Discrepancy and Some Applications in TCS

Romain Bourneuf
LaBRI & ENS de Lyon

June 10, 2026

Discrepancy

Discrepancy studies the irregularities of distributions.

Discrepancy

Discrepancy studies the irregularities of distributions.

Today: Given a distribution μ on a set V that simultaneously satisfies some properties $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_p$ on expectation, how close can a single element from V be from satisfying simultaneously all $\mathcal{P}_i$?

Discrepancy

Discrepancy studies the irregularities of distributions.

Today: Given a distribution μ on a set V that simultaneously satisfies some properties $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_p$ on expectation, how close can a single element from V be from satisfying simultaneously all $\mathcal{P}_i$?
What is the best element of V for $\mathcal{P}_1, \ldots, \mathcal{P}_p$ and how good is it?

Discrepancy

Discrepancy studies the irregularities of distributions.

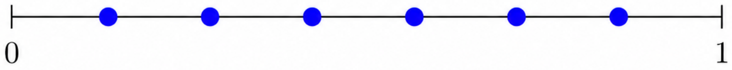
Today: Given a distribution μ on a set V that simultaneously satisfies some properties $\mathcal{P}_1, \mathcal{P}_2, \ldots, \mathcal{P}_p$ on expectation, how close can a single element from V be from satisfying simultaneously all $\mathcal{P}_i$?
What is the best element of V for $\mathcal{P}_1, \ldots, \mathcal{P}_p$ and how good is it?

What is the most uniform set of n points in $[0, 1]$?

Discrepancy studies the irregularities of distributions.

Today: Given a distribution μ on a set V that simultaneously satisfies some properties $\mathcal{P}_1, \mathcal{P}_2, \dots, \mathcal{P}_p$ on expectation, how close can a single element from V be from satisfying simultaneously all $\mathcal{P}_i$?
What is the best element of V for $\mathcal{P}_1, \dots, \mathcal{P}_p$ and how good is it?

What is the most uniform set of n points in $[0, 1]$?

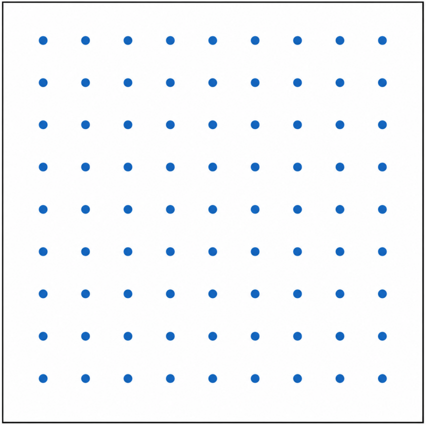


In the plane

What is the most uniform set of n points in $[0, 1]^2$?

In the plane

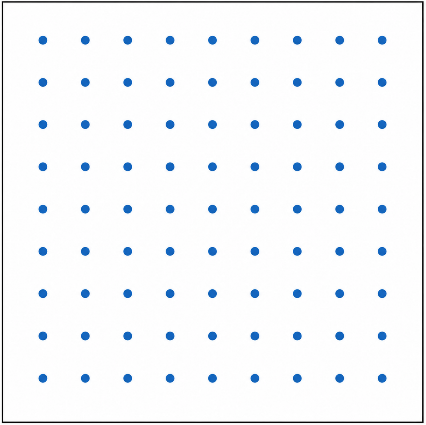
What is the most uniform set of n points in $[0, 1]^2$?



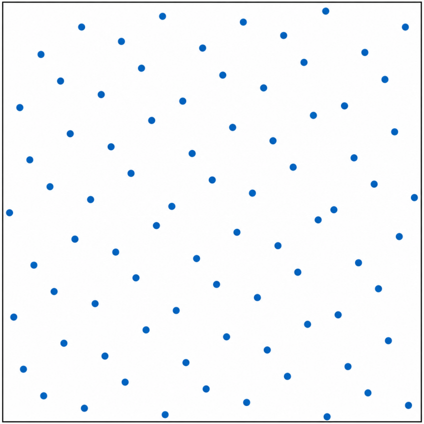
Grid

In the plane

What is the most uniform set of n points in $[0, 1]^2$?



Grid



Van der Corput

Rectangle discrepancy in the plane

How to measure the discrepancy?

How to measure the discrepancy?

Definition

$P \subseteq [0,1]^2$. For every rectangle $R \subseteq [0,1]^2$,

$$\mathrm{disc}(P, R) = \left| \frac{|P \cap R|}{|P|} - \mathrm{Area}(R) \right| .$$

How to measure the discrepancy?

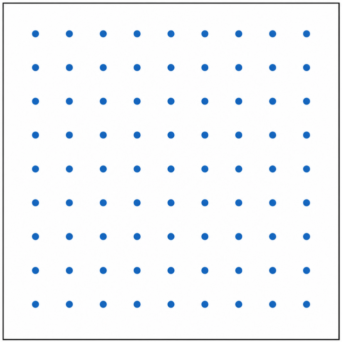
Definition

$P \subseteq [0,1]^2$. For every rectangle $R \subseteq [0,1]^2$,

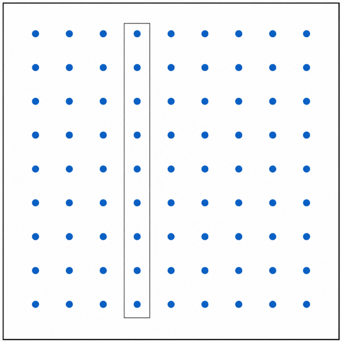
$$\mathrm{disc}(P,R) = \left| \frac{|P \cap R|}{|P|} - \mathrm{Area}(R) \right|.$$

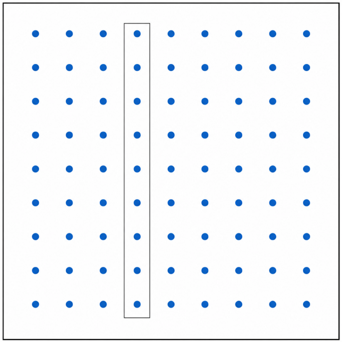
$$\mathrm{disc}(P) = \max_R \mathrm{disc}(P,R).$$

