A Gaussian Fixed Point Random Walk

Yang P. Liu
Department of Mathematics, Stanford University, Stanford, CA, USA

Ashwin Sah
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA

Mehtaab Sawhney
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA, USA

Abstract

In this note, we design a discrete random walk on the real line which takes steps $0, \pm 1$ (and one with steps in $\{\pm 1, 2\}$) where at least 96% of the signs are ± 1 in expectation, and which has $\mathcal{N}(0,1)$ as a stationary distribution. As an immediate corollary, we obtain an online version of Banaszczyk's discrepancy result for partial colorings and $\pm 1, 2$ signings. Additionally, we recover linear time algorithms for logarithmic bounds for the Komlós conjecture in an oblivious online setting.

2012 ACM Subject Classification Mathematics of computing → Probabilistic algorithms

Keywords and phrases Discrepancy, Partial Coloring

Digital Object Identifier 10.4230/LIPIcs.ITCS.2022.101

Related Version *Full Version*: https://arxiv.org/abs/2104.07009

Funding *Yang P. Liu*: Yang P. Liu was supported by the Department of Defense (DoD) through the National Defense Science and Engineering Graduate Fellowship (NDSEG) Program.
Ashwin Sah: Sah was supported by NSF Graduate Research Fellowship Program DGE-1745302.
Mehtaab Sawhney: Sawhney was supported by NSF Graduate Research Fellowship Program DGE-1745302.

Acknowledgements The authors thank Ryan Alweiss, Arun Jambulapati, Allen Liu, and Mark Sellke for earlier discussions on this topic. Furthermore we thank Ghaith Hiary for discussions regarding computing theta series. Finally we thank Nikhil Bansal for comments on the manuscript.

1 Introduction

In the (oblivious) online vector discrepancy problem an adversary fixes vectors $\{v_i\}_{i\in[t]}$ in advance and the objective is to assign signs $\epsilon_i \in \{-1, 1\}$ based only on vectors $v_1, \ldots, v_i$ to maintain that $\|\sum_{i\le t'} \epsilon_i v_i\|_\infty$ is small at all times $t' \in [t]$. Vector balancing includes a number of different problems in discrepancy theory including Spencer's [17] work on set discrepancy. Spencer's "six standard deviations suffice" result states that given vectors $v_1, \ldots, v_n \in \{0,1\}^n$ there exists a ± 1-signing such that $\|\sum_{i\le n} \epsilon_i v_i\|_\infty \le 6\sqrt{n}$. Conjecturally, however, the restriction to $\{0,1\}^n$ vectors can be relaxed to a norm condition. In particular, the Komlós conjecture states that given $v_1, \ldots, v_t$, each of at most unit length, there exists a sequence of signs $\epsilon_1, \ldots, \epsilon_t$ such that $\|\sum_{i\le t} \epsilon_i v_i\|_\infty = O(1)$. Despite substantial effort, the Komlós conjecture is still open and the best known bounds due to Banaszczyk [4] give the existence of a sequence of signs so that $\epsilon_1, \ldots, \epsilon_t$ such that $\|\sum_{i\le t} \epsilon_i v_i\|_\infty = O(\sqrt{\min(\log n, \log t)})$. However, these original proofs were by their nature non-algorithmic.

More recent research in theoretical computer science has focused on developing algorithmic versions of these results starting with the Bansal [5] and Lovett and Meka [15] polynomial-time algorithms for Spencer's [17] "six standard deviations suffice". Since then, there have been several other constructive discrepancy minimization algorithms [16, 14, 6, 8, 7, 13]. Notably for our purposes, Bansal, Dadush, and Garg [6] and Bansal, Dadush, Garg, and Lovett [7] have made the work of Banaszczyk [4] algorithmic. However in all cases these algorithms require all vectors to be known at the start and hence do not extend to the online setting.

13th Innovations in Theoretical Computer Science Conference (ITCS 2022).
Editor: Mark Braverman; Article No. 101; pp. 101:1–101:10
Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

In the online setting, significant work has been devoted to the case where v_i are drawn from a fixed (and known) distribution $\mathfrak{p}$ supported on $[-1,1]^n$. In the setting where $\mathfrak{p}$ is uniform on $[-1,1]^n$, Bansal and Spencer [11] showed one can maintain $\max_{t'\le t}\|\sum_{i\le t'}\epsilon_i v_i\|_\infty \le O(\sqrt{n}\log t)$. In the more general setting where $\mathfrak{p}$ is a general distribution supported on $[-1,1]^n$, Aru, Narayanan, Scott, and Venkatesan [3] achieved a bound of $O_n(\sqrt{\log t})$ (where the implicit dependence on n is super-exponential) and Bansal, Jiang, Meka, Singla, and Sinha [9] (building on work of Bansal, Jiang, Singla, and Sinha [10]) achieved an ℓ_∞ guarantee of $O(\sqrt{n}\log(nt)^4)$.

