McEliece and Niederreiter Cryptosystems that Resist Quantum Fourier Sampling Attacks

Hang Dinh
Cristopher Moore
Alexander Russell

SFI WORKING PAPER: 2011-06-021

SANTA FE INSTITUTE

McEliece and Niederreiter Cryptosystems That Resist Quantum Fourier Sampling Attacks

Hang Dinh
Indiana University South Bend
hdinh@cs.iusb.edu

Cristopher Moore
University of New Mexico
and Santa Fe Institute
moore@cs.unm.edu

Alexander Russell
University of Connecticut
acr@cse.uconn.edu

June 9, 2011

Abstract

Quantum computers can break the RSA, El Gamal, and elliptic curve public-key cryptosystems, as they can efficiently factor integers and extract discrete logarithms. This motivates the development of *post-quantum* cryptosystems: classical cryptosystems that can be implemented with today's computers, that will remain secure even in the presence of quantum attacks.

In this article we show that the McEliece cryptosystem over *rational Goppa codes* and the Niederreiter cryptosystem over *classical Goppa codes* resist precisely the attacks to which the RSA and El Gamal cryptosystems are vulnerable—namely, those based on generating and measuring coset states. This eliminates the approach of strong Fourier sampling on which almost all known exponential speedups by quantum algorithms are based. Specifically, we show that the natural case of the Hidden Subgroup Problem to which McEliece-type cryptosystems reduce cannot be solved by strong Fourier sampling, or by any measurement of a coset state. To do this, we extend recent negative results on quantum algorithms for Graph Isomorphism to subgroups of the automorphism groups of linear codes.

This gives the first rigorous results on the security of the McEliece-type cryptosystems in the face of quantum adversaries, strengthening their candidacy for post-quantum cryptography. We also strengthen some results of Kempe, Pyber, and Shalev on the Hidden Subgroup Problem in S_n.

1 Introduction

If and when quantum computers are built, common public-key cryptosystems such as RSA, El Gamal, and elliptic curve cryptography will no longer be secure. Given that fact, the susceptibility or resistance of other well-studied public-key cryptosystems to quantum attacks is of fundamental interest. We present evidence for the strength of McEliece-type cryptosystems against quantum attacks, demonstrating that the quantum Fourier sampling attacks that cripple RSA and El Gamal do not apply to the McEliece or Niederreiter cryptosystems as long as the underlying code satisfies certain algebraic properties. While there are known classical attacks on these systems for the case of rational Goppa codes, our results also apply to the Niederreiter cryptosystem with classical Goppa codes, which to our knowledge is still believed to be classically secure. While our results do not rule out other quantum (or classical) attacks, they do demonstrate security precisely against the types of quantum algorithms that have proven so powerful for number theory. We also strengthen some results of Kempe et al. [9] on subgroups of S_n reconstructible by Fourier sampling.

McEliece-type cryptosystems. The McEliece cryptosystem is a public-key cryptosystem proposed by McEliece in 1978 [13], conventionally built over Goppa codes. A dual variant of the system, proposed by Niederreiter [16], can provide slightly improved efficiency with equivalent security [10]. This dual system can additionally be used to construct a digital signature scheme [3], a shortcoming of the original system.

There are two basic types of attacks known against the McEliece-type cryptosystems: decoding attacks, and direct attacks on the private key. The former appears challenging, considering that the general decoding problem is NP-hard; indeed, historical confidence in the security of the McEliece system relies on the idea that this hardness can be retained for scrambled version of specific codes. Decoding attacks remain challenging quantumly for quantum computers are not known to be able to efficiently solve NP-hard problems. The latter—direct attacks on the key—can be successful on certain classes of linear codes, and is our focus. In a McEliece-type cryptosystem, the private key of a user Alice consists of three matrices: a $k \times n$ matrix M over a finite field $\mathbb{F}_{q^\ell}$, a $k \times k$ invertible matrix S over the field $\mathbb{F}_q$, and an $n \times n$ permutation matrix P. In the McEliece version, M is a generator matrix of a q-ary $[n,k]$-linear code (hence, $\ell = 1$), while in Niederreiter's dual system, M is a parity check matrix of a q-ary linear code of length n. The matrices S and P are selected randomly. Alice's public key consists of the matrix $M^* = SMP$. An adversary may attack the private key, attempting to recover the secret row "scrambler" S and the secret permutation P from M^* and M, assuming he already knew M.[1] As pointed out in [4], it crucial to keep S and P secret for the security of the McEliece system.

The security of these McEliece-type systems have received considerable attention in the literature, often focusing on particular choices for the underlying codes. Various classes of Goppa codes have received the greatest attention: along these lines, Sidelnokov and Shestakov's attack [23] can efficiently compute the matrices S and MP from the public matrix $M^* = SMP$ if the underlying code is a generalized Reed-Solomon code.[2] While this attack can reveal the structure of an alternative code, it does not reveal the secret permutation. An attack in which the secret permutation is revealed was proposed by Loidreau and Sendrier [11], using the Support Splitting Algorithm [21]. However, this attack only works with a very limited subclass of classical binary Goppa codes, namely those with a binary Goppa polynomial.

Although the McEliece-type cryptosystems are efficient and still considered classically secure, at least with classical binary Goppa codes [4], they are rarely used in practice because of its comparatively large public key (see remark 8.33 in [14]). The discovery of successful quantum attacks on RSA and El Gamal, however, has changed the landscape: as suggested by Ryan [20] and Bernstein et al. [2], if "post-quantum" security guarantees can be made for the McEliece cryptosystem, this may compensate for its comparatively expensive computational demands.

Quantum Fourier sampling. 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 [22] and Simon's algorithm [24]. 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

[1]Recovering the secret scrambler and the secret permutation is different from the Code Equivalence problem. The former finds a transformation between two equivalent codes, while the latter decides whether two linear codes are equivalent.

[2]We remark that the class of generalized Reed-Solomon codes is essentially equal to the class of rational Goppa codes.

Fourier basis $\{|\rho, i, j\rangle\}$ for the space $\mathbb{C}[G]$, where ρ is an irrep[3] 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 [6, 15, 7] have suggested that the approach is inherently limited when the underlying groups are rich enough. In particular, Moore, Russell, and Schulman [15] 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. [7] strengthened these results, demonstrating that even entangled measurements on $o(\log n!)$ coset states yield essentially no information.

Kempe and Shalev [8] 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.[4] 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 [9].