Rectangle discrepancy in the plane



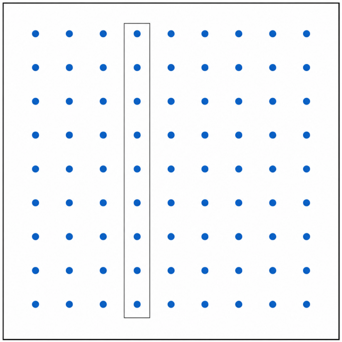
Rectangle discrepancy in the plane

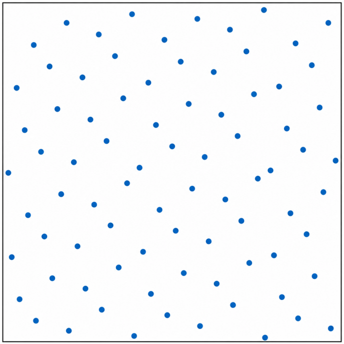




$$\text{disc}(P) = \Omega(\sqrt{n})$$

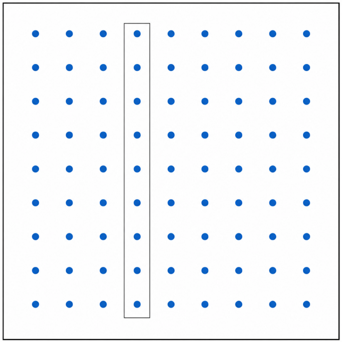
Rectangle discrepancy in the plane



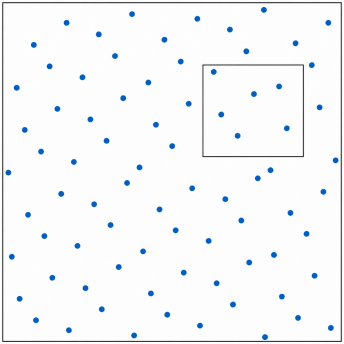


$\text{disc}(P) = \Omega(\sqrt{n})$

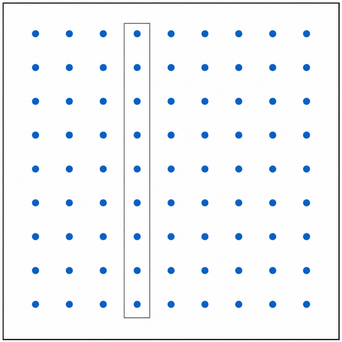
Rectangle discrepancy in the plane



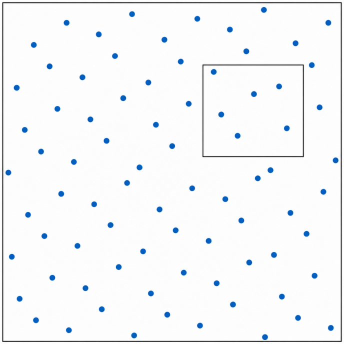
$$\text{disc}(P) = \Omega(\sqrt{n})$$



Rectangle discrepancy in the plane

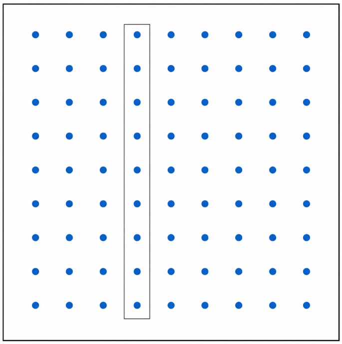


$\text{disc}(P) = \Omega(\sqrt{n})$

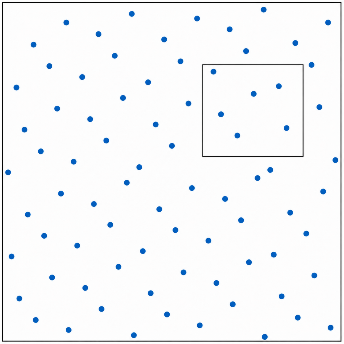


$\text{disc}(P) = O(\log n)$

Rectangle discrepancy in the plane



$\text{disc}(P) = \Omega(\sqrt{n})$



$\text{disc}(P) = O(\log n)$

Theorem [Schmidt '72]

Every set of n points in $[0,1]^2$ has discrepancy $\Omega(\log n)$.

Uniform matrices

What is the most uniform $n \times n$ $\{-1, 1\}$-matrix?

Uniform matrices

What is the most uniform $n \times n$ $\{-1, 1\}$-matrix?

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

Uniform matrices

What is the most uniform $n \times n$ $\{-1, 1\}$-matrix?

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$H_{k+1} = \begin{bmatrix} H_k & H_k \\ H_k & -H_k \end{bmatrix}$$

Uniform matrices

What is the most uniform $n \times n$ $\{-1, 1\}$-matrix?

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad H_{k+1} = \begin{bmatrix} H_k & H_k \\ H_k & -H_k \end{bmatrix}$$

Proposition

H_n is a $2^n \times 2^n$ matrix with the following properties:

Uniform matrices

What is the most uniform $n \times n$ $\{-1, 1\}$-matrix?

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad H_{k+1} = \begin{bmatrix} H_k & H_k \\ H_k & -H_k \end{bmatrix}$$

Proposition

H_n is a $2^n \times 2^n$ matrix with the following properties:

- Every '1'-monochromatic rectangle in H_n contains at most 2^n entries, and every '-1'-monochromatic rectangle in H_n contains at most 2^{n-1} entries.

What is the most uniform $n \times n$ $\{-1, 1\}$-matrix?

$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad H_{k+1} = \begin{bmatrix} H_k & H_k \\ H_k & -H_k \end{bmatrix}$$

Proposition

H_n is a $2^n \times 2^n$ matrix with the following properties:

- Every '1'-monochromatic rectangle in H_n contains at most 2^n entries, and every '-1'-monochromatic rectangle in H_n contains at most 2^{n-1} entries.
- $H_n(H_n)^\intercal = 2^n I$.

Uniform matrices

What is the most uniform $n \times n$ $\{-1, 1\}$-matrix?

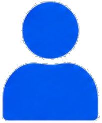
$$H_1 = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \qquad H_{k+1} = \begin{bmatrix} H_k & H_k \\ H_k & -H_k \end{bmatrix}$$

Proposition

H_n is a $2^n \times 2^n$ matrix with the following properties:

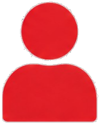
- Every '1'-monochromatic rectangle in H_n contains at most 2^n entries, and every '-1'-monochromatic rectangle in H_n contains at most 2^{n-1} entries.
- $H_n(H_n)^\intercal = 2^n I$.
- After replacing the '1' entries by '0' entries and the '-1' entries by '1' entries, we get the matrix of the inner products in $\{0, 1\}^n$ modulo 2 (ordered lexicographically).