In this work we focus on the online setting where the only guarantee is $\|v_i\|_2 \le 1$. The only previous work in this oblivious online setting is the following result of Alweiss, Liu, and Sawhney [1].

▶ **Theorem 1** ([1, Theorem 1.1]). *For any vectors $v_1, v_2, \cdots, v_t \in \mathbb{R}^n$ with $\|v_i\|_2 \le 1$ for all $i \in [t]$, there exists an online algorithm* $\textsc{Balance}(v_1,\cdots,v_t,\delta)$ *which maintains $\|\sum_{i\le t'}\epsilon_i v_i\|_\infty = O\left(\log(nt/\delta)\right)$ for all $t' \in [t]$ with probability at least $1-\delta$.*

The proof in [1] relies on a coupling procedure which compares the distribution of $\sum_{i\le t}\epsilon_i v_i$ to a Gaussian at each stage via a stochastic domination argument and then deduces the necessary tail bounds. In this work, we recover Theorem 1 (in fact with a slightly improved dependence) as well as the following corollary.

▶ **Corollary 2.** *For any vectors $v_1, v_2, \cdots, v_t \in \mathbb{R}^n$ with $\|v_i\|_2 \le 1$ for all $i \in [t]$, there exists an online algorithm which assigns $\epsilon_i \in \{\pm 1, 2\}$ and maintains $\|\sum_{i\le t'}\epsilon_i v_i\|_\infty = O(\sqrt{\log(nt/\delta)})$ for all $t' \in [t]$ with probability at least $1-\delta$.*

This result essentially recovers the best known bound on the Komlós conjecture due to Banaszczyk [4] in an online algorithmic fashion, with the slight defect of requiring a +2-signing option. Furthermore due to the online nature of the algorithm, the algorithm will run in essentially input-sparsity time which is substantially faster than the Gram-Schmidt walk [7] which gives an algorithmic proof of the result of [4] (without the defect of requiring a +2-signing option). Additionally, our signings achieve small discrepancy for all prefixes, while the Gram-Schmidt walk only achieves small discrepancy for the entire sum.

1.1 Approach

Our results are based on the observation that there exists Markov chains on $\mathbb{R}$ with transition steps of $0, \pm 1$ (where most of the steps are ± 1) or $\pm 1, 2$ such that $\mathcal{N}(0,1)$ is a stationary distribution (as well as $\mathcal{N}(0,\sigma^2)$ for appropriate values of σ). Now, rotational invariance of normal distributions in $\mathbb{R}^n$ allows us to use these one-dimensional walks to give a random signing of an input vector v with $\|v\|_2 \le 1$ (with signs either $\{0,\pm 1\}$ or $\{\pm 1, 2\}$) that preserves the n-dimensional normal distribution exactly. Note that the argument crucially uses that we are maintaining the normal distribution exactly, as opposed to inductively maintaining a weaker subgaussian property (that was used in previous works [7, 1]).

1.2 Additional Remarks

We note that no one-dimensional Markov chains with steps ± 1 can preserve $\mathcal{N}(0,1)$ exactly, as $\sum_{n\in\mathbb{Z}}(-1)^n e^{-n^2/2} \ne 0$ and therefore $\mathcal{N}(0,1)$ fails the natural "parity constraint" that the total mass on even integers and odd integers is the same. Our arguments suggest that this is essentially the only constraint. More precisely, an extension of the arguments presented likely demonstrates that for any distinct integers $\{a,b,c\}$ where there is at least one positive and one negative integer and there is no "modular constraint", i.e. $\gcd(b-a, c-b) = 1$, then there is a Markov chain on $\mathbb{R}$ with signs $\{a, b, c\}$ and stationary distribution $\mathcal{N}(0, \sigma^2)$ as long as σ^2 is a sufficiently large constant (in terms of a, b, c).

1.3 Organization

The remainder of the paper is organized as follows. In Section 2 we construct the required Markov chain on $\mathbb{R}$ with transition steps of $0, \pm 1$ such that $\mathcal{N}(0, \sigma^2)$ is a stationary distribution. In Section 3 we extend this to a walk with transition steps of $\pm 1, 2$ as long as $\sigma \geq 1$. Finally, in Section 4 we deduce the various algorithmic consequences.

