K_x to be the set of all k-colorings such that $x + S$ is colorful. It is easy to see that each K_x is closed. The finite statement proved above asserts that any finite collection of the K_x has a non-empty intersection. It follows, by compactness, that all K_x, $x \in \mathbf{R}$, have a non-empty intersection. Any element in this intersection is a coloring desired by the theorem. □

Exercise

1.5.1 Show that there exists a positive constant c such that the following holds. For every sufficiently large n, there is a graph on n points which does not contain the following two objects: a triangle and an independent set of size $c\sqrt{n}\log n$. (An independent set is a set of vertices, no two of which are connected by an edge.)

1.6 Janson's inequality

Let $t_1, \dots, t_n$ be jointly independent boolean random variables. In Corollary 1.9 we established a large deviation inequality for the polynomial $t_1 + \cdots + t_n$. In many applications, it is also of interest to obtain large deviation inequalities for more general polynomials $P(t_1, \dots, t_n)$ of the boolean variables $t_1, \dots, t_n$. One particularly important case is that of a *boolean polynomial*

$$X := \sum_{A \in \mathcal{A}} \prod_{j \in A} t_j,$$

where $\mathcal{A}$ is some collection of non-empty subsets of $[1, n]$. Observe that boolean polynomials are automatically positive and monotone increasing, and hence any two boolean polynomials are positively correlated via the FKG inequality (Theorem 1.19). More generally, if X and Y are boolean polynomials, then $f(X)$ and $f(Y)$ will be positively correlated whenever f is a monotone increasing or decreasing function. In particular, we see that

$$\mathbf{E}\big(e^{-s(X+Y)}\big) \geq \mathbf{E}\big(e^{-sX}\big)\mathbf{E}\big(e^{-sY}\big) \tag{1.34}$$

for any real number s. Using this fact, the exponential moment method, and some additional convexity arguments, Janson [190] derived a powerful bound for the lower tail probability $\mathbf{P}(X \leq \mathbf{E}(X) - T)$:

Theorem 1.28 (Janson's inequality) *Let $t_1, \dots, t_n, \mathcal{A}, X$ be as above. Then for any $0 \leq T \leq \mathbf{E}(X)$ we have the lower tail estimate*

$$\mathbf{P}(X \leq E(X) - T) \leq \exp\left(-\frac{T^2}{2\Delta}\right)$$

where

$$\Delta = \sum_{A,B\in\mathcal{A}:A\cap B\neq\emptyset} \mathbf{E}\left(\prod_{j\in A\cup B} t_j\right).$$

In particular, we have

$$\mathbf{P}(X=0) \le \exp\left(-\frac{\mathbf{E}(X)^2}{2\Delta}\right).$$

Remark 1.29 Informally, Janson's inequality asserts that if $\Delta = O(\mathbf{E}(X)^2)$, then $X = \Omega(\mathbf{E}(X))$ with large probability. In the case where $\mathcal{A}$ is just the collection of singletons $\{1\}, \ldots, \{n\}$, then $X = t_1 + \cdots + t_n$, $\Delta = \mathbf{E}(X)$, and the above claim is then essentially (the lower half of) Corollary 1.9.

The quantity Δ is somewhat inconvenient to work with directly. Using the independence of the t_j, one can rewrite it as

$$\Delta = \sum_{A\in\mathcal{A}} \mathbf{E}\left(\prod_{j\in A} t_j\right) \sum_{B\in\mathcal{A}:A\cap B\neq\emptyset} \mathbf{E}\left(\prod_{j\in B\setminus A} t_j\right).$$

Since $\mathbf{E}(X) = \sum_{A\in\mathcal{A}} \mathbf{E}(\prod_{j\in A} t_j)$, we thus have

$$\Delta \le \mathbf{E}(X) \sup_{A\in\mathcal{A}} \sum_{B\in\mathcal{A}:A\cap B\neq\emptyset} \mathbf{E}\left(\prod_{j\in B\setminus A} t_j\right). \tag{1.35}$$

We record a particular consequence of this estimate concerning quadratic boolean polynomials that we shall use shortly.