Communication complexity of the inner product



Alice

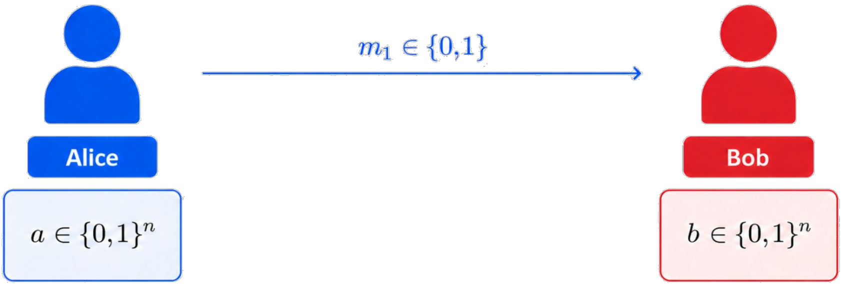
$a \in \{0,1\}^n$



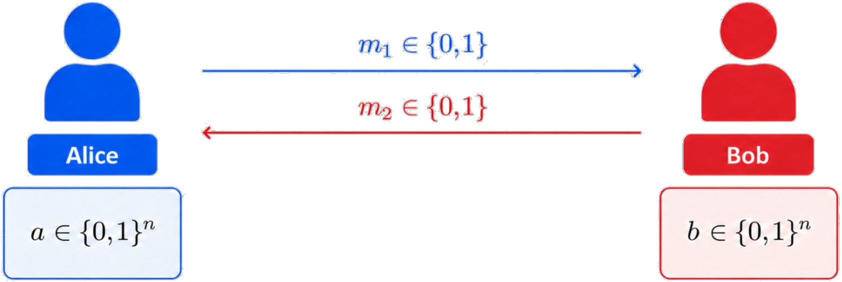
Bob

$b \in \{0,1\}^n$

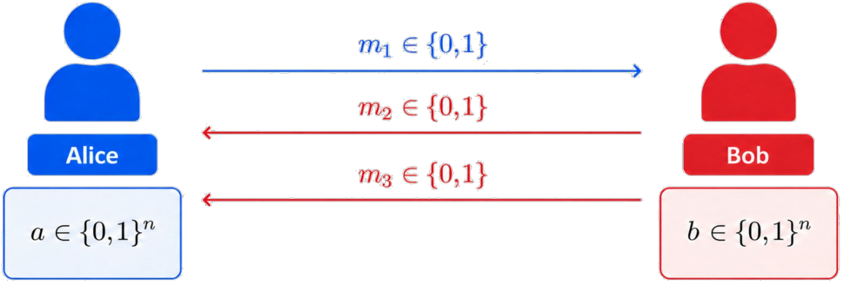
Communication complexity of the inner product



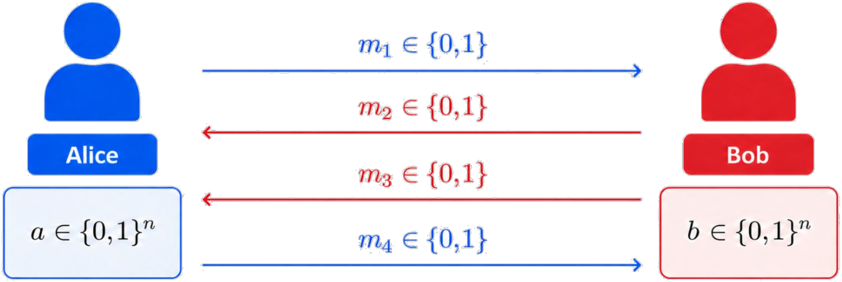
Communication complexity of the inner product



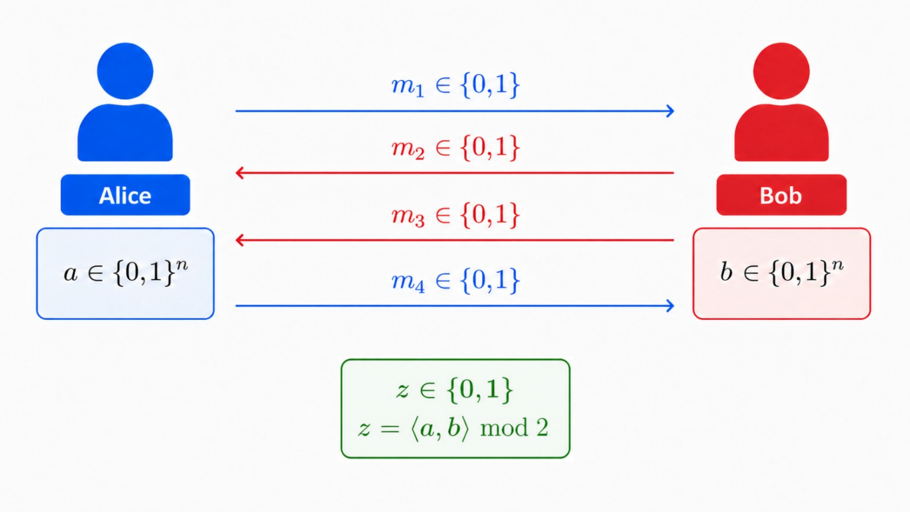
Communication complexity of the inner product



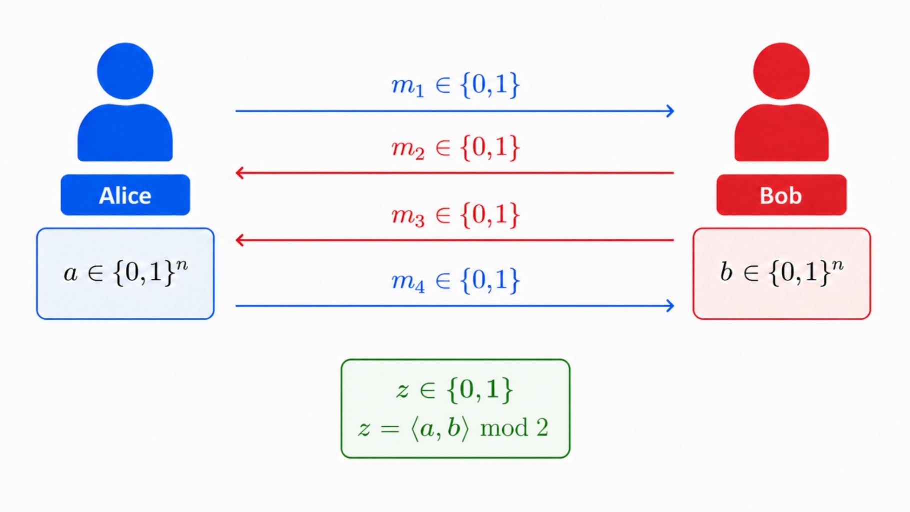
Communication complexity of the inner product



Communication complexity of the inner product



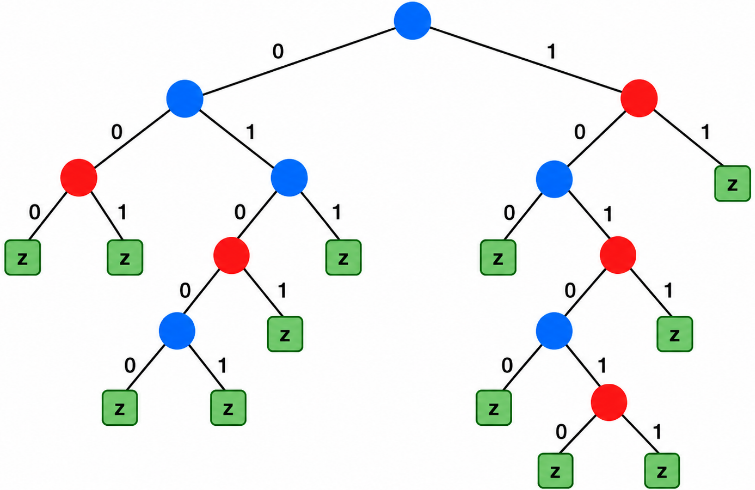
Communication complexity of the inner product



