Quantum Information and Computation, Vol. 15, No. 3&4 (2015) 0260–0294
© Rinton Press

LIMITATIONS OF SINGLE COSET STATES AND QUANTUM ALGORITHMS FOR CODE EQUIVALENCE

HANG DINH
Indiana University South Bend
htdinh@iusb.edu

CRISTOPHER MOORE
Santa Fe Institute
moore@santafe.edu

ALEXANDER RUSSELL
University of Connecticut
acr@cse.uconn.edu

Received December 17, 2013
Revised July 19, 2014

Quantum computers can break the RSA, El Gamal, and elliptic curve public-key cryptosystems, as they can efficiently factor integers and extract discrete logarithms. The power of such quantum attacks lies in *quantum Fourier sampling*, an algorithmic paradigm based on generating and measuring coset states. In this article we extend previous negative results of quantum Fourier sampling for Graph Isomorphism, which corresponds to hidden subgroups of order two (over S_n), to several cases corresponding to larger hidden subgroups. For one case, we strengthen some results of Kempe, Pyber, and Shalev on the Hidden Subgroup Problem over the symmetric group. In another case, we show the failure of quantum Fourier sampling on the Hidden Subgroup Problem over the general linear group $\mathsf{GL}_2(\mathbb{F}_q)$. The most important case corresponds to Code Equivalence, the problem of determining whether two given linear codes are equivalent to each other up to a permutation of the coordinates. Our results suggest that for many codes of interest—including generalized Reed Solomon codes, alternant codes, and Reed-Muller codes—solving these instances of Code Equivalence via Fourier sampling appears to be out of reach of current families of quantum algorithms.

Keywords: Quantum Fourier sampling, hidden subgroup problem, Code Equivalence.

Communicated by: R Jozsa & M Mosca

1 Introduction

Quantum Fourier Sampling (QFS) is the key ingredient in nearly all known efficient quantum algorithms for algebraic problems, including Shor's algorithms for factorization and discrete logarithm [30] and Simon's algorithm [33]. Shor's algorithm relies on quantum Fourier sampling over the cyclic group $\mathbb{Z}_N$, while Simon's algorithm uses quantum Fourier sampling over $\mathbb{Z}_2^n$. In general, these algorithms solve instances of the *Hidden Subgroup Problem* (HSP) over a finite group G. Given a function f on G whose level sets are left cosets of some unknown subgroup $H < G$, i.e., such that f is constant on each left coset of H and distinct on different left cosets, they find a set of generators for the subgroup H.

The standard approach to this problem treats f as a black box and applies f to a uniform superposition over G, producing the coset state $|cH\rangle = (1/\sqrt{|H|})\sum_{h\in H}|ch\rangle$ for a random c. We then measure $|cH\rangle$ in a Fourier basis $\{|\rho,i,j\rangle\}$ for the space $\mathbb{C}[G]$, where ρ is an irrep[a]of G and i,j are row and column indices of a matrix $\rho(g)$. In the *weak* form of Fourier sampling, only the representation name ρ is measured, while in the *strong* form, both the representation name and the matrix indices are measured, the latter in a chosen basis. This produces probability distributions from which classical information can be extracted to recover the subgroup H. Moreover, since $|cH\rangle$ is block-diagonal in the Fourier basis, the optimal measurement of the coset state can always be described in terms of strong Fourier sampling.

Understanding the power of Fourier sampling in nonabelian contexts has been an ongoing project, and a sequence of negative results [10, 22, 11] have suggested that the approach is inherently limited when the underlying groups are rich enough. In particular, Moore, Russell, and Schulman [22] showed that over the symmetric group, even the strong form of Fourier sampling cannot efficiently distinguish the conjugates of most order-2 subgroups from each other or from the trivial subgroup. That is, for any $\sigma \in S_n$ with large support, and most $\pi \in S_n$, if $H = \{1, \pi^{-1}\sigma\pi\}$ then strong Fourier sampling, and therefore any measurement we can perform on the coset state, yields a distribution which is exponentially close to the distribution corresponding to $H = \{1\}$. This result implies that Graph Isomorphism cannot be solved by the naive reduction to strong Fourier sampling. Hallgren et al. [11] strengthened these results, demonstrating that even entangled measurements on $o(\log n!)$ coset states yield essentially no information. However, both the results obtained by Moore et al. [22] for single-register Fourier sampling and those obtained by Hallgren et al. [11] for multi-register Fourier sampling apply specifically to subgroups of order two.

Kempe and Shalev [15] showed that weak Fourier sampling of single coset states in S_n cannot distinguish the trivial subgroup from larger subgroups H with polynomial size and non-constant minimal degree.[b]They conjectured, conversely, that if a subgroup $H < S_n$ can be distinguished from the trivial subgroup by weak Fourier sampling, then the minimal degree of H must be constant. Their conjecture was later proved by Kempe, Pyber, and Shalev [16].

We emphasize that previous results on limitations of strong Fourier sampling did not handle subgroups of order more than two. Our major contribution is to develop some tools to handle these cases. Additionally, we use these tools to investigate the Hidden Subgroup Problem instances that arise from Code Equivalence—the problem of deciding whether two given linear codes are equal up to a fixed permutation on the codeword coordinates.

Petrank and Roth [26] showed that Code Equivalence is unlikely to be NP-complete, but is at least as hard as Graph Isomorphism. We consider a search version of Code Equivalence: Given $k \times n$ generator matrices (or check matrices) M and M' of two equivalent linear q-ary codes, find a pair of matrices (S,P), where S is an invertible square matrix over $\mathbb{F}_q$ and P is a permutation matrix, such that $M' = SMP$. This search version of Code Equivalence has an immediate presentation as a *hidden subgroup problem*, suggesting that one might be able to develop an efficient quantum algorithm for it via the quantum Fourier transform. In this article, however, we show that for many families of linear codes, the resulting instance of the hidden subgroup problem requires entangled measurements of the coset states and, hence, appears to be beyond the reach of current methods.

[a] Throughout the paper, we write “irrep” as short for “irreducible representation.”

[b] The minimal degree of a permutation group H is the minimal number of points moved by a non-identity element of H.

1.1 A Sketch of Our Results

To state our results, we say that a subgroup $H < G$ is *indistinguishable by strong Fourier sampling* (over G) if the conjugate subgroups $g^{-1}Hg$ cannot be distinguished from each other (or from the trivial subgroup) by measuring the coset state in an arbitrary Fourier basis. A precise definition is presented in Section 3.2. Since the optimal measurement of a coset state can always be expressed as an instance of strong Fourier sampling, these results imply that no measurement of a single coset state yields any useful information about H. Based on the strategy of Moore et al. [22], we first develop a general framework, formalized in Theorem 1, to determine indistinguishability of a subgroup by strong Fourier sampling. We emphasize that the results of Moore et al. [22] cover the case where the subgroup has order two. Our principal contribution is to show how to extend their methods to more general subgroups.

We then apply this general framework to three classes of groups: the symmetric group S_n, the finite general linear group $\mathsf{GL}_2(\mathbb{F}_q)$, and the wreath product $(\mathsf{GL}_k(\mathbb{F}_q) \times S_n) \wr \mathbb{Z}_2$. For the symmetric group, we extend the results of [22] to larger subgroups of S_n. Specifically, we show that any subgroup $H < S_n$ with minimal degree $m \geq \Theta(\log|H|) + \omega(\log n)$ is indistinguishable by strong Fourier sampling over S_n. This partially extends the results of Kempe et al. [16], which apply only to weak Fourier sampling.

For the general linear group $\mathsf{GL}_2(\mathbb{F}_q)$, we gave the first negative result regarding the power of strong Fourier sampling over $\mathsf{GL}_2(\mathbb{F}_q)$. In particular, we show that any subgroup $H < \mathsf{GL}_2(\mathbb{F}_q)$ that does not contain non-identity scalar matrices and has order $|H| \leq q^{\delta}$ for some $\delta < 1/2$ is indistinguishable by strong Fourier sampling. Examples of such subgroups are those generated by a constant number of triangular unipotent matrices.

The case $G = (\mathsf{GL}_k(\mathbb{F}_q) \times S_n) \wr \mathbb{Z}_2$ corresponds to the hidden subgroup problem reduced from the search version of Code Equivalence as discussed above. We apply our general framework to this class of wreath products to bound the distinguishability of the hidden subgroup $K < G$ that represents the Code Equivalence instance given by the $k \times n$ matrix M. Note, here, that matrix M has entries from a finite field $\mathbb{F}_{q^{\ell}}$, where $\ell = 1$ when M is a generator matrix of a q-ary linear code. Our bound, given in Corollary 1 of Theorem 4, depends on the column rank[c]of the matrix M as well as the minimal degree and the size of the *automorphism group* $\mathrm{Aut}(M)$, where $\mathrm{Aut}(M)$ is defined in Subsection 6 as the set of all permutations P on the columns of M such that $M = SMP$ for some $S \in \mathsf{GL}_k(\mathbb{F}_q)$. In particular, the subgroup K is indistinguishable by strong Fourier sampling if M has column rank at least $k - o(\sqrt{n})/\ell$, and the automorphism group $\mathrm{Aut}(M)$ has minimal degree $\Omega(n)$ and size $\mathrm{e}^{o(n)}$.

We call a family of linear codes an *HSP-hard instance of Code Equivalence* if they have a generator matrix or check matrix for which the subgroup K is indistinguishable. Using the aforementioned bound on the distinguishability of subgroup K, we go on to identify three families of linear codes that are HSP-hard instances of Code Equivalence, including rational Goppa codes (or Generalized Reed Solomon codes), alternant codes, and Reed-Muller codes. The case of q-ary Goppa codes – a subclass of alternant codes – has been covered in our preliminary results [7].

1.2 Ramifications for Code-based Cryptosystems

As typical code-based cryptosystems are directly related to the Code Equivalence problem, it is interesting to explore what our results imply about the possibility of quantum attacks on these cryptosystems.

[c] The column rank of M is understood to be over the field $F_{q^{\ell}}$.

The most popular forms of code-based cryptosystems are based on either the McEliece cryptosystem [20] or the Niederreiter cryptosystem [24], and are conventionally built over binary Goppa codes. The private key of a McEliece (resp., Niederreiter) cryptosystem is a triple (S, M, P), where S is an invertible matrix over $\mathbb{F}_q$, P is a permutation matrix, and M is a $k \times n$ generator matrix (resp., check matrix) for a q-ary error-correcting code that permits efficient decoding. The public key is the matrix $M' = SMP$. If both M and M' are known to an adversary, the problem of recovering S and P (the remainder of the secret key) is precisely the version of Code Equivalence described above. If M and M' have full rank, which is always the case when M and M' are generator matrices, then given P we can find S by linear algebra. Thus the potentially hard part of the problem is finding the hidden permutation P.

We call an adversary apprised of both M and M' a *known-code* adversary. In our previous article [7], we noted that our results on Goppa codes imply that the natural quantum attack available to a known-code adversary yields hard cases of the hidden subgroup problem, and asserted that this should bolster our confidence in the post-quantum security of the McEliece cryptosystem.

However, the classical *support splitting algorithm* (SSA) of Sendrier [28] can efficiently solve Code Equivalence for the family of linear codes with small hull, which includes Goppa codes. Thus for McEliece/Niederreiter cryptosystems based on Goppa codes, the known-code adversary is too powerful: it can break the cryptosystem classically. Therefore, the hardness of the corresponding instances of the HSP has little bearing on the post-quantum security of these code-based cryptosystems.

The situation is similar in many ways to the status of Graph Isomorphism. There is a natural reduction from Graph Isomorphism to the HSP on the symmetric group, but a series of results (e.g., Hallgren et al. [12], Moore et al. [23]) have shown that the resulting instances of the HSP require highly-entangled measurements, and that known families of such measurements cannot succeed. Thus the miracle of Shor's algorithms for factoring and discrete log, where we can solve these problems simply by looking at the symmetries of a certain function, does not seem to apply to Graph Isomorphism. Any efficient quantum algorithm for it would have to involve significantly new ideas.

On the other hand, many cases of Graph Isomorphism are easy classically, including graphs with bounded eigenvalue multiplicity [1] and constant degree [18]. Many of these classical algorithms work by finding a *canonical labeling* of the graph [2, 3], giving each vertex a unique label based on local quantities. These labeling schemes use the details of the graph, and not just its symmetries—precisely what the reduction to the HSP leaves out. Analogously, the support splitting algorithm labels each coordinate of the code by the weight enumerator of the hull of the code punctured at that coordinate. For most codes, including Goppa codes, this creates a labeling that is unique or nearly unique, allowing us to determine the permutation P.

There are families of instances of Graph Isomorphism that defeat known methods, due to the fact that no local or spectral property appears to distinguish the vertices from each other. In particular, no polynomial-time algorithm is known for isomorphism of strongly regular graphs. (On the other hand, these graphs are highly structured, yielding canonical-labeling algorithms that, while still exponential, are faster than those known for general graphs [34].) In the same vein, we might hope that there are families of codes where the coordinates are hard to distinguish from each other using linear-algebraic properties. In that case, the corresponding McEliece cryptosystem might be hard classically even for known-code adversaries, and the reduction to the HSP would be relevant to their post-quantum security.

Along these lines, Sidelnikov [31] proposed a variant of the McEliece cryptosystem using binary

Reed-Muller codes. Since there is a single Reed-Muller code of given rate and block length, the code M is known to the adversary and the security of the system is directly related to the Code Equivalence problem. Additionally, since Reed-Muller codes are self-dual, they coincide with their hulls so that the weight enumerators used by the SSA are exponentially large, making them a hard case for that classical algorithm.

As mentioned above, our results apply directly to Reed-Muller codes, and thus frustrate the natural quantum Fourier sampling approach to the corresponding instances of Code Equivalence. As virtually all known exponential speed-ups of quantum algorithms for algebraic problems derive from Fourier sampling, this suggests that new ideas would be necessary to exploit quantum computing for breaking the Sidelnikov system. Note, however, that this result does not rule out classical attacks on the Sidelnikov system. In particular, a classical algorithm of Minder and Shokrollahi [21] solves the Code Equivalence problem for binary Reed-Muller codes in quasipolynomial time, at least in the low-rate setting where Reed-Muller codes have the best performance, yielding a direct attack on the Sidelnikov system.

Recently, Sendrier and Simos [29] considered a general version of Code Equivalence, called Linear Code Equivalence, which is to decide whether two linear codes are identical up to a linear isometry of the Hamming distance. This problem is related to the security of code-based cryptosystems in a general form. They showed then that Linear Code Equivalence can also be reduced to (Permutation) Code Equivalence, but the corresponding instance of Code Equivalence is a hard case for the classical algorithm SSA, at least for q-ary codes with $q \geq 5$. Additionally, using the results of our preliminary work, Sendrier and Simos [29] pointed out that the instance of Code Equivalence reduced from Linear Code Equivalence is also HSP-hard. Based on this hardness of Code Equivalence, they improved Girault's zero-knowledge protocol, which is also a candidate for post-quantum cryptography.

1.3 Summary of Technical Ideas

Let G be a finite group. We wish to establish general criteria for indistinguishability of subgroups $H < G$ by strong Fourier sampling. We begin with the general strategy, developed in [22], that controls the resulting probability distributions in terms of the representation-theoretic properties of G. In order to handle richer subgroups, however, we have to overcome some technical difficulties. Our principal contribution here is a "decoupling" lemma that allows us to handle the cross terms arising from pairs of nontrivial group elements.

Roughly, the approach (presented in Section 3.2) identifies two disjoint subsets, Small and Large, of irreps of G. The set Large consists of all irreps whose dimensions are no smaller than a certain threshold D. While D should be as large as possible, we also need to choose D small enough so that the set Large is large. In contrast, the representations in Small must have small dimension (much smaller than $\sqrt{D}$), and the set Small should be small or contain few irreps that appear in the decomposition of the tensor product representation $\rho \otimes \rho^*$ for any $\rho \in Large$. In addition, any irrep ρ outside $Small$ must have small normalized character $|\chi_\rho(h)|/d_\rho$ for any nontrivial element $h \in H$. If two such sets exist, and if $|H|$ is sufficiently small, we establish that H is indistinguishable by strong Fourier sampling over G.

In the case $G = S_n$, as in [22] we define Small as the set Λ_c of all Young diagrams whose top row or left column has length at least $(1-c)n$, and define Large by setting $D = n^{dn}$, for appropriate constants $0 < c, d < 1$. We show that any irrep outside Small has large dimension and therefore small normalized characters.

In the case $G = \mathsf{GL}_2(\mathbb{F}_q)$, we choose Small as the set of all linear representations and set the threshold $D = q-1$. The key lemma we need to prove is then that for any nonlinear irrep ρ of $\mathsf{GL}_2(\mathbb{F}_q)$, the decomposition of $\rho \otimes \rho^*$ contains at most two inequivalent linear representations. (Lemma 8).