Our contributions. To state our results, we say that a subgroup $H < G$ is *indistinguishable by strong Fourier sampling* 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, Russell, and Schulman [15], we first develop a general framework, formalized in Theorem 4, to determine indistinguishability of a subgroup by strong Fourier sampling. We emphasize that their results 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 a class of semi-direct products $(\mathsf{GL}_k(\mathbb{F}_q) \times S_n) \wr \mathbb{Z}_2$, bounding the distinguishability for the HSP corresponding to the private-key attack on a McEliece-type cryptosystem, i.e., the problem of determining a secret scrambler S and a secret permutation P from $M^* = SMP$ and M. Our bound, given in Corollary 9 of Theorem 8, depends on the column rank[5] 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 4.2 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 general, our result indicates that McEliece-type cryptosystems resist known attacks based on 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 $e^{o(n)}$. In particular, generator matrices of rational Goppa codes and canonical parity check matrices of classical Goppa codes have good values for these quantities (see Lemma 10). The result is most interesting for classical Goppa codes, which are considered classically secure; the McEliece system over *rational* Goppa

[3] Throughout the paper, we write "irrep" as short for "irreducible representation."
[4] The minimal degree of a permutation group H is the minimal number of points moved by a non-identity element of H.
[5] The column rank of M is understood as over the field F_{q^ℓ}. Recall that the entries of the matrix M are in F_{q^ℓ}.

codes is subject to the Sidelnokov-Shestakov [23] attack.

While our main application is the security of the McEliece cryptosystem, we show in addition that our general framework is applicable to other classes of groups with simpler structure, including the symmetric group and the finite general linear group[6] $\mathsf{GL}_2(\mathbb{F}_q)$. For the symmetric group, we extend the results of [15] 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. [9], which apply only to weak Fourier sampling.

Remark Our results show that the natural reduction of McEliece to a hidden subgroup problem yields negligible information about the secret key. Thus they rule out the direct analogue of the quantum attack that breaks, for example, RSA. Of course, our results do not rule out other quantum (or classical) attacks. Neither do they establish that a quantum algorithm for the McEliece cryptosystem would violate a natural hardness assumption, as do recent lattice cryptosystem constructions whose hardness is based on the Learning With Errors problem (e.g. Regev [18]). Nevertheless, they indicate that any such algorithm would have to involve significant new ideas beyond than those that have been proposed so far.

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 [15], 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 [15] 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 in Lemma 16 that any irrep outside SMALL has large dimension and therefore small normalized characters.

For the case $G = (\mathsf{GL}_k(\mathbb{F}_q) \times S_n) \wr \mathbb{Z}_2$ corresponding to McEliece-type cryptosystems, 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 a classical Goppa code, $\mathrm{Aut}(M)$ lies inside the automorphism group of a rational Goppa code, which can be controlled using Stichtenoth's Theorem [25].

[6]The case of $\mathsf{GL}_2(\mathbb{F}_q)$ is put in the Appendix for lack of space.

2 Hidden Subgroup Attacks on McEliece Cryptosystems

As mentioned in the Introduction, we consider an attack attempting to recover the secret scrambler S and permutation P from M and M^*. We frame the problem such an attacker needs to solve as follows:

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).

In the case where the matrix M is a generator matrix of a linear code over $\mathbb{F}_q$, the decision version of this problem is known as the CODE EQUIVALENCE problem, which is at least as hard as GRAPH ISOMORPHISM, although it is unlikely to be NP-complete [17]. 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:

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 algoroithms 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 $b\in\mathbb{Z}_2$,

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

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. 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,

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

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 2. *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 3. *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 $G \wr \mathbb{Z}_2$:*

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

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. The *Quantum Fourier transform* over G is the unitary operator, denoted F_G, that transforms a vector in $\mathbb{C}[G]$ from the point-mass 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. If X is a subgroup of G, then Π_X^ρ is a projection operator. That is, $(\Pi_X^\rho)^\dagger = \Pi_X^\rho$ and $(\Pi_X^\rho)^2 = \Pi_X^\rho$.

Quantum Fourier Sampling (QFS) is a standard procedure based on the Quantum Fourier Transform to solve the Hidden Subgroup Problem (HSP) (see [12] 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

$$P_{\alpha H}(\rho) = \|\widehat{\alpha H}(\rho)\|_F^2 = \frac{d_\rho |H|}{|G|} \operatorname{Tr}\left((\Pi_{\alpha H}^\rho)^\dagger \Pi_{\alpha H}^\rho\right) = \frac{d_\rho |H|}{|G|} \operatorname{Tr}\left(\Pi_H^\rho\right)$$

where $\operatorname{Tr}(A)$ denotes the trace of a matrix A, and $\|A\|_F := \sqrt{\operatorname{Tr}(A^\dagger A)}$ is the Frobenius norm of A. The last equality is due to the fact that $\Pi_{\alpha H}^\rho = \rho(\alpha)\Pi_H^\rho$ and that Π_H^ρ is an orthogonal projector.

Since there is no point in measuring the rows [6], we are only concerned with measuring the columns. As pointed out in [15], 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

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

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,

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

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 $\operatorname{Tr}\left(\Pi^\rho_H\right) = \operatorname{Tr}\left(\Pi^\rho_{H^g}\right)$, the probability distributions obtained by QFS for recovering the hidden subgroup H^g are

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

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. 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}:

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

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 $\mathscr{D}_H \leq \log^{-\omega(1)}|G|$.

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 $\varepsilon > 0$,

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

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 [8]. 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

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

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

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 $|\widehat{G}|$.

The detailed proof of this theorem is relegated to the Appendix A.

4 Applications of the Main Theorem

In this section, we present applications of Theorem 4 to analyze strong Fourier sampling over certain non-abelian groups, including the symmetric group and the wreath product corresponding to the McEliece-type cryptosystems. Another application to the HSP over the groups $\mathsf{GL}_2(\mathbb{F}_q)$ appears in Appendix C.

4.1 Strong Fourier Sampling over S_n

We focus now on the case where G is the symmetric group S_n. 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 columns. 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 [15], we shall apply Roichman's upper bound [19] on normalized characters:

Theorem 5 (Roichman's Theorem [19]). *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,*

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

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 $\frac{\lambda_1}{n} \ge 1-c$ or $\frac{\lambda_1'}{n} \ge 1-c$, 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 $\alpha > 0$,

$$\left|\frac{\chi_\lambda(\pi)}{d_\lambda}\right| \le \mathrm{e}^{-\alpha\cdot \mathrm{supp}(\pi)}. \tag{4}$$

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 [15].

Lemma 6 (Lemma 6.2 in [15]). *Let $p(n)$ denote the number of integer partitions of n. Then $|\Lambda_c| \le 2cn \cdot p(cn)$, 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 $p(n)$ [5, pg. 45]: $p(n) \approx \mathrm{e}^{\pi\sqrt{2n/3}}/(4\sqrt{3}n) = \mathrm{e}^{O(\sqrt{n})}n^{-1}$ as $n \to \infty$.

Now we are ready to prove the main result of this section, an application of Theorem 4.