Corollary 1.30 *Let $t_1, \ldots, t_n$ be as above, and let $X = \sum_{1\le i\le j\le n: i\sim j} t_i t_j$, where $i \sim j$ is some symmetric relation on* $[1, n]$. *Then we have*

$$\mathbf{P}(X=0) \le \exp\left(-\frac{\mathbf{E}(X)}{2 + 4\sup_i \sum_{j:i\sim j} \mathbf{E}(t_j)}\right).$$

Proof We take $\mathcal{A} := \{\{i, j\} : i \sim j\}$. For any $A \in \mathcal{A}$, it is easy to verify that

$$\sum_{B\in\mathcal{A}:A\cap B\neq\emptyset} \mathbf{E}\left(\prod_{j\in B\setminus A} t_j\right) \le 1 + 2\sup_i \sum_{j:i\sim j} \mathbf{E}(t_j)$$

and so the claim follows from (1.35) and Theorem 1.28. □

Before presenting the proof of Theorem 1.28, let us give an application. This application again concerns complementary bases of primes, but this time of order 2 rather than 1. The following result (which should be compared with Theorems 1.16 and 1.22) in the case $k = 2$ was recently proved by Vu [376].

Theorem 1.31 *For any $k \geq 2$, P has a complementary base $B \in \mathbf{Z}^+$ of order k with $|B \cap [1, n]| = O(\log n)$ for all large n.*

Proof It suffices to establish the claim when $k = 2$. To construct B we shall again use the probabilistic method. More precisely, we let $B \subset \mathbf{Z}^+$ be a random set with the events $n \in B$ being independent with probability

$$\mathbf{P}(n \in B) = \min\left(\frac{c}{n}, 1\right)$$

for all $n \in \mathbf{Z}^+$, where c is a positive constant to be determined. As before, we will not discuss the measure-theoretic issues associated with requiring infinitely-many independent random variables, as they can be dealt with by a suitable finitization of this argument. Let t_n be the boolean random variable $t_n := \mathbf{I}(n \in B)$. By Corollary 1.10 we have

$$\mathbf{P}(|B \cap [1, m]| \leq 10c \log m) = 1 - O\left(\frac{1}{m^2}\right)$$

for all large m, and hence by the Borel–Cantelli lemma (Lemma 1.2) we have with probability 1 that

$$|B \cap [1, m]| = O_c(\log m) \text{ for all sufficiently large } m > 1. \tag{1.36}$$

Now for each $n \in \mathbf{Z}^+$, consider the counting function

$$\begin{aligned} r_{P+B+B}(n) &= |\{(p, i, j) \in P \times B \times B : n = p + i + j\}| \\ &= \sum_{p<n} \sum_{i+j=n-p} t_i t_j. \end{aligned}$$

This is of course a random variable for each n. In view of (1.36), it will suffice to show that with probability 1, we have $r_{P+B+B}(n) \neq 0$ for all but finitely many n. From the Borel–Cantelli lemma, it thus suffices to show that

$$\mathbf{P}(r_{P+B+B}(n) = 0) = O\left(\frac{1}{n^2}\right)$$

for all large n, if c is chosen large enough.

Fix n to be large. It will be convenient to work with a reduced version of $r_{P+B+B}(n)$, namely the boolean polynomial

$$Y_n := \sum_{i>j\geq n^{2/3}: i+j \in n-P} t_i t_j.$$

Clearly we have $Y_n \leq r_{P+B+B}(n)$, and so it suffices to show that

$$\mathbf{P}(Y_n = 0) = O\left(\frac{1}{n^2}\right).$$

We now apply Corollary 1.30 (using the relation $i \sim j$ if $i \neq j$ and $i + j \in n - P$) to give

$$\mathbf{P}(Y_n = 0) \leq \exp\left(-\frac{\mathbf{E}(Y_n)}{2 + 4\sup_{i \geq n^{2/3}} \sum_{j \geq n^{2/3}: i+j \in n-P} \mathbf{E}(t_j)}\right).$$