For the case $G = (\mathsf{GL}_k(\mathbb{F}_q) \times S_n) \wr \mathbb{Z}_2$ reduced from Code Equivalence, the normalized characters on the hidden subgroup K depend on the minimal degree of the automorphism group $\mathrm{Aut}(M) < S_n$. If we choose Small as the set of all irreps constructed from tensor product representations $\tau \times \lambda$ of $\mathsf{GL}_k(\mathbb{F}_q) \times S_n$ with $\lambda \in \Lambda_c$, then the "small" features of Λ_c will induce the "small" features of this set Small. Finally, $|K|$ depends on $|\mathrm{Aut}(M)|$ and the column rank of M. When M is a generator matrix of a rational Goppa code or a canonical parity check matrix of an alternant code, $\mathrm{Aut}(M)$ lies inside the automorphism group of a rational Goppa code, which can be controlled using Stichtenoth's Theorem [35].

2 Standard Reduction from Code Equivalence to HSP

As mentioned in the introduction, we consider a search version of Code Equivalence that recovers a "scrambler" S and permutation P from matrices M and M'. We generalize this search version into the following problem:

Definition 1 (Scrambler-Permutation Problem)**.** *Given two $k \times n$ matrices M and M' with entries in a finite field containing $\mathbb{F}_q$ such that $M' = SMP$ for some $S \in \mathsf{GL}_k(\mathbb{F}_q)$ and some $n \times n$ permutation matrix P, find such a pair (S,P).*

This problem can be immediately recast as a Hidden Subgroup Problem (described below). We begin with a presentation of the problem as a Hidden Shift Problem:

Definition 2 (Hidden Shift Problem)**.** *Let G be a finite group and Σ be a finite set. Given two functions $f_0 : G \to \Sigma$ and $f_1 : G \to \Sigma$ with the promise that there is an element $s \in G$ for which $f_1(x) = f_0(sx)$ for all $x \in G$, the problem is to determine such s by making queries to f_0 and f_1. An element s with this property is called a* left shift *from f_0 to f_1 (or, simply, a* shift*).*

The Scrambler-Permutation Problem can be immediately reduced to the Hidden Shift Problem over the group $G = \mathsf{GL}_k(\mathbb{F}_q) \times S_n$ by defining functions f_0 and f_1 on $\mathsf{GL}_k(\mathbb{F}_q) \times S_n$ so that for all $(S,P) \in \mathsf{GL}_k(\mathbb{F}_q) \times S_n$,

$$f_0(S,P) = S^{-1}MP\,, \qquad f_1(S,P) = S^{-1}M'P\,. \tag{1}$$

Here and from now on, we identify each $n \times n$ permutation matrix with its corresponding permutation in S_n. Evidently, $SMP = M'$ if and only if (S^{-1},P) is a shift from f_0 to f_1.

Next, following the standard approach to developing quantum algorithms for such problems, we reduce this Hidden Shift Problem on a group G to the Hidden Subgroup Problem on the wreath product $G \wr \mathbb{Z}_2 = G^2 \rtimes \mathbb{Z}_2$. Given two functions f_0 and f_1 on G, we define the function $f : G \wr \mathbb{Z}_2 \to \Sigma \times \Sigma$ as follows: for $(x,y) \in G^2$ and $b \in \mathbb{Z}_2$,

$$f((x,y),b) \overset{\text{def}}{=} \begin{cases} (f_0(x),\, f_1(y)) & \text{if } b = 0 \\ (f_1(y),\, f_0(x)) & \text{if } b = 1 \end{cases} \tag{2}$$

Now we would like to see that the Hidden Shift Problem is equivalent to determining the subgroup whose cosets are distinguished by f. Recall that a function f on a group G *distinguishes the right* *cosets* of a subgroup $H < G$ if for all $x, y \in G$, $f(x) = f(y) \iff yx^{-1} \in H$.

Definition 3. *Let f be a function on a group G. We say that f is* injective under right multiplication *if for all $x, y \in G$, $f(x) = f(y) \iff f(yx^{-1}) = f(1)$. Define the subset $G|_f \subseteq G$ as the level set containing the identity,*

$$G|_f \overset{\text{def}}{=} \{g \in G \mid f(g) = f(1)\} .$$

Proposition 1. *Let f be a function on a group G. If f distinguishes the right cosets of a subgroup $H < G$, then f must be injective under right multiplication and $G|_f = H$. Conversely, if f is injective under right multiplication, then $G|_f$ is a subgroup and f distinguishes the right cosets of the subgroup $G|_f$.*

Hence, the function f defined in (2) can distinguish the right cosets of some subgroup if and only if it is injective under right multiplication.

Lemma 1. *The function f defined in* (2) *is injective under right multiplication if and only if (1) f_0 is injective under right multiplication and (2) $f_1(x) = f_0(sx)$ for some s.*

The proof of this lemma is straightforward, so we omit it here.

Proposition 2. *Assume f_0 is injective under right multiplication. Let $H_0 = G|_{f_0}$ and s be a shift. Then the function f defined in* (2) *distinguishes right cosets of the following subgroup of $G \wr \mathbb{Z}_2$:*

$$G \wr \mathbb{Z}_2|_f = \big((H_0, s^{-1}H_0 s), 0\big) \cup \big((H_0 s, s^{-1}H_0), 1\big) ,$$

which has size $2|H_0|^2$. The set of all shifts from f_0 to f_1 is $H_0 s$.

If we can determine the hidden subgroup $K = G \wr \mathbb{Z}_2|_f$, we can find a shift by selecting an element of the form $((g_1, g_2), 1)$ from K. Then g_1 must belong to $H_0 s$, and so is a shift from f_0 to f_1.

Application to the Scrambler-Permutation problem. Returning to the Hidden Shift Problem over $G = \mathsf{GL}_k(\mathbb{F}_q) \times S_n$ corresponding to the Scrambler-Permutation problem, it is clear that the function f_0 defined in (1) is injective under right multiplication, and that

$$H_0 = \mathsf{GL}_k(\mathbb{F}_q) \times S_n|_{f_0} = \left\{(S, P) \in \mathsf{GL}_k(\mathbb{F}_q) \times S_n \mid S^{-1}MP = M\right\} .$$

The automorphism group of M is the projection of H_0 onto S_n, i.e.,

$$\mathrm{Aut}(M) = \left\{P \in S_n \mid \exists S : S^{-1}MP = M\right\} .$$

Note that each $P \in \mathrm{Aut}(M)$ has the same number of preimages $S \in \mathsf{GL}_k(\mathbb{F}_q)$ in this projection.

3 Quantum Fourier Sampling (QFS)

3.1 Preliminaries and Notation

Fix a finite group G, abelian or non-abelian, and let $\widehat{G}$ denote the set of irreducible unitary representations, or "irreps" for short, of G. For each irrep $\rho \in \widehat{G}$, let V_ρ denote a vector space over $\mathbb{C}$ on which ρ

acts so that ρ is a group homomorphism from G to the general linear group over V_ρ, and let d_ρ denote the dimension of V_ρ. For each ρ, we fix an orthonormal basis $B_\rho = \left\{\mathbf{b}_1, \ldots, \mathbf{b}_{d_\rho}\right\}$ for V_ρ. Then we can represent each $\rho(g)$ as a $d_\rho \times d_\rho$ unitary matrix whose j^{th} column is the vector $\rho(g)\mathbf{b}_j$.

Viewing the vector space $\mathbb{C}[G]$ as the regular representation of G, we can decompose $\mathbb{C}[G]$ into irreps as the direct sum $\bigoplus_{\rho \in \widehat{G}} V_\rho^{\oplus d_\rho}$. This has a basis $\{|\rho,i,j\rangle : \rho \in \widehat{G}, 1 \le i,j \le d_\rho\}$, where $\{|\rho,i,j\rangle \mid 1 \le i \le d_\rho\}$ is a basis for the j^{th} copy of V_ρ. Up to normalization, $|\rho,i,j\rangle$ corresponds to the i,j entry of the irrep ρ.

Definition 4. *The* Quantum Fourier transform *over G is the unitary operator, denoted F_G, that transforms a vector in $\mathbb{C}[G]$ from the basis $\{|g\rangle \mid g \in G\}$ into the basis given by the decomposition of $\mathbb{C}[G]$. For all $g \in G$,*

$$F_G\,|g\rangle = \sum_{\rho,i,j} \sqrt{\frac{d_\rho}{|G|}}\,\rho(g)_{i,j}\,|\rho,i,j\rangle\ ,$$

where $\rho(g)_{ij}$ is the (i,j)-entry of the matrix $\rho(g)$. Alternatively, we can view $F_G\,|g\rangle$ as a block diagonal matrix consisting of the block $\sqrt{d_\rho/|G|}\,\rho(g)$ for each $\rho \in \widehat{G}$.

Notation. For each subset $X \subseteq G$, define $|X\rangle = (1/\sqrt{|X|})\sum_{x\in X}|x\rangle$, which is the uniform superposition over X. For each $X \subseteq G$ and $\rho \in \widehat{G}$, define the operator $\Pi_X^\rho \stackrel{\text{def}}{=} \frac{1}{|X|}\sum_{x\in X}\rho(x)$, and let $\widehat{X}(\rho)$ denote the $d_\rho \times d_\rho$ matrix block at ρ in the quantum Fourier transform of $|X\rangle$, i.e.,

$$\widehat{X}(\rho) \stackrel{\text{def}}{=} \sqrt{\frac{d_\rho}{|G||X|}}\sum_{x\in X}\rho(x) = \sqrt{\frac{d_\rho|X|}{|G|}}\,\Pi_X^\rho\ .$$

Fact 1. *If X is a subgroup of G, then Π_X^ρ is a projection operator [22].*

Quantum Fourier Sampling (QFS) is a standard procedure based on the Quantum Fourier Transform to solve the Hidden Subgroup Problem (HSP) (see [17] for a survey). An instance of the HSP over G consists of a black-box function $f : G \to \{0,1\}^*$ such that $f(x) = f(y)$ if and only if x and y belong to the same left coset of H in G, for some subgroup $H \le G$. The problem is to recover H using the oracle $O_f : |x,y\rangle \mapsto |x, y \oplus f(x)\rangle$. The general QFS procedure for this is the following:

1. Prepare a 2-register quantum state, the first in a uniform superposition of the group elements and the second with the value zero: $|\psi_1\rangle = (1/\sqrt{|G|})\sum_{g\in G}|g\rangle\,|0\rangle$.

2. Query f, i.e., apply the oracle O_f, resulting in the state

$$|\psi_2\rangle = O_f\,|\psi_1\rangle = \frac{1}{\sqrt{|G|}}\sum_{g\in G}|g\rangle\,|f(g)\rangle = \frac{1}{\sqrt{|T|}}\sum_{\alpha\in T}|\alpha H\rangle\,|f(\alpha)\rangle$$

   where T is a transversal of H in G.

3. Measure the second register of $|\psi_2\rangle$, resulting in the state $|\alpha H\rangle\,|f(\alpha)\rangle$ with probability $1/|T|$ for each $\alpha \in T$. The first register of the resulting state is then $|\alpha H\rangle$ for some uniformly random $\alpha \in G$.

4. Apply the quantum Fourier transform over G to the coset state $|\alpha H\rangle$ observed at step 3:
$$F_G\,|\alpha H\rangle = \sum_{\rho\in\widehat{G}, 1\le i,j\le d_\rho} \widehat{\alpha H}(\rho)_{i,j}\,|\rho,i,j\rangle\ .$$
5. (Weak) Observe the representation name ρ. (Strong) Observe ρ and matrix indices i,j.
6. Classically process the information observed from the previous step to determine the subgroup H.

Probability distributions produced by QFS. For a particular coset αH, the probability of measuring the representation ρ in the state $F_G\,|\alpha H\rangle$ is

$$P_{\alpha H}(\rho) = \|\widehat{\alpha H}(\rho)\|_F^2 = \frac{d_\rho |H|}{|G|}\mathrm{Tr}\left((\Pi^\rho_{\alpha H})^\dagger \Pi^\rho_{\alpha H}\right) = \frac{d_\rho |H|}{|G|}\mathrm{Tr}\left(\Pi^\rho_H\right)$$

where $\mathrm{Tr}(A)$ denotes the trace of a matrix A, and $\|A\|_F := \sqrt{\mathrm{Tr}(A^\dagger A)}$ is the Frobenius norm of A. The last equality is due to the fact that $\Pi^\rho_{\alpha H} = \rho(\alpha)\Pi^\rho_H$ and that Π^ρ_H is an orthogonal projector.

Since there is no point in measuring the rows [10], we are only concerned with measuring the columns. As pointed out in [22], the optimal von Neumann measurement on a coset state can always be expressed in this form for some basis B_ρ. Conditioned on observing ρ in the state $F_G\,|\alpha H\rangle$, the probability of measuring a given $\mathbf{b}\in B_\rho$ is $\|\widehat{\alpha H}(\rho)\mathbf{b}\|^2$. Hence the conditional probability that we observe the vector $\mathbf{b}$, given that we observe the representation ρ, is then

$$P_{\alpha H}(\mathbf{b}\mid\rho) = \frac{\|\widehat{\alpha H}(\rho)\mathbf{b}\|^2}{P_{\alpha H}(\rho)} = \frac{\|\Pi^\rho_{\alpha H}\mathbf{b}\|^2}{\mathrm{Tr}\left(\Pi^\rho_H\right)} = \frac{\|\Pi^\rho_H\mathbf{b}\|^2}{\mathrm{Tr}\left(\Pi^\rho_H\right)}$$

where in the last equality, we use the fact that as $\rho(\alpha)$ is unitary, it preserves the norm of the vector $\Pi^\rho_H\mathbf{b}$.

The coset representative α is unknown and is uniformly distributed in T. However, both distributions $P_{\alpha H}(\rho)$ and $P_{\alpha H}(\mathbf{b}\mid\rho)$ are independent of α and are the same as those for the state $F_G\,|H\rangle$. Thus, in Step 5 of the QFS procedure above, we observe $\rho\in\widehat{G}$ with probability $P_H(\rho)$, and conditioned on this event, we observe $\mathbf{b}\in B_\rho$ with probability $P_H(\mathbf{b}\mid\rho)$.

If the hidden subgroup is trivial, $H=\{1\}$, the conditional probability distribution on B_ρ is uniform,

$$P_{\{1\}}(\mathbf{b}\mid\rho) = \frac{\|\Pi^\rho_{\{1\}}\mathbf{b}\|^2}{\mathrm{Tr}\left(\Pi^\rho_{\{1\}}\right)} = \frac{\|\mathbf{b}\|^2}{d_\rho} = \frac{1}{d_\rho}\,.$$

3.2 Distinguishability by QFS

We fix a finite group G and consider quantum Fourier sampling over G in the basis given by $\{B_\rho\}$. For a subgroup $H<G$ and for $g\in G$, let H^g denote the conjugate subgroup $g^{-1}Hg$. Since $\mathrm{Tr}\left(\Pi^\rho_H\right) = \mathrm{Tr}\left(\Pi^\rho_{H^g}\right)$, the probability distributions obtained by QFS for recovering the hidden subgroup H^g are

$$P_{H^g}(\rho) = \frac{d_\rho|H|}{|G|}\mathrm{Tr}\left(\Pi^\rho_H\right) = P_H(\rho) \quad\text{and}\quad P_{H^g}(\mathbf{b}\mid\rho) = \frac{\|\Pi^\rho_{H^g}\mathbf{b}\|^2}{\mathrm{Tr}\left(\Pi^\rho_H\right)}\,.$$

As $P_{H^g}(\rho)$ does not depend on g, weak Fourier sampling can not distinguish conjugate subgroups. Our goal is to point out that for certain nontrivial subgroup $H<G$, strong Fourier sampling can not

efficiently distinguish the conjugates of H from each other or from the trivial one. Recall that the distribution $P_{\{1\}}(\cdot \mid \rho)$ obtained by performing strong Fourier sampling on the trivial hidden subgroup is the same as the uniform distribution U_{B_ρ} on the basis B_ρ. Thus, our goal can be boiled down to showing that the probability distribution $P_{H^g}(\cdot \mid \rho)$ is likely to be close to the uniform distribution U_{B_ρ} in total variation, for a random $g \in G$ and an irrep $\rho \in \widehat{G}$ obtained by weak Fourier sampling.

