Limitations of quantum coset states for graph isomorphism

Sean Hallgren* Cristopher Moore† Martin Rötteler‡ Alexander Russell§
Pranab Sen¶

Abstract

It has been known for some time that graph isomorphism reduces to the hidden subgroup problem (HSP). What is more, most exponential speedups in quantum computation are obtained by solving instances of the HSP. A common feature of the resulting algorithms is the use of quantum coset states, which encode the hidden subgroup. An open question has been how hard it is to use these states to solve graph isomorphism. It was recently shown by Moore, Russell, and Schulman [MRS05] that only an exponentially small amount of information is available from one, or a pair of coset states. A potential source of power to exploit are entangled quantum measurements that act jointly on many states at once.

We show that entangled quantum measurements on at least $\Omega(n \log n)$ coset states are necessary to get useful information for the case of graph isomorphism, matching an information theoretic upper bound. This may be viewed as a negative result because in general it seems hard to implement a given highly entangled measurement. Our main theorem is very general and also rules out using joint measurements on few coset states for some other groups, such as $\mathrm{GL}(n, \mathbb{F}_{p^m})$ and G^n where G is finite and satisfies a suitable property.

1 Introduction

Most exponential speedups that have been achieved in quantum computing are obtained by solving some instances of the Hidden Subgroup Problem (HSP). In particular, the problems underlying Shor's algorithms for factoring and discrete logarithm [Sho97], as well as Simon's problem [Sim94], can be naturally generalized to the HSP: given a function $f : G \to S$ from a group G to a set S that is constant on left cosets of some subgroup $H \leq G$ and distinct on different cosets, find a set of generators for H. Ideally, we would like to find H in time polynomial in the input size, i. e. $\log |G|$. The abelian HSP [Kit95, BH97, ME98], i. e., when G is an abelian group, lies at the heart of efficient quantum algorithms for important number-theoretic problems like factoring, discrete logarithm, Pell's equation, unit group of a number field etc. [Sho97, Hal02, Hal05, SV05].

It has been known for some time that graph isomorphism reduces to the HSP over the symmetric group [Bea97, EHK99a], a non-abelian group. While the non-abelian HSP has received much attention as a result, efficient algorithms are known only for some special classes of groups [IMS03, HRT03, FIM$^+$03,

*Department of Computer Science and Engineering, Pennsylvania State University, University Park, PA, U.S.A. Email: hallgren@cse.psu.edu

†Department of Computer Science, University of New Mexico, Albuquerque, NM, U.S.A. Email: moore@cs.unm.edu

‡NEC Laboratories America, Inc., Princeton, NJ, U.S.A. Email: mroetteler@nec-labs.com

§Department of Computer Science and Engineering, University of Connecticut, Storrs, CT, U.S.A. Email: acr@cse.uconn.edu

¶School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai, India. Email: pgdsen@tcs.tifr.res.in

MRRS07, Gav04, BCD05, ISS07]. On the other hand, the HSP presents a systematic way to try and approach the graph isomorphism problem, and this approach is rooted in developing a deeper understanding of how far techniques and tools that have worked in the abelian case can be applied. To the best of our knowledge, the only other approach to solve graph isomorphism on a quantum computer is by creating a uniform superposition of all graphs isomorphic to a given graph. It has been proposed to create this superposition via quantum sampling of Markov chains [AT03], however, very little is known about this.

One of the key features of a quantum computer is that it can compute functions in superposition. This fact alone does not lend itself to exponential speedups, for instance for unstructured search problems it merely leads to a polynomial speedup [Gro96, BBBV97]. On the other hand, the quantum states resulting from HSP instances have far more structure since they capture some periodicity aspects of the function f. Evaluating the function f in superposition and ignoring the function value results in a random *coset state*. Coset states are quantum states of the form $|gH\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |gh\rangle$, in other words, a coset state is a uniform superposition over the elements of the left coset gH. The challenge in using coset states lies in the fact that g is a random element of the group, beyond our control, that is, we only have the mixed state $\sigma_H^G = \frac{1}{|G|} \sum_{g \in G} |gH\rangle\langle gH|$ and we have to determine H from it. Though it is conceivable that some advantage can be had by making use of the function values, currently there are no proposals for using function values in any meaningful way.

How much information can be extracted from coset states? The most general way to extract classical information from quantum states are POVMs [NC00]. A fixed POVM operates on a fixed number k of coset states at once. This induces a probability distribution over the set of classical outcomes associated with the POVM. A potential source of power with no classical analog is that the probability distribution induced by a POVM acting jointly on k coset states (i. e., an entangled POVM) may have significantly more information about the hidden subgroup than a distribution induced by a POVM that acts on these k coset states just one state at a time. The goal of this paper is to determine how small k can be made such that a polynomial amount of information about H can be obtained from any POVM on k coset states. More precisely, we want to know how small k can be so that there exists a POVM on k coset states that gives polynomially large total variation distance between every pair of candidate hidden subgroups. Note that this POVM can have many classical outcomes, and it may have to be repeated a polynomial number of times if we want to identify the actual hidden subgroup H with constant probability.

In this paper, we show that for many groups G this number k, which we call the *jointness* of a measurement, has to be quite large, sometimes as large as $\Omega(\log |G|)$. This matches the information theoretic $O(\log |G|)$ upper bound for general groups [EHK99b] to within constant multiplicative factors. Our result may be viewed as a negative result because in general it seems hard to implement a given highly entangled measurement, including the ones that arise from the HSP. Note that the time required to perform a generic measurement entangled across k states increases exponentially with k. Following are three instances of highly entangled measurements that have been proposed for the HSP but which are not known to be efficiently implementable: the measurement proposed by Ettinger, Høyer and Knill [EHK99b], the pretty good measurement approach [BCD06, BCD05], and the missing harmonic measurement [MR05a, AMR07].

For abelian groups the picture simplifies dramatically. Indeed, in this case a POVM operating on one coset state (i. e., with jointness $k = 1$) exists that gives a polynomial amount of information about the hidden subgroup. Moreover, this measurement is efficiently implementable using the quantum Fourier transform over the group. The Fourier based approach extends to some non-abelian groups as well, e. g., dihedral, affine, Heisenberg groups and more generally Gel'fand pairs, and shows that for these groups there are measurements on single coset states that give polynomially large information about the hidden subgroup [EH00, MRRS07, RRS09].

Except for the general information theoretic upper bound, only a few examples of measurements operating jointly on more than one coset state (i. e., jointness $k > 1$) are known that give a polynomial amount of information about the hidden subgroup. Kuperberg [Kup05] gave a measurement for the dihedral group operating jointly on $2^{O(\sqrt{\log |G|})}$ coset states that also takes $2^{O(\sqrt{\log |G|})}$ time to implement. Bacon et al. [BCD05] gave an efficiently implementable measurement for the Heisenberg group operating jointly on two coset states, and similar efficient measurements for some other groups operating jointly on a constant number of coset states.

The case of the symmetric group S_n has been much harder to understand. First it was shown that some restricted measurements related to the abelian case cannot solve the problem [HRT03]. Next the non-abelian aspects of the group were attacked by Grigni et al. [GSVV04] who showed that for hidden subgroups in S_n, measuring the Fourier transform of a single coset state using random choices of bases for the representations of S_n gives exponentially little information. They left open the question whether a clever choice of basis for each representation space can indeed give enough information about the hidden subgroup. Recently, major progress was made by Moore, Russell and Schulman [MRS05] who answered this question in the negative for $k = 1$ by showing that any measurement on a single coset state of S_n gives exponentially little information, i. e., any algorithm for the HSP in S_n that measures one coset state at a time requires at least $\exp(\Omega(n))$ coset states. Subsequently, Moore and Russell [MR05b] extended this result by showing that any algorithm that jointly measures two coset states at a time gives sub-polynomial amount of information, more precisely, any algorithm for the HSP in S_n that measures two coset states at a time requires at least $\exp(\Omega(\sqrt{n}/\log n))$ coset states. However, their techniques fail for algorithms that jointly measure three or more coset states at a time, and they left the $k \geq 3$ case open.

In this paper, we show that no quantum measurement on $k = o(n \log n)$ coset states can extract polynomial amount of information about the hidden subgroup in S_n. Thus, any algorithm operating on coset states that solves the hidden subgroup problem in S_n in polynomial time has to either use $\Omega(n \log n)$ *qubits* of work space irrespective of the amount of classical work space, or has to make joint measurements on $\Omega(n \log n)$ coset states. Our lower bound matches the information theoretic upper bound of Ettinger, Høyer and Knill [EHK04] to within constant multiplicative factors. Our results apply to the hidden subgroups arising out of the reduction from isomorphism of rigid graphs, and rules out any efficient quantum algorithm that tries to solve graph isomorphism via the standard reduction to the HSP in S_n, using less than $\Theta(n \log n)$ qubits of work space and measurements that act jointly on less than $\Theta(n \log n)$ coset states at a time.

As the results we present frustrate certain natural approaches to efficiently solving these hidden subgroup problems, they suggest an attractive family of hardness assumptions on which one might base cryptographic constructions. An immediate consequence of Shor's efficient algorithms for integer factorization and discrete logarithm is that most public-key cryptosystems in use are manifestly insecure in the face of quantum adversaries. One strategy for remedying this is to construct cryptosystems from (presumably) difficult instances of hidden subgroup problems. One such proposal [Reg04a, Reg04b] hinges on the assumed hardness of the dihedral hidden subgroup problem. The results of this article, however, demonstrate a quantitative difference between the the dihedral hidden subgroup problem, for which single-register Fourier sampling suffices to (information-theoretically) determine the subgroup, and more "nonabelian" groups for which rich joint measurements are required. Indeed, a recent article [MRV07] applies the conclusions of this article as evidence of a proposed cryptosystem related to the HSP over $\mathrm{GL}(n, \mathbb{F}_q) \wr \mathbb{Z}_2$.

Our lower bound on the jointness of a measurement for the HSP holds for a more general setting: Given a group G, suppose we want to decide if the hidden subgroup is a conjugate of an a priori known order two subgroup H, or the identity subgroup. We show a lower bound on the jointness k of any measurement on coset states of the hidden subgroup that distinguishes between the above two cases. Our main theorem uses

only properties of G that can be read off from the values of the characters at the two elements of H. We also prove a *transfer lemma* that allows us to transfer lower bounds proved for the HSP in a group G to the HSP in some other group $\tilde{G}$ that is related to G in a suitable way. Using our main theorem and the transfer lemma, we show lower bounds on the jointness of measurements for the HSP in groups $\mathrm{PSL}(2, \mathbb{F}_{p^m})$, $\mathrm{GL}(n, \mathbb{F}_{p^m})$, and groups of the form G^n, where G a constant-sized group satisfying a suitable property.

Recently, Childs and Wocjan [CW07] proposed a *hidden shift* approach to graph isomorphism. They established a lower bound for the total number of hidden shift states required and also showed that a single hidden shift state contains exponentially little information about the isomorphism. Our results generalize both their bounds and imply $o(n \log n)$ hidden shift states contain exponentially little information about the isomorphism.

The chief technical innovation required to prove our main theorem is an improved upper bound for the second moment of the probability of observing a particular measurement outcome as we vary over different candidate hidden subgroups. In particular, we give a new and improved analysis of the projection lengths of vectors of the form $\mathbf{b} \otimes \mathbf{b}$ onto homogeneous spaces of irreducible representations of a group. The earlier works [MRS05, MR05b] tried to bound these projection lengths using simple geometric methods. As a result, their methods failed beyond $k = 2$ for the symmetric group. Instead, we make crucial use of the representation-theoretic structure of the projection operators as well as the structure of the vectors, in order to prove upper bounds on the projection lengths better than those obtainable by mere geometry. This allows us to prove a general theorem that applies with large k for many groups.

Finally, we also prove a simple lower bound on the total number of coset states required by any algorithm to solve the HSP in a group G. This lower bound gives a simple proof of the fact that distinguishing a hidden reflection from the identity subgroup in the dihedral group D_n requires $\Omega(\log n)$ coset states.

Subsequent work: Subsequent to the conference version of this work [HMR$^+$06], Moore, Russell and Śniady [MRŚ07] showed that a sieve approach à la Kuperberg [Kup05] for solving the HSP problem arising from the reduction of rigid-graph isomorphism requires superpolynomially many coset states. The sieve approach is one example of a systematic procedure for performing highly entangled measurements on coset states, and Moore, Russell and Śniady's work rules it out as an efficient technique for the HSP corresponding to rigid-graph isomorphism.

2 Preliminaries

2.1 Basic facts about quantum states

In this paper, all groups and sets are finite, all Hilbert spaces finite dimensional, and all measurements will have finitely many outcomes. A general quantum state in $\mathbb{C}^n$ is modeled by a so-called *density matrix*, which is an $n \times n$ Hermitian positive semidefinite matrix with unit trace. The most general way to obtain classical information from a quantum state is via a generalized measurement, also known as a positive operator-valued measure or POVM [NC00]. The elements of a POVM $\mathcal{M}$ in $\mathbb{C}^n$ are finitely many $n \times n$ Hermitian positive semidefinite operators E_i which have to satisfy the completeness condition $\sum_i E_i = \mathbb{1}_n$. If the state of the quantum system is given by the density matrix σ, then the probability p_i to observe outcome labeled i is given by the *Born rule* $p_i = \mathrm{Tr}(\sigma E_i)$.

The *total variation distance*, also known as the ℓ_1*-distance*, between two vectors $v, w \in \mathbb{C}^n$ is defined as $\|v - w\|_1 := \sum_{i=1}^n |v_i - w_i|$. The *trace norm* of a square matrix A is defined as $\|A\|_{\mathrm{tr}} := \mathrm{Tr}\sqrt{A^\dagger A}$. Observe that for any vector $v \in \mathbb{C}^n$, $\|\mathrm{diag}(v)\|_{\mathrm{tr}} = \|v\|_1$, where $\mathrm{diag}(v)$ is the $n \times n$ matrix with

v on the diagonal and zeroes elsewhere. The trace distance between two quantum states ρ, σ in the same Hilbert space is an upper bound on the total variation distance between the two probability distributions $\mathcal{M}[\rho]$, $\mathcal{M}[\sigma]$ obtained by performing the same POVM $\mathcal{M}$ on ρ, σ, that is, for any POVM $\mathcal{M}$, $\|\mathcal{M}[\rho] - \mathcal{M}[\sigma]\|_1 \leq \|\rho - \sigma\|_{\mathrm{tr}}$. Moreover, given any two states ρ, σ, there is in fact a two-outcome POVM $\mathcal{M}'$ such that $\|\mathcal{M}'[\rho] - \mathcal{M}'[\sigma]\|_1 = \|\rho - \sigma\|_{\mathrm{tr}}$ [AKN98].

The POVM $\mathcal{M}'$ of the previous paragraph generally depends upon the two quantum states ρ, σ to be distinguished. The following fact states the existence of a *single* POVM $\mathcal{F}$ that distinguishes somewhat well between any pair of states of an ensemble [RRS09]. Observe that typically $\mathcal{F}$ will do a much worse job at distinguishing between a specific pair of states ρ, σ from the ensemble compared to the tailor made POVM $\mathcal{M}'$ for ρ, σ. This is unavoidable in general, as evidence by the work of Moore, Russell and Schulman [MRS05] and the present work on hidden subgroups of the symmetric group.

Fact 1. *Let $\mathcal{E} := \{\rho_i\}_i$ be a finite ensemble of quantum states in $\mathbb{C}^n$. Then there is a POVM $\mathcal{F}$ such that $\|\mathcal{F}[\rho_i] - \mathcal{F}[\rho_j]\|_1 \geq c\, n^{-1/2} \|\rho_i - \rho_j\|_{\mathrm{tr}}$ for every pair of states $\rho_i, \rho_j \in \mathcal{E}$, where c is a universal constant, $0 < c < 1$.*

The most general quantum operation on a quantum state is described by a so-called *completely positive trace preserving superoperator*. Concretely, it is a $\mathbb{C}$-linear map $\mathcal{A}$ from $n \times n$ complex matrices to $m \times m$ complex matrices that sends quantum states in $\mathbb{C}^n$ to quantum states in $\mathbb{C}^m$ and continues to be so when tensored with the identity quantum operation. Equivalently, $\mathcal{A}$ can be described by a finite ensemble of operators $\{E_i\}_i$ from $\mathbb{C}^n$ to $\mathbb{C}^m$ such that $\sum_i E_i^\dagger E_i = \mathbb{1}_n$; then, $\mathcal{A}[X] := \sum_i E_i X E_i^\dagger$, where X is an $n \times n$ matrix and $\mathcal{A}[X]$ is its image which is an $m \times m$ matrix. It can be shown that a quantum operation $\mathcal{A}$ cannot increase the trace distance between two quantum states, that is, for any two quantum states ρ, σ in the domain of $\mathcal{A}$, $\|\mathcal{A}[\rho] - \mathcal{A}[\sigma]\|_{\mathrm{tr}} \leq \|\rho - \sigma\|_{\mathrm{tr}}$ [NC00].

2.2 Quantum Fourier transform and HSP

We collect some standard facts from representation theory of finite groups; see e.g. the book by Serre [Ser77] for more details. For a group G, we use $\mathbb{C}[G]$ to denote its group algebra, namely, the $|G|$-dimensional vector space spanned by formal $\mathbb{C}$-linear combinations of group elements. For a group element $g \in G$, we use $|g\rangle$ to denote the formal $\mathbb{C}$-linear combination that is 1 at g and 0 at $g' \neq g$. In other words, $\mathbb{C}[G]$ can be thought of as the $\mathbb{C}$-linear space of all functions from G to $\mathbb{C}$ $|g\rangle$ under pointwise addition, and $|g\rangle$ can be thought of as the function that is 1 at g and 0 elsewhere. The vectors $|g\rangle$, $g \in G$ form an orthonormal basis for $\mathbb{C}[G]$ under the standard inner product on $\mathbb{C}[G]$; thus, $\mathbb{C}[G]$ is a $|G|$-dimensional Hilbert space. The group algebra $\mathbb{C}[G]$ also has a multiplicative structure inherited from the group law of G respecting $\mathbb{C}$-linearity.