By construction of the t_j, and Proposition 1.54 from the Appendix, we have for any $i \geq n^{2/3}$

$$\begin{aligned}\sum_{j \geq n^{2/3}: i+j \in n-P} \mathbf{E}(t_j) &= \sum_{p \leq n-i-n^{2/3}} \min\left(\frac{c}{n-i-p}, 1\right) \\ &= O(c).\end{aligned}$$

On the other hand, from linearity of expectation (1.3) and independence, we have

$$\begin{aligned}\mathbf{E}(Y_n) &= \sum_{i > j \geq n^{2/3}: i+j \in n-P} \mathbf{E}(t_i t_j) \\ &= \sum_{i > j \geq n^{2/3}: i+j \in n-P} \frac{c^2}{ij} \\ &= c^2 \sum_{p \leq n-2n^{2/3}} \sum_{i > j \geq n^{2/3}: i+j = n-p} \frac{1}{ij} \\ &= c^2 \sum_{p \leq n-2n^{2/3}} \Omega\left(\frac{\log(n-p)}{n-p}\right) \\ &= \Omega(c^2 \log n),\end{aligned}$$

where in the last line we again used Proposition 1.54 from the Appendix. Putting all of these estimates together we obtain

$$\mathbf{P}(Y_n = 0) \leq \exp(-\Omega(c \log n))$$

and the claim follows by choosing c to be suitably large. □

Now we are going to prove Theorem 1.28.

Proof of Theorem 1.28 We shall use the exponential moment method. By a limiting argument we may assume that $\mathbf{P}(t_j = 0), \mathbf{P}(t_j = 1) > 0$ for all j. We introduce the moment generating function $F(t) := \mathbf{E}(e^{-tX})$ for any $t > 0$. By (1.16) we have

$$\mathbf{P}(X \leq \mathbf{E}(X) - T) \leq \frac{F(t)}{e^{-t(\mathbf{E}(X)-T)}}.$$

Taking logarithms, we see that we only need to establish the inequality

$$\log F(t) + t(\mathbf{E}(X) - T) \le -\frac{T^2}{2\Delta}$$

for some $t > 0$. Unlike the situation in Theorem 1.8, the summands in X are not necessarily independent, so we cannot factorize $F(t) = \mathbf{E}(e^{-tX})$ easily. Janson found a beautiful argument to get around this difficulty. Since $F(0) = 1$, we see from the fundamental theorem of calculus that

$$\log F(t) = \int_0^t \frac{F'(s)}{F(s)}\, ds.$$

Direct calculation shows that

$$\begin{aligned} F'(s) &= -\mathbf{E}(Xe^{-sX}) \\ &= -\sum_{A\in\mathcal{A}} \mathbf{E}\left(e^{-sX}\prod_{j\in A} t_j\right) \\ &= -\sum_{A\in\mathcal{A}} \mathbf{E}(e^{-sX}|E_A)\mathbf{P}(E_A), \end{aligned}$$

where E_A is the event that $t_j = 1$ for all $j \in A$. Thus it suffices to show that

$$\sum_{A\in\mathcal{A}} \mathbf{P}(E_A)\int_0^t \frac{\mathbf{E}(e^{-sX}|E_A)}{F(s)}\, ds - t(\mathbf{E}(X) - T) \ge \frac{T^2}{2\Delta}$$

for some $t > 0$.

