Optimal Online Discrepancy Minimization

Thomas Rothvoss

Joint work with Janardhan Kulkarni and Victor Reis

UNIVERSITY *of* WASHINGTON

Online discrepancy

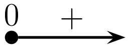
- **Input:** Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ revealed one at a time
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^{t} x_i v_i$ short for all $t \in [T]$

Online discrepancy

- **Input:** Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ revealed one at a time
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^{t} x_i v_i$ short for all $t \in [T]$

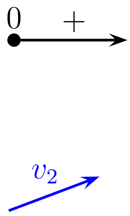
Online discrepancy

- **Input:** Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ revealed one at a time
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^{t} x_i v_i$ short for all $t \in [T]$



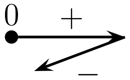
Online discrepancy

- **Input:** Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ revealed one at a time
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^t x_i v_i$ short for all $t \in [T]$



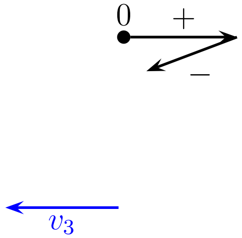
Online discrepancy

- **Input:** Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ revealed one at a time
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^{t} x_i v_i$ short for all $t \in [T]$



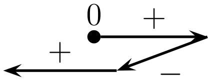
Online discrepancy

- **Input:** Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ revealed one at a time
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^{t} x_i v_i$ short for all $t \in [T]$



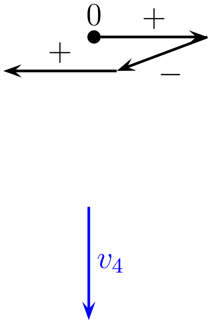
Online discrepancy

- **Input:** Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ revealed one at a time
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^t x_i v_i$ short for all $t \in [T]$



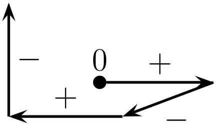
Online discrepancy

- **Input:** Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ revealed one at a time
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^{t} x_i v_i$ short for all $t \in [T]$



Online discrepancy

- **Input:** Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \le 1$ revealed one at a time
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^{t} x_i v_i$ short for all $t \in [T]$



A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\|\sum_{i=1}^{T} x_i v_i\|_2 \geq \sqrt{T}$.

A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\| \sum_{i=1}^{T} x_i v_i \|_2 \geq \sqrt{T}$.

Proof.

- Pick v_t orthogonal to $w_{t-1} := \sum_{i=1}^{t-1} x_i v_i$ with $\|v_t\|_2 = 1$

A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\|\sum_{i=1}^{T} x_i v_i\|_2 \geq \sqrt{T}$.

Proof.

- Pick v_t orthogonal to $w_{t-1} := \sum_{i=1}^{t-1} x_i v_i$ with $\|v_t\|_2 = 1$

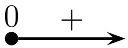
A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\|\sum_{i=1}^{T} x_i v_i\|_2 \geq \sqrt{T}$.

Proof.

- Pick v_t orthogonal to $w_{t-1} := \sum_{i=1}^{t-1} x_i v_i$ with $\|v_t\|_2 = 1$



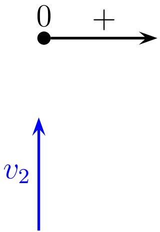
A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\| \sum_{i=1}^{T} x_i v_i \|_2 \geq \sqrt{T}$.

Proof.

- Pick v_t orthogonal to $w_{t-1} := \sum_{i=1}^{t-1} x_i v_i$ with $\|v_t\|_2 = 1$



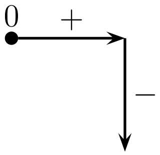
A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\|\sum_{i=1}^{T} x_i v_i\|_2 \geq \sqrt{T}$.

Proof.

- Pick v_t orthogonal to $w_{t-1} := \sum_{i=1}^{t-1} x_i v_i$ with $\|v_t\|_2 = 1$



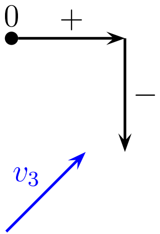
A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\|\sum_{i=1}^{T} x_i v_i\|_2 \geq \sqrt{T}$.

Proof.

- Pick v_t orthogonal to $w_{t-1} := \sum_{i=1}^{t-1} x_i v_i$ with $\|v_t\|_2 = 1$



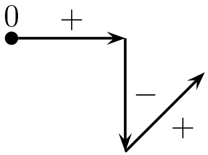
A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\|\sum_{i=1}^{T} x_i v_i\|_2 \geq \sqrt{T}$.

Proof.

- Pick v_t orthogonal to $w_{t-1} := \sum_{i=1}^{t-1} x_i v_i$ with $\|v_t\|_2 = 1$



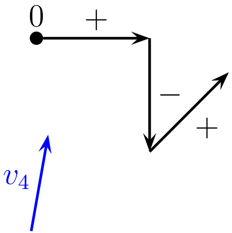
A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\|\sum_{i=1}^{T} x_i v_i\|_2 \geq \sqrt{T}$.

Proof.

- Pick v_t orthogonal to $w_{t-1} := \sum_{i=1}^{t-1} x_i v_i$ with $\|v_t\|_2 = 1$



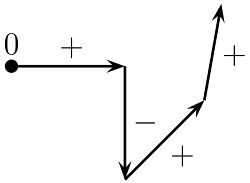
A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\| \sum_{i=1}^{T} x_i v_i \|_2 \geq \sqrt{T}$.

Proof.

- Pick v_t orthogonal to $w_{t-1} := \sum_{i=1}^{t-1} x_i v_i$ with $\|v_t\|_2 = 1$
- Easy: $\|w_t\|_2^2 = \|w_{t-1}\|_2^2 + 2x_t \langle w_{t-1}, v_t \rangle + \|v_t\|_2^2 = t$



A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\| \sum_{i=1}^{T} x_i v_i \|_2 \geq \sqrt{T}$.

Proof.

- Pick v_t orthogonal to $w_{t-1} := \sum_{i=1}^{t-1} x_i v_i$ with $\|v_t\|_2 = 1$
- Easy: $\|w_t\|_2^2 = \|w_{t-1}\|_2^2 + 2x_t \langle w_{t-1}, v_t \rangle + \|v_t\|_2^2 = t$

Observations:

- Player also has strategy to keep $\| \sum_{i=1}^{T} x_i v_i \|_2 \leq \sqrt{T}$

A lower bound

Theorem (Folklore)

An **adaptive adversary** *has a strategy so that for any player* $\|\sum_{i=1}^{T} x_i v_i\|_2 \geq \sqrt{T}$.

Proof.

- Pick v_t orthogonal to $w_{t-1} := \sum_{i=1}^{t-1} x_i v_i$ with $\|v_t\|_2 = 1$
- Easy: $\|w_t\|_2^2 = \|w_{t-1}\|_2^2 + 2x_t \langle w_{t-1}, v_t \rangle + \|v_t\|_2^2 = t$

Observations:

- Player also has strategy to keep $\|\sum_{i=1}^{T} x_i v_i\|_2 \leq \sqrt{T}$
- **Randomization** does not help player

Oblivious adversary

- **Input:** Adversary fixes vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ **in advance**. Then reveals vectors one at a time.
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^{t} x_i v_i$ approximately balanced for all $t \in [T]$

Oblivious adversary

- **Input:** Adversary fixes vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ **in advance**. Then reveals vectors one at a time.
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^{t} x_i v_i$ approximately balanced for all $t \in [T]$

Thoughts: How is this easier for us? Maybe arriving vector v_t still orthogonal to current position. Also would not make a difference for existing potential function arguments!

Oblivious adversary

- **Input:** Adversary fixes vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ **in advance**. Then reveals vectors one at a time.
- **Goal:** Find signs $x_1, \ldots, x_T \in \{-1, 1\}$ **online** so that $\sum_{i=1}^{t} x_i v_i$ approximately balanced for all $t \in [T]$

Thoughts: How is this easier for us? Maybe arriving vector v_t still orthogonal to current position. Also would not make a difference for existing potential function arguments!
But now **randomization** can help!