We use the term irrep to denote an irreducible complex unitary representation of a finite group G and denote by $\widehat{G}$ a complete set of inequivalent irreps. For any unitary representation ρ of G, let ρ^* denote the representation obtained by entry-wise conjugating the unitary matrices $\rho(g)$, where $g \in G$. Note that the definition of ρ^* depends upon the choice of the basis used to concretely describe the matrices $\rho(g)$. If ρ is an irrep of G so is ρ^*, but in general ρ^* may be inequivalent to ρ. Let V_ρ denote the vector space of ρ, define $d_\rho := \dim V_\rho$, and notice that $V_\rho = V_{\rho^*}$. The group elements $|g\rangle$, where $g \in G$ form an orthonormal basis of $\mathbb{C}^{|G|}$. Since $\sum_{\rho \in \widehat{G}} d_\rho^2 = |G|$, we can consider another orthonormal basis called the *Fourier basis* of $\mathbb{C}^{|G|}$ indexed by $|\rho, i, j\rangle$, where $\rho \in \widehat{G}$ and i, j run over the row and column indices of ρ. The quantum Fourier transform over G, QFT_G is the following linear transformation:

$$|g\rangle \mapsto \sum_{\rho \in \widehat{G}} \sqrt{\frac{d_\rho}{|G|}} \sum_{i,j=1}^{d_\rho} \rho_{ij}(g) |\rho, i, j\rangle.$$

It follows from Schur's orthogonality relations (see e.g. [Ser77, Chapter 2, Proposition 4, Corollary 3]) that QFT_G is a unitary transformation in $\mathbb{C}^{|G|}$, where the domain space is $\mathbb{C}[G] \cong \mathbb{C}^{|G|}$ and the range space is the space spanned by formal linear combinations of triples (ρ, i, j) which turns out to be isomorphic to $\mathbb{C}^{|G|}$. For a subgroup $H \leq G$ and irrep $\rho \in \widehat{G}$, define $\rho(H) := \frac{1}{|H|} \sum_{h \in H} \rho(h)$. It follows from Schur's lemma (see e.g. [Ser77, Chapter 2, Proposition 4]) that $\rho(H)$ is an orthogonal projection to the subspace of V_ρ consisting of vectors that are point-wise fixed by every $\rho(h)$, $h \in H$. Define $r_\rho(H) := \mathrm{rank}(\rho(H))$; then $r_\rho(H) = \frac{1}{|H|} \sum_{h \in H} \chi_\rho(h)$, where χ_ρ denotes the character of ρ. Notice that $r_\rho(H) = r_{\rho^*}(H)$. For any subset $S \leq G$ define $|S\rangle := \frac{1}{\sqrt{|S|}} \sum_{s \in S} |s\rangle$ to be the uniform superposition over the elements of S. The *standard method* of attacking the HSP in G using coset states [GSVV04] starts by first forming the uniform superposition $\frac{1}{\sqrt{|G|}} \sum_{g \in G} |g\rangle|0\rangle$. It then queries f to get the superposition $\frac{1}{\sqrt{|G|}} \sum_{g \in G} |g\rangle|f(g)\rangle$. Ignoring the second register the reduced state on the first register becomes the density matrix $\sigma_H^G = \frac{1}{|G|} \sum_{g \in G} |gH\rangle\langle gH|$, that is the reduced state is a uniform mixture over all left coset states of H in G. It can be shown (see e.g. [RRS09]) that applying QFT_G to σ_H^G gives us the density matrix

$$\frac{|H|}{|G|} \bigoplus_{\rho \in \widehat{G}} \bigoplus_{i=1}^{d_\rho} |\rho, i\rangle\langle\rho, i| \otimes \rho^*(H),$$

where $\rho^*(H)$ operates on the space of column indices of ρ. When measuring this state, we obtain an irrep ρ with probability $\frac{d_\rho |H| r_{\rho(H)}}{|G|}$. Conditioned on measuring ρ we obtain a uniform distribution $1/d_\rho$ on the row indices. The reduced state on the space of column indices after having observed an irrep ρ and a row index i is then given by the state $\rho^*(H)/r_\rho(H)$, and a basic task for a hidden subgroup finding algorithm is how to extract information about H from it.

If the the hidden subgroup is the trivial subgroup $\{1\}$, the probability of measuring ρ is given by the so-called *Plancherel distribution* $\mathcal{P}(\rho) := \frac{d_\rho^2}{|G|}$. This distribution will be useful to us later on in the proof of the main theorem.

The following observation is crucial for the HSP case: since the states σ_H^G are simultaneously block diagonal in the Fourier basis for any $H \leq G$, the elements of any POVM $\mathcal{M}$ operating on these states can without loss of generality be assumed to have the same block structure. From this it is clear that any measurement to identify H without loss of generality first applies the quantum Fourier transform QFT_G to σ_H^G, measures the name ρ of an irrep, the index i of a row, and then measures the reduced state on the column space of ρ using a POVM $\mathcal{M}_\rho$ in $\mathbb{C}^{d_\rho}$. This POVM $\mathcal{M}_\rho$ may depend on ρ but is independent of i.

Furthermore, $\mathcal{M}_\rho$ can be assumed to be a frame, i. e., a collection $\mathcal{B}_\rho := \{(a_b, b)\}$, where $b \in \mathbb{C}^{d_\rho}$ with $\|b\| = 1$ and $0 \leq a_b \leq 1$ such that $\sum_{b \in \mathcal{B}} a_b |b\rangle\langle b| = \mathbb{1}_{d_\rho}$ i.e. a frame is a POVM with rank one elements. Orthonormal bases are special cases of frames in which $a_b = 1$ for all $b \in \mathcal{B}_\rho$. We can assume that the POVM on the column space is a frame because any POVM can be refined to a frame such that for any quantum state, the probabilities according to the original POVM are certain sums, independent of the state measured, of probabilities according to the frame, and thus, refining the POVM to a frame cannot decrease the total variation distance between probability distributions obtained by applying the POVM to a pair of quantum states.

If the the hidden subgroup is the trivial subgroup $\{1\}$, after observing an irrep ρ and a row index i, the reduced state on the space of column indices of ρ is the totally mixed state $\frac{\mathbb{1}_{d_\rho}}{d_\rho}$. The probability of observing a vector b in frame $\mathcal{B}_\rho$ is given by the so-called *natural distribution* on $\mathcal{B}_\rho$ defined by $\mathcal{N}(b \mid \rho) := \frac{d_\rho^2}{|G|}$. This distribution will be useful to us later on in the proof of the main theorem.

The above description was for single register quantum Fourier sampling. Fourier sampling on k registers can be defined analogously. Here one starts off with k independent copies of the coset state σ_H^G, i. e., with the state $(\sigma_H^G)^{\otimes k} \cong \sigma_{H^k}^{G^k}$ and applies $\mathrm{QFT}_G^{\otimes k}$ to it. Here G^k, H^k denote the k-fold direct product of G, H respectively. Note that since $\widehat{G^k} \cong \widehat{G}^{\otimes k}$, we have that $\mathrm{QFT}_{G^k} = \mathrm{QFT}_G^{\otimes k}$. We can express an irrep $\boldsymbol{\rho}$ of G^k as $\boldsymbol{\rho} = \otimes_{i=1}^k \rho_i$, $\rho_i \in \widehat{G}$; observe that $V_{\boldsymbol{\rho}} = \otimes_{i=1}^k V_{\rho_i}$. We adopt the convention that multiregister vectors and representations are denoted in boldface type. After applying $\mathrm{QFT}_G^{\otimes k}$, we measure the name $\boldsymbol{\rho}$ of an irrep of G^k, i. e, irreps $\rho_1, \ldots, \rho_k$ of G. After that, we measure a row index of $\boldsymbol{\rho}$ i. e., row indices of $\rho_1, \ldots, \rho_k$, and then measure the resulting reduced state in the column space of $\boldsymbol{\rho}$ using a frame $\boldsymbol{\mathcal{B}}$ of $V_{\boldsymbol{\rho}}$. The frame $\boldsymbol{\mathcal{B}}$ used depends on the observed $\boldsymbol{\rho}$ but not on the observed row indices. Notice that only the application of the frame $\boldsymbol{\mathcal{B}}$ may be an entangled measurement, the application of $\mathrm{QFT}_G^{\otimes k}$ and measurement of $\boldsymbol{\rho}$ together with a row index of $\boldsymbol{\rho}$ are single register operations.

2.3 Graph isomorphism and HSP

The usual reduction of deciding isomorphism of two n-vertex graphs to HSP in S_{2n} actually embeds the problem into a proper subgroup of S_{2n}, namely, $S_n \wr S_2$ [EHK99a]. The elements of $S_n \wr S_2$ are tuples of the form (π, σ, b) where $\pi, \sigma \in S_n$ and $b \in \mathbb{Z}_2$ with the multiplication rule $(\pi_1, \sigma_1, 0) \cdot (\pi_2, \sigma_2, b) := (\pi_1\pi_2, \sigma_1\sigma_2, b)$ and $(\pi_1, \sigma_1, 1) \cdot (\pi_2, \sigma_2, b) := (\pi_1\sigma_2, \sigma_1\pi_2, 1 \oplus b)$. The embedding of $S_n \wr S_2$ in S_{2n} treats $\{1, \ldots, 2n\}$ as a union of $\{1, \ldots, n\} \cup \{n+1, \ldots, 2n\}$ with π, σ permuting the first and second sets respectively when $b = 0$, and π permuting the first set onto the second and and σ permuting the second set onto the first when $b = 1$. There is an element of the form $(\pi, \pi^{-1}, 1)$, called an *involutive swap*, in the hidden subgroup iff the two graphs are isomorphic.

Additionally, if the two graphs are rigid, i. e., have no non-trivial automorphisms, then the hidden subgroup is trivial if they are non-isomorphic, and is generated by $(\pi, \pi^{-1}, 1)$ if they are isomorphic where π is the unique isomorphism from the first graph onto the second. This element $(\pi, \pi^{-1}, 1)$ is of order two, and is a conjugate in $S_n \wr S_2$ of $h := (e, e, 1)$ where $e \in S_n$ is the identity permutation. Viewed as an element of S_{2n}, $h = (1, n+1)(2, n+2)\cdots(n, 2n)$. The set of conjugates of h in $S_n \wr S_2$ is the set of all involutive swaps $(\pi, \pi^{-1}, 1)$, $\pi \in S_n$, and corresponds exactly to all the isomorphisms possible between the two graphs. Also $S_n \wr S_2$ is the smallest group containing all involutive swaps as a single conjugacy class. This algebraic property makes $S_n \wr S_2$ ideal for the study of isomorphism of rigid graphs as a hidden subgroup problem. Note that graph automorphism, i. e., deciding if a given graph has a non-trivial automorphism, is Turing equivalent classically to isomorphism of rigid graphs [KST93].

In this paper, we consider the following problem: Given that the hidden subgroup in $S_n \wr S_2$ is either generated by an involutive swap or is trivial, decide which case is true. Graph automorphism as well as rigid-graph isomorphism reduces to this problem. We show that any efficient algorithm using coset states that solves this problem needs to make measurements entangled across $\Omega(n \log n)$ states (Corollary 6). Note that any lower bound for this problem for a coset state based algorithm holds true even when the involutive swaps are considered as elements of S_{2n} rather than $S_n \wr S_2$. This is because of the following general *transfer lemma*.

Lemma 1 (Transfer lemma)**.** *Let G be a finite group and suppose that either $G \leq \tilde{G}$ or $G \cong \tilde{G}/N$, $N \trianglelefteq \tilde{G}$ holds. Then lower bounds for coset state based algorithms for the HSP in G transfer to the same bounds for the HSP in $\tilde{G}$ and vice versa, as long as the hidden subgroups involved are contained in G.*

Proof. Let $H \leq G$. The case $G \leq \tilde{G}$ follows from the observation that $\mathbb{C}[\tilde{G}] = \bigoplus_{\tilde{g} \in \tilde{G}/G} L_{\tilde{g}} \cdot \mathbb{C}[G]$, where $\tilde{G}/G$ denotes a system of left coset representatives of G in $\tilde{G}$ and $L_{\tilde{g}}$ stands for left multiplication

by $\tilde{g}$. Then, $\sigma_H^{\tilde{G}} = \bigoplus_{\tilde{g}\in\tilde{G}/G} L_{\tilde{g}} \cdot \sigma_H^G \cdot L_{\tilde{g}}^\dagger$, and so any coset state based algorithm without loss of generality performs the same operations on each block of the orthogonal direct sum. The case $G \cong \tilde{G}/N$ follows from the observation that $\mathbb{C}[G]$ is isometric to the subspace of $\mathbb{C}[\tilde{G}]$ spanned by coset states of N namely states of the form $|\tilde{g}N\rangle$, $\tilde{g} \in \tilde{G}$. There is a subgroup $\tilde{H} \leq \tilde{G}$, $N \trianglelefteq \tilde{H}$ such that $\tilde{H}/N \cong H$. Hence, $\sigma_H^G \cong \sigma_{\tilde{H}}^{\tilde{G}}$. Thus, any coset state based algorithm without loss of generality performs the same operations on σ_H^G and $\sigma_{\tilde{H}}^{\tilde{G}}$. □

Lemma 1 says that algorithms working on coset states can get no advantage by changing the ambient group while keeping the hidden subgroup the same. However, there are efficient algorithms for HSP instances that work by changing the hidden subgroup, for example the self-reducibility techniques introduced in [FIM$^+$03] that reduce the HSP for a subgroup H to the HSP for a larger subgroup HN, where $N \trianglelefteq G$, see also [ISS07]. An example for this is discussed later in this paper in a remark following Corollary 13. Lemma 1 does not apply to these settings.

Childs and Wocjan [CW07] showed an $\Omega(n)$ lower bound for the total number of hidden shift states required to solve graph isomorphism, and also proved that a single hidden shift state contains exponentially little information about the isomorphism. However, their results do not rule out an algorithm that makes joint measurements on, say, two states at a time and uses a total of $O(n)$ hidden shift states. Since the hidden shift state corresponding to the shift (π, π^{-1}), where $\pi \in S_{n/2}$ is exactly the coset state for the hidden subgroup generated by the involutive swap $(\pi, \pi^{-1}, 1)$ in $S_{n/2} \wr S_2$, Lemma 1 and Corollary 6 of our paper show that any efficient algorithm using hidden shift states to solve the graph isomorphism problem needs to make measurements entangled across $\Omega(n \log n)$ states, generalizing their results.

3 The main theorem

Let G be a group and $h \in G$ be an involution, that is, $H := \{1, h\}$ is an order two subgroup of G. We let $H^g := gHg^{-1}$ denote the conjugate of H by $g \in G$. Let k be a positive integer. Fix a POVM $\mathcal{M}$ on $\mathbb{C}[G]^{\otimes k} \cong \mathbb{C}[G^k]$. Let $\mathcal{M}_{H^g}$, $\mathcal{M}_{\{1\}}$ denote the classical probability distributions obtained by measuring the states $\sigma_{H^g}^{\otimes k}$, $\sigma_{\{1\}}^{\otimes k}$ respectively according to $\mathcal{M}$. We will show that the average total variation distance between $\mathcal{M}_{H^g}$ and $\mathcal{M}_{\{1\}}$ over conjugates H^g, $g \in G$ is at most 2^k times a quantity that depends purely on the pair (G, H). In the next section, we will show that this quantity is exponentially small for many pairs (G, H) of interest, including when $G = S_n \wr S_2$ and H is generated by an involutive swap, i. e., the case relevant to isomorphism of rigid graphs.

Theorem 1 (Main theorem)**.** *Let G be a finite group and $H := \{1, h\}$ be an order two subgroup of G. Let $k \geq 1$ be an integer. Fix a POVM $\mathcal{M}$ on $\mathbb{C}[G]^{\otimes k}$ and let $\mathcal{M}_{H^g}$, $\mathcal{M}_{\{1\}}$ denote the classical probability distributions obtained by measuring the states $\sigma_{H^g}^{\otimes k}$, $\sigma_{\{1\}}^{\otimes k}$ respectively according to $\mathcal{M}$. For $0 < \varepsilon < 1$, define the set*

$$\mathcal{S}_\varepsilon := \left\{ \tau \in \widehat{G} : \frac{|\chi_\tau(h)|}{d_\tau} \geq \varepsilon \right\}.$$

Define

$$\delta_1 := \left(\varepsilon + \frac{1}{|G|} \cdot \left(\sum_{\tau\in\mathcal{S}_\varepsilon} d_\tau |\chi_\tau(h)| \right) \cdot \left(\sum_{\nu\in\widehat{G}} d_\nu \right) \right)^{1/2}, \; \delta_2 := \left(\varepsilon + \left(\sum_{\tau\in\mathcal{S}_\varepsilon} d_\tau^2 \right) \cdot \left(\frac{|\widehat{G}|}{|G|} \right)^{1/2} \right)^{1/2}.$$

Then

$$\mathrm{E}_g \left[\left\| \mathcal{M}_{H^g} - \mathcal{M}_{\{1\}} \right\|_1 \right] \leq 2^k \delta_1 \leq 2^k \delta_2,$$

where the expectation is taken over the uniform distribution on $g \in G$.

