Electronic Colloquium on Computational Complexity, Report No. 129 (2019)

Fourier and Circulant Matrices are Not Rigid

Zeev Dvir* Allen Liu†

Abstract

The concept of matrix rigidity was first introduced by Valiant in [Val77]. Roughly speaking, a matrix is rigid if its rank cannot be reduced significantly by changing a small number of entries. There has been extensive interest in rigid matrices as Valiant showed in [Val77] that rigidity can be used to prove arithmetic circuit lower bounds.

In a surprising result, Alman and Williams showed that the (real valued) Hadamard matrix, which was conjectured to be rigid, is actually not very rigid. This line of work was extended by [DE17] to a family of matrices related to the Hadamard matrix, but over finite fields. In our work, we take another step in this direction and show that for any abelian group G and function $f : G \to \mathbb{C}$, the matrix given by $M_{xy} = f(x-y)$ for $x, y \in G$ is not rigid. In particular, we get that complex valued Fourier matrices, circulant matrices, and Toeplitz matrices are all not rigid and cannot be used to carry out Valiant's approach to proving circuit lower bounds. Our results also hold when we consider matrices over a fixed finite field instead of the complex numbers. This complements a recent result of Goldreich and Tal [GT16] who showed that Toeplitz matrices are nontrivially rigid (but not enough for Valiant's method). Our work differs from previous non-rigidity results in that those works considered matrices whose underlying group of symmetries was of the form $\mathbb{F}_p^n$ with p fixed and n tending to infinity, while in the families of matrices we study, the underlying group of symmetries can be any abelian group and, in particular, the cyclic group $\mathbb{Z}_N$, which has very different structure. Our results also suggest natural new candidates for rigidity in the form of matrices whose symmetry groups are highly non-abelian.

Our proof for matrices over $\mathbb{C}$ has four parts. The first extends the results of [AW16, DE17] to generalized Hadamard matrices over the complex numbers via a new proof technique. The second part handles the $N \times N$ Fourier matrix when N has a particularly nice factorization that allows us to embed smaller copies of (generalized) Hadamard matrices inside of it. The third part uses results from number theory to bootstrap the non-rigidity for these special values of N and extend to all sufficiently large N. The fourth and final part involves using the non-rigidity of the Fourier matrix to show that the group algebra matrix, given by $M_{xy} = f(x-y)$ for $x, y \in G$, is not rigid for any function f and abelian group G. Once we complete the proof for matrices over $\mathbb{C}$, we introduce a few additional tools for extending our results to finite fields.

*Department of Computer Science and Department of Mathematics, Princeton University. Email: zeev.dvir@gmail.com.
Research supported by NSF CAREER award DMS-1451191 and NSF grant CCF-1523816.
†Department of Mathematics, MIT. Email: cliu568@mit.edu

ISSN 1433-8092

1 Introduction

1.1 Background

A major goal in complexity theory is to prove lower bounds on the size and depth of arithmetic circuits that compute certain functions. One specific problem that remains open despite decades of effort is to find functions for which we can show super-linear size lower bounds for circuits of logarithmic depth. In [Val77], Valiant introduced the notion of matrix rigidity as a possible method of proving such lower bounds for arithmetic circuits. More precisely, over a field $\mathbb{F}$, an $m \times n$ matrix M is said to be (r, s)-rigid if any $m \times n$ matrix of rank at most r differs from M in at least s entries. Valiant showed that for any linear function $f : \mathbb{F}^n \to \mathbb{F}^n$ that can be computed by an arithmetic circuit of size $O(n)$ and depth $O(\log n)$, the corresponding matrix can be reduced to rank $O(\frac{n}{\log \log n})$ by changing $O(n^{1+\epsilon})$ entries for any $\epsilon > 0$. Thus, to prove a circuit lower bound for a function f, it suffices to lower bound the rigidity of the corresponding matrix at rank $O(\frac{n}{\log \log n})$. We call a matrix Valiant-rigid if it is $\left(O(\frac{n}{\log \log n}), O(n^{1+\epsilon})\right)$-rigid for some $\epsilon > 0$, i.e. sufficiently rigid for Valiant's method to yield circuit lower bounds. Over any infinite field, Valiant shows that almost all $n \times n$ matrices are $(r, (n-r)^2)$-rigid for any r, while over a finite field one can get a similar result with a logarithmic loss in the sparsity parameter. Despite extensive work, explicit constructions of rigid matrices have remained elusive.

Over infinite (or very large) fields, there are ways to construct highly rigid matrices using either algebraically independent entries or entries that have exponentially large description (see [Lok06,KLPS14,Lok00]) [1]. However, these constructions are not considered to be fully explicit as they do not tell us anything about the computational complexity of the corresponding function. Ideally, we would be able to construct rigid 0, 1-matrices, but even a construction where the entries are in a reasonably simple field (such as the Fourier matrix) would be a major breakthrough. The best known constructions of such matrices are $(r, O(\frac{n^2}{r} \log \frac{n}{r}))$-rigid (see [SSS97, Fri93]). There has also been work towards constructing semi-explicit rigid matrices, which require $O(n)$ bits of randomness (instead of the usual $O(n^2)$), as such a construction would still yield circuit lower bounds through Valiant's approach [2]. The best result in this realm (see [GT16]) shows that random Toeplitz matrices are $(r, \frac{n^3}{r^2 \log n})$-rigid with high probability. Note that both of these bounds become trivial when r is $\frac{n}{\log \log n}$.

Many well-known families of matrices, such as Hadamard matrices and Fourier transform matrices, have been conjectured to be Valiant-rigid [L$^+$09]. However, a recent line of works (see [AW16, DE17]) shows that certain well-structured matrices are not rigid. Alman and Williams show in [AW16] that the $2^n \times 2^n$ Hadamard matrix, given by $H_{xy} = (-1)^{x \cdot y}$ as x and y range over $\{0, 1\}^n$, is not Valiant-rigid over $\mathbb{Q}$. Along similar lines, Dvir and Edelman show in [DE17] that group algebra matrices for the additive group $\mathbb{F}_p^n$, given by $M_{xy} = f(x - y)$ where $f : \mathbb{F}_p^n \to \mathbb{F}_p$ and x, y range over $\mathbb{F}_p^n$, are not Valiant-rigid over $\mathbb{F}_p$ (where we view p as fixed and n goes to infinity). The Hadamard matrix and the group algebra matrices for $\mathbb{F}_p^n$ satisfy the property that for any $\epsilon > 0$, there exists an $\epsilon' > 0$ such that it is possible to change at most $N^{1+\epsilon}$ entries and reduce the rank to $N^{1-\epsilon'}$ (where N denotes the size of the matrix). The proofs of both results rely on constructing a matrix determined by a polynomial $P(x, y)$ that agrees with the given matrix on almost all entries and then arguing that the constructed matrix has low rank.

1.2 Our Contribution

In this paper, we show that several broad families of matrices, including Fourier, circulant and Toeplitz matrices[3], are all not Valiant-rigid. The families of matrices we consider in our work have very different underlying group structure than those considered in previous works. Both [AW16, DE17] analyze matrices

[1] It remains open to construct a matrix that is Valiant-rigid, even if we only require that the entries live in a number field of dimension polynomial in the size of the matrix.

[2] Note however, that it is easy to construct rigid matrices with $O(n^{1+\epsilon})$ bits of randomness for any $\epsilon > 0$ (for example by taking a random matrix with at most n^ϵ non-zeros per row) but this is not sufficient for Valiant's approach.

[3] It is not hard to see that rigidity of circulant and Toeplitz matrices is essentially the same question so for the sake of consistency with our (group theoretic) approach we will primarily consider circulant matrices.

constructed from an underlying group of the form $\mathbb{F}_p^n$ with p fixed and n tending to infinity. Fourier and circulant matrices, which we focus on, are analogs of the Hadamard and group algebra matrices[4] for a cyclic group $\mathbb{Z}_N$. Since any abelian group can be decomposed into simple building blocks of the form $\mathbb{Z}_N$, our results extend to all abelian groups (see details below). While most natural constructions of matrices are highly symmetric, our results suggest that matrices that are symmetric under abelian groups are not rigid and that perhaps we should look toward less structured matrices, or matrices whose symmetry group is non-abelian, as candidates for rigidity.

We now move into a more technical overview of our paper. Define the regular-rigidity of a matrix A, $\mathsf{r}_A(r)$, as the minimum value of s such that it is possible to change at most s entries in each row and column of A to obtain a matrix of rank at most r. The notion of regular-rigidity is weaker than the usual notion of rigidity (and is also weaker than the commonly used notion of row-rigidity) as if A is an $n \times n$ matrix and A is (r, ns)-rigid then $\mathsf{r}_A(r) \geq s$. Note that this actually makes our results stronger as we will show that the matrices we consider are not regular-rigid.

In general, matrices that we deal with will be over $\mathbb{C}$ except in Sections 7-8 where we extend our results to matrices over finite fields. The $d^n \times d^n$ generalized Hadamard matrix $H_{d,n}$ has rows and columns indexed by $\mathbb{Z}_d^n$ and entries $H_{xy} = \omega^{x \cdot y}$ where $\omega = e^{\frac{2\pi i}{d}}$. Throughout this paper, we use the term Hadamard matrix to refer to any generalized Hadamard matrix. We use F_N to denote the $N \times N$ Fourier transform matrix. Our main result, that all Fourier matrices are not rigid enough to carry out Valiant's approach, is stated below.

Theorem 1.1 (Fourier Matrices are Not Rigid)**.** *Let F_N denote the $N \times N$ Fourier transform matrix. For any fixed $0 < \epsilon < 0.1$ and N sufficiently large,*

$$\mathsf{r}_{F_N}\left(\frac{N}{2^{\epsilon^6 (\log N)^{0.35}}}\right) \leq N^{15\epsilon}$$

One key idea in our work is the observation that, if a large family of matrices $\mathcal{A}$ are all diagonalizable by a single matrix M then, the rigidity of *any* matrix $A \in \mathcal{A}$ implies the rigidity of the single matrix M. This situation happens, e.g., when $\mathcal{A}$ is the family of circulant matrices and M is the Fourier matrix. This simple, yet crucial observation allows us to deduce the non-rigidity of a larger family of matrices.

Theorem 1.2. *[Circulant Matrices are not Rigid] Let $0 < \epsilon < 0.1$ be fixed. For all sufficiently large N, if M is an $N \times N$ circulant matrix over $\mathbb{C}$,*

$$\mathsf{r}_{F_N}\left(\frac{N}{2^{\epsilon^6 (\log N)^{0.35}}}\right) \leq N^{15\epsilon}$$

Remark. *We will show later in Section 7 that the same result holds if we let M be an $N \times N$ circulant matrix with entries over a finite field $\mathbb{F}_q$ and require that N is sufficiently large compared to ϵ, q.*

Also notice that since any Toeplitz matrix of size at most $\frac{N}{2}$ can be embedded in an $N \times N$ circulant matrix, the above implies an analogous result for all Toeplitz matrices. While [GT16] shows nontrivial rigidity lower bounds for rank much smaller than N, our results imply that there are actually no nontrivial rigidity lower bounds for rank close to N.

With a bit more work, it is possible to prove the non-rigidity of group algebra matrices for any abelian group.

Theorem 1.3. *Let $0 < \epsilon < 0.1$ be fixed. Let G be an abelian group and $f : G \to \mathbb{C}$ be a function. Let M be a matrix with rows and columns indexed by elements $x, y \in G$ and entries $M_{xy} = f(x - y)$. If $|G|$ is sufficiently large then*

$$\mathsf{r}_M\left(\frac{2|G|}{2^{\epsilon^8 (\log |G|)^{0.32}}}\right) \leq |G|^{38\epsilon}$$

Remark. *Similar to the previous theorem, we prove an analogous result for matrices with entries over a finite field $\mathbb{F}_q$ in Section 8.*

[4]While group algebra matrices are supposed to be defined as $M_{xy} = f(x - y)$, we will work with $M_{xy} = f(x + y)$ in the body of our paper for technical reasons. Note that the two definitions differ only in a permutation of the rows and thus have the same rigidity.

1.3 Proof Overview

We now take a more detailed look at the techniques used in the proof of Theorem 1.1.

1.3.1 Generalized Hadamard Matrices

The first step in the proof of Theorem 1.1 is proving the following result that all Hadamard matrices are not rigid.

Theorem 1.4 (Hadamard Matrices are not Rigid)**.** *For fixed d and $0 < \epsilon < 0.1$, there exists an ϵ' such that for all sufficiently large n, $r_{H_{d,n}}\left(d^{n(1-\epsilon')}\right) \leq d^{n\epsilon}$*

Note that Theorem 1.4 generalizes the main result of [AW16] (which only deals with $d = 2$). Also, given any $d^n \times d^n$ matrix of the form $M_{xy} = f(x - y)$ with $f : \mathbb{Z}_d^n \to \mathbb{C}$, we can permute its rows so that it is diagonalized by $H_{d,n}$. Thus, we can apply the diagonalization trick mentioned above and obtain the following result, which extends the work in [DE17] to matrices over $\mathbb{C}$.

Corollary 1.5. *Let f be a function from $\mathbb{Z}_d^n \to \mathbb{C}$ and let M be a $d^n \times d^n$ matrix with $M_{xy} = f(x-y)$. Then for any fixed d and $0 < \epsilon < 0.1$, there exists an $\epsilon' > 0$ such that for all sufficiently large n, $r_M\left(d^{n(1-\epsilon')}\right) \leq d^{n\epsilon}$*

1.3.2 Fourier Matrices

Equipped with the machinery for Hadamard matrices, we can complete the proof of Theorem 1.1. Our proof consists of two steps. First we show that for integers N of a very special form, the $N \times N$ Fourier matrix is not rigid because it can be decomposed into submatrices with Hadamard-type structure. We say an integer N is well-factorable if it is a product of distinct primes $q_1, \dots, q_l$ such that for all i, $q_i - 1$ has no large prime power divisors. We will make this notion more precise later, but informally, the first step is as follows:

Theorem 1.6. *Let F_N denote the $N \times N$ Fourier transform matrix. For any fixed $0 < \epsilon < 0.1$ and well-factorable integer N, we have*

$$r_{F_N}\left(\frac{N}{2^{\epsilon^6 (\log N)^{0.36}}}\right) \leq N^{7\epsilon}$$

The main intuition is that if N is a product of distinct primes $q_1, \dots, q_l$, then within the Fourier matrix F_N, we can find submatrices whose rows and columns can be indexed by $\mathbb{Z}_{q_1}^* \otimes \cdots \otimes \mathbb{Z}_{q_l}^*$. This multiplicative structure can be replaced by the additive structure of $\mathbb{Z}_{q_1-1} \otimes \cdots \otimes \mathbb{Z}_{q_l-1}$. We can then factor each additive group $\mathbb{Z}_{q_i-1}$ into prime power components. If $q_1 - 1, \dots, q_l - 1$ all have no large prime power divisors, we expect prime powers to be repeated many times when all of the terms are factored. This allows us to find submatrices with $\mathbb{Z}_d^l$ additive structure for which we can apply tools such as Theorem 1.4 and Corollary 1.5 to reduce the rank while changing a small number of entries. We then bound the rank and total number of entries changed over all submatrices to deduce that F_N is not rigid.