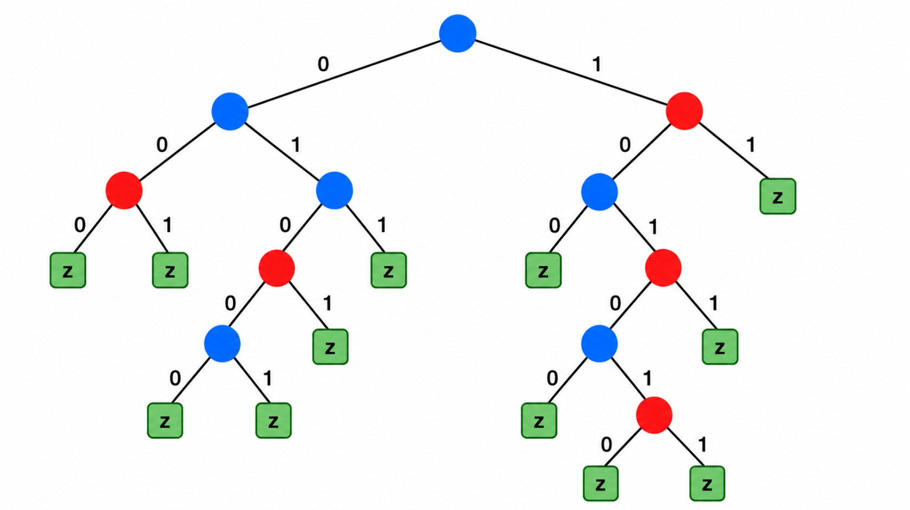
Cost: Number of bits sent in the worst case

A strategy for Alice and Bob

A strategy for Alice and Bob



A strategy for Alice and Bob



Key observation: The leaves give a partition of H_n into monochromatic rectangles.

Bound on the communication complexity

Let $N = 2^n$.
The matrix of the inner products in $\{0,1\}^n$ modulo 2 contains $\frac{N(N+1)}{2}$ entries 0 and $\frac{N(N-1)}{2}$ entries 1.

Bound on the communication complexity

Let $N = 2^n$.
The matrix of the inner products in $\{0,1\}^n$ modulo 2 contains $\frac{N(N+1)}{2}$ entries 0 and $\frac{N(N-1)}{2}$ entries 1.
$\rightarrow$ We need at least $N-1$ monochromatic rectangles of 1 and $\frac{N+1}{2}$ monochromatic rectangles of 0 to cover all entries.

Bound on the communication complexity

Let $N = 2^n$.
The matrix of the inner products in $\{0,1\}^n$ modulo 2 contains $\frac{N(N+1)}{2}$ entries 0 and $\frac{N(N-1)}{2}$ entries 1.
$\rightarrow$ We need at least $N-1$ monochromatic rectangles of 1 and $\frac{N+1}{2}$ monochromatic rectangles of 0 to cover all entries.
$\rightarrow$ The number of leaves of the tree is at least $\frac{3N-1}{2} > N$, so the tree has depth $> \log_2(N) = n$.

Bound on the communication complexity

Let $N = 2^n$.
The matrix of the inner products in $\{0,1\}^n$ modulo 2 contains $\frac{N(N+1)}{2}$ entries 0 and $\frac{N(N-1)}{2}$ entries 1.
$\rightarrow$ We need at least $N-1$ monochromatic rectangles of 1 and $\frac{N+1}{2}$ monochromatic rectangles of 0 to cover all entries.
$\rightarrow$ The number of leaves of the tree is at least $\frac{3N-1}{2} > N$, so the tree has depth $> \log_2(N) = n$.

Theorem

The communication complexity of the inner product modulo 2 is $n+1$.

Bound on the communication complexity

Let $N = 2^n$.
The matrix of the inner products in $\{0,1\}^n$ modulo 2 contains $\frac{N(N+1)}{2}$ entries 0 and $\frac{N(N-1)}{2}$ entries 1.
$\rightarrow$ We need at least $N-1$ monochromatic rectangles of 1 and $\frac{N+1}{2}$ monochromatic rectangles of 0 to cover all entries.
$\rightarrow$ The number of leaves of the tree is at least $\frac{3N-1}{2} > N$, so the tree has depth $> \log_2(N) = n$.

Theorem

The communication complexity of the inner product modulo 2 is $n+1$.

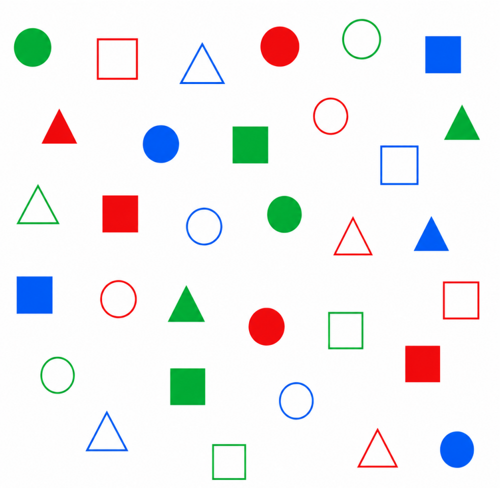
Remark: A similar result holds if we just require the output bit to be correct with probability $1/2 + \varepsilon$.

Towards combinatorial discrepancy

How to split a set of objects into 2 as evenly as possible?

Towards combinatorial discrepancy

How to split a set of objects into 2 as evenly as possible?



Combinatorial discrepancy

A *set system* $\mathcal{S}$ is a family $S_1, \ldots, S_m$ of subsets of a *universe* V of size n.

Combinatorial discrepancy

A *set system* $\mathcal{S}$ is a family $S_1, \ldots, S_m$ of subsets of a *universe* V of size n.

A *coloring* of $\mathcal{S}$ is $\chi : V \to \{-1, 1\}$.

Combinatorial discrepancy

A *set system* $\mathcal{S}$ is a family $S_1, \ldots, S_m$ of subsets of a *universe* V of size n.

A *coloring* of $\mathcal{S}$ is $\chi : V \to \{-1, 1\}$.

For $S \in \mathcal{S}$ and $\chi : V \to \{-1, 1\}$,

$$\chi(S) = \sum_{v \in S} \chi(v)$$

Combinatorial discrepancy

A *set system* $\mathcal{S}$ is a family $S_1, \ldots, S_m$ of subsets of a *universe* V of size n.

A *coloring* of $\mathcal{S}$ is $\chi : V \to \{-1, 1\}$.

For $S \in \mathcal{S}$ and $\chi : V \to \{-1, 1\}$,

$$\chi(S) = \sum_{v \in S} \chi(v) = |\chi^{-1}(1) \cap S| - |\chi^{-1}(-1) \cap S|.$$

Combinatorial discrepancy

A *set system* $\mathcal{S}$ is a family $S_1, \ldots, S_m$ of subsets of a *universe* V of size n.

A *coloring* of $\mathcal{S}$ is $\chi : V \to \{-1, 1\}$.

For $S \in \mathcal{S}$ and $\chi : V \to \{-1, 1\}$,

$$\chi(S) = \sum_{v \in S} \chi(v) = |\chi^{-1}(1) \cap S| - |\chi^{-1}(-1) \cap S|.$$

$$\mathrm{disc}(\mathcal{S}, \chi) = \max_{S \in \mathcal{S}} |\chi(S)|.$$

Combinatorial discrepancy

A *set system* $\mathcal{S}$ is a family $S_1, \ldots, S_m$ of subsets of a *universe* V of size n.

