Notes on Matrix Rigidity

Yangxinyu Xie

June 10, 2020

Contents

Chapter 1

Motivation and Definition

Definition 1.0.1. The **density** of a matrix A is the number of nonzero elements drawn from a field $\mathbb{F}$, denoted by dens(A).

Definition 1.0.2. The **rigidity** of a matrix A is the function $\mathcal{R}_A^{\mathbb{F}}(r) : \{1, ..., N\} \to \{0, 1, ..., N^2\}$ defined by

$$\mathcal{R}_A^{\mathbb{F}}(r) := \min\{i | \exists\, B, \text{dens}(B) = i, \text{rank}(A + B) \leq r\} \tag{1.0.0.1}$$

Valiant motivated the definition of matrix rigidity from the analysis of circuit complexity and proved that if A is a rigid matrix, i.e. $\mathcal{R}_A^{\mathbb{F}}(\epsilon N) = N^{1+\delta}$ for some $\epsilon, \delta > 0$, then the linear program to compute Ax cannot be a circuit of $+$ gates of size $O(N)$ and depth $O(n = \log N)$. Moreover, Valiant gave a non-constructive proof for the following.

Theorem 1.0.3 ([Val77])**.**

1. *For an infinite field $\mathbb{F}$, for all N, there exists a $N \times N$ matrix A such that $\mathcal{R}_A^{\mathbb{F}}(r) = (N - r)^2$.*
2. *For a finite field $\mathbb{F}$ with c elements, for all N, there exists a $N \times N$ matrix A such that for all $r < N - \sqrt{2N \log_c 2 + \log_2 N}$,*

$$\mathcal{R}_A^{\mathbb{F}}(r) \geq \frac{(N - r)^2 - 2N \log_c 2 - \log_2 N}{2 \log_c N + 1} \tag{1.0.0.2}$$

Open Question 1.0.4. Find *explicit* matrices A.

Chapter 2

Explicit Lower Bounds

For explicit lower bounds, most of the proofs consist of two steps: first, we show that most sub-matrices of the given matrix M has large or full rank r; second, if $\mathcal{R}_A^{\mathbb{F}}(r)$ is small, we are likely to get a sub-matrix that remains intact. The first step implies that it is plausible to find rigid matrices with high regularity.

Definition 2.0.1 ([BCS97])**.** A matrix A is called **totally regular** if and only if every minor of A is invertible.

A relaxation of this definition will suffice for our purpose.

Definition 2.0.2. A matrix A is called **almost totally regular** if and only if every $r \times r$ minor of A has rank $\Omega(r)$.

Another notion of regularity based on expectation is introduced by Pudlak, called densely regular.

Definition 2.0.3 ([Pud94])**.** Let A be an $N \times N$ matrix, $0 \le \epsilon, \delta, \eta \le 1$. We say that A is (ϵ, δ, η)–**densely regular**, if for every k with $\eta N \le k \le N$, there are nonempty sets of k elements subset $\mathcal{X}, \mathcal{Y} \in [1, n]^k$ such that for every $i, j = 1, ..., N$

$$\delta \mathbb{P}[i \in X] \le k/N \quad \text{and} \quad \delta \mathbb{P}[j \in Y] \le k/N \tag{2.0.0.1}$$

where $X \in \mathcal{X}, Y \in \mathcal{Y}$ are chosen with some probability distributions and such that for random $X \in \mathcal{X}, Y \in \mathcal{Y}$ the mean value of the rank of the matrix determined by X and Y is at least ϵk.

Again, a relaxation of this will also be sufficient to us.

Definition 2.0.4 ([Che05])**.** Let A be an $N \times N$ matrix. We say that A is ϵ-**densely regular**, if there is a constant $0 < \epsilon < 1$ such that for every k with $0 \le k \le N$, a $k \times k$ minor of A picked uniformly at random has an expected rank at least ϵk.

However, as pointed out by Lokam [Lok00], any proof relying on the second step cannot produce a lower bound better than $\Omega((n^2/r)\log(n/r))$. Moreover, due to the existence of linear size superconcentrators, the first step is far from sufficient to show a desirable rigidity.

Proposition 2.0.5 ([Val77])**.** *For each N there is an $N \times N$ totally regular matrix A such that*

$$\mathcal{R}_A^{\mathbb{F}}(\frac{N \log\log\log N}{\log\log N}) \leq N^{1+O(\frac{1}{\log\log N})} \tag{2.0.0.2}$$

Nevertheless, finding explicit matrices with high regularity is still a reasonable start.

2.1. Totally Regular Matrices

2.1.1 A Combinatorial Lemma

Lemma 2.1.1 ([SSS97])**.** *If fewer than*

$$\mu(N,r) = (N-r+1)(N-(r-1)^{1/r}N^{1-1/r}) \tag{2.1.1.1}$$

entries of an $N \times N$ matrix A is marked, then there is an $r \times r$ submatrix that remains intact.

Proof. Think of A as the adjacency matrix of a bipartite graph $G_{N,N}$ which contains an edge (i,j) if and only if $A_{i,j}$ has not been marked. Hence, having an $r \times r$ submatrix that remains intact is equivalent to having a $K_{r,r}$ complete bipartite subgraph in $G_{N,N}$. Hence, we have that $G_{N,N}$ cannot have more than $n^2 - \mu(n,r)$ edges, or it will contain a $K_{r,r}$ complete bipartite subgraph (see [Juk11], Theorem 2.10). $\square$

Corollary 2.1.2. *Let $r \geq \log^2 n$ and let N be sufficiently large. If fewer than*

$$\frac{N(N-r+1)}{2r}\log\frac{N}{r-1} \tag{2.1.1.2}$$

changes are made to an $N \times N$ matrix A, then there exists an $r \times r$ submatrix that remains intact.

It is easy to see that if any $r \times r$ minor has rank $\Omega(r)$, we have that the rigidity of A is $\Omega(\frac{N^2}{r}\log\frac{N}{r})$.

2.1.2 Cauchy Matrix

Definition 2.1.3. Let $x_1, ..., x_n, y_1, ..., y_n$ be elements of a field $\mathbb{F}_N$ with the property that

$$\prod_{i \neq j}(x_i - x_j) \neq 0, \quad \prod_{i \neq j}(y_i - y_j) \neq 0, \quad \prod_{i,j}(x_i + y_j) \neq 0 \tag{2.1.2.1}$$

we define the **Cauchy matrix** by

$$C := \Big(\frac{1}{x_i + y_j}\Big)_{1\leq i,j\leq n} \tag{2.1.2.2}$$

Hence, for every $1 \leq r \leq n$, each of its $r \times r$-submatrix has the determinant

$$\frac{\prod_{i\neq j}(x_i - x_j)\prod_{i\neq j}(y_i - y_j)}{\prod_{i,j}(x_i + y_j)} \tag{2.1.2.3}$$

which is nonzero. In other words, the Cauchy matrix is totally regular.

The following theorem is thus a direct result of Corollary 2.1.2.

Theorem 2.1.4 ([SSS97])**.** *Let $\mathbb{F}_N$ be a sequence of fields and let (C_N) be a sequence of Cauchy matrices where $C_N \in \mathbb{F}_N^{N\times N}$. Then if $\log^2 N \leq r \leq N/2$, we have*

$$\mathcal{R}_{C_N}^{\mathbb{F}_N} \geq \Big(\frac{N^2}{4r}\log\frac{N}{r-1}\Big) \tag{2.1.2.4}$$

2.1.3 Fourier Transform Matrix

Another type of totally regular matrices is the **Discrete Fourier Transform Matrix**. Hence, the following result is also a direct consequence of Corollary 2.1.2. Also note that Discrete Fourier Transform Matrix is a type of Vandermonde Matrix.

Theorem 2.1.5 ([Lok00], [Lok09])**.** *Let $F = (\omega_i^{j-1})_{i,j=0}^{N-1}$, where ω is a primitive nth root of unity. Then, as N ranges over all prime numbers and $\log^2 N \leq r \leq N/2$,*

$$\mathcal{R}_F(r) \geq \frac{N^2}{4(r+1)}\log\frac{N}{r} \tag{2.1.3.1}$$

Theorem 2.1.6 ([DL19])**.** *Let F denote the $N \times N$ Fourier transform matrix. For any fixed $0 < \epsilon < 0.1$ and N sufficiently large,*

$$\mathcal{R}_F\Big(\frac{N}{2^{\epsilon^6 n^{0.35}}}\Big) \leq N^{15\epsilon} \tag{2.1.3.2}$$

where $n = \log N$.

2.1.4 Asymptotically Good Error Correcting Codes

Lemma 2.1.7 (Tsfasman-Vladut-Zink Bound [TVZ82], as stated in [Lok09])**.** *Let q be the square of a prime. Then for every rate p, there exists an infinite sequence $[n_t, k_t, \mathcal{R}_t], t = 1, 2, ...$ of codes over $\mathbb{F}_q$ such that the asymptotic rate $p = \lim_{t\to\infty} k_t/n_t$ and the asymptotic relative distance $\delta = \lim_{t\to\infty} d_t/n_t$ such that*

$$p \geq 1 - \delta - \frac{1}{\sqrt{q} - 1} \tag{2.1.4.1}$$

For infinitely many N, there exists a $[2N, N, d]$-code with $d \geq (1-\epsilon)N$ where $\epsilon = 2/(\sqrt{q}-1)$.

For a given N, let Γ be the $[2N, N, d]$-code as in Corollary 2.1.8, whose generator matrix has the form $(I_N|A)$ where I_N is the $N \times N$ identity matrix.

Theorem 2.1.9 ([SSS97])**.** *Let A be an $N \times N$ matrix as defined above. Then, for $\max(\log^2 N, \epsilon N) \leq r \leq N/4$,*

$$\mathcal{R}_A^{\mathbb{F}_q}(r) \geq \frac{N^2}{8r} \log \frac{N}{2r-1} \tag{2.1.4.2}$$

Proof. We first show that for all $2r \times 2r$ submatrix of A, the rank must be at least r. Suppose on the contrary, then let B be a $2r \times 2r$ submatrix of A with r dependent rows. Then a linear combination of these r rows of the generator matrix gives a code word of weight

$$r + N - 2r = N - r \leq (1-\epsilon)N - 1 \leq d \tag{2.1.4.3}$$

where the first r comes from the identity matrix and the $N - 2r$ comes from the fact that these r rows are dependent. Hence, we reach a contradiction.
By Corollary 2.1.2, we obtain the desired result. □

2.2. Densely Regular Matrices

Theorem 2.2.1. *For every positive ϵ, δ, there exists a positive α such that for any field $\mathbb{F}$, every $\eta, 0 < \eta \leq 1$, and any $N \times N$ (ϵ, δ, η)–densely regular matrix A,*

$$\mathcal{R}_A^{\mathbb{F}}(r) \geq \alpha \frac{N^2}{r} \tag{2.2.0.1}$$

for $\epsilon\eta N/2 \leq r \leq \epsilon N/2$

Proof. Let $\epsilon\eta N/2 \leq r \leq \epsilon N/2$ and sets $\mathcal{X}, \mathcal{Y}$ be given and set $k = \lceil 2r/\epsilon \rceil$. For random variables $X \in \mathcal{X}, Y \in \mathcal{Y}$,, we have that for a coordinate (i, j),

$$\mathbb{P}[(i,j) \in X \times Y] \leq \frac{k^2}{\delta^2 N^2} \tag{2.2.0.2}$$

Let Z be the minimal set of coordinates changed to reduce the rank of A to r and let z be the number of coordinates in the set $Z \cap (X \times Y)$. Then, we have

$$\mathbb{E}[z] = \sum_{(i,j) \in Z} \mathbb{P}[(i,j) \in X \times Y] \leq |Z| \frac{k^2}{\delta^2 N^2} \tag{2.2.0.3}$$

Notice that on the other hand, the mean value of the rank of the matrix determined by X and Y

is at least ϵk, which implies

$$\mathbb{E}[z] \geq \epsilon k - r \geq 2r - r = r \tag{2.2.0.4}$$

Therefore, we have

$$r \leq \mathbb{E}[z] \leq |Z| \frac{k^2}{\delta^2 N^2} \tag{2.2.0.5}$$

and thus by plugging $k = \lceil 2r/\epsilon \rceil$

$$|Z| \geq \frac{r\delta^2 N^2}{k^2} = \Omega(\frac{N^2}{r}) \tag{2.2.0.6}$$

□

The proof for ϵ-densely regular matrix is very similar.

