2018 IEEE 59th Annual Symposium on Foundations of Computer Science

Balancing Vectors in Any Norm

Daniel Dadush*, Aleksandar Nikolov†, Kunal Talwar‡, Nicole Tomczak-Jaegermann§
* *Centrum Wiskunde & Informatica, Amsterdam, Netherlands*
Email: dadush@cwi.nl
† *University of Toronto, Toronto, Ontario,Canada*
Email: anikolov@cs.toronto.edu
‡ *Google Brain*
Email: kunal@google.com
† *University of Alberta Edmonton, Alberta, Canada*
Email: nicole.tomczak@ualberta.ca

***Abstract*—In the vector balancing problem, we are given symmetric convex bodies C and K in $\mathbb{R}^n$, and our goal is to determine the minimum number $\beta \geq 0$, known as the vector balancing constant from C to K, such that for any sequence of vectors in C there always exists a signed combination of them lying inside βK. Many fundamental results in discrepancy theory, such as the Beck-Fiala theorem (Discrete Appl. Math '81), Spencer's "six standard deviations suffice" theorem (Trans. Amer. Math. Soc '85) and Banaszczyk's vector balancing theorem (Random Structures & Algorithms '98) correspond to bounds on vector balancing constants.**

The above theorems have inspired much research in recent years within theoretical computer science. In this work, we show that all vector balancing constants admit "good" approximate characterizations, with approximation factors depending only polylogarithmically on the dimension n. First, we show that a volumetric lower bound due to Banaszczyk is tight within a $O(\log n)$ factor. Our proof is algorithmic, and we show that Rothvoss's (FOCS '14) partial coloring algorithm can be analyzed to obtain these guarantees. Second, we present a novel convex program which encodes the "best possible way" to apply Banaszczyk's vector balancing theorem for bounding vector balancing constants from above, and show that it is tight within an $O(\log^{2.5} n)$ factor. This also directly yields a corresponding polynomial time approximation algorithm both for vector balancing constants, and for the hereditary discrepancy of any sequence of vectors with respect to an arbitrary norm.

***Keywords*-Discrepancy; Convex Geometry; Gaussian measure; M-ellipsoid; K-convexity.**

I. Introduction

The discrepancy of a set system is defined as the minimum, over the set of ± 1 colorings of the elements, of the imbalance between the number of $+1$ and -1 elements in the most imbalanced set. Classical combinatorial discrepancy theory studies bounds on the discrepancy of set systems, in terms of their structure. The tools developed for deriving bounds on the discrepancy of set systems have found many applications in mathematics and computer science [1], [2], from the study of pseudorandomness, to communication complexity, and most recently, to approximation algorithms and privacy. Here we study a geometric generalization of combinatorial discrepancy, known as vector balancing, which captures some of the most powerful techniques in the area, and is of intrinsic interest.

Vector Balancing: In many instances, the best known techniques for finding good bounds in combinatorial discrepancy were derived by working with more general vector balancing problems, where convex geometric techniques can be applied. Given symmetric convex bodies $C, K \subseteq \mathbb{R}^n$, the vector balancing constant of C into K is defined as

$$\mathrm{vb}(C,K) := \sup \left\{ \min_{x \in \{-1,1\}^N} \Big\| \sum_{i=1}^{N} x_i u_i \Big\|_K : N \in \mathbb{N}, u_1, \dots, u_N \in C \right\},$$

where $\|x\|_K := \min \{ s \geq 0 : x \in sK \}$ is the norm induced by K.

As an example, one may consider Spencer's "six standard deviations" theorem [3], independently obtained by Gluskin [4], which states that every set system on n points and n sets can be colored with discrepancy at most $O(\sqrt{n})$. In the vector balancing context, the more general statement is that $\mathrm{vb}(B_\infty^n, B_\infty^n) = O(\sqrt{n})$ (also proved in [3], [4]), where we use the notation $B_p^n = \left\{ x \in \mathbb{R}^n : \|x\|_p \leq 1 \right\}$, $p \in [1, \infty]$, to denote the unit ball of the ℓ_p norm. To encode Spencer's theorem, we simply represent the set system using its incidence matrix $U \in \{0,1\}^{n \times n}$, where $U_{ji} = 1$ if element i is in set j and 0 otherwise. Here the columns of U have ℓ_∞ norm 1, and thus the sign vector $x \in \{-1,1\}^n$ satisfying $\|Ux\|_\infty = O(\sqrt{n})$ indeed yields the desired coloring.