We now exploit the fact that some of the factors of e^{-sX} are independent of E_A. For each $A \in \mathcal{A}$, we split X as $Y_A + Z_A$, which are the boolean polynomials

$$Y_A := \sum_{B\in\mathcal{A}:A\cap B\neq\emptyset}\ \prod_{j\in B} t_j; \quad Z_A = \sum_{B\in\mathcal{A}:A\cap B=\emptyset}\ \prod_{j\in B} t_j.$$

By (1.34) (conditioning on the variables in E_A), we conclude

$$\mathbf{E}(e^{-sX}|E_A) \ge \mathbf{E}(e^{-sY_A}|E_A)\mathbf{E}(e^{-sZ_A}|E_A).$$

On the other hand, Z_A is independent from E_A and is bounded from above by X; thus

$$\mathbf{E}(e^{-sZ_A}|E_A) = \mathbf{E}(e^{-sZ_A}) \ge \mathbf{E}(e^{-sX}) = F(s).$$

Combining all these estimates, we have reduced to showing that

$$\sum_{A\in\mathcal{A}} \mathbf{P}(E_A)\int_0^t \mathbf{E}(e^{-sY_A}|E_A)\, ds - t(\mathbf{E}(X) - T) \ge \frac{T^2}{2\Delta}$$

for some $t > 0$.

Next, we exploit the convexity of the function $x \mapsto e^{-sx}$ via Jensen's inequality (Exercise 1.2.4), concluding that

$$\mathbf{E}(e^{-sY_A}|E_A) \geq e^{-s\mathbf{E}(Y_A|E_A)}.$$

From linearity of expectation we have $\sum_{A\in\mathcal{A}} \mathbf{P}(E_A) = \mathbf{E}(X)$, and so another application of Jensen's inequality gives

$$\sum_{A\in\mathcal{A}} \mathbf{P}(E_A)e^{-s\mathbf{E}(Y_A|E_A)} \geq \mathbf{E}(X)e^{-s\sum_{A\in\mathcal{A}} \frac{\mathbf{P}(E_A)}{\mathbf{E}(X)}\mathbf{E}(Y_A|E_A)}.$$

On the other hand, from the definition of conditional probability we have

$$\sum_{A\in\mathcal{A}} \mathbf{P}(E_A)\mathbf{E}(Y_A|E_A) = \sum_{A\in\mathcal{A}} \sum_{B\in\mathcal{A}:A\cap B\neq\emptyset} \mathbf{E}\left(\mathbf{I}(E_A)\prod_{j\in B} t_j\right) = \Delta.$$

We thus have

$$\sum_{A\in\mathcal{A}} \mathbf{P}(E_A)\int_0^t \mathbf{E}(e^{-sY_A}|E_A)\,ds - t(\mathbf{E}(X) - T) \tag{1.37}$$

$$\geq \mathbf{E}(X)\int_0^t e^{-s\Delta/\mathbf{E}(X)}\,ds - t(\mathbf{E}(X) - T)$$

$$= \frac{\mathbf{E}(X)^2}{\Delta}\left(1 - e^{-t\Delta/\mathbf{E}(X)}\right) - t(\mathbf{E}(X) - T). \tag{1.38}$$

If we set $t := T/\Delta$, then $t\Delta/\mathbf{E}(X) = T/\mathbf{E}(X) \leq 1$, and we have

$$\begin{aligned} 1 - e^{-t\Delta/\mathbf{E}(X)} &= 1 - e^{-T/\mathbf{E}(X)} \\ &\geq T/\mathbf{E}(X) - T^2/2\mathbf{E}(X)^2 \end{aligned}$$

and hence

$$\begin{aligned} \sum_{A\in\mathcal{A}} \mathbf{P}(E_A)\int_0^t \mathbf{E}(e^{-sY_A}|E_A)\,ds - t(\mathbf{E}(X) - T) \\ \geq \frac{T\mathbf{E}(X)}{\Delta} - \frac{T^2}{2\Delta} - \frac{T}{\Delta}(\mathbf{E}(X) - T) \\ = \frac{T^2}{2\Delta} \end{aligned}$$

as desired. □

Remark 1.32 Choosing $t = T/\Delta$ might be convenient, but may not be optimal. One can have a slightly better bound by optimizing the right hand side of (1.38) over t.

Remark 1.33 The proof of Janson's inequality is not symmetric. In other words, it cannot be extended to give a bound for the upper tail probability $\mathbf{P}(X \geq \mu + T)$. This probability will be addressed in the next section.