2.2.1 Vandermonde Matrix

Proposition 2.2.2. *Let $V = (x_i^{j-1})_{i,j=1}^N$ be a Vandermonde matrix with distinct x_i over some field. V is $(1, 1/2, 0)$-densely regular.*

By theorem 2.2.1, we prove the $\Omega(N^2/r)$ lower bound for Vandermonde matrices. An alternative proof is given by Shparlinsky.

2.2.2 Hadamard Matrix

Definition 2.2.3. A matrix $H = (h_{i,j}) \in \mathbb{C}^{N \times N}$ is called a **(generalised) Hadamard matrix** if $|h_{i,j}| = 1$ for all $i, j \in [n]$ and $HH^* = NI_N$ where H^* is the conjugate transpose of H and I_N is the $N \times N$ identity matrix.

In other words, a matrix H is a (generalised) Hadamard matrix if $|h_{i,j}| = 1$ for all $i, j \in [n]$ and the rows of H are pairwise orthogonal.

Definition 2.2.4. The **Frobenius norm** of a matrix $A \in \mathbb{C}^{N \times N}$ is

$$||A||_F := \Big(\sum_{i,j} |a_{i,j}|^2 \Big)^{1/2} \tag{2.2.2.1}$$

Definition 2.2.5. The **trace** of a matrix $A \in \mathbb{C}^{N \times N}$ is the sum of its eigenvalues.

$$\text{Tr}(A) = \sum_{i=1}^N \lambda_(A) \tag{2.2.2.2}$$

Definition 2.2.6. The i**th singular value** $\sigma_i(A)$ is defined by

$$\sigma_i(A) := \sqrt{\lambda_i(AA^*)}, 1 \leq i \leq n \tag{2.2.2.3}$$

where λ_i denotes the ith largest eigenvalue of AA^*.

We recall some fact about singular value decomposition and Frobenius norm. The proof can be found in chapter 2.4 in [GVL12].

Proposition 2.2.7. *For any matrix $A \in \mathbb{C}^{N\times N}$,*

- *there exists unitary matrices $U, V \in \mathbb{C}^{N\times N}$ such that*

$$U^*AV = \text{diag}(\sigma_1(A), \sigma_2(A), ..., \sigma_N(A)) \tag{2.2.2.4}$$

- $||A||_F^2 = \sigma_1^2(A) + \sigma_2^2(A) + ... + \sigma_N^2(A)$.

Proposition 2.2.8. *If $A \in \mathbb{R}^{N\times N}$ is symmetric, then*

$$\frac{\text{Tr}(A)^2}{||A||_F^2} \leq \text{rank}(A) \tag{2.2.2.5}$$

Proof. Let $B = AA^*$. Notice that

$$\text{Tr}(B) = \sum_{i=1}^{N} \lambda_i(B) = ||A||_F^2 \tag{2.2.2.6}$$

Because A is symmetric,

$$\sum_{i=1}^{N} \lambda_i(B) = \text{Tr}(B) = \sum_{i=1}^{N} \lambda_i(A^2) = \sum_{i=1}^{N} \lambda_i^2(A) \tag{2.2.2.7}$$

Moreover, B has only rank(B) =rank(A) non-zero eigenvalues, which are all positive. Assume without loss of generality $\lambda_1^2(A) \geq \lambda_2^2(A) \geq ... \geq \lambda_N^2(A)$. Then, $\sum_{i=1}^{N} \lambda_i^2(A) = \sum_{i=1}^{\text{rank}(A)} \lambda_i^2(A)$. Hence, by Cauchy-Schwarz inequality, we have

$$||A||_F^2 = \sum_{i=1}^{\text{rank}(A)} \lambda_i^2(A) \geq \frac{\left(\sum_{i=1}^{\text{rank}(A)} \lambda_i(A)\right)^2}{\text{rank}(A)} \geq \frac{\text{Tr}(A)^2}{\text{rank}(A)} \tag{2.2.2.8}$$

□

Proposition 2.2.9 ([KR98])**.** *Let H be an $N \times N$ generalised Hadamard matrix. Let G be a random $q \times N$ submatrix of H and let A be a random $q \times q$ submatrix of G. Then $\mathbb{E}[\text{rank}(A)] \geq r/8$.*

Proof. Let $B = AA^*$. Then B is a positive definite symmetric matrix in $\mathbb{R}^{q\times q}$. Recall that $h_{i,j} = 1$, which implies all entries of B on the main diagonal equals to q. Thus Tr$(B) = q^2$. By proposition

2.2.8, we obtain that $\text{rank}(A) \leq r$ for some positive integer r implies

$$||B||_F^2 \geq \frac{\text{Tr}(B)^2}{\text{rank}(B)} \geq \frac{q^4}{r} \tag{2.2.2.9}$$

Let

$$\epsilon_j = \begin{cases} 1 & \text{if the } j\text{th column of } H \text{ is in } H_0 \\ 0 & \text{otherwise} \end{cases} \tag{2.2.2.10}$$

we have

$$\mathbb{E}[\epsilon_{j_1}\epsilon_{j_2}] = \begin{cases} \frac{q}{N} & \text{if } j_1 = j_2 \\ \frac{q(q-1)}{N(N-1)} & \text{if } j_1 \neq j_2 \end{cases} \tag{2.2.2.11}$$

Now, notice that $b_{i,j} = \sum_{k=1}^q a_{i,k}a_{j,k}^* = \sum_{k=1}^q g_{i,k}g_{j,k}^*\epsilon_k$ and $b_{i,j}^* = \sum_{l=1}^q a_{i,l}^*a_{j,l} = \sum_{k=1}^q g_{i,k}g_{j,k}^*\epsilon_l$

$$\begin{aligned} ||B||_F^2 &= \sum_{1\leq i,j\leq q} |b_{i,j}|^2 = \sum_{1\leq i,j\leq q} b_{i,j}b_{i,j}^* = \sum_{1\leq i\leq q, 1\leq j\leq n}\ \sum_{1\leq k,l\leq q} g_{i,k}g_{j,k}g_{i,l}^*g_{j,l}^*\epsilon_k\epsilon_l \\ &= \sum_{1\leq k,l\leq q}\Big(\epsilon_k\epsilon_l \sum_{1\leq i\leq q,1\leq j\leq n} g_{i,k}g_{j,k}g_{i,l}^*g_{j,l}^*\Big) \end{aligned} \tag{2.2.2.12}$$

Thus,

$$\begin{aligned} \mathbb{E}[||B||_F^2] &= \sum_{1\leq k,l\leq q}\Big(\mathbb{E}[\epsilon_k\epsilon_l] \sum_{1\leq i\leq q,1\leq j\leq n} g_{i,k}g_{j,k}g_{i,l}^*g_{j,l}^*\Big) \\ &= \frac{q(q-1)}{N(N-1)} \sum_{1\leq k,l\leq q}\ \sum_{1\leq i\leq q,1\leq j\leq n} g_{i,k}g_{j,k}g_{i,l}^*g_{j,l}^* + \Big(\frac{q}{N} - \frac{q(q-1)}{N(N-1)}\Big) \sum_{1\leq i\leq q,1\leq j\leq n} g_{i,k}g_{j,k}g_{i,l}^*g_{j,l}^* \\ &= \frac{q(q-1)}{N(N-1)}||GG^*||_F^2 + \Big(\frac{q}{N} - \frac{q(q-1)}{N(N-1)}\Big) \sum_{k=1}^q\ \sum_{1\leq i\leq q,1\leq j\leq n} g_{i,k}g_{i,k}^*g_{j,k}g_{j,k}^* \end{aligned} \tag{2.2.2.13}$$

Notice that $GG^* = NI_q$ where I_q is the $q \times q$ identity matrix.

$$\begin{aligned} \mathbb{E}[||B||_F^2] &= \frac{q(q-1)}{N(N-1)}||GG^*||_F^2 + \Big(\frac{q}{N} - \frac{q(q-1)}{N(N-1)}\Big) \sum_{k=1}^q\ \sum_{1\leq i\leq q,1\leq j\leq n} g_{i,k}g_{i,k}^*g_{j,k}g_{j,k}^* \\ &= \frac{q(q-1)}{N(N-1)}N^2q^2 + \Big(\frac{q}{N} - \frac{q(q-1)}{N(N-1)}\Big)Nq^2 \\ &= q^2(q + (N-q)\frac{q-1}{N-1}) \leq 2q^3 \end{aligned} \tag{2.2.2.14}$$

By Chebyshev's inequality, we have

$$\mathbb{P}[||B||_F^2 \geq \frac{q^4}{r}] \leq \frac{r}{q^4}\mathbb{E}[||B||_F^2] \leq \frac{2r}{q} \tag{2.2.2.15}$$

Thus, we have $\mathbb{P}[\text{rank}(A) \leq r] \leq \mathbb{P}[||B||_F^2 \geq \frac{q^4}{r}] \leq 2r/q$. Choosing $r = q/4$, we have$\mathbb{P}[\text{rank}(A) \leq q/4] \leq 1/2$. Again, using Chebyshev's inequality,

$$\mathbb{E}[\text{rank}(A)] \geq \frac{q}{4}\mathbb{P}[\text{rank}(A) \geq \frac{q}{4}] = \frac{q}{4}(1 - \mathbb{P}[\text{rank}(A) \geq \frac{q}{4}]) \geq \frac{q}{8} \tag{2.2.2.16}$$

□

Corollary 2.2.10. *Let H be an $N \times N$ generalised Hadamard matrix, then H is $1/8$-densely regular.*

Theorem 2.2.11 ([AW17])**.** *For every field $\mathbb{F}$, for every sufficiently small $\epsilon > 0$, and for all n, we have*

$$\mathcal{R}_H(2^{n-f(\epsilon)n}) \leq 2^{n(1+\epsilon)} \tag{2.2.2.17}$$

2.3. Averaging Argument

Another set of proofs utilises averaging argument to select some number of rows that has small changes and then show that the remaining part of these rows has high rank.