A *coloring* of $\mathcal{S}$ is $\chi : V \to \{-1, 1\}$.

For $S \in \mathcal{S}$ and $\chi : V \to \{-1, 1\}$,

$$\chi(S) = \sum_{v \in S} \chi(v) = |\chi^{-1}(1) \cap S| - |\chi^{-1}(-1) \cap S|.$$

$$\mathrm{disc}(\mathcal{S}, \chi) = \max_{S \in \mathcal{S}} |\chi(S)|.$$

$$\mathrm{disc}(\mathcal{S}) = \min_{\chi} \mathrm{disc}(\mathcal{S}, \chi).$$

Combinatorial discrepancy

A *set system* $\mathcal{S}$ is a family $S_1, \ldots, S_m$ of subsets of a *universe* V of size n.

A *coloring* of $\mathcal{S}$ is $\chi : V \to \{-1, 1\}$.

For $S \in \mathcal{S}$ and $\chi : V \to \{-1, 1\}$,

$$\chi(S) = \sum_{v \in S} \chi(v) = |\chi^{-1}(1) \cap S| - |\chi^{-1}(-1) \cap S|.$$

$$\text{disc}(\mathcal{S}, \chi) = \max_{S \in \mathcal{S}} |\chi(S)|.$$

$$\text{disc}(\mathcal{S}) = \min_{\chi} \text{disc}(\mathcal{S}, \chi).$$

Given $A \in \mathbb{R}^{m \times n}$, $\text{disc}(A) = \min_{x \in \{-1,1\}^n} \|Ax\|_\infty$.

Upper bounds

Proposition

$\forall \mathcal{S}, \mathrm{disc}(\mathcal{S}) \leq n/2$, and $\mathrm{disc}(2^{[n]}) = n/2$.

Upper bounds

Proposition

$\forall \mathcal{S}, \mathrm{disc}(\mathcal{S}) \leq n/2$, and $\mathrm{disc}(2^{[n]}) = n/2$.

Proposition

$\forall \mathcal{S}, \mathrm{disc}(\mathcal{S}) = O(\sqrt{n \log m})$.

Upper bounds

Proposition

$\forall \mathcal{S}, \mathrm{disc}(\mathcal{S}) \leq n/2$, and $\mathrm{disc}(2^{[n]}) = n/2$.

Proposition

$\forall \mathcal{S}, \mathrm{disc}(\mathcal{S}) = O(\sqrt{n \log m})$.

Proof: Take $\chi : V \to \{-1, 1\}$ uniformly at random.

Upper bounds

Proposition

$\forall \mathcal{S}, \text{disc}(\mathcal{S}) \leq n/2$, and $\text{disc}(2^{[n]}) = n/2$.

Proposition

$\forall \mathcal{S}, \text{disc}(\mathcal{S}) = O(\sqrt{n \log m})$.

Proof: Take $\chi : V \to \{-1, 1\}$ uniformly at random.
By Chernoff, for every $S \in \mathcal{S}$, $\mathbb{P}[|\chi(S)| \geq O(\sqrt{n \log m})] \leq \frac{1}{2m}$.

Upper bounds

Proposition

$\forall \mathcal{S}, \text{disc}(\mathcal{S}) \leq n/2$, and $\text{disc}(2^{[n]}) = n/2$.

Proposition

$\forall \mathcal{S}, \text{disc}(\mathcal{S}) = O(\sqrt{n \log m})$.

Proof: Take $\chi : V \to \{-1, 1\}$ uniformly at random.
By Chernoff, for every $S \in \mathcal{S}$, $\mathbb{P}[|\chi(S)| \geq O(\sqrt{n \log m})] \leq \frac{1}{2m}$.
By union bound, $\mathbb{P}[\text{disc}(S, \chi) \geq O(\sqrt{n \log m})] \leq \frac{1}{2}$.

Upper bounds

Proposition

$\forall \mathcal{S}, \text{disc}(\mathcal{S}) \leq n/2$, and $\text{disc}(2^{[n]}) = n/2$.

Proposition

$\forall \mathcal{S}, \text{disc}(\mathcal{S}) = O(\sqrt{n \log m})$.

Proof: Take $\chi : V \rightarrow \{-1, 1\}$ uniformly at random.
By Chernoff, for every $S \in \mathcal{S}$, $\mathbb{P}[|\chi(S)| \geq O(\sqrt{n \log m})] \leq \frac{1}{2m}$.
By union bound, $\mathbb{P}[\text{disc}(S, \chi) \geq O(\sqrt{n \log m})] \leq \frac{1}{2}$.
$\rightarrow \mathbb{P}[\text{disc}(S, \chi) \leq O(\sqrt{n \log m})] \geq \frac{1}{2} > 0$.

Upper bounds

Proposition

$\forall \mathcal{S}, \mathrm{disc}(\mathcal{S}) \leq n/2$, and $\mathrm{disc}(2^{[n]}) = n/2$.

Proposition

$\forall \mathcal{S}, \mathrm{disc}(\mathcal{S}) = O(\sqrt{n \log m})$.

Proof: Take $\chi : V \to \{-1, 1\}$ uniformly at random.
By Chernoff, for every $S \in \mathcal{S}$, $\mathbb{P}[|\chi(S)| \geq O(\sqrt{n \log m})] \leq \frac{1}{2m}$.
By union bound, $\mathbb{P}[\mathrm{disc}(S, \chi) \geq O(\sqrt{n \log m})] \leq \frac{1}{2}$.
$\to \mathbb{P}[\mathrm{disc}(S, \chi) \leq O(\sqrt{n \log m})] \geq \frac{1}{2} > 0$.

Theorem [Spencer '85]

If $m \geq n$ then $\mathrm{disc}(\mathcal{S}) \leq O(\sqrt{n \log \frac{2m}{n}})$.
If $m = n$ then $\mathrm{disc}(\mathcal{S}) \leq 6\sqrt{n}$.

Lower bounds

Proposition

There exists a set system $\mathcal{S}$ with $m = n$ sets and $\text{disc}(\mathcal{S}) = \Omega(\sqrt{n})$.

Lower bounds

Proposition

There exists a set system $\mathcal{S}$ with $m = n$ sets and $\text{disc}(\mathcal{S}) = \Omega(\sqrt{n})$.

Proof: Let $\mathcal{S}$ be the set system with "incidence matrix" H_n, set $N = 2^n$.

Lower bounds

Proposition

There exists a set system $\mathcal{S}$ with $m = n$ sets and $\operatorname{disc}(\mathcal{S}) = \Omega(\sqrt{n})$.

Proof: Let $\mathcal{S}$ be the set system with "incidence matrix" H_n, set $N = 2^n$. Every element is in $N/2$ sets, every two elements share $N/4$ sets.

Lower bounds

Proposition

There exists a set system $\mathcal{S}$ with $m = n$ sets and $\text{disc}(\mathcal{S}) = \Omega(\sqrt{n})$.

