# Probabilistic Upper Bounds

There are a few different ways to prove the upper bound $t(d, N) = O(d^2 \log(N/d))$. We only present a couple of them here.

## 1 Concatenating a random code with the identity code

Let $q, N, d, n$ be integers such that $q > d$, and $q^n \geq N$. Let $C_{\text{in}}$ be the identity code $\text{ID}_q$. Let $C_{\text{out}}$ be a random code of length $n$, size $N$, alphabet $[q]$ constructed as follows. We randomly select each codeword $\mathbf{c}$ of $C_{\text{out}}$ by picking uniformly a random symbol from $[q]$ for each position of $\mathbf{c}$ independently. Let $\mathbf{M} = C_{\text{out}} \circ C_{\text{in}}$. We bound the probability that $\mathbf{M}$ is *not* $d$-disjunct. Let $j_0, \dots, j_d$ be a fixed set of $d+1$ columns of $\mathbf{M}$. Then,

$$\text{Prob}[\text{codeword } \mathbf{M}^{j_0} \text{ is covered by } \mathbf{M}^{j_1}, \dots, \mathbf{M}^{j_d}] \leq (d/q)^n.$$

Thus, by the union bound

$$\text{Prob}[\mathbf{M} \text{ is not } d\text{-disjunct}] \leq (d+1)\binom{N}{d+1}(d/q)^n.$$

The following proposition is thus proved.

**Proposition 1.1.** *Let $q, N, d, n$ be integers such that $q > d$ and $q^n \geq N$. If*

$$(d+1)\binom{N}{d+1}(d/q)^n < 1$$

*then there exists a $d$-disjunct matrix with $qn$ rows and $N$ columns.*

**Corollary 1.2.** *We have the following upper bound for $t(d, N)$:*

$$t(d, N) = O(d^2 \log(N/d)).$$

*Proof.* Pick $q = 2d$, and $n = 2(d+2)\log(N/(d+1)) + 1$. Without loss of generality we can assume $e \leq N/(d+1)$ and $(d+1) \leq (N/(d+1))^2$. Observe that

$$(d+1)\binom{N}{d+1} \leq (d+1)(Ne/(d+1))^{d+1} \leq (N/(d+1))^{2+2(d+1)} = 2^{2(d+2)\log(N/(d+1))} < 2^n = (q/d)^n.$$

□

**Open Problem 1.3.** The upperbound $O(d^2 \log(N/d))$ is only slightly larger than the best known lower bound $\Omega(d^2 2 \log N / \log d)$ that we have proved in the previous lecture. Closing this gap is the major open question in group testing theory.

## 2 Connection to the $k$-RESTRICTION problem

### 2.1 The $k$-RESTRICTION problem

Alon-Moshkovitz-Safra [1] pointed out that constructing $d$-disjunct matrices is a special case of the $k$-RESTRICTION problem, and showed that their derandomization result on the *$k$-restriction* problem can be used to construct $d$-disjunct matrices in time $N^{O(d)}$; unfortunately, this is not polynomial time for super-constant $d$. We present their construction here.

We will only need the binary version of the $k$-RESTRICTION problem, which is defined as follows. The input to the problem is an alphabet $\Sigma = \{0, 1\}$, a length $N$, and a set of $m$ possible *demands* $f_i : \Sigma^k \to \{0, 1\}$, $1 \leq i \leq m$. For every demand $f_i$, there is at least one vector $\mathbf{a} = (a_1, \dots, a_k) \in \Sigma^k$ such that $f_i(\mathbf{a}) = 1$ (in words, $f_i$ "demands" vector $\mathbf{a}$). In English, every demand $f_i$ "demands" a non-empty subset of vectors from $\Sigma^k$. A feasible solution to the problem is a subset $T \subseteq \Sigma^N$ such that: for any choice of $k$ indices $1 \leq j_1 < j_2 < \cdots < j_k \leq N$, and any demand $f_i$, there is some vector $\mathbf{v} = (v_1, \dots, v_N) \in T$ such that the projection of $\mathbf{v}$ onto those $k$ indices satisfies demand $f_i$; namely, $f_i(v_{j_1}, v_{j_2}, \dots, v_{j_k}) = 1$. The objective is to find a feasible solution $T$ with minimum cardinality. Alon, Moshkovitz, and Safra gave a couple of algorithmic solutions to the $k$-restriction problems. In order to describe their results, we need a few more concepts.

Given a distribution $\mathcal{D} : \Sigma^N \to [0, 1]$, the *density* of a $k$-RESTRICTION problem with respect to $\mathcal{D}$ is

$$\epsilon := \min_{\substack{1 \leq j_1 < \cdots < j_k \leq N \\ 1 \leq i \leq m}} \left\{ \operatorname*{Prob}_{\mathbf{v} \leftarrow \mathcal{D}} \left[ f_i(v_{j_1}, v_{j_2}, \dots, v_{j_k}) = 1 \right] \right\}$$

In words, for any $k$ positions $j_1, \dots, j_k$, and any demand $f_i$, if we pick a vector $\mathbf{v}$ from $\Sigma^N$ at random according to $\mathcal{D}$ then the projection of $\mathbf{v}$ onto those $k$ positions satisfies $f_i$ with probability at least $\epsilon$.

**Exercise 1.** Show that every $k$-RESTRICTION problem has density at least $1/2^k$ with respect to the uniform distribution.

**Exercise 2.** Show that, every $k$-RESTRICTION problem with density $\epsilon$ with respect to some probability distribution $\mathcal{D}$ has a solution of size at most $\left\lceil \frac{k \ln N + lnm}{\epsilon} \right\rceil$.