2.3.1 Vandermonde Matrix

Theorem 2.3.1 (Shparlinsky, see [Lok00])**.** *Let $V = (x_i^{j-1})_{i,j=1}^N$ be a Vandermonde matrix with distinct x_i over some field. Then*

$$\mathcal{R}_V(r) \geq \frac{(N-r)^2}{r+1} \tag{2.3.1.1}$$

Proof. Let r be given and let $s = \mathcal{R}_V(r)$. By averaging argument, we can select $r+1$ consecutive columns such that the total number of changes within these columns are at most $s(r+1)/(N-r)$. Then we select the rows that do not contain any changes in these these columns, which gives us at least $N - s(r+1)/(N-r)$ rows. Hence, we constructed a submatrix S of size $(r+1) \times (N - s(r+1)/(N-r))$. Because the rank of this submatrix is at most r, we have that there exists a nonzero vector g such that $S\text{g} = 0$. In other words, we obtain a polynomial $\sum_{t=0}^{r} g_t x^t = 0$ with at least $(N - s(r+1)/(N-r))$ roots. On the other hand, this polynomial can have at most r roots. Therefore,

$$r \geq (N - s(r+1)/(N-r)) \tag{2.3.1.2}$$

which gives $s \geq (N-r)^2)/(r+1)$. □

2.3.2 Hadamard Matrix

Proposition 2.3.2 (Alon, see [Juk11]). *Every non-trivial linear combination of any k rows of a Hadamard matrix $H = (h_{i,j}) \in \mathbb{C}^{N\times N}$ has at least N/k nonzero entries.*

Proof. Let A be a $k \times n$ submatrix of H and let $y = x^T A$ for some nonzero vector $x \in \mathbb{R}^k$. Let S be the set of the coordinates of the non-zero entries in y and let $s = |S|$. We need to show that $s \geq n/k$.
Assume without loss of generality that $x_1 = \max_{i\in[k]} |x_i|$. Let a^i denote the ith row of A. Because the rows of A are mutually orthogonal, we have

$$kx_1^2 N \geq \sum_{i=1}^k x_i^2 N = \sum_{i=1}^k \langle x_i a^i, x_i a^i\rangle = \langle \sum_{i=1}^k x_i a^i, \sum_{i=1}^k x_i a^i \rangle \tag{2.3.2.1}$$

Notice that $\sum_{i=1}^k x_i a^i = x^T A = y$, we have

$$\langle \sum_{i=1}^k x_i a^i, \sum_{i=1}^k x_i a^i \rangle = \langle y, y\rangle = \sum_{j=1}^N y_j^2 = \sum_{j=1}^N y_j^2 = \sum_{j\in S} y_j^2 = \sum_{j\in S} |y_j|^2 \tag{2.3.2.2}$$

Using Cauchy-Schwarz inequality, we have

$$\sum_{j\in S} |y_j|^2 \geq \frac{1}{s}(\sum_{j\in S} |y_j|)^2 = \frac{1}{s}(\sum_{j=1}^N |y_j|)^2 \tag{2.3.2.3}$$

On the other hand, because $|a_{i,j}| = 1$, we have

$$\begin{aligned} \sum_{j=1}^N |y_j| &\geq \sum_{j=1}^N y_j a_{1,j} = \sum_{j=1}^N \langle x, a^j\rangle a_{1,j} = \sum_{j=1}^N \sum_{i=1}^k x_i a_{i,j} a_{1,j} \\ &= \sum_{i=1}^k x_i \sum_{j=1}^N a_{i,j} a_{1,j} = \sum_{i=1}^k x_i \langle a^i, a^1\rangle = x_1 \langle a^1, a^1\rangle = x_1 N \end{aligned} \tag{2.3.2.4}$$

This gives us

$$kx_1^2 N \geq \frac{1}{s}(x_1 N)^2 \tag{2.3.2.5}$$

Thus, $s \geq N/k$. □

Notice that the following two corollaries holds for real-valued Hadamard matrices.

Corollary 2.3.3 (Alon, see [Juk01]). *If $t > (1 - 1/r)N$, then every $r \times t$ sub-matrix H' of an $N \times N$ Hadamard matrix $H \in \mathbb{R}^{N\times N}$ has rank r.*

Proof. For the sake of contradiction, we assume the opposite that rank$(H') < r$. Hence, there exists a nonzero vector $x \in \mathbb{R}^r$ such that $x^t H' = 0$. Because $t > (1 - 1/r)N$, this contradicts with proposition 2.3.2 that any nonzero linear combination of these r rows of H has at least N/r nonzero entries. $\square$

Corollary 2.3.4 (Alon, see [Juk01]). *If fewer than $(n/r)^2$ entries of an $N \times N$ Hadamard matrix $H \in \mathbb{R}^{N\times N}$ are changed), then the rank of the resulting matrix remains at least r.*

Proof. By averaging argument, we can choose (n/r) rows that has fewer than (n/r) changes in total. Therefore, the number of columns that remain intact in these (n/r) rows is greater than $(1-1/r)N$. By 2.3.3, we complete the proof. $\square$

Lemma 2.3.5 ([Lok95]). *For any $u \times v$ submatrix H_0 if an $N \times N$ generalised Hadamard matrix H*, rank$(H_0) \geq uv/N$.

Proof. Let $A \in \mathbb{C}^{k\times k}$ for some $k > 0$. Let $\lambda_1(A)$ be the largest eigenvalue of AA^*. We thus have

$$\frac{||A||_F^2}{\lambda_1(A)} = \frac{\sum_{i=1}^{N} \lambda_i(A)}{\lambda_1(A)} \tag{2.3.2.6}$$

Notice that AA^* has exactly rank(AA^*) =rank(A) nonzero entries, all of which are positive, which implies

$$\frac{||A||_F^2}{\lambda_1(A)} = \frac{\sum_{i\in[N],\lambda_i(A)>0} \lambda_i(A)}{\lambda_1(A)} \leq \text{rank}(A) \tag{2.3.2.7}$$

On the other hand, H_0 is a submatrix of H, we have $\lambda_1(H_0) \leq \lambda_1(H)$. Thus

$$\frac{||H_0||_F^2}{\lambda_1(H_0)} \geq \frac{||H_0||_F^2}{\lambda_1(H)} = \frac{uv}{N} \tag{2.3.2.8}$$

Notice that the last equality follows from the fact that $H_0H_0^* = vI_u$. Therefore, we have rank$(H_0) \geq ||H_0||_F^2/\lambda_1(H_0) \geq uv/N$. $\square$

Theorem 2.3.6 ([dW06]). *If $r \leq N/2$, then $\mathcal{R}_H(r) \geq N^2/4r$.*

Proof. Let r be given and let $s = \mathcal{R}_H(r)$. By averaging argument, we can select $2r$ rows that has fewer than $2rs/N$ changes. If $2rs/N \geq N$, we have $s \geq N^2/(2r)$ and we are done. If $2rs/N < N$, we then have that by lemma 2.3.5, for the submatrix H_0 that contains the $N - 2rs/N$ intact columns of these $2r$ rows,

$$r \geq \text{rank}(H_0) \geq \frac{2r(N-2rs/N)}{N} \tag{2.3.2.9}$$

which implies $s \geq N^2/4r$. $\square$

The same bound can be proved with a much simpler argument for a special type of Hadamard matrix called the **Sylvester matrix**, which is recursively defined as follows:

- $S_1 := (1)$.

- $S_{2n} := \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \otimes H_n = \begin{bmatrix} H_n & H_n \\ H_n & -H_n \end{bmatrix}$

Theorem 2.3.7 ([Mid05])**.** *If $S(N)$ is a Sylvester matrix and $r \leq N/2$ is a power of 2, then*

$$\mathcal{R}_{S(N)}(r) \geq \frac{N^2}{4r} \tag{2.3.2.10}$$

Proof. Let r be given and let $s = \mathcal{R}_S(r)$. Assume on the contrary that $s < N^2/4r$. If we divide S into $(N/2r)^2$ grids of size $2r \times 2r$, then by averaging argument, there exists a grid that has fewer than

$$s \cdot \frac{(2r)^2}{N^2} < \frac{N^2}{4r} \cdot \frac{(2r)^2}{N^2} = r \tag{2.3.2.11}$$

changes. Notice that each grid has full rank because it is exactly a Sylvester matrix of size $2r \times 2r$. Then this grid still has rank more than $2r - r = r$ after these r changes. Hence, the rank of S after these s changes will be more than r, which gives us a contradiction. $\square$

Remark 2.3.8. Notice that this simple proof can be applied to any totally regular matrix, for example, the Discrete Fourier Transform Matrix.

Chapter 3

Somewhat-Explicit Lower Bounds

3.1. The Shoup–Smolensky Dimensions

Definition 3.1.1 ([Mor96])**.** Let K be a field extension of $\mathbb{F}$, and let $t_1, ..., t_n \in K$. The set $\{t_1, ..., t_n\}$ is **algebraically independent** over $\mathbb{F}$ if $f(t_1, ..., t_n) \neq 0$ for all nonzero polynomials $f \in \mathbb{F}[x_1, ..., x_n]$.

Theorem 3.1.2 ([Lok00])**.** *Let $V = (x_i^{j-1})_{i,j=1}^{N}$ be a Vandermonde matrix where x_i are algebraically independent over $\mathbb{Q}$. Then*

$$\mathcal{R}_V(r) \geq \frac{N(N - cr^2)}{2} \tag{3.1.0.1}$$

where $c > 0$ is an absolute constant.

Theorem 3.1.3 ([Lok06])**.** *Let A be an $N \times N$ matrix over C and $0 \leq r \leq n$. Suppose, $D_{Nr}(A) = \binom{N^2}{Nr}$,i. e., all products of Nr distinct entries of A are linearly independent over $\mathbb{Q}$. Then,*

$$\mathcal{R}_A(r) \geq N(N - 16r) \tag{3.1.0.2}$$

3.2. Random Toeplitz/Hankel Matrices

Let $m, k \in \mathbb{N}, 16 \leq k \leq m$. Let $A \in \mathbb{F}_2^{m \times m}$ be the random matrix

$$\begin{bmatrix} a_1 & a_2 & \cdots & a_m \\ a_{k+1} & a_{k+2} & \cdots & a_{k+m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{(m-1)k+1} & a_{(m-1)k+2} & \cdots & a_{(m-1)k+m} \end{bmatrix}$$

where $a_1, a_2, ..., a_{(m-1)k+m}$ are uniform independent random bits, and let $S \in \mathbb{F}_2^{m\times m}$ be some fixed matrix. We aim to show the following lemma

Lemma 3.2.1 ([GT18])**.** $\mathbb{P}_A[\text{rank}(S + A) \leq m/2] \leq 2^{-km/16}$.

Let $B = S + A$ and let B_i denote the ith row of B. If rank$(B) \leq m/2$, then we can find a basis $B_{i_1}, B_{i_2}, ..., B_{i_{\text{rank}(B)}}$ of the row space spanned by B in the following constructive fashion:

1. Let i_1 be the index of the first nonzero row of B.
2. For each t, let i_t be the index of the first row of B that cannot be spanned by $B_{i_1}, ..., B_{i_{t-1}}$.