The second step of our proof that Fourier matrices are not rigid involves extending Theorem 1.6 to all values of N. The diagonalization trick gives that $N \times N$ circulant matrices are not rigid when N is well-factorable. We then show that for $N' < \frac{N}{2}$, we can rescale the columns of the $N' \times N'$ Fourier matrix and embed it into an $N \times N$ circulant matrix. As long as N' is not too much smaller than N (say $N' > \frac{N}{(\log N)^2}$), we get that the $N' \times N'$ Fourier matrix is not rigid. Thus, for each well-factorable N and all N' in the range $\frac{N}{(\log N)^2} < N' < \frac{N}{2}$, the $N' \times N'$ Fourier transform matrix is not rigid. We then use a number theoretic result of [BH98] to show that the gaps between well-factorable integers are not too large. Thus, the above intervals cover all integers as N runs over all well-factorable numbers, finishing the proof.

1.4 Organization

In Section 2, we introduce notation and prove several basic results that we will use throughout the paper. In Section 3, we show that Hadamard and several closely related families of matrices are not rigid. In Section

4, we show that $N \times N$ Fourier matrices are not rigid when N satisfies certain number-theoretic properties. In Section 5, we complete the proof that all Fourier matrices are not rigid. We then deduce that all Toeplitz matrices are not rigid. In Section 6, we use the results from the previous section to show that group algebra matrices for abelian groups are not rigid. Through Sections 2-6, we work with matrices over $\mathbb{C}$ for ease of exposition. In Section 7 and Section 8, we sketch how to modify the proofs in the previous sections to deal with "missing" roots of unity over a finite field. Finally, in Section 9, we discuss a few open questions and possible directions for future work.

2 Preliminaries

Throughout this paper, we let $d \geq 2$ be an integer and $\omega = e^{\frac{2\pi i}{d}}$ be a primitive d^{th} root of unity. When we consider an element of $\mathbb{Z}_d^n$, we will view it as an n-tuple with entries in the range $[0, d-1]$. When we say a list of d^n elements $x_1, \dots, x_{d^n}$ is indexed by $\mathbb{Z}_d^n$, we mean that each x_i is labeled with an element of $\mathbb{Z}_d^n$ such that all labels are distinct and the labels of $x_1, \dots, x_{d^n}$ are in lexicographical order.

2.1 Basic Notation

We will frequently work with tuples, say $I = (i_1, \dots, i_n) \in \mathbb{Z}_d^n$. Below we introduce some notation for dealing with tuples that will be used later on.

Definition 2.1. *For a tuple I, we let I^i denote its i^{th} entry. For instance if $I = (i_1, \dots, i_n)$ then $I^k = i_k$.*

Definition 2.2. *For an n-tuple $I = (i_1, i_2, \dots, i_n)$, define the polynomial over n variables $x^I = x_1^{i_1} \dots x_n^{i_n}$.*

Definition 2.3. *For ω a d^{th} root of unity and an n-tuple $I = (i_1, i_2, \dots, i_n) \in \mathbb{Z}_d^n$, we define $\omega^{[I]} = (\omega^{i_1}, \dots, \omega^{i_n})$.*

Definition 2.4. *For a function $f : \mathbb{Z}_d^n \to \mathbb{C}$, define the n-variable polynomial P_f as*

$$P_f = \sum_{I \in \mathbb{Z}_d^n} f(I) x^I$$

Definition 2.5. *For an n-tuple $I = (i_1, i_2, \dots, i_n)$, we define the set $\mathsf{perm}(I)$ to be a set of n-tuples consisting of all distinct permutations of the entries of I. Similarly, for a set of n-tuples S, we define $\mathsf{perm}(S)$ to be the set of all n-tuples that can be obtained by permuting the entries of some element of S.*

Definition 2.6. *We say a set $S \subseteq \mathbb{Z}_d^n$ is symmetric if for any $I \in S$, $\mathsf{perm}(I) \subseteq S$.*

Definition 2.7. *For a set of n-tuples S, let $\mathsf{red}(S)$ denote the set of equivalence classes under permutation of entries in S. Let $\mathsf{rep}(S)$ be a set of n-tuples formed by taking one representative from each equivalence class in $\mathsf{red}(S)$ (note $\mathsf{rep}(S)$ is not uniquely determined but this will not matter for our use).*

Note that if $\mathsf{rep}(S) = \{I_1, \dots, I_k\}$, then the sets $\mathsf{perm}(I_1), \mathsf{perm}(I_2), \dots, \mathsf{perm}(I_k)$ are disjoint and their union contains S. If the set S is symmetric then their union is exactly S.

2.2 Special Families of Matrices

We now define notation for working with a few special families of matrices.

Definition 2.8. *An $N \times N$ matrix M is called a Toeplitz matrix if M_{ij} depends only on $i - j$. An $N \times N$ matrix M is called a Hankel matrix if M_{ij} depends only on $i + j$. Note that the rows of any Toeplitz matrix can be permuted to obtain a Hankel matrix so any non-rigidity results we show for one family also hold for the other.*

Definition 2.9. *For an abelian group G and a function $f : G \to \mathbb{C}$, let $M_G(f)$ denote the $|G| \times |G|$ matrix (over $\mathbb{C}$) whose rows and columns are indexed by elements $x, y \in G$ and whose entries are given by $M_{xy} = f(x+y)$. When it is clear what G is from context, we will simply write $M(f)$. We let V_G denote the family of matrices $M_G(f)$ as f ranges over all functions from G to $\mathbb{C}$. We call V_G the family of adjusted group algebra matrices for the group G. When G is a cyclic group, we call the matrices in V_G adjusted-circulant.*

Compared to the usual group algebra (and circulant) matrices defined by $M_{xy} = f(x-y)$, the matrix $M_G(f)$ differs only in a permutation of the rows. In the proceeding sections, we will work with $M_G(f)$ for technical reasons, but it is clear that the same non-rigidity results hold for the usual group algebra matrices. Similarly, we will use adjusted-circulant and Hankel matrices as it is clear that the same non-rigidity results hold for circulant and Toeplitz matrices. Also note that adjusted-circulant matrices are a special case of Hankel matrices.

Definition 2.10. *Let $H_{d,n}$ denote the $d^n \times d^n$ Hadamard matrix, i.e. the matrix whose rows and columns are indexed by n-tuples $I, J \in \mathbb{Z}_d^n$ and whose entries are $H_{IJ} = \omega^{I \cdot J}$ where $\omega = e^{\frac{2\pi i}{d}}$. When $n = 1$, we define $F_d = H_{d,1}$ and call F_d a Fourier matrix.*

2.3 Matrix Rigidity

Here, we review basic notation for matrix rigidity.

Definition 2.11. *For a matrix M and a real number r, we define $\mathsf{R}_M(r)$ to be the smallest number s for which there exists a matrix A with at most s nonzero entries and a matrix B of rank at most r such that $M = A + B$. If $\mathsf{R}_M(r) \geq s$, we say M is (r, s)-rigid.*

Definition 2.12. *For a matrix M and a real number r, we define $\mathsf{r}_M(r)$ to be the smallest number s for which there exists a matrix A with at most s nonzero entries in each row and column and a matrix B of rank at most r such that $M = A + B$. If $\mathsf{r}_M(r) \geq s$, we say M is (r, s)-regular rigid.*

It is clear that if a matrix is (r, ns)-rigid, then it must be (r, s)-regular rigid. In proceeding sections, we will show that various matrices are not $(\frac{N}{\log \log N}, N^{\epsilon})$-regular rigid for any $\epsilon > 0$ and this will imply that Valiant's method for showing circuit lower bounds in [Val77] cannot be applied.

2.4 Preliminary Results

Next, we mention several basic results that will be useful in the proofs later on.

Claim 2.13. *$H_{d,n} = \underbrace{F_d \otimes \cdots \otimes F_d}_{n}$ where $\otimes$ denotes the Kronecker product.*

Proof. This can easily be verified from the definition. $\square$

Claim 2.14. *$H_{d,n} H_{d,n}^* = d^n I$ where $H_{d,n}^*$ is the conjugate transpose of $H_{d,n}$ and I is the identity matrix.*

Proof. We verify that $F_d F_d^* = dI$, and then using the previous claim, we deduce that $H_{d,n} H_{d,n}^* = d^n I$. $\square$

Claim 2.15. *Let $f : \mathbb{Z}_d^n \to \mathbb{C}$ be a function. Let ω be a d^{th} root of unity and set $P_f = \sum_{I \in \mathbb{Z}_d^n} f(I) x^I$. Let $D = H_{d,n} M_{\mathbb{Z}_d^n}(f) H_{d,n}$. Then D is a diagonal matrix with diagonal entries $d^n P_f(\omega^{[J]})$ as J ranges over $\mathbb{Z}_d^n$.*

Proof. First, we analyze the product $M_{\mathbb{Z}_d^n}(f) H_{d,n}$. This is a $d^n \times d^n$ matrix and its rows and columns can naturally be indexed by tuples $I, J \in \mathbb{Z}_d^n$. The entry with row indexed by I and column indexed by J is

$$\sum_{I' \in \mathbb{Z}_d^n} f(I + I') \omega^{I' \cdot J} = \omega^{-I \cdot J} \sum_{I' \in \mathbb{Z}_d^n} f(I + I') \omega^{(I' + I) \cdot J} = \omega^{-I \cdot J} P_f(\omega^{[J]})$$

Therefore, the columns of $M_{\mathbb{Z}_d^n}(f) H_{d,n}$ are multiples of the columns of $H_{d,n}^*$. In fact, the column of $M_{\mathbb{Z}_d^n}(f) H_{d,n}$ indexed by J is $P_f(\omega^{[J]})$ times the corresponding column of $H_{d,n}^*$. Since $H_{d,n} H_{d,n}^* = d^n I$, D must be a diagonal matrix whose entries on the diagonal are $d^n P_f(\omega^{[J]})$ as J ranges over $\mathbb{Z}_d^n$. $\square$

Plugging $n = 1$ into the above gives:

Claim 2.16. *Let M be a $d \times d$ adjusted-circulant matrix. Then $F_d M F_d$ is a diagonal matrix.*

Claim 2.15 gives us a characterization of the rank of matrices of the form $M_{\mathbb{Z}_d^n}(f)$.

Claim 2.17. *Let $f : \mathbb{Z}_d^n \to \mathbb{C}$ be a function. Let ω be a d^{th} root of unity and say $P_f = \sum_{I \in \mathbb{Z}_d^n} f(I) x^I$ has C roots among the set $\{(\omega^{i_1}, \dots, \omega^{i_n}) | (i_1, \dots, i_n) \in \mathbb{Z}_d^n\}$. Then $\mathsf{rank}(M_{\mathbb{Z}_d^n}(f)) = d^n - C$.*

Proof. Consider the product $D = H_{d,n} M_{\mathbb{Z}_d^n}(f) H_{d,n}$. Note that $H_{d,n}$ is clearly invertible by Claim 2.14. Therefore, it suffices to compute the rank of D. By Claim 2.15, D must be a diagonal matrix whose entries on the diagonal are $d^n P_f(\omega^{[J]})$ as J ranges over $\mathbb{Z}_d^n$. The rank of D is the number of nonzero diagonal entries which is simply $d^n - C$ $\square$

As mentioned in the introduction, we can relate the rigidity of a matrix to the rigidity of matrices that it diagonalizes.

Lemma 2.18. *If $B = A^* D A$ where D is a diagonal matrix and $\mathsf{r}_A(r) \leq s$ then $\mathsf{r}_B(2r) \leq s^2$. The same inequality holds also for $B' = ADA$.*

Proof. Let E be the matrix with at most s nonzero entries in each row and column such that $A - E$ has rank at most r. We have

$$B - E^* D E = A^* D (A - E) + (A^* - E^*) D E$$

Since $\mathsf{rank}(A - E) \leq r$, $\mathsf{rank}(B - E^* D E) \leq 2r$. Also, $E^* D E$ has at most s^2 nonzero entries in each row and column so $\mathsf{r}_B(2r) \leq s^2$. The second part can be proved in the exact same way with A^* replaced by A. $\square$

In light of Lemma 2.18, Claim 2.16, and Claim 2.15, proving non-rigidity for $d \times d$ circulant matrices reduces to proving non-rigidty for F_d and proving non-rigidity for group algebra matrices for $\mathbb{Z}_d^n$ reduces to proving non-rigidity for $H_{d,n}$. Below, we show that these statements are actually equivalent.

Claim 2.19. *It is possible to rescale the rows and columns of $H_{d,n}$ to get a matrix of the form $M_{\mathbb{Z}_d^n}(f)$ for some symmetric function $f : \mathbb{Z}_d^n \to \mathbb{C}$. In particular, it is possible to rescale the rows and columns of F_d to get an adjusted-circulant matrix.*

Proof. Let ζ be such that $\zeta^2 = \omega$. Multiply each row of $H_{d,n}$ by $\zeta^{(I \cdot I)}$ and each column by $\zeta^{(J \cdot J)}$ to get a matrix H'. We have

$$H'_{IJ} = \zeta^{(I+J) \cdot (I+J)}$$

For a tuple $x = (x_1, \dots, x_d) \in \mathbb{Z}_d^n$, we define $f(x) = \zeta^{x_1^2 + \cdots + x_d^2}$. To complete the proof, it suffices to show that $f : \mathbb{Z}_d^n \to \mathbb{C}$ is well defined. To do this, we will show that ζ^{x^2} depends only on the residue of $x \mod d$. If d is odd, we can choose ζ to be a d^{th} root of unity and the claim is clear. If d is even $\zeta^{(x+d)^2} = \zeta^{x^2} \zeta^{2dx + d^2}$ but since $2dx + d^2$ is a multiple of $2d$, $\zeta^{2dx+d^2} = 1$ and thus $\zeta^{(x+d)^2} = \zeta^{x^2}$. $\square$

3 Non-rigidity of Generalized Hadamard Matrices

In this section, we show that the Hadamard matrix $H_{d,n}$ becomes highly non-rigid for large values of n. The precise result is stated below.

Theorem 3.1. *Let $N = d^n$ for positive integers d, n. Let $0 < \epsilon < 0.1$ and assume $n \geq \frac{d^2 (\log d)^2}{\epsilon^4}$. Then $\mathsf{r}_{H_{d,n}}(N^{1 - \frac{\epsilon^4}{d^2 \log d}}) \leq N^{\epsilon}$.*

First we prove a few lemmas about symmetric polynomials that we will use in the proof of Theorem 3.1.

