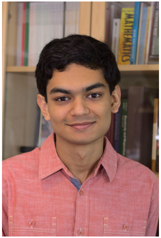
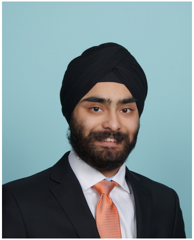
A Gaussian Fixed Point Random Walk

Yang P. Liu (Stanford University)
Joint with Ashwin Sah (MIT), Mehtaab Sawhney (MIT)

February 4, 2022

Online Vector Balancing

- Vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ arrive one at a time.
- Assign signs $\varepsilon_1, \ldots, \varepsilon_T \in \{-1, 1\}$ to maintain small *discrepancy*: keep the quantity $\max_{t \in [T]} \left\| \sum_{i=1}^{t} \varepsilon_i v_i \right\|_\infty$ small.

Online Vector Balancing

- Vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ arrive one at a time.
- Assign signs $\varepsilon_1, \ldots, \varepsilon_T \in \{-1, 1\}$ to maintain small *discrepancy*: keep the quantity $\max_{t\in[T]} \left\|\sum_{i=1}^{t} \varepsilon_i v_i\right\|_\infty$ small.
- Numerous applications including randomized controlled trials (Harshaw, Sävje, Spielman, Zhang (2019)) and online envy minimization algorithms (Jiang, Kulkarni, Singla (2019)).
- In general, algorithmic discrepancy theory has been an active field (Bansal (2010), Lovett, Meka (2012)).

Example Online Vector Balancing

	Vector	Sign	Partial Sum
v_1			
v_2			
v_3			
v_4			
v_5			

Example Online Vector Balancing

	Vector	Sign	Partial Sum
v_1	(-1, 1, 1, -1)		
v_2			
v_3			
v_4			
v_5			

Example Online Vector Balancing

	Vector	Sign	Partial Sum
v_1	(-1, 1, 1, -1)	1	(-1, 1, 1, -1)
v_2			
v_3			
v_4			
v_5			

Example Online Vector Balancing

	Vector	Sign	Partial Sum
v_1	(-1, 1, 1, -1)	1	(-1, 1, 1, -1)
v_2	(-1, -1, -1, 1)		
v_3			
v_4			
v_5			

Example Online Vector Balancing

	Vector	Sign	Partial Sum
v_1	(-1, 1, 1, -1)	1	(-1, 1, 1, -1)
v_2	(-1, -1, -1, 1)	1	(-2, 0, 0, 0)
v_3			
v_4			
v_5			

Example Online Vector Balancing

	Vector	Sign	Partial Sum
v_1	(-1, 1, 1, -1)	1	(-1, 1, 1, -1)
v_2	(-1, -1, -1, 1)	1	(-2, 0, 0, 0)
v_3	(-1, 1, 1, 1)		
v_4			
v_5			

Example Online Vector Balancing

	Vector	Sign	Partial Sum
v_1	(-1, 1, 1, -1)	1	(-1, 1, 1, -1)
v_2	(-1, -1, -1, 1)	1	(-2, 0, 0, 0)
v_3	(-1, 1, 1, 1)	-1	(-1, -1, -1, -1)
v_4			
v_5			

Example Online Vector Balancing

	Vector	Sign	Partial Sum
v_1	(-1, 1, 1, -1)	1	(-1, 1, 1, -1)
v_2	(-1, -1, -1, 1)	1	(-2, 0, 0, 0)
v_3	(-1, 1, 1, 1)	-1	(-1, -1, -1, -1)
v_4	(-1, 1, -1, -1)		
v_5			

Example Online Vector Balancing

	Vector	Sign	Partial Sum
v_1	(-1, 1, 1, -1)	1	(-1, 1, 1, -1)
v_2	(-1, -1, -1, 1)	1	(-2, 0, 0, 0)
v_3	(-1, 1, 1, 1)	-1	(-1, -1, -1, -1)
v_4	(-1, 1, -1, -1)	-1	(0, -2, 0, 0)
v_5			