Let an index set $\mathcal{I} = \{i_1, ..., i_r\}, r \leq m/2$ be given and set $\mathcal{J} = [m] \setminus \mathcal{I}$. We have that

$$\forall j \in J, B_j \in \text{span}\{B_i : i \in I, i < j\} \tag{3.2.0.1}$$

Let an arbitrary $j \in J$ be given. Notice that if we fix the random bits $a_1, ..., a_{j-1}k$, the jth row is completely undetermined because the first entry of the jth row is $a_{(j-1)k+1}$.

Claim 3.2.2. Let $\mathcal{I}' = \mathcal{I} \cap [j-1]$ and $p = |\mathcal{I}'|$ and fix a vector $\mathbf{c} \in \{0,1\}^p$. We have that

$$\mathbb{P}[B_j = \sum_{i\in\mathcal{I}'} c_i B_i] = 2^{-m}. \tag{3.2.0.2}$$

Proof. Notice that for all $h \in [m]$, $B_{j,h} = S_{j,h} + a_{(j-1)k+h}$ where $S_{j,h}$ is fixed but $a_{(j-1)k+h}$ is not. Hence, since $\sum_{i\in\mathcal{I}'} c_i B_i$ is fixed, we have that the hth bit of this linear combination will be equal to the hth bit of B_j with probability exactly $1/2$ since $a_{(j-1)k+h}$ is uniformly chosen from $\{0, 1\}$. Also notice that each of these probabilities are independent since $a_{(j-1)k+h}$ are independently chosen. □

Let $\Delta = \lceil m/k \rceil$, then we can select an increasing sequence of $|J|/\Delta$ indices in J such that each two indices differ by at least Δ. Let $j_1, j_2, ..., j_t$ be such a sequence of indices where $t \geq |J|/\Delta$. For each $l \in [t]$, let E_l be the event that j_lth row is spanned by the rows indexed by $\mathcal{I} \cap [j_l - 1]$.

Claim 3.2.3. For all $l \in [t], \mathbb{P}[E_l|E_1, E_2, ..., E_{l-1}] \leq 2^{-m/2}$.

Proof. Notice that $j_l \geq j_{l-1} + \Delta =\geq j_{l-1} + \lceil m/k \rceil$. That is, $(j_l - 1)k \geq (j_{l-1} - 1)k + m$. On the other hand, given fixed bits $a_1, ..., a_{j_l-1}$, we can determine the rows $B_{j_1}, ..., B_{j_{l-1}}$ but $B_{j_l} = (a_{(j_l-1)k+1}, ..., a_{(j_l-1)k+m})$. Hence, we have

$$\mathbb{P}[E_l|E_1, E_2, ..., E_{l-1}] \leq \mathbb{P}[E_l|a_1, ..., a_{j_l-1}] \tag{3.2.0.3}$$

By claim 3.2.2, for a fixed linear combination, we have $\mathbb{P}[E_l|a_1, ..., a_{j_l-1}, \mathbf{c}] = 2^{-m}$. Let $\mathcal{I}' = \mathcal{I} \cap [j-1]$ and $p = |\mathcal{I}'|$. Because there are 2^p different values for $\mathbf{c}$, and recall that $p \le r \le \text{rank}(B) \le m/2$, by union bound, $\mathbb{P}[E_l|a_1, ..., a_{j_l-1}] \le 2^p 2^{-m} \le 2^{-m/2}$. ☐

We are now in a good shape to prove the following lemma.

Lemma 3.2.4. *Let E be the event that*

$$\forall j \in \mathcal{J}, B_j \in \textit{span}\{B_i : i \in \mathcal{I}, i < j\} \tag{3.2.0.4}$$

for a given index set $\mathcal{I}$. Then $\mathbb{P}[E] \le 2^{-mk/8}$.

Proof. Recall that $(j_1, j_2, ..., j_t)$ is a sequence of indices where $t \ge |J|/\Delta$ and each two indices are Δ apart. For each $l \in [t]$, let E_l be the event that j_lth row is spanned by the rows indexed by $\mathcal{I} \cap [j_l - 1]$. Hence,

$$\mathbb{P}[E] \le \mathbb{P}[E_1, E_2, ..., E_{t-1}, E_t] = \mathbb{P}[E_1]\mathbb{P}[E_2|E_1]...\mathbb{P}[E_t|E_1, E_2, ..., E_{t-1}] \le \left(2^{-m/2}\right)^t \tag{3.2.0.5}$$

Notice that

$$t \ge |J|/\Delta \ge \frac{m/2}{\lceil m/k \rceil} \ge k/4 \tag{3.2.0.6}$$

Therefore, $\mathbb{P}[E] \le \left(2^{-m/2}\right)^t \le 2^{-mk/8}$. ☐

proof of lemma 3.2.1. We can simply apply union bound among all possible choices of $\mathcal{I}$, which is less than 2^m. Hence, because we chose $k \ge 16, \mathbb{P}_A[\text{rank}(S + A) \le m/2] \le 2^m \mathbb{P}[E] \le 2^{-km/16}$. ☐

In this section we show a proof for Hankel matrices in the field $\mathbb{F}_2$ of the following theorem. The main idea of this proof is based on an averaging argument very similar to the one used to prove theorem 2.3.7.

Theorem 3.2.5 ([GT18])**.** *Let T be a random Toeplitz/Hankel matrix of size $N \times N$. Then, for every $r \in [\sqrt{n}, n/32]$, with probability $1 - o(1)$, the matrix T has rigidity $\Omega(\frac{n^3}{r^2 \log n})$.*

Chapter 4

Paturi-Pudlák Dimensions

We first introduce the definition of sparsity.

Definition 4.0.1 (Sparsity)**.** A vector $v \in \mathbb{F}^n$ is s**-sparse** if the number of non-zero coordinates in v is at most s. A matrix $A \in \mathbb{F}^{M \times N}$ is s**-sparse** if it has s nonzero entires. A is s**-row sparse** if each of its row is s-sparse. A subspace $V \in \mathbb{F}^M$ is s**-sparse** if it is the *column space* of a s-row sparse matrix B.

4.1. Friedman's result

In [Fri93], Friedman defined that an $N \times N$ matrix is A is (s, t)-rigid if for any s-sparse matrix B, we have that rank$(A + B) \geq t$.

Theorem 4.1.1 ([Fri93])**.** *For any constant $C_1 > 0$ there is a constant $C_2 > 0$ such that the following holds. Let $\mathbb{F}$ be a finite field of q elements. Let A be an $N \times N$ matrix such that the first $N/2$ rows are the basis of a linear error-correcting code in $\mathbb{F}^N$ of minimum distance $\geq C_1 N$. If B is any $N \times N$ matrix over $\mathbb{F}$ with at most k non-zero entries in each row, where $k \leq N/C_2$, then we have*

$$\text{rank}(A + B) \geq \frac{N}{C_2 k}(\log_q k + \log_q(q - 1)) \tag{4.1.0.1}$$

Let r denote the rank of $A + B$, notice that this theorem implies

$$k \geq \frac{N}{r}(\log_q k + \log_q(q - 1)) \geq \frac{N}{r} \log_q \frac{N}{r} \tag{4.1.0.2}$$

for small finite fields and $k \geq q$. Many people claim that this implies an $\Omega(\frac{N^2}{r} \log_q \frac{N}{r})$ lower bound of matrix rigidity. However, it is not very obvious from this theorem as the converse of the statement only ensures that some rows of B will have more than k entries.

proof of 4.1.1. Let $A_{N/2}$ denote the first $N/2$ rows of A and $B_{N/2}$ the first $N/2$ rows of B. We set $D_{N/2} \in \mathbb{F}_q^{N/2\times N}$ by $D_{N/2} = A + B$ and let r denote the rank of $_{N/2}$. Let S denote the linear space spanned by all vectors $w \in \mathbb{F}_q^{N/2}$ such that

$$w \cdot D_{N/2} = 0 \tag{4.1.0.3}$$

We see that S is a subspace of $\mathbb{F}_q^{N/2}$ with dimension $n/2 - r$.

Claim 4.1.2. Suppose t is an integer such that the size of a Hamming sphere of radius $t/2$ in $\mathbb{F}_q^{N/2}$ is at least q^r. Then there is a vector $w \in S$ with weight at most t.

Proof. Suppose on the contrary that the weight of w is greater than t for all $w \in S$. Let l denote the size of a Hamming sphere of radius $t/2$ in $\mathbb{F}_q^{N/2}$. Then

$$|S| \cdot l > |S| \cdot q^r = q^{N/2-r} \cdot q^r = q^{N/2} \tag{4.1.0.4}$$

However, we know that there are at most $q^{N/2}$ points in $\mathbb{F}_q^{N/2}$. That is

$$q^{N/2} \geq |S| \cdot l \tag{4.1.0.5}$$

Hence, we complete the proof by contradiction. ∎

Now, let $w \in S$ be a vector of weight t as defined in the claim above. We then have

$$0 = w \cdot D_{N/2} = w \cdot A_{N/2} + w \cdot B_{N/2} \tag{4.1.0.6}$$

Because all the rows in $A_{N/2}$ are independent, we have $w \cdot A_{N/2} \neq 0$ and thus $w \cdot B_{N/2} \neq 0$. Since each row of B has at most k non-zero entries and w has weight t, we have that the weight of $w \cdot B_{N/2}$ is at most tk. On the other hand, since the code represented by $A_{N/2}$ has minimum distance C_1N, we have that the weight of $w \cdot A_{N/2}$ is at least C_1N. Therefore, we must have

$$tk \geq C_1N \tag{4.1.0.7}$$

Take $t_0 = \lceil C_1N/k \rceil$, we then have the size of a Hamming sphere of radius $t/2$ in $\mathbb{F}_q^{N/2}$ is at most q^t because the weight of w must be greater than t_0 to achieve $0 = w \cdot D_{N/2}$. Then

$$q^r \geq \binom{n/2}{t_0/2}(q-1)^{t_0/2} \tag{4.1.0.8}$$

which is

$$r \geq \log_q \left[\binom{N/2}{t_0/2} (q-1)^{t_0/2} \right] \geq \log_q \binom{N/2}{t_0/2} + \frac{t_0}{2} \log_q (q-1) \geq \frac{t_0}{2} \frac{N}{k} \log_q k \tag{4.1.0.9}$$

Choosing $C_2 \sim 1/C_1$, we have

$$r \geq \log_q \left[\binom{N/2}{t_0/2} (q-1)^{t_0/2} \right] \geq \frac{n}{C_2 k} (\log_q k + \log_q (q-1)) \tag{4.1.0.10}$$

□

4.2. Strong Rigidity and Paturi-Pudlák Dimensions

In [PP06], Paturi and Pudlák introduced two dimensions that refine the notion of rigidity studied by Friedman.

In 4.1.1, Friedman also gave the following notion of strong rigidity.