In fact, vector balancing was studied earlier, and independently from combinatorial discrepancy. In 1963 Dvoretzky posed the general problem of determining $\mathrm{vb}(K, K)$ for a given symmetric convex body K. The more general version with two different bodies was introduced by Barany and Grinberg [5] who proved that for any symmetric convex body K in $\mathbb{R}^n$, $\mathrm{vb}(K,K) \leq n$. In addition to Spencer's

2575-8454/18/$31.00 ©2018 IEEE
DOI 10.1109/FOCS.2018.00010

theorem, as described above, many other fundamental discrepancy bounds, as well as conjectured bounds, can be stated in terms of vector balancing constants. The Beck-Fiala theorem, which bounds the discrepancy of any t-sparse set system (i.e. one in which each element appears in at most t sets) by $2t-1$, can be recovered from the bound $\mathrm{vb}(B_1^n, B_\infty^n) < 2$ [6]. The Beck-Fiala conjecture, which asks whether the bound for t-sparse set systems can be improved to $O(\sqrt{t})$, is generalized by the Kómlos conjecture [7], which asks whether $\mathrm{vb}(B_2^n, B_\infty^n) = O(1)$. One of the most important vector balancing bounds is due to Banaszczyk [8], who proved that for any convex body $K \subseteq \mathbb{R}^n$ of Gaussian measure $1/2$, one has the bound $\mathrm{vb}(B_2^n, K) \leq 5$. In particular, this implies the bound of $\mathrm{vb}(B_2^n, B_\infty^n) = O(\sqrt{\log n})$ for the Kómlos conjecture.

Hereditary Discrepancy.: While vector balancing gives useful worst-case bounds, one is often interested in understanding the discrepancy guarantees one can get for instances derived from a fixed set of vectors, known as hereditary discrepancy. Given vectors $(u_i)_{i=1}^N$ in $\mathbb{R}^n$, the discrepancy and hereditary discrepancy with respect to a symmetric convex body $K \subseteq \mathbb{R}^n$ are defined as:

$$\mathrm{disc}((u_i)_{i=1}^N, K) := \min_{\varepsilon_1,\ldots,\varepsilon_N \in \{-1,1\}} \Big\| \sum_{i=1}^N \varepsilon_i u_i \Big\|_K;$$
$$\mathrm{hd}((u_i)_{i=1}^N, K) := \max_{S \subseteq [N]} \mathrm{disc}((u_i)_{i\in S}, K).$$

When convenient, we will also use the notation $\mathrm{hd}(U,K) := \mathrm{hd}((u_i)_{i=1}^N, K)$, where $U := (u_1,\ldots,u_N) \in \mathbb{R}^{n\times N}$, and $\mathrm{disc}(U_S, K) := \mathrm{disc}((u_i)_{i\in S}, K)$ for any subset $S \subseteq [N]$. In the context of set systems, ℓ_∞ hereditary discrepancy corresponds to the worst-case discrepancy of any element induced subsystem, which gives a robust notion of discrepancy, and can be seen as a measure of the complexity of the set system. As an interesting example, a set system has ℓ_∞ hereditary discrepancy 1 if and only if its incidence matrix is totally unimodular [9].