The Self-balancing Walk (Alweiss et al)

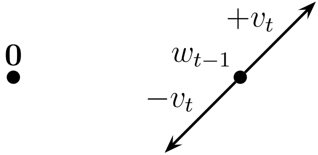
(1) Set $w_0 := \mathbf{0}$ and $s := \Theta(\sqrt{\log(nT)})$
(2) FOR $t = 1$ TO T
   (3) With probability $\frac{1}{2} - \frac{\langle w_{t-1}, v_t \rangle}{2s}$ set $x_t := 1$, otherwise $x_t := -1$
   (4) Update $w_t := w_{t-1} + x_t v_t$

The Self-balancing Walk (Alweiss et al)

(1) Set $w_0 := \mathbf{0}$ and $s := \Theta(\sqrt{\log(nT)})$
(2) FOR $t = 1$ TO T
  (3) With probability $\frac{1}{2} - \frac{\langle w_{t-1}, v_t \rangle}{2s}$ set $x_t := 1$, otherwise $x_t := -1$
  (4) Update $w_t := w_{t-1} + x_t v_t$

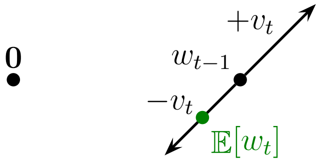
The Self-balancing Walk (Alweiss et al)

(1) Set $w_0 := \mathbf{0}$ and $s := \Theta(\sqrt{\log(nT)})$
(2) FOR $t = 1$ TO T
  (3) With probability $\frac{1}{2} - \frac{\langle w_{t-1}, v_t \rangle}{2s}$ set $x_t := 1$, otherwise $x_t := -1$
  (4) Update $w_t := w_{t-1} + x_t v_t$



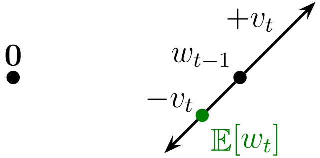
The Self-balancing Walk (Alweiss et al)

(1) Set $w_0 := \mathbf{0}$ and $s := \Theta(\sqrt{\log(nT)})$
(2) FOR $t = 1$ TO T
  (3) With probability $\frac{1}{2} - \frac{\langle w_{t-1}, v_t \rangle}{2s}$ set $x_t := 1$, otherwise $x_t := -1$
  (4) Update $w_t := w_{t-1} + x_t v_t$



The Self-balancing Walk (Alweiss et al)

(1) Set $w_0 := \mathbf{0}$ and $s := \Theta(\sqrt{\log(nT)})$
(2) FOR $t = 1$ TO T
    (3) With probability $\frac{1}{2} - \frac{\langle w_{t-1}, v_t \rangle}{2s}$ set $x_t := 1$, otherwise $x_t := -1$
    (4) Update $w_t := w_{t-1} + x_t v_t$

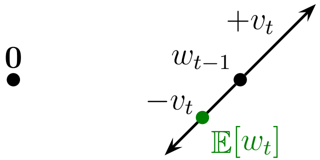


Theorem (Alweiss, Liu, Sawhney STOC 2021)

For any t, w_t is $\Theta(\sqrt{\log(nT)})$-subgaussian in every direction.

The Self-balancing Walk (Alweiss et al)

(1) Set $w_0 := \mathbf{0}$ and $s := \Theta(\sqrt{\log(nT)})$
(2) FOR $t = 1$ TO T
   (3) With probability $\frac{1}{2} - \frac{\langle w_{t-1}, v_t \rangle}{2s}$ set $x_t := 1$, otherwise $x_t := -1$
   (4) Update $w_t := w_{t-1} + x_t v_t$



Theorem (Alweiss, Liu, Sawhney STOC 2021)

For any t, w_t is $\Theta(\sqrt{\log(nT)})$-subgaussian in every direction.

Example:

- $\mathbb{E}[\|w_T\|_2] \leq O(\sqrt{n\log(nT)})$
- $\mathbb{E}[\max_t \|w_t\|_\infty] \leq O(\log(nT))$

Main contribution

Theorem (Kulkarni, Reis, R.'23)

There is an **online coloring strategy** that for any sequence $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$, fixed **in advance**, revealed one at a time, finds **random signs** $x_1, \ldots, x_T \in \{-1, 1\}$ so that: for each t, $\sum_{i=1}^{t} x_i v_i$ is $O(1)$**-subgaussian**.

Main contribution

Theorem (Kulkarni, Reis, R.'23)

There is an **online coloring strategy** that for any sequence $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$, fixed **in advance**, revealed one at a time, finds **random signs** $x_1, \ldots, x_T \in \{-1, 1\}$ so that: for each t, $\sum_{i=1}^t x_i v_i$ is $O(1)$**-subgaussian**.

- Improves over the $\Theta(\sqrt{\log(nT)})$-subgaussian bound by [ALS'21] ...

Main contribution

Theorem (Kulkarni, Reis, R.'23)

There is an **online coloring strategy** that for any sequence $v_1, \dots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$, fixed **in advance**, revealed one at a time, finds **random signs** $x_1, \dots, x_T \in \{-1, 1\}$ so that: for each t, $\sum_{i=1}^t x_i v_i$ is $O(1)$**-subgaussian**.

- Improves over the $\Theta(\sqrt{\log(nT)})$-subgaussian bound by [ALS'21] ...
- ...but not quite constructive (best running time we can prove is $2^{2^{\mathrm{poly}(nT)}}$).

Subgaussianity

Lemma (Subgaussianity)

For a random variable X the following is equivalent (up to constant factors)

(i) $\Pr[|X| \geq t] \leq 2\exp(-t^2/\sigma^2)$ *for all* $t \geq 0$.

Subgaussianity

Lemma (Subgaussianity)

For a random variable X the following is equivalent (up to constant factors)

(i) $\Pr[|X| \geq t] \leq 2\exp(-t^2/\sigma^2)$ *for all* $t \geq 0$.

(ii) $\mathbb{E}[\exp(\frac{X^2}{\sigma^2})] \leq 2$

Subgaussianity

Lemma (Subgaussianity)

For a random variable X the following is equivalent (up to constant factors)

(i) $\Pr[|X| \geq t] \leq 2\exp(-t^2/\sigma^2)$ *for all* $t \geq 0$.

(ii) $\mathbb{E}[\exp(\frac{X^2}{\sigma^2})] \leq 2$

If $\mathbb{E}[X = 0]$ *then also equivalent to*

(iii) $\mathbb{E}[\exp(\lambda X)] \leq \exp(\lambda^2\sigma^2)$ *for all* $\lambda \geq 0$.

Subgaussianity

Lemma (Subgaussianity)

For a random variable X the following is equivalent (up to constant factors)

(i) $\Pr[|X| \geq t] \leq 2\exp(-t^2/\sigma^2)$ *for all* $t \geq 0$.

(ii) $\mathbb{E}[\exp(\frac{X^2}{\sigma^2})] \leq 2$ $\leftarrow$

If $\mathbb{E}[X = 0]$ *then also equivalent to*

(iii) $\mathbb{E}[\exp(\lambda X)] \leq \exp(\lambda^2\sigma^2)$ *for all* $\lambda \geq 0$.