Lemma 3.2. *Let T_m denote the set of tuples in $\mathbb{Z}_d^n$ such that at least m entries are equal to 0. Say $\mathsf{rep}(T_m) = \{I_1, \ldots, I_k\}$. Consider the polynomials $P_1(x_1, \ldots, x_n), \ldots, P_k(x_1, \ldots, x_n)$ defined by*

$$P_i(x_1, \ldots, x_n) = \sum_{I \in \mathsf{perm}(I_i)} x^I$$

For any complex numbers $y_1, \ldots, y_m$, and any polynomial $Q(x_{m+1}, \ldots x_n)$ that is symmetric and degree at most $d-1$ in each of its variables, there exist coefficients $c_1, \ldots, c_k$ such that

$$Q(x_{m+1}, \ldots, x_n) = \sum c_i P_i(y_1, \ldots y_m, x_{m+1}, \ldots, x_n)$$

Proof. It suffices to prove the statement for all Q of the form

$$\sum_{I'' \in \mathsf{perm}(I')} x^{I''}$$

where $I' \in \mathbb{Z}_d^{n-m}$. We will prove this by induction on the degree. Clearly one of the I_i is $(0, 0 \ldots 0)$, so one of the polynomials $P_i(x_1, \ldots, x_n)$ is constant. This finishes the case when Q has degree 0. Now we do the induction step. Note that we can extend I' to an element of T_m by setting the first m entries equal to 0. Call this extension I and say that $I \in \mathsf{perm}(I_i)$. We have

$$\sum_{I'' \in \mathsf{perm}(I')} x^{I''} = P_i(y_1, \ldots, y_m, x_{m+1}, \ldots, x_n) - R(y_1, \ldots, y_m, x_{m+1}, \ldots x_n)$$

$R(y_1, \ldots, y_m, x_{m+1}, \ldots x_n)$, when viewed as a polynomial in $x_{m+1}, \ldots, x_n$ (since $y_1, \ldots, y_m$ are complex numbers that we can plug in), is symmetric and of lower degree than the left hand side. Thus, using the induction hypothesis, we can write R in the desired form. This completes the induction step. $\square$

The key ingredient in the proof of Theorem 3.1 is the following lemma which closely resembles the main result in [DE17], but deals with matrices over $\mathbb{C}$.

Lemma 3.3. *Let $f : \mathbb{Z}_d^n \to \mathbb{C}$ be a symmetric function on the n variables. Let $N = d^n$. Let $0 < \epsilon < 0.1$ and assume $n \geq \frac{d^2(\log d)^2}{\epsilon^4}$. Then $\mathsf{r}_{M(f)}(N^{1-\frac{\epsilon^4}{d^2 \log d}}) \leq N^\epsilon$.*

Let $\delta = \epsilon^2$, $m = \lceil n\left(\frac{1-\delta}{d}\right)\rceil$ and let S denote the set of all tuples $(i_1, i_2, \ldots, i_n) \in \mathbb{Z}_d^n$ such that the entries indexed $1, 2, \ldots m$ are equal to 0, the entries indexed $m+1, \ldots, 2m$ are equal to 1 and in general for $0 \leq i \leq d-1$, the entries indexed $im+1, \ldots, (i+1)m$ are equal to i. Note $|S| = d^{n-dm} \approx d^{\delta n} = N^{\epsilon^2}$ (since $n - dm$ is approximately δn).

The main idea will be to change f in a small number of locations so that it has many zeros in the set $\{\omega^{[I]} | I \in \mathbb{Z}_d^n\}$ in order to make use of Claim 2.17. More precisely, first we will change f to f' by changing its values in at most N^ϵ places so that f' is still symmetric in all of the variables and

$$P_{f'}\left(\omega^{[I]}\right) = 0 \ \forall I \in S$$

Note that although the size of S is small, the fact that f' is symmetric implies that f' also vanishes on $\mathsf{perm}(S)$, which covers almost all of $\mathbb{Z}_d^n$. Once we have shown the above, we quantitatively bound the number of entries changed between $M(f)$ and $M(f')$ and also the rank of $M(f')$ to complete the proof of Lemma 3.3. To do the first part, we need the following sub-lemma.

Lemma 3.4. *Let T denote the set of all tuples $(i_1, i_2, \ldots, i_n) \in \mathbb{Z}_d^n$ such that at least $n\big(1-\delta\big)$ of the entries are 0. By changing the values of f only on elements of T, we can obtain f' satisfying*

$$P_{f'}\left(\omega^{[I]}\right) = 0 \ \forall I \in S \tag{1}$$

Proof. We interpret (1) as a system of linear equations where the unknowns are the values of f' at various points. Let $\mathsf{rep}(T) = \{J_1, J_2, \ldots, J_k\}$ for $J_1, J_2, \ldots J_k \in T$. Since we must maintain that f' is symmetric, there are essentially k variables each corresponding to an equivalence class of tuples under permutations. Each equivalence class is of the form $\mathsf{perm}(J_j)$ and we denote the corresponding variable by m_j. The system of equations in (1) can be rewritten in the form

$$\sum_{j=1}^{k} m_j \sum_{J \in \mathsf{perm}(J_j)} \omega^{I \cdot J} + \sum_{J' \notin T} f(J') \omega^{I \cdot J'} = 0 \ \forall I \in S$$

If we let $\mathsf{rep}(S) = \{I_1, I_2, \ldots, I_l\}$, the system has exactly l distinct equations corresponding to each element of $\mathsf{rep}(S)$ due to our symmetry assumptions. Let M denote the $l \times k$ coefficient matrix represented by $M_{ij} = \sum_{J \in \mathsf{perm}(J_j)} \omega^{I_i \cdot J}$. To show that the system has a solution, it suffices to show that the column span of M is full. This is equivalent to showing that for each $i = 1, 2, \ldots l$ there exist coefficients $a_1, a_2, \ldots, a_k$ such that

$$\sum_{j=1}^{k} a_j \cdot \sum_{J \in \mathsf{perm}(J_j)} \omega^{I_i \cdot J} \neq 0$$

$$\sum_{j=1}^{k} a_j \cdot \sum_{J \in \mathsf{perm}(J_j)} \omega^{I_{i'} \cdot J} = 0 \ \forall i' \neq i$$

Fix an index i_0. We can view each equation above as a polynomial in $\omega^{[I_i]}$ given by

$$P(x_1, \ldots, x_n) = \sum_{j=1}^{k} a_j \sum_{J \in \mathsf{perm}(J_j)} x^J$$

and the problem becomes equivalent to constructing a polynomial that vanishes on $\omega^{[I_i]}$ if and only if $i \neq i_0$. Note that only the entries $x_{dm+1}, \ldots, x_n$ matter as we have $x_1 = \cdots = x_m = 1, \ldots, x_{(d-1)m+1} = \cdots = x_{dm} = \omega^{d-1}$ for all points we consider.

For $I_i = (i_1, i_2, \ldots i_n)$, let I'_i denote the sub-tuple $(i_{dm+1}, \ldots, i_n)$. The problem is equivalent to constructing a polynomial

$$Q(x_{dm+1}, \ldots, x_n) = P(1, 1, \ldots, \omega^{d-1}, \ldots, \omega^{d-1}, x_{dm+1}, \ldots x_n)$$

such that Q vanishes on $\omega^{[I'_i]}$ if and only if $i \neq i_0$.

Lemma 3.2 implies that by choosing the coefficients $a_1, \ldots, a_k$, we can make Q be any polynomial that is symmetric in $x_{dm+1}, \ldots, x_n$ and degree at most $d-1$ in each of the variables.

Now consider the polynomial

$$Q_{i_0}(x_{dm+1}, \ldots, x_n) = \sum_{I' \in \mathsf{perm}(I'_{i_0})} \left(\frac{x_{dm+1}^d - 1}{x_{dm+1} - \omega^{I'0}} \right) \cdots \left(\frac{x_n^d - 1}{x_n - \omega^{I'(n-dm)}} \right)$$

(note this is a polynomial with coefficients in $\mathbb{C}$ since each of the factors reduces to a degree $d-1$ polynomial)

It is clear that the above polynomial is symmetric in all of the variables and satisfies the degree constraint so we know we can choose suitable coefficients $a_1, \ldots, a_k$. We claim that the polynomial we construct does not vanish on $\omega^{[I'_{i_0}]}$ but vanishes on $\omega^{[I'_i]}$ for $i \neq i_0$. Indeed, the product

$$\left(\frac{x_{dm+1}^d - 1}{x_{dm+1} - \omega^{I'0}} \right) \cdots \left(\frac{x_n^d - 1}{x_n - \omega^{I'(n-dm)}} \right)$$

is 0 if and only if $(x_{dm+1}, \ldots, x_n) \neq I'$. However, there is exactly one $I' \in \mathsf{perm}(I'_{i_0})$ with $I' = I'_{i_0}$ and none with $I' = I'_i$ for $i \neq i_0$ since $I_1, I_2, \ldots, I_l$ are representatives of distinct equivalence classes under permutation of entries. This means that the polynomial Q_{i_0} we constructed has the desired properties and completes the proof that the system is solvable. $\square$

Proof of Lemma 3.3. Since $M(f) = (M(f) - M(f')) + M(f')$, to complete the proof of Lemma 3.3, it suffices to bound the number of nonzero entries in $M(f) - M(f')$ and the rank of $M(f')$.

The number of nonzero entries in each row and column of $(M(f) - M(f'))$ is at most $|T|$. This is exactly the number of elements of $\mathbb{Z}_d^n$ with at least $n\big(1-\delta\big)$ entries equal to 0. Using standard tail bounds on the binomial distribution, the probability of a random n-tuple having at least that many 0s is at most

$$e^{-nD(1-\delta||\frac{1}{d})} = e^{-n\left((1-\delta)\log(d(1-\delta))+\delta\log(\frac{d\delta}{d-1})\right)} = d^{-n(1-\delta)}e^{-n\left((1-\delta)\log(1-\delta)+\delta\log(\frac{d\delta}{d-1})\right)}$$

where $D(\cdot||\cdot)$ denotes KL-divergence. For $\delta < 0.01$, the above is at most $d^{-n(1-\sqrt{\delta})}$ and thus we change at most $d^{\epsilon n}$ entries in each row and column.

By Claim 2.17, the rank of $M(f')$ is at most $d^n - |\mathsf{perm}(S)|$. Equivalently, this is the number of n-tuples such that some element in $\{0, 1, \ldots, d-1\}$ appears less than $\big(\frac{1-\delta}{d}\big)n$ times. We use Hoeffding's inequality and then union bound over the d possibilites to get the probability that a randomly chosen n-tuple in $\mathbb{Z}_d^n$ is outside S is at most

$$de^{-2\frac{\delta^2 n}{d^2}} = e^{-2\frac{\delta^2 n}{d^2}+\log d}$$

When $n > \frac{d^2(\log d)^2}{\delta^2}$, the above is at most $d^{-\frac{\epsilon^4 n}{d^2\log d}}$ and thus the rank of $M(f')$ is at most $d^{\left(1-\frac{\epsilon^4}{d^2\log d}\right)n}$, completing the proof of Lemma 3.3.

$\square$

Proof of Theorem 3.1. Applying Claim 2.19 and Lemma 3.3 we immediately get the desired. $\square$

Using Theorem 3.1, Lemma 2.18, and Claim 2.15, we get the following result which extends Lemma 3.3 to matrices where f is not symmetric.

Corollary 3.5. *For any function $f : \mathbb{Z}_d^n \to \mathbb{C}$ and any $0 < \epsilon < 0.1$ such that $n \geq \frac{d^2(\log d)^2}{\epsilon^4}$, we have*

$$\mathsf{r}_{M(f)}(2N^{1-\frac{\epsilon^4}{d^2\log d}}) \leq N^{2\epsilon}$$

where $N = d^n$.

4 Non-rigidity for Fourier Matrices of Well-Factorable Size

Our goal in this section is to show that we can find infinitely many values of N for which the Fourier matrix F_N is highly non-rigid. The integers N we analyze will be products of many distinct primes q_i with the property that $q_i - 1$ is very smooth (has all prime factors small). For these values of N, we can decompose the matrix F_N into several submatrices that are closely related to Hadamard matrices. We then apply the results from the previous section to show that each submatrix is non-rigid and aggregate over the submatrices to conclude that F_N is non-rigid.

We first show precisely how to construct N. We rely on the following number theoretic result, found in [BH98], that allows us to find a large set of primes q_i for which $q_i - 1$ is very smooth.

Definition 4.1. *For a positive integer m, let $P^+(m)$ denote the largest prime factor of m. For a fixed positive integer a, let*

$$\pi_a(x, y) = |\{p | a < p \leq x, P^+(p-a) \leq y\}|$$

where p ranges over all primes. In other words, $\pi_a(x, y)$ is the number of primes at most x such that $p - a$ is y-smooth.

Theorem 4.2 ([BH98])[5]. *There exist constants x_0, C such that for $\beta = 0.2961$, $x > x_0$ and $y \geq x^\beta$ we have*

$$\pi_1(x,y) > \frac{x}{(\log x)^C}$$

Throughout the remainder of this paper, set $C_0 = C + 1$ where C is the constant in Theorem 4.2. The properties that we want N to have are stated in the following two definitions.

Definition 4.3. *We say a prime q is (α, x)-good if the following properties hold.*

- $\frac{x}{(\log x)^{C_0}} \leq q \leq x$
- *All prime powers dividing $q-1$ are at most x^α*

Definition 4.4. *We say an integer N is (l, α, x)-factorable if the following properties hold.*

- *$N = q_1 \dots q_l$ where $q_1, \dots, q_l$ are distinct primes*
- *$q_1, \dots, q_l$ are all (α, x)-good*

To show the existence of (l, α, x)-factorable integers, it suffices to show that there are many (α, x)-good primes. This is captured in the following lemma.

Lemma 4.5. *For a fixed constant C_0, any parameter $\alpha > 0.2961$, and sufficiently large x (possibly depending on α), there are at least $\frac{10x}{(\log x)^{C_0}}$ distinct (α, x)-good primes.*

Proof of Lemma 4.5. Let $y = x^\beta$ where $\beta = 0.2961$. By Theorem 4.2, for sufficiently large x, we can find at least $\lceil \frac{x}{(\log x)^C} - \frac{x}{(\log x)^{C_0}} \rceil$ primes $p_1, \dots, p_l$ between $\frac{x}{(\log x)^{C_0}}$ and x such that all prime factors of $p_i - 1$ are at most x^β. Eliminate all of the p_i such that one of the prime powers in the prime factorization of $p_i - 1$ is more than x^α. Note that there are at most x^β distinct primes that divide $p_i - 1$ for some i. Thus, there are at most $x^\beta \log x$ different prime powers bigger than x^α that divide some $p_i - 1$. Each of these prime powers can divide at most $x^{1-\alpha}$ of the elements $\{p_1, \dots, p_l\}$, so in total, we eliminate at most $x^{1-\alpha+\beta} \log x$ of the p_i. Thus, for sufficiently large x, the number of (α, x)-good primes is at least

$$\frac{x}{(\log x)^C} - \frac{x}{(\log x)^{C_0}} - x^{1-\alpha+\beta} \log x \geq \frac{x}{2(\log x)^C}$$

□

For simplicity, we will set $\alpha = 0.3$ by default.

