Introduction to Coding Theory **IST-UL, 2025/2026**

Simple applications of the probabilistic method

Lecturer: João Ribeiro

Introduction

The probabilistic method is an extremely useful tool in mathematics for proving the existence of combinatorial objects (e.g., graphs, codes, functions...) with desired properties. This is done by taking a probabilistic viewpoint on the construction of such objects, even when the problem at hand seems to have nothing to do with probability at all. Surprisingly, injecting probability into a problem sometimes makes it much easier to tackle. However, this also has the downside of typically making the existence proofs *non-constructive*, in the sense that the proof does not tell us an efficient way of constructing an object that satisfies that property (nevertheless, sometimes you can get around this – the interplay between randomness and computation is a fascinating area of mathematics and theoretical computer science).

These notes are meant to guide you through a couple of simple applications of the probabilistic method. We will see some more examples in the lectures and in the problem sets. If you would like to explore more, the following are great references:

- The Probabilistic Method, by Noga Alon and Joel Spencer.
- These lecture notes by Jiří Matoušek and Jan Vondrák. I highly recommend you read at least Section 1 on probability theory to refresh your memory and (re)learn some basic inequalities.

1 Ramsey graphs

The probabilistic method was first used in the 1940s by Szele and Erdős. In this section we will focus on the result that Erdős proved using this method.

A graph $G = (V, E)$ is *K-Ramsey* if it does not contain any clique or independent set of size K. Clearly, if G is K-Ramsey then it is also K'-Ramsey for all $K' > K$, so Ramsey graphs with smaller K are harder to find. Therefore, a first natural challenge is to understand, for a given number of vertices N, what is the smallest K for which a K-Ramsey graph exists.

Erdős gave an elegant answer to this problem [Erd47].

Theorem 1 *For any integer $N \geq 2$, there exists a K-Ramsey graph on N vertices with $K \leq 2(1 + \log N)$.*

Proof: Construct a graph G on N vertices as follows: for every pair of vertices (u, v), add an edge between u and v independently with probability $1/2$. We will show that this graph will be K-Ramsey for $K \leq 2 \log N$ with positive probability, which implies the existence of such a graph.

Fix any set S of K vertices. The vertices in S form a clique if and only if all $\binom{K}{2}$ edges between vertices in S exist, and they form an independent set if and only if none of these edges exist. Then, the probability that the vertices in S form a clique or an independent set is exactly $2 \cdot 2^{-\binom{K}{2}}$. Since there are $\binom{N}{K}$ possible choices for S, the probability that there exists such a set S of size K that forms a clique or an independent set is at most[1]

$$2 \cdot \binom{N}{K} \cdot 2^{-\binom{K}{2}}.$$

Therefore, we are guaranteed a positive probability of G being K-Ramsey so long as

$$2 \cdot \binom{N}{K} \cdot 2^{-\binom{K}{2}} < 1.$$

Since $\binom{N}{K} \leq N^K$, it suffices to have

$$N^K < 2^{\frac{K(K-1)}{2}-1}.$$

It is easy to verify that this holds whenever

$$N \leq 2^{K/2-1}.$$

Rearranging, we get that we can take $K = 2(1 + \log N)$. ∎

This is a remarkably short proof. On the other hand, it does not really tell us how we could "efficiently"[2] construct a K-Ramsey graph with small K. It turns out that Ramsey graphs are very closely related to important objects in theoretical computer science and cryptography, called *two-source randomness dispersers* and *randomness extractors* (in fact, two-source dispersers are equivalent to *bipartite* Ramsey graphs), so efficiently constructing objects such as Ramsey graphs is something theoretical computer scientists care about.

Coming up with efficient constructions of (families of) K-Ramsey graphs (and stronger objects) with K closely matching the existential result of Erdős is a much greater challenge that has been the focus of plenty of research over the past few decades. This culminated in breakthrough results of Chattopadhyay and Zuckerman [CZ19] and Cohen [Coh21].

2 Balancing vectors

Here is another application of the probabilistic method, where we leverage linearity of expectation.

[1] Here we are using the simple (but very useful!) *union bound*: $\Pr[A \vee B] \leq \Pr[A] + \Pr[B]$.

[2] There is more than one interesting notion of efficiency here. A construction is "explicit" if we can produce the adjacency matrix of the graph in time polynomial in N. It is "strongly explicit" if given any two vertices u and v we can decide if they are adjacent in time polynomial in $\log N$. The latter notion is the more useful one in computer science, and is what I mean by "efficient" here.

Theorem 2 *Let $v_1, \ldots, v_t \in \mathbb{R}^n$ satisfying[3] $\|v_i\|_2 = 1$ for all $i = 1, \ldots, t$. Then, there exist $\alpha_1, \ldots, \alpha_t \in \{-1, 1\}$ such that*

$$\|\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_t v_t\|_2 \leq \sqrt{t}.$$

Proof: Consider choosing each $\alpha_1, \alpha_2, \ldots, \alpha_t$ independently and uniformly at random from $\{-1, 1\}$.

First, note that

$$\begin{aligned}\|\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_t v_t\|_2^2 &= \langle \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_t v_t, \alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_t v_t \rangle \\ &= \sum_{i,j} \alpha_i \alpha_j \langle v_i, v_j \rangle.\end{aligned}$$

Now we look at the expected value of this expression (recalling that the α_i's are independent and uniformly distributed over $\{-1, 1\}$). By linearity of expectation, we have

$$\mathrm{E}\left[\sum_{i,j} \alpha_i \alpha_j \langle v_i, v_j \rangle\right] = \sum_{i,j} \mathrm{E}[\alpha_i \alpha_j] \langle v_i, v_j \rangle.$$

When $i = j$ we have $\alpha_i \alpha_j = \alpha_i^2 = 1$ always, so $\mathrm{E}[\alpha_i \alpha_j] = 1$ in this case. When $i \neq j$, we have $\mathrm{E}[\alpha_i \alpha_j] = \mathrm{E}[\alpha_i] \cdot \mathrm{E}[\alpha_j] = 0$ by independence of α_i and α_j. Therefore, the expected value of the squared Euclidean norm above is

$$\sum_i \langle v_i, v_i \rangle = \sum_i \|v_i\|_2^2 = \sum_i 1 = t.$$

Since knowing that $\mathrm{E}[X] = t$ implies that $\Pr[X \leq t] > 0$ for any random variable X, we conclude that there exists a choice of $\alpha_1, \ldots, \alpha_t$ such that

$$\|\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_t v_t\|_2^2 \leq t.$$

Taking square roots on both sides concludes the proof. ∎

References

[Coh21] Gil Cohen. Two-source dispersers for polylogarithmic entropy and improved Ramsey graphs. *SIAM Journal on Computing*, 50(3):STOC16–30–STOC16–67, 2021.

[CZ19] Eshan Chattopadhyay and David Zuckerman. Explicit two-source extractors and resilient functions. *Annals of Mathematics*, 189(3):653–705, 2019.

[Erd47] Paul Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4):292–294, 1947.

[3] Here, $\|v\|_2 = \left(\sum_{j=1}^n v_j^2\right)^{1/2}$ is the standard Euclidean norm.