Proof: Let $\mathcal{S}$ be the set system with "incidence matrix" H_n, set $N = 2^n$.
Every element is in $N/2$ sets, every two elements share $N/4$ sets.
Let $\chi : V \to \{-1, 1\}$.

Lower bounds

Proposition

There exists a set system $\mathcal{S}$ with $m = n$ sets and $\text{disc}(\mathcal{S}) = \Omega(\sqrt{n})$.

Proof: Let $\mathcal{S}$ be the set system with "incidence matrix" H_n, set $N = 2^n$.
Every element is in $N/2$ sets, every two elements share $N/4$ sets.
Let $\chi : V \to \{-1, 1\}$.

$$\sum_{S \in \mathcal{S}} \text{disc}(S, \chi)^2$$

Lower bounds

Proposition

There exists a set system $\mathcal{S}$ with $m = n$ sets and $\text{disc}(\mathcal{S}) = \Omega(\sqrt{n})$.

Proof: Let $\mathcal{S}$ be the set system with "incidence matrix" H_n, set $N = 2^n$.
Every element is in $N/2$ sets, every two elements share $N/4$ sets.
Let $\chi : V \to \{-1, 1\}$.

$$\sum_{S \in \mathcal{S}} \text{disc}(S, \chi)^2 = \sum_{S} \left(\sum_{v \in S} \chi(v) \right)^2$$

Lower bounds

Proposition

There exists a set system $\mathcal{S}$ with $m = n$ sets and $\text{disc}(\mathcal{S}) = \Omega(\sqrt{n})$.

Proof: Let $\mathcal{S}$ be the set system with "incidence matrix" H_n, set $N = 2^n$.
Every element is in $N/2$ sets, every two elements share $N/4$ sets.
Let $\chi : V \to \{-1, 1\}$.

$$\sum_{S \in \mathcal{S}} \text{disc}(S, \chi)^2 = \sum_{S} \left(\sum_{v \in S} \chi(v) \right)^2$$

$$= \sum_{u,v} \chi(u)\chi(v) \cdot |\{S \in \mathcal{S} : u, v \in S\}|$$

Lower bounds

Proposition

There exists a set system $\mathcal{S}$ with $m = n$ sets and $\text{disc}(\mathcal{S}) = \Omega(\sqrt{n})$.

Proof: Let $\mathcal{S}$ be the set system with "incidence matrix" H_n, set $N = 2^n$. Every element is in $N/2$ sets, every two elements share $N/4$ sets.
Let $\chi : V \to \{-1, 1\}$.

$$\sum_{S \in \mathcal{S}} \text{disc}(S, \chi)^2 = \sum_{S} \left(\sum_{v \in S} \chi(v) \right)^2$$

$$= \sum_{u,v} \chi(u)\chi(v) \cdot |\{S \in \mathcal{S} : u, v \in S\}|$$

$$= N/4 \left(\sum_{v \in V} \chi(v)^2 + \left(\sum_{v \in V} \chi(v) \right)^2 \right)$$

Lower bounds

Proposition

There exists a set system $\mathcal{S}$ with $m = n$ sets and $\mathrm{disc}(\mathcal{S}) = \Omega(\sqrt{n})$.

Proof: Let $\mathcal{S}$ be the set system with "incidence matrix" H_n, set $N = 2^n$.
Every element is in $N/2$ sets, every two elements share $N/4$ sets.
Let $\chi : V \to \{-1, 1\}$.

$$\begin{aligned}
\sum_{S \in \mathcal{S}} \mathrm{disc}(S, \chi)^2 &= \sum_{S} \left(\sum_{v \in S} \chi(v) \right)^2 \\
&= \sum_{u,v} \chi(u)\chi(v) \cdot |\{S \in \mathcal{S} : u, v \in S\}| \\
&= N/4 \left(\sum_{v \in V} \chi(v)^2 + \left(\sum_{v \in V} \chi(v) \right)^2 \right) \\
&\geq N^2/4.
\end{aligned}$$

Proposition

There exists a set system $\mathcal{S}$ with $m = n$ sets and $\mathrm{disc}(\mathcal{S}) = \Omega(\sqrt{n})$.

Proof: Let $\mathcal{S}$ be the set system with "incidence matrix" H_n, set $N = 2^n$.
Every element is in $N/2$ sets, every two elements share $N/4$ sets.
Let $\chi : V \to \{-1, 1\}$.

$$\begin{aligned}
\sum_{S \in \mathcal{S}} \mathrm{disc}(S, \chi)^2 &= \sum_{S} \left(\sum_{v \in S} \chi(v) \right)^2 \\
&= \sum_{u,v} \chi(u)\chi(v) \cdot |\{S \in \mathcal{S} : u, v \in S\}| \\
&= N/4 \left(\sum_{v \in V} \chi(v)^2 + \left(\sum_{v \in V} \chi(v) \right)^2 \right) \\
&\geq N^2/4.
\end{aligned}$$

$\to$ There exists $S \in \mathcal{S}$ with $\mathrm{disc}(S, \chi) = \Omega(\sqrt{n})$.

More discrepancy

More discrepancy

Definition

The *hereditary discrepancy* of $\mathcal{S}$ is $\text{herdisc}(\mathcal{S}) = \max_{U \subseteq V} \text{disc}(\mathcal{S}_{|U})$.

More discrepancy

Definition

The *hereditary discrepancy* of $\mathcal{S}$ is $\text{herdisc}(\mathcal{S}) = \max_{U \subseteq V} \text{disc}(\mathcal{S}_{|U})$.
The *linear discrepancy* of $\mathcal{S}$ is

$$\text{lindisc}(\mathcal{S}) = \max_{w:V \to [-1,1]} \min_{\chi:V \to \{-1,1\}} \max_{S \in \mathcal{S}} |w(S) - \chi(S)|.$$

More discrepancy

Definition

The *hereditary discrepancy* of $\mathcal{S}$ is $\mathrm{herdisc}(\mathcal{S}) = \max_{U \subseteq V} \mathrm{disc}(\mathcal{S}_{|U})$.
The *linear discrepancy* of $\mathcal{S}$ is

$$\mathrm{lindisc}(\mathcal{S}) = \max_{w:V \to [-1,1]} \min_{\chi:V \to \{-1,1\}} \max_{S \in \mathcal{S}} |w(S) - \chi(S)|.$$

Definition

The *hereditary discrepancy* of $A \in \mathbb{R}^{m \times n}$ is $\mathrm{herdisc}(A) = \max_{A' \sqsubseteq A} \mathrm{disc}(A')$.