Remarks:
1. As made clearer in Corollary 3 below, if k is chosen small enough so that $2^k \delta_1$ is exponentially small, we get a lower bound of k on the jointness of any measurement on coset states solving the HSP in G.
2. The parameter ε above can be thought of as a *smoothness* parameter. An irrep $\rho \in \widehat{G}$ is *smooth* if $|\chi_\rho / d_\rho| < \varepsilon$. Note that the trivial irrep of G is non-smooth for every choice of ε. It is not too difficult to show that smooth irreps contribute at most $2^k \varepsilon$ to the expected total variation distance. The majority of the work required in proving Theorem 1 goes into bounding the contribution of the non-smooth irreps of G, namely, the irreps in the set $\mathcal{S}_\varepsilon$. In order to apply Theorem 1 to a concrete example and prove a small upper bound on the expected total variation distance, one has to make a clever choice of ε. For some group-subgroup pairs (G, H) like $G = S_n \wr S_2$ and H is generated by an involutive swap, it is possible to make a clever choice of ε so that 'most' irreps of G become smooth and δ_1 becomes exponentially small. In this case, we get a lower bound of $\Omega(\log |G|)$ on the jointness of any efficient measurement on coset states solving the HSP in G. But for some other groups like abelian G this is not possible. If G is abelian, for any choice of ε, all irreps of G are non-smooth and δ_1 is a constant. Indeed, single register measurements suffice to solve the HSP in G.
3. The expectation over all conjugates of H cannot be avoided if one wishes to prove a small upper bound on the total variation distance. This is because for any fixed conjugate subgroup H^g,

$$\left\| \sigma_{H^g}^{\otimes k} - \sigma_{\{1\}}^{\otimes k} \right\|_{\mathrm{tr}} \geq \left\| \sigma_{H^g} - \sigma_{\{1\}} \right\|_{\mathrm{tr}} \geq 1,$$

and so there exists a POVM $\mathcal{M}$ with $\left\| \mathcal{M}_{H^g} - \mathcal{M}_{\{1\}} \right\|_1 \geq 1$. The second inequality follows from the fact that coset states for any two different hidden subgroups have trace distance at least 1 [RRS09]. The point of Theorem 1 is that if we fix any POVM $\mathcal{M}$ *a priori*, it will give small total variation distance between an average σ_{H^g} and $\sigma_{\{1\}}$.

We now prove a corollary to Theorem 1.

Corollary 2. *Assume the conditions of Theorem 1. Let $\mathcal{A}$ be a quantum operation on $\mathbb{C}[G]^{\otimes k}$, and let $\mathcal{A}[\sigma_{H^g}^{\otimes k}]$, $\mathcal{A}[\sigma_{\{1\}}^{\otimes k}]$ denote the resulting quantum states obtained by applying $\mathcal{A}$ on $\sigma_{H^g}^{\otimes k}$, $\sigma_{\{1\}}^{\otimes k}$ respectively. Suppose the output of $\mathcal{A}$ consists of at most l qubits and an arbitrary number of classical bits. Then*

$$\mathrm{E}_g \left[\left\| \mathcal{A}[\sigma_{H^g}^{\otimes k}] - \mathcal{A}[\sigma_{\{1\}}^{\otimes k}] \right\|_{\mathrm{tr}} \right] \leq 2^{C+k+(l/2)} \, \delta_1 \leq 2^{C+k+(l/2)} \, \delta_2,$$

where the expectation is taken over the uniform distribution on $g \in G$, and C is a universal constant, $C > 0$.

Proof. Let σ_1, σ_2 be two quantum states in the same Hilbert space and $p_1, p_2 \geq 0$. Suppose $p_1 \geq p_2$. Then,

$$\|p_1 \sigma_1 - p_2 \sigma_2\|_{\mathrm{tr}} \leq \|p_1(\sigma_1 - \sigma_2)\|_{\mathrm{tr}} + \|(p_1 - p_2)\sigma_2\|_{\mathrm{tr}} = p_1 \|\sigma_1 - \sigma_2\|_{\mathrm{tr}} + |p_1 - p_2|.$$

Let $\mathcal{M}$ be a POVM. Now,

$$\|p_1 \mathcal{M}[\sigma_1] - p_2 \mathcal{M}[\sigma_2]\|_1 = \|p_1(\mathcal{M}[\sigma_1] - \mathcal{M}[\sigma_2]) + (p_1 - p_2)\mathcal{M}[\sigma_2]\|_1 \geq \frac{p_1}{2} \|\mathcal{M}[\sigma_1] - \mathcal{M}[\sigma_2]\|_1 .$$

The inequality above follows by considering those outcomes of $\mathcal{M}$ that have at least as much probability for σ_1 as for σ_2, and the fact that $(p_1 - p_2)\mathcal{M}[\sigma_2]$ is a vector with non-negative entries. Also,

$\|p_1\mathcal{M}[\sigma_1] - p_2\mathcal{M}[\sigma_2]\|_1 \geq |p_1 - p_2|$. Now suppose $\|\mathcal{M}[\sigma_1] - \mathcal{M}[\sigma_2]\|_1 \geq c\,2^{-l/2}\,\|\sigma_1 - \sigma_2\|_{\mathrm{tr}}$, where $1 > c > 0$. Then,

$$\begin{aligned}
\|p_1\mathcal{M}[\sigma_1] - p_2\mathcal{M}[\sigma_2]\|_1 &\geq \frac{|p_1 - p_2|}{2} + \frac{p_1}{4}\,\|\mathcal{M}[\sigma_1] - \mathcal{M}[\sigma_2]\|_1 \\
&\geq \frac{c\,2^{-l/2}|p_1 - p_2|}{4} + \frac{c\,p_1 2^{-l/2}}{4}\,\|\sigma_1 - \sigma_2\|_{\mathrm{tr}} \\
&\geq \frac{c\,2^{-l/2}\,\|p_1\sigma_1 - p_2\sigma_2\|_{\mathrm{tr}}}{4}.
\end{aligned}$$

Let b range over the possible values of the classical output of $\mathcal{A}$. Fix some b. Let $\rho_{H^g;b}$ denote the state of the quantum output of $\mathcal{A}$ conditioned on the classical output being b, when the input state is $\sigma_{H^g}^{\otimes k}$; $\rho_{\{1\};b}$ is defined similarly. Consider the ensemble $\mathcal{E}_b := \{\rho_{H^g;b}\}_{g\in G} \cup \{\rho_{\{1\};b}\}$. Applying Fact 1 to ensemble $\mathcal{E}_b$, we see that there exists a POVM $\mathcal{M}_b$ independent of g such that

$$\left\|\mathcal{M}_b[\rho_{H^g;b}] - \mathcal{M}_b[\rho_{\{1\};b}]\right\|_1 \geq c\,2^{-l/2}\left\|\rho_{H^g;b} - \rho_{\{1\};b}\right\|_{\mathrm{tr}}$$

for all $g \in G$, where c is a universal constant, $0 < c < 1$.

Let $\mathcal{A}[\sigma_{H^g}^{\otimes k}](b)$, $\mathcal{A}[\sigma_{\{1\}}^{\otimes k}](b)$ denote the probabilities of b in the states $\mathcal{A}[\sigma_{H^g}^{\otimes k}]$, $\mathcal{A}[\sigma_{\{1\}}^{\otimes k}]$ respectively. Define a POVM $\mathcal{A}'$ that first applies $\mathcal{A}$ and then measures its quantum output with $\mathcal{M}_b$ conditioned on its classical output being b. Then,

$$\begin{aligned}
\left\|\mathcal{A}[\sigma_{H^g}^{\otimes k}] - \mathcal{A}[\sigma_{\{1\}}^{\otimes k}]\right\|_{\mathrm{tr}} &= \left\|\sum_b \mathcal{A}[\sigma_{H^g}^{\otimes k}](b)\cdot|b\rangle\langle b| \otimes \rho_{H^g;b} - \sum_b \mathcal{A}[\sigma_{\{1\}}^{\otimes k}](b)\cdot|b\rangle\langle b| \otimes \rho_{\{1\};b}\right\|_{\mathrm{tr}} \\
&= \sum_b \left\|\mathcal{A}[\sigma_{H^g}^{\otimes k}](b)\cdot\rho_{H^g;b} - \mathcal{A}[\sigma_{\{1\}}^{\otimes k}](b)\cdot\rho_{\{1\};b}\right\|_{\mathrm{tr}} \\
&\leq c^{-1}2^{l/2}\sum_b \left\|\mathcal{A}[\sigma_{H^g}^{\otimes k}](b)\cdot\mathcal{M}_b[\rho_{H^g;b}] - \mathcal{A}[\sigma_{\{1\}}^{\otimes k}](b)\cdot\mathcal{M}_b[\rho_{\{1\};b}]\right\|_1 \\
&= c^{-1}2^{l/2}\left\|\mathcal{A}'[\sigma_{H^g}^{\otimes k}] - \mathcal{A}'[\sigma_{\{1\}}^{\otimes k}]\right\|_1.
\end{aligned}$$

Applying Theorem 1 to the POVM $\mathcal{A}'$ gives

$$\mathrm{E}_g\left[\left\|\mathcal{A}'[\sigma_{H^g}^{\otimes k}] - \mathcal{A}'[\sigma_{\{1\}}^{\otimes k}]\right\|_{\mathrm{tr}}\right] \leq 2^k\,\delta_1 \leq 2^k\,\delta_2,$$

where the expectation is taken over the uniform distribution on $g \in G$. Combining the above two inequalities and setting $C := \log c^{-1}$ completes the proof of the present corollary. □

Remark: In general, the exponential dependence on l in the statement of Corollary 2 seems necessary. For several group-subgroup pairs, for example when $G := S_n \wr S_2$ and H is generated by an involutive swap as in Section 2.3, we have that $\delta_2 = 2^{-\Theta(\log|G|)}$. However, if l is allowed to be larger than $\log|G|$, then the quantum work space can hold a coset state. Since coset states for any two different hidden subgroups have trace distance at least 1 [RRS09], taking the quantum operation from states in $\mathbb{C}[G]^{\otimes k}$ to states in $\mathbb{C}[G]$ that traces out the first $k-1$ registers as $\mathcal{A}$ in Corollary 2 will give trace distance at least 1. This is true even if $k = 1$. Thus, we seem to require an exponential dependence on l in the statement of Corollary 2.

We now define the class of POVMs for which our lower bound results apply.

Definition 1. *A POVM $\mathcal{F}$ on t coset states with* jointness k *and l* qubits of work space *consists of a sequence of quantum operations, that is, completely positive trace preserving superoperators $(\mathcal{M}_i)_{i\in[t']}$, $t' \leq t$, where each $\mathcal{M}_i$ operates on at most l qubits and on a fresh set of at most k coset states. The final operation $\mathcal{M}_{t'}$ does not produce any quantum output, that is, it is a POVM by itself. The total number of coset states operated upon by $\mathcal{F}$ is at most t. The work space qubits are in addition to the qubits required to store the t coset states and are initialized to $|0\rangle$; also, they are the only additional qubits in the circuit for $\mathcal{F}$, treating the operations $\mathcal{M}_i$ as black boxes. Note that the black boxes implementing $\mathcal{M}_i$ are free to use arbitrary amount of classical and quantum computational resources internally, and the circuit for $\mathcal{F}$ can have arbitrary amount of classical work space outside the black boxes for $\mathcal{M}_i$. The outcome of POVM $\mathcal{F}$ is a sequence of length t' corresponding to the classical outputs of $\mathcal{M}_i$. The choice of $\mathcal{M}_i$ may depend on the observed outcomes of $\mathcal{M}_1, \ldots, \mathcal{M}_{i-1}$. If required, further classical postprocessing may be done on the outcome of $\mathcal{F}$.*

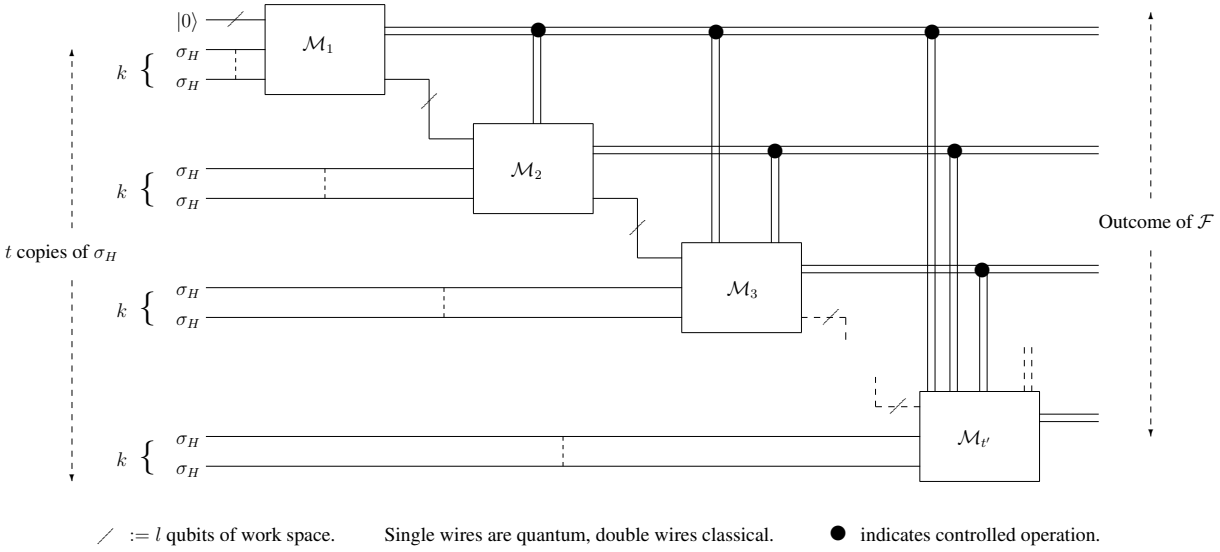

Figure 1: Schematic diagram of a POVM $\mathcal{F}$ on t copies of coset state σ_H with jointness k and l qubits of work space.

After defining POVMs with limited jointness and quantum work space, we can now prove the following corollary of Theorem 1.

Corollary 3. *Suppose $\mathcal{F}$ is a POVM on t coset states with jointness k and l qubits of work space. Let $\mathcal{F}_{H^g}$, $\mathcal{F}_{\{1\}}$ be the probability distributions obtained by applying $\mathcal{F}$ on $\sigma_{H^g}^{\otimes t}$, $\sigma_{\{1\}}^{\otimes t}$. Then for at least a fraction of $1 - t\,2^{C+k+(l/2)}\delta_1^{1/2}$ conjugate subgroups H^g, $g \in G$, $\left\|\mathcal{F}_{H^g} - \mathcal{F}_{\{1\}}\right\|_1 \leq \delta_1^{1/2}$.*

Proof. For $1 \leq i \leq t'$, let $\mathcal{F}_{H^g;i}$, $\mathcal{F}_{\{1\};i}$ denote the states of the classical and quantum work space in the circuit for $\mathcal{F}$ after the application of $\mathcal{M}_1, \ldots, \mathcal{M}_i$, when the input coset states are σ_{H^g}, $\sigma_{\{1\}}$ respectively. Observe that $\mathcal{F}_{H^g} = \mathcal{F}_{H^g;t'}$ and $\mathcal{F}_{\{1\}} = \mathcal{F}_{\{1\};t'}$.

Let $\mathbf{b}$ range over possible values of the classical outputs of the quantum operations $\mathcal{M}_1, \ldots, \mathcal{M}_i$. Suppose $1 \leq i < t'$. Let $\mathcal{M}_{i+1;\mathbf{b}}$ be the $(i+1)$th quantum operation that is applied on the quantum work space and k fresh coset states, conditioned on the classical output of $\mathcal{M}_1, \ldots, \mathcal{M}_i$ being $\mathbf{b}$. Let $\mathcal{M}'_{i+1}$ denote

the overall controlled quantum operation corresponding to applying $\mathcal{M}_{i+1;\mathbf{b}}$ when the classical output of $\mathcal{M}_1, \ldots, \mathcal{M}_i$ is $\mathbf{b}$. Define a quantum operation $\mathcal{M}''_{i+1}$ on $\mathbb{C}[G]^{\otimes k}$ as

$$\mathcal{M}''_{i+1}[X] := \mathcal{M}'_{i+1}[\mathcal{F}_{\{1\};i} \otimes X]$$

where X is an operator on $\mathbb{C}[G]^{\otimes k}$. Define another quantum operation $\mathcal{M}'''_{H^g;i+1}$ on the classical outputs of $\mathcal{M}_1, \ldots, \mathcal{M}_i$ together with the quantum output of $\mathcal{M}_i$ as

$$\mathcal{M}'''_{H^g;i+1}[Y] := \mathcal{M}'_{i+1}[Y \otimes \sigma_{H^g}]$$

where Y is an operator on the Hilbert space corresponding to the classical outputs of $\mathcal{M}_1, \ldots, \mathcal{M}_i$ and the quantum output of $\mathcal{M}_i$.