Definition 5. *We define the* distinguishability *of a subgroup H (using strong Fourier sampling over G), denoted $\mathscr{D}_H$, to be the expectation of the squared L_1-distance between $P_{H^g}(\cdot \mid \rho)$ and U_{B_ρ}:*

$$\mathscr{D}_H \overset{\text{def}}{=} \mathbb{E}_{\rho,g}\left[\|P_{H^g}(\cdot \mid \rho) - U_{B_\rho}\|_1^2\right],$$

where ρ is drawn from $\widehat{G}$ according to the distribution $P_H(\rho)$, and g is chosen from G uniformly at random. We say that the subgroup H is indistinguishable *if $\mathscr{D}_H \le \log^{-\omega(1)}|G|$.*

Note that if $\mathscr{D}_H$ is small, then the total variation distance between $P_{H^g}(\cdot \mid \rho)$ and U_{B_ρ} is small with high probability due to Markov's inequality: for all $\varepsilon > 0$,

$$\Pr_g\left[\|P_{H^g}(\cdot \mid \rho) - U_{B_\rho}\|_{t.v.} \ge \varepsilon/2\right] = \Pr_g\left[\|P_{H^g}(\cdot \mid \rho) - U_{B_\rho}\|_1^2 \ge \varepsilon^2\right] \le \mathscr{D}_H/\varepsilon^2 .$$

In particular, if the subgroup H is indistinguishable by strong Fourier sampling, then for all constant $c > 0$,

$$\|P_{H^g}(\cdot \mid \rho) - U_{B_\rho}\|_{t.v.} < \log^{-c}|G|$$

with probability at least $1 - \log^{-c}|G|$ in both g and ρ. Our notion of indistinguishability is the direct analogue of that of Kempe and Shalev [15]. Focusing on weak Fourier sampling, they say that H is indistinguishable if $\|P_H(\cdot) - P_{\{1\}}(\cdot)\|_{t.v.} < \log^{-\omega(1)}|G|$.

Our main theorem below will serve as a general guideline for bounding the distinguishability of H. For this purpose we define, for each $\sigma \in \widehat{G}$, the *maximal normalized character of σ on H* as

$$\overline{\chi}_\sigma(H) \overset{\text{def}}{=} \max_{h \in H\setminus\{1\}} \frac{|\chi_\sigma(h)|}{d_\sigma} .$$

For each subset $S \subset \widehat{G}$, let

$$\overline{\chi}_{\overline{S}}(H) = \max_{\sigma \in \widehat{G}\setminus S} \overline{\chi}_\sigma(H) \quad \text{and} \quad d_S = \max_{\sigma \in S} d_\sigma .$$

In addition, for each reducible representation ρ of G, we let $I(\rho)$ denote the set of irreps of G that appear in the decomposition of ρ into irreps.

Theorem 1. *(Main Theorem) Suppose S is a subset of $\widehat{G}$. Let $D > d_S^2$ and $L = L_D \subset \widehat{G}$ be the set of all irreps of dimension at least D. Let*

$$\Delta = \Delta_{S,L} = \max_{\rho \in L}\left|S \cap I(\rho \otimes \rho^*)\right| . \tag{3}$$

Then the distinguishability of H is bounded by

$$\mathscr{D}_H \le 4|H|^2\left(\overline{\chi}_{\overline{S}}(H) + \Delta\frac{d_S^2}{D} + \frac{|\overline{L}|D^2}{|G|}\right) .$$

Intuitively, the set S consists of irreps of small dimension, and L consists of irreps of large dimension. Moreover, we wish to have that the size of S is small while the size of L is large, so that most irreps are likely in L. In the cases where there are relatively few irreps, i.e., $|S| \ll D$ and $|\widehat{G}| \ll |G|$, we can simply upper bound Δ by $|S|$ and upper bound $|\overline{L}|$ by $|\widehat{G}|$.

We discuss the proof of this theorem in Section 4. Applications of Theorem 1 will be presented in Section 5 (for the symmetric group), in Section 7 (for the general linear group $\mathsf{GL}_2(\mathbb{F}_q)$), and in Section 6 (for Code Equivalence).

4 Bounding Distinguishability

We now sketch the proof for the main theorem (Theorem 1). Fixing a nontrivial subgroup $H < G$, we want to upper bound $\mathscr{D}_H$. Let us start with bounding the expectation over the random group element $g \in G$, for a fixed irrep $\rho \in \widehat{G}$:

$$E_H(\rho) \stackrel{\text{def}}{=} \mathbb{E}_g\left[\|P_{H^g}(\cdot \mid \rho) - U_{B_\rho}\|_1^2\right] .$$

Obviously we always have $E_H(\rho) \leq 4$. More interestingly, we have

$$\begin{aligned} E_H(\rho) &= \mathbb{E}_g\left[\left(\sum_{\mathbf{b}\in B_\rho}\left|P_{H^g}(\mathbf{b} \mid \rho) - \frac{1}{d_\rho}\right|\right)^2\right] \\ &\leq \mathbb{E}_g\left[d_\rho \sum_{\mathbf{b}\in B_\rho}\left(P_{H^g}(\mathbf{b} \mid \rho) - \frac{1}{d_\rho}\right)^2\right] \qquad \text{(by Cauchy-Schwarz)} \\ &= d_\rho \sum_{\mathbf{b}\in B_\rho} \mathrm{Var}_g[P_{H^g}(\mathbf{b} \mid \rho)] \qquad (\text{since } \mathbb{E}_g[P_{H^g}(\mathbf{b} \mid \rho)] = \frac{1}{d_\rho}) \\ &= \frac{d_\rho}{\mathrm{Tr}(\Pi_H^\rho)^2} \sum_{\mathbf{b}\in B_\rho} \mathrm{Var}_g\left[\|\Pi_{H^g}^\rho \mathbf{b}\|^2\right] . \end{aligned} \tag{4}$$

The equation $\mathbb{E}_g[P_{H^g}(\mathbf{b} \mid \rho)] = 1/d_\rho$ can be shown using *Schur's lemma* as in Proposition 3 below.

Proposition 3. *Let $H < G$ and g be chosen from G uniformly at random. Then for $\rho \in \widehat{G}$ and $\mathbf{b} \in B_\rho$,*

$$\mathbb{E}_g[P_{H^g}(\mathbf{b} \mid \rho)] = 1/d_\rho .$$

Proof: Schur's lemma asserts that if ρ is irreducible, the only matrices which commute with $\rho(g)$ for all g are the scalars. Hence,

$$\mathbb{E}_g\left[\Pi_{H^g}^\rho\right] = \frac{1}{|G|}\sum_{g\in G} \rho^\dagger(g)\Pi_H^\rho \rho(g) = \frac{\mathrm{Tr}(\Pi_H^\rho)}{d_\rho}\mathbf{1}_{d_\rho} ,$$

which implies that

$$\mathbb{E}_g\left[\|\Pi_{H^g}^\rho \mathbf{b}\|^2\right] = \mathbb{E}_g\left[\langle \mathbf{b}, \Pi_{H^g}^\rho \mathbf{b}\rangle\right] = \left\langle \mathbf{b}, \mathbb{E}_g\left[\Pi_{H^g}^\rho\right]\mathbf{b}\right\rangle = \frac{\mathrm{Tr}(\Pi_H^\rho)}{d_\rho} .$$

□

From (4), we are motivated to bound the variance of $\|\Pi_{H^g}^{\rho}\mathbf{b}\|^2$ when g is chosen uniformly at random. We provide an upper bound that depends on the projection of the vector $\mathbf{b}\otimes\mathbf{b}^*$ onto irreducible subspaces of $\rho\otimes\rho^*$, and on maximal normalized characters of σ on H for all irreps σ appearing in the decomposition of $\rho\otimes\rho^*$. Recall that the representation $\rho\otimes\rho^*$ is typically reducible and can be written as an orthogonal direct sum of irreps $\rho\otimes\rho^* = \bigoplus_{\sigma\in\widehat{G}} a_\sigma\sigma$, where $a_\sigma \geq 0$ is the multiplicity of σ. Then $I(\rho\otimes\rho^*)$ consists of σ with $a_\sigma > 0$, and we let $\Pi_\sigma^{\rho\otimes\rho^*}$ denote the projection operator whose image is $a_\sigma\sigma$, that is, the subspace spanned by all copies of σ. Our upper bound given in Lemma 2 below generalizes the bound given in Lemma 4.3 of [22], which only applies to subgroups H of order 2.

Lemma 2. *(Decoupling Lemma) Let ρ be an irrep of G. Then for any vector $\mathbf{b}\in V_\rho$,*

$$\mathrm{Var}_g\left[\|\Pi_{H^g}^{\rho}\mathbf{b}\|^2\right] \leq \sum_{\sigma\in I(\rho\otimes\rho^*)} \overline{\chi}_\sigma(H)\left\|\Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2 .$$

Proof of Lemma 2 Fix a vector $\mathbf{b}\in V_\rho$. To simplify notations, we shall write Π_g as shorthand for $\Pi_{H^g}^{\rho}$, and write $g\mathbf{b}$ for $\rho(g)\mathbf{b}$. For any $g\in G$, we have

$$\begin{aligned}\|\Pi_g\mathbf{b}\|^2 &= \left\langle \Pi_g\mathbf{b}, \Pi_g\mathbf{b}\right\rangle = \left\langle \mathbf{b}, \Pi_g\mathbf{b}\right\rangle \\ &= \frac{1}{|H|}\left(\langle \mathbf{b},\mathbf{b}\rangle + \sum_{h\in H\setminus\{1\}}\left\langle \mathbf{b}, g^{-1}hg\mathbf{b}\right\rangle\right) .\end{aligned}$$

Let $S_g = \sum_{h\in H\setminus\{1\}}\left\langle \mathbf{b}, g^{-1}hg\mathbf{b}\right\rangle$. Then

$$\mathrm{Var}_g\left[\|\Pi_g\mathbf{b}\|^2\right] = \frac{\mathrm{Var}_g[S_g]}{|H|^2} = \frac{\mathbb{E}_g\left[S_g^2\right] - \mathbb{E}_g[S_g]^2}{|H|^2} .$$

To bound the variance, we upper bound S_g^2 for all $g\in G$. Since S_g is real, applying Cauchy-Schwarz inequality, we have

$$S_g^2 = \left|\sum_{h\in H\setminus\{1\}}\left\langle \mathbf{b}, g^{-1}hg\mathbf{b}\right\rangle\right|^2 \leq (|H|-1)\left(\sum_{h\in H\setminus\{1\}}\left|\left\langle \mathbf{b}, g^{-1}hg\mathbf{b}\right\rangle\right|^2\right) .$$

As in Lemma 4.2 of [22], one can express the second moment of the inner product $\left\langle \mathbf{b}, g^{-1}hg\mathbf{b}\right\rangle$ in terms of the projection of $\mathbf{b}\otimes\mathbf{b}^*$ into the irreducible constituents of the tensor product representation $\rho\otimes\rho^*$. Specifically, for any $h\in G$, we have

$$\mathbb{E}_g\left[\left|\left\langle \mathbf{b}, g^{-1}hg\mathbf{b}\right\rangle\right|^2\right] = \sum_{\sigma\in I(\rho\otimes\rho^*)}\frac{\chi_\sigma(h)}{d_\sigma}\left\|\Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2 .$$

It follows that

$$\mathrm{Var}_g\left[\|\Pi_{H^g}^{\rho}\mathbf{b}\|^2\right] \leq \frac{|H|-1}{|H|^2}\sum_{h\in H\setminus\{1\}}\mathbb{E}_g\left[\left|\left\langle \mathbf{b}, g^{-1}hg\mathbf{b}\right\rangle\right|^2\right]$$

$$\leq \mathbb{E}_{h \in H \setminus \{1\}} \left[\sum_{\sigma \in I(\rho \otimes \rho^*)} \frac{\chi_\sigma(h)}{d_\sigma} \left\| \Pi_\sigma^{\rho \otimes \rho^*}(\mathbf{b} \otimes \mathbf{b}^*) \right\|^2 \right]$$
$$\leq \sum_{\sigma \in I(\rho \otimes \rho^*)} \overline{\chi}_\sigma(H) \left\| \Pi_\sigma^{\rho \otimes \rho^*}(\mathbf{b} \otimes \mathbf{b}^*) \right\|^2 .$$

□

Returning to our goal of bounding $E_H(\rho)$ using the bound in Lemma 2, the strategy will be to separate irreps appearing in the decomposition of $\rho \otimes \rho^*$ into two groups, those with small dimension and those with large dimension, and treat them differently. If d_σ is large, we shall rely on bounding $\overline{\chi}_\sigma(H)$. If d_σ is small, we shall control the projection given by $\Pi_\sigma^{\rho \otimes \rho^*}$ using the following lemma which was proved implicitly in [22]:

Lemma 3. *For any irrep σ, we have $\sum_{\mathbf{b} \in B_\rho} \left\| \Pi_\sigma^{\rho \otimes \rho^*}(\mathbf{b} \otimes \mathbf{b}^*) \right\|^2 \leq d_\sigma^2$.*

Proof of Lemma 3 Let L_σ be the subspace of $\rho \otimes \rho^*$ consisting of all copies of σ. Since B_ρ is orthonormal, the vectors $\left\{ \mathbf{b} \otimes \mathbf{b}^* \mid \mathbf{b} \in B_\rho \right\}$ are mutually orthogonal in $\rho \otimes \rho^*$. Thus,

$$\sum_{\mathbf{b} \in B_\rho} \left\| \Pi_\sigma^{\rho \otimes \rho^*}(\mathbf{b} \otimes \mathbf{b}^*) \right\|^2 \leq \dim L_\sigma .$$

Note that $\dim L_\sigma$ is equal to d_σ times the multiplicity of σ in $\rho \otimes \rho^*$. On the other hand, we have

$$\begin{aligned} \text{multiplicity of } \sigma \text{ in } \rho \otimes \rho^* &= \langle \chi_\sigma, \chi_\rho \chi_{\rho^*} \rangle = \langle \chi_\sigma \chi_\rho, \chi_{\rho^*} \rangle \\ &= \text{multiplicity of } \rho^* \text{ in } \sigma \otimes \rho \\ &\leq \frac{\dim(\sigma \otimes \rho)}{\dim \rho^*} = d_\sigma , \end{aligned}$$

Hence,

$$\sum_{\mathbf{b} \in B_\rho} \left\| \Pi_\sigma^{\rho \otimes \rho^*}(\mathbf{b} \otimes \mathbf{b}^*) \right\|^2 \leq d_\sigma^2 .$$

□

The method discussed above for bounding $E_H(\rho)$ is culminated into Lemma 4 below.

Lemma 4. *Let $\rho \in \widehat{G}$ be arbitrary and $S \subset \widehat{G}$ be any subset of irreps that does not contain ρ. Then*

$$E_H(\rho) \leq 4|H|^2 \left(\overline{\chi}_{\overline{S}}(H) + |S \cap I(\rho \otimes \rho^*)| \frac{d_S^2}{d_\rho} \right) .$$

Proof of Lemma 4 Combining Inequality (4) and Lemmas 2 give

$$E_H(\rho) \leq \frac{d_\rho}{\mathrm{Tr}(\Pi_H^\rho)^2} \sum_{\sigma \in I(\rho \otimes \rho^*)} \overline{\chi}_\sigma(H) \sum_{\mathbf{b} \in B_\rho} \left\| \Pi_\sigma^{\rho \otimes \rho^*}(\mathbf{b} \otimes \mathbf{b}^*) \right\|^2 .$$

Now we split additive items in the above upper bound into two groups separated by the set S. For the first group (large dimension),

$$\begin{aligned}\sum_{\sigma\in\overline{S}\cap\widehat{G}^{\rho\otimes\rho^*}}\overline{\chi}_\sigma(H)\sum_{\mathbf{b}\in B_\rho}\left\|\Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2 &\leq \overline{\chi}_{\overline{S}}(H)\sum_{\mathbf{b}\in B_\rho}\underbrace{\sum_{\sigma\in I(\rho\otimes\rho^*)}\left\|\Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2}_{\leq 1}\\ &\leq \overline{\chi}_{\overline{S}}(H)d_\rho\,.\end{aligned}$$

For the second group (small dimension),

$$\begin{aligned}\sum_{\sigma\in S\cap I(\rho\otimes\rho^*)}\overline{\chi}_\sigma(H)\sum_{\mathbf{b}\in B_\rho}\left\|\Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2 &\leq \sum_{\sigma\in S\cap I(\rho\otimes\rho^*)}\overline{\chi}_\sigma(H)d_\sigma^2 && \text{(by Lemma 3)}\\ &\leq \sum_{\sigma\in S\cap I(\rho\otimes\rho^*)}d_\sigma^2 && (\text{since } \overline{\chi}_\sigma(H)\leq 1)\\ &\leq |S\cap I(\rho\otimes\rho^*)|\,d_S^2\,.\end{aligned}$$

Summing the last bounds for the two groups yields

$$E_H(\rho)\leq\left(\frac{d_\rho}{\mathrm{Tr}(\Pi_H^\rho)}\right)^2\left(\overline{\chi}_{\overline{S}}(H)+|S\cap I(\rho\otimes\rho^*)|\,\frac{d_S^2}{d_\rho}\right)\,.$$

On the other hand, since $E_H(\rho)\leq 4$, we can assume $H^2\overline{\chi}_{\overline{S}}(H)\leq 1$, and thus $\overline{\chi}_{\overline{S}}(H)\leq\frac{1}{|H|^2}\leq\frac{1}{2|H|}$. Hence, we have

$$\frac{\mathrm{Tr}(\Pi_H^\rho)}{d_\rho}=\frac{1}{|H|}\left(1+\sum_{h\in H\setminus\{1\}}\frac{\chi_\rho(h)}{d_\rho}\right)\geq\frac{1}{|H|}-\overline{\chi}_\rho(H)\geq\frac{1}{2|H|}\,,$$

where the last inequality is due to $\overline{\chi}_\rho(H)\leq\overline{\chi}_{\overline{S}}(H)\leq\frac{1}{2|H|}$. This completes the proof. □

To apply this lemma, we should choose the subset S such that $d_S^2\ll d_\rho$, that is, S should consist of small dimensional irreps. Then applying Lemma 4 for all irreps ρ of large dimension, we can prove our general main theorem straightforwardly.