Definition 4.2.1 (Strong Rigidity)**.** Let $V \subseteq \mathbb{F}^M$ be a subspace. V is (s,t)**-strongly rigid** if for any s-sparse subspace $U \in \mathbb{F}^M$ with $\dim(U) \leq \dim(V)$,

$$\dim(V \cap U) \leq \dim(V) - t \tag{4.2.0.1}$$

Definition 4.2.2 (Inner Dimension)**.** Let $V \subseteq \mathbb{F}^M$ be a subspace, and s be a positive integer less than M. We defined the **inner dimension** $d_V(s)$ **of V** by

$$d_V(s) := \max\{\dim(V \cap U) | U \in \mathbb{F}^M, \dim(U) \leq \dim(V), U \text{ is } s\text{-sparse}\} \tag{4.2.0.2}$$

Notice that the above definition is based on strong rigidity. To see the relation, we define $\rho_A(s)$ for a matrix $A \in \mathbb{F}^{M \times N}$.

$$\rho_A(s) := \min_B \{\text{rank}(A - B) : B \text{ is } s\text{-row sparse}\} \tag{4.2.0.3}$$

Proposition 4.2.3. *Let $A \in \mathbb{F}^{M \times N}$ with $M > N$ and $0 < s \leq N$ be given. Let V be the row space of A. Then*

$$\text{rank}(A) - d_V(s) \leq \rho_A(s) \tag{4.2.0.4}$$

Proof. Let B be the matrix that matches $\rho_A(s)$, i.e., rank$(A - B) = \rho_A(s)$. Let U be the row space of B and let W be the row space of $A - B$. Then we have $\dim(W) = \text{rank}(A - B) = \rho_A(s)$ and thus

$$\dim(V \cup U) \leq \dim(U) + \dim(W) = \dim(U) + \rho_A(s) \tag{4.2.0.5}$$

Hence, we obtain

$$\begin{aligned} \mathrm{rank}(A) - d_V(s) = \dim(V) - d_V(s) &= \dim(V \cap U) + \dim(V \cup U) - \dim(U) - d_V(s) \\ &\leq \dim(V \cap U) + \dim(U) + \rho_A(s) - \dim(U) - d_V(s) \\ &= \dim(V \cap U) - d_V(s) + \rho_A(s) \end{aligned} \tag{4.2.0.6}$$

Because B is s-row sparse, we have that U is s-sparse. Thus, $d_V(s) \geq \dim(V \cap U)$. Hence, we obtain $\mathrm{rank}(A) - d_V(s) \leq \rho_A(s)$. □

Definition 4.2.4 (Outer Dimension)**.** Let $V \subseteq \mathbb{F}^M$ be a subspace, and s be a positive integer less than M. We defined the **outer dimension** $D_V(s)$ **of V** by

$$D_V(s) := \max\{\dim(U) | U \in \mathbb{F}^M, V \subseteq U, U \text{ is } s\text{-sparse}\} \tag{4.2.0.7}$$

4.2.1 A Simple Bound

Proposition 4.2.5. *Let $V \subseteq \mathbb{F}^M$ be a subspace and s be a positive integer less than M. Then,*

$$d_V(s) + D_V(s) \geq 2\dim(V) \tag{4.2.1.1}$$

Proof. Let $V \subseteq \mathbb{F}^M$ be a subspace such that $V \subseteq U, U$ is s-sparse and $\dim(U) = D_V(s)$. Let $m = \dim(V)$ and W be an m-dimensional subspace of $\mathbb{F}^M$ such that $W \subseteq U$. Hence, $\dim(V \cap W) \leq d_V(s)$ and thus

$$\begin{aligned} 2\dim(V) &= \dim(V) + \dim(W) \\ &= \dim(V \cap W) + \dim(V \cup W) \\ &\leq d_V(s) + \dim(U) = d_V(s) + D_V(s) \end{aligned} \tag{4.2.1.2}$$

□

Theorem 4.2.6. *Let C be an $[n, k, d]$ linear code over $\mathbb{F}_2$. Then for $s \leq d/2$,*

$$\begin{aligned} D_C(s) &\geq k + \frac{d}{2s}\log(\frac{2sk}{d}) \\ d_C(s) &\leq k - \frac{d}{2s}\log(\frac{2sk}{d}) \end{aligned} \tag{4.2.1.3}$$

The proof is essentially the same as that of theorem 4.1.1. We use the following auxiliary lemma.

Lemma 4.2.7. *Let C be an $[n, k, d]$ linear code over $\mathbb{F}_2$. Then for $s \leq d/2$, then there exists a $[D_C(s), k, d/s]$-code.*

Proof. Consider the subspace $W \subseteq \mathbb{F}_2^n$ with $\dim(W) = D_C(s), C \subseteq W$ and W is s-sparse. Let $D = D_C(s)$ and $\{w_1, ..., w_D\}$ be a basis of W where each $w_i, i \in [D]$ is s-sparse. Hence, for any $x \in C$, we have that there exists a $y \in \mathbb{F}_2^D$ such that

$$x = \sum_{i=1}^{D} y_i w_i \tag{4.2.1.4}$$

Let E be the set of all such y for all x. Then, we have $\dim(E) = \dim(C) = k$. Let $y' \in E$ be a nonzero vector with minimum weight. Because $x' = \sum_{i=1}^{D} y'_i w_i$ has weight at least d and each w_i is s-sparse, we have that at least d/s coordinates of y' is nonzero. Hence, we obtain that E is a $[D_C(s), k, d/s]$-code. □

Proof of theorem 4.2.6. Using the sphere packing bound on the $[D = D_C(s), k, d/s]$-code E we just constructed. We have that the Hamming balls of radius at most $d/2s$ at each vector in B do not intersect with each other. Hence,

$$\sum_{j=1}^{d/2s} \binom{D}{j} \leq 2^{D-k} \tag{4.2.1.5}$$

Notice that

$$\sum_{j=1}^{d/2s} \binom{D}{j} \geq \binom{D}{d/2s} \geq \binom{k}{d/2s} \geq (2sk/d)^{d/2s} \tag{4.2.1.6}$$

Hence,

$$D - k \geq \frac{d}{2s} \log(\frac{2sk}{d}) \tag{4.2.1.7}$$

Let $U \subseteq \mathbb{F}_2^n$ be a subspace with $\dim(U) = k$, U is s-sparse and $\dim(C \cap U) = d_C(s)$. Then $F = C \cap U$ is simply a $[n, d_C(s), d]$ code. Applying lemma 4.2.7 again, we obtain a $[D_F(s), d_C(s), d/s]$-code. Hence, because $D_F(s) \leq \dim(U) = k$, equation 4.2.1.7 implies

$$k - d_C(s) \geq D_F(s) - d_C(s) \geq \frac{d}{2s} \log(\frac{2sk}{d}) \tag{4.2.1.8}$$

which means $d_C(s) \leq k - (d/2s)\log(2sk/d)$. □

4.2.2 Connection Between Outer and Inner Dimensions

Theorem 4.2.8 ([DGW19]). *Let t and k be positive integers and let $0 < \epsilon < 1$. If $A \in \mathbb{F}^{M \times N}$ is a matrix whose columns space $V \subseteq \mathbb{F}^M$ has an outer dimension*

$$D_V(tk + N\epsilon^k) \geq \frac{N}{1 - \epsilon} \tag{4.2.2.1}$$

then for some $N' \geq N\epsilon^k$, A contains a submatrix $B \in \mathbb{F}^{M \times N'}$ whose columns space $U \subseteq \mathbb{F}^M$ has an inner dimension

$$d_U \leq \text{rank}(B) - \epsilon N' \tag{4.2.2.2}$$

4.3. Row Rigidity

In this section, we introduce the definition of *row rigidity* from [DGW19], specifically for rectangular matrices. In the next section, we will discuss the link between rigidity and data structure lower bounds.

Definition 4.3.1 (Rigidity for Rectangular Matrices)**.** A matrix $A \in \mathbb{F}^{M \times N}$ is said to be (r, s)**-rigid** if for any matrix $B \in \mathbb{F}^{M \times N}$ with dens$(B) \leq s$, we have $A + B$ has rank at least r. A matrix $A \in \mathbb{F}^{M \times N}$ is said to be (r, s)**-strongly rigid** if for any invertible matrix $C \in \mathbb{F}^{N \times N}$, we have $A \times C$ is (r, s) rigid.

Definition 4.3.2 (Row Rigidity)**.** A matrix $A \in \mathbb{F}^{M \times N}$ is said to be (r, s)**-row rigid** if for any matrix s-sparse $B \in \mathbb{F}^{M \times N}$, we have $A + B$ has rank at least r.

Definition 4.3.3 (Strong Row Rigidity)**.** A matrix $A \in \mathbb{F}^{M \times N}$ is said to be (r, s)**-strongly row rigid** if for any invertible matrix $C \in \mathbb{F}^{N \times N}$, we have $A \times C$ is (r, s)-row rigid.

4.3.1 Strong rigidity is equivalent to small inner dimension

The following lemma shows that the definition of strong row rigidity of rectangular matrices is equivalent to the strong rigidity of subspaces as defined by Friedman. In particular, we limit our attention to matrices with more rows than columns, i.e. $M > N$.

Lemma 4.3.4. *Let matrix $A \in \mathbb{F}^{M \times N}$ have rank N and let $V \subseteq \mathbb{F}^M$ be its columns space. Then the following are equivalent:*

1. *A is (r, s)-strongly row rigid.*
2. *$d_V(s) \leq$ rank$(A) - r$.*
3. *V is not contained in a subspace of the form $E \cup F$ where $E, F \subseteq \mathbb{F}^M$ are subspaces with $\dim(E) \leq N, \dim(F) < r$ and E is s-sparse.*

Proof. $(1 \Rightarrow 2)$: Suppose $d_V(t) >$ rank$(A) - r$. By the definition of inner dimension 4.2.2, there exists a subspace $U \in \mathbb{F}^M, \dim(U) \leq \dim(V) =$ rank$(A) = N, U$ is s-sparse and $\dim(U \cap V) >$ rank$(A) - r$. In other words, there exists a subspace $W \in \mathbb{F}^M$ with $\dim(W) < r$ such that $V = U \cup W$. Let $C \in \mathbb{F}^{M \times N}$ be a s-row sparse basis matrix of U and $B \in \mathbb{F}^{M \times N}$ a basis matrix of

W. There exists an invertible matrix $T \in \mathbb{F}^{N\times N}$ such that

$$A = C \times T + B \tag{4.3.1.1}$$

This implies $A \times T^{-1} = C + B \times T^{-1}$ is not (r,s)-row rigid because rank$(A \times T^{-1} - C)$ = rank$(B\times T^{-1}) < r$. This means that A is not (r,s)-strongly row rigid and leads to a contradiction.
$(2 \Rightarrow 3)$: Because $d_V(s) \leq$ rank$(A) - r$, for all subspaces $E \in \mathbb{F}^M$ such that $\dim(E) \leq N = \dim(V)$ and E is s-sparse, we have $\dim(V \cap E) \leq N - r$. That is, for any subspace $F \in \mathbb{F}^M$ with $V \subseteq E \cup F$, we have

$$\begin{aligned} N - r \geq \dim(V \cap E) &= \dim(V) + \dim(E) - \dim(V \cup E) \\ &\geq \dim(V) + \dim(E) - \dim(E \cup F) \\ &= \dim(V) + \dim(E) - \dim(E) - \dim(F) + \dim(E \cap F) \\ &= N - \dim(F) + \dim(E \cap F) \end{aligned} \tag{4.3.1.2}$$

This implies $\dim(F) \geq r$.
$(3 \Rightarrow 1)$: Take any invertible $T \in \mathbb{F}^{N\times N}$, if we write

$$A \times T = C + B \tag{4.3.1.3}$$

where $C \in \mathbb{F}^{M\times N}$ be a s-row sparse and $B \in \mathbb{F}^{M\times N}$. Let E be the columns space of C and F the columns space of B. Because T is invertible and, rank$(A \times T)$ = rank$(A) = \dim(V)$. As we have just seen, $\dim(F) \geq r$, hence rank$(B) > r$ and thus A must be (r,s)-strongly row rigid. □

Remark 4.3.5. Notice that by proposition 4.2.3, A is $(\rho_A(s), s)$-strongly row rigid if and only if $d_V(s) =$ rank$(A) - \rho_A(s)$.