Theorem 7. *Let H be a nontrivial subgroup of S_n with minimal degree m, i.e., $m = \min_{\pi\in H\setminus\{1\}} \mathrm{supp}(\pi)$. Then for sufficiently large n, $\mathscr{D}_H \le O(|H|^2\mathrm{e}^{-\alpha m})$.*

Proof. Let $2c < d < 1/2$ be constants. We will apply Theorem 4 by setting $S = \Lambda_c$ and $D = n^{dn}$. By Lemma 6, we have $d_S \le n^{cn}$. 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 (4), we have $\overline{\chi}_\mu(H) \le \mathrm{e}^{-\alpha m}$ for all $\mu \in \widehat{S_n}\setminus S$. Hence, $\overline{\chi}_{\overline{S}}(H) \le \mathrm{e}^{-\alpha m}$. To bound the second term in the upper bound of Theorem 4, as $\Delta \le |S|$, it suffices to bound:

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

for sufficiently large n, so long as $\gamma < d - 2c$. Now bounding the last term in the upper bound of Theorem 4: Since $|\overline{L_D}| \le |\widehat{S_n}| = p(n)$ and $n! > n^n\mathrm{e}^{-n}$ by Stirling's approximation,

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

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

□

Theorem 7 generalizes Moore, Russell, and Schulman's result [15] 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. [9], observe that since $\log|S_n| = \Theta(n\log n)$, the subgroup H is indistinguishable by strong Fourier sampling if $|H|^2\mathrm{e}^{-\alpha m} \le (n\log n)^{-\omega(1)}$ or, equivalently, if $m \ge (2/\alpha)\log|H| + \omega(\log n)$.

4.2 Applications to McEliece-type Cryptosystems

Our main application of Theorem 4 is to show the limitations of strong Fourier sampling in attacking the McEliece-type cryptosystems. Throughout this section, we fix parameters n,k,q of a McEliece-type cryptosystem, and fix the underlying $k \times n$ matrix M of the system. Here, M can be a generator matrix or a parity check matrix of the q-ary linear code used in the 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 canonical quantum attack against this McEliece cryptosystem involves 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

$$K = ((H_0, s^{-1}H_0 s), 0) \cup ((H_0 s, s^{-1}H_0), 1) \tag{5}$$

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

$$H_0 = \left\{ (A,P) \in \mathsf{GL}_k(\mathbb{F}_q) \times S_n \mid A^{-1}MP = M \right\} . \tag{6}$$

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

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

Note that $\mathrm{Aut}(M)$ is a subgroup of the symmetric group S_n and each element $(A,P) \in H_0$ must have $P \in \mathrm{Aut}(M)$. This allows us to control the maximal normalized characters on K through the minimal degree of $\mathrm{Aut}(M)$. Then applying Theorem 4, we show that

Theorem 8. *Assume $q^{k^2} \le n^{an}$ 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 (5) has $\mathscr{D}_K \le O(|K|^2 \mathrm{e}^{-\delta m})$, where $\delta > 0$ is a constant.*

The proof of Theorem 8 follows the technical ideas discussed in the Introduction. The details appear in Appendix B. As $q^{k^2} \le n^{an}$, 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 $|K|^2 \mathrm{e}^{-\delta m} \le (n \log n)^{-\omega(1)}$. 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 $\mathrm{Fix}(M)$ is the set of scramblers fixing M, i.e., $\mathrm{Fix}(M) \stackrel{\text{def}}{=} \left\{ S \in \mathsf{GL}_k(\mathbb{F}_q) \mid SM = M \right\}$. To bound the size of $\mathrm{Fix}(M)$, we record an easy fact which can be obtained by the orbit-stabilizer formula:

Fact. Let r be the column rank of M. Then $|\mathrm{Fix}(M)| \le (q^{\ell k} - q^{\ell r})(q^{\ell k} - q^{\ell(r+1)}) \cdots (q^{\ell k} - q^{\ell(k-1)}) \le q^{\ell k(k-r)}$.

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

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

□

Corollary 9. *Assume $q^{k^2} \le n^{0.2n}$ 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 (5) has $\mathscr{D}_K \le |\mathrm{Aut}(M)|^4 q^{4\ell k(k-r)} \mathrm{e}^{-\Omega(n)}$. In particular, the subgroup K is indistinguishable if, further, $|\mathrm{Aut}(M)| \le \mathrm{e}^{o(n)}$ and $r \ge k - o(\sqrt{n})/\ell$.*

Application to the McEliece cryptosystem. Consider a McEliece cryptosystem using a q-ary linear $[n,k]$-code C, with parameters satisfying $q^{k^2} \le n^{0.2n}$. Since the automorphism group of the code C equals the automorphism group of its generator matrix, we can conclude that this McEliece cryptosystem resists the standard quantum Fourier sampling attack if the code C is *(i) well-scrambled*, i.e., it has a generator matrix of rank at least $k - o(\sqrt{n})$, and is *(ii) well-permuted*, i.e., its automorphism group has minimal degree at least $\Omega(n)$ and has size at most $e^{o(n)}$. Recall that in terms of security, the Niederreiter system using $(n-k) \times n$ parity check matrices over $\mathbb{F}_q$ of the same code C is equivalent to the McEliece system using the code C [10].

Application to Goppa codes. We would like to point out that if M is a generator matrix of a rational Goppa code or a canonical parity check matrix of a classical Goppa code, it will give good bounds in Corollary 9. Specifically, we consider a matrix M over a finite field $\mathbb{F}_{q^\ell} \supset \mathbb{F}_q$ of the following form:

$$M = \begin{pmatrix} v_1 f_1(\alpha_1) & v_2 f_1(\alpha_2) & \cdots & v_n f_1(\alpha_n) \\ v_1 f_2(\alpha_1) & v_2 f_2(\alpha_2) & \cdots & v_n f_2(\alpha_n) \\ \vdots & \vdots & \ddots & \vdots \\ v_1 f_k(\alpha_1) & v_2 f_k(\alpha_2) & \cdots & v_n f_k(\alpha_n) \end{pmatrix} \tag{7}$$

where $v_1, \ldots, v_n$ are nonzero elements in the field $\mathbb{F}_{q^\ell}$, $(\alpha_1, \ldots, \alpha_n)$ is a list of distinct points in the projective line $\mathbb{F}_{q^\ell} \cup \{\infty\}$, and $f_1, \ldots, f_k$ are $\mathbb{F}_{q^\ell}$-linearly independent polynomials in $\mathbb{F}_{q^\ell}[X]$ of degree less than k (by convention, $f_i(\infty)$ is the X^{k-1}-coefficient of $f_i(X)$). Note that such a matrix M is a generator matrix of a rational Goppa $[n,k]$-code over the field $\mathbb{F}_{q^\ell}$, and is also a parity check matrix of a classical Goppa $[n, \ge n - \ell k]$-code over $\mathbb{F}_q$. To apply Corollary 9, we show the following properties of the matrix M:

Lemma 10. *The matrix M in the form of* (7) *has full rank (i.e., its column rank equals k), and* $\mathrm{Aut}(M)$ *has minimal degree at least $n-2$, and* $|\mathrm{Aut}(M)| \le q^{3\ell}$.