4.1 Proof of the Main Theorem

We now present the formal proof of the main theorem.

Proof of Theorem 1: Suppose S is a subset of $\widehat{G}$. Let $D>d_S^2$, $L=L_D\subset\widehat{G}$ be the set of all irreps of dimension at least D, and $\Delta=\Delta_{S,L}=\max_{\rho\in L}\left|S\cap I(\rho\otimes\rho^*)\right|$. Our goal is to show that

$$\mathscr{D}_H\leq 4|H|^2\left(\overline{\chi}_{\overline{S}}(H)+\Delta\frac{d_S^2}{D}+\frac{|\overline{L}|D^2}{|G|}\right)\,. \tag{5}$$

For any $\rho\in L$, since $d_\rho\geq D>d_S^2$, we must have $\rho\notin S$. By Lemma 4,

$$E_H(\rho)\leq 4|H|^2\left(\overline{\chi}_{\overline{S}}(H)+\Delta\frac{d_S^2}{D}\right)\quad\text{for all } \rho\in L\,.$$

Combining this with the fact that $E_H(\rho)\leq 4$ for all $\rho\notin L$, we obtain

$$\mathscr{D}_H=\mathbb{E}_\rho[E_H(\rho)]\leq 4|H|^2\left(\overline{\chi}_{\overline{S}}(H)+\Delta\frac{d_S^2}{D}\right)+4\mathrm{Pr}_\rho[\rho\notin L]\,.$$

To complete the proof, it remains to bound $\Pr_\rho[\rho \notin L]$. Since $\mathrm{Tr}(\Pi_H^\rho) \leq d_\rho$, we have

$$P(\rho) = \frac{d_\rho |H|}{|G|} \mathrm{Tr}(\Pi_H^\rho) \leq \frac{d_\rho^2 |H|}{|G|} .$$

Since $d_\rho < D$ for all $\rho \in \widehat{G} \setminus L$, it follows that

$$\Pr_\rho[\rho \notin L] = \sum_{\rho \notin L} P(\rho) \leq \frac{|\overline{L}| D^2 |H|}{|G|} \leq \frac{|\overline{L}| D^2 |H|^2}{|G|} .$$

This completes the proof of (5). □

5 Strong Fourier Sampling over S_n

We focus now on applying the main theorem to the case where G is the symmetric group S_n. As mentioned in the introduction, this case is motivated by the work of Kempe and Shalev [15] and Kempe et al. [16], which established similar results regarding the general hidden subgroup problem over S_n, but with weak Fourier sampling. The basic techniques used in this case, which control the normalized characters and dimensions of representations of the symmetric group, will be adopted in the case of Code Equivalence.

Recall that each irrep of S_n is in one-to-one correspondence to an integer partition $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_t)$ of n often given by a *Young diagram* of t rows in which the i^{th} row contains λ_i boxes. The conjugate representation of λ is the irrep corresponding to the partition $\lambda' = (\lambda'_1, \lambda'_2, \ldots, \lambda'_{t'})$, obtained by flipping the Young diagram λ about the diagonal.

As in [22], we shall apply Roichman's upper bound [27] on normalized characters:

Theorem 2 (Roichman's Theorem [27])**.** *There exist constant $b > 0$ and $0 < q < 1$ so that for $n > 4$, for every $\pi \in S_n$, and for every irrep λ of S_n,*

$$\left| \frac{\chi_\lambda(\pi)}{d_\lambda} \right| \leq \left(\max\left(q, \frac{\lambda_1}{n}, \frac{\lambda'_1}{n} \right) \right)^{b \cdot \mathrm{supp}(\pi)}$$

where $\mathrm{supp}(\pi) = \#\{k \in [n] \mid \pi(k) \neq k\}$ is the support of π.

This bound works well for unbalanced Young diagrams. In particular, for a constant $0 < c < 1/4$, let Λ_c denote the collection of partitions λ of n with the property that either $\frac{\lambda_1}{n} \geq 1 - c$ or $\frac{\lambda'_1}{n} \geq 1 - c$, i.e., the Young diagram λ contains at least $(1-c)n$ rows or contains at least $(1-c)n$ columns. Then, Roichman's upper bound implies that for every $\pi \in S_n$ and $\lambda \notin \Lambda_c$, and a universal constant $\alpha > 0$,

$$\left| \frac{\chi_\lambda(\pi)}{d_\lambda} \right| \leq \mathrm{e}^{-\alpha \cdot \mathrm{supp}(\pi)} . \tag{6}$$

On the other hand, both $|\Lambda_c|$ and the maximal dimension of representations in Λ_c are small, as shown in the following Lemma of [22].

Lemma 5 (Lemma 6.2 in [22])**.** *Let $p(n)$ denote the number of integer partitions of n. Then $|\Lambda_c| \leq 2cn \cdot p(cn)$, and $d_\mu < n^{cn}$ for any $\mu \in \Lambda_c$.*

To give a more concrete bound for the size of Λ_c, we record the asymptotic formula for the partition function [8, pg. 45]: $p(n) \approx \mathrm{e}^{\pi\sqrt{2n/3}}/(4\sqrt{3}n) = \mathrm{e}^{O(\sqrt{n})}/n$ as $n \to \infty$.

Now we are ready to prove the main result of this section, an application of Theorem 1.

Theorem 3. *Let H be a nontrivial subgroup of S_n with minimal degree m, i.e.,*

$$m = \min_{\pi \in H\setminus\{1\}} \operatorname{supp}(\pi)\,.$$

Then for sufficiently large n, $\mathscr{D}_H \leq O(|H|^2 \mathrm{e}^{-\alpha m})$.

Proof: Let $2c < d < 1/2$ be constants. We will apply Theorem 1 by setting $S = \Lambda_c$ and $D = n^{dn}$. By Lemma 5, we have $d_S \leq n^{cn}$. Hence, the condition $2c < d$ guarantees that $D > d_S^2$. First, we need to bound the maximal normalized character $\overline{\chi}_{\overline{S}}(H)$. By (6), we have $\overline{\chi}_\mu(H) \leq \mathrm{e}^{-\alpha m}$ for all $\mu \in \widehat{S_n} \setminus S$. Hence, $\overline{\chi}_{\overline{S}}(H) \leq \mathrm{e}^{-\alpha m}$. To bound the second term in the upper bound of Theorem 1, as $\Delta \leq |S|$, it suffices to bound:

$$|S| \cdot \frac{d_S^2}{D} \leq 2cn \cdot p(cn) \cdot \frac{n^{2cn}}{n^{dn}} \leq \mathrm{e}^{O(\sqrt{n})} \cdot n^{(2c-d)n} \leq n^{-\gamma n}/2$$

for sufficiently large n, so long as $\gamma < d - 2c$. Now bounding the last term in the upper bound of Theorem 1: Since $|\overline{L_D}| \leq |\widehat{S_n}| = p(n)$ and $n! > n^n \mathrm{e}^{-n}$ by Stirling's approximation,

$$\frac{|\overline{L_D}|D^2}{|S_n|} \leq \frac{p(n)n^{2dn}}{n!} \leq \frac{\mathrm{e}^{O(\sqrt{n})}n^{2dn}}{n^n \mathrm{e}^{-n}} \leq \mathrm{e}^{O(n)} n^{(2d-1)n} \leq n^{-\gamma n}/2$$

for sufficiently large n, so long as $\gamma < 1 - 2d$. By Theorem 1, $\mathscr{D}_H \leq 4|H|^2(\mathrm{e}^{-\alpha m} + n^{-\gamma n})$.

□

Theorem 3 generalizes Moore, Russell, and Schulman's result [22] on strong Fourier sampling over S_n, which only applied in the case $|H| = 2$. To relate our result to the results of Kempe et al. [16], observe that since $\log|S_n| = \Theta(n\log n)$, the subgroup H is indistinguishable by strong Fourier sampling if $|H|^2\mathrm{e}^{-\alpha m} \leq (n\log n)^{-\omega(1)}$ or, equivalently, if $m \geq (2/\alpha)\log|H| + \omega(\log n)$.

6 Applications to Code Equivalence

Our main application of Theorem 1 is to show the limitations of strong Fourier sampling in solving Code Equivalence. Throughout this section, we fix a $k \times n$ matrix M, which can be a generator matrix or a parity check matrix of a q-ary linear code used in some code-based cryptosystem. Note that the entries of M are in a finite field $\mathbb{F}_{q^\ell} \supset \mathbb{F}_q$ (when M is a generator matrix of a q-ary linear code, we must have $\ell = 1$).

Recall that the Code Equivalence instance with input M can be reduced to the HSP over the wreath product group $(\mathsf{GL}_k(\mathbb{F}_q) \times S_n) \wr \mathbb{Z}_2$; the hidden subgroup in this case is

$$K = ((H_0, s^{-1}H_0 s), 0) \cup ((H_0 s, s^{-1}H_0), 1) \tag{7}$$

for some hidden element $s \in \mathsf{GL}_k(\mathbb{F}_q) \times S_n$. Here, H_0 is a subgroup of $\mathsf{GL}_k(\mathbb{F}_q) \times S_n$ given by

$$H_0 = H_0(M) \overset{\text{def}}{=} \left\{(A,P) \in \mathsf{GL}_k(\mathbb{F}_q) \times S_n \mid A^{-1}MP = M\right\}. \tag{8}$$

To understand the structure of the subgroup $H_0(M)$, we define the *automorphism group* of M as

$$\mathrm{Aut}(M) \overset{\mathrm{def}}{=} \left\{P \in S_n \mid SMP = M \text{ for some } S \in \mathsf{GL}_k(\mathbb{F}_q)\right\} .$$

Note that $\mathrm{Aut}(M)$ is a subgroup of the symmetric group S_n and that each element $(A,P) \in H_0$ must have $P \in \mathrm{Aut}(M)$. This structure allows us to control the maximal normalized characters on K through the minimal degree of $\mathrm{Aut}(M)$.

6.1 Controlling Normalized Characters

In this part, we will discuss how to control the maximal normalized characters on the subgroup K defined in (7), which is necessary in order to apply the main theorem.

6.1.1 Normalized Characters for $G \wr \mathbb{Z}_2$

Firstly, we consider the wreath product $G \wr \mathbb{Z}_2$, for a general group G, and a subgroup of the form

$$K = ((H_0, s^{-1}H_0 s), 0) \cup ((H_0 s, s^{-1}H_0), 1) < G \wr \mathbb{Z}_2$$

for some subgroup $H_0 < G$ and some element $s \in G$. The irreducible characters of $G \wr \mathbb{Z}_2$ can be naturally constructed as induced characters:

1. Each unordered pair of two non-isomorphic irreps $\sigma, \rho \in \widehat{G}$ gives rise to an irrep of $G \wr \mathbb{Z}_2$, denoted $\{\rho, \sigma\}$, with character given by:

$$\chi_{\{\rho,\sigma\}}((x,y),b) = \begin{cases} \chi_\rho(x)\chi_\sigma(y) + \chi_\rho(y)\chi_\sigma(x) & \text{if } b = 0 \\ 0 & \text{if } b = 1 \,. \end{cases}$$

   The dimension of representation $\{\rho, \sigma\}$ is equal to $\chi_{\{\rho,\sigma\}}((1,1),0) = 2d_\rho d_\sigma$.

2. Each irrep $\rho \in \widehat{G}$ gives rise to two irreps of $G \wr \mathbb{Z}_2$, denoted $\{\rho\}$ and $\{\rho\}'$, with characters given by:

$$\chi_{\{\rho\}}((x,y),b) = \begin{cases} \chi_\rho(x)\chi_\rho(y) & \text{if } b = 0 \\ \chi_\rho(xy) & \text{if } b = 1 \end{cases}$$

$$\chi_{\{\rho\}'}((x,y),b) = \begin{cases} \chi_\rho(x)\chi_\rho(y) & \text{if } b = 0 \\ -\chi_\rho(xy) & \text{if } b = 1 \,. \end{cases}$$

   Both representations $\{\rho\}$ and $\{\rho\}'$ have the same dimension equal d_ρ^2.

Clearly, the number of irreps of $G \wr \mathbb{Z}$ is equal to $|\widehat{G}|^2/2 + 3|\widehat{G}|/2$, which is no more than $|\widehat{G}|^2$ as long as G has at least three irreps. Now it is easy to determine the maximal normalized characters on subgroup K.

Proposition 4. *For non-isomorphic irreps $\rho, \sigma \in \widehat{G}$,*

$$\overline{\chi}_{\{\rho,\sigma\}}(K) \leq \overline{\chi}_\rho(H_0)\overline{\chi}_\sigma(H_0) \,.$$

For irrep $\rho \in \widehat{G}$,

$$\overline{\chi}_{\{\rho\}}(K) = \overline{\chi}_{\{\rho\}'}(K) = \max\left\{\overline{\chi}_\rho(H_0)^2, 1/d_\rho\right\} .$$

Hence, to bound the maximal normalized characters on K, we can turn to bounding the normalized characters on the subgroup H_0 and the dimension of an irrep of G.

6.1.2 Normalized Characters for $(\mathsf{GL}_k(\mathbb{F}_q) \times S_n) \wr \mathbb{Z}_2$

Recall that for the HSP case reduced from Code Equivalence, we have $G = \mathsf{GL}_k(\mathbb{F}_q) \times S_n$ and every element $(A,P) \in H_0$ has $P \in \mathrm{Aut}(M)$.

For $\tau \in \widehat{\mathsf{GL}_k(\mathbb{F}_q)}$ and $\lambda \in \widehat{S_n}$, let $\tau \times \lambda$ denote the tensor product as a representation of $\mathsf{GL}_k(\mathbb{F}_q) \times S_n$. Those tensor product representations $\tau \times \lambda$ are all irreps of $\mathsf{GL}_k(\mathbb{F}_q) \times S_n$.

For all $(A,P) \in \mathsf{GL}_k(\mathbb{F}_q) \times S_n$, since $\overline{\chi}_{\tau\times\lambda}(A,P) = \overline{\chi}_\tau(A)\overline{\chi}_\lambda(P)$ and $\overline{\chi}_\tau(A) \leq 1$, we have $\overline{\chi}_{\tau\times\lambda}(A,P) \leq \overline{\chi}_\lambda(P)$. If $(A,P) \in H_0$ then $P \in \mathrm{Aut}(M)$. Hence,

$$\overline{\chi}_{\tau\times\lambda}(H_0) \leq \overline{\chi}_\lambda(\mathrm{Aut}(M))\,.$$

As in the treatment for the symmetric group, we can bound the maximal normalized character $\overline{\chi}_\lambda(\mathrm{Aut}(M))$ based on the minimum support of non-identity elements in $\mathrm{Aut}(M)$, for any $\lambda \in \widehat{S_n} \setminus \Lambda_c$.

To complete bounding the maximal normalized characters on the subgroup K, it remains to bound the dimension of a representation $\tau \times \lambda$ of the group $\mathsf{GL}_k(\mathbb{F}_q) \times S_n$ with $\lambda \in \widehat{S_n} \setminus \Lambda_c$. Since the dimension of $\tau \times \lambda$ is

$$d_{\tau\times\lambda} = d_\tau d_\lambda \geq d_\lambda\,,$$

we prove the following lower bound for d_λ.

Lemma 6. *Let $0 < c \leq 1/6$ be a constant. Then there is a constant $\beta > 0$ depending only on c such that for sufficiently large n and for $\lambda \in \widehat{S_n} \setminus \Lambda_c$,*

$$d_\lambda \geq \mathrm{e}^{\beta n}\,.$$

Proof of Lemma 6 Consider an integer partition of n, $\lambda = (\lambda_1, \ldots, \lambda_t)$, with both λ_1 and t less than $(1-c)n$. Let $\lambda' = (\lambda'_1, \ldots, \lambda'_{\lambda_1})$ be the conjugate of λ, where $t = \lambda'_1 \geq \lambda'_2 \geq \ldots \geq \lambda'_{\lambda_1}$ and $\sum_i \lambda'_i = n$. WLOG, assume $\lambda'_1 \leq \lambda_1$. We label all the cells of the Young diagram of shape λ as $c_1, \ldots, c_n$ such that c_i is the i^{th} cell from the left of the first row, for $1 \leq i \leq \lambda_1$ (see Figure 1 below for an example).

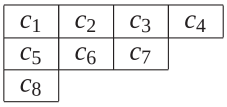

Fig. 1. Young diagram $\lambda = (4,3,1)$, where $\mathrm{hook}(c_1) = 6$, $\mathrm{hook}(c_2) = 4$, $\mathrm{hook}(c_3) = 3$, $\mathrm{hook}(c_4) = 1$.