Example Online Vector Balancing

	Vector	Sign	Partial Sum
v_1	(-1, 1, 1, -1)	1	(-1, 1, 1, -1)
v_2	(-1, -1, -1, 1)	1	(-2, 0, 0, 0)
v_3	(-1, 1, 1, 1)	-1	(-1, -1, -1, -1)
v_4	(-1, 1, -1, -1)	-1	(0, -2, 0, 0)
v_5	(-1, -1, 1, 1)		

Example Online Vector Balancing

	Vector	Sign	Partial Sum
v_1	(-1, 1, 1, -1)	1	(-1, 1, 1, -1)
v_2	(-1, -1, -1, 1)	1	(-2, 0, 0, 0)
v_3	(-1, 1, 1, 1)	-1	(-1, -1, -1, -1)
v_4	(-1, 1, -1, -1)	-1	(0, -2, 0, 0)
v_5	(-1, -1, 1, 1)	-1	(1, -1, -1, -1)

Komlós Setting

- Focus on the case where $\|v_i\|_2 \leq 1$.
- Oblivious adversary: v_i are fixed beforehand and do not change based on the randomness of the algorithm.

Komlós Setting

- Focus on the case where $\|v_i\|_2 \leq 1$.
- Oblivious adversary: v_i are fixed beforehand and do not change based on the randomness of the algorithm.
- Generalizes the offline setting where all the v_i are revealed at the beginning.
- $\Omega(\sqrt{T})$ is a lower bound against adaptive adversaries: the adversary picks the next v_i to be orthogonal to previous partial sum.

Previous Results

Komlós Conjecture

Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \le 1$ for all $i \in [T]$ there are signs $\varepsilon_1, \ldots, \varepsilon_T \in \{-1, 1\}$ such that $\left\|\sum_{i=1}^T \varepsilon_i v_i\right\|_\infty \le O(1)$.

Previous Results

Komlós Conjecture

Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ for all $i \in [T]$ there are signs $\varepsilon_1, \ldots, \varepsilon_T \in \{-1, 1\}$ such that $\left\|\sum_{i=1}^{T} \varepsilon_i v_i\right\|_\infty \leq O(1)$.

Komlós Conjecture for Prefixes

Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \leq 1$ for all $i \in [T]$ there are signs $\varepsilon_1, \ldots, \varepsilon_T \in \{-1, 1\}$ such that $\left\|\sum_{i=1}^{t} \varepsilon_i v_i\right\|_\infty \leq O(1)$ for all $t \in [T]$.

Previous Results

Komlós Conjecture

Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \le 1$ for all $i \in [T]$ there are signs $\varepsilon_1, \ldots, \varepsilon_T \in \{-1, 1\}$ such that $\left\|\sum_{i=1}^T \varepsilon_i v_i\right\|_\infty \le O(1)$.

Komlós Conjecture for Prefixes

Given vectors $v_1, \ldots, v_T \in \mathbb{R}^n$ with $\|v_i\|_2 \le 1$ for all $i \in [T]$ there are signs $\varepsilon_1, \ldots, \varepsilon_T \in \{-1, 1\}$ such that $\left\|\sum_{i=1}^t \varepsilon_i v_i\right\|_\infty \le O(1)$ for all $t \in [T]$.

- Best known bound of $O(\sqrt{\log T})$ for prefixes (Banaszczyk (1998)).
- Polynomial time algorithm achieving $O(\sqrt{\log T})$ for total sum (Bansal, Dadush, Garg (2016), Bansal, Dadush, Garg, Lovett (2018)).
- Online bound of $O(\log T)$ for prefixes (Alweiss, L., Sawhney (2021)).

Our Results

Theorem (Partial Colorings)

In the Komlós setting there is an online algorithm against oblivious adversaries that selects signs $\varepsilon_i \in \{-1, 0, 1\}$ with at most 4% of signs chosen as 0 that achieves discrepancy $O(\sqrt{\log T})$ with high probability.

Theorem ($\{-1, 1, 2\}$ Colorings)

In the Komlós setting there is an online algorithm against oblivious adversaries that selects signs $\varepsilon_i \in \{-1, 1, 2\}$ that achieves discrepancy $O(\sqrt{\log T})$ with high probability.

