Proc. 14th ACM-SIAM Symposium on Discrete Algorithms*, pages 489–498, 2003.

[5] Mark Ettinger and Peter Høyer. On quantum algorithms for noncommutative hidden subgroups. Preprint, quant-ph/9807029 (1998).

[6] Mark Ettinger and Peter Høyer and Emmanuel Knill. The quantum query complexity of the hidden subgroup problem is polynomial. *Information Processing Letters*, to appear.

[7] Katalin Friedl, Gábor Ivanyos, Frédéric Magniez, Miklos Santha, and Pranab Sen. Hidden translation and orbit coset in quantum computing. *Proc. 35th ACM Symposium on Theory of Computing*, 2003.

[8] William Fulton and Joe Harris. *Representation Theory: A First Course.* Number 129 in Graduate Texts in Mathematics. Springer-Verlag, 1991.

[9] Michelangelo Grigni, Leonard J. Schulman, Monica Vazirani, and Umesh Vazirani. Quantum mechanical algorithms for the nonabelian hidden subgroup problem. *Proc. 33rd ACM Symposium on Theory of Computing*, pages 68–74, 2001.

[10] Lisa Hales and Sean Hallgren. Quantum Fourier sampling simplified. *Proc. 31st Annual ACM Symposium on Theory of Computing*, 1999.

[11] Lisa Hales and Sean Hallgren. An improved quantum Fourier transform algorithm and applications. *Proc. 41st Annual Symposium on Foundations of Computer Science*, 2000.

[12] Sean Hallgren, Alexander Russell, and Amnon Ta-Shma. Normal subgroup reconstruction and quantum computation using group representations. *Proc. 32nd ACM Symposium on Theory of Computing*, pages 627–635, 2000.

[13] Peter Høyer. Efficient quantum transforms. Preprint, quant-ph/9702028 (1997).

[14] Yoshifumi Inui and François Le Gall. An efficient algorithm for the hidden subgroup problem over a class of semi-direct product groups. Proc. EQIS 2004.

[15] Lawrence Ip. Shor's algorithm is optimal. Preprint, 2004.

[16] Gábor Ivanyos, Frédéric Magniez, and Miklos Santha. Efficient quantum algorithms for some instances of the non-abelian hidden subgroup problem. *Int. J. Found. Comput. Sci.* 14(5): 723–740, 2003.

[17] Richard Jozsa. Quantum factoring, discrete logarithms and the hidden subgroup problem. Preprint, quant-ph/0012084 (2000).

[18] Julia Kempe and Aner Shalev. The hidden subgroup problem and permutation group theory. Preprint, quant-ph/0406046 (2004).

[19] S. V. Kerov. *Asymptotic representation theory of the symmetric group and its applications in analysis.* Translated by N. V. Tsilevich. Volume 219 in Translations of Mathematical Monographs. American Mathematical Society, 2003.

[20] Greg Kuperberg. A subexponential-time quantum algorithm for the dihedral hidden subgroup problem. Preprint, quant-ph/0302112 (2003).

[21] Cristopher Moore, Daniel Rockmore, and Alexander Russell. Generic quantum Fourier transforms. *Proc. 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 778–787, 2004.

[22] Cristopher Moore, Daniel Rockmore, Alexander Russell, and Leonard Schulman. The value of basis selection in Fourier sampling: hidden subgroup problems for affine groups. *Proc. 15th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1113–1122, 2004.

[23] Oded Regev. Quantum computation and lattice problems. *Proc. 43rd Symposium on Foundations of Computer Science*, pages 520–530, 2002.

[24] Martin Roetteler and Thomas Beth. Polynomial-time solution to the hidden subgroup problem for a class of non-abelian groups. Preprint, quant-ph/9812070 (1998).

[25] Yuval Roichman. Upper bound on the characters of the symmetric groups. *Inventiones Mathematicae*, 125:451–485, 1996.

[26] Steven Roman. *Advanced Linear Algebra.* Number 135 in Graduate Texts in Mathematics. Springer, 1992.

[27] Jean-Pierre Serre. *Linear Representations of Finite Groups.* Number 42 in Graduate Texts in Mathematics. Springer-Verlag, 1977.

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

[29] Daniel R. Simon. On the power of quantum computation. *SIAM Journal on Computing*, 26(5):1474–1483, 1997.

[30] A. M. Vershik and S. V. Kerov. Asymptotic behavior of the maximum and generic dimensions of irreducible representations of the symmetric group. Funk. Anal. i Prolizhen, 19(1):25–36, 1985; English translation, Funct. Anal. Appl., 19:21–31, 1989.