Definition 4.6. *A prime is said to be x-good if it is $(0.3, x)$-good. An integer N is said to be (l, x)-factorable if it is $(l, 0.3, x)$-factorable.*

Lemma 4.5 implies that for all sufficiently large x and $l \leq \frac{x}{(\log x)^{C_0}}$ (where C_0 is an absolute constant), we can find (l, x)-factorable integers. We now show that if we choose x sufficiently large and N to be (l, x)-factorable for some $\frac{x}{(\log x)^{C_0+100}} \leq l \leq \frac{x}{(\log x)^{C_0+10}}$, then F_N is highly non-rigid.

Theorem 4.7. *Let $0 < \epsilon < 0.1$ be some constant. For x sufficiently large and N a (l, x)-factorable number with $\frac{x}{(\log x)^{C_0+100}} \leq l \leq \frac{x}{(\log x)^{C_0+10}}$, we must have*

$$r_{F_N}\left(\frac{N}{2^{\epsilon^6 (\log N)^{0.36}}}\right) \leq N^{7\epsilon}$$

In order to prove Theorem 4.7, we will first prove a series of preliminary results that characterize the structure of Fourier and Hadamard matrices.

[5] [BH98] proves the same inequality with $\pi_a(x,y)$ for any integer a where x_0 may depend on a and C is an absolute constant.

4.1 Structure of Hadamard and Fourier Matrices

Lemma 4.8. *Let $n = x_1x_2\dots x_j$ for pairwise relatively prime positive integers $x_1,\dots,x_j$. There exists a permutation of the rows and columns of F_n, say F' such that*

$$F' = F_{x_1} \otimes \cdots \otimes F_{x_j}$$

where $\otimes$ denotes the Kronecker product.

Proof. Let γ be a primitive n^{th} root of unity. For $i = 1,2,\dots j$, let $\gamma_i = \gamma^{c_i \frac{n}{x_i}}$ where c_i is chosen such that $c_i\frac{n}{x_i} \equiv 1 \mod x_i$. Note this is possible since $x_1,\dots,x_j$ are pairwise relatively prime. γ_i is a primitive ${x_i}^{\text{th}}$ root of unity.

Now by the Chinese remainder theorem, there is a ring isomorphism between $\mathbb{Z}_n$ and $\mathbb{Z}_{x_1} \times \cdots \times \mathbb{Z}_{x_j}$. We can thus view F_n as a matrix whose rows and columns are indexed by elements of $\mathbb{Z}_{x_1} \times \cdots \times \mathbb{Z}_{x_j}$ and such that the entry $F_{n|AB}$ corresponding to tuples $A = (a_1,\dots,a_j)$ and $B = (b_1,\dots,b_j)$ is γ^c where c is the unique element of $\mathbb{Z}_n$ with $c \equiv a_ib_i \mod x_i$ for all i.

For each matrix F_{x_i} its rows and columns are labeled with elements of $\mathbb{Z}_{x_i}$ and its entries are $F_{x_i|ab} = \gamma_i^{a\cdot b}$. Thus in the Kronecker product, the rows and columns are labeled with elements of $\mathbb{Z}_{x_1} \times \cdots \times \mathbb{Z}_{x_j}$ such that the entry corresponding to tuples $(a_1,\dots,a_j)$ and $(b_1,\dots,b_j)$ is

$$\gamma_1^{a_1b_1}\dots\gamma_j^{a_jb_j} = \gamma^{c_1a_1b_1\frac{n}{x_1}+\cdots+c_ja_jb_j\frac{n}{x_j}}$$

For each x_i, we compute the residue of the exponent in the above expression $\mod x_i$. The term $c_ia_ib_i\frac{n}{x_i}$ is congruent to a_ib_i by definition and all other terms are 0 so the sum is congruent to $a_ib_i \mod x_i$. Thus, for some permutation of the rows and columns of F_n, it is equal to the Kronecker product $F_{x_1} \otimes \cdots \otimes F_{x_j}$, as desired. □

Lemma 4.9. *Let $M = A \otimes B$ where A is an $m \times m$ matrix and B is an $n \times n$ matrix. For any two integers r_1, r_2 we have*

$$\mathsf{r}_M(r_1n + r_2m) \le \mathsf{r}_A(r_1)\mathsf{r}_B(r_2)$$

Proof. The proof of this lemma is similar to the proof of Lemma 2.18. There are matrices E, F with atmost $\mathsf{r}_A(r_1)$ and $\mathsf{r}_B(r_2)$ nonzero entries respectively such that $\mathsf{rank}(A + E) \le r_1$ and $\mathsf{rank}(B + F) \le r_2$. We will now show that $\mathsf{rank}(M - E \otimes F) \le r_1n + r_2m$. Indeed

$$M - E \otimes F = (A + E) \otimes B - E \otimes (B + F)$$

and the right hand side of the above has rank at most $r_1n + r_2m$ since rank multiplies under the Kronecker product. Clearly $E \otimes F$ has at most $\mathsf{r}_A(r_1)\mathsf{r}_B(r_2)$ nonzero entries in each row and column so we are done. □

Lemma 4.10. *Consider the matrix*

$$A = (\underbrace{F_{t_1} \otimes \cdots \otimes F_{t_1}}_{a_1}) \otimes \cdots \otimes (\underbrace{F_{t_n} \otimes \cdots \otimes F_{t_n}}_{a_n})$$

Let $0 < \epsilon < 0.1$ be some chosen parameter and D be some sufficiently large constant (possibly depending on ϵ). Assume $t_1 \le t_2 \cdots \le t_n$ and $a_i \ge \max\left(\frac{t_i^2(\log t_i)^2}{\epsilon^{10}}, D\right)$ for all i. Let $P = t_1^{a_1}\dots t_n^{a_n}$ and $L = \lceil 2\log\log P\rceil$. Then

$$\mathsf{r}_A\left(P^{1-\frac{\epsilon^6}{10Lt_n^2\log t_n}}\right) \le P^{5\epsilon}$$

Proof. First, we consider the case when there exists an integer B such that $B \leq t_1^{a_1}, \dots, t_n^{a_n} \leq B^2$. Note that $(\underbrace{F_{t_i} \otimes \cdots \otimes F_{t_i}}_{a_i}) = H_{t_i,a_i}$. By Theorem 3.1, for each i there exists a matrix E_i such that E_i has at most $t_i^{\epsilon a_i}$ nonzero entries in each row and column and $\mathsf{rank}(H_{t_i,a_i} - E_i) \leq t_i^{a_i\left(1-\frac{\epsilon^4}{t_i^2 \log t_i}\right)}$. Let $A_i = H_{t_i,a_i} - E_i$.

$$(\underbrace{F_{t_1} \otimes \cdots \otimes F_{t_1}}_{a_1}) \otimes \cdots \otimes (\underbrace{F_{t_n} \otimes \cdots \otimes F_{t_n}}_{a_n}) = (E_1 + A_1) \otimes \cdots \otimes (E_n + A_n) = \sum_{S \subset [n]} \left(\bigotimes_{i \in S} A_i\right) \otimes \left(\bigotimes_{i' \notin S} E_{i'}\right)$$

$$= \sum_{S \subset [n], |S| \geq \epsilon n} \left(\bigotimes_{i \in S} A_i\right) \otimes \left(\bigotimes_{i' \notin S} E_{i'}\right) + \sum_{S \subset [n], |S| < \epsilon n} \left(\bigotimes_{i \in S} A_i\right) \otimes \left(\bigotimes_{i' \notin S} E_{i'}\right)$$

Let the first term above be N_1 and the second term be N_2. We bound the rank of N_1 and the number of nonzero entries in each row and column of N_2. Note that by grouping terms in the sum for N_1, we can find matrices E_S for all $S \subset [n]$ with $|S| = \epsilon n$ and write

$$N_1 = \sum_{S \subset [n], |S| = \epsilon n} \left(\bigotimes_{i \in S} A_i\right) \otimes E_S$$

Now we have

$$\mathsf{rank}(N_1) \leq \sum_{S \subset [n], |S| = \epsilon n} P \prod_{i \in S} \frac{1}{t_i^{\frac{a_i \epsilon^4}{t_i^2 \log t_i}}} \leq \binom{n}{\epsilon n} \frac{P}{\left(B^{\frac{\epsilon^4}{t_n^2 \log t_n}}\right)^{\epsilon n}} \leq \frac{(n)^{\epsilon n}}{\left(\frac{\epsilon n}{3}\right)^{\epsilon n}} \frac{P}{\left(B^{\frac{\epsilon^4}{t_n^2 \log t_n}}\right)^{\epsilon n}} \leq \left(\frac{3}{\epsilon}\right)^{\epsilon n} \frac{P}{B^{\frac{\epsilon^5 n}{t_n^2 \log t_n}}}$$

Note that $B^{\frac{\epsilon^4}{t_n^2 \log t_n}} \geq t_n^{\frac{a_n \epsilon^4}{2t_n^2 \log t_n}} \geq \max\left(t_n^{0.5 \log t_n}, t_n^{\frac{D\epsilon^4}{2t_n^2 \log t_n}}\right)$. Either the first term is larger than $\epsilon^{\frac{-100}{\epsilon}}$ or t_n is bounded above by some function of ϵ in which case if we choose D sufficiently large, the second term will be larger than $\epsilon^{\frac{-100}{\epsilon}}$. In any case we get

$$\mathsf{rank}(N_1) \leq \frac{P}{\left(\frac{\epsilon}{3} B^{\frac{\epsilon^4}{t_n^2 \log t_n}}\right)^{\epsilon n}} \leq \frac{P}{B^{\frac{\epsilon^5 n}{2t_n^2 \log t_n}}} \leq P^{1 - \frac{\epsilon^5}{4t_n^2 \log t_n}}$$

Now we bound the number of nonzero entries in each row and column of N_2. This number is at most

$$2^n B^{2\epsilon n} P^{\epsilon} \leq 2^n P^{3\epsilon} \leq P^{4\epsilon}$$

Thus, when we have $B \leq t_1^{a_1}, \dots, t_n^{a_n} \leq B^2$,

$$\mathsf{r}_A\left(P^{1 - \frac{\epsilon^5}{4t_n^2 \log t_n}}\right) \leq P^{4\epsilon}$$

Now we move on to the case where we no longer have control over the range of values $t_1^{a_1}, \dots, t_n^{a_n}$. Fix $k = 2^D$ and consider the intervals $I_1 = [k, k^2), I_2 = [k^2, k^4), \dots, I_j = [k^{2^{j-1}}, k^{2^j}), \dots$ and so on. Note

$$A = \bigotimes_{i \in [L]} \left(\bigotimes_{t_j^{a_j} \in I_i} (\underbrace{F_{t_j} \otimes \cdots \otimes F_{t_j}}_{a_j})\right)$$

For an integer i, let $P_i = \prod_{t_j^{a_j} \in I_i} t_j^{a_j}$. Let T be the set of indices $i \in [L]$ such that $P_i \geq P^{\frac{\epsilon}{2L}}$. Then

$$A = \left(\bigotimes_{i \in T} \left(\bigotimes_{t_j^{a_j} \in I_i} (\underbrace{F_{t_j} \otimes \cdots \otimes F_{t_j}}_{a_j})\right)\right) \otimes \left(\bigotimes_{i \notin T} \left(\bigotimes_{t_j^{a_j} \in I_i} (\underbrace{F_{t_j} \otimes \cdots \otimes F_{t_j}}_{a_j})\right)\right) = B \otimes C$$

(where naturally B denotes the first term and C denotes the second)

Note that the dimension of the matrix C is at most $\left(P^{\frac{\epsilon}{2L}}\right)^L = P^{\frac{\epsilon}{2}}$. We now apply Lemma 4.9 repeatedly to bound the rigidity of B. Let $B_i = \left(\bigotimes_{t_j^{a_j} \in I_i} \underbrace{(F_{t_j} \otimes \cdots \otimes F_{t_j})}_{a_j}\right)$.

$$\mathsf{r}_B\left(\left(\prod_{i\in T} P_i\right)\left(\sum_{i\in T}\frac{1}{P_i^{\frac{\epsilon^5}{4t_n^2 \log t_n}}}\right)\right) \le \left(\prod_{i\in T} P_i\right)^{4\epsilon}$$

From the above statements about B, C we deduce that

$$\mathsf{r}_A\left(P^{1-\frac{\epsilon^6}{10Lt_n^2 \log t_n}}\right) \le \mathsf{r}_A\left(LP^{1-\frac{\epsilon^6}{8Lt_n^2 \log t_n}}\right) \le P^{5\epsilon}$$

□

4.2 Proof of Theorem 4.7

To complete the proof of Theorem 4.7, we will break F_N into submatrices, show that each submatrix is non-rigid using techniques from the previous section, and then combine our estimates to conclude that F_N is non-rigid. Recall that N is (l,x)-factorable with $\frac{x}{(\log x)^{C_0+100}} \le l \le \frac{x}{(\log x)^{C_0+10}}$, meaning $N = q_1 q_2 \dots q_l$ for some distinct primes $q_1, \dots, q_l$ where $q_i - 1$ has no large prime power divisors for all i. Let γ be a primitive N^{th} root of unity.

Definition 4.11. *For a subset $S \subset [l]$ define $\mathsf{mult}_N(S) = \prod_{s\in S} q_s$ and $\mathsf{fact}_N(S) = \prod_{s \in S}(q_s - 1)$.*

Definition 4.12. *For all $S \subset [l]$ we will define T_S as the subset of $[N] \times [N]$ indexed by (i,j) such that*

$$\begin{aligned} ij &\not\equiv 0 \mod q_s \ \forall s \in S \\ ij &\equiv 0 \mod q_s \ \forall s \notin S \end{aligned}$$

Note that as S ranges over all subsets of $[l]$, the sets T_S form a partition of $[N] \times [N]$.

For each S, we will divide the set T_S into submatrices such that when filled with the corresponding entries of F_N, we can apply Lemma 4.10 to show that each submatrix is nonrigid. The key intuition is that for a given prime q_i, once we restrict to nonzero residues, the multiplicative subgroup actually has the additive structure of $\mathbb{Z}_{q_i - 1}$. Since $q_i - 1$ is smooth, $\mathbb{Z}_{q_i-1}$ is a direct sum of cyclic groups of small order.

Definition 4.13. *For all $S \subset [l]$, we define the $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ matrix $M(S)$ as follows. Let R_S be the set of residues modulo $\mathsf{mult}_N(S)$ that are relatively prime to $\mathsf{mult}_N(S)$. Note that $|R_S| = \mathsf{fact}_N(S)$. Each row and each column of $M(S)$ is indexed by an element of R_S and the entry in row i and column j is $\theta^{i\cdot j}$ where θ is a primitive $\mathsf{mult}_N(S)$ root of unity. The exact order of the rows and columns will not matter for our uses. Note that replacing θ with θ^k for k relatively prime to $\mathsf{mult}_N(S)$ simply permutes the rows so it does not matter which root of unity we choose.*

Lemma 4.14. *Consider the set of entries in F_N indexed by elements of T_S. We can partition this set into $\prod_{s\notin S}(2q_s - 1)$ submatrices each of size $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ that are equivalent to $M(S)$ up to some permutation of rows and columns.*

Proof. In T_S, for each prime q_s with $s \notin S$, there are $2q_s - 1$ choices for what i and j are $\mod q_s$. Now fix the choice of $i, j \mod q_s$ for all $s \notin S$. Say we restrict to indices with $i \equiv c_1 \mod \prod_{s\notin S} q_s$ and $j \equiv c_2 \mod \prod_{s \notin S} q_s$.