1.4 Notation

Throughout this paper let $\mathcal{N}(\mu, \sigma^2)$ denote the Gaussian random variable with mean μ and variance σ^2. Furthermore, let $\mathrm{nnz}(\{v_i\}_{i\in S})$ denote the total number of non-zero entries of the vectors $\{v_i\}_{i\in S}$.

2 $0, \pm 1$ walk

▶ **Definition 3.** *Given $\sigma > 0$ and $f \in [-1/2, 1/2]$, consider the following random walk on $f + \mathbb{Z}$. For $n \geq 1$ the state $n + f$ moves to $n + 1 + f$ with probability $p_\sigma(n + f)$ and to $n - 1 + f$ otherwise, and the state $-n + f$ moves to $-n - 1 + f$ with probability $p_\sigma(n - f)$ and to $-n + 1 + f$ otherwise. Finally, the state f moves to $1 + f$ with probability $p_\sigma(f)$, to state $-1 + f$ with probability $p_\sigma(-f)$, and stays at f with probability $r_\sigma(f)$. Here*

$$p_\sigma(x) = \sum_{j\geq 1} (-1)^{j-1} \exp\left(-\frac{j^2 + 2xj}{2\sigma^2}\right)$$
$$r_\sigma(f) = \sum_{j=-\infty}^{\infty} (-1)^j \exp\left(-\frac{j^2 + 2fj}{2\sigma^2}\right)$$

for all $x \in \mathbb{R}$.

These series clearly absolutely converge. We prove that these indeed correspond to consistent probabilities giving a walk, and additionally show that this walk preserves the discrete Gaussian distribution on $f + \mathbb{Z}$ (i.e., $\mathcal{N}(0, \sigma^2)|_{f+\mathbb{Z}}$).

▶ **Lemma 4.** *For $\sigma > 0$ and $f \in [-1/2, 1/2]$, we have that $p_\sigma(n \pm f) \in (0, 1)$ for all $n \geq 0$, that $p_\sigma(f) + r_\sigma(f) + p_\sigma(-f) = 1$, that $r_\sigma(f) \in [0, 1]$, and that furthermore*

$$r_\sigma(f) \leq e^{-\sigma^2}$$

if $\sigma \geq 1/2$. Additionally, $\mathcal{N}(0, \sigma^2)|_{f+\mathbb{Z}}$ is stationary under a step of random walk defined in Definition 3 with parameters σ, f.

Proof. First, note that $\exp(-(j^2 + 2xj)/\sigma^2)$ is strictly decreasing on integers $j \geq 1$ as long as $x \geq -1/2$. Therefore $p_\sigma(x)$ is given by an alternating series with strictly decreasing terms, and we immediately deduce

$$0 < p_\sigma(x) \leq \exp\left(-\frac{j^2 + 2xj}{2\sigma^2}\right) < 1.$$

Since $n+f, n-f \geq -1/2$ for $n \geq 0$, we see that $p_\sigma(n \pm f) \in (0,1)$, as desired. Second, note that

$$p_\sigma(-f) + r_\sigma(f) + p_\sigma(f) = 1$$

holds as trivially everything except the $j=0$ term of the sum for $r_\sigma(f)$ cancels. Third, we have for $u = \exp(-1/(2\sigma^2))$ and $v = \sqrt{-1}\exp(-f/(2\sigma^2))$ that $|u| < 1$ and $v \neq 0$, hence the Jacobi triple product identity (see [2] for a short but slick proof) yields

$$\begin{aligned} r_\sigma(f) = \sum_{j=-\infty}^{\infty} u^{j^2} v^{2j} &= \prod_{j=1}^{\infty}(1-u^{2j})(1+u^{2j-1}v^2)(1+u^{2j-1}v^{-2}) \\ &= \prod_{j=1}^{\infty}(1-e^{-j/\sigma^2})(1-e^{-(2j+2f-1)/(2\sigma^2)})(1-e^{-(2j-2f-1)/(2\sigma^2)}). \end{aligned} \tag{1}$$

Since $f \in [-1/2, 1/2]$ we see each term is nonnegative and clearly less than 1, so $r_\sigma(f) \in [0,1]$ is immediate. Therefore we indeed have a well-defined walk. In fact, we see that

$$r_\sigma(f) \leq r_\sigma(0) \leq \prod_{j=1}^{\infty}(1-e^{-j/\sigma^2})^3 \leq \prod_{j=1}^{\lfloor\sigma^2\rfloor}(1-e^{-j/\sigma^2})^3 \leq (1-e^{-1})^{3\lfloor\sigma^2\rfloor}.$$

This is at most $\exp(-\sigma^2)$ for $\sigma \geq 2$, and we can further numerically check that $r_\sigma(0) \leq \exp(-\sigma^2)$ for $\sigma \in [1/2, 2]$.