More discrepancy

Definition

The *hereditary discrepancy* of $\mathcal{S}$ is $\mathrm{herdisc}(\mathcal{S}) = \max_{U \subseteq V} \mathrm{disc}(\mathcal{S}_{|U})$.
The *linear discrepancy* of $\mathcal{S}$ is

$$\mathrm{lindisc}(\mathcal{S}) = \max_{w:V \to [-1,1]} \min_{\chi:V \to \{-1,1\}} \max_{S \in \mathcal{S}} |w(S) - \chi(S)|.$$

Definition

The *hereditary discrepancy* of $A \in \mathbb{R}^{m \times n}$ is $\mathrm{herdisc}(A) = \max_{A' \sqsubseteq A} \mathrm{disc}(A')$.
The *linear discrepancy* of A is

$$\mathrm{lindisc}(A) = \max_{w \in [-1,1]^n} \min_{x \in \{-1,1\}^n} \|A(w - x)\|_\infty .$$

More discrepancy

Definition

The *hereditary discrepancy* of $\mathcal{S}$ is $\text{herdisc}(\mathcal{S}) = \max_{U \subseteq V} \text{disc}(\mathcal{S}_{|U})$.
The *linear discrepancy* of $\mathcal{S}$ is

$$\text{lindisc}(\mathcal{S}) = \max_{w: V \to [-1,1]} \min_{\chi: V \to \{-1,1\}} \max_{S \in \mathcal{S}} |w(S) - \chi(S)|.$$

Definition

The *hereditary discrepancy* of $A \in \mathbb{R}^{m \times n}$ is $\text{herdisc}(A) = \max_{A' \sqsubseteq A} \text{disc}(A')$.
The *linear discrepancy* of A is

$$\text{lindisc}(A) = \max_{w \in [-1,1]^n} \min_{x \in \{-1,1\}^n} \|A(w - x)\|_\infty .$$

Proposition

Let $A \in \mathbb{R}^{m \times n}$ and $w \in (\mathbb{R}^+)^n$. There exists $x \in \mathbb{N}^n$ such that

$$\|A(w - x)\|_\infty \leq \text{lindisc}(A).$$

Hereditary discrepancy and linear discrepancy

Proposition

$\forall A$, $\text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A)$.

Hereditary discrepancy and linear discrepancy

Proposition

$\forall A$, $\mathrm{lindisc}(A) \leq 2 \cdot \mathrm{herdisc}(A)$.

Goal: $\forall w \in [-1,1]^n, \exists x \in \{-1,1\}^n : \|A(w-x)\|_\infty \leq 2 \cdot \mathrm{herdisc}(A)$.

Hereditary discrepancy and linear discrepancy

Proposition

$\forall A$, $\text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A)$.

Goal: $\forall w \in [-1,1]^n, \exists x \in \{-1,1\}^n : \|A(w-x)\|_\infty \leq 2 \cdot \text{herdisc}(A)$.
Let $w \in [-1,1]^n$.

Hereditary discrepancy and linear discrepancy

Proposition

$\forall A$, $\text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A)$.

Goal: $\forall w \in [-1,1]^n, \exists x \in \{-1,1\}^n : \|A(w-x)\|_\infty \leq 2 \cdot \text{herdisc}(A)$.
Let $w \in [-1,1]^n$. Let $y_1 \in \{-1,1\}^n$ be the sign vector of w.

Hereditary discrepancy and linear discrepancy

Proposition

$\forall A$, $\text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A)$.

Goal: $\forall w \in [-1,1]^n, \exists x \in \{-1,1\}^n : \|A(w-x)\|_\infty \leq 2 \cdot \text{herdisc}(A)$.
Let $w \in [-1,1]^n$. Let $y_1 \in \{-1,1\}^n$ be the sign vector of w.
Then, $2w - y_1 \in [-1,1]^n$ so there exists $y_2 \in \{-1,1\}^n$ such that $\|A(2w - y_1 - y_2)\|_\infty \leq \text{lindisc}(A)$.

Hereditary discrepancy and linear discrepancy

Proposition

$\forall A$, $\text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A)$.

Goal: $\forall w \in [-1,1]^n, \exists x \in \{-1,1\}^n : \|A(w-x)\|_\infty \leq 2 \cdot \text{herdisc}(A)$.
Let $w \in [-1,1]^n$. Let $y_1 \in \{-1,1\}^n$ be the sign vector of w.
Then, $2w - y_1 \in [-1,1]^n$ so there exists $y_2 \in \{-1,1\}^n$ such that $\|A(2w - y_1 - y_2)\|_\infty \leq \text{lindisc}(A)$.
Let $z = \frac{y_1+y_2}{2} \in \{-1,0,1\}^n$.
Then, $\|A(w-z)\|_\infty \leq \text{lindisc}(A)/2$.

Hereditary discrepancy and linear discrepancy

Proposition

$\forall A$, $\text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A)$.

Goal: $\forall w \in [-1,1]^n, \exists x \in \{-1,1\}^n : \|A(w-x)\|_\infty \leq 2 \cdot \text{herdisc}(A)$.
Let $w \in [-1,1]^n$. Let $y_1 \in \{-1,1\}^n$ be the sign vector of w.
Then, $2w - y_1 \in [-1,1]^n$ so there exists $y_2 \in \{-1,1\}^n$ such that $\|A(2w - y_1 - y_2)\|_\infty \leq \text{lindisc}(A)$.
Let $z = \frac{y_1+y_2}{2} \in \{-1,0,1\}^n$.
Then, $\|A(w-z)\|_\infty \leq \text{lindisc}(A)/2$.
There exists $x \in \{-1,1\}^n$ so that $\|A(z-x)\|_\infty \leq \text{herdisc}(A)$.

Hereditary discrepancy and linear discrepancy

Proposition

$\forall A$, $\text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A)$.

Goal: $\forall w \in [-1,1]^n, \exists x \in \{-1,1\}^n : \|A(w-x)\|_\infty \leq 2 \cdot \text{herdisc}(A)$.
Let $w \in [-1,1]^n$. Let $y_1 \in \{-1,1\}^n$ be the sign vector of w.
Then, $2w - y_1 \in [-1,1]^n$ so there exists $y_2 \in \{-1,1\}^n$ such that $\|A(2w - y_1 - y_2)\|_\infty \leq \text{lindisc}(A)$.
Let $z = \frac{y_1+y_2}{2} \in \{-1,0,1\}^n$.
Then, $\|A(w-z)\|_\infty \leq \text{lindisc}(A)/2$.
There exists $x \in \{-1,1\}^n$ so that $\|A(z-x)\|_\infty \leq \text{herdisc}(A)$.
Then, $\|A(w-x)\|_\infty \leq \text{lindisc}(A)/2 + \text{herdisc}(A)$.

Hereditary discrepancy and linear discrepancy

Proposition

$\forall A$, $\text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A)$.

Goal: $\forall w \in [-1,1]^n, \exists x \in \{-1,1\}^n : \|A(w-x)\|_\infty \leq 2 \cdot \text{herdisc}(A)$.
Let $w \in [-1,1]^n$. Let $y_1 \in \{-1,1\}^n$ be the sign vector of w.
Then, $2w - y_1 \in [-1,1]^n$ so there exists $y_2 \in \{-1,1\}^n$ such that $\|A(2w - y_1 - y_2)\|_\infty \leq \text{lindisc}(A)$.
Let $z = \frac{y_1+y_2}{2} \in \{-1,0,1\}^n$.
Then, $\|A(w-z)\|_\infty \leq \text{lindisc}(A)/2$.
There exists $x \in \{-1,1\}^n$ so that $\|A(z-x)\|_\infty \leq \text{herdisc}(A)$.
Then, $\|A(w-x)\|_\infty \leq \text{lindisc}(A)/2 + \text{herdisc}(A)$.
$\rightarrow \text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A)$.