Beyond set systems, hereditary discrepancy can also usefully bound the worst-case "error" required for rounding a fractional LP solution to an integral one. More precisely, given any solution $y \in \mathbb{R}^n$ to a linear programming relaxation $Ax \leq b$, $x \in [0,1]^n$, with $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, of a binary IP, and given any norm $\|\cdot\|$ on $\mathbb{R}^m$ measuring "constraint violation", one can ask what guarantees can be given on $\min_{x\in\{0,1\}^n} \|A(y-x)\|$? Using a well-known reduction of Lovász, Spencer and Vesztergombi [10], this error can be bounded by $\mathrm{hd}(A,K)$ where K is the unit ball of $\|\cdot\|$. Furthermore, this reduction guarantees that x agrees with y on its integer coordinates. Note that we have the freedom to choose the norm $\|\cdot\|$ so that the error bounds meaningfully relate to the structure of the problem. Indeed, much work has been done on the achievable "error profiles" one can obtain algorithmically, e.g. for which $\Delta \in \mathbb{R}_{>0}^m$ we can always find $x \in \{0,1\}^m$ satisfying $|A(y-x)| \leq \Delta$, $\forall y \in [0,1]^m$? Note that the feasibility of an error profile can be recovered from a bound of 1 on the hereditary discrepancy with respect to the weighted ℓ_∞ norm $\|y-x\|_\Delta = \max_{i\in[m]} |y_i - x_i|/\Delta_i$. Indeed, in many instances, this is (at least implicitly) how these bounds are proved. These error profile bounds have been fruitfully leveraged for problems where small "additive violations" to the constraints are either allowed or can be repaired. In particular, they were used in the recent $O(\log n)$-additive approximation for bin packing [11], an additive approximation scheme for the train delivery problem [12], and additive approximations of the degree bounded matroid basis problem [13].

Discrepancy Minimization.: The original proofs of many of the aforementioned discrepancy upper bounds were existential, and did not come with efficient algorithms capable of constructing the requisite low discrepancy colorings. Starting with the breakthrough work of Bansal [14], who gave a constructive version of Spencer's theorem using random walk and semidefinite programming techniques, nearly all known bounds have been made algorithmic in the last eight years.

One of the most important discrepancy minimization techniques is Beck's partial coloring method, which covers most of the above discrepancy results apart from Banaszczyk's vector balancing theorem. This method was first primarily applied to ℓ_∞ discrepancy minimization problems of the form

$$\min_{x\in\{-1,1\}^n} \Big\| \sum_{i=1}^n x_i v_i \Big\|_\infty, \text{ where } (v_i)_{i=1}^n \in \mathbb{R}^m.$$

As before, the goal is not to solve such problems near-optimally but instead to find solutions satisfying a guaranteed error bound. The partial coloring method solves this problem in phases, where at each phase it "colors" (i.e. sets to ± 1) at least a constant fraction of the remaining uncolored variables. This yields $O(\log n)$ partial coloring phases, where the discrepancy of the full coloring is generally bounded by the sum of discrepancies incurred in each phase. The existence of low discrepancy partial colorings was initially established via the pigeon hole principle and arguments based on the probabilistic and the entropy methods. In particular, the entropy method gave a general sufficient condition for the feasibility of any error profile (as above) with respect to partial colorings. This method was made constructive by Lovett and Meka [15] using random walk techniques. The partial coloring method was generalized by Giannopoulos [16] to the general vector balancing setting using Gaussian measure. Precisely, he showed that if a symmetric convex body $K \subseteq \mathbb{R}^n$ has Gaussian measure at least 2^{-cn}, for c small enough, then for any sequence of vectors $v_1,\ldots,v_n \in B_2^n$, there exists a partial coloring $x \in \{-1,0,1\}^n$, having support at least $n/2$, such that $\sum_{i=1}^n x_i v_i \in O(1)K$. This method was made constructive by Rothvoss [17], using a random projection algorithm, and later by Eldan and Singh [18] who used the solution of a random linear maximization problem. An important difference between the constructive and existential partial coloring methods, is that the constructive methods only guarantee that the "uncolored" coordinates of a partial coloring x are in $(-1,1)$ instead of equal to 0. This relaxation seems to make the constructive methods more robust, i.e. the conditions needed for such "fractional" partial colorings are somewhat milder, without having noticeable drawbacks in most applications.

The main alternative to the partial coloring method comes from Banaszczyk's vector balancing theorem [8]. Banaszczyk's method proves the existence of a full coloring when K has Gaussian measure $1/2$, in contrast to Giannopoulos's result which gives a partial coloring but requires measure only 2^{-cn}. Banaszczyk's method was only very recently made constructive in the sequence of works [19]–[21]. In particular, [20] showed an equivalence of Banaszczyk's theorem to the existence of certain subgaussian signing distributions, and [21] gave a random walk-based algorithm to build such distributions.