Definition

The **subgaussian norm** of a random variable is

$$\|X\|_{\psi_2} := \inf\left\{t > 0 : \mathbb{E}[\exp(X^2/t^2)] \leq 2\right\}.$$

Subgaussianity (2)

Definition

The **subgaussian norm** of a random variable is

$$\|X\|_{\psi_2} := \inf \left\{ t > 0 : \mathbb{E}[\exp(X^2/t^2)] \le 2 \right\}.$$

Examples:

- If $X \sim N(0, \sigma^2)$, then $\|X\|_{\psi_2} \asymp \sigma$
- If $X \sim \{-\sigma, \sigma\}$ uniformly, then $\|X\|_{\psi_2} \asymp \sigma$

Subgaussianity (2)

Definition

The **subgaussian norm** of a random variable is

$$\|X\|_{\psi_2} := \inf \left\{ t > 0 : \mathbb{E}[\exp(X^2/t^2)] \leq 2 \right\}.$$

Examples:

- If $X \sim N(0, \sigma^2)$, then $\|X\|_{\psi_2} \asymp \sigma$
- If $X \sim \{-\sigma, \sigma\}$ uniformly, then $\|X\|_{\psi_2} \asymp \sigma$

Fact:

- $\|\cdot\|_{\psi_2}$ is in fact a **norm**, i.e. $\|X + Y\|_{\psi_2} \leq \|X\|_{\psi_2} + \|Y\|_{\psi_2}$ even for non-independent random variables.

Subgaussianity (2)

Definition

The **subgaussian norm** of a random variable is

$$\|X\|_{\psi_2} := \inf\left\{t > 0 : \mathbb{E}[\exp(X^2/t^2)] \leq 2\right\}.$$

Examples:

- If $X \sim N(0, \sigma^2)$, then $\|X\|_{\psi_2} \asymp \sigma$
- If $X \sim \{-\sigma, \sigma\}$ uniformly, then $\|X\|_{\psi_2} \asymp \sigma$

Fact:

- $\|\cdot\|_{\psi_2}$ is in fact a **norm**, i.e. $\|X + Y\|_{\psi_2} \leq \|X\|_{\psi_2} + \|Y\|_{\psi_2}$ even for non-independent random variables.

Definition

For a **random vector** $X \in \mathbb{R}^n$, set

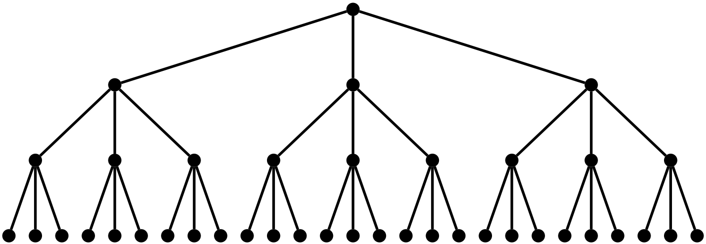
$$\|X\|_{\psi_{2,\infty}} := \sup_{w \in S^{n-1}} \|\langle X, w \rangle\|_{\psi_2}.$$

Reduction

Claim. Online algorithm reduces to the following:

Theorem

Given any **rooted tree** $\mathcal{T} = (V, E)$,

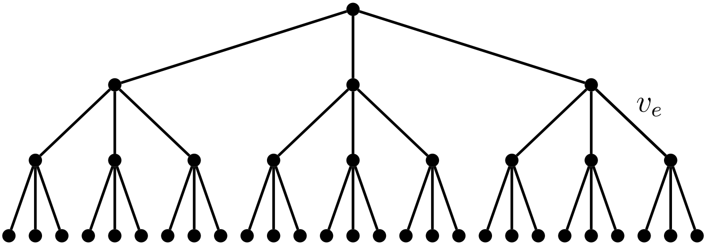


Reduction

Claim. Online algorithm reduces to the following:

Theorem

Given any **rooted tree** $\mathcal{T} = (V, E)$, *edges labeled with* $v_e \in B_2^n$.

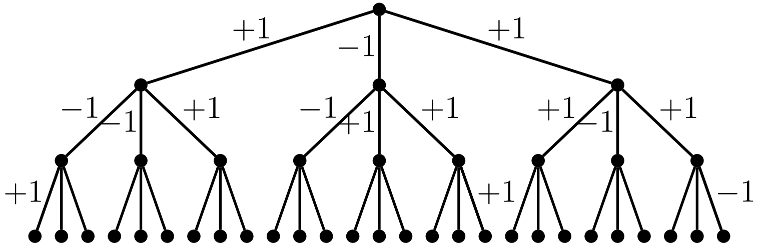


Reduction

Claim. Online algorithm reduces to the following:

Theorem

Given any **rooted tree** $\mathcal{T} = (V, E)$*, edges labeled with* $v_e \in B_2^n$*. There is a distribution* $\mathcal{D}$ *over signs* $x \in \{-1, 1\}^E$

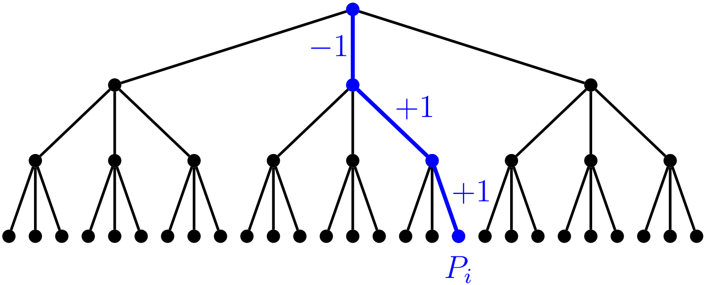


Reduction

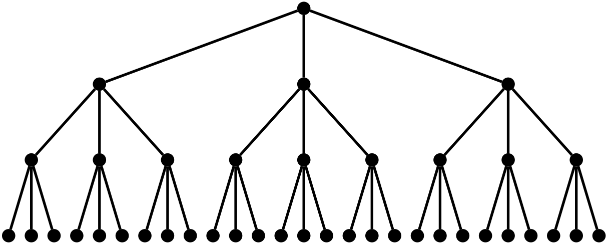
Claim. Online algorithm reduces to the following:

Theorem

Given any **rooted tree** $\mathcal{T} = (V, E)$*, edges labeled with* $v_e \in B_2^n$*. There is a distribution* $\mathcal{D}$ *over signs* $x \in \{-1, 1\}^E$ *so that* $\sum_{e \in P_i} x_e v_e$ *is* $O(1)$**-subgaussian** *for each* **root-leaf path** P_i.

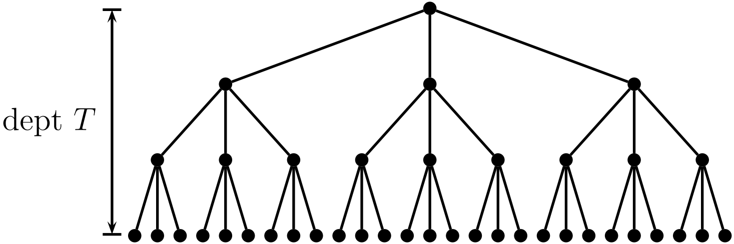


Reduction (2)



Reduction:

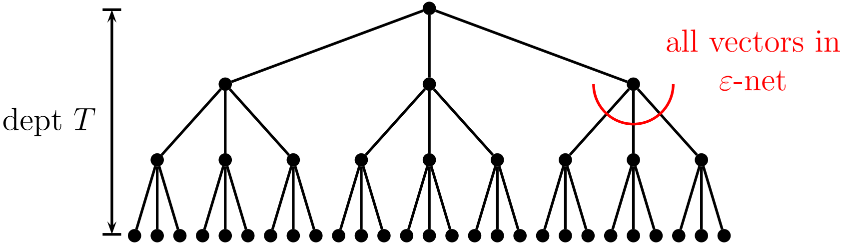
Reduction (2)



Reduction:

- Construction of $\mathcal{T}$: Tree has depth T

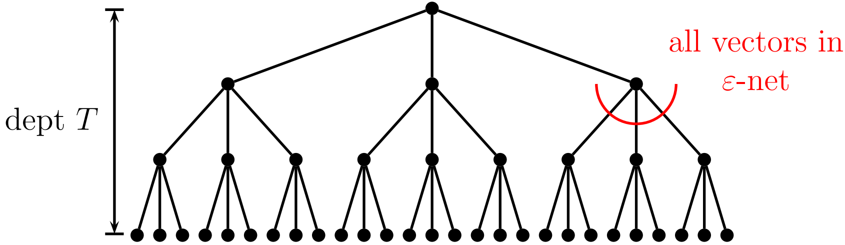
Reduction (2)