We are left with a $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ matrix, call it A, where i and j run over all residues modulo $\mathsf{mult}_N(S)$ that are relatively prime to $\mathsf{mult}_N(S)$. Naturally, label all rows and columns of this matrix by what the corresponding indices i and j are modulo $\mathsf{mult}_N(S)$. For a row labeled a and a column labeled b, we compute the entry A_{ab}. The value is $\gamma^{a'\cdot b'}$ where a' is the unique element of $\mathbb{Z}_N$ such that $a' \equiv a \mod \mathsf{mult}_N(S)$ and $a' \equiv c_1 \mod \prod_{s\notin S} q_s$ and b' is defined similarly. We have

$$a' \cdot b' \equiv ab \mod \mathsf{mult}_N(S)$$

$$a' \cdot b' \equiv c_1 c_2 \equiv 0 \mod \prod_{s\notin S} q_s$$

Therefore

$$a'b' \equiv k \prod_{s\notin S} q_s ab \mod \mathsf{mult}_N(S)$$

where k is defined as an integer such that $k\prod_{s\notin S} q_s \equiv 1 \mod \mathsf{mult}_N(S)$. Note that k clearly exists since $\prod_{s\notin S} q_s$ and $\mathsf{mult}_N(S)$ are relatively prime. Since $\gamma^{k\prod_{s\notin S} q_s}$ is a primitive $\mathsf{mult}_N(S)$ root of unity, the matrix A is equivalent to $M(S)$ up to some permutation, as desired. $\square$

Lemma 4.15. *For a subset $S \subset [l]$ with $|S| = k$ and $M(S)$ (as defined in Definition 4.13) a $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ matrix as described above. we have*

$$r_{M(S)}\left(\frac{\mathsf{fact}_N(S)}{2^{\epsilon^6} x^{0.37}}\right) \le (\mathsf{fact}_N(S))^{6\epsilon}$$

as long as $k \ge \frac{x}{(\log x)^{C_0+200}}$

Proof. WLOG $S = \{1, 2, \dots, k\}$. Consider the factorizations of $q_1 - 1, \dots, q_k - 1$ into prime powers. For each prime power $p_i^{e_i} \le x^{0.3}$, let $c(p_i^{e_i})$ be the number of indices j for which $p_i^{e_i}$ appears (exactly) in the factorization of $q_j - 1$. Consider all prime powers $p_i^{e_i}$ for which $c(p_i^{e_i}) < x^{0.62}$.

$$\prod_{t, c(t) \le x^{0.62}} t^{c(t)} \le \left((x^{0.3})^{x^{0.62}}\right)^{x^{0.3}} \le x^{x^{0.92}}$$

Now consider all prime powers say $t_1, \dots, t_n$ for which $c(t_i) \ge x^{0.62}$. Let $P = t_1^{c(t_1)} \dots t_n^{c(t_n)}$. From the above we know that as long as x is sufficiently large

$$P \ge \frac{\mathsf{fact}_N(S)}{x^{x^{0.92}}} \ge (\mathsf{fact}_N(S))^{(1-\epsilon)} \frac{\left(\frac{x}{(\log x)^{C_0+1}}\right)^{\epsilon k}}{x^{x^{0.92}}} \ge (\mathsf{fact}_N(S))^{(1-\epsilon)} \tag{2}$$

We will use the prime powers t_i and Theorem 3.1 to show that $M(S)$ is not rigid. Note that we can associate each row and column of $M(S)$ to a k-tuple $(a_1, \dots, a_k)$ where $a_i \in \mathbb{Z}_{q_i - 1}$ as follows. First, it is clear that each row and column of $M(S)$ can be associated to a k-tuple $(z_1, \dots, z_k) \in \mathbb{Z}^*_{q_1} \times \dots \times \mathbb{Z}^*_{q_k}$. Now $\mathbb{Z}^*_{q_i}$ can be viewed as a cyclic group on $q_i - 1$ elements. This allows us to create a bijection between the rows and columns of $M(S)$ and elements of $\mathbb{Z}_{q_1-1} \times \dots \times \mathbb{Z}_{q_k-1}$.

Also note that for a row indexed by $A = (a_1, \dots, a_k)$ and a column indexed by $B = (b_1, \dots, b_k)$, the entry $M(S)_{AB}$ is dependent only on $A + B$. We will now decompose $M(S)$ into several $P \times P$ submatrices. In particular, we can write $q_i - 1 = d_i T_i$ where T_i is a product of some subset of $\{t_1, \dots, t_n\}$ and d_i is relatively prime to T_i. We have $T_1 T_2 \dots T_k = P$. For each $A', B' \in \mathbb{Z}_{d_1} \times \dots \times \mathbb{Z}_{d_k}$, we can construct a $P \times P$ submatrix $M(S, A', B')$ consisting of all entries $M(S)_{AB}$ of $M(S)$ such that $A \equiv A', B \equiv B'$ (where the equivalence is over $\mathbb{Z}_{d_1} \times \dots \times \mathbb{Z}_{d_k}$). This gives us d^2 different submatrices where $d = d_1 \dots d_k$. Naturally, we can associate each row and column of a submatrix $M(S, A', B')$ with an element of $\mathbb{Z}_{T_1} \times \cdots \times \mathbb{Z}_{T_k}$ such that for a row labeled I and a column labeled J, the entry $M(S, A', B')_{IJ}$ only depends on $I + J$. In particular, this means that $X\left(M(S, A', B')\right) X$ is diagonal where $X = F_{T_1} \otimes \cdots \otimes F_{T_k}$. Now, using Lemma 4.8, we can rewrite

$$X = (\underbrace{F_{t_1} \otimes \cdots \otimes F_{t_1}}_{c(t_1)}) \otimes \cdots \otimes (\underbrace{F_{t_n} \otimes \cdots \otimes F_{t_n}}_{c(t_n)})$$

Since for x sufficiently large, $c(t_i) \geq x^{0.62} \geq \frac{t_i^2 (\log t_i)^2}{\epsilon^{10}}$, we can use Lemma 4.10 and get that

$$\mathsf{r}_X\left(P^{1-\frac{\epsilon^6}{20(\log\log P)x^{0.62}}}\right) \leq P^{5\epsilon}$$

Let E be the matrix of changes to reduce the rank of X according to the above. We have that E has at most P^{ϵ} nonzero entries in each row and column, and

$$\mathsf{rank}(X - E) \leq P^{1-\frac{\epsilon^6}{20(\log\log P)x^{0.62}}}$$

We can write $M(S)$ in block form as

$$\begin{bmatrix} M(S,A_1,B_1) & M(S,A_1,B_2) & \dots & M(S,A_1,B_d) \\ M(S,A_2,B_1) & M(S,A_2,B_2) & \dots & M(S,A_1,B_d) \\ \vdots & \vdots & \ddots & \vdots \\ M(S,A_d,B_1) & M(S,A_d,B_2) & \dots & M(S,A_d,B_d) \end{bmatrix}$$

where $A_1, \ldots, A_d$ and $B_1, \ldots, B_d$ range over the elements of $\mathbb{Z}_{d_1} \times \cdots \times \mathbb{Z}_{d_k}$. We can rearrange the above as

$$\begin{bmatrix} M(S,A_1,B_1) & \dots & M(S,A_1,B_d) \\ \vdots & \ddots & \vdots \\ M(S,A_d,B_1) & \dots & M(S,A_d,B_d) \end{bmatrix} = \begin{bmatrix} XD_{11}X & \dots & XD_{1d}X \\ \vdots & \ddots & \vdots \\ XD_{d1}X & \dots & XD_{dd}X \end{bmatrix}$$

where the D_{ij} are diagonal matrices. Now consider the matrix

$$E(S) = \begin{bmatrix} ED_{11}E & \dots & ED_{1d}E \\ \vdots & \ddots & \vdots \\ ED_{d1}E & \dots & ED_{dd}E \end{bmatrix}$$

We have

$$M(S) - E(S) = \begin{bmatrix} XD_{11}X - ED_{11}E & \dots & XD_{1d}X - ED_{1d}E \\ \vdots & \ddots & \vdots \\ XD_{d1}X - ED_{d1}E & \dots & XD_{dd}X - ED_{dd}E \end{bmatrix} =$$

$$\begin{bmatrix} XD_{11}(X-E) & \dots & XD_{1d}(X-E) \\ \vdots & \ddots & \vdots \\ XD_{d1}(X-E) & \dots & XD_{dd}(X-E) \end{bmatrix} + \begin{bmatrix} (X-E)D_{11}E & \dots & (X-E)D_{1d}E \\ \vdots & \ddots & \vdots \\ (X-E)D_{d1}E & \dots & (X-E)D_{dd}E \end{bmatrix}$$

In the above expression, each of the two terms has rank at most

$$dP^{1-\frac{\epsilon^6}{20(\log\log P)x^{0.62}}} = \frac{\mathsf{fact}_N(S)}{P^{\frac{\epsilon^6}{20(\log\log P)x^{0.62}}}} \leq \frac{1}{2}\left(\frac{\mathsf{fact}_N(S)}{2^{\epsilon^6 x^{0.37}}}\right)$$

Note that when computing the rank, we only multiply by d (and not d^2) because the small blocks are all multiplied by the same low rank matrix on either the left or right. The number of nonzero entries in each row and column of $E(S)$ is at most $P^{5\epsilon}d = \frac{\mathsf{fact}_N(S)}{P^{1-5\epsilon}}$. Since $P \geq (\mathsf{fact}_N(S))^{1-\epsilon}$, we conclude

$$\mathsf{r}_{M(S)}\left(\frac{\mathsf{fact}_N(S)}{2^{\epsilon^6 x^{0.37}}}\right) \leq (\mathsf{fact}_N(S))^{6\epsilon}$$

□

We are now ready to complete the analysis of the non-rigidity of the Fourier transform matrix F_N.

Proof of Theorem 4.7. Set the threshold $m = x^{0.365}$ and $k_0 = l - m$. The sets T_S, as S ranges over all subsets of $[l]$, form a partition of $[N] \times [N]$. For each $S \subset [l]$ with $|S| \geq k_0$, we will divide T_S into $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ submatrices using Lemma 4.14 and change entries to reduce the rank of every submatrix according to Lemma 4.15. We will not touch the entries in sets T_S for $|S| < k_0$. Call the resulting matrix M'. We now estimate the rank of M' and then the maximum number of entries changed in any row or column.

We remove all rows and columns corresponding to integers divisible by at least $\frac{m}{2}$ of the primes $q_1, \dots, q_l$. The number of rows and columns removed is at most

$$N\left(\sum_{S \subset [l], |S| = \frac{m}{2}} \prod_{i \in S} \frac{1}{q_i}\right) \leq \frac{N}{\left(\frac{x}{(\log x)^{C_0}}\right)^{\frac{m}{2}}} \binom{l}{\frac{m}{2}} < N\left(\frac{l}{\frac{x}{(\log x)^{C_0}}}\right)^{\frac{m}{2}} \leq \frac{N}{(\log x)^{x^{0.365}}}$$

The remaining entries must be subdivided into matrices of the form $M(S)$ for various subsets $S \subset [l]$, $|S| \geq k_0$. Say $q_1 < q_2 < \cdots < q_l$. The number of such submatrices is at most

$$\frac{N^2}{((q_1 - 1)\dots(q_{k_0} - 1))^2} \leq (q_{k_0+1} \dots q_l)^2 \left(\frac{q_1 \dots q_{k_0}}{(q_1 - 1)\dots(q_{k_0} - 1)}\right)^2 \leq 3(q_{k_0+1} \dots q_l)^2 \leq 3x^{2m}$$

Each one of the submatrices has rank at most

$$\frac{N}{2^{\epsilon^6 x^{0.37}}}$$

so in total the rank is at most

$$N\frac{3x^{2m}}{2^{\epsilon^6 x^{0.37}}} \leq \frac{N}{2^{\epsilon^6 x^{0.369}}}$$

Combining the two parts we easily get

$$\mathsf{rank}(M') \leq \frac{N}{2^{\epsilon^6 x^{0.365}}}$$

Now we bound the number of entries changed. The number of entries changed in each row or column is at most

$$\frac{N}{((q_1 - 1)\dots(q_{k_0} - 1))} N^{6\epsilon} \leq (q_{k_0+1} \dots q_l)\left(\frac{q_1 \dots q_{k_0}}{(q_1 - 1)\dots(q_{k_0} - 1)}\right) N^{6\epsilon} \leq 3N^{6\epsilon + 1.1\frac{m}{l}} \leq N^{7\epsilon}$$

As $2^{\epsilon^6 x^{0.365}} \geq 2^{\epsilon^6 (\log N)^{0.36}}$ for sufficiently large x, we conclude

$$\mathsf{r}_{F_N}\left(\frac{N}{2^{\epsilon^6 (\log N)^{0.36}}}\right) \leq N^{7\epsilon}$$

□

5 Non-rigidity of All Circulant Matrices

In the previous section, we showed that there exists an infinite set of Fourier matrices that are not Valiant-rigid. In this section, we will bootstrap the results from Section 4 to show that in fact, all sufficiently large Fourier matrices are not rigid.

The first ingredient will be a stronger form of Lemma 4.5. Recall that a prime q is defined to be x-good if $\frac{x}{(\log x)^{C_0}} \leq q \leq x$ and all prime powers dividing $q-1$ are at most $x^{0.3}$ and that an integer N is defined to be (l,x)-factorable if it can be written as the product of l distinct x-good primes.

Lemma 5.1. *For all sufficiently large integers K, there exist l, x, N such that $\frac{x}{(\log x)^{C_0+100}} \leq l \leq \frac{x}{(\log x)^{C_0+10}}$, N is (l,x)-factorable, and $K < N < K(\log K)^2$.*

Proof. Call an N well-factorable if it is (l,x)-factorable for some x and $\frac{x}{(\log x)^{C_0+100}} \leq l \leq \frac{x}{(\log x)^{C_0+10}}$. Let N_0 be the largest integer that is well-factorable with $N_0 \leq K$. Say N_0 is (l,x)-factorable.

We have $N_0 = q_1 \dots q_l$ where $q_1, \dots, q_l$ are distinct, x-good primes. If $l < \lfloor \frac{x}{(\log x)^{C_0+10}} \rfloor$ then by Lemma 4.5, we can find another x-good prime q_{l+1}. We can then replace N_0 with $q_{l+1}N_0$. $q_{l+1}N_0 > K$ by the maximality of N_0 and also $q_{l+1}N_0 \leq N_0 x \leq N_0(\log N_0)^2$ so $q_{l+1}N_0$ satisfies the desired properties.