The dimension of λ is determined by the *hook length formula* [8, p.50]:

$$d_\lambda = \frac{n!}{\mathrm{Hook}(\lambda)}\,, \qquad \mathrm{Hook}(\lambda) = \prod_{i=1}^{n} \mathrm{hook}(c_i)\,,$$

where $\mathrm{hook}(c_i)$ is the number of cells appearing in either the same column or the same row as the cell c_i, excluding those that are above or to the left of c_i. In particular,

$$\mathrm{hook}(c_i) = \lambda_1 - i + \lambda'_i \qquad \text{for } 1 \leq i \leq \lambda_1.$$

If $\lambda_1 \leq cn$, we have $\mathrm{hook}(c_i) \leq t + \lambda_1 \leq 2cn$ for all $1 \leq i \leq n$, thus

$$d_\lambda \geq \frac{n!}{(2cn)^n} \geq \frac{n^n}{\mathrm{e}^n(2cn)^n} \geq \left(\frac{3}{\mathrm{e}}\right)^n \geq \mathrm{e}^{\beta n}\,,$$

where β is any constant such that $0 \le \beta \le \ln 3 - 1$.

Now we consider the case $cn < \lambda_1 < (1-c)n$. Let $\tilde{\lambda} = (\lambda_2, \ldots, \lambda_t)$, this is an integer partition of $n - \lambda_1$ whose Young diagram is obtained by removing the first row of λ. Applying the hook length formula for $\tilde{\lambda}$ and the fact that $d_{\tilde{\lambda}} \ge 1$ gives us:

$$\mathrm{Hook}(\tilde{\lambda}) = \frac{(n-\lambda_1)!}{d_{\tilde{\lambda}}} \le (n-\lambda_1)!\,.$$

Then we have

$$\begin{aligned}
\mathrm{Hook}(\lambda) &= \mathrm{Hook}(\tilde{\lambda}) \prod_{i=1}^{\lambda_1} \mathrm{hook}(c_i) \\
&\le (n-\lambda_1)! \prod_{i=1}^{\lambda_1} (\lambda_1 - i + \lambda_i') \\
&= (n-\lambda_1)!\lambda_1! \prod_{i=1}^{\lambda_1} \left(1 + \frac{\lambda_i' - 1}{\lambda_1 - i + 1}\right) \\
&\le (n-\lambda_1)!\lambda_1! \exp\left(\sum_{i=1}^{\lambda_1} \frac{\lambda_i' - 1}{\lambda_1 - i + 1}\right) \quad \text{(since } 1 + x \le \mathrm{e}^x \text{ for all } x\text{).}
\end{aligned}$$

To upper bound the exponent in the last equation, we use Chebyshev's sum inequality, which states that for any increasing sequence $a_1 \ge a_2 \ge \ldots \ge a_k$ and any decreasing sequence $b_1 \le b_2 \le \ldots \le b_k$ of real numbers, we have $k \sum_{i=1}^k a_i b_i \le \left(\sum_{i=1}^k a_i\right)\left(\sum_{i=1}^k b_i\right)$. Since the sequence $\{\lambda_i' - 1\}_{i=1}^{\lambda_1}$ is decreasing and the sequence $\{1/(\lambda_1 - i + 1)\}_{i=1}^{\lambda_1}$ is increasing, we get

$$\sum_{i=1}^{\lambda_1} \frac{\lambda_i' - 1}{\lambda_1 - i + 1} \le \frac{\sum_{i=1}^{\lambda_1}(\lambda_i' - 1)}{\lambda_1} \left(\sum_{i=1}^{\lambda_1} \frac{1}{\lambda_1 - i + 1}\right) = \frac{n - \lambda_1}{\lambda_1} \left(\sum_{i=1}^{\lambda_1} \frac{1}{i}\right).$$

For any integer $N > 0$ we have

$$\sum_{i=1}^{N} \frac{1}{i} = \left(\sum_{i=1}^{\lfloor \sqrt{N} \rfloor} \frac{1}{i}\right) + \left(\sum_{i=\lfloor \sqrt{N} \rfloor + 1}^{N} \frac{1}{i}\right) \le \left\lfloor \sqrt{N} \right\rfloor + \frac{N - \lfloor \sqrt{N} \rfloor}{\sqrt{N}} \le 2\sqrt{N}\,.$$

Hence,

$$\mathrm{Hook}(\lambda) \le (n-\lambda_1)!\lambda_1! \exp\left(\frac{2(n-\lambda_1)}{\sqrt{\lambda_1}}\right).$$

It follows that

$$\begin{aligned}
d_\lambda &\ge \frac{n!}{(n-\lambda_1)!\lambda_1! \exp\left(\frac{2(n-\lambda_1)}{\sqrt{\lambda_1}}\right)} \\
&= \binom{n}{\lambda_1} \exp\left(-\frac{2n}{\sqrt{\lambda_1}} + 2\sqrt{\lambda_1}\right) \\
&\ge \left(\frac{n}{\lambda_1}\right)^{\lambda_1} \exp\left(-\frac{2n}{\sqrt{\lambda_1}} + 2\sqrt{\lambda_1}\right)
\end{aligned}$$

$$\geq \left(\frac{1}{1-c}\right)^{\lambda_1} \exp\left(-\frac{2n}{\sqrt{\lambda_1}} + 2\sqrt{\lambda_1}\right) \qquad \text{(since } \lambda_1 < (1-c)n\text{)}$$
$$\geq \left(\frac{1}{1-c}\right)^{cn} \exp\left(-2\sqrt{n/c} + 2\sqrt{nc}\right) \qquad \text{(since } \lambda_1 \geq cn\text{)}.$$

Choosing any constant β such that $0 < \beta < c \ln \frac{1}{1-c}$, we have $d_\lambda \geq \mathrm{e}^{\beta n}$ for sufficiently large n. □

Remark The lower bound in Lemma 6 is essentially tight. To see this, consider the hook of width $(1-c)n$ and of depth cn. This hook has dimension roughly equal to $\binom{n}{cn}$, which is no more than $(e/c)^{cn}$.

6.2 Applying the Main Theorem

Now applying Theorem 1, we show that

Theorem 4. *Assume $q^{k^2} \leq n^{an}$ for some constant $0 < a < 1/4$. Let m be the minimal degree of the automorphism group $\mathrm{Aut}(M)$. Then for sufficiently large n, the subgroup K defined in (7) has $\mathscr{D}_K \leq O(|K|^2 \mathrm{e}^{-\delta m})$, where $\delta > 0$ is a constant.*

The proof of Theorem 4 follows the technical ideas discussed in the introduction, using the aforementioned tools for controlling the maximal normalized characters on K.

Proof of Theorem 4 To apply Theorem 1, let $0 < c < \min\{1/6, 1/4 - a\}$ be a constant and S be the set of irreps of $(\mathsf{GL}_k(\mathbb{F}_q) \times S_n) \wr \mathbb{Z}_2$ of the forms $\{\tau \times \lambda, \eta \times \mu\}$, $\{\tau \times \lambda\}$, $\{\tau \times \lambda\}'$ with $\tau, \eta \in \widehat{\mathsf{GL}_k(\mathbb{F}_q)}$ and $\lambda, \mu \in \Lambda_c$, where Λ_c is mentioned in Section 5. Firstly, we need upper bounds for $\overline{\chi}_{\overline{S}}(K)$, $|S|$, and d_S.

Since $\mathrm{Aut}(M)$ has minimal degree m, by Inequality (6) in Section 5, we have for all $\lambda \in \widehat{S_n} \setminus \Lambda_c$,

$$\overline{\chi}_\lambda(\mathrm{Aut}(M)) \leq \mathrm{e}^{-\alpha m}.$$

Combining with Lemma 6 yields

$$\overline{\chi}_{\overline{S}}(K) \leq \max\left\{\mathrm{e}^{-2\alpha m}, \mathrm{e}^{-\beta n}\right\} \leq \mathrm{e}^{-\delta m},$$

for some constant $\delta > 0$ (we can set $\delta = \min\{2\alpha, \beta\}$).

Since $\left|\widehat{\mathsf{GL}_k(\mathbb{F}_q)}\right| \leq \left|\mathsf{GL}_k(\mathbb{F}_q)\right| \leq q^{k^2}$ and by Lemma 5, we have

$$|S| \leq \left|\widehat{\mathsf{GL}_k(\mathbb{F}_q)}\right|^2 |\Lambda_c|^2 \leq q^{2k^2} \mathrm{e}^{O(\sqrt{n})}.$$

To bound d_S, we start with bounding the dimension of each representation in S. A representation $\{\tau \times \lambda, \eta \times \mu\}$ in S has dimension

$$2d_{\tau\times\lambda} d_{\eta\times\mu} = 2d_\tau d_\lambda d_\eta d_\mu \leq 2d_\tau d_\eta n^{2cn} \leq 2q^{k^2} n^{2cn},$$

where the first inequality follows Lemma 5. The last inequality holds because for any $\tau \in \widehat{\mathsf{GL}_k(\mathbb{F}_q)}$,

$$d_\tau^2 \leq \sum_{\rho \in \widehat{\mathsf{GL}_k(\mathbb{F}_q)}} d_\rho^2 = |\mathsf{GL}_k(\mathbb{F}_q)|.$$

Similarly, a representation $\{\tau \times \lambda\}$ or $\{\tau \times \lambda\}'$ in S has dimension $d_{\tau\times\lambda}^2 \leq q^{k^2} n^{2cn}$. Hence, the maximal dimension of a representation in the set S is

$$d_S \leq 2q^{k^2} n^{2cn}.$$

Let $4a+4c<d<1$ be a constant and let γ_1 be any constant such that $0<\gamma_1<d-4c-4a$. Now we set the dimension threshold $D=n^{dn}$. From the upper bounds on $|S|$ and d_S, we get

$$\begin{aligned} |S|\frac{d_S^2}{D} &\leq 4q^{4k^2} \mathrm{e}^{O(\sqrt{n})} n^{(4c-d)n} \\ &\leq 4\mathrm{e}^{O(\sqrt{n})} n^{(4a+4c-d)n} && \text{(since } q^{k^2} \leq n^{an}) \\ &\leq n^{-\gamma_1 n} && \text{for sufficiently large } n. \end{aligned}$$

Let L be the set of all irreps of $(\mathsf{GL}_k(\mathbb{F}_q)\times S_n)\wr\mathbb{Z}_2$ of dimension at least D. Bounding $|L|$ by the number of irreps of $(\mathsf{GL}_k(\mathbb{F}_q)\times S_n)\wr\mathbb{Z}_2$, which is no more than square of the number of irreps of $\mathsf{GL}_k(\mathbb{F}_q)\times S_n$, we have

$$|L| \leq \left|\widehat{\mathsf{GL}_k(\mathbb{F}_q)}\right|^2 \left|\widehat{S_n}\right|^2 \leq \left|\mathsf{GL}_k(\mathbb{F}_q)\right|^2 p(n)^2.$$

Hence, for sufficiently large n,

$$\begin{aligned} \frac{|L|D^2}{\left|(\mathsf{GL}_k(\mathbb{F}_q)\times S_n)\wr\mathbb{Z}_2\right|} &\leq \frac{\left|\mathsf{GL}_k(\mathbb{F}_q)\right|^2 p(n)^2 n^{2dn}}{2\left|(\mathsf{GL}_k(\mathbb{F}_q)\right|^2 |S_n|^2} = \frac{p(n)^2 n^{2dn}}{2(n!)^2} \\ &\leq \frac{\mathrm{e}^{O(\sqrt{n})} n^{2dn}}{2n^{2n}\mathrm{e}^{-2n}} \\ &\leq \mathrm{e}^{O(n)} n^{2(d-1)n} \leq n^{-\gamma_2 n} \qquad \text{so long as } \gamma_2 < 2(1-d). \end{aligned}$$

By Theorem 1, we have

$$\mathscr{D}_K \leq 4|K|^2(\mathrm{e}^{-\delta m} + n^{-\gamma_1 n} + n^{-\gamma_2 n}) \leq 4|K|^2(\mathrm{e}^{-\delta m} + n^{-\gamma n}),$$

for some constant $\gamma > 0$. This completes the proof.

□

As $q^{k^2} \leq n^{an}$, we have $\log\left|(\mathsf{GL}_k(\mathbb{F}_q)\times S_n)\wr\mathbb{Z}_2\right| = O(\log n! + \log q^{k^2}) = O(n\log n)$. Hence, the subgroup K is indistinguishable if $|K|^2\mathrm{e}^{-\delta m} \leq (n\log n)^{-\omega(1)}$. The size of the subgroup K is given by $|K| = 2|H_0|^2$, and $|H_0| = |\mathrm{Aut}(M)| \times |\mathrm{Fix}(M)|$, where

$$\mathrm{Fix}(M) \overset{\text{def}}{=} \left\{S \in \mathsf{GL}_k(\mathbb{F}_q) \mid SM = M\right\}$$

is the set of scramblers fixing M. To bound the size of $\mathrm{Fix}(M)$, we record an easy fact which can be obtained by the orbit-stabilizer formula:

Fact 2. *Let r be the column rank of M. Then* $|\mathrm{Fix}(M)| \leq q^{\ell k(k-r)}$.

Proof: WLOG, assume the first r columns of M are $\mathbb{F}_{q^\ell}$-linearly independent, and each remaining column is an $\mathbb{F}_{q^\ell}$-linear combination of the first r columns. Let N be the $k\times r$ matrix consisting of the first r columns of M. Then we can decompose M as $M = (N \mid NA)$, where A is an $r \times (n-r)$ matrix with entries in $\mathbb{F}_{q^\ell}$. Clearly, $\mathrm{Fix}(M) = \mathrm{Fix}(N)$. Consider the action of $\mathsf{GL}_k(\mathbb{F}_{q^\ell})$ on the set of $k \times r$ matrices over $\mathbb{F}_{q^\ell}$. Under this action, the stabilizer of N contains $\mathrm{Fix}(N)$, and the orbit of the matrix N, denoted $\mathrm{Orb}(N)$, consists of all $k \times r$ matrices over $\mathbb{F}_{q^\ell}$ whose columns are $\mathbb{F}_{q^\ell}$-linearly independent. Thus, $|\mathrm{Orb}(N)| = (q^{\ell k}-1)(q^{\ell k}-q^\ell)\ldots(q^{\ell k}-q^{\ell(r-1)})$. By the orbit-stabilizer formula, we have

$$\begin{aligned}|\mathrm{Fix}(N)| \leq \frac{|\mathsf{GL}_k(\mathbb{F}_{q^\ell})|}{|\mathrm{Orb}(N)|} &= \frac{(q^{\ell k}-1)(q^{\ell k}-q^\ell)\ldots(q^{\ell k}-q^{\ell(k-1)})}{(q^{\ell k}-1)(q^{\ell k}-q^\ell)\ldots(q^{\ell k}-q^{\ell(r-1)})} \\ &= (q^{\ell k}-q^{\ell r})(q^{\ell k}-q^{\ell(r+1)})\cdots(q^{\ell k}-q^{\ell(k-1)}) \leq q^{\ell k(k-r)}\,.\end{aligned}$$

□

Corollary 1. *Assume $q^{k^2} \leq n^{0.2n}$ and the automorphism group $\mathrm{Aut}(M)$ has minimal degree $\Omega(n)$. Let r be the column rank of M. Then the subgroup K defined in (7) has $\mathscr{D}_K \leq |\mathrm{Aut}(M)|^4 q^{4\ell k(k-r)} \mathrm{e}^{-\Omega(n)}$. In particular, the subgroup K is indistinguishable if, further, $|\mathrm{Aut}(M)| \leq \mathrm{e}^{o(n)}$ and $r \geq k - o(\sqrt{n})/\ell$.*

The constraint $q^{k^2} \leq n^{0.2n}$ implies $\log|\mathsf{GL}_k(\mathbb{F}_q)| = O(n\log n)$, so Alice only needs to flip $O(n\log n)$ bits to pick a random S from $\mathsf{GL}_k(\mathbb{F}_q)$. Thus she needs only $O(n\log n)$ coin flips overall to generate her private key.

6.3 HSP-hard Instances of Code Equivalence

Definition 6. *If M is a generator matrix or parity matrix of linear code C, and the subgroup K defined in* (7) *is indistinguishable, then we say that the code C is an* HSP-hard instance of Code Equivalence, *or* HSP-hard *for short.*

In other words, we say a linear code C is *HSP-hard* if strong quantum Fourier sampling, or more generally any measurement of a coset state, reveals negligible information about the permutation between C and any code equivalent to C. Our results above suggest that the HSP-hardness of a linear code C may be related to its (permutation) automorphism group. To make this precise, let us recall some definitions.

Definition 7. *Let C be a q-ary $[n,k]$-linear code. The* automorphism group *of C, denoted* $\mathrm{Aut}(C)$, *is the set of $n \times n$* monomial matrices *Q over $\mathbb{F}_q$ such that $SGQ = G$ for some generator matrix G of C and some $k \times k$ invertible matrix S over $\mathbb{F}_q$ [19, pg. 238]. The* permutation automorphism group *of C, denoted* $\mathrm{PAut}(C)$, *is the subgroup of* $\mathrm{Aut}(C)$ *consisting of all permutation matrices in* $\mathrm{Aut}(C)$, *i.e.,* $\mathrm{PAut}(C) = \mathrm{Aut}(C) \cap S_n$.