Reduction:

- Construction of $\mathcal{T}$: Tree has depth T, each interior nodes has outgoing edge for every vector in a fine **ε-net** of B_2^n

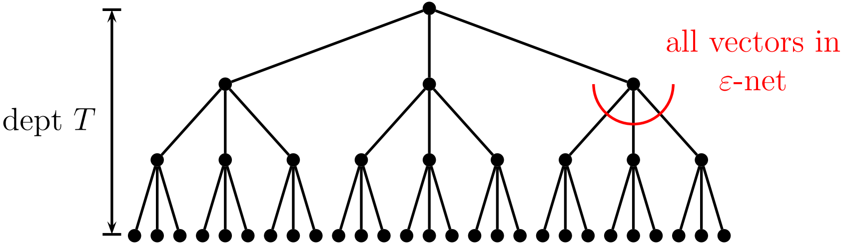
Reduction (2)



Reduction:

- Construction of $\mathcal{T}$: Tree has depth T, each interior nodes has outgoing edge for every vector in a fine **ε-net** of B_2^n
- Fix distribution $\mathcal{D}$ over $\{-1, 1\}^E$ giving $O(1)$-subgaussian root-leaf sums

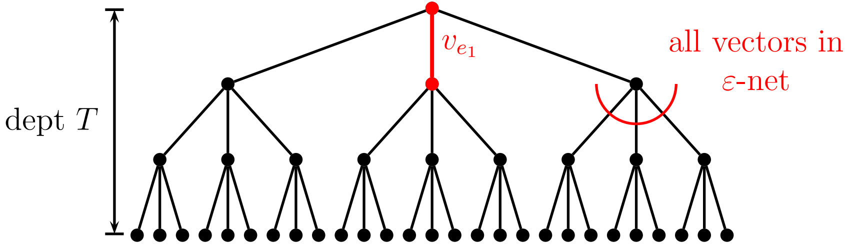
Reduction (2)



Reduction:

- Construction of $\mathcal{T}$: Tree has depth T, each interior nodes has outgoing edge for every vector in a fine **ε-net** of B_2^n
- Fix distribution $\mathcal{D}$ over $\{-1, 1\}^E$ giving $O(1)$-subgaussian root-leaf sums
- **Our strategy:** We draw $x \sim \mathcal{D}$.

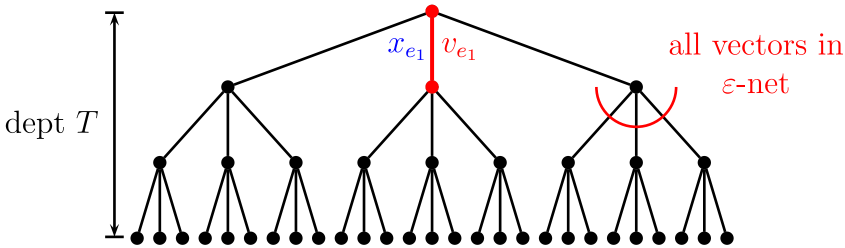
Reduction (2)



Reduction:

- Construction of $\mathcal{T}$: Tree has depth T, each interior nodes has outgoing edge for every vector in a fine **ε-net** of B_2^n
- Fix distribution $\mathcal{D}$ over $\{-1, 1\}^E$ giving $O(1)$-subgaussian root-leaf sums
- **Our strategy:** We draw $x \sim \mathcal{D}$. Then adversary reveals root-leaf path edge-by-edge. We reveal corresponding sign x_e.

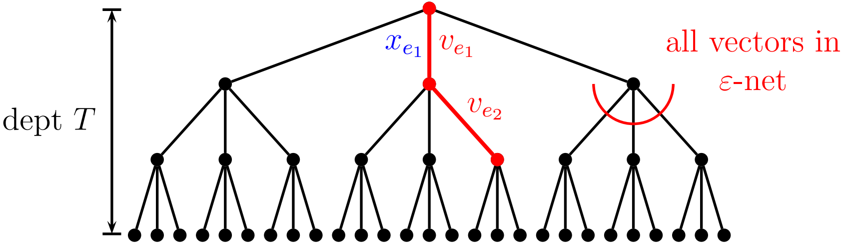
Reduction (2)



Reduction:

- Construction of $\mathcal{T}$: Tree has depth T, each interior nodes has outgoing edge for every vector in a fine **ε-net** of B_2^n
- Fix distribution $\mathcal{D}$ over $\{-1, 1\}^E$ giving $O(1)$-subgaussian root-leaf sums
- **Our strategy:** We draw $x \sim \mathcal{D}$. Then adversary reveals root-leaf path edge-by-edge. We reveal corresponding sign x_e.

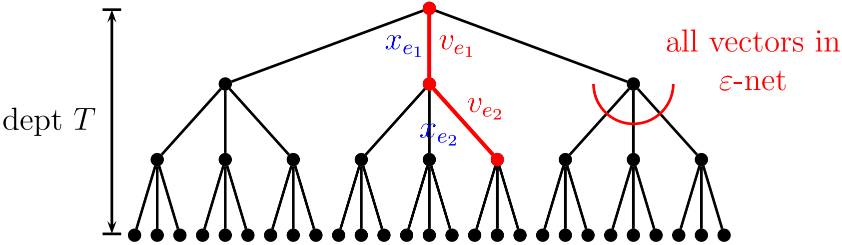
Reduction (2)



Reduction:

- Construction of $\mathcal{T}$: Tree has depth T, each interior nodes has outgoing edge for every vector in a fine **ε-net** of B_2^n
- Fix distribution $\mathcal{D}$ over $\{-1, 1\}^E$ giving $O(1)$-subgaussian root-leaf sums
- **Our strategy:** We draw $x \sim \mathcal{D}$. Then adversary reveals root-leaf path edge-by-edge. We reveal corresponding sign x_e.

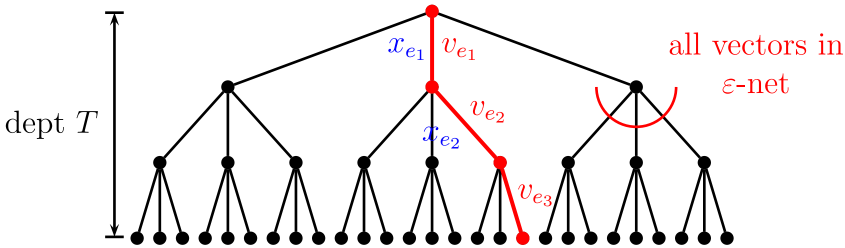
Reduction (2)



Reduction:

- Construction of $\mathcal{T}$: Tree has depth T, each interior nodes has outgoing edge for every vector in a fine **ε-net** of B_2^n
- Fix distribution $\mathcal{D}$ over $\{-1, 1\}^E$ giving $O(1)$-subgaussian root-leaf sums
- **Our strategy:** We draw $x \sim \mathcal{D}$. Then adversary reveals root-leaf path edge-by-edge. We reveal corresponding sign x_e.

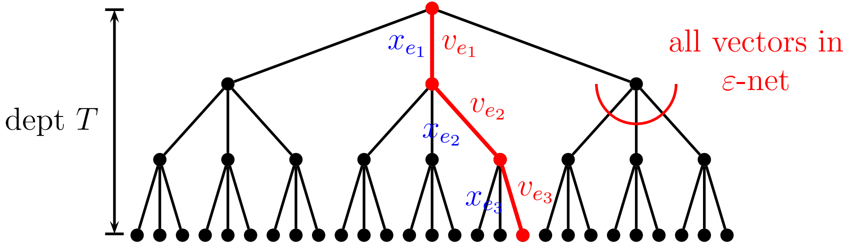
Reduction (2)



Reduction:

- Construction of $\mathcal{T}$: Tree has depth T, each interior nodes has outgoing edge for every vector in a fine **ε-net** of B_2^n
- Fix distribution $\mathcal{D}$ over $\{-1,1\}^E$ giving $O(1)$-subgaussian root-leaf sums
- **Our strategy:** We draw $x \sim \mathcal{D}$. Then adversary reveals root-leaf path edge-by-edge. We reveal corresponding sign x_e.