Our Results

Theorem (Partial Colorings)

In the Komlós setting there is an online algorithm against oblivious adversaries that selects signs $\varepsilon_i \in \{-1, 0, 1\}$ with at most 4% of signs chosen as 0 that achieves discrepancy $O(\sqrt{\log T})$ with high probability.

Theorem ($\{-1, 1, 2\}$ Colorings)

In the Komlós setting there is an online algorithm against oblivious adversaries that selects signs $\varepsilon_i \in \{-1, 1, 2\}$ that achieves discrepancy $O(\sqrt{\log T})$ with high probability.

- Recovers $O(\log T)$ bound for online discrepancy with $\{-1, 1\}$ signs.
- Falls short of matching Banaszczyk's bound online because of the additional 0 (respectively 2) signs allowed.
- Optimistic that $O(\sqrt{\log T})$ online (and prefix) discrepancy with only $\{-1, 1\}$ colors is achievable.

Main Algorithm

Theorem (Gaussian Fixed Point Walk)

There is a Markov chain on $\mathbb{R}$ with steps in $\{-1, 0, 1\}$ (or $\{-1, 1, 2\}$), and at most 4% probability of picking 0 every step, whose stationary distribution is $\mathcal{N}(0, 1)$.

Main Algorithm

Theorem (Gaussian Fixed Point Walk)

There is a Markov chain on $\mathbb{R}$ with steps in $\{-1, 0, 1\}$ (or $\{-1, 1, 2\}$), and at most 4% probability of picking 0 every step, whose stationary distribution is $\mathcal{N}(0, 1)$.

Online Discrepancy Algorithm

For simplicity consider case where $\|v_i\|_2 = 1$ for all i.
Initialize a starting vector $\mathbb{R}^n \ni w_0 \sim \mathcal{N}(0, I)$. When v_i arrives:

Main Algorithm

Theorem (Gaussian Fixed Point Walk)

There is a Markov chain on $\mathbb{R}$ with steps in $\{-1, 0, 1\}$ (or $\{-1, 1, 2\}$), and at most 4% probability of picking 0 every step, whose stationary distribution is $\mathcal{N}(0, 1)$.

Online Discrepancy Algorithm

For simplicity consider case where $\|v_i\|_2 = 1$ for all i.
Initialize a starting vector $\mathbb{R}^n \ni w_0 \sim \mathcal{N}(0, I)$. When v_i arrives:

- Decompose $\mathcal{N}(0, I)$ in n-dimensions into one-dimensional Gaussians in the direction of v_i.
- Let ε_i be the step size of the Gaussian Fixed Point Walk in the theorem above.
- $w_i \leftarrow w_{i-1} + \varepsilon_i v_i$.

Why the Algorithm Works

Online Discrepancy Algorithm

For simplicity consider case where $\|v_i\|_2 = 1$ for all i.
Initialize a starting vector $\mathbb{R}^n \ni w_0 \sim \mathcal{N}(0, I)$. When v_i arrives:

- Decompose $\mathcal{N}(0, I)$ in n-dimensions into one-dimensional Gaussians in the direction of v_i.
- Let ε_i be the step size of the Gaussian Fixed Point Walk in the theorem above.
- $w_i \leftarrow w_{i-1} + \varepsilon_i v_i$.

Why the Algorithm Works

Online Discrepancy Algorithm

For simplicity consider case where $\|v_i\|_2 = 1$ for all i.
Initialize a starting vector $\mathbb{R}^n \ni w_0 \sim \mathcal{N}(0, I)$. When v_i arrives:

- Decompose $\mathcal{N}(0, I)$ in n-dimensions into one-dimensional Gaussians in the direction of v_i.
- Let ε_i be the step size of the Gaussian Fixed Point Walk in the theorem above.
- $w_i \leftarrow w_{i-1} + \varepsilon_i v_i$.