We remark that if the code C is binary, i.e., $q = 2$, its permutation automorphism group coincides with its automorphism group. Moreover, the notion of permutation automorphism group agrees with the notion of automorphism group used by van Lint [36, pg.51] and Stichtenoth [35].

Observe that in the case the matrix M is a generator matrix of the code C, we have $\mathrm{Aut}(M) = \mathrm{PAut}(C)$ and M has full rank. Thus, Corollary 1 immediately gives us:

Corollary 2. *Let C be a q-ary $[n,k]$-linear code such that $q^{k^2} \leq n^{0.2n}$. If $|\mathrm{PAut}(C)| \leq \mathrm{e}^{o(n)}$ and the minimal degree of $\mathrm{PAut}(C)$ is $\Omega(n)$, then C is HSP-hard.*

In the remaining part of this section, we will present a few specific cases of linear codes that are HSP-hard. Most of these cases have been used or could be used in code-based cryptosystems.

6.3.1 Rational Goppa Codes (or Generalized Reed-Solomon Codes)

The first case of an HSP-hard linear code is rational Goppa codes, also known as Generalized Reed-Solomon (GRS) codes described in Subsection 6.3.2. Our tool to show the HSP-hardness of rational Goppa codes is due to a theorem of Stichtenoth [35].

Theorem 5. *[Stichtenoth [35]] Suppose $2 \leq k \leq n-2$. Then the (permutation) automorphism group of any rational Goppa $[n,k]$-code over a field F is isomorphic to a subgroup of $\mathsf{PGL}_2(F)$, the projective linear group over F.*

In particular, we will use Stichtenoth [35]'s theorem to control the permutation automorphism group of a rational Goppa code as follows.

Lemma 7. *Suppose $2 \leq k \leq n-2$. Let C be a rational Goppa code of length n and dimension k over a field F, then $\mathrm{PAut}(C) \leq |F|^3$ and $\mathrm{PAut}(C)$ has minimal degree at least $n-2$.*

Proof: By Theorem 5, we have $\mathrm{PAut}(C) \subseteq \mathsf{PGL}_2(F)$. Thus,

$$|\mathrm{PAut}(C)| \leq |\mathsf{PGL}_2(F)| \leq |F|^3 .$$

The lower bound of $n-2$ on the minimal degree of $\mathrm{PAut}(C)$ is obtained by the observation that any transformation in $\mathsf{PGL}_2(F)$ that fixes at least three distinct projective lines must be the identity. □

The following theorem, which shows the HSP-hardness of rational Goppa codes, follows immediately from Corollary 2 and Lemma 7.

Theorem 6. *Let C be a q-ary rational Goppa code of length n and dimension k such that $q^{k^2} \leq n^{0.2n}$. Then C is an HSP-hard instance of Code Equivalence.*

Rational Goppa codes were proposed to be used in a variant of the McEliece cryptosystem by Janwa and Moreno [14]. However, this McEliece variant is broken by Sidelnikov and Shestakov [32]'s structural attack. Therefore, the HSP-hardness of rational Goppa codes has little bearing on the post-quantum security of such a code-based cryptosystem.

6.3.2 Alternant Codes

Recall that alternant codes are subfield subcodes of the generalized Reed-Solomon (GRS) codes. Formally, let $\mathbf{v} = (v_1, \ldots, v_n)$, where v_i's are nonzero elements of the field $\mathbb{F}_{q^\ell}$, and let $\alpha = (\alpha_1, \ldots, \alpha_n)$, where α_i's are distinct elements of $\mathbb{F}_{q^\ell}$. For each integer $k_0 < n$, the GRS code $\mathsf{GRS}_{k_0}(\alpha, \mathbf{v})$ over the field $\mathbb{F}_{q^\ell}$ is defined as follows: [d]

$$\mathsf{GRS}_{k_0}(\alpha,\mathbf{v}) := \left\{ (v_1 f(\alpha_1),\ldots,v_n f(\alpha_n)) \mid f(X) \in \mathbb{F}_{q^\ell}[X],\ \deg f < k_0 \right\}.$$

Then the *alternant code* $\mathscr{A}_{k_0}(\alpha,\mathbf{v})$ is the q-ary code consisting of all codewords from $\mathsf{GRS}_{k_0}(\alpha,\mathbf{v})$ with components in $\mathbb{F}_q$. In other words, $\mathscr{A}_{k_0}(\alpha,\mathbf{v})$ is the intersection of $\mathsf{GRS}_{k_0}(\alpha,\mathbf{v})$ and $\mathbb{F}_q^n$. Readers can see [19, Ch.12] for backgrounds on alternant codes.

Fact 3. $\mathsf{GRS}_{k_0}(\alpha,\mathbf{v})$ *has $k_0 \times n$ generator matrices over a finite field $\mathbb{F}_{q^\ell}$ of the following form:*

$$M_{k_0}(\alpha,\mathbf{v}) := \begin{pmatrix} v_1 f_1(\alpha_1) & v_2 f_1(\alpha_2) & \cdots & v_n f_1(\alpha_n) \\ v_1 f_2(\alpha_1) & v_2 f_2(\alpha_2) & \cdots & v_n f_2(\alpha_n) \\ \vdots & \vdots & \ddots & \vdots \\ v_1 f_{k_0}(\alpha_1) & v_2 f_{k_0}(\alpha_2) & \cdots & v_n f_{k_0}(\alpha_n) \end{pmatrix} \tag{9}$$

where $f_1,\ldots,f_{k_0}$ are $\mathbb{F}_{q^\ell}$-linearly independent polynomials in $\mathbb{F}_{q^\ell}[X]$ of degree less than k_0.

Fact 4. *The dual code of* $\mathsf{GRS}_{k_0}(\alpha,\mathbf{v})$ *is* $\mathsf{GRS}_{n-k_0}(\alpha,\mathbf{v})$.

It follows that $\mathsf{GRS}_{k_0}(\alpha,\mathbf{v})$, and therefore $\mathscr{A}_{k_0}(\alpha,\mathbf{v})$, has an $(n-k_0)\times n$ parity check matrix of the form $M_{n-k_0}(\alpha,\mathbf{v})$. With this fact and using the result of Stichtenoth [35] again, we can show the HSP-hardness of alternant codes as follows.

Theorem 7. *Suppose $q^{(n-k_0)^2} \le n^{0.2n}$ and $q^{3\ell} \le \mathrm{e}^{o(n)}$. Then the alternant code $\mathscr{A}_{k_0}(\alpha,\mathbf{v})$ is HSP-hard.*

Proof: Let $k = n-k_0$ and consider the case where the $k\times n$ matrix M is of the form $M_{n-k_0}(\alpha,\mathbf{v})$. Thus, M is a parity check matrix of the code $\mathscr{A}_{k_0}(\alpha,\mathbf{v})$.

On the other hand, M is also a generator matrix of the code $\mathsf{GRS}_{n-k_0}(\alpha,\mathbf{v})$, which implies that M has full rank and that $\mathrm{Aut}(M) = \mathrm{PAut}(\mathsf{GRS}_{n-k_0}(\alpha,\mathbf{v}))$. Since $\mathsf{GRS}_{n-k_0}(\alpha,\mathbf{v})$ is a rational Goppa code over $\mathbb{F}_{q^\ell}$, by Lemma 7, we have $\mathrm{PAut}(\mathsf{GRS}_{n-k_0}(\alpha,\mathbf{v}))$ has size at most $q^{3\ell}$ and has minimal degree at least $n-2$.

The proof is then completed by applying Corollary 1. □

Note that if M is a parity check matrix of the form M_{n-k_0}, a parity check matrix $\overline{M}$ over the subfield $\mathbb{F}_q$ for $\mathscr{A}_{k_0}(\alpha,\mathbf{v})$ can be obtained by replacing each element in M with the corresponding column vector in $\mathbb{F}_q^\ell$. Thus, $\overline{M}$ has $\ell(n-k_0)$ rows, which implies that $\mathscr{A}_{k_0}(\alpha,\mathbf{v})$ has dimension $k' \ge n-\ell(n-k_0)$ (also $k' \le k_0$) [19, pg. 334]. The sufficient condition specified in Theorem 7 for $\mathscr{A}_{k_0}(\alpha,\mathbf{v})$ to be HSP-hard requires $\ell(n-k_0)$ to be small, and thus, the dimension of $\mathscr{A}_{k_0}(\alpha,\mathbf{v})$ to be large.

In code-based cryptography, many subclasses of alternant codes have been proposed to be used. One important subclass is binary Goppa codes, which is used in McEliece and Niederreiter cryptosystems. Others subclasses are *generalized Srivastava codes* [25], *quasi-monoidic alternant codes* [5], *quasi-cyclic alternant codes* (a subclass of these codes, namely quasi-cyclic Goppa codes, has been proposed for use to reduce key sizes [6, 9, 4]). Technically, since alternant codes have an efficient

[d] More precisely, α_i can be ∞, in which case, $f(\alpha_i)$ is evaluated by the convention: $f(\infty)$ is the X^{k-1}-coefficient of $f(X)$.

decoding algorithms [19, §9], any alternant code can be used to build a code-based cryptosystem. Theorem 7 suggests that a cryptosystem using alternant codes, with appropriate parameters, can resist quantum attacks that use the standard QFS method to recover secret permutations. Of course, such a cryptosystem is still subject to other quantum (or classical) attacks. Therefore, the HSP-hardness of alternant codes does not guarantee that any cryptosystem using an alternant code, including the McEliece and Niederreiter systems, are secure against quantum and classical adversaries.

6.3.3 Reed-Muller Codes

We now consider binary Reed-Muller codes, which are used in the Sidelnikov cryptosystem [31]. Let m and r be positive integers with $r < m$, and let $n = 2^m$. Fix an ordered list $(\alpha_1, \ldots, \alpha_n)$ of all 2^m binary vectors of length m, i.e., $\mathbb{F}_2^m = \{\alpha_1, \ldots, \alpha_n\}$. The r^{th}-order binary Reed-Muller code of length n, denoted $\mathsf{RM}(r,m)$, consists of codewords of the form $(f(\alpha_1), \ldots, f(\alpha_n))$, where $f \in \mathbb{F}_2[X_1, \ldots, X_m]$ ranges over all binary polynomials on m variables of degree at most r. The code $\mathsf{RM}(r,m)$ has dimension equal to the number of monomials of degree at most r,

$$k = \sum_{j=0}^{r} \binom{m}{j} .$$

To apply Corollary 2, we first need to choose r such that $k^2 \le 0.2\,m2^m$. If $r < 0.1\,m$ then $k < r\binom{m}{0.1m} < r2^{0.47\,m}$, and $k^2 \le 0.2\,m2^m$ for sufficiently large m.

Next, we examine the permutation automorphism group of the Reed-Muller codes. Recall that for any binary code, the permutation automorphism group coincides with the automorphism group. Let $\mathsf{GL}_m(\mathbb{F}_2)$ denote the set of invertible $m \times m$ matrices over $\mathbb{F}_2$. It is known [31, 19] that $\mathrm{Aut}(\mathsf{RM}(r,m))$ coincides with the general affine group of the space $\mathbb{F}_2^m$. In other words, $\mathrm{Aut}(\mathsf{RM}(r,m))$ consists of all affine permutations of the form $\sigma_{A,\beta}(x) = Ax + \beta$ where $A \in \mathsf{GL}_m(\mathbb{F}_2)$ and $\beta \in \mathbb{F}_2^m$. Hence the size of $\mathrm{Aut}(\mathsf{RM}(r,m))$ is

$$|\mathrm{Aut}(\mathsf{RM}(r,m))| = |\mathsf{GL}_m(\mathbb{F}_2)| \cdot |\mathbb{F}_2^m| \le 2^{m^2+m} = 2^{O(\log^2 n)} \le \mathrm{e}^{o(n)} .$$

Finally, we compute the minimal degree of $\mathrm{Aut}(\mathsf{RM}(r,m))$ as follows.

Proposition 5. *The minimal degree of* $\mathrm{Aut}(\mathsf{RM}(r,m))$ *is exactly* $2^{m-1} = n/2$.

Proof: The minimal degree is at most 2^{m-1}, since there is an affine transformation with support 2^{m-1}. For example, let A be the $m \times m$ binary matrix with 1s on the diagonal and the $(1,m)$-entry and 0s elsewhere. Then $\sigma_{A,0}$ fixes the subspace spanned by the first $m-1$ standard basis vectors. Its support is the complement of this subspace, which has size 2^{m-1}.

Conversely, if $\sigma_{A,\beta}$ fixes a set S that spans $\mathbb{F}_2^m$, then $\sigma_{A,\beta}$ must be the identity. To see this, let $x_0 \in S$ and consider the translated set $S' = S - x_0$. Then $Ay = y$ for any $y \in S'$, since

$$y + x_0 = \sigma_{A,\beta}(y + x_0) = Ay + \sigma_{A,\beta}(x_0) = Ay + x_0 .$$

If S spans $\mathbb{F}_2^m$ then so does S', in which case $A = \mathbf{1}$. Then $\beta = 0$, since otherwise $\sigma_{\mathbf{1},\beta}$ doesn't fix anything, and $\sigma_{\mathbf{1},0}$ is the identity.

Moreover, any set S of size greater than 2^{m-1} spans $\mathbb{F}_2^m$. To see this, let B be a maximal subset of S consisting of linearly independent vectors. Since B spans S, we have $|S| \le 2^{|B|}$. Thus if $|S| > 2^{m-1}$ we have $|B| = m$, so B and therefore S span $\mathbb{F}_2^m$. Thus no nonidentity affine transformation can fix more than 2^{m-1} points, and the minimal degree is at least 2^{m-1}. $\square$

We have proved the following:

Theorem 8. *Reed-Muller codes* $\mathsf{RM}(r,m)$ *with* $r \leq 0.1\,m$ *and* m *sufficiently large are HSP-hard.*

In the original proposal of Sidelnikov [31], r is taken to be a small constant, where the Reed-Muller codes have low rate. It is worth noting that the attack of Minder and Shokrollahi [21] becomes infeasible in the high-rate case where r is large, due to the difficulty of finding minimum-weight codewords, while Theorem 8 continues to apply. However, as those authors point out, taking large r degrades the performance of Reed-Muller codes, and presumably opens the Sidelnikov system to other classical attacks.

7 Strong Fourier Sampling over $\mathsf{GL}_2(\mathbb{F}_q)$

Now we supplement applications of the main theorem (Theorem 1) with the case of the finite general linear group $G = \mathsf{GL}_2(\mathbb{F}_q)$, whose structure and irreps are well known [8, §5.2]. Our lower bounds complement previous work of Ivanyos [13] on the hidden subgroup problem over general linear groups. We remark also that our negative results may also have applications to quantum-resistant cryptosystems: In a talk [37] at MIT in 2007, Umesh Vazirani outlined a one-way function whose security is related to solving the hidden subgroup problem over the general linear group.

7.1 Irreducible Representations of $\mathsf{GL}_2(\mathbb{F}_q)$

We will first present preliminary background on the structure of $\mathsf{GL}_2(\mathbb{F}_q)$ followed by description of its irreps. We refer readers to [8, §5.2] for the missing technical details in this part.

Viewing $\mathsf{GL}_2(\mathbb{F}_q)$ as the group of all $\mathbb{F}_q$-linear invertible endomorphisms of the quadratic extension $\mathbb{F}_{q^2}$ of $\mathbb{F}_q$, we have a large subgroup of $\mathsf{GL}_2(\mathbb{F}_q)$ that is isomorphic to $\mathbb{F}_{q^2}^*$ via the identification:

$$\left\{ f_\xi \mid \xi \in \mathbb{F}_{q^2}^* \right\} \simeq \mathbb{F}_{q^2}^*\,, \quad f_\xi \leftrightarrow \xi$$

where $f_\xi : \mathbb{F}_{q^2} \to \mathbb{F}_{q^2}$ is the $\mathbb{F}_q$-linear map given by $f_\xi(v) = \xi v$ for all $v \in \mathbb{F}_{q^2}$.

To turn each map f_ξ into a matrix form, we fix a basis $\{1,\gamma\}$ of $\mathbb{F}_{q^2}$ as a vector space over $\mathbb{F}_q$. For each $\xi \in \mathbb{F}_{q^2}$, writing $\xi = \xi_{x,y} = x + \gamma y$ for some $x, y \in \mathbb{F}_q$, then the map f_ξ corresponds to the matrix $\begin{pmatrix} x & \gamma^2 y \\ y & x \end{pmatrix}$, since $f_\xi(1) = x + \gamma y$ and $f_\xi(\gamma) = \gamma^2 y + \gamma x$. Hence, we can rewrite the above identification as

$$\left\{ \begin{pmatrix} x & \gamma^2 y \\ y & x \end{pmatrix} \mid x, y \in \mathbb{F}_q, x \neq 0 \text{ or } y \neq 0 \right\} \simeq \mathbb{F}_{q^2}^*\,, \quad \begin{pmatrix} x & \gamma^2 y \\ y & x \end{pmatrix} \leftrightarrow \xi_{x,y} = x + \gamma y\,.$$

For example, if q is odd, choose a generator ε of $\mathbb{F}_q^*$, then ε must be non-square in $\mathbb{F}_q$, which implies that $\left\{1, \sqrt{\varepsilon}\right\}$ form a basis of $\mathbb{F}_{q^2}$ as a vector space over $\mathbb{F}_q$. In such a case, we can define $\xi_{x,y} = x + y\sqrt{\varepsilon}$.