Reduction (2)



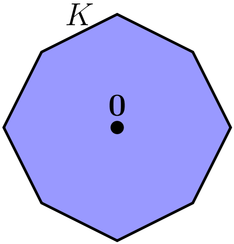
Reduction:

- Construction of $\mathcal{T}$: Tree has depth T, each interior nodes has outgoing edge for every vector in a fine **ε-net** of B_2^n
- Fix distribution $\mathcal{D}$ over $\{-1,1\}^E$ giving $O(1)$-subgaussian root-leaf sums
- **Our strategy:** We draw $x \sim \mathcal{D}$. Then adversary reveals root-leaf path edge-by-edge. We reveal corresponding sign x_e.

Banaszczyk Theorem

Theorem (Banaszczyk 1998)

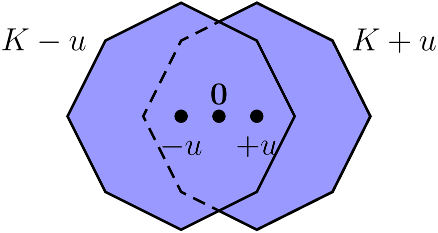
For any convex body $K \subseteq \mathbb{R}^n$ with $\gamma_n(K) \geq \frac{1}{2}$ and vector $u \in \mathbb{R}^n$ with $\|u\|_2 \leq \frac{1}{5}$, there is a convex body $(K * u) \subseteq (K + u) \cup (K - u)$ with $\gamma_n(K * u) \geq \gamma_n(K)$.



Banaszczyk Theorem

Theorem (Banaszczyk 1998)

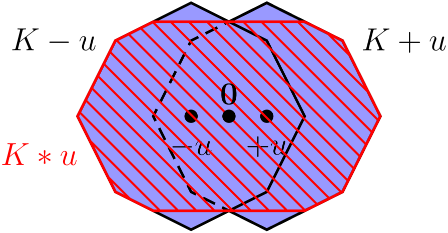
For any convex body $K \subseteq \mathbb{R}^n$ *with* $\gamma_n(K) \geq \frac{1}{2}$ *and vector* $u \in \mathbb{R}^n$ *with* $\|u\|_2 \leq \frac{1}{5}$*, there is a convex body* $(K * u) \subseteq (K + u) \cup (K - u)$ *with* $\gamma_n(K * u) \geq \gamma_n(K)$*.*



Banaszczyk Theorem

Theorem (Banaszczyk 1998)

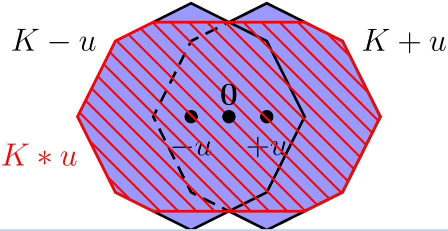
For any convex body $K \subseteq \mathbb{R}^n$ *with* $\gamma_n(K) \geq \frac{1}{2}$ *and vector* $u \in \mathbb{R}^n$ *with* $\|u\|_2 \leq \frac{1}{5}$, *there is a convex body* $(K * u) \subseteq (K + u) \cup (K - u)$ *with* $\gamma_n(K * u) \geq \gamma_n(K)$.



Banaszczyk Theorem

Theorem (Banaszczyk 1998)

For any convex body $K \subseteq \mathbb{R}^n$ with $\gamma_n(K) \geq \frac{1}{2}$ and vector $u \in \mathbb{R}^n$ with $\|u\|_2 \leq \frac{1}{5}$, there is a convex body $(K * u) \subseteq (K + u) \cup (K - u)$ with $\gamma_n(K * u) \geq \gamma_n(K)$.



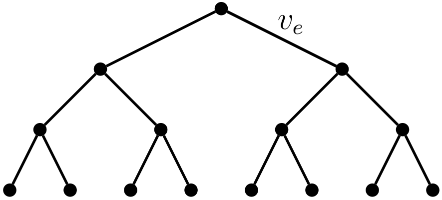
Theorem (Banaszczyk 1998)

For any convex body $K \subseteq \mathbb{R}^n$ with $\gamma_n(K) \geq \frac{1}{2}$ and vectors $v_1, \ldots, v_m$ with $\|v_i\|_2 \leq \frac{1}{5}$, there are signs $x \in \{-1, 1\}^m$ so that $\sum_{i=1}^m x_i v_i \in K$.

Signs for a labelled tree

Theorem

Let $\mathcal{T} = (V, E)$ be a **rooted tree** *where each edge $e \in E$ is assigned a vector $v_e \in \mathbb{R}^n$ with $\|v_e\|_2 \leq 1$. Let $K \subseteq \mathbb{R}^n$ be a* **convex body** *with $\gamma_n(K) \geq 1 - \frac{1}{2|E|}$. Then there are signs $x \in \{-1, 1\}^E$ so that $\sum_{e \in P_i} x_e v_e \in 5K \; \forall i \in V$, where $P_i \subseteq E$ are the edges on the path from the root to i.*



Signs for a labelled tree

Theorem

Let $\mathcal{T} = (V, E)$ be a **rooted tree** *where each edge $e \in E$ is assigned a vector $v_e \in \mathbb{R}^n$ with $\|v_e\|_2 \leq 1$. Let $K \subseteq \mathbb{R}^n$ be a* **convex body** *with $\gamma_n(K) \geq 1 - \frac{1}{2|E|}$. Then there are signs $x \in \{-1, 1\}^E$ so that $\sum_{e \in P_i} x_e v_e \in 5K \; \forall i \in V$, where $P_i \subseteq E$ are the edges on the path from the root to i.*

Not yet what we need:

- This only finds one sign vector, not a distribution!
- K needs to be huge!!

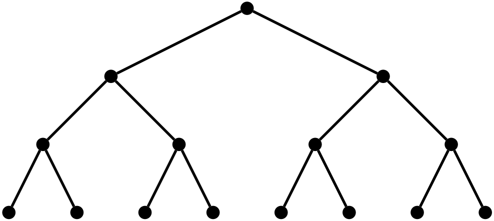
Proof similar to:

- [Banaszczyk 2012] $\mathcal{T}$ is path
- [Bansal, Jiang, Meka, Singla, Sinha 2022] vertices are labelled rather than edges

Signs for a labelled tree

Proof.

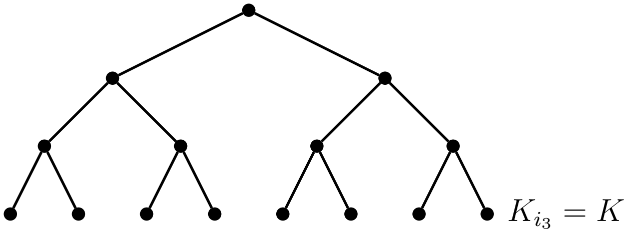
- Recursively define $K_i \subseteq \mathbb{R}^n$: for leaf $i \in V$, define $K_i := K$, for non-leaf $K_i := \bigcap_{j \in \text{children}(i)} (K_j * \frac{1}{5} v_{\{i,j\}}) \cap K$



Signs for a labelled tree

Proof.