- By induction, the distribution of w_i is $\mathcal{N}(0, I)$ every step.
- Distribution of $\sum_{i=1}^{t} \varepsilon_i v_i$ is the difference of two (coupled) Gaussians with distribution $\mathcal{N}(0, I)$.
- Hence $\left\|\sum_{i=1}^{t} \varepsilon_i v_i\right\|_\infty \leq O(\sqrt{\log T})$ with high probability.

Building the Gaussian Fixed Point Walk

- **Goal**: Build a Markov chain on $\mathbb{R}$ whose stationary distribution is $\mathcal{N}(0,1)$, and whose steps are $\{-1,0,1\}$ (or $\{-1,1,2\}$).

Building the Gaussian Fixed Point Walk

- **Goal**: Build a Markov chain on $\mathbb{R}$ whose stationary distribution is $\mathcal{N}(0,1)$, and whose steps are $\{-1,0,1\}$ (or $\{-1,1,2\}$).
- **Observation**: Treat $f+\mathbb{Z}$ separately for each $f \in [-1/2, 1/2)$.
- We focus on $f=0$: build Markov chain on $\mathbb{Z}$ with stationary proportional to $\exp(-x^2/2)$ for $x \in \mathbb{Z}$.

Parity constraint

Lemma (Parity constraint)

No Markov chain on $\mathbb{Z}$ *with steps* $\{-1, 1\}$ *with stationary distribution proportional to* $\exp(-x^2/2)$.

Proof.

If such a chain exists, total mass on even integers and odd integers is the same. But $\sum_{x \text{ even}} \exp(-x^2/2) \neq \sum_{x \text{ odd}} \exp(-x^2/2)$. □

Full Walk Construction

Define transition probabilities $m(x)$ (move) and s (stay) as

$$m(x) := \sum_{j\geq 1}(-1)^{j-1}\exp\left(\frac{-j^2+2xj}{2}\right) \quad \text{and} \quad s := \sum_{j\in\mathbb{Z}}(-1)^j\exp\left(\frac{-j^2}{2}\right).$$

Full Walk Construction

Define transition probabilities $m(x)$ (move) and s (stay) as

$$m(x) := \sum_{j\geq 1}(-1)^{j-1}\exp\left(\frac{-j^2+2xj}{2}\right) \quad \text{and} \quad s := \sum_{j\in\mathbb{Z}}(-1)^{j}\exp\left(\frac{-j^2}{2}\right).$$

Transition probabilities for $x \in \mathbb{Z}$

- For $x \geq 1$ move $+1$ with prob. $m(x)$, -1 with prob. $1 - m(x)$.
- For $x \leq -1$, move $+1$ with prob. $1 - m(-x)$, -1 with prob. $m(-x)$.
- For $x = 0$, move $+1$ with prob. $m(0)$, -1 with prob. $m(0)$, and stay with prob. s.

Full Walk Construction

Define transition probabilities $m(x)$ (move) and s (stay) as

$$m(x) := \sum_{j \geq 1} (-1)^{j-1} \exp\left(\frac{-j^2 + 2xj}{2}\right) \quad \text{and} \quad s := \sum_{j \in \mathbb{Z}} (-1)^j \exp\left(\frac{-j^2}{2}\right).$$

Transition probabilities for $x \in \mathbb{Z}$

- For $x \geq 1$ move $+1$ with prob. $m(x)$, -1 with prob. $1 - m(x)$.
- For $x \leq -1$, move $+1$ with prob. $1 - m(-x)$, -1 with prob. $m(-x)$.
- For $x = 0$, move $+1$ with prob. $m(0)$, -1 with prob. $m(0)$, and stay with prob. s.

- Unique for walks where only $x = 0$ can stay put (take steps of size 0).
- Compute $m(x), s$ via direct algebra for walks where only $x = 0$ can stay put.