Taking an expectation over the uniform distribution on $g \in G$, we get

$$\begin{aligned}
\mathrm{E}_g[\left\|\mathcal{F}_{H^g;i+1} - \mathcal{F}_{\{1\};i+1}\right\|_{\mathrm{tr}}] &= \mathrm{E}_g\left[\left\|\mathcal{M}'_{i+1}[\mathcal{F}_{H^g;i} \otimes \sigma_{H^g}^{\otimes k}] - \mathcal{M}'_{i+1}[\mathcal{F}_{\{1\};i} \otimes \sigma_{\{1\}}^{\otimes k}]\right\|_{\mathrm{tr}}\right] \\
&\leq \mathrm{E}_g\left[\left\|\mathcal{M}'_{i+1}[\mathcal{F}_{H^g;i} \otimes \sigma_{H^g}^{\otimes k}] - \mathcal{M}'_{i+1}[\mathcal{F}_{\{1\};i} \otimes \sigma_{H^g}^{\otimes k}]\right\|_{\mathrm{tr}}\right] \\
&\quad + \mathrm{E}_g\left[\left\|\mathcal{M}'_{i+1}[\mathcal{F}_{\{1\};i} \otimes \sigma_{H^g}^{\otimes k}] - \mathcal{M}'_{i+1}[\mathcal{F}_{\{1\};i} \otimes \sigma_{\{1\}}^{\otimes k}]\right\|_{\mathrm{tr}}\right] \\
&= \mathrm{E}_g\left[\left\|\mathcal{M}'''_{H^g;i+1}[\mathcal{F}_{H^g;i}] - \mathcal{M}'''_{H^g;i+1}[\mathcal{F}_{\{1\};i}]\right\|_{\mathrm{tr}}\right] \\
&\quad + \mathrm{E}_g\left[\left\|\mathcal{M}''_{i+1}[\sigma_{H^g}^{\otimes k}] - \mathcal{M}''_{i+1}[\sigma_{\{1\}}^{\otimes k}]\right\|_{\mathrm{tr}}\right] \\
&\leq \mathrm{E}_g\left[\left\|\mathcal{F}_{H^g;i} - \mathcal{F}_{\{1\};i}\right\|_{\mathrm{tr}}\right] + 2^{C+k+(l/2)}\,\delta_1,
\end{aligned}$$

where the second inequality follows from the fact that a quantum operation cannot increase the trace distance between density matrices and by Corollary 2.

Thus, by induction on i we see that

$$\mathrm{E}_g[\left\|\mathcal{F}_{H^g} - \mathcal{F}_{\{1\}}\right\|_1] = \mathrm{E}_g[\left\|\mathcal{F}_{H^g;t'} - \mathcal{F}_{\{1\};t'}\right\|_{\mathrm{tr}}] \leq t'\, 2^{C+k+(l/2)}\delta_1 \leq t\, 2^{C+k+(l/2)}\delta_1.$$

Applying Markov's inequality to the above expectation finishes the proof of the present corollary. □

The remainder of the section is devoted to proving Theorem 1. We first give some notation that will be useful for the proofs of various lemmas. Our notation and setup is inspired to a large extent by the notation in [MR05b].

As argued in the previous section, we can assume without loss of generality that $\mathcal{M}$ first applies $\mathrm{QFT}_G^{\otimes k}$ to $\sigma_{H^g}^{\otimes k}$, measures the name of an irrep of G^k, $\boldsymbol{\rho}^*$ together with a row index of $\boldsymbol{\rho}^*$, and then measures the resulting reduced state in the column space of $\boldsymbol{\rho}^*$ using a frame $\boldsymbol{\mathcal{B}}$ of $V_{\boldsymbol{\rho}^*} = V_{\boldsymbol{\rho}}$. If $\boldsymbol{\rho} = \otimes_{i=1}^k \rho_i$, $\rho_i \in \widehat{G}$, then $\boldsymbol{\rho}^* = \otimes_{i=1}^k \rho_i^*$. The frame $\boldsymbol{\mathcal{B}}$ used depends on the observed $\boldsymbol{\rho}^*$ but not on the observed row indices.

Suppose the hidden subgroup is H^g for some $g \in G$. It is easy to see that the probability that $\mathcal{M}$ measures $\boldsymbol{\rho}^*$ is given by

$$\mathcal{M}_{H^g}(\boldsymbol{\rho}^*) = \frac{d_{\boldsymbol{\rho}^*}|H^g|^k \cdot r_{\boldsymbol{\rho}^*}((H^g)^k)}{|G|^k} = \frac{2^k d_{\boldsymbol{\rho}} r_{\boldsymbol{\rho}}(H^k)}{|G|^k}.$$

Notice that $\mathcal{M}_{H^g}(\boldsymbol{\rho}^*) = \mathcal{M}_H(\boldsymbol{\rho})$. Let $\boldsymbol{\mathcal{B}} = \{a_{\mathbf{b}}, \mathbf{b}\}$, where $0 \leq a_{\mathbf{b}} \leq 1$ and $\sum_{\mathbf{b}} a_{\mathbf{b}}|\mathbf{b}\rangle\langle\mathbf{b}| = \mathbb{1}_{V_{\boldsymbol{\rho}}}$. Then the reduced state in the column space of $\boldsymbol{\rho}^*$ is $\frac{\boldsymbol{\rho}((H^g)^k)}{r_{\boldsymbol{\rho}}(H^k)}$, if $r_{\boldsymbol{\rho}} \neq 0$. Hence, the probability of observing a

particular $\mathbf{b}$ conditioned on having observed $\boldsymbol{\rho}^*$ is

$$\mathcal{M}_{H^g}(\mathbf{b} \mid \boldsymbol{\rho}^*) = \frac{a_{\mathbf{b}} \left\langle \mathbf{b} \left| \boldsymbol{\rho}((H^g)^k) \right| \mathbf{b} \right\rangle}{r_{\boldsymbol{\rho}}(H^k)},$$

if $r_{\boldsymbol{\rho}}(H^k) \neq 0$, and 0 otherwise. Similarly, if the hidden subgroup is the identity subgroup then

$$\mathcal{M}_{\{1\}}(\boldsymbol{\rho}^*) = \frac{d_{\boldsymbol{\rho}}^2}{|G|^k} =: \mathcal{P}(\boldsymbol{\rho}),$$

where $\mathcal{P}(\cdot)$ is the *Plancherel distribution* on irreps of G^k. Also

$$\mathcal{M}_{\{1\}}(\mathbf{b} \mid \boldsymbol{\rho}^*) = \frac{a_{\mathbf{b}}}{d_{\boldsymbol{\rho}}} =: \mathcal{N}(\mathbf{b} \mid \boldsymbol{\rho}^*),$$

where $\mathcal{N}(\cdot \mid \boldsymbol{\rho}^*)$ is the *natural distribution* corresponding to the frame $\boldsymbol{\mathcal{B}}$. Henceforth in the paper, we will use the following shorthand for expectations: $\mathrm{E}_{\boldsymbol{\rho}}[\cdot]$, $\mathrm{E}_{\mathbf{b}}[\cdot]$ and $\mathrm{E}_g[\cdot]$ denote expectations over the Plancherel distribution on irreps, natural distribution on frame vectors and uniform distribution on elements of G respectively.

We define a function $X : \widehat{G}^{\otimes k} \times \boldsymbol{\mathcal{B}} \times G \to [-1, 1]$ as

$$X(\boldsymbol{\rho}, \mathbf{b}, g) := \left\langle \mathbf{b} \left| \boldsymbol{\rho}((H^g)^k) \right| \mathbf{b} \right\rangle - \frac{1}{2^k},$$

where $\boldsymbol{\mathcal{B}}$ is a frame for $V_{\boldsymbol{\rho}}$. The significance of the function X is made clear in Lemma 2 below.

Lemma 2. $\left\| \mathcal{M}_{H^g} - \mathcal{M}_{\{1\}} \right\|_1 < 2^k \cdot \mathrm{E}_{\boldsymbol{\rho},\mathbf{b}}[|X(\boldsymbol{\rho}, \mathbf{b}, g)|]$.

Proof. If the hidden subgroup is H^g for some $g \in G$, the probability of observing an irrep $\boldsymbol{\rho}^* \in \widehat{G}^{\otimes k}$, row index $\mathbf{i} \in [d_{\boldsymbol{\rho}}]$ and frame vector $\mathbf{b} \in \boldsymbol{\mathcal{B}}$ is given by

$$\mathcal{M}_{H^g}(\boldsymbol{\rho}, \mathbf{i}, \mathbf{b}) = \mathcal{M}_H(\boldsymbol{\rho}) \cdot \frac{1}{d_{\boldsymbol{\rho}}} \cdot \mathcal{M}_{H^g}(\mathbf{b} \mid \boldsymbol{\rho}^*) = \frac{2^k a_{\mathbf{b}} \left\langle \mathbf{b} \left| \boldsymbol{\rho}((H^g)^k) \right| \mathbf{b} \right\rangle}{|G|^k}.$$

The above equality holds even when $r_{\boldsymbol{\rho}}(H^k) = 0$.

If the hidden subgroup is $\{1\}$, the probability of observing an irrep $\boldsymbol{\rho}^* \in \widehat{G}^{\otimes k}$, row index $\mathbf{i} \in [d_{\boldsymbol{\rho}}]$ and frame vector $\mathbf{b} \in \boldsymbol{\mathcal{B}}$ is given by

$$\mathcal{M}_{\{1\}}(\boldsymbol{\rho}, \mathbf{i}, \mathbf{b}) = \mathcal{P}(\boldsymbol{\rho}) \cdot \frac{1}{d_{\boldsymbol{\rho}}} \cdot \mathcal{N}(b \mid \boldsymbol{\rho}) = \frac{a_{\mathbf{b}}}{|G|^k}.$$

Thus,

$$\begin{aligned}
\left\| \mathcal{M}_{H^g} - \mathcal{M}_{\{1\}} \right\|_1 &= \sum_{\boldsymbol{\rho} \in \widehat{G}^{\otimes k}} \sum_{\mathbf{i}=1}^{d_{\boldsymbol{\rho}}} \sum_{\mathbf{b} \in \boldsymbol{\mathcal{B}}} \left| \mathcal{M}_{H^g}(\boldsymbol{\rho}, \mathbf{i}, \mathbf{b}) - \mathcal{M}_{\{1\}}(\boldsymbol{\rho}, \mathbf{i}, \mathbf{b}) \right| \\
&= \sum_{\boldsymbol{\rho} \in \widehat{G}^{\otimes k}} \sum_{\mathbf{i}=1}^{d_{\boldsymbol{\rho}}} \sum_{\mathbf{b} \in \boldsymbol{\mathcal{B}}} \left| \frac{2^k a_{\mathbf{b}} \left\langle \mathbf{b} \left| \boldsymbol{\rho}((H^g)^k) \right| \mathbf{b} \right\rangle}{|G|^k} - \frac{a_{\mathbf{b}}}{|G|^k} \right| \\
&= 2^k \sum_{\boldsymbol{\rho} \in \widehat{G}^{\otimes k}} \frac{d_{\boldsymbol{\rho}}^2}{|G|^k} \sum_{\mathbf{b} \in \boldsymbol{\mathcal{B}}} \frac{a_{\mathbf{b}}}{d_{\boldsymbol{\rho}}} \left| \left\langle \mathbf{b} \left| \boldsymbol{\rho}((H^g)^k) \right| \mathbf{b} \right\rangle - \frac{1}{2^k} \right| \\
&= 2^k \mathrm{E}_{\boldsymbol{\rho},\mathbf{b}}[|X(\boldsymbol{\rho}, \mathbf{b}, g)|].
\end{aligned}$$

□

By Lemma 2, it suffices to upper bound $\mathrm{E}_{\boldsymbol{\rho},\mathbf{b},g}[|X(\boldsymbol{\rho},\mathbf{b},g)|]$ in order to bound $\mathrm{E}_g[\left\|\mathcal{M}_{H^g} - \mathcal{M}_{\{1\}}\right\|_1]$. By concavity of square-root, $\mathrm{E}_{\boldsymbol{\rho},\mathbf{b},g}[|X(\boldsymbol{\rho},\mathbf{b},g)|] \leq \sqrt{\mathrm{E}_{\boldsymbol{\rho},\mathbf{b},g}[X(\boldsymbol{\rho},\mathbf{b},g)^2]}$. So in the remainder of this section, we shall try to upper bound $\mathrm{E}_{\boldsymbol{\rho},\mathbf{b},g}[X(\boldsymbol{\rho},\mathbf{b},g)^2]$. For this, we require some additional notation. For a non-empty subset $I \subseteq [k]$, define $\boldsymbol{\rho}^I := (\otimes_{i\in I}\rho_i) \otimes (\otimes_{i'\in[k]\setminus I}\mathbb{1}_{d_{\rho_{i'}}})$, where $\mathbb{1}_{d_{\rho_{i'}}}$ denotes the identity representation of G of degree equal to that of $\rho_{i'}$. For non-empty subsets $I_1, I_2 \subseteq [k]$, define $\boldsymbol{\rho}^{I_1,I_2} := \boldsymbol{\rho}^{I_1} \otimes \boldsymbol{\rho}^{I_2}$. For a representation $\boldsymbol{\theta} = \otimes_{i=1}^n \theta_i$ of G^n, θ_i representation of G, we use $\boldsymbol{\theta}(g)$ as a shorthand for $\otimes_{i=1}^n \theta_i(g)$. For an irrep $\tau \in \widehat{G}$, we use $a_\tau^{\boldsymbol{\theta}}$ to denote the multiplicity of τ in the Clebsch-Gordan decomposition of $\boldsymbol{\theta}$, i. e. the number of times τ occurs in $\boldsymbol{\theta}$ when $\boldsymbol{\theta}$ is viewed as a representation of G embedded as the diagonal subgroup of G^n. We let $\Pi_\tau^{\boldsymbol{\theta}}$ denote the orthogonal projection from $V_{\boldsymbol{\theta}}$ onto the homogeneous component of τ in the above decomposition.

We start the process of upper bounding the second moment $\mathrm{E}_{\boldsymbol{\rho},\mathbf{b},g}[X(\boldsymbol{\rho},\mathbf{b},g)^2]$ by proving Lemma 3, which originally appeared in an equivalent form as [MR05b, Lemma 11]. The lemma gives us a way to express the second moment of X in terms of projections of 'coupled' frame vectors $\mathbf{b}\otimes\mathbf{b}$ onto homogeneous components corresponding to irreps $\tau \in \widehat{G}$. The advantage of doing this is that we can now distinguish between the roles of the smooth and non-smooth irreps, since the right hand side of the equality of Lemma 3 depends upon the ratio $\frac{\chi_\tau(h)}{d_\tau}$. It is easy to see that the total contribution of all the smooth irreps to the second moment of X will be upper bounded by ε. Thus, the main task that remains is to upper bound the contribution of the non-smooth irreps to the second moment of X. This idea of distinguishing between smooth and non-smooth irreps goes back to [MRS05].

Lemma 3.

$$\mathrm{E}_g[X(\boldsymbol{\rho},\mathbf{b},g)^2] = \frac{1}{4^k} \sum_{I_1,I_2\neq\{\}} \sum_{\tau\in\widehat{G}} \frac{\chi_\tau(h)}{d_\tau} \left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2.$$

Proof. Since

$$X(\boldsymbol{\rho},\mathbf{b},g) = \frac{1}{2^k}\left(\langle \mathbf{b}\,|\mathbb{1}_{d_\rho}|\,\mathbf{b}\rangle + \sum_{I\neq\{\}} \langle \mathbf{b}\,|\boldsymbol{\rho}^I(ghg^{-1})|\,\mathbf{b}\rangle\right) - \frac{1}{2^k} = \frac{1}{2^k}\sum_{I\neq\{\}} \langle \mathbf{b}\,|\boldsymbol{\rho}^I(ghg^{-1})|\,\mathbf{b}\rangle,$$

we get

$$\begin{aligned}
\mathrm{E}_g[X(\boldsymbol{\rho},\mathbf{b},g)^2] &= \mathrm{E}_g\left[\frac{1}{4^k}\sum_{I_1,I_2\neq\{\}}\left\langle \mathbf{b}\left|\boldsymbol{\rho}^{I_1}(ghg^{-1})\right|\mathbf{b}\right\rangle\cdot\left\langle \mathbf{b}\left|\boldsymbol{\rho}^{I_2}(ghg^{-1})\right|\mathbf{b}\right\rangle\right] \\
&= \frac{1}{4^k}\sum_{I_1,I_2\neq\{\}}\mathrm{E}_g\left[\left\langle \mathbf{b}\otimes\mathbf{b}\left|\boldsymbol{\rho}^{I_1,I_2}(ghg^{-1})\right|\mathbf{b}\otimes\mathbf{b}\right\rangle\right] \\
&= \frac{1}{4^k}\sum_{I_1,I_2\neq\{\}}\mathrm{E}_g\left[\left\langle \mathbf{b}\otimes\mathbf{b}\left|\bigoplus_{\tau\in\widehat{G}} a_\tau^{\boldsymbol{\rho}^{I_1,I_2}}\tau(ghg^{-1})\right|\mathbf{b}\otimes\mathbf{b}\right\rangle\right] \\
&= \frac{1}{4^k}\sum_{I_1,I_2\neq\{\}}\left\langle \mathbf{b}\otimes\mathbf{b}\left|\bigoplus_{\tau\in\widehat{G}} a_\tau^{\boldsymbol{\rho}^{I_1,I_2}}\mathrm{E}_g[\tau(ghg^{-1})]\right|\mathbf{b}\otimes\mathbf{b}\right\rangle \\
&= \frac{1}{4^k}\sum_{I_1,I_2\neq\{\}}\left\langle \mathbf{b}\otimes\mathbf{b}\left|\bigoplus_{\tau\in\widehat{G}} a_\tau^{\boldsymbol{\rho}^{I_1,I_2}}\frac{\chi_\tau(h)}{d_\tau}\mathbb{1}_{V_\tau}\right|\mathbf{b}\otimes\mathbf{b}\right\rangle \\
&= \frac{1}{4^k}\sum_{I_1,I_2\neq\{\}}\sum_{\tau\in\widehat{G}}\frac{\chi_\tau(h)}{d_\tau}\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2.
\end{aligned}$$

The fifth equality above follows from Schur's lemma. □

Lemma 3 takes care of the smooth irreps. However, we have to do something in order to bound $\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2$ for non-smooth irreps τ. The papers [MRS05, MR05b] tried to bound it using the following simple geometric argument: If $\boldsymbol{\mathcal{B}}$ is an orthonormal basis for $V_{\boldsymbol{\rho}}$, then $\{\mathbf{b}\otimes\mathbf{b}\}_{\mathbf{b}\in\boldsymbol{\mathcal{B}}}$ is an orthonormal set in $V_{\boldsymbol{\rho}}\otimes V_{\boldsymbol{\rho}}$. Hence the expectation, over the uniform distribution on $\boldsymbol{\mathcal{B}}$, of the above quantity is upper bounded by $\frac{\mathrm{rank}(\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}})}{d_{\boldsymbol{\rho}}}$. If $\boldsymbol{\mathcal{B}}$ is a POVM rather than an orthonormal basis, a similar argument can be made. This simple method works for $k=1,2$ for the symmetric group, but fails for $k\geq 3$. This is because $\mathrm{rank}(\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}})$ becomes larger than $d_{\boldsymbol{\rho}}$. The problem with the simple method is that $\mathrm{rank}(\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}})$ can be potentially as large as $d_{\boldsymbol{\rho}}^2$. This is where we need new ideas as compared to those in [MRS05, MR05b]. We use the fact that the projection $\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}$ is not arbitrary, but is rather the projection onto the homogeneous component corresponding to an irrep of G. There is an explicit representation-theoretic formula for such a projection operator (see e.g. [Ser77, Chapter 2, Theorem 8]). Using this formula allows us to 'decouple' $\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})$ into an expression involving only $\boldsymbol{\rho}^{I_1}$ and $\mathbf{b}$, and $\boldsymbol{\rho}^{I_2}$ and $\mathbf{b}$, that is, it allows us to remove the tensor product. This 'decoupling' gets around the problem that the rank of the projector can be larger than $d_{\boldsymbol{\rho}}$ whereas the size of the basis $\boldsymbol{\mathcal{B}}$ is only $d_{\boldsymbol{\rho}}$. It allows us to apply a standard corollary of Schur's orthogonality relations and finally bound the length of the projection of $\mathbf{b}\otimes\mathbf{b}$ by a small quantity.

