EXERCISES (DISCREPANCY AND SDPS)

(1) The **Beck Fiala Theorem** states that if $S$ is a set system with degree at most $t$, i.e. each element is contained in at most $t$ sets, then $\text{disc}(S) \leq 2t - 1$.

(a) Consider the set system with elements as $n$ arbitrary points in a 2d plane, and the sets as all the axis-parallel rectangles. Use the Beck Fiala theorem to show that this system has discrepancy $O(\log n)$.

(b) (Tricky) Prove the Beck Fiala Theorem.

   Hint: The number of sets with cardinality greater than $t$ is strictly less then $n$. Now, apply iterative rounding to the linear program with constraints $Ax = 0$ and $x \in [-1, 1]$, where $A$ is the incidence matrix of $S$.

(2) **Entropy Method**: Recall that the partial coloring lemma.

**Lemma 1.** Let $S$ be a set system on an $n$-point set $V$, and let a number $\Delta_S > 0$ be given for each set $S \in S$. Suppose $\Delta_S$ satisfy the condition

$$\sum_{S \in S} g\left(\frac{\Delta_S}{\sqrt{|S|}}\right) \leq \frac{n}{5}$$

where

$$g(\lambda) = \begin{cases} Ke^{-\lambda^2/9} & \text{if } \lambda > 0.1 \\ K \ln(\lambda^{-1}) & \text{if } \lambda \leq 0.1 \end{cases}$$

and $K$ is some absolute constant. Then there is a partial coloring $X$ that assigns $\pm 1$ to at least $n/2$ variables (and 0 to the rest), and satisfies $|X(S)| \leq \Delta_S$ for each $S \in S$.

(a) Use this to show that a degree $t$ set system on $n$ elements has discrepancy $O(\sqrt{t \log n})$.

   Hint: Show that there is a partial coloring with discrepancy $O(\sqrt{t})$. Use that the number of sets of cardinality more than $k$ is at most $nt/k$.

(b) (Optional) Show that set system formed by axis-parallel rectangles has discrepancy $O(\log^{2.5} n)$.

(3) **Lower bound for Discrepancy:**

(a) Let $A$ be the incidence matrix of a set system $S$ on $n$ elements and $m$ sets. Show that

$$\text{disc}(S) \leq \sqrt{\frac{n}{m}} \lambda_{\text{min}}$$

where $\lambda_{\text{min}}$ is the minimum eigenvalue of the $n \times n$ matrix $A^t A$.

   Hint: For any matrix $A$, $\text{disc}(S) = \min_{x \in \{-1, +1\}^n} |Ax|_\infty$ and $\|Ax\|_2^2 \geq \lambda_{\text{min}} \|x\|_2^2$.

(b) Use the idea in the hint above to show the following: Let $S$ be a set system with $n$ sets and $n$ elements, such that each element lies in exactly $t$ sets, and every two elements $i, j$ are contained in exactly $s$ sets. Show the $\text{disc}(S) \geq \sqrt{t - s}$. 


(4) **Lower bound for Hereditary Discrepancy:** We will show that for any matrix $S$

\[
\text{herdisc}(S) \geq \max_k \max_{B} \frac{1}{2} \det(B)^{1/k}
\]

where $B$ ranges over all $k \times k$ submatrices of $S$.

Let $A$ be an $n \times n$ matrix. Recall that for any matrix $A$ and $x \in [-1,1]^n$, there is a $\tilde{x} \in \{-1,+1\}^n$ such that

\[
\|A(x - \tilde{x})\|_\infty \leq 2 \text{herdisc}(A).
\]

Consider the convex body

\[
B(t) = \{y \in \mathbb{R}^n : \|Ay\|_\infty \leq t\}.
\]

Note that $B(t)$ is symmetric about the origin. Let $t_0 = 2 \text{herdisc}(A)$.

(a) Show that for every $x \in [-1,1]^n$, the body $B(t_0)$ centered at $x$ contains at least one of the points $\{-1,+1\}^n$.

(b) Show that the volume of $B(t_0)$ is at least $2^n$.

Hint: Use $B(t_0)$ to tile $\mathbb{R}^n$.

(c) Show that $\text{herdisc}(A) \geq \frac{1}{2} \det(A)^{1/n}$.

Hint: $B(1)$ is the inverse image of the cube $[-1,1]^n$ under the linear map $y \rightarrow Ay$, and hence $\text{vol}(B(1)) = 2^n / \det(A)$.

(d) Show how the above implies (1).

(5) **SDPs and Covariance matrices:** Let $v_1, \ldots, v_m$ be vectors in $\mathbb{R}^n$. Let $g$ be standard $n$-dimensional Gaussian vector in $\mathbb{R}^n$ (i.e. each coordinate chosen independently from the distribution $N(0,1)$). Let $x_i = \langle g, v_i \rangle$. We saw that each $x_i$ is distributed as $N(0,\|v_i\|_2^2)$.

Show that for every $1 \leq i, j \leq m$ it also holds that

\[
E_g[x_i x_j] = \langle v_i, v_j \rangle.
\]

This observation lies at the heart of many recent SDP based algorithms.

(6) **Anti-concentration via energy:** Consider the standard $\pm 1$ random walk on a line that starts at the origin, and terminates on reaching one of the points $-k$ or $k$. Show that the probability that the walk has not terminated by time $ck^2$ is at most $1/c$.

Hint: Let $X_t$ denote the position of the walk at time $t$. Consider the random variable $Y_t$ defined as $Y_t = X_t^2$ until the walk terminates, and $Y_t = Y_{t-1} + 1$ thereafter. Apply a Markov type inequality to $Y_t$. 