Exercises

1.6.1 By refining the argument, show that the complementary base B constructed in the proof of Theorem 1.31 has (with high probability) the property that $r_{P+B+B}(n) = \Omega(\log n)$ for all sufficiently large n.

1.6.2 Define a random graph $G(n, p)$ on the vertex set $[1, n]$ as follows. For each pair i, j $(1 \leq i < j \leq n)$ draw an edge between i and j with probability p, independently.

(a) Prove that if $p = o(n^{-1})$, then with probability $1 - o(1)$, $G(n, p)$ does not contain a triangle.

(b) Assume that $p = n^{-1+\epsilon}$ for some small positive constant ϵ. Bound the probability that G does not contain a triangle.

1.6.3 Prove that for any $k \geq 2$ there is a basis B of order k with with $|B \cap [1, n]| = O(n^{1/2} \log^{1/k} n)$ for all large n.

1.7 Concentration of polynomials

In previous sections, we often considered a polynomial $Y = Y(t_1, \ldots, t_n)$ of n independent random variables $t_1, \ldots, t_n$, and wished to control the tail distribution of Y. For instance Chernoff's inequality shows that the polynomial $t_1 + \cdots + t_n$ is concentrated around its mean, while Janson's inequality shows that the values of certain polynomials (especially those of low degree) could very rarely be significantly less than the mean.

In this section, we present some further results of this type, that assert that certain polynomials with small degrees are strongly concentrated. These results can be seen as generalizing Chernoff's bound, and also provide (in certain cases) the missing half (upper tail bound) of Janson's inequality.

To motivate the results, let us first give a classical result which works for any function Y (not just a polynomial) provided that the Lipschitz constant of Y is small.

Lemma 1.34 (Lipschitz concentration inequality) *Let $Y : \{0, 1\}^n \to \mathbf{R}$ be a function such that $|Y(t) - Y(t')| \leq K$ whenever $t, t' \in \{0, 1\}^n$ differ in only one coordinate. Then if $t_1, \ldots, t_n$ are independent boolean variables, we have*

$$\mathbf{P}(|Y(t_1, \ldots, t_n) - \mathbf{E}(Y(t_1, \ldots, t_n))| \geq \lambda K \sqrt{n}) \leq 2e^{-\lambda^2/2}$$

for all $\lambda > 0$.

Remark 1.35 This inequality asserts that if each t_i can only influence the random variable $Y(t_1, \ldots, t_n)$ by at most $O(K)$, then $Y(t_1, \ldots, t_n)$ itself is concentrated in an interval of length $O(K\sqrt{n})$ around its mean. It should be compared with Hoeffding's inequality, which deals with the case $Y(t_1, \ldots, t_n) := t_1 + \cdots + t_n$, and also with Corollary 1.30.

Proof By dividing Y by K we may renormalize $K = 1$. Introduce the partially-conditioned random variables $Y_0, Y_1(t_1), \ldots, Y_n(t_1, \ldots, t_n) = Y(t_1, \ldots, t_n)$ by $Y_j(t_1, \ldots, t_j) := \mathbf{E}(Y|t_1, \ldots, t_j)$; thus Y_j is the conditional expectation of Y with the first j boolean variables t_j fixed. In particular $Y_0 = \mathbf{E}(Y)$ and $Y_n = Y(t_1, \ldots, t_n)$. We can thus write

$$Y(t_1, \ldots, t_n) - \mathbf{E}(Y(t_1, \ldots, t_n)) = X_1 + \cdots + X_n$$

where $X_j := Y_j - Y_{j-1}$. One then easily verifies (using the Lipschitz property) that $|X_j| \leq 1$ and $X_1, \ldots, X_n$ form a martingale difference sequence in the sense of Exercise 1.3.6. The claim then follows from Azuma's inequality (1.24). □

The above lemma is very useful when one has uniform Lipschitz control on Y, for instance if $Y = Y(t_1, \ldots, t_n)$ is a polynomial for which the partial derivatives $\frac{\partial Y}{\partial t_i}$ are small for *all* $t_1, \ldots, t_n$ in the unit cube. However in many applications (especially to thin bases), these partial derivatives will only be small on the *average*. Fortunately there are analogs of the above lemma which apply in this case, though they also require some average control on higher derivatives of Y. To state the results we need some notation. Let $Y = Y(t_1, \ldots, t_n)$ be a polynomial of n real variables. We say that Y is *totally positive* if all of its coefficients are non-negative, and furthermore that Y is *regular* if all the coefficients are between zero and one. We also say that Y is *simplified* if all of its monomials are square-free (i.e. do not contain any factor of t_i^2), and *homogeneous* if all the monomials have the same degree. Thus for instance a boolean polynomial is automatically regular and simplified, though not necessarily homogeneous. Given any multi-index $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbf{Z}_+^n$, we define the partial derivative $\partial^\alpha Y$ as

$$\partial^\alpha Y := \left(\frac{\partial}{\partial t_1}\right)^{\alpha_1} \cdots \left(\frac{\partial}{\partial t_n}\right)^{\alpha_n} Y(t_1, \ldots, t_n),$$

and denote the order of α as $|\alpha| := \alpha_1 + \cdots + \alpha_n$. For any order $d \geq 0$, we denote $\mathbf{E}_d(Y) := \max_{\alpha:|\alpha|=d} \mathbf{E}(\partial^\alpha Y)$; thus for instance $\mathbf{E}_0(Y) = \mathbf{E}(Y)$, and $\mathbf{E}_d(Y) = 0$ if d exceeds the degree of Y. These quantities are vaguely reminiscent of Sobolev norms for the random variable Y. We also define $\mathbf{E}_{\geq d}(Y) := \max_{d' \geq d} \mathbf{E}_{d'}(Y)$.

The following result is due to Kim and Vu [203].

Theorem 1.36 *Let $k \geq 1$, and let $Y = Y(t_1, \ldots, t_n)$ be a totally positive polynomial of n independent boolean variables $t_1, \ldots, t_n$. Then there exists a constant $C_k > 0$ depending only on k such that*

$$\mathbf{P}\big(|Y - \mathbf{E}(Y)| \geq C_k \lambda^{k-1/2} \sqrt{\mathbf{E}_{\geq 0}(Y)\mathbf{E}_{\geq 1}(Y)}\big) = O_k\big(e^{-\lambda/4 + (k-1)\log n}\big)$$

for all $\lambda > 0$.

Informally Theorem 1.36 asserts that when the derivatives of Y are smaller on average than Y itself, and the degree of Y is small, then Y is concentrated around