Full Walk Construction

Define transition probabilities $m(x)$ (move) and s (stay) as

$$m(x) := \sum_{j \geq 1} (-1)^{j-1} \exp\left(\frac{-j^2 + 2xj}{2}\right) \quad \text{and} \quad s := \sum_{j \in \mathbb{Z}} (-1)^j \exp\left(\frac{-j^2}{2}\right).$$

Transition probabilities for $x \in \mathbb{Z}$

- For $x \geq 1$ move $+1$ with prob. $m(x)$, -1 with prob. $1 - m(x)$.
- For $x \leq -1$, move $+1$ with prob. $1 - m(-x)$, -1 with prob. $m(-x)$.
- For $x = 0$, move $+1$ with prob. $m(0)$, -1 with prob. $m(0)$, and stay with prob. s.

- Unique for walks where only $x = 0$ can stay put (take steps of size 0).
- Compute $m(x), s$ via direct algebra for walks where only $x = 0$ can stay put.
- Need to check: walk is well defined, eg. check $s \in [0, 1]$.

Bounding s, Triple Product Formula

$$s := \sum_{j \in \mathbb{Z}} (-1)^j \exp\left(\frac{-j^2}{2}\right).$$

Numerically, $s \leq .0361$, so at most 3.7% of signs are 0 whp.

Bounding s, Triple Product Formula

$$s := \sum_{j\in\mathbb{Z}}(-1)^j \exp\left(\frac{-j^2}{2}\right).$$

Numerically, $s \le .0361$, so at most 3.7% of signs are 0 whp.

Theorem (Jacobi Triple Product Formula)

For complex numbers $|u| < 1, v \neq 0$ we have

$$\sum_{j\in\mathbb{Z}} u^{j^2} v^{2j} = \prod_{j\geq 1}(1-u^{2j})(1+u^{2j-1}v^2)(1+u^{2j-1}v^{-2}).$$

Bounding s, Triple Product Formula

$$s := \sum_{j \in \mathbb{Z}} (-1)^j \exp\left(\frac{-j^2}{2}\right).$$

Numerically, $s \leq .0361$, so at most 3.7% of signs are 0 whp.

Theorem (Jacobi Triple Product Formula)

For complex numbers $|u| < 1, v \neq 0$ we have

$$\sum_{j \in \mathbb{Z}} u^{j^2} v^{2j} = \prod_{j \geq 1} (1 - u^{2j})(1 + u^{2j-1} v^2)(1 + u^{2j-1} v^{-2}).$$

Bounding s

Take $u = \exp(-1/2), v = \sqrt{-1}$ to get

$$s = \sum_{j \in \mathbb{Z}} (-1)^j \exp\left(\frac{-j^2}{2}\right) = \prod_{j \geq 1} (1 - \exp(-j))(1 - \exp(-(2j-1)/2))^2.$$

Conclusion

Relations to Other Works

- (Alweiss, L., Sawhney (2021)) Algorithm/analysis based on proving that distribution of partial sum w_i is *spread* by a Gaussian $\mathcal{N}(0, O(\log T)I)$.
  - Gaussian Fixed Point Walk is a "limit" of this.
- (Chewi, Gerber, Rigollet, Turner (2021)) Builds walk in $\mathbb{R}^2$ with steps of length 1 with stationary distribution $\mathcal{N}(0, I)$. Applications to discrepancy for two-dimensional vector colorings.

Conclusion

Relations to Other Works

- (Alweiss, L., Sawhney (2021)) Algorithm/analysis based on proving that distribution of partial sum w_i is *spread* by a Gaussian $\mathcal{N}(0, O(\log T)I)$.
  - Gaussian Fixed Point Walk is a "limit" of this.
- (Chewi, Gerber, Rigollet, Turner (2021)) Builds walk in $\mathbb{R}^2$ with steps of length 1 with stationary distribution $\mathcal{N}(0, I)$. Applications to discrepancy for two-dimensional vector colorings.

Future Directions

- $O(\sqrt{\log T})$ discrepancy bound with signs $\{-1, 1\}$?
- Easier: polynomial time $O(\sqrt{\log T})$ discrepancy for all prefixes?
- Other applications of Gaussian Fixed Point Walk?

The End

- Paper available at `https://arxiv.org/pdf/2104.07009.pdf`
- Email: `yangpliu@stanford.edu`
- Homepage: `yangpliu.github.io`