A. Approximating Vector Balancing and Hereditary Discrepancy

Given the powerful tools mentioned above, a natural question is whether they can be extended to get nearly optimal bounds for any vector balancing or hereditary discrepancy problem. More precisely, we will be interested in the following computational and mathematical questions:

1) Given vectors $(u_i)_{i=1}^N$ and a symmetric convex body K in $\mathbb{R}^n$, can we (a) efficiently compute a coloring whose K-discrepancy is approximately bounded by $\mathrm{hd}((u_i)_{i=1}^N, K)$? (b) efficiently approximate $\mathrm{hd}((u_i)_{i=1}^N, K)$?
2) Given two symmetric convex bodies $C, K \subseteq \mathbb{R}^n$, does $\mathrm{vb}(C,K)$ admit a "good" characterization? Namely, are there simple certificates which certify nearly tight upper and lower bounds on $\mathrm{vb}(C,K)$?

To begin, a few remarks are in order. Firstly, question 2 can be inefficiently encoded as question 1b, by letting $(u_i)_{i=1}^N$ denote a sufficiently fine net of C. Thus "good" characterizations for hereditary discrepancy transfer over to vector balancing, and, for this reason, we restrict for now the discussion to the former. For question 1a, one may be tempted to ask whether we can directly compute a coloring whose K-discrepancy is approximately $\mathrm{disc}((u_i)_{i=1}^N, K)$ instead of $\mathrm{hd}((u_i)_{i=1}^N, K)$. Unfortunately, even for $K = B_\infty^n$ and $(u_i)_{i=1}^n \in [-1,1]^n$, it was shown in [22] that it is NP-hard to distinguish whether $\mathrm{disc}((u_i)_{i=1}^n, B_\infty^n)$ is 0 or $\Omega(\sqrt{n})$ (note that $O(\sqrt{n})$ is guaranteed by Spencer's theorem), thus one cannot hope for any non-trivial approximation guarantee in this context.

We now discuss prior work on these questions and then continue with our main results.

B. Prior work.

For both questions 1 and and 2 above, prior work has mostly dealt with the case of ℓ_∞ or ℓ_2 discrepancy. Bounds on vector balancing constants from ℓ_p to ℓ_q for some p and q have also been studied, as described earlier, but without a unified approach. The question of obtaining near-optimal results for general vector balancing and hereditary discrepancy problems has not been studied before.

In terms of coloring algorithms, Bansal [14] gave a partial coloring based random walk algorithm which on $U \in \mathbb{R}^{m\times n}$, produces a full coloring with ℓ_∞ discrepancy $O(\sqrt{\log m \log \mathrm{rk}(U)}\,\mathrm{hd}(U, B_\infty^m))$, where $\mathrm{rk}(U)$ is the rank of U. Recently, Larsen [23] gave an algorithm for the ℓ_2 norm achieving discrepancy $O(\sqrt{\log(\mathrm{rk}(U))}\,\mathrm{hd}(U, B_2^m))$.

In terms of certifying lower bounds on $\mathrm{hd}(U, B_\infty^m))$, the main tool has been the so-called determinant lower bound of [10], where it was shown that

$$\mathrm{hd}(U, B_\infty^m) \geq \mathrm{detLB}(U) := \max_k \max_B \frac{1}{2} |\det(B)|^{1/k}$$

with the maximum over $k \times k$ submatrices B of U. Matoušek [24] built upon the results of [14] to show that

$$\mathrm{hd}(U, B_\infty^m) \leq O(\sqrt{\log m} \log \mathrm{rk}(U)\, \mathrm{detLB}(U)).$$

For certifying tight upper bounds, [25], [26] showed that the γ_2 norm of U, defined by

$$\gamma_2(U) := \min\{\|A\|_{2\to\infty}\|B\|_{1\to 2} : U = AB, \\ A \in \mathbb{R}^{m\times k}, B \in \mathbb{R}^{k\times n}, k \in \mathbb{N}\}$$

where $\|A\|_{2\to\infty}$ is the maximum ℓ_2 norm of any row of A, and $\|B\|_{1\to 2}$ is the maximum ℓ_2 norm of any column of B, satisfies

$$\Omega(\gamma_2(U)/\log(\mathrm{rk}(U))) \leq \mathrm{detLB}(U) \\ \leq \mathrm{hd}(U, B_\infty^m) \leq O(\sqrt{\log m}\gamma_2(U))$$

which implies a $O(\sqrt{\log m} \log \mathrm{rk}(U))$ approximation to ℓ_∞ hereditary discrepancy. For ℓ_2, it was shown in [25] that a relaxation of γ_2 yields an $O(\log \mathrm{rk}(U))$-approximation to $\mathrm{hd}(U, B_2^m)$. We note that part of the strategy of [25], [26] is to replace the ℓ_∞ norm via an averaged version of ℓ_2, where one optimizes over the averaging coefficients, which makes the ℓ_2 norm by itself an easier special case.