We now consider the case where $l = \lfloor \frac{x}{(\log x)^{C_0+10}} \rfloor$. First, if $q_1, \dots, q_l$ are not the l largest x-good primes then we can replace one of them say q_1 with $q_1' > q_1$. The number $N' = q_1' q_2 \dots q_l$ is well-factorable and between N_0 and $N_0(\log x)^{C_0}$. Using the maximality of N_0, we deduce that N' must be in the desired range.

On the other hand if $q_1, \dots, q_l$ are the l largest x-good primes, we know they are actually all between $\frac{3x}{(\log x)^{C_0}}$ and x. This is because by Lemma 4.5, there are at least $\frac{10x}{(\log x)^{C_0}}$ distinct x-good primes. Let $x' = 2x$. The above implies that $q_1, \dots, q_l$ are x'-good and clearly $\frac{x'}{(\log x')^{C_0+100}} \leq l \leq \frac{x'}{(\log x')^{C_0+10}}$. Furthermore, $\frac{x'}{(\log x')^{C_0+10}} > \frac{x}{(\log x)^{C_0+10}} + 1$ so $l = \lfloor \frac{x}{(\log x)^{C_0+10}} \rfloor < \lfloor \frac{x'}{(\log x')^{C_0+10}} \rfloor$ and we can now repeat the argument from the first case. $\square$

We can now complete the proof that all circulant matrices are not rigid.

Theorem 5.2. *Let $0 < \epsilon < 0.1$ be a given parameter. For all sufficiently large N, if M is an $N \times N$ adjusted-circulant (or Hankel) matrix*

$$\mathsf{r}_M\left(\frac{N}{2^{\epsilon^6(\log N)^{0.35}}}\right) \leq N^{15\epsilon}$$

Proof. First we analyze circulant matrices of size N_0 where N_0 is (l,x)-factorable for some $\frac{x}{(\log x)^{C_0+100}} \leq l \leq \frac{x}{(\log x)^{C_0+10}}$. Theorem 4.7 and Lemma 2.18 imply that for M_0 an $N_0 \times N_0$ circulant matrix where N_0 satisfies the previously mentioned properties,

$$\mathsf{r}_{M_0}\left(\frac{2N_0}{2^{\epsilon^6(\log N_0)^{0.36}}}\right) \leq N_0^{14\epsilon}$$

Now for a circulant matrix M of arbitrary size $N \times N$, note that it is possible to embed an M in the upper left corner of a circulant matrix of any size at least $2N$. By Lemma 5.1, there exists an N_0 that is (l,x)-factorable for some $\frac{x}{(\log x)^{C_0+100}} \leq l \leq \frac{x}{(\log x)^{C_0+10}}$ such that

$$\frac{N_0}{(\log N_0)^2} \leq N \leq \frac{N_0}{2}$$

We deduce

$$\mathsf{r}_M^{\mathbb{F}_q}\left(\frac{2N_0}{2^{\epsilon^6(\log N_0)^{0.36}}}\right) \leq N_0^{14\epsilon}$$

Rewriting the bounds in terms of N we get

$$\mathsf{r}_M^{\mathbb{F}_q}\left(\frac{N}{2^{\epsilon^6(\log N)^{0.35}}}\right) \leq N^{15\epsilon}$$

□

Remark. *Note that our proof actually shows something slightly stronger, namely that the changes to reduce the rank of a circulant matrix are actually fixed linear combinations of the entries. See Definition 8.1 and Claim 8.2 for a more precise statement.*

From the above and Claim 2.19, we immediately deduce that all Fourier matrices are not rigid.

Theorem 5.3. *Let $0 < \epsilon < 0.1$ be a given parameter. For all sufficiently large N,*

$$\mathsf{r}_{F_N}\left(\frac{N}{2^{\epsilon^6(\log N)^{0.35}}}\right) \leq N^{15\epsilon}$$

6 Non-rigidity of Group Algebra Matrices for Abelian Groups

Using the results from the previous section, we can show that group algebra matrices for any abelian group are not Valiant-rigid.

Theorem 6.1. *Let $0 < \epsilon < 0.1$ be fixed. Let G be an abelian group and $f : G \to \mathbb{C}$ be a function. Let $M = M_G(f)$ be the adjusted group algebra matrix. If $|G|$ is sufficiently large then*

$$\mathsf{r}_M\left(\frac{2|G|}{2^{\epsilon^8(\log|G|)^{0.32}}}\right) \leq |G|^{38\epsilon}$$

Proof. By the fundamental theorem of finite abelian groups, we can write $G = \mathbb{Z}_{n_1} \oplus \cdots \oplus \mathbb{Z}_{n_a}$. In light of Lemma 2.18, it suffices to bound the rigidity of $F = F_{n_1} \otimes \cdots \otimes F_{n_a}$.

WLOG, $n_1 \leq n_2 \leq \cdots \leq n_a$. We will choose k to be a fixed, sufficiently large positive integer. By Theorem 5.3, we can ensure that for $N > k$

$$\mathsf{r}_{F_N}\left(\frac{N}{2^{\epsilon^6(\log N)^{0.35}}}\right) \leq N^{15\epsilon}$$

Consider the ranges $I_1 = [k, k^2), I_2 = [k^2, k^4), \dots I_j = [k^{2^{j-1}}, k^{2^j}) \dots$ and so on. Let S_j be a multiset defined by $S_j = I_j \cap \{n_1, \dots, n_a\}$. Fix a j and say the elements of S_j are $x_1 \leq \cdots \leq x_b$. By Theorem 5.3, for each x_i, there are matrices E_{x_i} and A_{x_i} such that $F_{x_i} = A_{x_i} + E_{x_i}$, E_{x_i} has at most $x_i^{15\epsilon}$ nonzero entries in each row and column, and

$$\mathsf{rank}(A_{x_i}) \leq \frac{x_i}{2^{\epsilon^6(\log x_i)^{0.35}}}$$

Now we can write

$$M_j = F_{x_1} \otimes \cdots \otimes F_{x_b} = (A_{x_1} + E_{x_1}) \otimes \cdots \otimes (A_{x_b} + E_{x_b}) = \sum_{S \subset [b]} \left(\bigotimes_{i \in S} A_{x_i}\right) \otimes \left(\bigotimes_{i' \notin S} E_{x'_i}\right)$$

$$= \sum_{S \subset [b], |S| \geq \epsilon b} \left(\bigotimes_{i \in S} A_{x_i}\right) \otimes \left(\bigotimes_{i' \notin S} E_{x'_i}\right) + \sum_{S \subset [b], |S| < \epsilon b} \left(\bigotimes_{i \in S} A_{x_i}\right) \otimes \left(\bigotimes_{i' \notin S} E_{x'_i}\right)$$

Let the first term above be N_1 and the second term be N_2. We will bound the rank of N_1 and the number of nonzero entries in each row and column of N_2. Note that by grouping the terms in the sum for N_1 we can write it in the form

$$\sum_{S\subset[b],|S|=\lceil \epsilon b\rceil} \bigotimes_{i\in S} A_{x_i} \otimes E_S$$

where for each S, E_S is some matrix. This implies that

$$\mathsf{rank}(N_1) \leq \binom{b}{\lceil \epsilon b\rceil} \frac{x_1 \dots x_b}{\left(2^{\epsilon^6 (\log x_1)^{0.35}}\right)^{\lceil \epsilon b\rceil}} \leq \frac{b^{\lceil \epsilon b\rceil}}{(\frac{\epsilon b}{3})^{\lceil \epsilon b\rceil}} \frac{x_1 \dots x_b}{\left(2^{\epsilon^6 (\log x_1)^{0.35}}\right)^{\lceil \epsilon b\rceil}} = x_1 \dots x_b \left(\frac{3}{\epsilon 2^{\epsilon^6 (\log x_1)^{0.35}}}\right)^{\lceil \epsilon b\rceil}$$

As long as k is sufficiently large, we have

$$\mathsf{rank}(N_1) \leq x_1 \dots x_b \left(\frac{3}{\epsilon 2^{\epsilon^6 (\log x_1)^{0.35}}}\right)^{\lceil \epsilon b\rceil} \leq x_1 \dots x_b \left(\frac{1}{2^{\epsilon^6 (\log x_1)^{0.34}}}\right)^{\lceil \epsilon b\rceil} \leq \frac{x_1 \dots x_b}{2^{\epsilon^7 (\log x_1 \dots x_b)^{0.33}}}$$

where in the last step we used the fact that $x_i \leq x_1^2$ for all i. The number of nonzero entries in each row or column of N_2 is at most

$$2^b x_b \dots x_{b-\lfloor \epsilon b\rfloor+1} (x_{b-\lfloor \epsilon b\rfloor} \dots x_1)^{15\epsilon} = 2^b (x_1 \dots x_b)^{15\epsilon} (x_b \dots x_{b-\lfloor \epsilon b\rfloor+1})^{1-15\epsilon} \leq (x_1 \dots x_b)^{18\epsilon}$$

Note in the last step above, we used the fact that $x_i \leq x_1^2$.

For each integer c between 2 and k, let n_c be the number of copies of c in the set $\{n_1, \dots, n_a\}$. If $n_c \geq \frac{k^2(\log k)^2}{\epsilon^4}$ then by Theorem 3.1, if we define $A_c = \underbrace{F_c \otimes \dots \otimes F_c}_{n_c}$ then

$$\mathsf{r}_{A_c}\left(c^{n_c(1-\frac{\epsilon^4}{k^2 \log k})}\right) \leq c^{n_c \epsilon}$$

Let $L = \lceil 2 \log\log |G| \rceil$ and ensure that $|G|$ is sufficiently large so that $L > k$. Let T be the set of integers c between 2 and k such that $c^{n_c} \geq |G|^{\frac{\epsilon}{2L}}$ (note that as long as $|G|$ is sufficiently large, all elements of T must satisfy $n_c \geq \frac{k^2(\log k)^2}{\epsilon^4}$). Let R be the set of indices j for which $\prod_{x\in S_j} x \geq |G|^{\frac{\epsilon}{2L}}$. Since S_j is clearly empty for $j \geq L$, the matrix F can be written as

$$F = \left(\bigotimes_{2\leq c<k} \left(\underbrace{F_c \otimes \dots \otimes F_c}_{n_c}\right)\right) \otimes \left(\bigotimes_{1\leq j\leq L} M_j\right)$$

Define

$$B = \left(\bigotimes_{c\notin T} \left(\underbrace{F_c \otimes \dots \otimes F_c}_{n_c}\right)\right) \otimes \left(\bigotimes_{j\notin R} M_j\right)$$

Note that the size of B is at most $\left(|G|^{\frac{\epsilon}{2L}}\right)^{k+L} \leq |G|^{\epsilon}$. Also $F = B \otimes D$ where

$$D = \left(\bigotimes_{c\in T} \left(\underbrace{F_c \otimes \dots \otimes F_c}_{n_c}\right)\right) \otimes \left(\bigotimes_{j\in R} M_j\right)$$

For any rank r, $\mathsf{r}_M(|B|r) \leq |B| \mathsf{r}_D(r)$. Applying Lemma 4.9 iteratively, we get

$$\mathsf{r}_D\left(\frac{|G|}{|B|}\left(\sum_{c\in T} \frac{1}{c^{n_c \frac{\epsilon^4}{k^2 \log k}}} + \sum_{j\in R} \frac{1}{2^{\epsilon^7 (\log \prod_{x\in S_j} x)^{0.33}}}\right)\right) \leq \left(\frac{|G|}{|B|}\right)^{18\epsilon}$$

Note that

$$\left(\sum_{c\in T}\frac{1}{c^{n_c\frac{\epsilon^4}{k^2\log k}}}+\sum_{j\in R}\frac{1}{2^{\epsilon^7(\log\prod_{x\in S_j}x)^{0.33}}}\right)\le\frac{k}{|G|^{\frac{\epsilon^5}{2Lk^2\log k}}}+\frac{L}{2^{\epsilon^8\left(\frac{\log|G|}{2L}\right)^{0.33}}}\le\frac{1}{2^{\epsilon^8(\log|G|)^{0.32}}}$$

Overall, we conclude

$$\mathsf{r}_F\left(\frac{|G|}{2^{\epsilon^8(\log|G|)^{0.32}}}\right)\le|B|\left(\frac{|G|}{|B|}\right)^{18\epsilon}\le|G|^{19\epsilon}$$

Since FMF is diagonal, Lemma 2.18 gives the desired. □

7 Finite Field Case

In this section, we sketch how to modify the proofs in the previous sections to deal with matrices over a finite field. The main difficulty that arises when attempting to extend the above methods to finite fields is that the entries of the corresponding Fourier matrix might not exist in the field. Furthermore, for a finite field $\mathbb{F}_q$ and integer k with $\gcd(k,q)>1$, there are no primitive k^{th} roots of unity over any extension of $\mathbb{F}_q$. The first lemma in this section allows us to lift to a field extension and then argue that if a matrix is highly non-rigid over some low-degree extension then it also cannot be rigid over the base field.

Lemma 7.1. *Consider a finite field $\mathbb{F}_q$ and some algebraic extension $\mathbb{F}_q[\gamma]$ where γ is some primitive d^{th} root of unity with $\gcd(q,d)=1$. If the degree of the minimal polynomial of γ is g then for any matrix $M\in\mathbb{F}_q^{n\times n}$ and any positive integer r,*

$$\mathsf{r}_M^{\mathbb{F}_q}(gr)\le\mathsf{r}_M^{\mathbb{F}_q[\gamma]}(r)$$

Proof. Say the conjugates of γ are $\gamma_1,\dots,\gamma_g$ where $\gamma_1=\gamma$. Let S be the set of all primitive d^{th} roots of unity. First we show that there exists an integer k such that $\gamma_1^k+\cdots+\gamma_g^k\neq 0$. Let the prime factorization of d be $p_1^{e_1}\dots p_m^{e_m}$ and set $c=p_1p_2\dots p_m$. Note that $\sum_{x\in S}x^{\frac{d}{c}}$ is equal to $\frac{d}{c}$ times the sum of all primitive c^{th} roots of unity. The sum of all $\phi(c)$ primitive c^{th} roots of unity is either 1 or -1 (and is thus nonzero) and thus $\sum_{x\in S}x^{\frac{d}{c}}$ is nonzero. Also, the set S can be partitioned into exactly $\frac{\phi(d)}{g}$ families of the form $\{\gamma_1^a,\dots,\gamma_g^a\}$ where a is some positive integer. In particular, the sum of the $\frac{d}{c}$-powers of the elements in one of these families must be nonzero so there must be some positive integer k such that $\gamma_1^k+\cdots+\gamma_g^k\neq 0$.