Hereditary discrepancy and linear discrepancy

Proposition

$\forall A$, $\text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A)$.

Goal: $\forall w \in [-1,1]^n, \exists x \in \{-1,1\}^n : \|A(w-x)\|_\infty \leq 2 \cdot \text{herdisc}(A)$.
Let $w \in [-1,1]^n$. Let $y_1 \in \{-1,1\}^n$ be the sign vector of w.
Then, $2w - y_1 \in [-1,1]^n$ so there exists $y_2 \in \{-1,1\}^n$ such that
$\|A(2w - y_1 - y_2)\|_\infty \leq \text{lindisc}(A)$.
Let $z = \frac{y_1+y_2}{2} \in \{-1,0,1\}^n$.
Then, $\|A(w-z)\|_\infty \leq \text{lindisc}(A)/2$.
There exists $x \in \{-1,1\}^n$ so that $\|A(z-x)\|_\infty \leq \text{herdisc}(A)$.
Then, $\|A(w-x)\|_\infty \leq \text{lindisc}(A)/2 + \text{herdisc}(A)$.
$\rightarrow \text{lindisc}(A) \leq 2 \cdot \text{herdisc}(A)$.

Remark: There are also efficient constructive proofs.

Beck-Fiala

Theorem [Beck, Fiala '81]

Let $\mathcal{S}$ be a set system where every $v \in V$ is in $\leq \Delta$ sets. Then, $\text{herdisc}(\mathcal{S}) \leq 2\Delta - 1$.

Beck-Fiala

Theorem [Beck, Fiala '81]

Let $\mathcal{S}$ be a set system where every $v \in V$ is in $\leq \Delta$ sets. Then, $\text{herdisc}(\mathcal{S}) \leq 2\Delta - 1$.

Remark: Beck and Fiala conjectured that $\text{herdisc}(\mathcal{S}) = O(\sqrt{\Delta})$. This is still a central open question in discrepancy theory.

Scheduling

Scheduling problem: m machines, n jobs, $t_{i,j}$ = time job i takes on machine j.

Scheduling

Scheduling problem: m machines, n jobs, $t_{i,j}$ = time job i takes on machine j.
Goal: Find minimum time needed to execute all the jobs.

Scheduling

Scheduling problem: m machines, n jobs, $t_{i,j}$ = time job i takes on machine j.
Goal: Find minimum time needed to execute all the jobs.

Known: polytime 2-approximation algorithm but no $(3/2 - \varepsilon)$-approximation unless P = NP.

Scheduling

Scheduling problem: m machines, n jobs, $t_{i,j}$ = time job i takes on machine j.
Goal: Find minimum time needed to execute all the jobs.

Known: polytime 2-approximation algorithm but no $(3/2 - \varepsilon)$-approximation unless P = NP.

Theorem

There is a polytime $O(1)$-approximation algorithm.

Approximation algorithm via discrepancy

Theorem

There is a polytime $O(1)$-approximation algorithm.

Approximation algorithm via discrepancy

Theorem

There is a polytime $O(1)$-approximation algorithm.

By binary search, guess $OPT \leq T < 2OPT$.

Approximation algorithm via discrepancy

Theorem

There is a polytime $O(1)$-approximation algorithm.

By binary search, guess $OPT \leq T < 2OPT$.
Find a solution $\overline{y}$ of the following LP:

$$\begin{aligned}
\sum_{j \in [m]} y_{i,j} &\geq 1, & \forall i \in [n] \\
\sum_{i \in [n]} y_{i,j} \cdot \frac{t_{i,j}}{T} &\leq 1, & \forall j \in [m] \\
y_{i,j} &\in [0,1], & \forall (i,j) \in [n] \times [m]
\end{aligned}$$

Approximation algorithm via discrepancy

Theorem

There is a polytime $O(1)$-approximation algorithm.

By binary search, guess $OPT \leq T < 2OPT$.
Find a solution $\overline{y}$ of the following LP:
Setting $\overline{w} = 7 \cdot \overline{y}$, we get:

$$
\begin{aligned}
\sum_{j \in [m]} w_{i,j} &\geq 7, && \forall i \in [n] \\
\sum_{i \in [n]} w_{i,j} \cdot \tfrac{t_{i,j}}{T} &\leq 7, && \forall j \in [m] \\
w_{i,j} &\in [0, 7], && \forall (i,j) \in [n] \times [m]
\end{aligned}
$$

Approximation algorithm via discrepancy

Theorem

There is a polytime $O(1)$-approximation algorithm.

By binary search, guess $OPT \leq T < 2OPT$.
Find a solution $\overline{y}$ of the following LP:
Setting $\overline{w} = 7 \cdot \overline{y}$, we get:
Every $w_{i,j}$ appears twice, with coefficients at most 1, so herdisc ≤ 3 by Beck-Fiala, so lindisc ≤ 6.

$$
\begin{aligned}
\sum_{j\in[m]} w_{i,j} \geq 7, \qquad & \forall i \in [n] \\
\sum_{i\in[n]} w_{i,j} \cdot \tfrac{t_{i,j}}{T} \leq 7, \ & \forall j \in [m] \\
w_{i,j} \in [0,7], \qquad & \forall (i,j) \in [n] \times [m]
\end{aligned}
$$

Approximation algorithm via discrepancy

Theorem

There is a polytime $O(1)$-approximation algorithm.

By binary search, guess $OPT \leq T < 2OPT$.
Find a solution $\overline{y}$ of the following LP:
Setting $\overline{w} = 7 \cdot \overline{y}$, we get:
Every $w_{i,j}$ appears twice, with coefficients at most 1, so herdisc ≤ 3 by Beck-Fiala, so lindisc ≤ 6.
Thus, there exists $\overline{x}$ such that:

$$\begin{aligned} \sum_{j \in [m]} x_{i,j} \geq 1, & \quad \forall i \in [n] \\ \sum_{i \in [n]} x_{i,j} \cdot \frac{t_{i,j}}{T} \leq 13, & \quad \forall j \in [m] \\ x_{i,j} \in \mathbb{N}, & \quad \forall (i,j) \in [n] \times [m] \end{aligned}$$

Approximation algorithm via discrepancy

Theorem

There is a polytime $O(1)$-approximation algorithm.

By binary search, guess $OPT \leq T < 2OPT$.
Find a solution $\overline{y}$ of the following LP:
Setting $\overline{w} = 7 \cdot \overline{y}$, we get:
Every $w_{i,j}$ appears twice, with coefficients at most 1, so herdisc ≤ 3 by Beck-Fiala, so lindisc ≤ 6.
Thus, there exists $\overline{x}$ such that:

$$\begin{aligned}
&\sum_{j\in[m]} x_{i,j} \geq 1, && \forall i \in [n] \\
&\sum_{i\in[n]} x_{i,j} \cdot \frac{t_{i,j}}{T} \leq 13, && \forall j \in [m] \\
&x_{i,j} \in \mathbb{N}, && \forall (i,j) \in [n] \times [m]
\end{aligned}$$

Thank you!