Proof. We can show that M has full rank directly by decomposing M as $M = AVD$, where $A = (a_{ij})$ is an $k \times k$ invertible matrix with entry a_{ij} being the the X^{j-1}-coefficient of polynomial $f_i(X)$; V is a $k \times n$ Vandermonde matrix with (i,j)-entry being α_j^{i-1}; and D is an $n \times n$ diagonal matrix with v_i in the (i,i)-entry. Then the rank of M equals the rank of the Vandermonde matrix V, which has full rank.

Now we can view M as a generator matrix of a rational Goppa $[n,k]$-code R over the field $\mathbb{F}_{q^\ell}$. Then we have $\mathrm{Aut}(M) \subset \mathrm{Aut}(R)$, where $\mathrm{Aut}(R) = \{P \in S_n \mid SMP = M \text{ for some } S \in \mathsf{GL}_k(\mathbb{F}_{q^\ell})\}$ is the automorphism group of the code R. By Stichtenoth's Theorem [25] (see Appendix D), $\mathrm{Aut}(R)$ is isomorphic to a subgroup of the projective linear group $\mathsf{PGL}_2(\mathbb{F}_{q^\ell})$. Thus, $|\mathrm{Aut}(M)| \le |\mathrm{Aut}(R)| \le |\mathsf{PGL}_2(\mathbb{F}_{q^\ell})| \le q^{3\ell}$.

To show that the minimal degree of $\mathrm{Aut}(M)$ is at least $n-2$, we view $\mathrm{Aut}(M) \subset \mathsf{PGL}_2(\mathbb{F}_{q^\ell})$, and observe that any transformation in $\mathsf{PGL}_2(\mathbb{F}_{q^\ell})$ that fixes at least three distinct projective lines must be the identity. □

Hence, classical Goppa codes or rational Goppa codes are good choices for the security of McEliece-type cryptosystems against standard quantum Fourier sampling attacks. Since the rational Goppa codes are broken (classically) by the Sidelnokov-Shestakov [23] structural attack, we shall focus on the classical Goppa codes, which remain secure given suitable choice of parameters.

Application to Neiderreiter systems with classical Goppa codes. Consider a classical q-ary Goppa code C constructed by a support list of distinct points $\alpha_1, \ldots, \alpha_n \in \mathbb{F}_{q^\ell}$ and a Goppa polynomial $g(X) \in \mathbb{F}_{q^\ell}[X]$ of degree k. This code has dimension $k' \ge n - \ell k$. More importantly, it has $k \times n$ parity check matrices in the form of (7) in which $v_j = 1/g(\alpha_j)$ (see [27]), we refer to those matrices as *canonical parity check matrices* of the classical Goppa code C. By Corollary 9 and Lemma 10, *the Niederreiter cryptosystem using* $k \times n$ *canonical parity check matrices of this code C resists the known quantum attack, provided* $q^{k^2} \leq n^{0.2n}$ *and* $q^{3\ell} \leq e^{o(n)}$. As pointed out in [4], this Niederreiter system is secure under the Sidelnokov-Shestakov attack. We remark, however, that the security of this Niederreiter cryptosystem may *not* be equivalent to that of the McEliece cryptosystem using the same code C, since the equivalence showed in [10] only applies to the Niederreiter cryptosystem using a parity check matrix over the subfield $\mathbb{F}_q$.

Setting the parameters. We discuss the parameters for classical Goppa codes that meet our security requirement. Traditionally, the code length is chosen as $n = q^\ell$, then our parameter setting requires only one constraint, $k^2 \leq 0.2n\ell$, which imposes that the code C must have large dimension, i.e., $k' \geq n - \ell k \geq n - \sqrt{0.2n}(\log_q n)^{3/2}$.

Now we compare our parameter setting with practical parameters suggestion. In most McEliece cyptosystems considered in practice, classical binary Goppa codes are used, that is $q = 2$ and $n = 2^\ell$. The code is also designed so that it has dimension $k' = n - \ell k$ and minimal distance $d \geq 2t + 1$, where $t \ll n$ is a predetermined parameter indicating the number of errors the code can correct. For those systems, the original parameters suggested by McEliece were $(n = 1024, k' \geq 524, t = 50)$, which would meet our requirement as long as the dimension k' is chosen to be slightly larger ($k' \geq 572$). The parameters $(n = 1024, k' = 524, t = 50)$, which can be broken in just 7 days by a cluster of 200 CPUs under Bernstein et al.'s attack [2], clearly do not meet our requirement. An optimal choice of parameters for the Goppa code which maximizes the adversary's work factor was recommended to be $(n = 1024, k' \geq 644, t = 38)$ (see Note 8.32 in [14]). Bernstein et al. [2] suggested two other sets of parameters, $(n = 2048, k' = 1751, t = 27)$ and $(n = 1632, k' = 1269, t = 34)$, that achieve the standard security against all known (classical) attacks. All of these parameters meet our requirement. Well, of course, these parameters were recommended for the orginal McEliece, or for the equivalent Neiderreiter system that uses parity check matrices over the subfield $\mathbb{F}_2$ with $n - k' = \ell k$ rows. However, if we view each element in $\mathbb{F}_{q^\ell}$ as a vector of dimension ℓ over the subfield $\mathbb{F}_q$, then a $k \times n$ canonical parity check matrix over $\mathbb{F}_{q^\ell}$ can be viewed as a $\ell k \times n$ parity check matrix over $\mathbb{F}_q$.

References

[1] Daniel J. Bernstein. List decoding for binary Goppa codes, 2008. Preprint.

[2] Daniel J. Bernstein, Tanja Lange, and Christiane Peters. Attacking and defending the McEliece cryptosystem. In *PQCrypto '08: Proceedings of the 2nd International Workshop on Post-Quantum Cryptography*, pages 31–46, Berlin, Heidelberg, 2008. Springer-Verlag. ISBN 978-3-540-88402-6.

[3] Nicolas Courtois, Matthieu Finiasz, and Nicolas Sendrier. How to achieve a mceliece-based digital signature scheme. In *Proceedings of the 7th International Conference on the Theory and Application of Cryptology and Information Security: Advances in Cryptology*, ASIACRYPT '01, pages 157–174, London, UK, 2001. Springer-Verlag.