### 2.2 Disjunct matrices as a $k$-RESTRICTION problem

We can view the problem of constructing a $d$-disjunct matrix as the problem of finding a solution to a special case of the $k$-RESTRICTION problem. We will set $k = d + 1$ and $m = d + 1$. There are $m = d + 1$ *demands* $f_i : \Sigma^{d+1} \to \{0, 1\}$ defined as follows:

$$f_i(\mathbf{a}) = f_i(a_1, \dots, a_{d+1}) = a_i \wedge \overline{a_1} \wedge \cdots \overline{a_{i-1}} \wedge \overline{a_{i+1}} \wedge \cdots \overline{a_{d+1}}.$$

It is not hard to see that any solution $T$ to the above instance of $k$-RESTRICTION form the rows of a $|T| \times N$ matrix $\mathbf{M}$ which is $d$-disjunct.

From the above two exercises, we know there exist $t \times N$ $d$-disjunct matrices with $O(2^d d \ln(Nd))$ rows. This number is too large compared to the random code concatenation technique we saw in the previous section. It turns out, however, that there are better distributions than the uniform distribution for the group testing problem. The density of the $k$-RESTRICTION problem can be much larger than $1/2^{d+1}$.

### 2.3 Derandomization

The *restriction* $\mathcal{D}_{j_1,\dots,j_k}$ of a distribution $\mathcal{D}$ on $\Sigma^N$ to indices $j_1, \dots, j_k$ is defined by

$$\mathcal{D}_{j_1,\dots,j_k}(\mathbf{a}) := \underset{\mathbf{v}\leftarrow\mathcal{D}}{\text{Prob}}\left[v_{j_1} = a_1 \wedge \cdots \wedge v_{j_k} = a_k\right]$$

Two distribution $\mathcal{P}, \mathcal{Q}$ on $\Sigma^N$ are said to be *$k$-wise $\epsilon$-close* if, for any $1 \leq j_1 < j_2 < \cdots < j_k \leq n$, $\|\mathcal{P}_{j_1,\dots,j_k} - Q_{j_1,\dots,j_k}\|_1 < \epsilon$. The *support* of a distribution on $\Sigma^N$ is the number of members of $\Sigma^N$ which have positive probabilities. Finally, a distribution $\mathcal{D}$ on $\Sigma^N$ is said to be *$k$-wise efficiently approximable* if the support of a distribution $\mathcal{P}$ which is $k$-wise $\epsilon$-close to $\mathcal{D}$ can be enumerated in time poly$(N, \frac{1}{\epsilon}, 2^k)$.

One of the main results from [1] is the following.

**Theorem 2.1** (Alon-Moshkovitz-Safra [1]). *Fix some $k$-wise efficiently approximable distribution $\mathcal{D}$ on $\Sigma^N$. For any $k$-restriction problem with density $\epsilon$ with respect to $\mathcal{D}$, there is an algorithm that, given an instance of the problem and a constant parameter $0 < \delta < 1$, obtains a solution $T$ of size at most $\lceil \frac{k \ln N + \ln m}{(1-\delta)\epsilon} \rceil$ in time poly$(m, N^k, 2^k, \frac{1}{\epsilon}, \frac{1}{\delta})$.*

A distribution $\mathcal{D}$ on $\Sigma^N = \{0,1\}^N$ is called a *product distribution* if all coordinates are independent Bernoulli variables. (Coordinate $i$ is 1 with probability $p_i$ for some fixed $p_i$.)

**Theorem 2.2** (Even-Goldreich-Luby-Nisan-Veličković [2]). *Any product distribution on $\Sigma^N$ is $k$-wise efficiently approximable.*

**Theorem 2.3.** *Given positive integers $n \geq d+1$, let $\epsilon_0 = \left(\frac{1}{d+1}\right)\left(\frac{d}{d+1}\right)^d$. Then, a $t \times N$ $d)$-disjunct matrix can be constructed so that $t = \left\lceil \frac{2(d+1)\ln N + 2\ln(d+1)}{\epsilon_0} \right\rceil$ in time poly$\left(N^d, \frac{1}{\epsilon}\right)$. In particular, since $\left(\frac{d}{d+1}\right)^d$ is a descreasing function in $d$ (whose limit is $1/e$), we know $\epsilon_0 \geq \frac{1}{e(d+1)}$, which means $t \leq 2e(d+1)^2 \ln N + 2e(d+1)\ln(d+1)$.*

*Proof.* From the previous section we know how to cast the problem of constructing a $d$-disjunct matrix as a special case of the $k$-RESTRICTION problem. Next, let $\mathcal{D}$ be the product distribution on $\{0,1\}^N$ defined by setting each coordinate to be 1 with probability $p$ to be determined. Then, $\mathcal{D}$ is $k$-wise efficiently approximable by Theorem 2.2. Fix $1 \leq j_1 < j_2 < \cdots < j_k \leq n$ and any demand $f_i$. Choose any vector $\mathbf{v}$ from $\{0,1\}^N$ according to $\mathcal{D}$. The projection of $\mathbf{v}$ onto coordinates $j_1, \dots, j_k$ "satisfies" $f_i$ with probability

$$\epsilon(p) = p\,(1-p)^d.$$

The density $\epsilon(p)$ is maximized at $p_0 = 1/(d+1)$. Set $\epsilon_0 = \epsilon(p_0)$, $\delta = 1/2$, and apply Theorem 2.1. □

## References

[1] N. ALON, D. MOSHKOVITZ, AND S. SAFRA, *Algorithmic construction of sets for $k$-restrictions*, ACM Trans. Algorithms, 2 (2006), pp. 153–177.

[2] G. EVEN, O. GOLDREICH, M. LUBY, N. NISAN, AND B. VELIČKOVIĆ, *Efficient approximation of product distributions*, Random Structures Algorithms, 13 (1998), pp. 1–16.