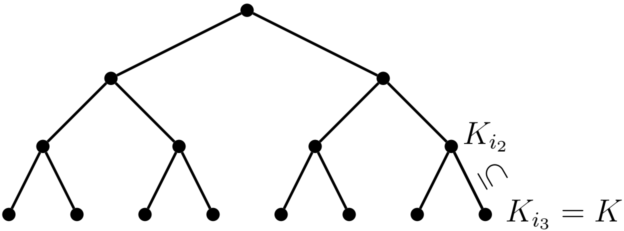
- Recursively define $K_i \subseteq \mathbb{R}^n$: for leaf $i \in V$, define $K_i := K$, for non-leaf $K_i := \bigcap_{j \in \text{children}(i)} (K_j * \frac{1}{5} v_{\{i,j\}}) \cap K$



Signs for a labelled tree

Proof.

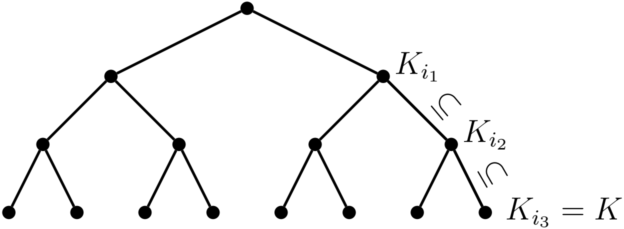
- Recursively define $K_i \subseteq \mathbb{R}^n$: for leaf $i \in V$, define $K_i := K$, for non-leaf $K_i := \bigcap_{j \in \text{children}(i)} (K_j * \frac{1}{5} v_{\{i,j\}}) \cap K$



Signs for a labelled tree

Proof.

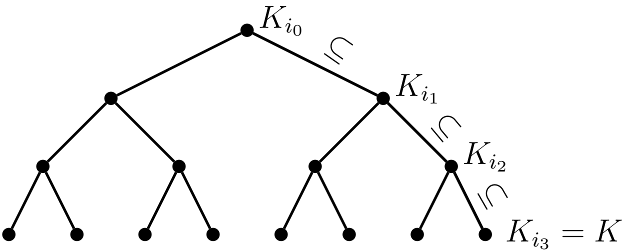
- Recursively define $K_i \subseteq \mathbb{R}^n$: for leaf $i \in V$, define $K_i := K$, for non-leaf $K_i := \bigcap_{j \in \text{children}(i)} (K_j * \frac{1}{5} v_{\{i,j\}}) \cap K$



Signs for a labelled tree

Proof.

- Recursively define $K_i \subseteq \mathbb{R}^n$: for leaf $i \in V$, define $K_i := K$, for non-leaf $K_i := \bigcap_{j \in \text{children}(i)} (K_j * \frac{1}{5} v_{\{i,j\}}) \cap K$

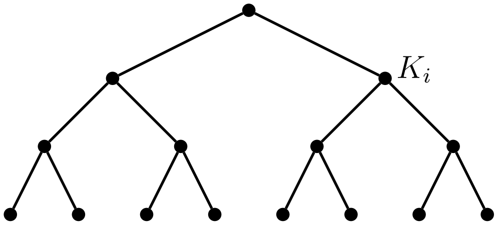


Signs for a labelled tree

Proof.

- Recursively define $K_i \subseteq \mathbb{R}^n$: for leaf $i \in V$, define $K_i := K$, for non-leaf $K_i := \bigcap_{j \in \text{children}(i)} (K_j * \frac{1}{5} v_{\{i,j\}}) \cap K$

Claim. $\gamma_n(K_i) \geq 1 - \frac{|\text{descendents of } i|}{2|E|}$.



Signs for a labelled tree

Proof.

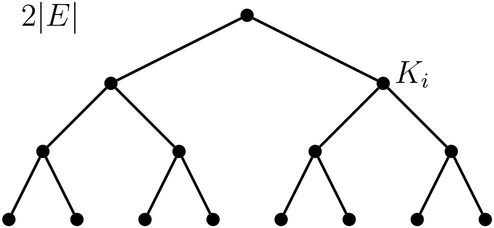
- Recursively define $K_i \subseteq \mathbb{R}^n$: for leaf $i \in V$, define $K_i := K$, for non-leaf $K_i := \bigcap_{j \in \text{children}(i)} (K_j * \frac{1}{5} v_{\{i,j\}}) \cap K$

Claim. $\gamma_n(K_i) \geq 1 - \frac{|\text{descendents of } i|}{2|E|}$.

Proof.

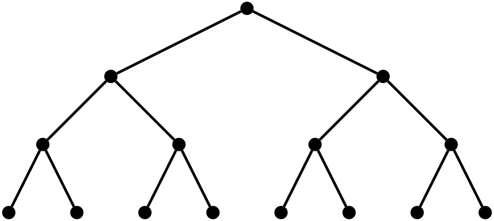
$$\gamma_n(K_i) \overset{\substack{\text{union} \\ \text{bound}}}{\geq} 1 - \sum_{j \text{ desc. of } i} \underbrace{\gamma_n\left(\mathbb{R}^n \setminus \left(K_j * \frac{v_{\{i,j\}}}{5}\right)\right)}_{\leq \frac{|\text{desc. of } j|}{2|E|} \text{ by ind.}} - \underbrace{\gamma_n(\mathbb{R}^n \setminus K)}_{\leq \frac{1}{2|E|}}$$

$$\geq 1 - \frac{|\text{desc. of } i|}{2|E|}$$



Signs for a labelled tree (2)

Claim. $\exists x \in \{-1,1\}^E$: $\sum_{e \in P_i} x_e v_e \in 5K_i$ for all $i \in V$.

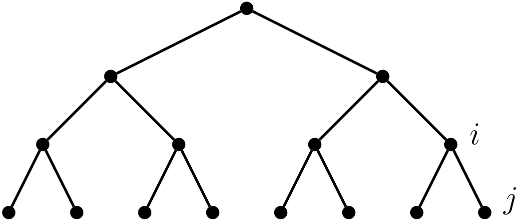


Signs for a labelled tree (2)

Claim. $\exists x \in \{-1,1\}^E$: $\sum_{e \in P_i} x_e v_e \in 5K_i$ for all $i \in V$.

Proof.

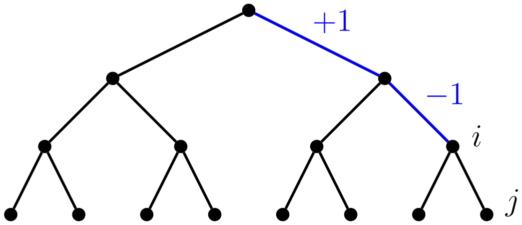
- Let $\{i,j\} \in E$.



Signs for a labelled tree (2)

Claim. $\exists x \in \{-1,1\}^E$: $\sum_{e \in P_i} x_e v_e \in 5K_i$ for all $i \in V$.
Proof.

- Let $\{i,j\} \in E$. Suppose by induction, we have already signs so that $\sum_{e \in P_i} x_e v_e \in 5K_i$.

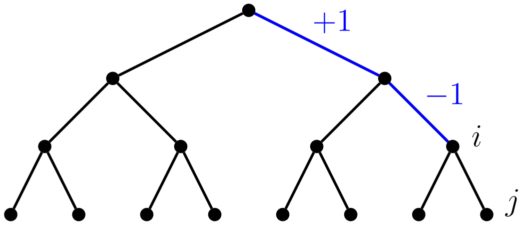


Signs for a labelled tree (2)

Claim. $\exists x \in \{-1,1\}^E$: $\sum_{e \in P_i} x_e v_e \in 5K_i$ for all $i \in V$.
Proof.