Conjugacy classes. The group $\mathsf{GL}_2(\mathbb{F}_q)$ has four types of conjugacy classes in Table 1, with representatives described as follows:

$$a_x = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix} \quad b_x = \begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix} \quad c_{x,y} = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix} \quad d_{x,y} = \begin{pmatrix} x & \gamma^2 y \\ y & x \end{pmatrix}\,. \tag{10}$$

Table 1. Conjugacy classes of $\mathsf{GL}_2(\mathbb{F}_q)$, where $[g]$ denotes the class of representative g.

class	$[a_x]$ $x \in \mathbb{F}_q^*$	$[b_x]$ $x \in \mathbb{F}_q^*$	$[c_{x,y}] = [c_{y,x}]$ $x,y \in \mathbb{F}_q^*, x \neq y$	$[d_{x,y}] = [d_{x,-y}]$ $x \in \mathbb{F}_q, y \in \mathbb{F}_q^*$
class size	1	q^2-1	q^2+q	q^2-q
number of classes	$q-1$	$q-1$	$\frac{(q-1)(q-2)}{2}$	$\frac{q(q-1)}{2}$

There are q^2-1 conjugacy classes, hence there are exactly q^2-1 irreps of $\mathsf{GL}_2(\mathbb{F}_q)$. We shall briefly describe below how to construct all those representations.

Linear representations. For each character $\alpha : \mathbb{F}_q^* \to \mathbb{C}^*$ of the cyclic group $\mathbb{F}_q^*$, we have a one-dimensional representation U_α of $\mathsf{GL}_2(\mathbb{F}_q)$ defined by:

$$U_\alpha(g) = \alpha(\det(g)) \qquad \forall g \in \mathsf{GL}(2,q)\,.$$

To compute $U_\alpha(d_{x,y})$, we shall use the following fact:

$$\det\begin{pmatrix} x & \gamma^2 y \\ y & x \end{pmatrix} = \mathrm{Norm}_{\mathbb{F}_{q^2}/\mathbb{F}_q}(\xi_{x,y}) = \xi_{x,y} \cdot \xi_{x,y}^q = \xi_{x,y}^{q+1}\,.$$

Recall that there are $q-1$ characters of $\mathbb{F}_q^* = \langle \varepsilon \rangle$ corresponding to $q-1$ places where the generator ε can be sent to. The linear representation U_{α_0}, where α_0 is the character sending ε to 1, is indeed the trivial representation, denoted U.

Irreducible representations by action on $\mathbb{P}^1(\mathbb{F}_q)$. $\mathsf{GL}_2(\mathbb{F}_q)$ acts transitively on the projective line $\mathbb{P}^1(\mathbb{F}_q)$ in the natural way:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot [x:y] = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = [ax+by : cx+dy]\,,$$

in which the stabilizer of the infinite point $[1:0]$ is the Borel subgroup B:

$$B = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \mid a,d \in \mathbb{F}_q^*,\ b \in \mathbb{F}_q \right\}.$$

The permutation representation of $\mathsf{GL}_2(\mathbb{F}_q)$ given by this action on $\mathbb{P}^1(\mathbb{F}_q)$ has dimension $q+1$ and decomposes into the trivial representation U and a q-dimensional representation V. The character of V is given as follows:

$$\chi_V(a_x) = q \quad \chi_V(b_x) = 0 \quad \chi_V(c_{x,y}) = 1 \quad \chi_V(d_{x,y}) = -1\,.$$

By checking $\langle \chi_V, \chi_V \rangle = 1$, we see that V is irreducible. Hence, for each of the $q-1$ characters α of $\mathbb{F}_q^*$, we have a q-dimensional irrep $V_\alpha = V \otimes U_\alpha$. Note that $V = V \otimes U$.

Irreducible representations induced from Borel subgroup B. For each pair of characters α, β of $\mathbb{F}_q^*$, there is a character of the subgroup B:

$$\phi_{\alpha,\beta} : B \to \mathbb{C}^* \quad \text{by} \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \mapsto \alpha(a)\beta(d)\,.$$

In other words, $\phi_{\alpha,\beta}$ is a one-dimensional representation of subgroup B. Let $W_{\alpha,\beta}$ be the representation of $\mathsf{GL}_2(\mathbb{F}_q)$ induced by $\phi_{\alpha,\beta}$. By computing characters, we have

- $W_{\alpha,\beta} = W_{\beta,\alpha}$,
- $W_{\alpha,\alpha} = U_\alpha \oplus V_\alpha$, and
- $W_{\alpha,\beta}$ is irreducible for $\alpha \neq \beta$. Each of these representations has dimension equal the index of B in $\mathsf{GL}_2(\mathbb{F}_q)$, i.e., $[\mathsf{GL}(2,q) : B] = q+1$.

There are $((q-1)^2 - (q-1))/2 = (q-1)(q-2)/2$ distinct irreps of this type.

Irreducible representations by characters of $\mathbb{F}^*_{q^2}$. Let $\varphi : \mathbb{F}^*_{q^2} \to \mathbb{C}^*$ be a character of the cyclic group $\mathbb{F}^*_{q^2}$. Since $\mathbb{F}^*_{q^2}$ can be viewed as a subgroup of $\mathsf{GL}_2(\mathbb{F}_q)$, we have the induced representation $\mathrm{Ind}\varphi$, which is not irreducible. However, it gives us a $(q-1)$-dimensional irrep with character given by

$$\chi_\varphi = \chi_{V \otimes W_{\alpha,1}} - \chi_{W_{\alpha,1}} - \chi_{\mathrm{Ind}\varphi} \quad \text{if } \varphi|_{\mathbb{F}^*_q} = \alpha.$$

Note that $\mathrm{Ind}\varphi \simeq \mathrm{Ind}\varphi^q$, thus $X_\varphi \simeq X_{\varphi^q}$. So, the characters φ of $\mathbb{F}^*_{q^2}$ with $\varphi \neq \varphi^q$ give a rise to the $\frac{1}{2}q(q-1)$ remaining irreps of $\mathsf{GL}_2(\mathbb{F}_q)$.

A summary of all irreducible characters of $\mathsf{GL}_2(\mathbb{F}_q)$ is given in Table 2 below.

Table 2. Character table of $\mathsf{GL}_2(\mathbb{F}_q)$, where α, β are characters of $\mathbb{F}^*_q$ ($\alpha \neq \beta$), and φ is a character of $\mathbb{F}^*_{q^2}$ with $\varphi^q \neq \varphi$; $a_x, b_x, c_{x,y}, d_{x,y}$ are elements of $\mathsf{GL}_2(\mathbb{F}_q)$ defined in (10), and $d_\rho = \chi_\rho(a_1)$ is the dimension of ρ.

Irrep ρ	d_ρ	$\chi_\rho(a_x)$	$\chi_\rho(b_x)$	$\chi_\rho(c_{x,y})$	$\chi_\rho(d_{x,y})$
U_α	1	$\alpha(x^2)$	$\alpha(x^2)$	$\alpha(xy)$	$\alpha(\xi_{x,y}^{q+1})$
V_α	q	$q\alpha(x^2)$	0	$\alpha(xy)$	$-\alpha(\xi_{x,y}^{q+1})$
$W_{\alpha,\beta}$	$q+1$	$(q+1)\alpha(x)\beta(x)$	$\alpha(x)\beta(x)$	$\alpha(x)\beta(y)+\alpha(y)\beta(x)$	0
X_φ	$q-1$	$(q-1)\varphi(x)$	$-\varphi(x)$	0	$-\varphi(\xi_{x,y}) - \varphi(\xi_{x,y}^q)$

From the character table of $\mathsf{GL}_2(\mathbb{F}_q)$, we can easily draw the following facts:

Fact 5. *Let σ be an irrep of $\mathsf{GL}_2(\mathbb{F}_q)$. Then*

- *For all $g \in \mathsf{GL}_2(\mathbb{F}_q)$, $|\chi_\sigma(g)| = d_\sigma$ if g is a scalar matrix, and $|\chi_\sigma(g)| \leq 2$ otherwise.*
- *If $d_\sigma > 1$, then $q-1 \leq d_\sigma \leq q+1$.*

7.2 Applying the Main Theorem to $\mathsf{GL}_2(\mathbb{F}_q)$

Let H be a subgroup of $\mathsf{GL}_2(\mathbb{F}_q)$. If H contains a non-identity scalar matrix, we have $\overline{\chi}_\sigma(H) = 1$ for all σ, making it impossible to find a set of irreps whose maximal normalized characters on H are small enough to apply our general main theorem (Theorem 1). For this reason, we shall assume that H does not contain scalar matrices except for the identity. An example of such a subgroup H is any group lying inside the subgroup of triangular unipotent matrices $\{T(b) \mid b \in \mathbb{F}_q\}$, where

$$T(b) := \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}.$$

From Fact 5, it is natural to choose the set S in Theorem 1 to be the set of linear (i.e., 1-dimensional) representations, and choose the dimensional threshold D to be $q-1$. However, since $\mathsf{GL}(2,q)$ has $q-1$

linear representations, i.e., $|S| = D$, we can't upper bound Δ by $|S|$. We prove the following lemma to provide a strong upper bound on Δ, which is, in this case, the maximal number of linear representations appearing in the decomposition of $\rho \otimes \rho^*$, for any nonlinear irrep ρ.

Lemma 8. *Let ρ be an irrep of $\mathsf{GL}(2,q)$. Then at most two linear representations appear in the decomposition of $\rho \otimes \rho^*$.*

To prove this lemma, we observe that if ρ is a linear irrep of $\mathsf{GL}(2,q)$ then $\rho \otimes \rho^*$ is the trivial representation. Therefore, we shall only consider the cases where ρ is non-linear.

Recall that the multiplicity of U_α in $\rho \otimes \rho^*$ is given by

$$\left\langle \chi_{\rho\otimes\rho^*}, \chi_{U_\alpha} \right\rangle = \frac{1}{|G|} \sum_{g\in G} |\chi_\rho(g)|^2 \chi_{U_\alpha}(g) = \frac{1}{|G|} \left(A(\rho,\alpha) + B(\rho,\alpha) + C(\rho,\alpha) + D(\rho,\alpha)\right),$$

where $A(\rho,\alpha), B(\rho,\alpha), C(\rho,\alpha), D(\rho,\alpha))$ are the sums of $|\chi_\rho(g)|^2 \chi_{U_\alpha}(g)$ over all elements g in the conjugacy classes with representatives of the form $a_x, b_x, c_{x,y}$ and $d_{x,y}$, respectively. That is, from the description of conjugacy classes in Table 1, we have

$$\begin{aligned}
A(\rho,\alpha) &= \sum_{x\in\mathbb{F}_q^*} |\chi_\rho(a_x)|^2 \chi_{U_\alpha}(a_x) \\
B(\rho,\alpha) &= (q^2-1) \sum_{x\in\mathbb{F}_q^*} |\chi_\rho(b_x)|^2 \chi_{U_\alpha}(b_x) \\
C(\rho,\alpha) &= \frac{1}{2}(q^2+q) \sum_{x,y\in\mathbb{F}_q^*, x\neq y} |\chi_\rho(c_{x,y})|^2 \chi_{U_\alpha}(c_{x,y}) \\
D(\rho,\alpha) &= \frac{1}{2}(q^2-q) \sum_{x,y\in\mathbb{F}_q, y\neq 0} |\chi_\rho(d_{x,y})|^2 \chi_{U_\alpha}(d_{x,y}) .
\end{aligned}$$

Our goal below will be to show that $\left\langle \chi_{\rho\otimes\rho^*}, \chi_{U_\alpha} \right\rangle = 0$ for all but two linear representations U_α and for all non-linear irreps ρ of $\mathsf{GL}_2(\mathbb{F}_q)$. We begin with the following lemma.

Lemma 9. *Let F be a finite field and $\phi : F^\times \to \mathbb{C}^*$ be a non-trivial character of the cyclic group $F^\times$, i.e., $\phi(x) \neq 1$ for some x. Then $\sum_{x\in F^\times} \phi(x) = 0$.*

Proof: Let n be the order of $F^\times$ and let τ be a generator of $F^\times$. Then $\tau^n = 1$ which implies $\phi(\tau)^n = 1$. Since ϕ is non-trivial, we must have $\phi(\tau) \neq 1$. Hence,

$$\sum_{x\in F^\times} \phi(x) = \sum_{k=0}^{n-1} \phi(\tau^k) = \sum_{k=0}^{n-1} \phi(\tau)^k = \frac{\phi(\tau)^n - 1}{\phi(\tau) - 1} = 0 .$$

□

Note that for any character α of $\mathbb{F}_q^*$, the map $\alpha^2 : \mathbb{F}_q^* \to \mathbb{C}^*$ defined by $\alpha^2(x) = \alpha(x^2)$ is also a character of $\mathbb{F}_q^*$. Hence, we have the following direct corollaries of Lemma 9.

Corollary 3. *Let α be a character of $\mathbb{F}_q^*$ such that α^2 is non-trivial. Then $\sum_{x\in\mathbb{F}_q^*} \alpha(x^2) = 0$.*

Corollary 4. *Let ρ be an irrep of $\mathsf{GL}_2(\mathbb{F}_q)$ and let α be a character of $\mathbb{F}_q^*$ such that α^2 is non-trivial. Then we always have $A(\rho,\alpha) = B(\rho,\alpha) = 0$.*

Proof: Observe that $|\chi_\rho(a_x)|$ and $|\chi_\rho(b_x)|$ do not depend on x, and $\chi_{U_\alpha}(a_x) = \chi_{U_\alpha}(b_x) = \alpha(x^2)$. Hence, to show $A(\rho,\alpha) = B(\rho,\alpha) = 0$, it suffices to use the fact that $\sum_{x\in\mathbb{F}_q^*} \alpha(x^2) = 0$. □

Remark There are at most two characters α of $\mathbb{F}_q^*$ such that α^2 is trivial. They are the trivial one, and the one that maps $\varepsilon \to \omega^{\frac{q-1}{2}}$ if q is odd, where $\omega = \mathrm{e}^{\frac{2\pi i}{q-1}}$ is a primitive $(q-1)^{\text{th}}$ root of unity, and ε is a chosen generator of the cyclic group $\mathbb{F}_q^*$. To see this, suppose $\alpha(\varepsilon) = \omega^k$, for some $k \in \{0,1,\ldots,q-2\}$. If $\alpha(\varepsilon)^2 = 1$, then $\omega^{2k} = 1$, which implies $q-1 \mid 2k$ because ω has order $q-1$. Hence $2k \in \{0, q-1\}$.

With this remark, Lemma 8 immediately follows Lemma 10 below.

Lemma 10. *Let ρ be a non-linear irrep of $\mathsf{GL}_2(\mathbb{F}_q)$ and let α be a character of $\mathbb{F}_q^*$ such that α^2 is trivial. Then U_α does not appear in the decomposition of $\rho \otimes \rho^*$.*

Proof: We will prove case by case of ρ that $C(\rho,\alpha) = D(\rho,\alpha) = 0$, which, together with Corollary 4, will complete the proof for the lemma.

Case $\rho = W_{\beta,\beta'}$. For this case, as $|\chi_{W_{\beta,\beta'}}(d_{x,y})| = 0$, we only need to show $C(W_{\beta,\beta'},\alpha) = 0$. Considering $x,y \in \mathbb{F}_q^*$ with $x \neq y$ and letting $z = x^{-1}y \neq 1$, we have

$$\begin{aligned} |\chi_{W_{\beta,\beta'}}(c_{x,y})|^2 &= [\beta(x)\beta'(y) + \beta(y)\beta'(x)][\beta(x^{-1})\beta'(y^{-1}) + \beta(y^{-1})\beta'(x^{-1})] \\ &= 2 + \beta(xy^{-1})\beta'(yx^{-1}) + \beta(yx^{-1})\beta'(xy^{-1}) \\ &= 2 + \beta(z^{-1})\beta'(z) + \beta(z)\beta'(z^{-1}) \,. \end{aligned}$$

This means $|\chi_{W_{\beta,\beta'}}(c_{x,y})|^2$ only depends on $z = x^{-1}y$. Now let $\gamma(z) = |\chi_{W_{\beta,\beta'}}(c_{x,y})|^2\alpha(z)$, we have

$$|\chi_{W_{\beta,\beta'}}(c_{x,y})|^2 \chi_{U_\alpha}(c_{x,y}) = |\chi_{W_{\beta,\beta'}}(c_{x,y})|^2 \alpha(x^2 z) = \gamma(z)\alpha(x^2).$$

Hence,

$$\begin{aligned} \sum_{x,y\in\mathbb{F}_q^*, x\neq y} |\chi_\rho(c_{x,y})|^2 \chi_{U_\alpha}(c_{x,y}) &= \sum_{x,z\in\mathbb{F}_q^*, z\neq 1} \gamma(z)\alpha(x^2) \\ &= \left(\sum_{x\in\mathbb{F}_q^*} \alpha(x^2)\right)\left(\sum_{z\in\mathbb{F}_q^*, z\neq 1} \gamma(z)\right) = 0 \end{aligned}$$

by Corollary 3, completing the proof for the case $\rho = W_{\beta,\beta'}$.