Moving to general norms.: While at first glance it may seem that the above techniques for ℓ_∞ do not apply to more general norms, this is in some sense deceptive. Notwithstanding complexity considerations, every norm can be isometrically embedded into ℓ_∞, where in particular any polyhedral norm with m facets can be embedded into B_∞^m. Vice versa, starting from a matrix $U \in \mathbb{R}^{m\times N}$, with $\mathrm{rk}(U) = n$ and rank factorization $U = AB$, it is direct to verify that $\mathrm{hd}(U, B_\infty^m) = \mathrm{hd}(B, K)$, where $K = \{x \in \mathbb{R}^n : \|Ax\|_\infty \leq 1\}$ is an n-dimensional symmetric polytope with m facets. Thus, for any $U \in \mathbb{R}^{n\times N}$, one can equivalently restate the guarantees of [14] as yielding colorings of discrepancy $O(\sqrt{\log m \log n}\, \mathrm{hd}(U,K))$ and of [26] as a $O(\sqrt{\log m}\log n)$ approximation to $\mathrm{hd}(U,K)$ for any n-dimensional symmetric polytope K with m facets. A natural question is therefore whether there exist corresponding coloring and approximation algorithms whose guarantees depend only polylogarithmically on the dimension of the norm and not on the complexity of its representation. The upper and lower bound tools mentioned above (the determinant lower bound and the γ_2 norm, respectively) are insufficient for this task, as the polylogarithmic dependence on m in the bounds is known to be inherent.

We note that polynomial bounds in n for general K can be achieved by simply approximating K by a sandwiching ellipsoid $E \subseteq K \subseteq \sqrt{n}E$ and applying the corresponding results for ℓ_2, which yield $O(\sqrt{n \log n})$ coloring and $O(\sqrt{n}\log n)$ approximations guarantees respectively. Interestingly, these guarantees are identical to what can be achieved by replacing K by a symmetric polytope with 3^n facets, giving a sandwiching factor of 2, and applying the ℓ_∞ results.

II. Results

Our main results are that polylogarithmic approximations to hereditary discrepancy and vector balancing constants with respect to arbitrary norms are indeed possible. In particular, given $U \in \mathbb{R}^{n\times N}$ and a symmetric convex body $K \subseteq \mathbb{R}^n$ (specified by an appropriate oracle), we give randomized polynomial time algorithms for computing colorings of discrepancy $O(\log n\, \mathrm{hd}(U,K))$ and approximating $\mathrm{hd}(U,K)$ up to an $O(\log^{2.5} n)$ factor. Furthermore, if K is a polyhedral norm with at most m facets, our approximation algorithm for $\mathrm{hd}(U,K)$ always achieves a tighter approximation factor than the γ_2 bound, and hence gives an $O(\min\{\log n\sqrt{\log m}, \log^{2.5} n\})$ approximation. To achieve these results, we first show that Rothvoss' partial coloring algorithm [17] is nearly optimal for general hereditary discrepancy by showing near-tightness with respect to a volumetric lower bound of Banaszczyk [27]. Second, we show that the "best possible way" to apply Banaszczyk's vector balancing theorem [8] for the purpose of bounding $\mathrm{hd}(U,K)$ from above can be encoded as a convex program, and prove that this bound is tight to within an $O(\log^{2.5} n)$ factor. As a consequence, we show that Banaszczyk's theorem is essentially "universal" for vector balancing. To analyze these approaches we rely on a novel combination of tools from convex geometry and discrepancy. In particular, we give a new way to prove lower bounds on Gaussian measure using only volumetric information, which could be of independent interest. Furthermore, we make a natural geometric conjecture which would imply that Rothvoss' algorithm is (in a hereditary sense) optimal for finding partial colorings in any norm, and prove the conjecture for the special case of ℓ_2.