- Let $\{i,j\} \in E$. Suppose by induction, we have already signs so that $\sum_{e \in P_i} x_e v_e \in 5K_i$.
- Then

$$\sum_{e \in P_i} x_e v_e \in 5K_i \subseteq 5(K_j * \frac{1}{5} v_{ij}) \subseteq 5((K_j + v_{ij}) \cup (K_j - v_{ij}))$$



Signs for a labelled tree (2)

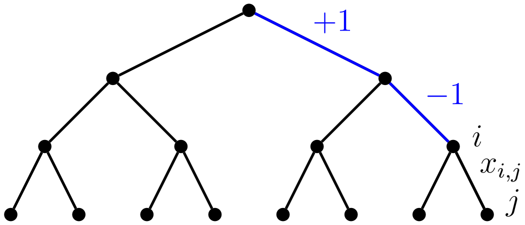
Claim. $\exists x \in \{-1,1\}^E$: $\sum_{e \in P_i} x_e v_e \in 5K_i$ for all $i \in V$.

Proof.

- Let $\{i,j\} \in E$. Suppose by induction, we have already signs so that $\sum_{e \in P_i} x_e v_e \in 5K_i$.
- Then

$$\sum_{e \in P_i} x_e v_e \in 5K_i \subseteq 5(K_j * \frac{1}{5} v_{ij}) \subseteq 5((K_j + v_{ij}) \cup (K_j - v_{ij}))$$

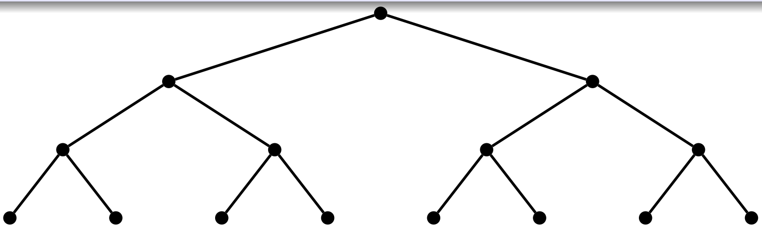
- So $\exists x_{ij} \in \{-1,1\}$ so that $\sum_{e \in P_j} x_e v_e \in 5K_j$



Sign distribution for a tree

Theorem

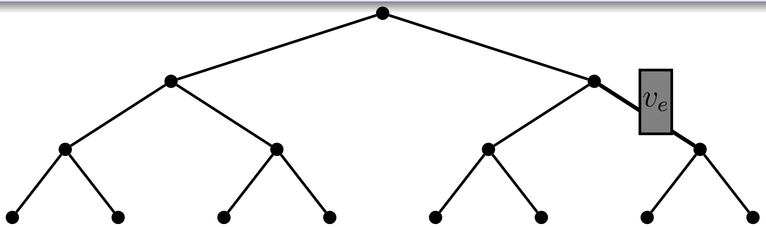
Given any **rooted tree** $\mathcal{T} = (V, E)$, *edges labeled with* $v_e \in B_2^n$. *There is a distribution* $\mathcal{D}$ *over signs* $x \in \{-1, 1\}^E$ *so that* $\sum_{e \in P_i} x_e v_e$ *is* $O(1)$**-subgaussian** *for each* **root-leaf path** P_i.



Sign distribution for a tree

Theorem

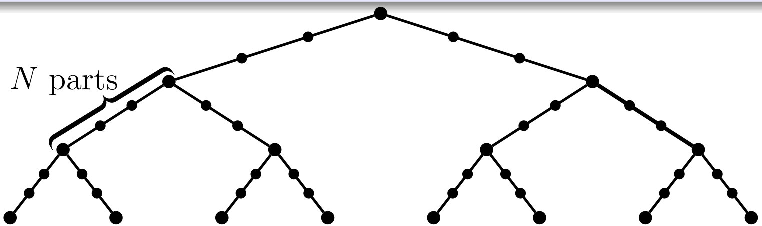
Given any **rooted tree** $\mathcal{T} = (V, E)$*, edges labeled with* $v_e \in B_2^n$*. There is a distribution* $\mathcal{D}$ *over signs* $x \in \{-1, 1\}^E$ *so that* $\sum_{e \in P_i} x_e v_e$ *is* $O(1)$**-subgaussian** *for each* **root-leaf path** P_i*.*



Sign distribution for a tree

Theorem

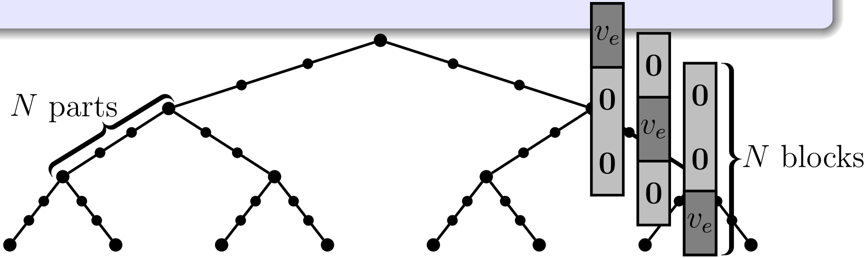
Given any **rooted tree** $\mathcal{T} = (V, E)$, *edges labeled with* $v_e \in B_2^n$. *There is a distribution* $\mathcal{D}$ *over signs* $x \in \{-1, 1\}^E$ *so that* $\sum_{e \in P_i} x_e v_e$ *is* $O(1)$**-subgaussian** *for each* **root-leaf path** P_i.



Sign distribution for a tree

Theorem

Given any **rooted tree** $\mathcal{T} = (V, E)$, *edges labeled with* $v_e \in B_2^n$. *There is a distribution* $\mathcal{D}$ *over signs* $x \in \{-1, 1\}^E$ *so that* $\sum_{e \in P_i} x_e v_e$ *is* $O(1)$**-subgaussian** *for each* **root-leaf path** P_i.

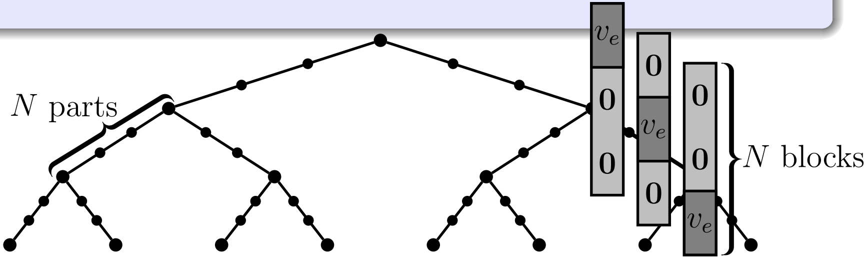


- Subdivide each edge into N edges. Replace $v_e \in \mathbb{R}^n$ by N copies in dim nN.

Sign distribution for a tree

Theorem

Given any **rooted tree** $\mathcal{T} = (V, E)$*, edges labeled with* $v_e \in B_2^n$*. There is a distribution* $\mathcal{D}$ *over signs* $x \in \{-1, 1\}^E$ *so that* $\sum_{e \in P_i} x_e v_e$ *is* $O(1)$**-subgaussian** *for each* **root-leaf path** P_i*.*



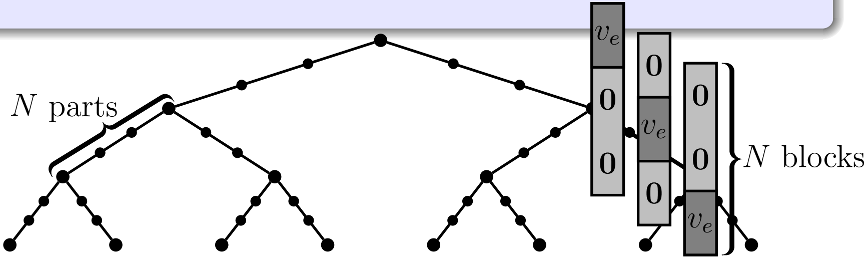
- Subdivide each edge into N edges. Replace $v_e \in \mathbb{R}^n$ by N copies in dim nN. Define the convex body

$$K := \left\{ y \in \mathbb{R}^{Nn} \mid \|Y\|_{\psi_2,\infty} \leq O(1) \text{ where } Y \sim \{y^{(1)}, \ldots, y^{(N)}\} \right\}$$