The reader may wonder if it is necessary to first 'couple' $\boldsymbol{\rho}^{I_1}$ and $\boldsymbol{\rho}^{I_2}$ together, as well as $\mathbf{b}$ and $\mathbf{b}$, and then 'decouple' some of the terms. The reason for this style of argument is the following. Not doing the 'coupling' process essentially amounts to setting the smoothness parameter $\varepsilon=0$ so that all irreps of G become non-smooth. This turns out to be a bad idea and the upper bounding procedure for the non-smooth irreps will give a constant upper bound even for $k=1$. Note that for abelian groups G, this is precisely what happens since all irreps of an abelian group are non-smooth. Hence, we require the 'coupling' process in order to reduce the problem of upper bounding the second moment of X to upper bounding the contribution

due to the non-smooth irreps only, via Lemma 3. With a judicious choice of the smoothness parameter ε, one can hope that very 'few' irreps are non-smooth, so their contribution cannot be too large. However, proving this turns out to be impossible via geometric arguments alone, as outlined in the previous paragraph. We have to realize that the 'coupling' process was the right thing to do for smooth irreps but a 'mistake' for non-smooth irreps. So we need a 'decoupling' process that works selectively for the non-smooth irreps only. In the rest of this section, we indicate how to achieve this goal.

We now state a few facts that will be used in our 'decoupling' arguments. The next fact is easy to show and was used in the simple geometric approach of [MRS05, MR05b] to bound $\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2$.

Fact 2. *Let W be a subspace of V. Let $\mathcal{B} := \{a_b, b\}$ be a frame for V. Let Π_W^V denote the orthogonal projection from V onto W. Then*

$$\mathrm{E}_b[\left\|\Pi_W^V(b)\right\|^2] = \frac{\dim W}{\dim V},$$

where the expectation is taken over the natural distribution on $\mathcal{B}$.

The following fact is a special case of [MR05b, Lemma 12], and can be easily proved by considering the regular representation of G^n.

Fact 3. *Let $\boldsymbol{\theta} := (\bigotimes_{i=1}^n \theta_i) \otimes \left(\bigotimes_{i'=1}^{n'} \mathbb{1}_{d_{i'}}\right)$ be a representation of $G^{n+n'}$, where $\theta_i \in \widehat{G}$ and $\mathbb{1}_{d_{i'}}$ is the identity representation of G of dimension $d_{i'}$. Suppose each θ_i is chosen independently from the Plancherel distribution on $\widehat{G}$. Fix $\tau \in \widehat{G}$. Let $a_\tau^{\boldsymbol{\theta}}$ denote the multiplicity of τ in the Clebsch-Gordan decomposition of $\boldsymbol{\theta}$ i. e. viewing $\boldsymbol{\theta}$ as a representation of G embedded as the diagonal subgroup of $G^{n+n'}$. Then*

$$\mathrm{E}_{\boldsymbol{\theta}}\left[\frac{a_\tau^{\boldsymbol{\theta}}}{d_{\boldsymbol{\theta}}}\right] = \frac{d_\tau}{|G|}.$$

The following fact is a standard result in representation theory (see e.g. [Ser77, Chapter 2, Proposition 4, Corollary 3]), and follows from Schur's orthogonality relations.

Fact 4. *Suppose $\tau \in \widehat{G}$ and $b \in V_\tau$, $\|b\| = 1$. Then, $\mathrm{E}_g[|\langle b|\tau(g)|b\rangle|^2] = \frac{1}{d_\tau}$.*

We start off the 'decoupling' process by the following lemma.

Lemma 4. *Fix $I_1, I_2 \subseteq [k]$, $I_1, I_2 \neq \{\}$, $\boldsymbol{\rho} \in \widehat{G}^{\otimes k}$, $\tau \in \widehat{G}$ and $\mathbf{b} \in V_{\boldsymbol{\rho}}$. Then,*

$$\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2 \leq \frac{d_\tau^2}{2}\left(\mathrm{E}_g[|\left\langle \mathbf{b}\left|\boldsymbol{\rho}^{I_1}(g)\right|\mathbf{b}\right\rangle|^2] + \mathrm{E}_g[|\left\langle \mathbf{b}\left|\boldsymbol{\rho}^{I_2}(g)\right|\mathbf{b}\right\rangle|^2]\right).$$

Proof.

$$\begin{aligned}
\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2 &= \left|\left\langle \mathbf{b}\otimes\mathbf{b}\left|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}\right|\mathbf{b}\otimes\mathbf{b}\right\rangle\right| = \left|\left\langle \mathbf{b}\otimes\mathbf{b}\left|d_\tau \mathrm{E}_g[\chi_\tau(g)^*\boldsymbol{\rho}^{I_1}(g)\otimes\boldsymbol{\rho}^{I_2}(g)]\right|\mathbf{b}\otimes\mathbf{b}\right\rangle\right| \\
&= d_\tau\left|\mathrm{E}_g\left[\chi_\tau(g)^*\left\langle \mathbf{b}\left|\boldsymbol{\rho}^{I_1}(g)\right|\mathbf{b}\right\rangle\cdot\left\langle \mathbf{b}\left|\boldsymbol{\rho}^{I_2}(g)\right|\mathbf{b}\right\rangle\right]\right| \\
&\leq d_\tau^2 \mathrm{E}_g\left[\left|\left\langle \mathbf{b}\left|\boldsymbol{\rho}^{I_1}(g)\right|\mathbf{b}\right\rangle\right|\cdot\left|\left\langle \mathbf{b}\left|\boldsymbol{\rho}^{I_2}(g)\right|\mathbf{b}\right\rangle\right|\right] \\
&\leq \frac{d_\tau^2}{2}\left(\mathrm{E}_g\left[\left|\left\langle \mathbf{b}\left|\boldsymbol{\rho}^{I_1}(g)\right|\mathbf{b}\right\rangle\right|^2\right] + \mathrm{E}_g\left[\left|\left\langle \mathbf{b}\left|\boldsymbol{\rho}^{I_2}(g)\right|\mathbf{b}\right\rangle\right|^2\right]\right).
\end{aligned}$$

The second equality follows from a standard result in representation theory describing the projection operator onto a homogeneous component corresponding to an irrep of G (see e.g. [Ser77, Chapter 2, Theorem 8]), the first inequality follows by bounding a character value by the dimension of the representation, and the second inequality follows from the fact that $|xy| \leq \frac{|x|^2+|y|^2}{2}$ for any pair of complex numbers x, y. □

We now prove a crucial lemma that allows us to prove good upper bounds on $\left\|\Pi_{\tau}^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2$.

Lemma 5. *Fix $I \subseteq [k]$, $I \neq \{\}$. Then,* $\mathrm{E}_{\boldsymbol{\rho},\mathbf{b},g}[|\left\langle \mathbf{b}\left|\boldsymbol{\rho}^I(g)\right|\mathbf{b}\right\rangle|^2] \leq \sum_{\tau\in\widehat{G}} \frac{d_\tau}{|G|}$.

Proof. Consider the Clebsch-Gordan decomposition of $\boldsymbol{\rho}^I$ i.e. treating $\boldsymbol{\rho}^I$ as a representation of G embedded in the diagonal of G^k. Let $a_\tau^{\boldsymbol{\rho}^I}$ denote the multiplicity of an irrep $\tau \in \widehat{G}$ in this decomposition. For an i, $1 \leq i \leq a_\tau^{\boldsymbol{\rho}^I}$ let τ_i denote the ith copy of τ in this decomposition. We let $\mathbf{b}_{\tau_i}$ denote the orthogonal projection of $\mathbf{b}$ onto this copy τ_i. If $\|\mathbf{b}_{\tau_i}\| > 0$, define $\widehat{\mathbf{b}}_{\tau_i}$ to be $\mathbf{b}_{\tau_i}$ normalized; otherwise, let $\widehat{\mathbf{b}}_{\tau_i}$ be an arbitrary unit vector in the copy τ_i. We now have

$$\begin{aligned}
\left|\left\langle \mathbf{b}\left|\boldsymbol{\rho}^I(g)\right|\mathbf{b}\right\rangle\right|^2 &= \left|\left\langle \mathbf{b}\left|\bigoplus_{\tau\in\widehat{G}}\bigoplus_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\tau_i(g)\right|\mathbf{b}\right\rangle\right|^2 = \left|\sum_{\tau\in\widehat{G}}\sum_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\left\langle \mathbf{b}_{\tau_i}\left|\tau_i(g)\right|\mathbf{b}_{\tau_i}\right\rangle\right|^2 \\
&= \left|\sum_{\tau\in\widehat{G}}\sum_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\|\mathbf{b}_{\tau_i}\|\cdot\|\mathbf{b}_{\tau_i}\|\left\langle \widehat{\mathbf{b}}_{\tau_i}\left|\tau_i(g)\right|\widehat{\mathbf{b}}_{\tau_i}\right\rangle\right|^2 \\
&\leq \left(\sum_{\tau\in\widehat{G}}\sum_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\|\mathbf{b}_{\tau_i}\|^2\right)\cdot\left(\sum_{\tau\in\widehat{G}}\sum_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\|\mathbf{b}_{\tau_i}\|^2\left|\left\langle \widehat{\mathbf{b}}_{\tau_i}\left|\tau_i(g)\right|\widehat{\mathbf{b}}_{\tau_i}\right\rangle\right|^2\right) \\
&= \sum_{\tau\in\widehat{G}}\sum_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\|\mathbf{b}_{\tau_i}\|^2\left|\left\langle \widehat{\mathbf{b}}_{\tau_i}\left|\tau_i(g)\right|\widehat{\mathbf{b}}_{\tau_i}\right\rangle\right|^2.
\end{aligned}$$

The inequality above follows from Cauchy-Schwartz, and the last equality is because $\sum_{\tau\in\widehat{G}}\sum_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\|\mathbf{b}_{\tau_i}\|^2 = \|\mathbf{b}\|^2 = 1$. Now,

$$\begin{aligned}
\mathrm{E}_{\boldsymbol{\rho},\mathbf{b},g}\left[|\left\langle \mathbf{b}\left|\boldsymbol{\rho}^I(g)\right|\mathbf{b}\right\rangle|^2\right] &\leq \mathrm{E}_{\boldsymbol{\rho},\mathbf{b},g}\left[\sum_{\tau\in\widehat{G}}\sum_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\|\mathbf{b}_{\tau_i}\|^2\left|\left\langle \widehat{\mathbf{b}}_{\tau_i}\left|\tau_i(g)\right|\widehat{\mathbf{b}}_{\tau_i}\right\rangle\right|^2\right] \\
&= \mathrm{E}_{\boldsymbol{\rho},\mathbf{b}}\left[\sum_{\tau\in\widehat{G}}\sum_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\|\mathbf{b}_{\tau_i}\|^2\,\mathrm{E}_g\left[\left|\left\langle \widehat{\mathbf{b}}_{\tau_i}\left|\tau_i(g)\right|\widehat{\mathbf{b}}_{\tau_i}\right\rangle\right|^2\right]\right] \\
&= \mathrm{E}_{\boldsymbol{\rho},\mathbf{b}}\left[\sum_{\tau\in\widehat{G}}\sum_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\frac{\|\mathbf{b}_{\tau_i}\|^2}{d_\tau}\right] = \mathrm{E}_{\boldsymbol{\rho}}\left[\sum_{\tau\in\widehat{G}}\sum_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\frac{\mathrm{E}_{\mathbf{b}}[\|\mathbf{b}_{\tau_i}\|^2]}{d_\tau}\right] \\
&= \mathrm{E}_{\boldsymbol{\rho}}\left[\sum_{\tau\in\widehat{G}}\sum_{i=1}^{a_\tau^{\boldsymbol{\rho}^I}}\frac{d_\tau}{d_\tau d_{\boldsymbol{\rho}}}\right] = \mathrm{E}_{\boldsymbol{\rho}}\left[\sum_{\tau\in\widehat{G}}\frac{a_\tau^{\boldsymbol{\rho}^I}}{d_{\boldsymbol{\rho}}}\right] = \sum_{\tau\in\widehat{G}}\mathrm{E}_{\boldsymbol{\rho}}\left[\frac{a_\tau^{\boldsymbol{\rho}^I}}{d_{\boldsymbol{\rho}}}\right] \\
&= \sum_{\tau\in\widehat{G}}\frac{d_\tau}{|G|}.
\end{aligned}$$

The second equality follows from Fact 4, the fourth equality follows from Fact 2 and the last equality follows from Fact 3. □

The next lemma ties up the above threads to prove an upper bound on the second moment of the function X independent of k.

Lemma 6.

$$\mathrm{E}_{\boldsymbol{\rho},\mathbf{b},g}[X(\boldsymbol{\rho},\mathbf{b},g)^2] < \varepsilon + \frac{1}{|G|}\cdot\left(\sum_{\nu\in\widehat{G}} d_\nu\right)\cdot\left(\sum_{\tau\in\mathcal{S}_\varepsilon} d_\tau|\chi_\tau(h)|\right).$$

Proof. First, note that

$$\begin{aligned}
\mathrm{E}_g[X(\boldsymbol{\rho},\mathbf{b},g)^2] &= \left|\frac{1}{4^k}\sum_{I_1,I_2\neq\{\}}\sum_{\tau\in\widehat{G}}\frac{\chi_\tau(h)}{d_\tau}\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2\right| \\
&\leq \frac{1}{4^k}\sum_{I_1,I_2\neq\{\}}\sum_{\tau\in\widehat{G}}\frac{|\chi_\tau(h)|}{d_\tau}\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2 \\
&< \frac{1}{4^k}\sum_{I_1,I_2\neq\{\}}\left(\varepsilon\cdot\sum_{\tau\in\widehat{G}\setminus\mathcal{S}_\varepsilon}\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2 + \sum_{\tau\in\mathcal{S}_\varepsilon}\frac{|\chi_\tau(h)|}{d_\tau}\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2\right) \\
&< \varepsilon + \frac{1}{4^k}\sum_{I_1,I_2\neq\{\}}\sum_{\tau\in\mathcal{S}_\varepsilon}\frac{|\chi_\tau(h)|}{d_\tau}\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2.
\end{aligned}$$

The equality follows from Lemma 3 and the fact that the quantity in the absolute value sign is non-negative, and the last inequality follows from the fact that

$$\sum_{\tau\in\widehat{G}\setminus\mathcal{S}_\varepsilon}\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2 \leq \|\mathbf{b}\otimes\mathbf{b}\|^2 = 1.$$

Fix $I_1, I_2 \subseteq [k]$, $I_1, I_2 \neq \{\}$. Then,

$$\begin{aligned}
&\mathrm{E}_{\boldsymbol{\rho},\mathbf{b}}\left[\sum_{\tau\in\mathcal{S}_\varepsilon}\frac{|\chi_\tau(h)|}{d_\tau}\left\|\Pi_\tau^{\boldsymbol{\rho}^{I_1,I_2}}(\mathbf{b}\otimes\mathbf{b})\right\|^2\right] \\
&\leq \mathrm{E}_{\boldsymbol{\rho},\mathbf{b}}\left[\sum_{\tau\in\mathcal{S}_\varepsilon}\frac{|\chi_\tau(h)|}{d_\tau}\cdot\frac{d_\tau^2}{2}\left(\mathrm{E}_g\left[\left|\langle\mathbf{b}\left|\boldsymbol{\rho}^{I_1}(g)\right|\mathbf{b}\rangle\right|^2\right] + \mathrm{E}_g\left[\left|\langle\mathbf{b}\left|\boldsymbol{\rho}^{I_2}(g)\right|\mathbf{b}\rangle\right|^2\right]\right)\right] \\
&= \left(\sum_{\tau\in\mathcal{S}_\varepsilon}\frac{d_\tau|\chi_\tau(h)|}{2}\right)\cdot\left(\mathrm{E}_{\boldsymbol{\rho},\mathbf{b},g}\left[\left|\langle\mathbf{b}\left|\boldsymbol{\rho}^{I_1}(g)\right|\mathbf{b}\rangle\right|^2\right] + \mathrm{E}_{\boldsymbol{\rho},\mathbf{b},g}\left[\left|\langle\mathbf{b}\left|\boldsymbol{\rho}^{I_2}(g)\right|\mathbf{b}\rangle\right|^2\right]\right) \\
&\leq \frac{1}{|G|}\cdot\left(\sum_{\tau\in\mathcal{S}_\varepsilon} d_\tau|\chi_\tau(h)|\right)\cdot\left(\sum_{\nu\in\widehat{G}} d_\nu\right).
\end{aligned}$$

The first inequality is due to Lemma 4 and the second inequality is due to Lemma 5. Combining the above two upper bounds proves the present lemma. □

We are now ready to prove the main theorem of the paper.