Case $\rho = V_\beta$. Since $|\chi_{V_\beta}(c_{x,y})| = 1$ and $\chi_{U_\alpha}(c_{x,y}) = \alpha(xy) = \alpha(x)\alpha(y)$,

$$\sum_{x,y\in\mathbb{F}_q^*, x\neq y} |\chi_{V_\beta}(c_{x,y})|^2 \chi_{U_\alpha}(c_{x,y}) = \sum_{x,y\in\mathbb{F}_q^*, x\neq y} \alpha(x)\alpha(y) = \left(\sum_{x\in\mathbb{F}_q^*} \alpha(x)\right)^2 - \sum_{x\in\mathbb{F}_q^*} \alpha(x^2) = 0$$

by Lemma 9 and Corollary 3. This shows $C(V_\beta, \alpha) = 0$.

Now we are going to show that $D(V_\beta, \alpha) = 0$, or equivalently, $\sum_{x,y\in\mathbb{F}_q, y\neq 0} \alpha(\xi_{x,y}^{q+1}) = 0$. We have

$$\sum_{\xi\in\mathbb{F}_{q^2}^*} \alpha(\xi^{q+1}) = \sum_{x,y\in\mathbb{F}_q, y\neq 0} \alpha(\xi_{x,y}^{q+1}) + \sum_{x\in\mathbb{F}_q^*} \alpha(\xi_{x,0}^{q+1}) = \sum_{x,y\in\mathbb{F}_q, y\neq 0} \alpha(\xi_{x,y}^{q+1})\,.$$

where in the last equality, we apply Corollary 3 and the fact that $\xi_{x,0}^{q+1} = x^{q+1} = x^2$ for all $x \in \mathbb{F}_q^*$.

Consider the map $\phi : \mathbb{F}_{q^2}^* \to \mathbb{C}^*$ given by $\phi(\xi) = \alpha(\xi^{q+1})$. Clearly, ϕ is a character of $\mathbb{F}_{q^2}^*$. Since α^2 is non-trivial and $\alpha^2(x) = \alpha(x^2) = \alpha(x^{q+1}) = \phi(x)$ for all $x \in \mathbb{F}_q^*$, the map ϕ is also non-trivial. By Lemma 9, we have $\sum_{\xi\in\mathbb{F}_{q^2}^*} \alpha(\xi^{q+1}) = 0$, which implies $D(V_\beta, \alpha) = 0$.

Case $\rho = X_\varphi$. As it is clear from the character table of $\mathsf{GL}_2(\mathbb{F}_q)$ that $C(X_\varphi, \alpha) = 0$, it remains to show $D(X_\varphi, \alpha) = 0$, or equivalently, $D_0 \overset{\text{def}}{=} \sum_{x,y\in\mathbb{F}_q, y\neq 0} |\varphi(\xi_{x,y}) + \varphi(\xi_{x,y}^q)|^2 \alpha(\xi_{x,y}^{q+1}) = 0$. We have

$$D_0 = \underbrace{\sum_{\xi\in\mathbb{F}_{q^2}^*} |\varphi(\xi) + \varphi(\xi^q)|^2 \alpha(\xi^{q+1})}_{D_1} - \underbrace{\sum_{x\in\mathbb{F}_q^*} |\varphi(\xi_{x,0}) + \varphi(\xi_{x,0}^q)|^2 \alpha(\xi_{x,0}^{q+1})}_{D_2}\,.$$

For $\xi \in \mathbb{F}_{q^2}^*$, we have

$$|\varphi(\xi) + \varphi(\xi^q)|^2 = (\varphi(\xi) + \varphi(\xi^q))(\varphi(\xi)^{-1} + \varphi(\xi^q)^{-1}) = 2 + \varphi(\xi^{q-1}) + \varphi(\xi^{1-q})\,.$$

Hence, since $x^{q-1} = 1$ for all $x \in \mathbb{F}_q^*$ and by Corollary 3,

$$D_2 = \sum_{x\in\mathbb{F}_q^*} (2 + \varphi(x^{q-1}) + \varphi(x^{1-q}))\alpha(x^{q+1}) = 3\sum_{x\in\mathbb{F}_q^*} \alpha(x^2) = 0\,.$$

The last thing we want to show is that $D_1 = 0$. Consider the map $\phi : \mathbb{F}_{q^2}^* \to \mathbb{C}^*$ given by $\phi(\xi) = \varphi(\xi^{q-1})\alpha(\xi^{q+1})$, which is apparently a character of $\mathbb{F}_{q^2}^*$. We shall see that it is non-trivial. Let ω be a generator of $\mathbb{F}_{q^2}^*$. Since $\omega^{q^2-1} = 1$, we have $\phi(\omega^{q+1}) = \alpha(\omega^{(q+1)^2}) = \alpha(\omega^{2(q+1)}) = \alpha^2(\omega^{q+1})$. On the other hand, ω^{q+1} is a generator for $\mathbb{F}_q^*$, because $\omega^{k(q+1)}$ with $k = 0, 1, \ldots, q-2$ are distinct, and $\omega^{(q-1)(q+1)} = 1$. Hence, if $\phi(\omega^{q+1}) = 1$, then $\alpha^2(x) = 1$ for all $x \in \mathbb{F}_q^*$. But since α^2 is non-trivial, we must have $\phi(\omega^{q+1}) \neq 1$, which means ϕ is non-trivial. Applying Lemma 9, we get $\sum_{\xi\in\mathbb{F}_{q^2}^*} \varphi(\xi^{q-1})\alpha(\xi^{q+1}) = 0$. Similarly, we also have $\sum_{\xi\in\mathbb{F}_{q^2}^*} \varphi(\xi^{1-q})\alpha(\xi^{q+1}) = 0$. Combining with the fact that $\sum_{\xi\in\mathbb{F}_{q^2}^*} \alpha(\xi^{q+1}) = 0$, which has been proved in the previous case, we have shown $D_1 = 0$, completing the proof.

□

Now applying Theorem 1 with S being the set of linear representations, and L being the set of non-linear irreps of $\mathsf{GL}_2(\mathbb{F}_q)$, we have:

Corollary 5. *Let H be a subgroup of $\mathsf{GL}_2(\mathbb{F}_q)$ that does not contain any scalar matrix other than the identity. Then $\mathscr{D}_H \leq 28|H|^2/q$.*

Proof of Corollary 5 Let S be the set of linear representations of $\mathsf{GL}_2(\mathbb{F}_q)$ and let $D = q-1$. Then in this case, L_D is the set of all non-linear irreps of $\mathsf{GL}_2(\mathbb{F}_q)$.

Since $\overline{\chi}_\sigma(H) \leq 2/(q-1)$ for all nonlinear irrep σ, we have

$$\overline{\chi}_{\overline{S}}(H) \leq 2/(q-1) \leq 0.5/|H|\,.$$

To bound the second term in the bound of 1, we have $\Delta \leq 2$ by Lemma 8 and $d_S = 1$, thus

$$\Delta\frac{d_S^2}{D} \leq 2/(q-1) \leq 3/q\,.$$

As $|G| = (q-1)^2 q(q+1)$ and $|\overline{L_D}| = |S| = q-1$, we have

$$\frac{|\overline{L_D}|D^2}{|G|} = \frac{(q-1)^3}{(q-1)^2 q(q+1)} = \frac{q-1}{q(q+1)} < 1/q\,.$$

By Theorem 1, $\mathscr{D}_H \leq 4|H|^2\,(7/q)$. □

In particular, H is indistinguishable by strong Fourier sampling over $\mathsf{GL}_2(\mathbb{F}_q)$ if $|H| \leq q^\delta$ for some $\delta < 1/2$, because in that case we have $\mathscr{D}_H \leq 28q^{2\delta-1} \leq \log^{-c}|\mathsf{GL}_2(\mathbb{F}_q)|$ for all constant $c > 0$.

Examples of indistinguishable subgroups of $\mathsf{GL}_2(\mathbb{F}_q)$. As a specific example, consider a cyclic subgroup H_b generated by a triangular unipotent matrix $T(b)$ for any $b \neq 0$. Since $T(b)^k = T(kb)$ for any integer $k \geq 0$, the order of H_b is the least positive integer k such that $kb = 0$. In particular, the order of H_b equals the characteristic of the finite field $\mathbb{F}_q$. Suppose $q = p^n$ for some prime number p and $n > 2$. Then $\mathbb{F}_q$ has characteristic p, and hence, $|H_b| = p$. By Corollary 5, we have $\mathscr{D}_{H_b} \leq O(p^{2-n})$.

Similarly, consider a subgroup $H_{a,b}$ generated by two distinct non-identity elements $T(a)$ and $T(b)$. Since elements of $H_{a,b}$ are of the form $T(ka+\ell b)$ for $k,\ell \in \{0,1,\ldots,p-1\}$, we have $|H_{a,b}| \leq p^2$. Thus, the distinguishability of $H_{a,b}$ using strong Fourier sampling over $\mathsf{GL}_2(\mathbb{F}_{p^n})$ is $O(p^{4-n})$. Clearly, both H_b and $H_{a,b}$ are indistinguishable, for n sufficiently large. More generally, any subgroup generated by a constant number of triangular unipotent matrices is indistinguishable.

References

[1] L. Babai, D. Grigoriev, and D. Mount. Isomorphism of directed graphs with bounded eigenvalue multiplicity. In *Proc. 14th Symposium on Theory of Computing*, pages 310–324, 1982.

[2] László Babai. On the complexity of canonical labeling of strongly regular graphs. *SIAM J. Computing*, 9(1):212–216, 1980.

[3] László Babai and Eugene M. Luks. Canonical labeling of graphs. In *Proc. 15th Symposium on Theory of Computing*, pages 171–183, 1983.

[4] M. Baldi and G. F. Chiaraluce. Cryptanalysis of a new instance of McEliece cryptosystem based on QC-LDPC codes. In *IEEE International Symposium on Information Theory*, March 2007.

[5] Paulo S. L. M. Barreto, Richard Lindner, and Rafael Misoczki. Monoidic codes in cryptography. In *Proceedings of the 4th international conference on Post-Quantum Cryptography*, PQCrypto'11, pages 179–199, Berlin, Heidelberg, 2011. Springer-Verlag. ISBN 978-3-642-25404-8.

[6] Thierry P. Berger, Pierre-Louis Cayrel, Philippe Gaborit, and Ayoub Otmani. Reducing key length of the McEliece cryptosystem. In *Proceedings of the 2nd International Conference on Cryptology in Africa: Progress in Cryptology*, AFRICACRYPT '09, pages 77–97, Berlin, Heidelberg, 2009. Springer-Verlag. ISBN 978-3-642-02383-5.

[7] Hang Dinh, Cristopher Moore, and Alexander Russell. McEliece and Niederreiter cryptosystems that resist quantum Fourier sampling attacks. In *Advances in Cryptology - CRYPTO 2011: 31st Annual Cryptology Conference Proceedings*, volume 6841 of *Lecture Notes in Computer Science*, pages 761–779. Springer-Verlag, August 2011.

[8] William Fulton and Joe Harris. *Representation Theory - A First Course.* Springer-Verlag, New York Inc., 1991.

[9] P. Gaborit. Shorter keys for code based cryptography. In *Proceedings of the 2005 International Workshop on Coding and Cryptography (WCC 2005)*, May 2005.

[10] Michelangelo Grigni, Leonard J. Schulman, Monica Vazirani, and Umesh Vazirani. Quantum mechanical algorithms for the nonabelian hidden subgroup problem. *Combinatorica*, 24(1):137–154, 2004.

[11] Sean Hallgren, Cristopher Moore, Martin Rötteler, Alexander Russell, and Pranab Sen. Limitations of quantum coset states for graph isomorphism. In *STOC '06: Proceedings of the thirty-eighth annual ACM symposium on Theory of computing*, pages 604–617, 2006.

[12] Sean Hallgren, Cristopher Moore, Martin Rötteler, Alexander Russell, and Pranab Sen. Limitations of quantum coset states for graph isomorphism. *Journal of the ACM*, 57(6), 2010.

[13] Gábor Ivanyos. Finding hidden borel subgroups of the general linear group. *Quantum Info. Comput.*, 12(7-8):661–669, July 2012. ISSN 1533-7146. URL `http://dl.acm.org/citation.cfm?id=2231016.2231026`.

[14] Heeralal Janwa and Oscar Moreno. McEliece public key cryptosystems using algebraic-geometric codes. *Des. Codes Cryptography*, 8(3):293–307, 1996.

[15] Julia Kempe and Aner Shalev. The hidden subgroup problem and permutation group theory. In *SODA '05: Proceedings of the sixteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 1118–1125, 2005.

[16] Julia Kempe, Laszlo Pyber, and Aner Shalev. Permutation groups, minimal degrees and quantum computing. *Groups, Geometry, and Dynamics*, 1(4):553–584, 2007. URL `http://xxx.lanl.gov/abs/quant-ph/0607204`.

[17] Chris Lomont. The hidden subgroup problem - review and open problems, 2004. URL `arXiv.org:quant-ph/0411037`.

[18] Eugene M. Luks. Isomorphism of graphs of bounded valence can be tested in polynomial time. *J. Computer and System Sciences*, 25(1):42–65, 1982.

[19] F.J. MacWilliams and N.J.A. Sloane. *The theory of error correcting codes*. North-Holland mathematical library. North-Holland Pub. Co., 1978. ISBN 9780444851932.

[20] R.J. McEliece. A public-key cryptosystem based on algebraic coding theory. *JPL DSN Progress Report*, pages 114–116, 1978.

[21] Lorenz Minder and Amin Shokrollahi. Cryptanalysis of the Sidelnikov cryptosystem. In *EUROCRYPT'07*, pages 347–360, 2007.

[22] Cristopher Moore, Alexander Russell, and Leonard J. Schulman. The symmetric group defies strong quantum Fourier sampling. *SIAM Journal of Computing*, 37:1842–1864, 2008.

[23] Cristopher Moore, Alexander Russell, and Piotr Śniady. On the impossibility of a quantum sieve algorithm for graph isomorphism. *SIAM J. Computing*, 39(6):2377–2396, 2010.

[24] Harald Niederreiter. Knapsack-type cryptosystems and algebraic coding theory. *Problems of Control and Information Theory. Problemy Upravlenija i Teorii Informacii*, 15(2):159–166, 1986.

[25] Edoardo Persichetti. Compact McEliece keys based on quasi-dyadic Srivastava codes. Cryptology ePrint Archive, Report 2011/179, 2011. `http://eprint.iacr.org/`.

[26] E. Petrank and R.M. Roth. Is code equivalence easy to decide? *IEEE Transactions on Information Theory*, 43(5):1602 – 1604, 1997. doi: 10.1109/18.623157.

[27] Yuval Roichman. Upper bound on the characters of the symmetric groups. *Invent. Math.*, 125 (3):451–485, 1996.

[28] Nicolas Sendrier. Finding the permutation between equivalent linear codes: the support splitting algorithm. *IEEE Transactions on Information Theory*, 46(4):1193 – 1203, 2000.

[29] Nicolas Sendrier and Dimitris E. Simos. The hardness of code equivalence over $\mathbb{F}_q$ and its application to code-based cryptography. In *Proceedings of the Fifth International Conference on Post-Quantum Cryptography (PQCrypto)*, 2013.

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

[31] V. M. Sidelnikov. A public-key cryptosystem based on binary Reed-Muller codes. *Discrete Mathematics and Applications*, 4(3):191–207, 1994.

[32] V. M. Sidelnikov and S. O. Shestakov. On insecurity of cryptosystems based on generalized Reed-Solomon codes. *Discrete Mathematics and Applications*, 2(4):439–444, 1992.

[33] Daniel R. Simon. On the power of quantum computation. *SIAM J. Comput.*, 26(5):1474–1483, 1997.

[34] Daniel A. Spielman. Faster isomorphism testing of strongly regular graphs. In *Proc. 28th Symposium on Theory of Computing*, pages 576–584, 1996.

[35] Henning Stichtenoth. On automorphisms of geometric Goppa codes. *Journal of Algebra*, 130: 113–121, 1990.

[36] J.H van Lint. *Introduction to coding theory*. Springer-Verlag, 2nd edition, 1992.

[37] Umesh Vazirani. Two challenges in quantum algorithms. TALK: Quantum Information Processing Seminar at MIT, October 2007. URL `https://calendar.csail.mit.edu/events/1626`.