Now we show that this walk preserves $\mathcal{N}(0,\sigma^2)|_{f+\mathbb{Z}}$. Note that

$$1 - p_\sigma(x) = \sum_{j\geq 0}(-1)^j \exp\left(-\frac{j^2+2xj}{2\sigma^2}\right).$$

Therefore

$$\begin{aligned} &p_\sigma(x-1)\exp\left(-\frac{(x-1)^2}{2\sigma^2}\right) + (1-p_\sigma(x+1))\exp\left(-\frac{(x+1)^2}{2\sigma^2}\right) \\ &\quad= \sum_{j\geq 1}(-1)^{j-1}\exp\left(-\frac{(j+x-1)^2}{2\sigma^2}\right) + \sum_{j\geq 0}(-1)^j\exp\left(-\frac{(j+x+1)^2}{2\sigma^2}\right) \\ &\quad= \exp\left(-\frac{x^2}{2\sigma^2}\right). \end{aligned}$$

Since the pdf of $\mathcal{N}(0,\sigma^2)|_{f+\mathbb{Z}}$ at $n+f$ is proportional to $\exp(-(n+f)^2/(2\sigma^2))$, we find that the random walk preserves this distribution at $n+f$ for all $n \neq 0$ (applying the above equation at values $x = n \pm f$). Furthermore, the final distribution is clearly still supported on $f+\mathbb{Z}$, therefore the probability at $n=0$ is also preserved as the total sum is 1. ◀

We immediately derive a walk which preserves $\mathcal{N}(0,\sigma^2)$ by piecing together all $f \in [-1/2, 1/2)$. Let J_x^σ be the random variable defined by writing $x = n+f$, where $f \in [-1/2, 1/2)$, and then performing a step according to Definition 3.

▶ **Lemma 5.** *If $Z = \mathcal{N}(0,\sigma^2)$ then $Z + J_Z^\sigma$ is distributed as $\mathcal{N}(0,\sigma^2)$.*

3 $\pm 1, 2$ walk

We now consider a variant of the above random walk with discrete ± 1 and 2 steps. Recall the definition of $p_\sigma(x)$ and $r_\sigma(f)$ from earlier. We will require the following numerical estimate which is deferred to Appendix A.

▶ **Lemma 6.** *If $\sigma \ge 1$ and $f \in [-1/2, 1/2]$ then*

$$p_\sigma(1+f) \ge r_\sigma(f) \exp\left(\frac{2f+1}{2\sigma^2}\right).$$

▶ Remark 7. This inequality is immediate for large σ as the left uniformly tends to $1/2$ and the right uniformly decays to zero.

▶ **Definition 8.** *Given $\sigma \ge 1$ and $f \in [-1/2, 1/2]$, consider the following random walk on $f + \mathbb{Z}$. For $n \ge 2$ the state $n + f$ moves to $n+1+f$ with probability $p_\sigma(n+f)$ and to $n-1+f$ otherwise. For $n \ge 1$ the state $-n+f$ moves to $-n-1+f$ with probability $p_\sigma(n-f)$ and to $-n+1+f$ otherwise. The state f moves to $1+f$ with probability $p_\sigma(f)$, to state $-1+f$ with probability $p_\sigma(-f)$, and moves to $2+f$ with probability $r_\sigma(f)$. Finally, for $n = 1$ the state $1+f$ moves to $2+f$ with probability $p_\sigma(1+f) - r_\sigma(f)\exp((2f+1)/(2\sigma^2))$ and to f otherwise.*

▶ **Lemma 9.** *For $\sigma \ge 1$ and $f \in [-1/2, 1/2]$, we have that the walk in Definition 8 is well-defined, and that $\mathcal{N}(0, \sigma^2)|_{f+\mathbb{Z}}$ is stationary under a step of the walk with parameters σ, f.*

Proof. That all probabilities are valid follows from Lemma 4, except that we need to additionally verify

$$p_\sigma(1+f) \ge r_\sigma(f) \exp\left(\frac{2f+1}{2\sigma^2}\right).$$

This is precisely Lemma 6.