Proof (of Theorem 1). The theorem is proved by using Lemmas 6, 2 and the concavity of the square-root function. The upper bound on δ_1 follows since the Cauchy-Schwartz inequality implies that $\sum_{\nu\in\widehat{G}} d_\nu \leq |\widehat{G}|^{1/2}\left(\sum_{\nu\in\widehat{G}} d_\nu^2\right)^{1/2} = |\widehat{G}|^{1/2}|G|^{1/2}$. □

Finally, we prove a simple lower bound irrespective of jointness on the total number of coset states t required by a measurement to distinguish a hidden subgroup H^g from the identity hidden subgroup.

Theorem 4. *Let G be a finite group and $H := \{1, h\}$ be an order two subgroup of G. Let $t \geq 1$ be an integer. Then,*

$$\left\|\mathrm{E}_g\left[\sigma_{H^g}^{\otimes t}\right] - \sigma_{\{1\}}^{\otimes t}\right\|_{\mathrm{tr}} < \frac{2^t}{|G|}\sum_{\tau\in\widehat{G}} d_\tau |\chi_\tau(h)|.$$

Proof. Let $\boldsymbol{\rho} \in \widehat{G}^{\otimes t}$, $I \subseteq [t]$, $I \neq \{\}$. Using arguments similar to those above, it is easy to see that

$$\begin{aligned}
\left\|\mathrm{E}_g[2^t\boldsymbol{\rho}((H^g)^t)] - \boldsymbol{\rho}(\{1\}^t)\right\|_{\mathrm{tr}} &= \left\|\mathrm{E}_g\left[\mathbb{1}_{d_{\boldsymbol{\rho}}} + \sum_{I\neq\{\}} \boldsymbol{\rho}^I(ghg^{-1})\right] - \mathbb{1}_{d_{\boldsymbol{\rho}}}\right\|_{\mathrm{tr}} \\
&= \left\|\sum_{I\neq\{\}} \mathrm{E}_g\left[\boldsymbol{\rho}^I(ghg^{-1})\right]\right\|_{\mathrm{tr}} \leq \sum_{I\neq\{\}} \left\|\mathrm{E}_g[\boldsymbol{\rho}^I(ghg^{-1})]\right\|_{\mathrm{tr}} \\
&= \sum_{I\neq\{\}} \left\|\bigoplus_{\tau\in\widehat{G}} \frac{\chi_\tau(h)}{d_\tau} \bigoplus_{j=1}^{a_\tau^{\boldsymbol{\rho}^I}} \mathbb{1}_{d_\tau}\right\|_{\mathrm{tr}} = \sum_{I\neq\{\}}\sum_{\tau\in\widehat{G}} a_\tau^{\boldsymbol{\rho}^I}|\chi_\tau(h)|.
\end{aligned}$$

Writing the density matrices in the Fourier basis and using Fact 3 we get,

$$\begin{aligned}
\left\|\mathrm{E}_g\left[\sigma_{H^g}^{\otimes t}\right] - \sigma_{\{1\}}^{\otimes t}\right\|_{\mathrm{tr}} &= \left\|\mathrm{E}_g\left[\frac{2^t}{|G|^t}\bigoplus_{\boldsymbol{\rho}}\bigoplus_{\mathbf{i}=1}^{d_{\boldsymbol{\rho}}} |\boldsymbol{\rho}^*,\mathbf{i}\rangle\langle\boldsymbol{\rho}^*,\mathbf{i}| \otimes \boldsymbol{\rho}((H^g)^t)\right]\right. \\
&\qquad \left. -\frac{1}{|G|^t}\bigoplus_{\boldsymbol{\rho}}\bigoplus_{\mathbf{i}=1}^{d_{\boldsymbol{\rho}}} |\boldsymbol{\rho}^*,\mathbf{i}\rangle\langle\boldsymbol{\rho}^*,\mathbf{i}| \otimes \boldsymbol{\rho}(\{1\}^t)\right\|_{\mathrm{tr}} \\
&= \left\|\frac{1}{|G|^t}\bigoplus_{\boldsymbol{\rho}}\bigoplus_{\mathbf{i}=1}^{d_{\boldsymbol{\rho}}} |\boldsymbol{\rho}^*,\mathbf{i}\rangle\langle\boldsymbol{\rho}^*,\mathbf{i}| \otimes \left(\mathrm{E}_g[2^t\boldsymbol{\rho}((H^g)^t)] - \boldsymbol{\rho}(\{1\}^t)\right)\right\|_{\mathrm{tr}} \\
&= \frac{1}{|G|^t}\sum_{\boldsymbol{\rho}} d_{\boldsymbol{\rho}} \left\|\mathrm{E}_g[2^t\boldsymbol{\rho}((H^g)^t)] - \boldsymbol{\rho}(\{1\}^t)\right\|_{\mathrm{tr}} \\
&\leq \frac{1}{|G|^t}\sum_{\boldsymbol{\rho}} d_{\boldsymbol{\rho}} \sum_{I\neq\{\}}\sum_{\tau\in\widehat{G}} a_\tau^{\boldsymbol{\rho}^I}|\chi_\tau(h)| \\
&= \sum_{I\neq\{\}}\sum_{\tau\in\widehat{G}} |\chi_\tau(h)| \left(\sum_{\boldsymbol{\rho}} \frac{d_{\boldsymbol{\rho}}^2}{|G|^t}\frac{a_\tau^{\boldsymbol{\rho}^I}}{d_{\boldsymbol{\rho}}}\right) = \sum_{I\neq\{\}}\sum_{\tau\in\widehat{G}} \frac{d_\tau|\chi_\tau(h)|}{|G|} \\
&< \frac{2^t}{|G|}\sum_{\tau\in\widehat{G}} d_\tau|\chi_\tau(h)|.
\end{aligned}$$

□

Corollary 5. *Any algorithm using a total of t coset states that distinguishes with constant probability between the case when the hidden subgroup is trivial and the case when the hidden subgroup is H^g for some $g \in G$ must satisfy $t = \Omega(\log(1/\eta))$.*

Proof. The algorithm can be viewed as a two-outcome POVM that outputs 1 with probability at least $2/3$ if the hidden subgroup is non-trivial, and 0 with probability at least $2/3$ if the hidden subgroup is trivial. Thus, the POVM distinguishes between the states $\mathrm{E}_g \left[\sigma_{H^g}^{\otimes t}\right]$ and $\sigma_{\{1\}}^{\otimes t}$ with constant total variation distance. Since the trace distance is always an upper bound on the total variation distance, invoking Theorem 4 completes the proof. □

The above corollary shows, for example, that any coset state based algorithm solving the HSP in $S_n \wr S_2$ needs a total number of $\Omega(n \log n)$ coset states. In the next section, we apply Theorem 1 to show a stronger result, namely, any algorithm solving the HSP in $S_n \wr S_2$ using polynomially many coset states needs to make measurements entangled across $\Omega(n \log n)$ coset states. However, Corollary 5 can sometimes prove non-trivial lower bounds on the total number of coset states for solving the HSP in groups G where Theorem 1 can only prove a constant lower bound on the order of entanglement. For example, the HSP in groups $G := A \rtimes \mathbb{Z}_2$, where A is an abelian group and $\mathbb{Z}_2$ acts on A by inversion can be solved by an algorithm using a total number of $O(\log |G|)$ coset states that measures one coset state at a time [EH00]. Using Corollary 5, one can show a matching $\Omega(\log |G|)$ lower bound on the total number of coset states when A is the cyclic group $\mathbb{Z}_n$, i. e., G is the dihedral group D_n. Using a different technique, Childs and Wocjan [CW07] in fact show an $\Omega(\log |G|)$ lower bound on the total number of coset states for the above groups for all abelian A.

4 Examples

4.1 Graph isomorphism and wreath products

The representation theory of the wreath product $G = S_n \wr S_2$ is well-known. The following is a summary of the necessary results, for more details we refer to Appendix A: the group $S_n \wr S_2$ has irreps $\kappa_{\lambda,\lambda'}$ of dimension $2d_\lambda d_{\lambda'}$, where $\lambda, \lambda' \in \widehat{S_n}$, $\lambda \neq \lambda'$. Define $h := (e, e, 1) \in G$, where e is the identity permutation in S_n. The character value of these irreps at h is zero. Furthermore, there are irreps ϑ_λ and ϑ'_λ of dimension d_λ^2, where $\lambda \in \widehat{S_n}$. The character values of ϑ_λ and ϑ'_λ at h are given by d_λ and $-d_\lambda$, respectively. The above irreps form a complete set of inequivalent irreps of G. The total number of irreps of G is $|\widehat{G}| = \binom{p(n)}{2} + 2p(n) \leq (p(n))^2$, where $p(n)$ denotes the number of partitions of n.

In order to apply Theorem 1 we choose $\varepsilon = n^{-n/5}$. Then

$$\mathcal{S}_\varepsilon = \left\{ \sigma \in \widehat{G} : \frac{|\chi_\sigma(h)|}{d_\sigma} \geq \varepsilon \right\} = \left\{ \vartheta_\lambda, \vartheta'_\lambda : d_\lambda \leq n^{n/5} \right\}.$$

Hence we obtain that

$$\sum_{\sigma \in \mathcal{S}_\varepsilon} d_\sigma^2 \leq 2 \cdot \sum_{\lambda \in \widehat{S_n}, d_\lambda \leq n^{n/5}} d_\lambda^4 \leq p(n) n^{4n/5} \leq n^{4n/5} e^{\nu\sqrt{n}}.$$

Here we have estimated the partition number as $p(n) = O(e^{\nu\sqrt{n}})$, where $\nu = \pi\sqrt{\frac{2}{3}}$. Using the notation of Theorem 1, we see that

$$\begin{aligned}
\delta_2^2 &:= \varepsilon + \left(\sum_{\sigma\in\mathcal{S}_\varepsilon} d_\sigma^2\right)\left(\frac{|\widehat{G}|}{|G|}\right)^{1/2} \\
&\leq n^{-n/5} + n^{4n/5}e^{\nu\sqrt{n}}\left(\frac{p(n)^2}{2(n!)^2}\right)^{1/2} \leq n^{-n/5} + \frac{n^{4n/5}e^{2\nu\sqrt{n}}}{\sqrt{2}n!} \\
&\leq n^{-n/6},
\end{aligned}$$

where we have used the fact that $n! \geq (n/e)^n$ for large n. Hence, we have proved the following corollary to Theorem 1:

Corollary 6. *Any algorithm operating on coset states that solves the hidden subgroup problem in $G = S_n \wr S_2$ in polynomial time has to make joint measurements on $k \geq 0.05n\log n$ coset states. The same is true for any algorithm that solves the hidden subgroup problem in S_n using coset states. Also, any efficient algorithm for isomorphism of two n-vertex graphs that uses the standard reduction to HSP in S_{2n} and then uses coset states to solve the HSP needs to make measurements entangled across $k \geq 0.05n\log n$ coset states.*

We remark that if we apply Theorem 1 to all the full-support involutions in S_{2n}, we only get a lower bound of $k = \Omega(n)$. This is because we use Roichman's [Roi96] upper bound on the normalized characters of S_{2n} in order to define $\mathcal{S}_\epsilon$, as in [MRS05], and Roichman's bound is always at least $e^{-O(n)}$. Since the involutive swaps form an exponentially small fraction of all the full-support involutions, it is possible that an average hidden full-support involution may be distinguishable from the hidden identity subgroup by a POVM with jointness $O(n)$ acting on $n^{O(1)}$-coset states. However, no such POVM is known and the best upper bound on the jointness required for this problem continues to be the $k = O(n\log n)$ information-theoretic one.

The arguments for the lower bound for $S_n \wr S_2$ given above extend in straightforward fashion to $G \wr S_2$, for many groups G. Since the hidden shift problem [CW07] in a group G with shift (g, g^{-1}), where $g \in G$, reduces to the HSP in $G \wr S_2$ with hidden subgroup generated by the involution $(g, g^{-1}, 1)$, our lower bound for HSP in $G \wr S_2$ implies a similar lower bound for the hidden shift problem in G. The following corollary shows that if the number of conjugacy classes in G is sufficiently small, solving the HSP in $G \wr S_2$ in polynomial time requires highly entangled measurements across many coset states of the hidden subgroup.

Corollary 7. *Fix a constant $\alpha < 1/2$. Suppose that the number of conjugacy classes in G satisfies $|\widehat{G}| < |G|^\alpha$. Fix a constant $\gamma < (1-2\alpha)/10$. Any algorithm operating on coset states that solves the hidden subgroup problem in $G \wr S_2$ in polynomial time has to make joint measurements on $k \geq \gamma\log|G|$ coset states.*

Proof. As also discussed in [MRS05], the connection between the representation theories of $S_n \wr S_2$ and S_n extends to $G \wr S_2$ and G for any group G. The group $G \wr S_2$ has irreps $\kappa_{\lambda,\lambda'}$ of dimension $2d_\lambda d_{\lambda'}$, where $\lambda, \lambda' \in \widehat{G}$, $\lambda \neq \lambda'$. Define $h := (e, e, 1) \in G \wr S_2$, where e is the identity element in G. The conjugacy class of h in $G \wr S_2$ is the set $\{(g, g^{-1}, 1)\}_{g\in G}$. The character value of these irreps at h is zero. Furthermore, there are irreps ϑ_λ and ϑ'_λ of $G \wr S_2$ of dimension d_λ^2, where $\lambda \in \widehat{G}$. The character values of ϑ_λ and ϑ'_λ at h are given by d_λ and $-d_\lambda$, respectively. The above irreps form a complete set of inequivalent irreps of $G \wr S_2$. The total number of irreps of $G \wr S_2$ is $\binom{|\widehat{G}|}{2} + 2|\widehat{G}| \leq |\widehat{G}|^2$.

In order to apply Theorem 1 we choose $\varepsilon = |G|^{(2\alpha-1)/5}$. Then

$$\mathcal{S}_\varepsilon = \left\{ \sigma \in \widehat{G \wr S_2} : \frac{|\chi_\sigma(h)|}{d_\sigma} \geq \varepsilon \right\} = \left\{ \vartheta_\lambda, \vartheta'_\lambda : \lambda \in \widehat{G}, d_\lambda \leq \frac{1}{\varepsilon} \right\}.$$

Hence we obtain that

$$\sum_{\sigma \in \mathcal{S}_\varepsilon} d_\sigma^2 \leq 2 \cdot \sum_{\lambda \in \widehat{G}, d_\lambda \leq 1/\varepsilon} d_\lambda^4 \leq 2|\widehat{G}|\varepsilon^{-4}.$$

Using the notation of Theorem 1, we see that

$$\begin{aligned} \delta_2^2 &:= \varepsilon + \left(\sum_{\sigma \in \mathcal{S}_\varepsilon} d_\sigma^2 \right) \left(\frac{|\widehat{G \wr S_2}|}{|G \wr S_2|} \right)^{1/2} \\ &\leq \varepsilon + 2|\widehat{G}|\varepsilon^{-4} \left(\frac{|\widehat{G}|^2}{2|G|^2} \right)^{1/2} \leq \varepsilon + \frac{2|\widehat{G}|^2}{\varepsilon^4 |G|} \\ &\leq \varepsilon + 2|G|^{2\alpha-1}\varepsilon^{-4} = 3|G|^{(2\alpha-1)/5}. \end{aligned}$$

Applying Theorem 1 allows us to conclude the statement of this corollary. □

4.2 The projective linear groups $\mathrm{PSL}(2, \mathbb{F}_q)$

The representation theory of the projective linear groups $G = \mathrm{PSL}(2, \mathbb{F}_q)$ over any finite field $\mathbb{F}_q$ is well-known. The following is a summary of the necessary results, for more details we refer to Appendix B. We treat the cases q even and q odd separately. In case q odd we have that $|\mathrm{PSL}(2, \mathbb{F}_q)| = \frac{q(q^2-1)}{2}$. There is one conjugacy class of $\frac{q(q\pm1)}{2}$ involutions (depending on whether $q \equiv 1$ or 3 modulo 4); let h denote a fixed member of this conjugacy class. The degrees of the irreps are given by $1, q, q \pm 1$, and $\frac{q\pm1}{2}$. The character values $|\chi(h)|$ can be upper bounded by 1, 1, 2, and 1, respectively. There is a total number of $|\widehat{G}| = \frac{q+5}{2}$ irreps.