[4] D. Engelbert, R. Overbeck, and A. Schmidt. A summary of McEliece-type cryptosystems and their security. *J. Math. Crypt.*, 1:151199, 2007.

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

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

[7] 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.

[8] 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.

[9] 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`.

[10] Yuan Xing Li, Robert H. Deng, and Xin Mei Wang. On the equivalence of McElieces and Niederreiters public-key cryptosystems. *IEEE Transactions on Information Theory*, 40(1):271273, 1994.

[11] Pierre Loidreau and Nicolas Sendrier. Weak keys in the McEliece public-key cryptosystem. *IEEE Transactions on Information Theory*, 47(3):1207–1212, 2001.

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

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

[14] A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone. *Handbook of applied cryptography*. CRC Press, 1996.

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

[16] 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.

[17] 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.

[18] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In *STOC '05: Proceedings of the thirty-seventh annual ACM symposium on Theory of computing*, pages 84–93, 2005.

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

[20] John A. Ryan. Excluding some weak keys in the McEliece cryptosystem. In *Proceedings of the 8th IEEE Africon*, pages 1–5, 2007.

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

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

[23] 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.

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

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

[26] Henning Stichtenoth. *Algebraic Function Fields and Codes*. Springer, 2nd edition, 2008.

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

Appendices

A Proof for the Main Theorem

We now present the proof for the main theorem (Theorem 4) in details.

A.1 Proof sketch

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

$$\begin{aligned} E_H(\rho) = \mathbb{E}_g\left[\left(\sum_{\mathbf{b}\in B_\rho}\left|P_{H^g}(\mathbf{b} \mid \rho) - \frac{1}{d_\rho}\right|\right)^2\right] &\leq \mathbb{E}_g\left[d_\rho \sum_{\mathbf{b}\in B_\rho}\left(P_{H^g}(\mathbf{b} \mid \rho) - \frac{1}{d_\rho}\right)^2\right] \quad \text{(by Cauchy-Schwarz)} \\ &= d_\rho \sum_{\mathbf{b}\in B_\rho} \mathrm{Var}_g[P_{H^g}(\mathbf{b} \mid \rho)] \qquad \left(\text{since } \mathbb{E}_g[P_{H^g}(\mathbf{b} \mid \rho)] = \frac{1}{d_\rho}\right) \\ &= \frac{d_\rho}{\mathrm{Tr}(\Pi_H^\rho)^2} \sum_{\mathbf{b}\in B_\rho} \mathrm{Var}_g\left[\|\Pi_{H^g}^\rho \mathbf{b}\|^2\right] . \end{aligned} \tag{8}$$

The equation $\mathbb{E}_g[P_{H^g}(\mathbf{b} \mid \rho)] = 1/d_\rho$ (Proposition 14 in Appendix A) can be shown using *Schur's lemma*.

From (8), 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 11 below generalizes the bound given in Lemma 4.3 of [15], which only applies to subgroups H of order 2.

Lemma 11. *(*Decoupling Lemma*) Let ρ be an irrep of G. Then for any vector $\mathbf{b} \in V_\rho$,*

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

Back to our goal of bounding $E_H(\rho)$ using the bound in Lemma 11, 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 [15] (its proof is also given in the Appendix):

Lemma 12. *For any irrep σ, we have $\sum_{\mathbf{b}\in B_\rho}\left\|\Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2 \leq d_\sigma^2$.*

The method discussed above for bounding $E_H(\rho)$ is culminated into Lemma 13 below.

Lemma 13. *Let $\rho \in \widehat{G}$ be arbitrary and $S \subset \widehat{G}$ be any subset of irreps that does not contain ρ. Then*

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

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 13 for all irreps ρ of large dimension, we can prove our general main theorem straightforwardly.

A.2 Detailed proofs

Proposition 14. *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,

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

which implies that

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

□

A.2.1 Proof of Decoupling Lemma

Proof of Lemma 11. Fix a vector $\mathbf{b} \in V_\rho$. To simplify notations, we shall write Π_g as shorthand for $\Pi^\rho_{H^g}$, and write $g\mathbf{b}$ for $\rho(g)\mathbf{b}$. For any $g \in G$, we have

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

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 [15], one can express the second moment of the inner product $\langle \mathbf{b}, g^{-1}hg\mathbf{b} \rangle$ 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

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

It follows that

$$\begin{aligned} \mathrm{Var}_g\left[\|\Pi_{H^g}^{\rho}\mathbf{b}\|^2\right] &\leq \frac{|H|-1}{|H|^2} \sum_{h\in H\setminus\{1\}} \mathbb{E}_g\left[\left|\langle \mathbf{b}, g^{-1}hg\mathbf{b} \rangle\right|^2\right] \\ &\leq \mathbb{E}_{h\in H\setminus\{1\}}\left[\sum_{\sigma \in I(\rho\otimes\rho^*)} \frac{\chi_\sigma(h)}{d_\sigma} \left\| \Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*) \right\|^2\right] \\ &\leq \sum_{\sigma \in I(\rho\otimes\rho^*)} \overline{\chi}_\sigma(H) \left\| \Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*) \right\|^2 . \end{aligned}$$

□

A.2.2 Proof of Lemma 12

Proof of Lemma 12. Let L_σ be the subspace of $\rho \otimes \rho^*$ consisting of all copies of σ. Since B_ρ is orthonormal, the vectors $\{\mathbf{b} \otimes \mathbf{b}^* \mid \mathbf{b} \in B_\rho\}$ 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,

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

□

A.2.3 Proof of Lemma 13

Proof of Lemma 13. Combining Inequality (8) and Lemmas 11 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),

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

For the second group (small dimension),

$$\begin{aligned}\sum_{\sigma\in S\cap I(\rho\otimes\rho^*)}\overline{\chi}_\sigma(H)\sum_{\mathbf{b}\in B_\rho}\left\|\Pi_\sigma^{\rho\otimes\rho^*}(\mathbf{b}\otimes\mathbf{b}^*)\right\|^2 &\le \sum_{\sigma\in S\cap I(\rho\otimes\rho^*)}\overline{\chi}_\sigma(H)d_\sigma^2 &&\text{(by Lemma 12)}\\ &\le \sum_{\sigma\in S\cap I(\rho\otimes\rho^*)}d_\sigma^2 &&(\text{since } \overline{\chi}_\sigma(H)\le 1)\\ &\le |S\cap I(\rho\otimes\rho^*)|\,d_S^2\,.\end{aligned}$$

Summing the last bounds for the two groups yields

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

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

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

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

A.2.4 Proof of Main Theorem

Proof of Theorem 4: For any $\rho\in L$, since $d_\rho\ge D>d_S^2$, we must have $\rho\notin S$. By Lemma 13,

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

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

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

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

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

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

$$\mathrm{Pr}_\rho[\rho\notin L]=\sum_{\rho\notin L}P(\rho)\le\frac{|\overline{L}|D^2|H|}{|G|}\le\frac{|\overline{L}|D^2|H|^2}{|G|}\,.$$

□

B Strong Fourier Sampling over $(\mathsf{GL}_k(\mathbb{F}_q) \times S_n) \wr \mathbb{Z}_2$

This section devotes to the proof of Theorem 8 which establishes the limitation of strong Fourier sampling in breaking the McEliece cryptosystem. The goal is to bound the distinguishability of the subgroup K defined in (5) of the wreath product $(\mathsf{GL}_k(\mathbb{F}_q) \times S_n) \wr \mathbb{Z}_2$.