To verify that $\mathcal{N}(0, \sigma^2)|_{f+\mathbb{Z}}$ is preserved under the walk defined in Definition 8, recall that $\mathcal{N}(0, \sigma^2)|_{f+\mathbb{Z}}$ is preserved under walk defined in Definition 3 by Lemma 4. This walk only differs in its probabilities that f goes to $f, 2+f$ and that $1+f$ goes to $f, 2+f$. Therefore the probabilities at $n+f$ for $n \in \mathbb{Z} \setminus \{0, 2\}$ are correct. Since the probabilities sum to 1, it is enough to check the probability at $2+f$ is correct. It therefore suffices to show that

$$\begin{aligned} r_\sigma(f)\exp\left(-\frac{f^2}{2\sigma^2}\right) + \left(p_\sigma(1+f) - r_\sigma(f)\exp\left(\frac{2f+1}{2\sigma^2}\right)\right)\exp\left(-\frac{(1+f)^2}{2\sigma^2}\right) \\ + (1 - p_\sigma(3+f))\exp\left(-\frac{(3+f)^2}{2\sigma^2}\right) = \exp\left(-\frac{(2+f)^2}{2\sigma^2}\right). \end{aligned}$$

We already verified in the proof of Lemma 4 that

$$p_\sigma(x-1)\exp\left(-\frac{(x-1)^2}{2\sigma^2}\right) + (1 - p_\sigma(x+1))\exp\left(-\frac{(x+1)^2}{2\sigma^2}\right) = \exp\left(-\frac{x^2}{2\sigma^2}\right).$$

Plugging in $x = 2+f$ gives the desired identity, upon canceling the terms containing $r_\sigma(f)$. ◀

Again, we immediately derive a walk which preserves $\mathcal{N}(0, \sigma^2)$ by piecing together all $f \in [-1/2, 1/2)$. Let R_x^σ be the random variable defined by writing $x = n + f$, where $f \in [-1/2, 1/2)$, and then performing a step according to Definition 3.

▶ **Lemma 10.** *If $\sigma \ge 1$ and $Z = \mathcal{N}(0, \sigma^2)$ then $Z + R_Z^\sigma$ is distributed as $\mathcal{N}(0, \sigma^2)$.*

4 Algorithmic Applications

We now derive a number of algorithmic consequences.

Algorithm 1 PARTIALCOLORING$_\sigma(v_1, \cdots, v_t)$.

$w_0 \leftarrow \mathcal{N}(0, \sigma^2 I_n)$
for $1 \le i \le t$ **do**
  $\sigma' \leftarrow \sigma/\|v_i\|_2$
  $x' \leftarrow \langle w_{i-1}, v_i\rangle/\|v_i\|_2$
  $w_i \leftarrow w_{i-1} + J_{x'}^{\sigma'} v_i$.
$w \leftarrow w_t - w_0$

Algorithm 2 BALANCING$_\sigma(v_1, \cdots, v_t)$.

$w_0 \leftarrow \mathcal{N}(0, \sigma^2 I_n)$
for $1 \le i \le t$ **do**
  $\sigma' \leftarrow \sigma/\|v_i\|_2$
  $x' \leftarrow \langle w_{i-1}, v_i\rangle/\|v_i\|_2$
  $w_i \leftarrow w_{i-1} + R_{x'}^{\sigma'} v_i$.
$w \leftarrow w_t - w_0$

In both BALACING$_\sigma$ and PARTIALCOLORING$_\sigma$, J and R are sampled independently every time. Additionally, note that BALACING$_\sigma$ is only well-defined when $\sigma \ge 1$. Finally, we clearly see that PARTIALCOLORING$_\sigma$ assigns a sign of ± 1 to each given vector online, or chooses to omit it (a sign of 0), while BALANCING$_\sigma$ does the same except that the sign 2 is the additional alternative.

Our first algorithm application is a (weak version) of the partial coloring lemma.

▶ **Theorem 11.** *Let $\|v_1\|_2, \dots, \|v_t\|_2 \le 1$ and $\delta \in (0, 1/2)$. With probability at least $1-\delta$ we have that $w_\ell - w_0$ in* PARTIALCOLORING$_1(v_1, \dots, v_t)$ *is $2\sqrt{2\log(2nt/\delta)}$-bounded for all times $\ell \in [t]$. Furthermore, with probability at least $1-\delta$ we have that $w_t - w_0$ is $2\sqrt{2\log(2n/\delta)}$-bounded. Finally, at least 96.3% of vectors are used with probability $1 - \exp(-\Omega(t))$.*

Proof. By Lemma 5 we immediately see that $w_i \sim \mathcal{N}(0, \sigma^2 I_n)$ for all $i \in [t]$. The discrepancy results follow by trivial Gaussian estimates. For example, we see that the jth coordinate of w_ℓ is $\sqrt{2\log(2nt/\delta)}$-bounded with probability at least $\delta/(2nt)$. Taking a union bound over $0 \le \ell \le t$ and $j \in [n]$ yields that $w_0, \dots, w_t$ are bounded with probability at least $1-\delta$. Therefore each difference is also bounded.