Let $s=\mathsf{r}_M^{\mathbb{F}_q[\gamma]}(r)$. There must be a matrix $E\in\mathbb{F}_q[\gamma]^{n\times n}$ with at most s nonzero entries in each row and column such that $\mathsf{rank}_{\mathbb{F}_q[\gamma]}(M-E)\le r$. Now consider a family of g matrices $E=E_1,\dots,E_g$ obtained by taking E and replacing γ with each of its conjugates. Define the matrix E' as follows

$$E'=\left(\frac{\gamma_1^k}{\gamma_1^k+\cdots+\gamma_g^k}E_1+\cdots+\frac{\gamma_g^k}{\gamma_1^k+\cdots+\gamma_g^k}E_g\right)$$

Note that $\gamma_1^k+\cdots+\gamma_g^k\in\mathbb{F}_q$ and also $\gamma_1^kE_1+\cdots+\gamma_g^kE_g\in\mathbb{F}_q^{n\times n}$ so $E'\in\mathbb{F}_q^{n\times n}$ and E' clearly has at most s nonzero entries in each row and column. Next

$$M-E'=\frac{1}{\gamma_1^k+\cdots+\gamma_g^k}\left(\gamma_1^k(M-E_1)+\cdots+\gamma_g^k(M-E_g)\right)$$

so $\mathsf{rank}_{\mathbb{F}_q}(M-E')\le gr$. Writing $M=(M-E')+E'$, we immediately get the desired conclusion.

□

Following the proof of Lemma 4.10, we can prove the following analog over finite fields.

Lemma 7.2. *Let $0 < \epsilon < 0.1$ be some chosen parameter, $\mathbb{F}_q$ be a fixed finite field, and D be some sufficiently large constant (possibly depending on ϵ and q). Consider positive integers $t_1 \le t_2 \cdots \le t_n$ with $\gcd(t_i, q) = 1$ for all i. Also assume $a_i \ge \max\left(\frac{t_i^2(\log t_i)^2}{\epsilon^{10}}, D\right)$ for all i. Let $P = t_1^{a_1} \dots t_n^{a_n}$ and $L = \lceil 2 \log\log P \rceil$.*

Consider the field extension $\mathbb{F}_q[\omega_1, \dots, \omega_n]$ where ω_i is a primitive $t_i{}^{th}$ root of unity. Let F_{t_i} be the $t_i \times t_i$ Fourier matrix with entries over the field extension. Let

$$A = (\underbrace{F_{t_1} \otimes \cdots \otimes F_{t_1}}_{a_1}) \otimes \cdots \otimes (\underbrace{F_{t_n} \otimes \cdots \otimes F_{t_n}}_{a_n})$$

Then we have

$$r_A^{\mathbb{F}_q[\omega_1,\dots,\omega_n]}\left(P^{1-\frac{\epsilon^6}{10Lt_n^2 \log t_n}}\right) \le P^{5\epsilon}$$

We also need a slight modification in the proof of Lemma 4.15. The parameters x, N will be set the same way as in Section 4. Recall Definition 4.13:

Definition (Restatement of 4.13)**.** *For all $S \subset [l]$, we define the $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ matrix $M(S)$ as follows. Let R_S be the set of residues modulo $\mathsf{mult}_N(S)$ that are relatively prime to $\mathsf{mult}_N(S)$. Note that $|R_S| = \mathsf{fact}_N(S)$. Each row and each column of $M(S)$ is indexed by an element of R_S and the entry in row i and column j is $\theta^{i \cdot j}$ where θ is a primitive $\mathsf{mult}_N(S)$ root of unity. The exact order of the rows and columns will not matter for our uses. Note that replacing θ with θ^k for k relatively prime to $\mathsf{mult}_N(S)$ simply permutes the rows so it does not matter which root of unity we choose.*

Remark. *Note if γ is a primitive N^{th} root of unity, the matrix $M(S)$ is defined over the extension $\mathbb{F}_q[\gamma]$ for all subsets S.*

Lemma 7.3. *Let γ be a primitive N^{th} root of unity. Let $\omega_1, \dots, \omega_a$ be roots of unity of order $t_1, \dots, t_a$ where $\{t_1, \dots, t_a\}$ is the set of all prime powers at most $x^{0.3}$. For a subset $S \subset [l]$ with $|S| = k$ and $M(S)$ (as defined in Definition 4.13) a $\mathsf{fact}_N(S) \times \mathsf{fact}_N(S)$ matrix. we have*

$$r_{M(S)}^{\mathbb{F}_q[\gamma,\omega_1,\dots,\omega_a]}\left(\frac{\mathsf{fact}_N(S)}{2^{\epsilon^6 x^{0.37}}}\right) \le (\mathsf{fact}_N(S))^{6\epsilon}$$

as long as $k \ge \frac{x}{(\log x)^{C_0+200}}$

Proof Sketch. The only necessary change in the proof is due to the fact that for an integer k with $\gcd(k, q) > 1$, primitive k^{th} roots of unity do not exist over an extension of $\mathbb{F}_q$. To deal with this, we will use a more precise bound than (2) where prime powers not relatively prime to q are also excluded from the product on the LHS.

WLOG $S = \{1, 2, \dots, k\}$. Consider the factorizations of $q_1 - 1, \dots, q_k - 1$ into prime powers. For each prime power $p_i^{e_i} \le x^{0.3}$, let $c(p_i^{e_i})$ be the number of indices j for which $p_i^{e_i}$ appears (exactly) in the factorization of $q_j - 1$. Also let p be the characteristic of the finite field $\mathbb{F}_q$ that we are working over (so q is a power of p). Note that

$$(q_1 - 1) \dots (q_k - 1) = \prod_t t^{c(t)} = p^{c(p)} p^{2c(p^2)} \dots p^{fc(p^f)} \prod_{gcd(t,p)=1} t^{c(t)}$$

where t ranges over all prime powers at most $x^{0.3}$ and p^f is the largest power of p that is at most $x^{0.3}$. For a power of p, say p^i, let $d(p^i)$ be the number of indices j such that $q_j - 1$ is divisible (not necessarily exactly divisible) by p^i. Let $L = \lfloor (1000 + C_0) \log_p \log x \rfloor$

$$\begin{aligned} p^{c(p)} p^{2c(p^2)} \dots p^{fc(p^f)} = p^{d(p)+d(p^2)+\dots+d(p^f)} \le p^{\sum_{i=1}^L d(p^i) + \sum_{i=L+1}^f d(p^i)} &\le p^{Lk + f\frac{x}{(\log x)^{1000+C_0}}} \\ &\le (\log x)^{(1000+C_0)k} x^{\frac{x}{(\log x)^{1000+C_0}}} \end{aligned}$$

Next, consider all prime powers $p_i^{e_i}$ for which $c(p_i^{e_i}) < x^{0.62}$.

$$\prod_{t, c(t) \le x^{0.62}} t^{c(t)} \le \left((x^{0.3})^{x^{0.62}} \right)^{x^{0.3}} \le x^{x^{0.92}}$$

Now WLOG say $t_1, \dots, t_n$ are the set of prime powers for which $\gcd(t_i, p) = 1$ and $c(t_i) \ge x^{0.62}$. Let $P = t_1^{c(t_1)} \dots t_n^{c(t_n)}$. From the above we know that as long as x is sufficiently large

$$P \ge \frac{\mathsf{fact}_N(S)}{x^{x^{0.92}} (\log x)^{(1000+C_0)k} x^{\frac{x}{(\log x)^{1000+C_0}}}} \ge (\mathsf{fact}_N(S))^{(1-\epsilon)} \frac{\left(\frac{x}{(\log x)^{C_0+1}}\right)^{\epsilon k}}{x^{x^{0.92}} (\log x)^{(1000+C_0)k} x^{\frac{x}{(\log x)^{1000+C_0}}}} \ge (\mathsf{fact}_N(S))^{(1-\epsilon)}$$

The remainder of the proof can be completed in the same way as Lemma 4.15. □

Using the above we can show the following analog of Theorem 4.7.

Theorem 7.4. *Let $\mathbb{F}_q$ be a fixed finite field and $0 < \epsilon < 0.1$ be some constant. Let x be sufficiently large and $N = q_1 q_2 \dots q_l$ be a (l, x)-factorable number with $\gcd(N, q) = 1$ and $\frac{x}{(\log x)^{C_0+100}} \le l \le \frac{x}{(\log x)^{C_0+10}}$. Let $t_1, \dots, t_a$ be the set of prime powers at most $x^{0.3}$ that are relatively prime to q. Let $\omega_1, \dots, \omega_a$ be primitive $t_1, \dots, t_a$ roots of unity and let γ be a primitive N^{th} root of unity. Then*

$$\mathsf{r}_{F_N}^{\mathbb{F}_q[\gamma, \omega_1, \dots, \omega_n]} \left(\frac{N}{2^{\epsilon^6 (\log N)^{0.36}}} \right) \le N^{7\epsilon}$$

Combining Theorem 7.4 with Lemma 7.1 we get our main theorem for circulant matrices over finite fields.

Theorem 7.5. *Let $0 < \epsilon < 0.1$ be a given parameter and $\mathbb{F}_q$ be a fixed finite field. For all sufficiently large N, if M is an $N \times N$ adjusted-circulant (or Hankel) matrix*

$$\mathsf{r}_M^{\mathbb{F}_q} \left(\frac{N}{2^{\epsilon^6 (\log N)^{0.35}}} \right) \le N^{15\epsilon}$$

Proof. First we analyze circulant matrices of size N_0 where N_0 is (l, x)-factorable for some $\frac{x}{(\log x)^{C_0+100}} \le l \le \frac{x}{(\log x)^{C_0+10}}$. Theorem 4.7 and Lemma 2.18 imply that for M_0 an $N_0 \times N_0$ circulant matrix where N_0 satisfies the previously mentioned properties,

$$\mathsf{r}_{M_0}^{\mathbb{F}_q[\gamma, \omega_1, \dots, \omega_a]} \left(\frac{2N_0}{2^{\epsilon^6 (\log N_0)^{0.36}}} \right) \le N_0^{14\epsilon}$$

Now we analyze the degree of the extension $\mathbb{F}_q[\gamma, \omega_1, \dots, \omega_a]$. Note $\mathbb{F}_q[\gamma, \omega_1, \dots, \omega_a] \subset \mathbb{F}_q[\eta]$ where η is a primitive root of unity of degree $C = N_0 \operatorname{lcm}(t_1, t_2, \dots t_a)$. The degree of the extension $\mathbb{F}_q[\eta]$ is the order of q modulo C. N_0 factors into a product of distinct x-good primes and all prime powers dividing $\operatorname{lcm}(t_1, t_2, \dots t_a)$ are at most $x^{0.3}$. Thus for any prime power r dividing C, the order of q modulo r divides $(x^{0.3})!$. Overall, the order of $q \mod C$ is at most

$$(x^{0.3})! < x^{0.3x^{0.3}} \le 2^{(\log N_0)^{0.31}}$$

Thus the degree of the extension $\mathbb{F}_q[\gamma, \omega_1, \dots, \omega_a]$ is at most $2^{(\log N_0)^{0.31}}$. By Lemma 7.1

$$\mathsf{r}_{M_0}^{\mathbb{F}_q} \left(\frac{N_0}{2^{\epsilon^6 (\log N_0)^{0.359}}} \right) \le N_0^{14\epsilon}$$

Now for a circulant matrix M of arbitrary size $N \times N$, note that it is possible to embed an M in the upper left corner of a circulant matrix of any size at least $2N$. By Lemma 5.1, there exists an N_0 that is (l, x)-factorable for some $\frac{x}{(\log x)^{C_0+100}} \le l \le \frac{x}{(\log x)^{C_0+10}}$ such that

$$\frac{N_0}{(\log N_0)^2} \le N \le \frac{N_0}{2}$$

We deduce

$$\mathsf{r}_M^{\mathbb{F}_q}\left(\frac{N_0}{2^{\epsilon^6(\log N_0)^{0.359}}}\right) \leq N_0^{14\epsilon}$$

Rewriting the bounds in terms of N we get

$$\mathsf{r}_M^{\mathbb{F}_q}\left(\frac{N}{2^{\epsilon^6(\log N)^{0.35}}}\right) \leq N^{15\epsilon}$$

□

8 Group Algebra Matrices over Finite Fields

We will now generalize Theorem 6.1 to matrices over a finite field $\mathbb{F}_q$. Write the underlying abelian group G as a direct product of cyclic groups $\mathbb{Z}_{n_1} \otimes \cdots \otimes \mathbb{Z}_{n_a}$. While for matrices with entries in $\mathbb{C}$, it sufficed to work with the Kronecker product of the Fourier matrices $F_{n_1} \otimes \cdots \otimes F_{n_a}$, we require slightly different techniques for rigidity over a fixed finite field as an extension containing all of the necessary roots of unity could have too high degree. Instead of working through Fourier matrices, we will work directly with the group algebra matrices themselves.

First note that Theorem 5.2 can be slightly strengthened so that to reduce the rank of any circulant matrices, the locations to be changed are fixed and the changes are fixed linear combinations of the entries of the circulant matrix. More precisely, we make the following definition.

Definition 8.1. *Given a group G with $|G| = n$, we say G is (r,s)-reducible over $\mathbb{F}_q$ if the following properties hold*

- *There exists a set of entries $S \subset [n] \times [n]$ such that S contains at most s nonzero entries in each row and column*
- *There are matrices $A, B \in \mathbb{F}_q^{n \times n}$ such that*

$$\mathsf{rank}(A), \mathsf{rank}(B) \leq r$$

- *There are matrices $E_1, \dots, E_n \in \mathbb{F}_q^{n\times n}$ with all nonzero entries in S and arbitrary matrices $Y_1, \dots, Y_n \in \mathbb{F}_q^{n\times n}$ and $Z_1, \dots, Z_n, \in \mathbb{F}_q^{n\times n}$ such that for any group algebra matrix of G, say M, with top row consisting of entries $x_1, \dots, x_n$,*

$$M = A(x_1Y_1 + \cdots + x_nY_n) + (x_1Z_1 + \cdots + x_nZ_n)B + (x_1E_1 + \cdots + x_nE_n)$$

If the group G is (r,s)-reducible over $\mathbb{F}_q$ we write

$$\mathsf{Non}_G^{\mathbb{F}_q}(r) \leq s$$

We call the matrices A, B (r,s)-reduction matrices and call the matrices $Y_1, \dots, Y_n, Z_1, \dots, Z_n, E_1, \dots, E_n$ (r,s)-reduction helpers. We write $Y_M = x_1Y_1 + \cdots + x_nY_n$ and similar for Z_M and E_M.

We now reformulate Theorem 5.2 below.