4.3.2 From Row Rigidity to General Rigidity

Theorem 4.3.6 ([DGW19])**.** *Let $A \in \mathbb{F}^{M\times N}$ be a rectangular matrix, $E \in \mathbb{F}^{L\times M}$ a $(t, \delta, 3/4)$-linear locally decodable code and set $B := EA$. Then,*

1. *If $A \in \mathbb{F}^{M\times N}$ is $(r, s+1)$-row rigid, then $B \in \mathbb{F}^{L\times N}$ is $(r, (\delta sL)/t)$-rigid.*

2. *If $A \in \mathbb{F}^{M\times N}$ is $(r, s+1)$-strongly row rigid, then $B \in \mathbb{F}^{L\times N}$ is $(r, (\delta sL)/t)$-strongly rigid.*

We need a few tools from locally decodable code to prove this theorem.

Notation 4.3.7. We use dist(u, v) to denote the Hamming distance between two vectors u, v.

Definition 4.3.8. A linear code $C : \mathbb{F}^M \to \mathbb{F}^N$ is said to be (t, δ, ϵ)**-locally decodable** if there exists a randomised decoding algorithm $\mathcal{A}$ such that for all $m \in \mathbb{F}^M$ and all $w \in \mathbb{F}^N$ such that $\text{dist}(C(m), w) \leq \delta$:

1. For every index $i \in [M]$
$$\mathbb{P}[\mathcal{A}(w, i) = m_i] \geq 1 - \epsilon, \tag{4.3.2.1}$$
where the probability is taken over the random coin tosses of the algorithm $\mathcal{A}$.

2. $\mathcal{A}$ makes at most t queries to w.

We abuse the notation and write $C \in \mathbb{F}^{M\times N}$ as its generating matrix.

Lemma 4.3.9 ([GKST02], [DS07])**.** *Let $C \in \mathbb{F}^{M\times N}$ be a $(t, \delta, 3/4)$-linear locally decodable code and let R be a set of rows of C with $|R| \geq (1 - \delta)M$. For any $i \in [N]$, there exists a set of t rows in R which spans the ith standard basis vector e_i.*

proof of theorem 4.3.6. Let $A \in \mathbb{F}^{M\times N}$ be $(r, s+1)$-row rigid and suppose that B is not $(r, (\delta sL)/t)$-rigid. Then we have $B = D + S$ where $D \in \mathbb{F}^{L\times N}$ has rank at most r and $S \in \mathbb{F}^{L\times N}$ has density $\text{dens}(S) \leq (\delta sL)/t$. Let S' be the set of row of S that are s/t-sparse. By averaging argument, we have that $|S'| \geq (1 - \delta)L$. Let D' be the corresponding rows in D. Because A is $(r, s + 1)$-row rigid, some rows A_i has Hamming distance at least $(s + 1)$ from the space generated by D.
On the other hand, by lemma 4.3.9, there exist t rows in D' and S' which spans A_i. This means that A_i has a Hamming distance at most $t \cdot (s/t) = s$ from the row space of L'.
Therefore, by contradiction, we must have B is $(r, (\delta sL)/t)$-rigid.
Let $A \in \mathbb{F}^{M\times N}$ be $(r, s + 1)$-strongly row rigid, then for all invertible matrix $T \in \mathbb{F}^{N\times N}$ such that AT is $(r, s + 1)$-row rigid. Notice that

$$E \times (A \times T) = (E \times A) \times T = B \times T \tag{4.3.2.2}$$

where $B \in (r, (\delta sL)/t)$. As we have just shown, $B \times T$ is $(r, (\delta sL)/t)$-rigid. Since T is arbitrary, we have that B is $(r, (\delta sL)/t)$-strongly rigid. □

The following corollary shows that we can obtain rigid square matrices from rigid rectangular matrices.

Corollary 4.3.10 (Rectangular Matrices to Square Matrices, [DGW19])**.** *For every constant $\alpha > 0$, and an $(r, s + 1)$-row rigid matrix $A \in \mathbb{F}^{M\times N}$, we can construct a square matrix $B \in \mathbb{F}^{L\times L}, L = M^{O(1/\alpha)}$, which is*

$$(r, \frac{L}{N} \cdot \frac{s}{(\log M)^{1+\alpha}})\text{-row rigid and } (r, \frac{L^2}{N} \cdot \frac{s}{(\log M)^{1+\alpha}})\text{-rigid.}$$

in polynomial time.

To prove this theorem we use the following result from locally decodable code without proof. This lemma allows us to construct linear locally decodable codes in polynomial time.

Lemma 4.3.11 ([Dvi11]). *For every $\alpha, \epsilon > 0$, there exists $\delta = \delta(\epsilon) > 0$ and an explicit family of $((\log N)^{1+\alpha}, \delta, \epsilon)$-linear locally decodable codes $C \in \mathbb{F}^{\mathbb{M}\times\mathbb{N}}$ for $M = N^{O(1/\alpha)}$.*

proof of corollary 4.3.10. Let $L = M^{O(1/\alpha)}$ be a multiple of N and δ some constant. We also let $C \in \mathbb{F}^{L\times M}$ be a $((\log M)^{1+\alpha}, \delta, 3/4)$-linear locally decodable code. Hence, we can construct a square matrix $B \in \mathbb{F}^{L\times L}$ by putting side by side (L/N) copies of $C \times A$.
It remains to show that B is rigid. By theorem 4.3.6, we have $C \times A$ is $(r, (\delta s L)/(\log M)^{1+\alpha})$-rigid. Hence,

$$\mathcal{R}_B(r) \geq \frac{L}{N} \cdot \frac{\delta s L}{(\log M)^{1+\alpha}}$$

and by averaging over L rows, B must be $(r, (\delta s L)/[N \cdot (\log M)^{1+\alpha}])$-row rigid. Setting $\delta = 1$ or some suitable constant, we get the desired result. $\square$

Remark 4.3.12 (Link to Friedman's result). Notice that in theorem 4.1.1, the first $N/2$ rows of A is $(r, (N/r)\log(N/r))$-row rigid. By corollary 4.3.10, we can construct a matrix $B \in \mathbb{F}^{N\times N}$ with

$$\mathcal{R}_B(r) = O\Big(\frac{N^2}{(\log N)r} \cdot \log(\frac{N}{r})\Big) \tag{4.3.2.3}$$

4.4. Linear Data Structure

Theorem 4.4.1 ([DGW19]). *A data structure lower bound of $t \geq \log^c n$ in the group (linear) model for computing a linear map $M \in \mathbb{F}^{m\times n}$, even against data structures with arbitrarily small linear space $s = (1+\epsilon)n$, yields an $(\epsilon n', d)$-row-rigid matrix $M' \in \mathbb{F}^{m\times n'}$ with $\epsilon n' \geq d \geq \Omega(\log^{c-1} n)$. Moreover, if M is explicit, then $M' \in \mathbf{P}^{\mathbf{NP}}$.*

Chapter 5

Rigid Sets

So far we have seen the trade-off between the values of dimensions and distance that can be obtained by explicit sets of size n. The study of rigid sets aim to investigate the trade-off between the values of *size* and *distance*, when the value of *dimension* is fixed.

Definition 5.0.1. For $x \in \mathbb{F}_2^N, U \subseteq \mathbb{F}_2^N$, we define the **Hamming distance from** x **to** U by

$$\text{dist}(x, U) = \min_{u \in U} |x + u| \tag{5.0.0.1}$$

where $|v|$ denotes the Hamming weight of v.

Definition 5.0.2 (Rigid Sets)**.** A set $S \subseteq \mathbb{F}_2^N$ is called (N, k, d)**-rigid** if for every linear subspace $U \subseteq \mathbb{F}_2^N, \dim(U) = k$, we have

$$\max_{s \in S} \text{dist}(s, U) \geq d \tag{5.0.0.2}$$

Let $A \in \mathbb{F}_2^{M \times N}, M = |S|$ be the matrix whose rows are the elements of S. Notice that A is (k, d)-rigid if and only if S is (N, k, d)-rigid.

Theorem 5.0.3 ([SY11])**.** *Let q be a prime power. For every $0 \leq d \leq O(N)$, there exists an explicit set $(N, N/2, d)$-rigid set $S \subseteq \mathbb{F}_q^N$ of size $2^{O(d)N/d}$.*

Notation 5.0.4. Let $I \subseteq [N]$ be a set of coordinates. For a vector $x \in \mathbb{F}^N$, we write $x|_I$ to denote the vector x restricted to the coordinates in I. Similarly, for a linear subspace $U \subseteq \mathbb{F}^N$, we write $U|_I$ to denote the linear subspace U restricted to the coordinates in I.