The fraction of vectors used being large follows from Chernoff's inequality and the fact that at every step, conditional on all previous choices, a vector is used with probability at least

$$\min_{f \in [-1/2, 1/2]} (1 - r_1(f)) \ge 0.9639. \qquad ◀$$

Our second algorithmic application recovers the online vector balancing results of Alweiss, Liu, and Sawhney [1, Theorems 1.1, 1.2].

▶ **Theorem 12.** *Let* $\|v_1\|_2, \ldots, \|v_t\|_2 \le 1$*,* $\delta \in (0, 1/2)$*, and set* $\sigma = \sqrt{\log(t/\delta)}$*. With probability at least* $1 - \delta$ *we have that* $w_\ell - w_0$ *in* $\textsc{PartialColoring}_\sigma(v_1, \cdots, v_t)$ *is* $2\sqrt{2\log(t/\delta)\log(2nt/\delta)}$*-bounded for all times* $\ell \in [t]$*. Furthermore, with probability at least* $1 - \delta$ *we have that* $w_t - w_0$ *is* $2\sqrt{2\log(t/\delta)\log(2n/\delta)}$*-bounded. Finally, all vectors are used with probability at least* $1 - \delta$*.*

Proof. The proof is essentially identical to that of Theorem 11. The only difference is that we see that at each step, a vector is not used with probability at most

$$\max_{f \in [-1/2, 1/2]} r_\sigma(f) \le e^{-\sigma^2} = \frac{\delta}{t}$$

due to our choice of σ, by the inequality in Lemma 4. A union bound shows that all vectors are used with probability at least $1 - \delta$. ◀

In fact, we can design an algorithm achieving the same bounds by using Algorithm 1 for any value of $\sigma \ge 1$ as follows. To do this, first run Algorithm 1, and then rerun Algorithm 1 on the vectors which were given a 0 sign until no vectors remain (note that this can still be done in an online manner). By Lemma 4, specifically $r_\sigma(f) \le e^{-\sigma^2}$, this process will terminate with probability $1 - \delta$ in $O(\sigma^{-2}\log(t/\delta))$ rounds. Each run produces a random vector with variance $O(\sigma^2)$ in every coordinate, hence the total variance is $O(\log(t/\delta))$ per coordinate as desired.

Finally we recover an online version of Banaszczyk [4], except using $\pm 1, 2$-signings. The proof is identical to that of Theorem 11 so we omit it.

▶ **Theorem 13.** *Let* $\|v_1\|_2, \ldots, \|v_t\|_2 \le 1$ *and* $\delta \in (0, 1/2)$*. With probability at least* $1 - \delta$ *we have that* $w_\ell - w_0$ *in* $\textsc{Balancing}_1(v_1, \ldots, v_t)$ *is* $2\sqrt{2\log(2nt/\delta)}$*-bounded for all times* $\ell \in [t]$*. Furthermore, with probability at least* $1 - \delta$ *we have that* $w_t - w_0$ *is* $2\sqrt{2\log(2n/\delta)}$*-bounded.*

All three algorithmic procedures are online.

4.1 Computational details

In the previous section the above idealized algorithms ignored the cost of computing $r_\sigma(f)$ and $p_\sigma(n \pm f)$ to sufficient precision in order to be used for algorithmic purposes. The key claim is that one can approximate the above sums within δ in $\mathrm{poly}(\log(\sigma/\delta))$-time.

In order to do so first note that we can truncate the sums $p_\sigma(n \pm f)$ and $r_\sigma(f)$ to values of $j \ge 1$ where $(j^2 + 2(n \pm f)j)/(2\sigma^2) = O(\log(\sigma/\delta))$. We now note that

$$\left| e^x - \sum_{j=0}^{m} \frac{x^j}{j!} \right| \le \frac{x^{m+1}}{(m+1)!} e^{\max(0,x)},$$

so taking $m = \Theta(\log(\sigma/\delta))$ gives a very good approximation to $\exp(-(j^2 + 2(n \pm f)j)/(2\sigma^2))$ in the range of terms considered. Now we can compute the desired sums by interpreting it as a sum of low degree (i.e. $O(\log(\sigma/\delta))$) polynomials on a sequence of integers, which can be evaluated quickly.