B.1 Normalized Characters for $G \wr \mathbb{Z}_2$

Firstly, we consider quantum Fourier sampling over the wreath product $G \wr \mathbb{Z}_2$, for a general group G, with a hidden 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 $s \in G$. Again, the first thing we need to understand is the maximal normalized characters on K. Recall that all irreducible characters of $G \wr \mathbb{Z}_2$ are constructed in the following ways:

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 15. *For non-isormorphic irreps $\rho, \sigma \in \widehat{G}$,*

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

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

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

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

B.2 Normalized Characters for $(\mathsf{GL}_k(\mathbb{F}_q) \times S_n) \wr \mathbb{Z}_2$

Recall that for the case of attacking McEliece cryptosystem, 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 $\mathsf{GL}_k(\mathbb{F}_q) \times S_n$. Since $\overline{\chi}_{\tau\times\lambda}(S_\pi, \pi) = \overline{\chi}_\tau(S_\pi)\overline{\chi}_\lambda(\pi)$ and $\overline{\chi}_\tau(S_\pi) \leq 1$ for all $\pi \in S_n$, we have

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

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 16. *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 16. 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 $c_1, \ldots, c_n$, in which c_i is the i^{th} cell from the left of the first row, for $1 \leq i \leq \lambda_1$.

The dimension of λ is determined by the *hook length formula*:

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

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 the 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 i, thus

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

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 $d_{\tilde{\lambda}} \geq 1$ gives us:

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

Then we have

$$\mathrm{Hook}(\lambda) = \mathrm{Hook}(\tilde{\lambda}) \prod_{i=1}^{\lambda_1} \mathrm{hook}(c_i) \leq (n-\lambda_1)! \prod_{i=1}^{\lambda_1} \mathrm{hook}(c_i)\,.$$

On the other hand, we have

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

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

$$\begin{aligned}
\sum_{i=1}^{\lambda_1} \frac{\lambda_i' - 1}{\lambda_1 - i + 1} &\leq \frac{\sum_{i=1}^{\lambda_1} (\lambda_i' - 1)}{\lambda_1} \left(\sum_{i=1}^{\lambda_1} \frac{1}{\lambda_1 - i + 1}\right) \\
&= \frac{n - \lambda_1}{\lambda_1} \left(\sum_{i=1}^{\lambda_1} \frac{1}{i}\right) \leq \frac{1}{c} \left(\sum_{i=1}^{\lambda_1} \frac{1}{i}\right) \qquad (\text{since } \lambda_1 > cn)\,.
\end{aligned}$$

Let r be a constant such that $1 < r/c < cn$. Bounding $1/i \leq 1$ for all $i \leq r/c$ and bounding $1/i \leq c/r$ for all $i > r/c$ yields

$$\sum_{i=1}^{\lambda_1} \frac{1}{i} \leq \frac{r}{c} + \frac{c\lambda_1}{r}\,.$$

Putting the pieces together, we get

$$\begin{aligned}
d_\lambda &\geq \frac{n!}{(n-\lambda_1)!\lambda_1! \mathrm{e}^{\lambda_1/r + r/c^2}} = \binom{n}{\lambda_1} \mathrm{e}^{-\lambda_1/r - r/c^2} \\
&\geq \left(\frac{n}{\lambda_1}\right)^{\lambda_1} \mathrm{e}^{-\lambda_1/r - r/c^2} \\
&\geq \left(\frac{\mathrm{e}^{-1/r}}{1-c}\right)^{\lambda_1} \mathrm{e}^{-r/c^2} \qquad (\text{since } \lambda_1 < (1-c)n)\,.
\end{aligned}$$

Let $0 < \delta < \ln \frac{1}{1-c}$ be a constant and choose r large enough so that $\mathrm{e}^{-1/r} \geq (1-c)\mathrm{e}^{\delta}$. Then

$$d_\lambda \geq \mathrm{e}^{\delta \lambda_1 - r/c^2} \geq \mathrm{e}^{\delta cn - r/c^2} \geq \mathrm{e}^{\beta n}\,.$$

□

Remark The lower bound in Lemma 16 is essentially tight. To see this, consider the hook of width $(1-c)n$ and of depth cn. This hook has dimension roughly equal $\binom{n}{cn}$, which is no more than $(e/c)^{cn}$.

B.3 Proof of Theorem 8

We are ready to prove Theorem 8.

Proof of Theorem 8. To apply Theorem 4, 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 4.1. 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 (4) in Section 4.1, 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 16 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 6, 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 6. The last inequality holds because $d_\tau^2 \leq \sum_{\rho \in \widehat{\mathsf{GL}_k(\mathbb{F}_q)}} d_\rho^2 = |\mathsf{GL}_k(\mathbb{F}_q)|$ for any $\tau \in \widehat{\mathsf{GL}_k(\mathbb{F}_q)}$. 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,

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

By Theorem 4, we have

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

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

□

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

C.1 Applying the Main Theorem

Now we supplement application of the main theorem (Theorem 4) with the case of the finite general linear group $G=\mathsf{GL}_2(\mathbb{F}_q)$, whose structure as well as irreps are well established [5, §5.2]. From the character table of $\mathsf{GL}_2(\mathbb{F}_q)$, which can be found in Section C.2 of Appendix C, we draw the following easy facts:

Fact. Let σ be an irrep of $\mathsf{GL}_2(\mathbb{F}_q)$. Then *(i)* For all $g\in\mathsf{GL}_2(\mathbb{F}_q)$, $|\chi_\sigma(g)|=d_\sigma$ if g is a scalar matrix, and $|\chi_\sigma(g)|\le 2$ otherwise. *(ii)* If $d_\sigma>1$, then $q-1\le d_\sigma\le q+1$.

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 theorem (Theorem 4). 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 $\left\{T(b)\mid b\in\mathbb{F}_q\right\}$, where $T(b):=\begin{pmatrix}1&b\\0&1\end{pmatrix}$.