In order to apply Theorem 1, we choose $\varepsilon = \frac{2}{q-1}$. Then

$$\mathcal{S}_\varepsilon = \{ \sigma \in \widehat{G} : \frac{|\chi_\sigma(h)|}{d_\sigma} \geq \varepsilon \} = \{ \mathbb{1} \}$$

contains only the trivial irrep. With this choice of the parameter ε we have that $\sum_{\sigma \in \boldsymbol{\mathcal{S}}_\varepsilon} d_\sigma^2 = 1$. and

$$\left(\frac{|\widehat{G}|}{|G|} \right)^{1/2} = \left(\frac{(q+5)/2}{q(q^2-1)/2} \right)^{1/2} = O(q^{-1}).$$

Hence, we can bound the parameter δ_2 used in Theorem 1 as follows:

$$\delta_2^2 := \varepsilon + \left(\sum_{\sigma \in \boldsymbol{\mathcal{S}}_\varepsilon} d_\sigma^2 \right) \left(\frac{|\widehat{G}|}{|G|} \right)^{1/2} \leq \frac{2}{q-1} + 1 \cdot O(q^{-1}) = O(q^{-1}).$$

The case $q = 2^n$, where $|\mathrm{PSL}(2, \mathbb{F}_{2^n})| = |\mathrm{SL}(2, \mathbb{F}_{2^n})| = q(q^2-1)$, can be treated similarly. There we use $\varepsilon = \frac{1}{q-1}$ which implies that $\delta_2^2 \leq O(q^{-1})$. Hence, using Theorem 1 we have shown the following result:

Corollary 8. *Let q be a prime power. Then any algorithm operating on coset states that solves the hidden subgroup problem in $G = \mathrm{PSL}(2, \mathbb{F}_q)$ in polynomial time has to make joint measurements on $k = \Omega(\log |G|) = \Omega(q)$ coset states.*

4.3 Special and general linear groups

Corollary 9. *Any algorithm that solves the HSP in* $\mathrm{SL}(2,\mathbb{F}_q)$ *or* $\mathrm{GL}(2,\mathbb{F}_q)$ *efficiently using coset states needs to make measurements entangled across* $k = \Omega(\log q)$ *registers.*

Proof. By Corollary 8 any algorithm solving the HSP in $\mathrm{PSL}(2,\mathbb{F}_q)$ efficiently using coset states needs to make measurements entangled across $k = \Omega(\log q)$ registers. The statement now follows from Lemma 1 by using the facts that $\mathrm{PSL}(2,\mathbb{F}_q) \cong \mathrm{SL}(2,\mathbb{F}_q)/\zeta(\mathrm{SL}(2,\mathbb{F}_q))$ and that $\mathrm{SL}(2,\mathbb{F}_q) \leq \mathrm{GL}(2,\mathbb{F}_q)$. □

Corollary 10. *Any algorithm that solves the HSP in* $\mathrm{GL}(n,\mathbb{F}_{p^m})$ *efficiently using coset states needs to make measurements entangled across* $k = \Omega(n(m\log p + \log n))$ *registers.*

Proof. Since $\mathrm{GL}(n,\mathbb{F}_{p^m})$ contains all $n \times n$ permutation matrices, a lower bound of $k = \Omega(n \log n)$ follows from Corollary 6 and Lemma 1. Also, we can use the embedding of $\mathrm{GL}(2,\mathbb{F}_{p^{nm}}) \leq \mathrm{GL}(2n,\mathbb{F}_{p^m})$ via $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} M_a & M_b \\ M_c & M_d \end{pmatrix}$, where for each $x \in \mathbb{F}_{p^{nm}}$ the matrix $M_x \in \mathrm{GL}(n,\mathbb{F}_{p^m})$ realizes multiplication by x with respect to a fixed basis of $\mathbb{F}_{p^{nm}}$ over $\mathbb{F}_{p^m}$. Hence by Lemma 1 we obtain that for the HSP in $\mathrm{GL}(2n,\mathbb{F}_{p^m})$ at least as much entanglement is necessary as in case of $\mathrm{GL}(2,\mathbb{F}_{p^{nm}})$. The latter has been bounded by $\Omega(nm\log p)$ in Corollary 9. □

4.4 Direct products of the form G^n

In this section we show that for a large class of finite groups G, efficient algorithms for HSP for direct products of the form G^n, where $n \geq 1$, require entangled measurements on at least $k = \Omega(n)$ coset states. Let G be a finite group and let $\widehat{G} = \{\sigma_1, \ldots, \sigma_m\}$ denote the irreducible representations of G. Recall that the centralizer $C(g)$ of an element $g \in G$ is the subgroup $C(g) := \{c \in G : cg = gc\}$. Let h be an involution in G, and let $\sigma \in \widehat{G}$. Then either $|\chi_\sigma(h)| = d_\sigma$ or $\frac{|\chi_\sigma(h)|}{d_\sigma} < 1 - \frac{2|C(h)|}{|G|}$ holds [Gal94]. We define $\varepsilon := (1 - \frac{2|C(h)|}{|G|})^t$, where $t = t(n)$ is a function of n to be determined later.

The irreps of G^n, where $n \geq 1$, are given by $\boldsymbol{\sigma} := \sigma_1 \otimes \ldots \otimes \sigma_n$, where $\sigma_i \in \widehat{G}$. We let $\Lambda := \{\sigma \in \widehat{G} : |\chi_\sigma(h)| = d_\sigma\}$, $\lambda := \sum_{\sigma\in\Lambda} d_\sigma^2$, and $\mu := \sum_{\sigma\in\widehat{G}\setminus\Lambda} d_\sigma^2 = |G| - \lambda$. The following property of the set

$$\boldsymbol{\mathcal{S}}_\varepsilon := \left\{ \boldsymbol{\sigma} \in \widehat{G}^n : \frac{|\chi_{\boldsymbol{\sigma}}(h,\ldots,h)|}{d_\sigma} \geq \varepsilon \right\}$$

holds for our choice of the parameter ε: if $\boldsymbol{\sigma} \in \boldsymbol{\mathcal{S}}_\varepsilon$ then necessarily at least $n-t$ positions σ_i have to be from Λ, i. e., have to satisfy $|\chi_{\sigma_i}(h)| = d_{\sigma_i}$. Indeed, otherwise we would have more than t positions σ_j in each of which $\frac{|\chi_{\sigma_j}(h)|}{d_{\sigma_j}} \leq 1 - \frac{2|C(h)|}{|G|}$, making the product less than ε. We next give an estimate for the quantity $\sum_{\boldsymbol{\sigma}\in\boldsymbol{\mathcal{S}}_\varepsilon} d_{\boldsymbol{\sigma}}^2$. For that we require the following lemma for estimating the tail of the binomial distribution.

Lemma 7. *Let* $\alpha, \beta > 0$*, let* $n \geq 1$*, and let* $t = n/c$*, where* $c > \frac{\alpha+\beta}{\beta}$*. Then*

$$\sum_{\ell=n-t}^{n} \binom{n}{\ell} \alpha^\ell \beta^{n-\ell} \leq \left(\alpha \left(\frac{ce(\alpha+\beta)}{\alpha} \right)^{1/c} \right)^n .$$

Proof. We have that

$$\begin{aligned}
\sum_{\ell=n-t}^{n} \binom{n}{\ell} \alpha^{\ell} \beta^{n-\ell} &= (\alpha+\beta)^n \sum_{\ell=n-t}^{n} \binom{n}{\ell} \left(\frac{\alpha}{\alpha+\beta}\right)^{\ell} \left(\frac{\beta}{\alpha+\beta}\right)^{n-\ell} \\
&\leq (\alpha+\beta)^n \binom{n}{n-t} \left(\frac{\alpha}{\alpha+\beta}\right)^{n-t} \\
&= \alpha^n \binom{n}{t} \left(\frac{\alpha+\beta}{\alpha}\right)^{t} \leq \alpha^n \left(\frac{ne(\alpha+\beta)}{t\alpha}\right)^{t} \\
&= \left(\alpha \left(\frac{ce(\alpha+\beta)}{\alpha}\right)^{1/c}\right)^{n},
\end{aligned}$$

where the first inequality follows from the union bound on probabilities and the second one from $\binom{n}{t} \leq \left(\frac{ne}{t}\right)^t$. □

Suppose we fix $\ell \geq n-t$ locations for putting in irreps from Λ. The contribution of this configuration to $\sum_{\boldsymbol{\sigma} \in \boldsymbol{\mathcal{S}}_\varepsilon} d_{\boldsymbol{\sigma}}^2$ is the sum of products of squares of dimensions of ℓ irreps from Λ and $n-\ell$ irreps from $\widehat{G} \setminus \Lambda$, which simplifies to $\lambda^{\ell} \mu^{n-\ell}$. Letting $\alpha := \lambda$, $\beta := \mu$, and $t = n/c$, with some constant c to be determined later, we obtain the following bound from Lemma 7:

$$\sum_{\boldsymbol{\sigma} \in \boldsymbol{\mathcal{S}}_\varepsilon} d_{\boldsymbol{\sigma}}^2 \leq \sum_{\ell=n-t}^{n} \binom{n}{\ell} \lambda^{\ell} \mu^{n-\ell} \leq \lambda^n \left(\left(\frac{ce|G|}{\lambda}\right)^{1/c}\right)^{n}.$$

Hence, for any given $\kappa > 0$ we can find a constant $c > 0$ such that $\sum_{\boldsymbol{\sigma} \in \boldsymbol{\mathcal{S}}_\varepsilon} d_{\boldsymbol{\sigma}}^2 \leq \lambda^n (1+\kappa)^n$ holds for all $n \geq c$. Note that $\sum_{\boldsymbol{\sigma} \in \boldsymbol{\mathcal{S}}_\varepsilon} d_{\boldsymbol{\sigma}} |\chi_{\boldsymbol{\sigma}}(h, \ldots, h)| \leq \sum_{\boldsymbol{\sigma} \in \boldsymbol{\mathcal{S}}_\varepsilon} d_{\boldsymbol{\sigma}}^2$. Also, observe that $\sum_{\boldsymbol{\rho} \in \widehat{G}^n} d_{\boldsymbol{\rho}} = \left(\sum_{\rho \in \widehat{G}} d_\rho\right)^n$. Now, we can bound the parameter δ_1 used in Theorem 1:

$$\begin{aligned}
\delta_1^2 &:= \varepsilon + \frac{1}{|G|^n} \left(\sum_{\boldsymbol{\sigma} \in \boldsymbol{\mathcal{S}}_\varepsilon} d_{\boldsymbol{\sigma}} |\chi_{\boldsymbol{\sigma}}(h, \ldots, h)|\right) \left(\sum_{\boldsymbol{\rho} \in \widehat{G}^n} d_{\boldsymbol{\rho}}\right) \\
&\leq \left(\left(1 - \frac{2|C(h)|}{|G|}\right)^{1/c}\right)^{n} + \frac{\lambda^n (1+\kappa)^n}{|G|^n} \left(\sum_{\rho \in \widehat{G}} d_\rho\right)^{n}.
\end{aligned}$$

For the following we make the assumption that $|G| > \lambda(1+\kappa) \left(\sum_{\rho \in \widehat{G}} d_\rho\right)$ holds. This implies that there exists a constant $\gamma_1 > 0$ such that $\delta_1 \leq \gamma_1^n$. Hence, we have proved the following corollary to Theorem 1.

Corollary 11. *Let G be a finite group and let $h \in G$ be an involution. Let $\widehat{G}$ denote the set of irreps of G and let $\Lambda := \{\sigma \in \widehat{G} : |\chi_\sigma(h)| = d_\sigma\}$. Suppose that $|G| > \left(\sum_{\sigma \in \Lambda} d_\sigma^2\right) \left(\sum_{\rho \in \widehat{G}} d_\rho\right)$ holds. Then any efficient algorithm operating on coset states that distinguishes between the case when the hidden subgroup is a conjugate of the subgroup $\langle (h, \ldots, h) \rangle \leq G^n$, and the case when the hidden subgroup is the identity subgroup in G^n, needs to make measurements entangled across $\Omega(n)$ registers.*

Recently, Alagic, Moore and Russell [AMR07] showed that any measurement on a single coset state gives exponentially little information about a hidden subgroup in the group G^n, where G is fixed and satisfies

a suitable condition. Their condition on G is weaker than our condition in Corollary 11, but they only prove lower bounds for algorithms measuring one coset state at a time. They also give several examples of families of groups satisfying their condition, including all non-abelian finite simple groups. In fact, the condition of Corollary 11 holds for all families of groups G considered in their paper, showing that efficient coset state based algorithms solving the HSP for their families of groups G^n need to make measurements entangled across $\Omega(n)$ registers.

From Corollary 11, it is easy to prove Corollary 12 via the Cauchy-Schwartz inequality.

Corollary 12. *Let G be a finite group and let $h \in G$ be an involution. Let $\widehat{G}$ denote the set of irreps of G and let $\Lambda := \{\sigma \in \widehat{G} : |\chi_\sigma(h)| = d_\sigma\}$. Suppose that $|G|^{1/2} > |\widehat{G}|^{1/2} \left(\sum_{\sigma\in\Lambda} d_\sigma^2\right)$ holds. Then any efficient algorithm operating on coset states that distinguishes between the case when the hidden subgroup is a conjugate of the subgroup $\langle(h,\ldots,h)\rangle \leq G^n$, and the case when the hidden subgroup is the identity subgroup in G^n, needs to make measurements entangled across $\Omega(n)$ registers.*

Using Corollary 12, we prove the following result.

Corollary 13. *Any efficient algorithm using only coset states that distinguishes between the case when the hidden subgroup is a conjugate of the subgroup $\langle(h,\ldots,h)\rangle \leq (S_m)^n$ where $h \in S_m$ is any involution and $m \geq 5$ is fixed, and the case when the hidden subgroup is the identity subgroup in $(S_m)^n$, needs to make measurements entangled across $\Omega(n)$ registers. The same holds also when $m = 4$ and $h = (1,2) \in S_4$.*

Proof. Let $G = S_m$, where $m \geq 5$, and let h be any involution in G. Recall that for $m \geq 5$ all irreps of S_m of degree greater than 1 are faithful [JK81, Theorem 2.1.13], and that the center of S_m is trivial. Since for faithful $\sigma \in \widehat{S_m}$ we have that $|\chi_\sigma(h)| = d_\sigma$ implies that h is in the center, we obtain that $|\chi_\sigma(h)| < d_\sigma$ for all $\sigma \in \widehat{S_m}$ with $d_\sigma > 1$. Hence $\Lambda = \{\mathbb{1}, \mathrm{alt}\}$ consists of the trivial and the alternating character only and we obtain that $\sum_{\sigma\in\Lambda} d_\sigma^2 = 2$. Since for $m \geq 5$ we have that $|G|^{1/2} = \sqrt{m!} > 2\sqrt{p(m)} = |\widehat{G}|^{1/2} \sum_{\sigma\in\Lambda} d_\sigma^2$, where $p(m)$ denotes the partition number of m, the statement for $m \geq 5$ follows from Corollary 12.

For $m = 4$ and $h = (1,2)$ we observe that the set Λ is again given by $\Lambda = \{\mathbb{1}, \mathrm{alt}\}$. We verify that the condition $|S_4|^{1/2} = \sqrt{24} > 2\sqrt{5} = |\widehat{S_4}|^{1/2} \sum_{\sigma\in\Lambda} d_\sigma^2$ holds. Hence the statement for this case also follows from Corollary 12. □

Remark: Corollary 13 implies a negative result for the HSP for conjugates of $H = \langle(h,\ldots,h)\rangle$ with $h = (1,2)$ over the groups $(S_4)^n$: it shows that any quantum algorithm which tries to solve the HSP over these groups needs to make use of POVMs entangled across $\Omega(n)$ copies of coset states σ_H. Note, however, that in fact there is an efficient algorithm for the HSP in $(S_4)^n$ using the *orbit coset* techniques of [FIM$^+$03]. This is based on the observation that the groups $(S_4)^n$ have derived series $(S_4)^n \rhd (A_4)^n \rhd (K_4)^n \rhd (\{e\})^n$, i. e., in the words of [FIM$^+$03] they are smoothly solvable since the length of the derived series is constant and the exponents of the factor groups are also constant.

Interestingly, this algorithm makes joint measurements on $n^{O(1)}$ states, but the states are not just coset states for the hidden subgroup H, but also coset states for various subgroups of the form HN, where $N \trianglelefteq (S_4)^n$. This example suggests that limiting oneself to HSP algorithms using only coset states of the hidden subgroup may be too restrictive, and one way to design efficient algorithms for the HSP making highly entangled measurements may be to use coset states for subgroups of G other than just the hidden subgroup H.

5 Acknowledgments

We thank Andrew Childs, Hirotada Kobayashi, Frédéric Magniez, Mario Szegedy, and Umesh Vazirani for helpful discussions and comments.

References

[AKN98] D. Aharonov, A. Kitaev, and N. Nisan. Quantum circuits with mixed states. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing (STOC'98)*, pages 20–30, 1998.

[AMR07] G. Alagic, C. Moore, and A. Russell. Quantum algorithms for Simon's problem over general groups. In *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'07)*, pages 1217–1224, 2007. also: ArXiv preprint quant-ph/0603251.

[AT03] D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing (STOC'03)*, pages 20–29, 2003. Also: ArXiv preprint quant–ph/0301023.

[BBBV97] C. Bennett, E. Bernstein, G. Brassard, and U. Vazirani. Strengths and weaknesses of quantum computing. *SIAM J. Comput.*, 26(5):1510–1523, 1997.

[BCD05] D. Bacon, A. Childs, and W. van Dam. From optimal measurement to efficient quantum algorithms for the hidden subgroup problem over semidirect product groups. In *Proceedings of the 46th Annual IEEE Symposium on Foundations of Computer Science (FOCS'05)*, 2005. Also: ArXiv preprint quant–ph/0504083.

[BCD06] D. Bacon, A. Childs, and W. van Dam. Optimal measurements for the dihedral hidden subgroup problem. *Chicago Journal of Theoretical Computer Science*, 2006. Also: ArXiv preprint quant–ph/0501044.

[Bea97] R. Beals. Quantum computation of Fourier transforms over the symmetric groups. In *Proceedings of the Symposium on Theory of Computing (STOC'97)*, pages 48–53, El Paso, Texas, 1997.

[BH97] G. Brassard and P. Høyer. An exact polynomial–time algorithm for Simon's problem. In *Proceedings of Fifth Israeli Symposium on Theory of Computing and Systems*, pages 12–33. ISTCS, IEEE Computer Society Press, 1997. Also: ArXiv preprint quant–ph/9704027.