In the implementation of the algorithms above, at time t if we are given a vector shorter than $1/(2t^2)$, we deterministically add it but ignore it for the purposes of maintaining a Gaussian distribution. These vectors have total length at most 1, so contribute only $O(1)$ discrepancy in each coordinate. For the remaining vectors, we have $\sigma \le 2t^2$. We thus can approximate the relevant probabilities to within $\delta/(2t^2)$ efficiently, and then sample appropriately. This will preserve the Gaussians in question up to total variation distance of at most

$$\sum_{t\geq 1}\frac{\delta}{2t^2}\leq \delta.$$

Therefore, the running time of all probability computations is $\operatorname{poly}(\log(t/\delta))$ at time t. Thus the modified versions of the algorithms in Theorems 11–13 run in $O\left(t\operatorname{poly}(\log(t/\delta))+n+\operatorname{nnz}(\{v_i\}_{i\in[t]})\right)$ time with discrepancy guarantees that are an absolute multiplicative factor worse. (The second term arises due to sampling the initial Gaussian point.) This running time essentially matches (up to logarithmic factors) the results of [1] and make progress towards input-sparsity time algorithms for discrepancy, a direction suggested by [12].

A variant of our algorithms which run in $O\left(t\operatorname{poly}(\log(t/\delta))+n\log t+\operatorname{nnz}(\{v_i\}_{i\in[t]})\right)$ time is achieved by "disregarding vectors" at time t which are shorter than $1/(2t^2)$ (as above) and otherwise grouping vectors by length into dyadic scales and running the algorithms separately with independent randomness on each of the scales. Note that when vector lengths are forced to live in a dyadic scale then sampling an appropriate Gaussian leads us to compute the above probabilities only when $\sigma\in[1,2]$ and hence directly evaluation of the series is efficient.

References

1 Ryan Alweiss, Yang P Liu, and Mehtaab Sawhney. Discrepancy minimization via a self-balancing walk. arXiv:2006.14009.

2 George E. Andrews. A simple proof of Jacobi's triple product identity. *Proc. Amer. Math. Soc.*, 16:333–334, 1965. doi:10.2307/2033875.

3 Juhan Aru, Bhargav Narayanan, Alex Scott, and Ramarathnam Venkatesan. Balancing sums of random vectors. *Discrete Anal.*, pages Paper No. 4, 17, 2018. doi:10.19086/da.3108.

4 Wojciech Banaszczyk. Balancing vectors and Gaussian measures of n-dimensional convex bodies. *Random Structures & Algorithms*, 12(4):351–360, 1998.

5 Nikhil Bansal. Constructive algorithms for discrepancy minimization. In *51th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2010, October 23-26, 2010, Las Vegas, Nevada, USA*, pages 3–10. IEEE Computer Society, 2010. doi:10.1109/FOCS.2010.7.

6 Nikhil Bansal, Daniel Dadush, and Shashwat Garg. An algorithm for Komlós conjecture matching Banaszczyk's bound. In Irit Dinur, editor, *IEEE 57th Annual Symposium on Foundations of Computer Science, FOCS 2016, 9-11 October 2016, Hyatt Regency, New Brunswick, New Jersey, USA*, pages 788–799. IEEE Computer Society, 2016. doi:10.1109/FOCS.2016.89.

7 Nikhil Bansal, Daniel Dadush, Shashwat Garg, and Shachar Lovett. The gram-schmidt walk: a cure for the Banaszczyk blues. In Ilias Diakonikolas, David Kempe, and Monika Henzinger, editors, *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing, STOC 2018, Los Angeles, CA, USA, June 25-29, 2018*, pages 587–597. ACM, 2018. doi:10.1145/3188745.3188850.

8 Nikhil Bansal and Shashwat Garg. Algorithmic discrepancy beyond partial coloring. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 914–926, 2017.

9 Nikhil Bansal, Haotian Jiang, Raghu Meka, Sahil Singla, and Makrand Sinha. Online discrepancy minimization for stochastic arrivals. In Dániel Marx, editor, *Proceedings of the 2021 ACM-SIAM Symposium on Discrete Algorithms, SODA 2021, Virtual Conference, January 10 - 13, 2021*, pages 2842–2861. SIAM, 2021. doi:10.1137/1.9781611976465.169.

10 Nikhil Bansal, Haotian Jiang, Sahil Singla, and Makrand Sinha. Online vector balancing and geometric discrepancy. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2020, pages 1139–1152, New York, NY, USA, 2020. Association for Computing Machinery. doi:10.1145/3357713.3384280.

11 Nikhil Bansal and Joel H Spencer. On-Line Balancing of Random Inputs. arXiv:1903.06898.

12 Daniel Dadush. URL: https://homepages.cwi.nl/~dadush/workshop/discrepancy-ip/open-problems.html.

13 Daniel Dadush, Aleksandar Nikolov, Kunal Talwar, and Nicole Tomczak-Jaegermann. Balancing vectors in any norm. In *2018 IEEE 59th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1–10. IEEE, 2018.