From the easy facts above for $\mathsf{GL}(2,q)$, it is natural to choose the set S in Theorem 4 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 17. *Let ρ be an irrep of $\mathsf{GL}(2,q)$. Then at most two linear representations appear in the decomposition of $\rho\otimes\rho^*$.*

The proof for this lemma can be found in Appendix C. Then applying Theorem 4 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 18. *Let H be a subgroup of $\mathsf{GL}_2(\mathbb{F}_q)$ that does not contain any scalar matrix other than the identity. Then $\mathscr{D}_H\le 28|H|^2/q$.*

Proof of Corollary 18. 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 $\overline{\chi}_\sigma(H) \le 2/(q-1)$ for all nonlinear irrep σ, we have

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

To bound the second term in the bound of 4, we have $\Delta \le 2$ by Lemma 17 and $d_S = 1$, thus

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

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 4, $\mathscr{D}_H \le 4|H|^2\,(7/q)$. ☐

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

Examples of indistinguishable subgroups. 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 $k \ge 0$, 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 18, we have $\mathscr{D}_{H_b} \le O(p^{2-n})$.

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 $|H_{a,b}| \le p^2$. 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.

C.2 Irreducible Representations of $\mathsf{GL}_2(\mathbb{F}_q)$

In this part, we will first present a preliminary background on the structure of $\mathsf{GL}_2(\mathbb{F}_q)$ followed by description of its irreps. We refer readers to [5, §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 $\{1, \sqrt{\varepsilon}\}$ 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. Group $\mathsf{GL}_2(\mathbb{F}_q)$ has four types of conjugacy classes in Table C.2, with representatives described as follows:

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

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
no. of classes	$q-1$	$q-1$	$\frac{(q-1)(q-2)}{2}$	$\frac{q(q-1)}{2}$

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

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 C.2 below.

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	$\alpha(xy)$	$-\alpha(\xi_{x,y}^{q+1})$
$W_{\alpha,\beta}$ $(\alpha\neq\beta)$	$q+1$	$(q+1)\alpha(x)\beta(x)$	$\alpha(x)\beta(x)$	$\alpha(x)\beta(y)+\alpha(y)\beta(x)$	0
X_φ	$q-1$	$(q-1)\varphi(x)$	$-\varphi(x)$	0	$-(\varphi(\xi_{x,y})+\varphi(\xi_{x,y}^q))$

Table 2: Character table of $\mathsf{GL}_2(\mathbb{F}_q)$, where α,β are characters of $\mathbb{F}_q^*$, and φ is a character of $\mathbb{F}_{q^2}^*$ with $\varphi^q \neq \varphi$, and $d_\rho = \chi_\rho(a_1)$ is the dimension of ρ.

C.3 Proof of Lemma 17

In the remaining of this section, we devote to prove Lemma 17, which states that there are at most two linear representations appearing in the decomposition of $\rho\otimes\rho^*$, for any irrep ρ of $\mathsf{GL}_2(\mathbb{F}_q)$. Obviously, if ρ is linear 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

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

where $A(\rho,\alpha),B(\rho,\alpha),C(\rho,\alpha),D(\rho,\alpha))$ are the sum of $|\chi_\rho(g)|^2\chi_{U_\alpha}(g)$ over all element 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 C.2,

$$A(\rho,\alpha) = \sum_{x\in\mathbb{F}_q^*} |\chi_\rho(a_x)|^2\chi_{U_\alpha}(a_x)$$

$$B(\rho,\alpha) = (q^2-1)\sum_{x\in\mathbb{F}_q^*} |\chi_\rho(b_x)|^2\chi_{U_\alpha}(b_x)$$

$$C(\rho,\alpha) = \frac{1}{2}(q^2+q)\sum_{x,y\in\mathbb{F}_q^*,x\neq y} |\chi_\rho(c_{x,y})|^2\chi_{U_\alpha}(c_{x,y})$$

$$D(\rho,\alpha) = \frac{1}{2}(q^2-q)\sum_{x,y\in\mathbb{F}_q,y\neq 0} |\chi_\rho(d_{x,y})|^2\chi_{U_\alpha}(d_{x,y})\,.$$

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 irrep ρ of $\mathsf{GL}_2(\mathbb{F}_q)$. We begin with the following lemma.

Lemma 19. *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 19.

Corollary 20. *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 21. *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 $\omega = e^{\frac{2\pi i}{q-1}}$ 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 17 will immediately follows Lemma 22 below.

Lemma 22. *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 21, 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

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

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 20, 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 19 and Corollary 20. 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 20 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 19, 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 20,

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

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 $k=0,1,\dots,q-2$ 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 19, 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.

□

D Rational Goppa Codes

This part summarizes definitions and key properties of rational Goppa codes that would be useful in our analysis. Following Stichtenoth [25], we shall describe Goppa codes in terms of algebraic function fields instead of algebraic curves. A complete treatment for this subject can be found in [26].

Let F be a finite field. A *rational function field* over F is a field extension $F(x)/F$ for some x transcendental over F. Each element $z \in F(x)$ can be viewed as a function whose evaluation at a base field element $a \in F$ is determined as follows: write $z = f(x)/g(x)$ for some polynomials $f(x), g(x) \in F[x]$, then

$$z(a) = \begin{cases} \frac{f(a)}{g(a)} \in F & \text{if } g(a)\neq 0 \\ \infty & \text{if } g(a)=0. \end{cases}$$

A *Rieman-Roch space*[7] in the rational function field $F(x)/F$ is a subset of $\mathbb{F}_q$ of the form

$$\mathscr{L}(r,g,h) = \left\{ \frac{f(x)g(x)}{h(x)} \,\middle|\, f(x)\in F[x],\ \deg f(x) \le r \right\}$$

for some nonzero polynomials $g(x), h(x) \in F[x]$ and some integer r. Note that $\mathscr{L}(r,g,h)$ is a vector space of dimension $r+1$ over F.

[7] In terms of algebraic function fields, a Rieman-Roch space is defined in the association with a *divisor* of the function field K/F, where a divisor is a finite sum $\sum_i n_i P_i$ with $n_i \in \mathbb{Z}$ and P_i's being *places* of the function field. In the rational function field $F(x)/F$, we can show that every divisor can be written as $rP_\infty + (z)$ for some integer r and some nonzero $z \in F(x)$, where P_∞ is the infinite place (defined in [26, pg. 9]), and (z) is the *principal divisor* of z. The space $\mathscr{L}(r,g,h)$ is indeed the Rieman-Roch space associated with the divisor $rP_\infty + (z)$ with $z = h(x)/g(x)$.