Lemma 5.0.5. *Let q be a prime power and $U \subseteq \mathbb{F}_q^N$ be a linear subspace with $\dim(U) = k$, then*

$$\mathbb{P}_{x \in \{0,1\}^N}[x \in U] \leq \frac{1}{2^{N-k}} \tag{5.0.0.3}$$

Proof. Let $I \subseteq [N]$ be the set of coordinates such that $U|_I = \mathbb{F}_q^N$ and $J = [N] \setminus I$. Hence, we note that a vector $x = x|_I + x|_J \in U$ is uniquely determined by $x|_I$ because $x|_J$ is the zero vector. Hence, for a random vector $x \in \{0,1\}^N$, there is at most 2^{k-N} chance that x is in U. $\square$

Lemma 5.0.6. *Let q be a prime power. For every $\epsilon > 0$, there exists a $\delta > 0$ such that for all linear subspaces $U \subseteq \mathbb{F}_q^N, \dim(U) \leq (1-\epsilon)N$, there exists a point $x \in \{0,1\}^n$ such that*

$$\text{dist}(x, U) \geq \delta N \tag{5.0.0.4}$$

Proof. Let U be given. We note that for a random vector $x \in \{0,1\}^N$,

$$\mathbb{P}[\text{dist}(x, U) \leq \delta N] = \mathbb{P}[\exists I \subseteq [N], |I| = (1-\delta)N \text{ such that } x|_I \in L|_I] \tag{5.0.0.5}$$

For a fixed set I with $|I| = (1-\delta)N$, we have by lemma 5.0.5, we have

$$\mathbb{P}[x|_I \in U|_I] = \frac{1}{2^{(1-\delta)N-(1-\epsilon)N}} = \frac{1}{2^{(\epsilon-\delta)N}} \tag{5.0.0.6}$$

Hence, by union bound on all possible set I of size $(1-\delta)N$, the probability

$$\mathbb{P}[\text{dist}(x, U) \leq \delta N] \leq \binom{N}{\delta N} \frac{1}{2^{(\epsilon-\delta)N}} \tag{5.0.0.7}$$

is negligible when δ is sufficiently smaller than ϵ. $\square$

proof of theorem 5.0.3. As a consequence of lemma 5.0.6, let δ be the constant that for all linear subspace $U \subseteq \mathbb{F}_q^N, \dim(U) = N/2$, there exists a point p in $\mathbb{F}_q^N$ that is more than δN-far from U. To obtain S, we first split the coordinates into cN/d disjoint sets $Z_1, Z_2, ..., Z_{\delta N/d}$, each of size d/δ. For each set Z_i, we let W_i be the set of all binary vectors $x_i \in \{0,1\}^N$ with support on this set Z_i. That is, each x_i has some value 1 on some coordinates in set Z_i and has 0 on every other coordinates. Let $S = \cup_i W_i$ consist of all these vectors. Hence,

$$|S| = 2^{O(d)} N/d \tag{5.0.0.8}$$

and every vector in $\mathbb{F}_q^N$ is the sum of at most $\delta N/d$ vectors in S.
Let a linear subspace $U \subseteq \mathbb{F}_q^N, \dim(U) = N/2$ be given. Suppose every vector in S is at most d-far from U. That is, any vector in S is the sum of one vector in U and at most d unit vectors. Because every vector v in $\mathbb{F}_q^N$ is the sum of at most $\delta N/d$ vectors in S, v must also be the sum of a vector in U and at most $d \times (\delta N/d) = \delta N$ unit vectors. Hence, no point p will be more than δN-far from U, which gives us a contradiction. $\square$

Corollary 5.0.7 ([APY09])**.** *For every $0 \leq d \leq O(N)$, there exists an explicit $(N, N/2, d)$-rigid set $S \subseteq \mathbb{F}_2^N$ of size $2^{O(d)} N/d$.*

5.1. Strong Rigid Sets

A new approach for constructing rigid sets by applying *U-polynomials* was introduced by Alon and Cohen [AC15].

Definition 5.1.1. For a subspace $U \subseteq \mathbb{F}_2^N$, the **U-polynomial** $p_U :\subseteq \mathbb{F}_2^N \to \mathbb{R}$ is defined as

$$p_{U,\rho}(x) = \frac{1}{W_\rho(U)} \cdot \sum_{u \in U} \rho^{|u|} \cdot (-1)^{\langle u,x \rangle} \tag{5.1.0.1}$$

where $W_\rho(U) = \sum_{u \in U} \rho^{|u|}$ is the **weight enumerator** of U with parameter $\rho \in (0,1)$.

Theorem 5.1.2 ([AC15])**.** *Let the parameter $\rho \in (0,1)$ and the linear subspace $U \subseteq \mathbb{F}_2^N$ be given. Then, for every $x \in \mathbb{F}_2^N$,*

$$\text{dist}(x, U) = \Omega(\log \frac{1}{p_{U^\perp,\rho}(x)}) \tag{5.1.0.2}$$

5.1.1 Fourier Analysis

Before we prove theorem 5.1.2, we first show the following.

Theorem 5.1.3 ([AC15])**.** *Let $U \subseteq \mathbb{F}_2^N$ be a linear subspace. Then, for any parameter $\rho \in (0,1)$ and any point $x \in \mathbb{F}_2^N$,*

$$\text{dist}(x, U) \geq \Big(\log \frac{1+\rho}{1-\rho}\Big)^{-1} \cdot \log \frac{1}{p_{U^\perp,\rho}(x)} \tag{5.1.1.1}$$

We will need a few definitions and tools from Fourier analysis and error correcting codes first.

Definition 5.1.4. We define the **inner product** $\langle \cdot, \cdot \rangle$ on pairs of function $f, g : \mathbb{F}_2^N \to \mathbb{R}$ by

$$\langle f, g \rangle = 2^{-N} \sum_{x \in \mathbb{F}_2^N} f(x) g(x) \tag{5.1.1.2}$$

Definition 5.1.5 (Fourier Expansion)**.** Every function $f : \mathbb{F}_2^N \to \mathbb{R}$ can be uniquely expressed as a multilinear polynomial,

$$f(x) = \sum_{\alpha \in \mathbb{F}_2^N} \hat{f}(\alpha) \chi_\alpha(x) \tag{5.1.1.3}$$

where $\chi_\alpha(x) = (-1)^{\langle \alpha, x \rangle}$. This expression is called the **Fourier expansion** of f, and the real number $\hat{f}(\alpha) = \langle f, \chi_\alpha \rangle$ is called the **Fourier coefficient** of f on S. Collectively, the coefficients are called the **Fourier spectrum** of f.

Definition 5.1.6 (Noise Operator)**.** For $0 \geq \rho \geq 1$ and $f : \mathbb{F}_2^N \to \mathbb{R}$, we define the **noise operator**

with parameter $\rho, T_\rho(f) : \mathbb{F}_2^N \to \mathbb{R}$ on the function f by

$$T_\rho(f)(x) = \sum_{y \in \mathbb{F}_2^N} \left(\frac{1-\rho}{2}\right)^{|y|} \cdot \left(\frac{1+\rho}{2}\right)^{N-|y|} \cdot f(x+y) \tag{5.1.1.4}$$

Proposition 5.1.7. $\widehat{T_\rho(f)}(\alpha) = \rho^{|\alpha|}\hat{f}(\alpha)$.

Definition 5.1.8. Let the parameter $\rho \in (0,1)$ and the linear subspace $U \subseteq \mathbb{F}_2^N$ be given. The function $\text{energy}_{U,\rho} : \mathbb{F}_2^N \to \mathbb{R}$ is defined as

$$\text{energy}_{U,\rho}(x) = \frac{1}{W_\rho(U)} \cdot \sum_{u \in U} \rho^{|u+x|} \tag{5.1.1.5}$$

Notice that $\text{energy}_{U,\rho}(x) \in (0,1]$ and $\text{energy}_{U,\rho}(x) = 1$ if and only if $x \in U$.

Theorem 5.1.9 (MacWilliam's Theorem [MS77])**.** *Let $U \subseteq \mathbb{F}_2^N$ be a linear subspace with $\dim U = k$. Then, for any parameter $\rho \in (0,1)$,*

$$W_\rho(U^\perp) = \frac{(1+\rho)^N}{2^k} \cdot W_{\frac{1-\rho}{1+\rho}}(U) \tag{5.1.1.6}$$

proof of theorem 5.1.3. We use $\mathbf{1}_U : \mathbb{F}_2^N \to \{0,1\}$ to denote the **indicator function** for U, i.e., $\mathbf{1}_U = 1$ if and only if $x \in U$. Then,

$$\begin{aligned} T_\rho(\mathbf{1}_U)(x) &= \sum_{y \in \mathbb{F}_2^N} \left(\frac{1-\rho}{2}\right)^{|y|} \cdot \left(\frac{1+\rho}{2}\right)^{N-|y|} \cdot \mathbf{1}_U(x+y) \\ &= \left(\frac{1+\rho}{2}\right)^{n} \cdot \sum_{y \in \mathbb{F}_2^N} \left(\frac{1-\rho}{1+\rho}\right)^{|y|} \cdot \mathbf{1}_U(x+y) \\ &= \left(\frac{1+\rho}{2}\right)^{n} \cdot \sum_{u \in U} \left(\frac{1-\rho}{1+\rho}\right)^{|x+u|} \\ &= \left(\frac{1+\rho}{2}\right)^{n} \cdot W_{\frac{1-\rho}{1+\rho}}(U) \cdot \text{energy}_{U,\frac{1-\rho}{1+\rho}}(x) \end{aligned} \tag{5.1.1.7}$$

Note that because U is a subspace, we have for $\alpha \notin U^\perp$, $\chi_\alpha(x) = 1$ for exactly half of the time and $\chi_\alpha(x) = -1$ for exactly the other half,

$$\langle \mathbf{1}_U, \chi_\alpha \rangle = \frac{1}{2^N}\Big(\sum_{x \in U} f(x)\chi_\alpha(x) + \sum_{x \notin U} f(x)\chi_\alpha(x)\Big) = \frac{1}{2^N}\sum_{x \in U} f(x)\chi_\alpha(x) = 0 \tag{5.1.1.8}$$

As for $\alpha \in U^\perp$,

$$\langle \mathbf{1}_U, \chi_\alpha \rangle = \frac{1}{2^N}\Big(\sum_{x \in U} f(x)\chi_\alpha(x) + \sum_{x \notin U} f(x)\chi_\alpha(x)\Big) = \frac{1}{2^N}\sum_{x \in U} \chi_\alpha(x) = \frac{1}{2^N} \cdot 2^k = 2^{k-N} \tag{5.1.1.9}$$

Therefore,

$$\widehat{\mathbf{1}_U}(\alpha) = \begin{cases} 2^{k-N}, & \alpha \in U^{\perp} \\ 0, & \text{otherwise} \end{cases} \tag{5.1.1.10}$$

By proposition 5.1.7,

$$\begin{aligned} T_\rho(\mathbf{1}_U)(x) &= \sum_{\alpha \in \mathbb{F}_2^N} \widehat{T_\rho(\mathbf{1}_U)}(\alpha) \cdot \chi_\alpha \\ &= \sum_{\alpha \in \mathbb{F}_2^N} \rho^{|\alpha|} \widehat{\mathbf{1}_U}(\alpha) \cdot \chi_\alpha \\ &= \sum_{\alpha \in U^{\perp}} \rho^{|\alpha|} \widehat{\mathbf{1}_U}(\alpha) \cdot \chi_\alpha \\ &= 2^{k-N} \sum_{\alpha \in U^{\perp}} \rho^{|\alpha|} \cdot \chi_\alpha \end{aligned} \tag{5.1.1.11}$$

By definition 5.1.1, we have

$$\begin{aligned} T_\rho(\mathbf{1}_U)(x) &= 2^{k-N} \sum_{\alpha \in U^{\perp}} \rho^{|\alpha|} \cdot \chi_\alpha \\ &= 2^{k-N} \cdot W_\rho(U^{\perp}) \cdot p_{U^{\perp},\rho}(x) \end{aligned} \tag{5.1.1.12}$$

By MacWilliam's Theorem, 5.1.9, we have

$$\begin{aligned} T_\rho(\mathbf{1}_U)(x) &= 2^{k-N} \cdot W_\rho(U^{\perp}) \cdot p_{U^{\perp},\rho}(x) \\ &= 2^{k-N} \cdot \frac{(1+\rho)^N}{2^k} \cdot W_{\frac{1-\rho}{1+\rho}}(U) \cdot p_{U^{\perp},\rho}(x) \\ &= \left(\frac{1+\rho}{2}\right)^N \cdot W_{\frac{1-\rho}{1+\rho}}(U) \cdot p_{U^{\perp},\rho}(x) \end{aligned} \tag{5.1.1.13}$$