14 Ronen Eldan and Mohit Singh. Efficient algorithms for discrepancy minimization in convex sets. *Random Struct. Algorithms*, 53(2):289–307, 2018. doi:10.1002/rsa.20763.

15 Shachar Lovett and Raghu Meka. Constructive discrepancy minimization by walking on the edges. *SIAM J. Comput.*, 44(5):1573–1582, 2015. doi:10.1137/130929400.

16 Thomas Rothvoss. Constructive discrepancy minimization for convex sets. *SIAM Journal on Computing*, 46(1):224–234, 2017.

17 Joel Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289(2):679–706, 1985. doi:10.2307/2000258.

A Proof of Lemma 6

Proof of Lemma 6. First note $r_\sigma(f) \le r_\sigma(0)$ and $r_\sigma(0) \le r_1(0)$ follow immediately from the Jacobi triple product identity Equation (1) and nonnegativity. Therefore it suffices to prove that

$$p_\sigma(1+f) \ge \exp(1/\sigma^2) r_1(0)$$

for all $\sigma \ge 1$ and $f \in [-1/2, 1/2]$.

First suppose that $\sigma \in [1,2]$. Then since $p_\sigma(1+f)$ is an alternating series with decreasing terms,

$$p_\sigma(1+f) \ge \exp\left(-\frac{1+2(1+f)}{2\sigma^2}\right) - \exp\left(-\frac{2+2(1+f)}{\sigma^2}\right).$$

Fixing σ, the right has derivative

$$-\frac{1}{\sigma^2}\exp\left(-\frac{1+2(1+f)}{2\sigma^2}\right) + \frac{2}{\sigma^2}\exp\left(-\frac{2+2(1+f)}{\sigma^2}\right),$$

which we can check is positive for f underneath some cutoff and negative above this cutoff. Therefore the earlier expression is minimized over $f \in [-1/2, 1/2]$ at some $f \in \{\pm 1/2\}$. Then, numerical checking shows that for each case $f \in \{\pm 1/2\}$ the resulting expression is minimized on $\sigma \in \{1, 2\}$ for similar reasons. We find the true minimum is at $f = 1/2$ and $\sigma = 1$, which gives

$$p_\sigma(1+f) \ge 0.12 \ge e r_1(0) \ge \exp(1/\sigma^2) r_1(0).$$

Now we suppose that $\sigma \ge 2$. Let $2k-1$ be the smallest odd integer larger than $\sigma - 1 - f$, which is clearly always a positive integer as $\sigma \ge 1$ and $f \le 1/2$. We know that $t \mapsto \exp(-t^2/(2\sigma^2))$ is convex for $t \ge \sigma$, hence $t \mapsto \exp(-(t^2 + 2(1+f)t)/(2\sigma^2))$ is certainly convex and decreasing for $t \ge \sigma - 1 - f$. Therefore the difference between the values at j and $j+1$ is at least the difference between the values at $j+1$ and $j+2$ when $j \ge 2k-1$, yielding

$$
\begin{aligned}
p_\sigma(1+f) &= \sum_{j\geq 1}(-1)^{j-1}\exp\left(-\frac{j^2+2(1+f)j}{2\sigma^2}\right)\\
&\geq \sum_{j\geq 2k-1}(-1)^{j-1}\exp\left(-\frac{j^2+2(1+f)j}{2\sigma^2}\right)\\
&\geq \frac{1}{2}\left(\sum_{j\geq 2k-1}(-1)^{j-1}\exp\left(-\frac{j^2+2(1+f)j}{2\sigma^2}\right)+\sum_{j\geq 2k}(-1)^{j}\exp\left(-\frac{j^2+2(1+f)j}{2\sigma^2}\right)\right)\\
&= \frac{1}{2}\exp\left(-\frac{(2k-1)^2+2(1+f)(2k-1)}{2\sigma^2}\right)\\
&\geq \frac{1}{2}\exp\left(-\frac{(\sigma+1-f)^2+2(1+f)(\sigma+1-f)}{2\sigma^2}\right)\\
&\geq \frac{1}{2}\exp\left(-\frac{4\sigma^2+16\sigma+15}{8\sigma^2}\right)\geq \frac{1}{2}\exp(-71/32)\exp(1/\sigma^2)\\
&\geq 0.05\exp(1/\sigma^2)\geq \exp(1/\sigma^2)r_1(0).
\end{aligned}
$$

◀