Definition. (A special case of Definition 2.2.1 in [26]) Let $g(x), h(x) \in F[x]$ be nonzero coprime polynomials, and let $r < n$ a nonnegative integer. Let $\gamma_1, \ldots, \gamma_n$ be n distinct elements in the field[8] F such that $g(\gamma_i) \neq 0$ and $h(\gamma_i) \neq 0$ for all i. Then a *rational Goppa code* associated with g, h and γ_i's is defined by

$$\mathscr{C}(\gamma_1, \ldots, \gamma_n, r, g, h) \stackrel{\text{def}}{=} \{(z(\gamma_1), \ldots, z(\gamma_n)) \mid z \in \mathscr{L}(r, g, h)\} \subset F^n .$$

Remark A classical binary Goppa code can be obtained by setting $F = \mathbb{F}_{2^m}$, $r = n - \deg g(x) - 1$, and $h(x) = \sum_{j=1}^{n} \prod_{i \neq j} (x - \gamma_i)$ and then intersecting the code $\mathscr{C}(\gamma_1, \ldots, \gamma_n, r, g, h)$ with the vector space $\mathbb{F}_2^n$ (see [1]). Generalized Reed-Solomon codes are a special case of rational Goppa codes in which the polynomials $g(x)$ and $h(x)$ are both constants.

Theorem 23. *(A special case of Corollary 2.2.3 in [26]) The code defined in Definition D is an $[n,k,d]$-linear code over F with dimension $k = r+1$ and minimum distance $d \geq n - r$. Consequentially, this code can correct at least $(n-r-1)/2$ errors.*

The rational Goppa code $\mathscr{C}(\gamma_1, \ldots, \gamma_n, r, g, h)$ has a generator matrix:

$$M_0 = \begin{pmatrix} \frac{g(\gamma_1)}{h(\gamma_1)} & \cdots & \frac{g(\gamma_n)}{h(\gamma_n)} \\ \gamma_1 \frac{g(\gamma_1)}{h(\gamma_1)} & \cdots & \gamma_n \frac{g(\gamma_n)}{h(\gamma_n)} \\ \vdots & \ddots & \vdots \\ \gamma_1^r \frac{g(\gamma_1)}{h(\gamma_1)} & \cdots & \gamma_n^r \frac{g(\gamma_n)}{h(\gamma_n)} \end{pmatrix} .$$

Proposition 24. *The matrix M_0 has full rank (over F), that is, its column rank equals $r+1$. Hence, every generator matrix of a rational Goppa code has full rank.*

Proof. It suffices to to show that the first $r+1$ columns of M_0 are linearly independent. Equivalently, we show that the matrix N_0 below has nonzero determinant:

$$N_0 = \begin{pmatrix} \frac{g(\gamma_1)}{h(\gamma_1)} & \cdots & \frac{g(\gamma_{r+1})}{h(\gamma_{r+1})} \\ \gamma_1 \frac{g(\gamma_1)}{h(\gamma_1)} & \cdots & \gamma_{r+1} \frac{g(\gamma_{r+1})}{h(\gamma_{r+1})} \\ \vdots & \ddots & \vdots \\ \gamma_1^r \frac{g(\gamma_1)}{h(\gamma_1)} & \cdots & \gamma_{r+1}^r \frac{g(\gamma_{r+1})}{h(\gamma_{r+1})} \end{pmatrix} = \begin{pmatrix} 1 & \cdots & 1 \\ \gamma_1 & \cdots & \gamma_{r+1} \\ \vdots & \ddots & \vdots \\ \gamma_1^r & \cdots & \gamma_{r+1}^r \end{pmatrix} \begin{pmatrix} \frac{g(\gamma_1)}{h(\gamma_1)} & & \\ & \ddots & \\ & & \frac{g(\gamma_{r+1})}{h(\gamma_{r+1})} \end{pmatrix} .$$

The first matrix in the above product is a Vandermonde matrix, which has nonzero determinant because γ_i's are distinct. The second matrix also has nonzero determinant because $g(\gamma_i) \neq 0$ for all i. Hence, N_0 has nonzero determinant. ☐

An important property of rational Goppa codes is that in general their automorphisms are induced by projective transformations of the projective line. We will make this precise below.

Definition. (See [27, pg. 53]) Let C be a code of length n. An *automorphism* of C is a permutation $\pi \in S_n$ which maps every word in C to a word in C by acting on the positions of the codewords. The set of all automorphisms of C forms a group called the **automorphism group** of C.

[8] In the case $r = \deg h(x) - \deg g(x)$, one can choose one of the points P_i's to be ∞. However, we rule out this case to keep the discussion simple.

In particular, an automorphism of $\mathscr{C}(\gamma_1,\ldots,\gamma_n,r,g,h)$ is a permutation $\pi \in S_n$ such that

$$\mathscr{C}(\gamma_1,\ldots,\gamma_n,r,g,h) = \mathscr{C}(\gamma_{\pi(1)},\ldots,\gamma_{\pi(n)},r,g,h)\,.$$

Remark Suppose M is a generator matrix for an $[n,k]$-linear code C over F. Then a permutation $\pi \in S_n$ is an automorphism of C if and only if there is an invertible matrix $A \in \mathsf{GL}_k(F)$ such that $AMP_\pi = P_\pi$, where P_π denotes the permutation matrix corresponding to π. If M has full rank, there is exactly one such matrix A for each automorphism π of C.

Theorem 25 (Stichtenoth [25])**.** *Suppose* $1 \le r \le n-3$. *Then the automorphism group of the rational Goppa code* $\mathscr{C}(\gamma_1,\ldots,\gamma_n,r,g,h)$ *is isomorphic to a subgroup of* $\mathrm{Aut}(F(x)/F)$.

Fact. The automorphism group of the rational function field $F(x)/F$ is isomorphic to the projective linear group over F. In notations, $\mathrm{Aut}(F(x)/F) \simeq \mathsf{PGL}_2(F)$.

Let $C = \mathscr{C}(\gamma_1,\ldots,\gamma_n,r,g,h)$ be a rational Goppa code. To give an intuition for how the automorphism group of C is embedded in $\mathsf{PGL}_2(F)$, consider a transformation $\sigma \in \mathsf{PGL}_2(F)$ and view each element $a \in F$ as the projective line $[a:1]$ (the point at infinity is written as $[1:0]$). Suppose σ transforms $[a:1]$ to the projective line $[b:1]$, then we shall write $\sigma a = b$. If σ transforms each line $[\gamma_i:1]$ to some line $[\gamma_j:1]$, then σ induces another rational Goppa code:

$$\mathscr{C}(\sigma\gamma_1,\ldots,\sigma\gamma_n,r,g,h)\,.$$

If, further, $\mathscr{C}(\sigma\gamma_1,\ldots,\sigma\gamma_n,r,g,h)$ equals the original code C, then σ induces an automorphism of C. Stichtenoth's theorem establishes that every automorphism of C is induced by such a transformation in $\mathsf{PGL}_2(F)$.