Claim 8.2. *For fixed $\epsilon > 0$ and all sufficiently large N,*

$$\mathsf{Non}_{\mathbb{Z}_N}^{\mathbb{F}_q}\left(\frac{N}{2^{\epsilon^6(\log N)^{0.35}}}\right) \leq N^{15\epsilon}$$

Proof. First consider when N_0 is (l, x)-factorable for some $\frac{x}{(\log x)^{C_0+100}} \le l \le \frac{x}{(\log x)^{C_0+10}}$. Let M_0 be a $N_0 \times N_0$ circulant matrix (i.e. a group algebra matrix for $\mathbb{Z}_{N_0}$) over $\mathbb{F}_q$ and say the entries in its top row are $x_1, \dots, x_{N_0}$. Let γ be a primitive $N_0{}^{\text{th}}$ root of unity and $t_1, \dots t_n$ be the set of prime powers at most $x^{0.3}$ that are relatively prime to q. Let $\omega_1, \dots, \omega_n$ be primitive $t_1, \dots, t_n$ roots of unity. By Theorem 7.4, there exists a matrix E over $\mathbb{F}_q[\gamma, \omega_1, \dots, \omega_n]$ with at most $N_0^{7\epsilon}$ nonzero entries in each row and column such that

$$\mathsf{rank}(F_{N_0} - E) \le \frac{N_0}{2^{\epsilon^6 (\log N_0)^{0.36}}}$$

Now write

$$M_0 = F_{N_0}^* D F_{N_0} = (F_{N_0} - E)^* D F_{N_0} + E^* D (F_{N_0} - E) + E^* D E$$

where D is a diagonal matrix whose entries are linear combinations of $x_1, \dots, x_{N_0}$. Note that all of the above matrices have entries contained in $\mathbb{F}_q[\eta]$ where η is a primitive root of unity of degree $C = N_0 \operatorname{lcm}(t_1, \dots, t_n)$. As argued before, the degree of the extension is at most $2^{(\log N_0)^{0.31}}$. Let the conjugates of η be $\eta_1 = \eta, \eta_2, \dots, \eta_m$. Let $F_{N_0}^1, \dots, F_{N_0}^m$ be obtained by taking F_{N_0} and replacing η with its conjugates. Define $D^1, \dots, D^m, E^1, \dots, E^m$ similarly. As in the proof of Lemma 7.1, there exists an integer k such that $\eta_1^k + \dots + \eta_m^k \neq 0$. We now have

$$M_0 = \frac{1}{\eta_1^k + \dots + \eta_m^k} \left(\sum_{i=1}^m \eta_i^k (F_{N_0}^i - E^i)^* D^i F_{N_0}^i + \sum_{i=1}^m \eta_i^k E^{i^*} D^i (F_{N_0}^i - E^i) + \sum_{i=1}^m \eta_i^k E^{i^*} D^i E^i \right)$$

Note that $\frac{1}{\eta_1^k + \dots + \eta_m^k} \in \mathbb{F}_q$ and all three of the sums are matrices whose entries are linear combinations of $x_1, \dots, x_{N_0}$ with coefficients in $\mathbb{F}_q$. The last term satisfies the desired sparsity constraint as it has at most $N_0^{14\epsilon}$ nonzero entries in each row and column and the locations of these entries are independent of M_0.

It remains to argue that the first two terms satisfy the desired rank constraint. Note that the span of the columns of $(F_{N_0}^1 - E^1), \dots, (F_{N_0}^m - E^m)$ has dimension at most

$$\frac{mN_0}{2^{\epsilon^6 (\log N_0)^{0.36}}} \le \frac{N_0}{2^{\epsilon^6 (\log N_0)^{0.359}}}$$

over $\mathbb{F}_q[\eta]^{N_0}$. Therefore, the dimension of the intersection of this subspace with $F_q^{N_0}$, say V, has dimension at most $\frac{N_0}{2^{\epsilon^6 (\log N_0)^{0.359}}}$. In particular we can write

$$\sum_{i=1}^m \eta_i^k (F_{N_0}^i - E^i)^* D^i F_{N_0}^i = x_1 C_1 + \dots + x_{N_0} C_{N_0}$$

for some fixed matrices $C_1, \dots, C_{N_0}$ with entries over $\mathbb{F}_q$. Also all columns of $C_1, \dots, C_{N_0}$ must be in V so each can be written as AY_i where A is a fixed matrix with rank at most $\frac{mN_0}{2^{\epsilon^6 (\log N_0)^{0.36}}}$. Thus there exists fixed matrices $Y_1, \dots, Y_{N_0} \in \mathbb{F}_q^{n \times n}$ and a matrix A satisfying the desired rank constraint such that

$$\sum_{i=1}^m \eta_i^k (F_{N_0}^i - E^i)^* D^i F_{N_0}^i = A(x_1 Y_1 + \dots + x_{N_0} Y_{N_0})$$

A similar argument shows that the second term can also be written in the desired form.

Now to extend to arbitrary N (not necessarily (l, x)-factorable), simply note that any circulant matrix of size N can be embedded into a circulant matrix of size at least $2N$ where all of the entries in the larger matrix are linear combinations of the entries of the original matrix. We can then apply Lemma 5.1 and complete the proof in the same way as Theorem 5.2. $\square$

Now we introduce the main technical result of this section that allows us to deal with direct products of groups.

Claim 8.3. *Say we have a set of groups $G_1, \dots, G_a$ such that $|G_i| = n_i$. Say that for each $1 \le i \le a$, G_i is (r_i, s_i)-reducible over $\mathbb{F}_q$. Let $G = G_1 \otimes \cdots \otimes G_a$ and $|G| = n = n_1 n_2 \dots n_a$. Then for any integer l and group algebra matrix M of G over $\mathbb{F}_q$, we have*

$$r_M^{\mathbb{F}_q}(r) \le s$$

where

$$r = \sum_{S \subset [a], |S| = l} 2^l \prod_{i \in S} \sqrt{r_i n_i} \prod_{i' \notin S} n_{i'}$$
$$s = \sum_{S \subset [a], |S| < l} 2^{|S|} \prod_{i \in S} n_i \prod_{i' \notin S} s_{i'}$$

Proof. For each $1 \le i \le a$, let A^i, B^i be the (r_i, s_i)-reduction matrices for the group G_i. We will write M as a sum of simpler "component" matrices. First, we can write M as follows

$$M = \begin{bmatrix} M_1 & \dots & M_{\frac{n}{n_1}} \\ \vdots & \ddots & \vdots \\ M_{\frac{n}{n_1}} & \dots & \end{bmatrix}$$

such that each M_i is a group algebra matrix of G_1. Now we can write

$$M = \begin{bmatrix} A^1 Y_{M_1} & \dots & A^1 Y_{M_{\frac{n}{n_1}}} & \\ \vdots & \vdots & \ddots & \vdots \\ A^1 Y_{M_{\frac{n}{n_1}}} & \dots & & \end{bmatrix} + \begin{bmatrix} Z_{M_1} B^1 & \dots & Z_{M_{\frac{n}{n_1}}} B^1 & \\ \vdots & \vdots & \ddots & \vdots \\ Z_{M_{\frac{n}{n_1}}} B^1 & \dots & & \end{bmatrix} + \begin{bmatrix} E_{M_1} & \dots & E_{M_{\frac{n}{n_1}}} & \\ \vdots & \vdots & \ddots & \vdots \\ E_{M_{\frac{n}{n_1}}} & \dots & & \end{bmatrix}$$

Call the three matrices above the first, second and third components respectively. Now in each of the matrices above, we can form blocks that correspond to group algebra matrices of G_2. We then decompose each into three parts and iterate the process. We end up with 3^a distinct matrices each of which corresponds to choosing one of the three components for each of $G_1, \dots, G_a$. For each of the 3^a distinct matrices, we label it M_I where $I = (i_1, \dots i_a) \in \{1,2,3\}^a$ is a tuple where i_j is 1 if the first component of G_j is chosen, i_j is 2 if the second component of G_j is chosen, and i_j is 3 if the third component of G_j is chosen.

Let $S^1(I), S^2(I), S^3(I) \subset [a]$ denote the subsets of locations where the entry of I is $1, 2$ or 3 respectively. Write

$$M = \sum_{\substack{I \in \{1,2,3\}^a \\ |S^3(I)| \le a - l}} M_I + \sum_{\substack{I \in \{1,2,3\}^a \\ |S^3(I)| > a - l}} M_I$$

We claim the first term is low-rank while the second term is sparse. For each $1 \le i \le a$, there exists a set of linearly independent vectors $v_1^i, \dots, v_{n_i - r_i}^i$ such that $v_j^i A^i = 0$ and a set of linearly independent vectors $u_1^i, \dots, u_{n_i - r_i}^i$ such that $B^i u_j^i = 0$ for all $1 \le j \le n_i - r_i$. We can complete the set $\{v_1^i, \dots, v_{n_i - r_i}^i\}$ to a basis $\{v_1^i, \dots, v_{n_i}^i\}$ and similar for $\{u_1^i, \dots, u_{n_i}^i\}$. Now say we are given a matrix M_I with $I \in \{1,2,3\}^a$. We claim

$$\mathsf{rank}(M_I) \le \prod_{i \in S^1(I)} r_i \prod_{i \in [a] \setminus S^1(I)} n_i$$

To see this, consider the basis of $\mathbb{F}_q^n$ consisting of the vectors $v_{j_1}^1 \otimes v_{j_2}^2 \otimes \cdots \otimes v_{j_a}^a$ where $(j_1, \dots, j_a) \in [n_1] \times \cdots \times [n_a]$. If for some index $i \in S^1(I)$, $j_i \le n_i - r_i$, then

$$\left(v_{j_1}^1 \otimes v_{j_2}^2 \otimes \cdots \otimes v_{j_a}^a\right) M_I = 0$$

Thus the number of vectors v in this basis such that $vM_I \neq 0$ is at most $\prod_{i \in S^1(I)} r_i \prod_{i \in [a] \setminus S^1(I)} n_i$. Similarly, we get

$$\mathsf{rank}(M_I) \leq \prod_{i \in S^2(I)} r_i \prod_{i \in [a] \setminus S^2(I)} n_i$$

Also note that if for two distinct tuples I, I', $S^1(I) \subset S^1(I')$, we get using the same argument above that

$$\mathsf{rank}(M_I + M_{I'}) \leq \prod_{i \in S^1(I)} r_i \prod_{i \in [a] \setminus S^1(I)} n_i$$

In the sum

$$\sum_{\substack{I \in \{1,2,3\}^a \\ |S^3(I)| \leq a-l}} M_I$$

we can essentially merge all of the terms and consider only the tuples I with $|S^3(I)| = a - l$. In particular,

$$\begin{aligned} \mathsf{rank}\left(\sum_{\substack{I \in \{1,2,3\}^a \\ |S^3(I)| \leq a-l}} M_I \right) &\leq \sum_{\substack{I \in \{1,2,3\}^a \\ |S^3(I)| = a-l}} \min \left(\prod_{i \in S^2(I)} r_i \prod_{i \in [a] \setminus S^2(I)} n_i, \prod_{i \in S^1(I)} r_i \prod_{i \in [a] \setminus S^1(I)} n_i \right) \\ &\leq \sum_{\substack{I \in \{1,2,3\}^a \\ |S^3(I)| = a-l}} \prod_{i \in [a] \setminus S^3(I)} \sqrt{r_i n_i} \prod_{i \in S^3(I)} n_i = \sum_{S \subset [a], |S| = l} 2^l \prod_{i \in S} \sqrt{r_i n_i} \prod_{i' \notin S} n_{i'} \end{aligned}$$

Now it remains to bound the sparsity of

$$\sum_{\substack{I \in \{1,2,3\}^a \\ |S^3(I)| > a-l}} M_I$$

Note that the number of nonzero entries in each row and column of M_I is at most $\prod_{i \in S^3(I)} s_i \prod_{i \in [a] \setminus S^3(I)} n_i$ and for each fixed subset $S^3(I)$, there are exactly $2^{|S^3(I)|}$ possible tuples I. Thus the number of nonzero entries in each row and column of the sum is at most

$$\sum_{\substack{I \in \{1,2,3\}^a \\ |S^3(I)| > a-l}} \prod_{i \in S^3(I)} s_i \prod_{i \in [a] \setminus S^3(I)} n_i = \sum_{S \subset [a], |S| < l} 2^{|S|} \prod_{i \in [a] \setminus S} s_i \prod_{i \in S} n_i$$

Overall we have shown how to write M as the sum of a matrix with the desired rank and a matrix with the desired sparsity, completing the proof.

$\square$

We are now ready to prove the main theorem about rigidity of group algebra matrices.

Theorem 8.4. *Let $\mathbb{F}_q$ be a fixed finite field and $\epsilon < 0.1$ be a fixed constant. Let G be an abelian group. As long as $|G|$ is sufficiently large, for any group algebra matrix M of G over $\mathbb{F}_q$, we have*

$$r_M^{\mathbb{F}_q}\left(\frac{|G|}{2^{\epsilon^{20} (\log |G|)^{0.3}}} \right) \leq |G|^{100\epsilon}$$

Proof. We can essentially follow the same method as the proof of Theorem 6.1 except using Claim 8.3 to deal with direct products of cyclic groups that are roughly the same size. $\square$

9 Final Remarks

Theorem 6.1 naturally raises the question of what happens when G is a non-abelian group. When G is non-abelian, it is no longer possible to diagonalize the matrix $M_G(f)$ but there is a change of basis matrix A such that $AM_G(f)A^*$ is block-diagonal where the diagonal blocks correspond to the irreducible representations of G. When all of the irreducible representations of G are small, it may be possible to use similar techniques to the ones used here. On the other hand, this suggests that perhaps $M_G(f)$ is a candidate for rigidity when all irreducible representations of G are large (for instance quasi-random groups [Gow08]).

References

[AW16] Josh Alman and Ryan Williams. Probabilistic rank and matrix rigidity. *CoRR*, abs/1611.05558, 2016.

[BH98] R. Baker and G. Harman. Shifted primes without large prime factors. *Acta Arithmetica*, 83(4):331–361, 1998.

[DE17] Zeev Dvir and Benjamin Edelman. Matrix rigidity and the croot-lev-pach lemma. *arXiv preprint arXiv:1708.01646*, 2017.

[Fri93] Joel Friedman. A note on matrix rigidity. *Combinatorica*, 13(2):235–239, 1993.

[Gow08] William T Gowers. Quasirandom groups. *Combinatorics, Probability and Computing*, 17(3):363–387, 2008.

[GT16] Oded Goldreich and Avishay Tal. Matrix rigidity of random toeplitz matrices. In *Proceedings of the forty-eighth annual ACM symposium on Theory of Computing*, pages 91–104. ACM, 2016.

[KLPS14] Abhinav Kumar, Satyanarayana V Lokam, Vijay M Patankar, and MN Jayalal Sarma. Using elimination theory to construct rigid matrices. *computational complexity*, 23(4):531–563, 2014.

[L$^+$09] Satyanarayana V Lokam et al. Complexity lower bounds using linear algebra. *Foundations and Trends® in Theoretical Computer Science*, 4(1–2):1–155, 2009.

[Lok00] Satyanarayana V Lokam. On the rigidity of vandermonde matrices. *Theoretical Computer Science*, 237(1-2):477–483, 2000.

[Lok06] Satyanarayana V Lokam. Quadratic lower bounds on matrix rigidity. In *International Conference on Theory and Applications of Models of Computation*, pages 295–307. Springer, 2006.

[SSS97] Mohammad Amin Shokrollahi, Daniel A Spielman, and Volker Stemann. A remark on matrix rigidity. *Information Processing Letters*, 64(6):283–285, 1997.

[Val77] Leslie G. Valiant. Graph-theoretic arguments in low-level complexity. In Jozef Gruska, editor, *Mathematical Foundations of Computer Science 1977*, pages 162–176, Berlin, Heidelberg, 1977. Springer Berlin Heidelberg.

ECCC ISSN 1433-8092

https://eccc.weizmann.ac.il