Combining equation 5.1.1.7 and 5.1.1.13, we have

$$\text{energy}_{U,\frac{1-\rho}{1+\rho}}(x) = p_{U^{\perp},\rho}(x) \tag{5.1.1.14}$$

Let $d = \text{dist}(x, U)$. Then there exists $w \in U$ such that $|x + w| = d$. By definition 5.1.8, we have

$$W_{\frac{1-\rho}{1+\rho}}(U) \cdot \text{energy}_{U,\frac{1-\rho}{1+\rho}}(x) = \sum_{u \in U} \left(\frac{1-\rho}{1+\rho}\right)^{|u+x|} \tag{5.1.1.15}$$

Because U is a subspace,

$$\sum_{u \in U} \left(\frac{1-\rho}{1+\rho}\right)^{|u+x|} = \sum_{u \in U} \left(\frac{1-\rho}{1+\rho}\right)^{|u+x+w|} \tag{5.1.1.16}$$

Using triangle inequality, we have $|u+x+w| \le |u| + |x+w|$. Thus,

$$\begin{aligned}\sum_{u\in U}\left(\frac{1-\rho}{1+\rho}\right)^{|u+x+w|} &\ge \sum_{u\in U}\left(\frac{1-\rho}{1+\rho}\right)^{|u|+|x+w|}\\ &= \left(\frac{1-\rho}{1+\rho}\right)^{d}\sum_{u\in U}\left(\frac{1-\rho}{1+\rho}\right)^{|u|} \\ &= \left(\frac{1-\rho}{1+\rho}\right)^{d}\cdot W_{\frac{1-\rho}{1+\rho}}(U)\end{aligned} \tag{5.1.1.17}$$

In summary, we obtain

$$p_{U^\perp,\rho}(x) = \text{energy}_{U,\frac{1-\rho}{1+\rho}}(x) = \frac{1}{W_{\frac{1-\rho}{1+\rho}}(U)}\cdot\sum_{u\in U}\left(\frac{1-\rho}{1+\rho}\right)^{|u+x+w|} \ge \left(\frac{1-\rho}{1+\rho}\right)^{d} \tag{5.1.1.18}$$

which concludes the proof. □

Proposition 5.1.10. *Let the parameter $\rho \in (0,1)$ and the linear subspace $U \subseteq \mathbb{F}_2^N, \dim U = N/2$ be given. Then*

$$W_\rho(U) \ge \left(\frac{1+\rho}{\sqrt{2}}\right)^N \tag{5.1.1.19}$$

Proof. □

Definition 5.1.11 (Strong Rigid Sets). A set $S \subseteq \mathbb{F}_2^N$ is called **strong** (N,k,d)**-rigid** if for every linear subspace $U \subseteq \mathbb{F}_2^N, \dim(U) = k$, we have

$$\mathbb{E}_{s\sim S}[\text{dist}(s,U)] \ge d \tag{5.1.1.20}$$

Definition 5.1.12 (Small Biased Sets). We say that a set $S \subseteq \mathbb{F}_2^N$ is ϵ-biased if for every nonzero $\alpha \in \mathbb{F}_2^N$,

$$|\mathbb{E}_{s\sim S}|[(-1)^{\langle\alpha,s\rangle}] \le \epsilon \tag{5.1.1.21}$$

Theorem 5.1.13 ([AC15]). *For every $0 \le d \le cN$ for some suitable constant $0 < c < 1$. If $S \subseteq \mathbb{F}_2^N$ is an $\exp(-d)$-biased set, then S is $(N, N/2, d)$-strong rigid.*

5.2. Linear Data Structure and Rigidity

Definition 5.2.1 (Systematic Linear Model). For a set $S \subseteq \mathbb{F}_2^N$, we define $T(S,k)$ by

$$T(Q,k) := \max_{v\in\mathbb{F}_2^N}\Big(\min\{d| \text{ can compute } \langle q,v\rangle \forall q \in Q\}\Big) \tag{5.2.0.1}$$

where we are only allowed to output a linear function of k precomputed linear functions of v along with any d bits of v.

Theorem 5.2.2 ([RR20]). *A set set $S \subseteq \mathbb{F}_2^N$ is (N, k, d)-rigid if and only if $T(S, k) \geq d$.*

Chapter 6

Complexity Theory

6.1. FNP

The complexity class FNP is the function-problem extension of the decision-problem class NP. Formally, a relation $R(x, y)$ is in FNP if there exists a non-deterministic polynomial-time Turing machine M such that for any input x, $M(x)$ outputs y such $R(x, y) = 1$ or rejects if no such y exists.

Theorem 6.1.1 ([AC19])**.** *There is an absolute constant $\delta > 0$ such for all prime powers $q = pr$ and all constants $\epsilon > 0$, there is a $\mathbf{P}^{\mathbf{NP}}$ machine M such that, for infinitely many N , on input 1^N, M outputs an $N \times N$ matrix $H_N \in \{0, 1\}^{N \times N}$ such that $\mathcal{R}_{H_N}(2^{n^{1/4-\epsilon}}) \geq \delta N^2$ over $\mathbb{F}_q$.*

Theorem 6.1.2 ([BHPT20])**.** *There is a constant $0 < \delta < 1$ such that there is an FNP-machine that for infinitely many N, on input 1^N outputs an $N \times N$ matrix that is $(\delta \cdot N^2, 2^{\log N/\Omega(\log \log N)})$-rigid.*

Bibliography

[AC15] Noga Alon and Gil Cohen. On rigid matrices and u-polynomials. *computational complexity*, 24(4):851–879, 2015.

[AC19] Josh Alman and Lijie Chen. Efficient construction of rigid matrices using an np oracle. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1034–1055, Nov 2019.

[APY09] Noga Alon, Rina Panigrahy, and Sergey Yekhanin. Deterministic approximation algorithms for the nearest codeword problem. In Irit Dinur, Klaus Jansen, Joseph Naor, and José Rolim, editors, *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques*, pages 339–351, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.

[AW17] Josh Alman and Ryan Williams. Probabilistic rank and matrix rigidity. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 641–652, New York, NY, USA, 2017. Association for Computing Machinery.

[BCS97] Peter Bürgisser, Michael Clausen, and Mohammad Amin Shokrollahi. *Linear Complexity*, pages 305–349. Springer Berlin Heidelberg, Berlin, Heidelberg, 1997.

[BHPT20] Amey Bhangale, Prahladh Harsha, Orr Paradise, and Avishay Tal. Rigid matrices from rectangular pcps. *arXiv preprint arXiv:2005.03123*, 2020.

[Che05] Mahdi Cheraghchi. On matrix rigidity and the complexity of linear forms. 2005.

[DGW19] Zeev Dvir, Alexander Golovnev, and Omri Weinstein. Static data structure lower bounds imply rigidity. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2019, pages 967–978, New York, NY, USA, 2019. Association for Computing Machinery.

[DL19] Zeev Dvir and Allen Liu. Fourier and circulant matrices are not rigid, 2019.

[DS07] Zeev Dvir and Amir Shpilka. Locally decodable codes with two queries and polynomial identity testing for depth 3 circuits. *SIAM Journal on Computing*, 36(5):1404–1434, 2007.

[Dvi11] Zeev Dvir. On matrix rigidity and locally self-correctable codes. *computational complexity*, 20(2):367–388, 2011.

[dW06] Ronald de Wolf. Lower bounds on matrix rigidity via a quantum argument. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *Automata, Languages and Programming*, pages 62–71, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

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

[GKST02] Oded Goldreich, Howard Karloff, Leonard J Schulman, and Luca Trevisan. Lower bounds for linear locally decodable codes and private information retrieval. In *Proceedings 17th IEEE Annual Conference on Computational Complexity*, pages 175–183. IEEE, 2002.

[GT18] Oded Goldreich and Avishay Tal. Matrix rigidity of random toeplitz matrices. *computational complexity*, 27(2):305–350, 2018.

[GVL12] Gene H Golub and Charles F Van Loan. *Matrix computations*, volume 3. JHU press, 2012.

[Juk01] Stasys Jukna. *Orthogonality and Rank Arguments*, pages 191–204. Springer Berlin Heidelberg, Berlin, Heidelberg, 2001.

[Juk11] Stasys Jukna. *Extremal Combinatorics: With Applications in Computer Science*, pages 197–212. Springer Berlin Heidelberg, Berlin, Heidelberg, 2011.

[KR98] B. S. Kashin and A. A. Razborov. Improved lower bounds on the rigidity of hadamard matrices. *Mathematical Notes*, 63(4):471–475, 1998.

[Lok95] Satyanarayana V. Lokam. Spectral methods for matrix rigidity with applications to size-depth tradeoffs and communication complexity. In *Proceedings of IEEE 36th Annual Foundations of Computer Science*, pages 6–15, Oct 1995.

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

[Lok06] Satyanarayana V. Lokam. Quadratic lower bounds on matrix rigidity. In Jin-Yi Cai, S. Barry Cooper, and Angsheng Li, editors, *Theory and Applications of Models of Computation*, pages 295–307, Berlin, Heidelberg, 2006. Springer Berlin Heidelberg.

[Lok09] Satyanarayana V. Lokam. Complexity lower bounds using linear algebra. *Found. Trends Theor. Comput. Sci.*, 4(1–2):1–155, January 2009.

[Mid05] Gatis Midrijanis. Three lines proof of the lower bound for the matrix rigidity, 2005.

[Mor96] Patrick Morandi. *Transcendental Extensions*, pages 173–224. Springer New York, New York, NY, 1996.

[MS77] Florence Jessie MacWilliams and Neil James Alexander Sloane. *The theory of error-correcting codes*, volume 16. Elsevier, 1977.

[PP06] R. Paturi and P. Pudlák. Circuit lower bounds and linear codes. *Journal of Mathematical Sciences*, 134(5):2425–2434, 2006.

[Pud94] P. Pudlák. Communication in bounded depth circuits. *Combinatorica*, 14(2):203–216, 1994.

[RR20] Sivaramakrishnan Natarajan Ramamoorthy and Cyrus Rashtchian. Equivalence of Systematic Linear Data Structures and Matrix Rigidity. In Thomas Vidick, editor, *11th Innovations in Theoretical Computer Science Conference (ITCS 2020)*, volume 151 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 35:1–35:20, Dagstuhl, Germany, 2020. Schloss Dagstuhl–Leibniz-Zentrum fuer Informatik.

[SSS97] M.A. Shokrollahi, D.A. Spielman, and V. Stemann. A remark on matrix rigidity. *Information Processing Letters*, 64(6):283 – 285, 1997.

[SY11] S. Saraf and S. Yekhanin. Noisy interpolation of sparse polynomials, and applications. In *2011 IEEE 26th Annual Conference on Computational Complexity*, pages 86–92, 2011.

[TVZ82] M. A. Tsfasman, S. G. Vladutx, and Th. Zink. Modular curves, shimura curves, and goppa codes, better than varshamov-gilbert bound. *Mathematische Nachrichten*, 109(1):21–28, 1982.

[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.