Sign distribution for a tree

Theorem

Given any **rooted tree** $\mathcal{T} = (V, E)$*, edges labeled with* $v_e \in B_2^n$*. There is a distribution* $\mathcal{D}$ *over signs* $x \in \{-1,1\}^E$ *so that* $\sum_{e \in P_i} x_e v_e$ *is* **$O(1)$-subgaussian** *for each* **root-leaf path** P_i*.*



- Subdivide each edge into N edges. Replace $v_e \in \mathbb{R}^n$ by N copies in dim nN. Define the convex body

$$K := \left\{ y \in \mathbb{R}^{Nn} \mid \|Y\|_{\psi_2,\infty} \leq O(1) \text{ where } Y \sim \{y^{(1)}, \ldots, y^{(N)}\} \right\}$$

- Suffices to prove $\gamma_{Nn}(K) \geq 1 - \frac{1}{poly(n,N)}$

Sign distribution for a tree (2)

- By union bound over $2^{O(n)}$ directions with $N := 2^{\Theta_C(n)}$, it suffices to show:

Claim. For C and N large enough,

$$\Pr_{g_1,\ldots,g_N \sim N(0,1)}\left[\mathop{\mathbb{E}}_{\ell\sim[N]}\left[\exp\left(\frac{g_\ell^2}{C^2}\right)\right] \le O(1)\right] \ge 1 - \frac{1}{N^{100}}$$

Sign distribution for a tree (2)

- By union bound over $2^{O(n)}$ directions with $N := 2^{\Theta_C(n)}$, it suffices to show:

Claim. For C and N large enough,

$$\Pr_{g_1,\ldots,g_N \sim N(0,1)} \left[\sum_{\ell \in [N]} \exp\left(\frac{g_\ell^2}{C^2}\right) \leq O(N) \right] \geq 1 - \frac{1}{N^{100}}$$

Sign distribution for a tree (2)

- By union bound over $2^{O(n)}$ directions with $N := 2^{\Theta_C(n)}$, it suffices to show:

Claim. For C and N large enough,

$$\Pr_{g_1,\dots,g_N \sim N(0,1)} \left[\sum_{\ell \in [N]} \exp\left(\frac{g_\ell^2}{C^2}\right) \leq O(N) \right] \geq 1 - \frac{1}{N^{100}}$$

- **Problem:** For $g \sim N(0,1)$, the random variables of the form $X := \exp(\frac{g^2}{C^2})$, have **heavy tails**, i.e. $\mathbb{E}[e^{\lambda X}] = \infty$ for all $\lambda > 0$. Chernov/Bernstein does not apply!

Sign distribution for a tree (2)

- By union bound over $2^{O(n)}$ directions with $N := 2^{\Theta_C(n)}$, it suffices to show:

Claim. For C and N large enough,

$$\Pr_{g_1,\ldots,g_N \sim N(0,1)} \left[\sum_{\ell \in [N]} \exp\left(\frac{g_\ell^2}{C^2}\right) \leq O(N) \right] \geq 1 - \frac{1}{N^{100}}$$

- **Problem:** For $g \sim N(0,1)$, the random variables of the form $X := \exp(\frac{g^2}{C^2})$, have **heavy tails**, i.e. $\mathbb{E}[e^{\lambda X}] = \infty$ for all $\lambda > 0$. Chernov/Bernstein does not apply!
- But **polynomial moments are bounded**: $\mathbb{E}[X^p] = \exp(\frac{p}{C^2} \cdot g^2) < \infty$ if $p < C^2/2$

Sign distribution for a tree (3)

Lemma (Variant of Rosenthal's Inequality)

Let $p \geq 2$, $c > 0$ *and let* $X_1, \ldots, X_N$ *independent centered RVs with* $\mathbb{E}[|X_i|^p] \leq O_p(1)$. *Then*

$$\Pr\left[\left|\sum_{\ell=1}^{N} X_\ell\right| \geq cN\right] \leq \frac{O_{c,p}(1)}{N^{p/2}}$$

- **Note:** Gives **polynomial** concentration, not exponential!
- Concludes the proof. □

Open problems

Polynomial time algorithm

Is there a **polynomial time online algorithm** that given $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_t\|_2 \leq 1$ one-by-one by an **oblivious adversary** keeps all signed prefix sums $O(1)$**-subgaussian**?

- We know $O(\sqrt{\log(nT)})$ [Alweiss, Liu, Sawhney 2021]

Open problems

Polynomial time algorithm

Is there a **polynomial time online algorithm** that given $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_t\|_2 \le 1$ one-by-one by an **oblivious adversary** keeps all signed prefix sums $O(1)$**-subgaussian**?

- We know $O(\sqrt{\log(nT)})$ [Alweiss, Liu, Sawhney 2021]

Oblivious Spencer

Given $v_1, \ldots, v_n \in [-1, 1]^n$ one-by-one (**obliviously**), can we find online signs $x_1, \ldots, x_n \in \{\pm 1\}$ so that $\|\sum_{i=1}^n x_i v_i\|_\infty \le O(\sqrt{n})$ w.h.p.?

Open problems

Polynomial time algorithm

Is there a **polynomial time online algorithm** that given $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_t\|_2 \leq 1$ one-by-one by an **oblivious adversary** keeps all signed prefix sums $O(1)$**-subgaussian**?

- We know $O(\sqrt{\log(nT)})$ [Alweiss, Liu, Sawhney 2021]

Oblivious Spencer

Given $v_1, \ldots, v_n \in [-1, 1]^n$ one-by-one (**obliviously**), can we find online signs $x_1, \ldots, x_n \in \{\pm 1\}$ so that $\|\sum_{i=1}^n x_i v_i\|_\infty \leq O(\sqrt{n})$ w.h.p.?

- True for offline version [Spencer 85]
- True, if $v_i \sim \{-1, 1\}^n$ at random [Bansal, Spencer '20]

Open problem (2)

Note: Prefix-discrepancy possibly easier than (oblivious) online

Prefix Beck-Fiala

Given $v_1, \dots, v_T \in \mathbb{R}^n$ with $\|v_i\|_1 \leq 1$, are there signs $x \in \{-1,1\}^T$ so that

$$\Big\| \sum_{i=1}^{t} x_i v_i \Big\|_\infty \leq O(1) \quad \forall t \in [T]$$

Open problem (2)

Note: Prefix-discrepancy possibly easier than (oblivious) online

Prefix Beck-Fiala

Given $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_1 \leq 1$, are there signs $x \in \{-1,1\}^T$ so that

$$\Big\| \sum_{i=1}^{t} x_i v_i \Big\|_\infty \leq O(1) \quad \forall t \in [T]$$

- Known is $O(\sqrt{\log(nT)})$-bound [Banaszczyk 2012]
- True, if only for **total sum** ($t = T$) [Beck-Fiala 81]
- False, if **online oblivious** (even for $n = 2$, online oblivious needs $\mathbb{E}[\max_{t \in [T]} \| \sum_{i=1}^{t} x_i v_i \|_\infty] \gtrsim \sqrt{\log(T)}$.

Open problem (2)

Note: Prefix-discrepancy possibly easier than (oblivious) online

Prefix Beck-Fiala

Given $v_1, \dots, v_T \in \mathbb{R}^n$ with $\|v_i\|_1 \leq 1$, are there signs $x \in \{-1,1\}^T$ so that

$$\Big\| \sum_{i=1}^{t} x_i v_i \Big\|_\infty \leq O(1) \quad \forall t \in [T]$$

- Known is $O(\sqrt{\log(nT)})$-bound [Banaszczyk 2012]
- True, if only for **total sum** $(t = T)$ [Beck-Fiala 81]
- False, if **online oblivious** (even for $n = 2$, online oblivious needs $\mathbb{E}[\max_{t \in [T]} \| \sum_{i=1}^{t} x_i v_i \|_\infty] \gtrsim \sqrt{\log(T)}$.

Thanks for your attention