[BZ99] Y. G. Berkovich and E. M. Zhmud. *Characters of Finite Groups, part 2*, volume 181 of *Translations of Mathematical Monographs*. American Mathematical Society, 1999.

[CW07] A. Childs and P. Wocjan. On the quantum hardness of solving isomorphism problems as nonabelian hidden shift problems. *Quantum Information and Computation*, 7(5&6):371–382, 2007. Also: ArXiv preprint quant–ph/0510185.

[EH00] M. Ettinger and P. Høyer. On quantum algorithms for noncommutative hidden subgroups. *Advances in Applied Mathematics*, 25(3):239–251, 2000.

[EHK99a] M. Ettinger, P. Høyer, and E. Knill. A quantum observable for the graph isomorphism problem. ArXiv preprint quant–ph/9901029, 1999.

[EHK99b] M. Ettinger, P. Høyer, and E. Knill. Hidden subgroup states are almost orthogonal. ArXiv preprint quant–ph/9901034, 1999.

[EHK04] M. Ettinger, P. Høyer, and E. Knill. The quantum query complexity of the hidden subgroup problem is polynomial. *Information Processing Letters*, 91(1):43–48, 2004. Also: ArXiv preprint quant–ph/0401083.

[FH91] W. Fulton and J. Harris. *Representation Theory: A First Course*, volume 129 of *Graduate Texts in Mathematics*. Springer, 1991.

[FIM+03] K. Friedl, G. Ivanyos, F. Magniez, M. Santha, and P. Sen. Hidden translation and orbit coset in quantum computing. In *Proceedings of the 35th Annual ACM Symposium on Theory of Computing (STOC'03)*, pages 1–9, 2003. Also: ArXiv preprint quant–ph/0211091.

[Gal94] P. X. Gallagher. Character values at involutions. *Proceeedings of the American Mathematical Society*, 120(3):657–659, 1994.

[Gav04] D. Gavinsky. Quantum solution to the hidden subgroup problem for Poly-Near-Hamiltonian groups. *Quantum Information and Computation*, 4(3):229–235, 2004.

[Gro96] L. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing (STOC'96)*, pages 212–219, 1996. Also: ArXiv preprint quant–ph/9605043.

[GSVV04] M. Grigni, L. Schulman, M. Vazirani, and U. Vazirani. Quantum mechanical algorithms for the nonabelian hidden subgroup problem. *Combinatorica*, pages 137–154, 2004.

[Hal02] S. Hallgren. Polynomial-time quantum algorithms for Pell's equation and the principal ideal problem. In *Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC'02)*, pages 653–658, 2002.

[Hal05] S. Hallgren. Fast quantum algorithms for computing the unit group and class group of a number field. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC'05)*, pages 468–474, 2005.

[HMR+06] R. Hallgren, C. Moore, M. Rötteler, A. Russell, and P. Sen. Limitations of quantum coset states for graph isomorphism. In *Proceedings of the 38th Annual ACM Symposium on Theory of Computing (STOC'06)*, pages 604–617, 2006.

[HRT03] S. Hallgren, A. Russell, and A. Ta-Shma. The hidden subgroup problem and quantum computation using group representations. *SIAM Journal on Computing*, 32(4):916–934, 2003.

[IMS03] G. Ivanyos, F. Magniez, and M. Santha. Efficient quantum algorithms for some instances of the non-abelian hidden subgroup problem. *International Journal of Foundations of Computer Science*, pages 723–740, 2003. Also: ArXiv preprint quant–ph/0102014.

[Isa76] I. M. Isaacs. *Character Theory of Finite Groups*. Academic Press, 1976.

[ISS07] G. Ivanyos, L. Sanselme, and M. Santha. An efficient quantum algorithm for the hidden subgroup problem in extraspecial groups. In *Proceedings of the 24th Annual Symposium on Theoretical Aspects of Computer Science (STACS'07)*, volume 4393 of *Lecture Notes in Computer Science*, pages 586–597. Springer-Verlag, 2007. Also: ArXiv preprint quant–ph/0701235.

[JK81] G. James and A. Kerber. *The Representation Theory of the Symmetric Group*. Addison-Wesley, Reading, 1981.

[Kit95] A. Yu. Kitaev. Quantum measurements and the abelian stabilizer problem. ArXiv preprint quant–ph/9511026, 1995.

[KST93] J. Köbler, U. Schöning, and J. Torán. *The Graph Isomorphism Problem*. Birkhäuser, 1993.

[Kup05] G. Kuperberg. A subexponential-time quantum algorithm for the dihedral hidden subgroup problem. *SIAM Journal on Computing*, 35(1):170–188, 2005. Also arxiv preprint quant–ph/0302112.

[LR92] John D. Lafferty and Daniel Rockmore. Fast Fourier analysis for SL_2 over a finite field and related numerical experiments. *Experimental Mathematics*, 1(2):*115*–139, 1992.

[ME98] M. Mosca and A. Ekert. The hidden subgroup problem and eigenvalue estimation on a quantum computer. In *Quantum Computing and Quantum Communications*, volume 1509 of *Lecture Notes in Computer Science*, pages 174–188. Springer-Verlag, 1998.

[MR05a] C. Moore and A. Russell. Explicit multiregister measurements for hidden subgroup problems. ArXiv preprint quant–ph/0504067, 2005.

[MR05b] C. Moore and A. Russell. The symmetric group defies strong Fourier sampling: Part II. ArXiv preprint quant–ph/0501066, 2005.

[MRRS07] C. Moore, D. Rockmore, A. Russell, and L. Schulman. The power of strong Fourier sampling: quantum algorithms for affine groups and hidden shifts. *SIAM Journal on Computing*, 37(3):938–958, 2007. Also: ArXiv preprint quant–ph/0503095.

[MRS05] C. Moore, A. Russell, and L. Schulman. The symmetric group defies strong Fourier sampling. In *Proceedings of the 46th Annual IEEE Symposium on the Foundations of Computer Science (FOCS'05)*, pages 479–488, 2005. Also: ArXiv preprint quant–ph/0501056.

[MRŚ07] C. Moore, A. Russell, and P. Śniady. On the impossibility of a quantum sieve algorithm for graph isomorphism. In *Proceedings of the 39th Annual ACM Symposium on Theory of Computing (STOC'07)*, pages 536–545, 2007. Also: ArXiv preprint quant-ph/0612089.

[MRV07] C. Moore, A. Russell, and U. Vazirani. A classical one-way function to confound quantum adversaries. ArXiv preprint quant–ph/0701115, 2007.

[NC00] M. Nielsen and I. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.

[Reg04a] O. Regev. New lattice-based cryptographic constructions. *Journal of the ACM*, 51(6):899–942, 2004.

[Reg04b] O. Regev. Quantum computation and lattice problems. *SIAM Journal on Computing*, 33(2):738–760, 2004.

[Roi96] Y. Roichman. Upper bound on the characters of the symmetric groups. *Inventiones Mathematicae*, 125:451–485, 1996.

[RRS09] J. Radhakrishnan, M. Rötteler, and P. Sen. Random measurement bases, quantum state distinction and applications to the hidden subgroup problem. *Algorithmica*, 55(3):490–516, 2009. Also: ArXiv preprints quant–ph/0503114 and quant–ph/0512085.

[Ser77] J. P. Serre. *Linear Representations of Finite Groups*. Springer, 1977.

[Sho97] P. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM Journal on Computing*, 26(5):1484–1509, 1997.

[Sim94] D. R. Simon. On the power of quantum computation. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science (FOCS'94)*, pages 116–123, Los Alamitos, CA, 1994. Institute of Electrical and Electronic Engineers Computer Society Press.

[SV05] A. Schmidt and U. Vollmer. Polynomial time quantum algorithm for the computation of the unit group of a number field. In *Proceedings of the 37th Annual ACM Symposium on Theory of Computing (STOC'05)*, pages 475–480, 2005.

A Representations of $S_n \wr S_2$

We describe the irreducible representations of the wreath product $S_n \wr S_2$, i. e., the group $(S_n \times S_n) \rtimes \mathbb{Z}_2$, where $\mathbb{Z}_2$ acts on $S_n \times S_n$ by swapping the two coordinates. We will also get formulas for the character values under these representations in terms of the character values of irreducible representations of S_n.

Let $\widehat{S_n} = \{\sigma_i : i = 1, \ldots, p(n)\}$ denote the irreducible representations of S_n, where $p(n)$ denotes the number of partitions of n. Denote the degree of $\sigma_i \in \widehat{S_n}$ by d_i. Letting

$$N := S_n \times S_n \quad \text{and} \quad G := (S_n \times S_n) \rtimes \mathbb{Z}_2$$

we have that $N \lhd G$ is a normal subgroup of index 2. The irreducible representations of N are given by $\widehat{N} = \{\sigma_i \otimes \sigma_j : i, j = 1, \ldots, p(n)\}$ and we define the shorthand $\phi_{i,j} := \sigma_i \otimes \sigma_j$. Define $t := (e, e, 1) \in G$, where e is the identity permutation in S_n. A transversal of N in G is given by $T = \{(e, e, 0), t\}$. Then t acts on $\widehat{N}$ as $(\sigma_i \otimes \sigma_j)^t = (\sigma_j \otimes \sigma_i)$. Hence we have that $\phi_{i,j}^t = \phi_{j,i}$. Since all $\phi_{i,j}$ are pair-wise inequivalent, we obtain the following two cases from Clifford's Theorem [Isa76].

(i) $i = j$. Then $\phi_{i,j} \cong \phi_{i,j}^t$. Hence $\phi_{i,j}$ has precisely 2 pairwise inequivalent extensions to G. One of these extensions is $\vartheta_i = \overline{\phi_{i,i}}$ in which the image of t permutes the tensor factors of $\mathbb{C}^{d_i} \otimes \mathbb{C}^{d_i}$, where $d_i = \deg(\sigma_i)$. Hence if $\{e_k : k = 1, \ldots, d_i\}$ denotes the standard basis of $\mathbb{C}^{d_i}$ then $\vartheta_i(t)$ is given by the matrix SWAP_{d_i} which maps $e_k \otimes e_\ell \mapsto e_\ell \otimes e_k$. The other extension ϑ_i' of $\phi_{i,i}$ to G is given by defining the image of t to be $\vartheta'(t) := -\vartheta_i(t)$. Note that both extensions have degree d_i^2. The character value $\mathrm{tr}(\vartheta_i(t))$ is given by the number of invariant tensors under the swap operation, i. e., $\mathrm{tr}(\vartheta_i(t)) = d_i$ and $\mathrm{tr}(\vartheta_i'(t)) = -d_i$.

(ii) $i \neq j$. Then $\phi_{i,j} \ncong \phi_{i,j}^t = \phi_{j,i}$. Hence $\kappa_{i,j} := \phi_{i,j} \uparrow_T G$ is irreducible. Moreover, we have that $(\phi_{i,j} \uparrow_T G) \downarrow N = \phi_{i,j} \oplus \phi_{j,i}$ and

$$(\phi_{i,j} \uparrow_T G)(t) = \begin{pmatrix} \mathbf{0}_{d_i d_j} & \mathbb{1}_{d_i d_j} \\ \mathbb{1}_{d_i d_j} & \mathbf{0}_{d_i d_j} \end{pmatrix}.$$

The facts relevant for this paper are summarized in Table 1. Overall, there are $\binom{p(n)}{2}$ pairwise inequivalent irreducible representations $\kappa_{i,j} \in \widehat{G}$, one for each pair i, j such that $i \neq j$. We have that the degree of $\kappa_{i,j}$ is given by $2d_id_j$. The character $\chi_{i,j}$ of $\kappa_{i,j}$ satisfies $\kappa_{i,j}(t) = 0$ for all $i \neq j$. Furthermore, there are $2p(n)$ pairwise inequivalent irreducible representations ϑ_i and ϑ'_i.

Irrep	Irrep on (π, μ, e)	Char. on (π, μ, e)	Irrep on t	Char. on t
ϑ_i	$\sigma_i(\pi) \otimes \sigma_i(\mu)$	$\chi_i(\pi)\chi_i(\mu)$	SWAP_{d_i}	d_i
ϑ'_i	$\sigma_i(\pi) \otimes \sigma_i(\mu)$	$\chi_i(\pi)\chi_i(\mu)$	$-\mathrm{SWAP}_{d_i}$	$-d_i$
$\kappa_{i,j}$	$\begin{pmatrix} \sigma_i(\pi) \otimes \sigma_j(\mu) & \mathbf{0}_{d_id_j} \\ \mathbf{0}_{d_id_j} & \sigma_j(\pi) \otimes \sigma_i(\mu) \end{pmatrix}$	$\chi_i(\pi)\chi_j(\mu) + \chi_j(\pi)\chi_i(\mu)$	$\begin{pmatrix} \mathbf{0}_{d_id_j} & \mathbb{1}_{d_id_j} \\ \mathbb{1}_{d_id_j} & \mathbf{0}_{d_id_j} \end{pmatrix}$	0

Table 1: Irreducible representations of the wreath products $G = (S_n \times S_n) \rtimes Z_2$. Shown are the images of elements of the form (π, μ, e) and $t = (e, e, 1)$ under the irreducible representations of G.

B Representations of $\mathrm{PSL}(2, \mathbb{F}_q)$

We briefly recall some facts from the representation theory of the projective linear groups $\mathrm{PSL}(2, \mathbb{F}_q)$, where q is a prime power. Good references on the complex representation theory of these groups are available, see e .g, [BZ99, FH91, LR92]. We treat the cases q odd and $q = 2^n$ separately and begin by describing the conjugacy classes of involutions and the irreducible representations of $\mathrm{PSL}(2, \mathbb{F}_q)$ for q odd. Recall that for q odd, the center of $\mathrm{SL}(2, \mathbb{F}_q)$ consists only of the identity matrix and the matrix

$$c := \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Once the characters of $\mathrm{SL}(2, \mathbb{F}_q)$ are known, we therefore have to filter out only those characters χ for which $\chi(c) = \chi(1)$ holds in order to obtain the irreducible representations of $\mathrm{PSL}(2, \mathbb{F}_q)$.

B.1 The case $\mathrm{PSL}(2, \mathbb{F}_q)$ where $q \equiv 1 \mod 4$

The involutions are given by conjugates of the residue class of $h = \overline{\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}} \in \mathrm{PSL}(2, \mathbb{F}_q)$, where the bar denotes the fact that we are using coset representatives with respect to the center $\langle c \rangle$ of $\mathrm{SL}(2, \mathbb{F}_q)$. There is a total of $\frac{q(q-1)}{2}$ many involutions that are conjugates of h. The characters and their values on h are summarized in Table 2.

B.2 The case $\mathrm{PSL}(2, \mathbb{F}_q)$ where $q \equiv 3 \mod 4$

Similar to the previous case all involutions are conjugate to the element h defined as above. However, now there are $\frac{q(q+1)}{2}$ involutions conjugate to h. The characters and their values on h are summarized below in Table 3.

Irrep name	Parameters	Number of irreps	Degree	Character value at h
$\mathbb{1}$	—	1	1	1
ψ	—	1	q	1
θ_k	$k = 2, 4, \ldots, \frac{q-1}{2}$	$\frac{q-1}{4}$	$q-1$	0
χ_j	$j = 2, 4, \ldots, \frac{q-5}{2}$	$\frac{q-5}{4}$	$q+1$	$2(-1)^{k/2}$
ζ_ℓ	$\ell = 1, 2$	2	$\frac{q+1}{2}$	$(-1)^{(q-1)/4}$

Table 2: Irreducible representations of $\mathrm{SL}(2, \mathbb{F}_q)$, where $q \equiv 1 \mod 4$. Shown are the irreducible representations, which come in several natural series, the total number of irreducible representations of a given degree, and the character value at the involution h.

Irrep name	Parameters	Number of irreps	Degree	Character value at h
$\mathbb{1}$	—	1	1	1
ψ	—	1	q	-1
θ_k	$k = 2, 4, \ldots, \frac{q-3}{2}$	$\frac{q-3}{4}$	$q-1$	$2(-1)^{k/2+1}$
χ_j	$j = 2, 4, \ldots, \frac{q-3}{2}$	$\frac{q-3}{4}$	$q+1$	0
η_ℓ	$\ell = 1, 2$	2	$\frac{q-1}{2}$	$(-1)^{\frac{q+1}{4}+1}$

Table 3: Irreducible representations of $\mathrm{SL}(2, \mathbb{F}_q)$, where $q \equiv 3 \mod 4$. Shown are the irreducible representations, which come in several natural series, the total number of irreducible representations of a given degree, and the character value at the involution h.

B.3 The case $\mathrm{PSL}(2, \mathbb{F}_q)$ where $q = 2^n$

This case behaves quite differently from the case q odd. First, observe that in this case the center is trivial, i. e., $\mathrm{PSL}(2, \mathbb{F}_{2^n}) = \mathrm{SL}(2, \mathbb{F}_{2^n})$. All involutions in $\mathrm{SL}(2, \mathbb{F}_{2^n})$ are conjugate to the element

$$h = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathrm{SL}(2, \mathbb{F}_q),$$

and there is a total number of $q^2 - 1$ of such involutions. The characters and their values on h are summarized in Table 4.

Irrep name	Parameters	Number of irreps	Degree	Character value on h
$\mathbb{1}$	—	1	1	1
ψ	—	1	q	0
θ_k	$k = 1, 2, \ldots, \frac{q}{2}$	$\frac{q}{2}$	$q-1$	-1
χ_j	$j = 1, 2, \ldots, \frac{q-2}{2}$	$\frac{q-2}{2}$	$q+1$	1

Table 4: Irreducible representations of $\mathrm{SL}(2, \mathbb{F}_{2^n})$. Shown are the irreducible representations, which come in several natural series, the total number of irreducible representations of a given degree, and the character value at the involution h.