Comparing to prior work, our coloring and hereditary discrepancy approximation algorithms give uniformly better (or at at least no worse) guarantees in almost every setting which has been studied. Furthermore our methods provide a unified approach for studying discrepancy in arbitrary norms, which we expect to have further applications.

Interestingly, our results imply a tighter relationship between vector balancing and hereditary discrepancy than one might initially expect. We observe that neither the volumetric lower bound we use nor our factorization based upper bound "see" the difference between vector balancing constants and hereditary discrepancy. More precisely, both bounds remain invariant when replacing $\mathrm{hd}(U,K)$ by $\mathrm{vb}(\mathrm{conv}\{\pm u_i : i \in [N]\}, K)$. This has the relatively non-obvious implication that

$$\mathrm{hd}(U,K) \le \mathrm{vb}(\mathrm{conv}\{\pm u_i : i \in [N]\}, K) \le O(\log n)\, \mathrm{hd}(U,K). \quad (1)$$

We believe it is an interesting question to understand whether a polylogarithmic separation indeed exists between the above quantities (we are currently unaware of any examples), as it would give a tangible geometric obstruction for tighter approximations.

III. Outline of the Proofs

Starting with hereditary discrepancy, to push beyond the limitations of prior approaches the first two tasks at hand are: (1) find a stronger lower bound and (2) develop techniques to avoid the "union bound". Fortunately, a solution to the first problem was already given by Banaszczyk [27], which we present in slightly adapted form below.

Lemma 1 (Volume Lower Bound)**.** *Let $U = (u_1, \dots, u_N) \in \mathbb{R}^{n\times N}$ and $K \subseteq \mathbb{R}^n$ be a symmetric convex body. For $S \subseteq [N]$, let U_S denote the submatrix of U consisting of the columns indexed by S. For $k \in [n]$, define*

$$\mathrm{volLB}^{\mathrm{h}}_k((u_i)_{i=1}^N, K) := \mathrm{volLB}^{\mathrm{h}}_k(U,K) := \max_{S\subseteq[N], |S|=k} \mathrm{vol}_k(\{x \in \mathbb{R}^k : U_S x \in K\})^{-1/k}. \quad (2)$$

Then, we have that

$$\mathrm{volLB}^{\mathrm{h}}((u_i)_{i=1}^N, K) := \mathrm{volLB}^{\mathrm{h}}(U,K) := \max_{k\in[n]} \mathrm{volLB}^{\mathrm{h}}_k(U,K) \le \mathrm{hd}(U,K). \quad (3)$$

A formal proof of the above is given in the full version, and follows from the argument in [27]. At a high level, the proof is a simple covering argument, where it is argued that for any subset S, $|S| = k$, every point in $[0,1]^k$ is at distance at most $\mathrm{hd}(U,K)$ from $\{0,1\}^k$ under the norm induced by $C := \{x \in \mathbb{R}^k : U_S x \in K\}$. Equivalently an $\mathrm{hd}(U,K)$ scaling of C placed around the points of $\{0,1\}^k$ cover $[0,1]^k$, and hence by a standard lattice argument must have volume at least that of $[0,1]^k$, namely 1. This yields the desired lower bound after rearranging.

We note that the volume lower bound extends in the obvious way to vector balancing. In particular, for two symmetric convex bodies $C, K \subseteq \mathbb{R}^n$, we define

$$\mathrm{volLB}^{\mathrm{h}}(C,K) := \sup \mathrm{volLB}((u_i)_{i=1}^k, K),$$

with the supremum over $k \in [n]$ and sequences $u_1, \dots, u_k \in C$. Then we have

$$\mathrm{volLB}^{\mathrm{h}}(C,K) \geq \mathrm{vb}(C,K). \tag{4}$$

This lower bound can be substantially stronger than the determinant lower bound for ℓ_∞ discrepancy. As a simple example, let $U \in \mathbb{R}^{2^n \times n}$ be the matrix having a row for each vector in $\{-1,1\}^n$. Since U has rank n, the determinant lower bound is restricted to $k \times k$ matrices for $k \in [n]$. Hadamard's inequality implies that for any $k \times k$ matrix B with ± 1 entries we have $|\det(B)|^{1/k} \leq \sqrt{k} \leq \sqrt{n}$. A moment's thought, however, reveals that for $x \in \mathbb{R}^n$, $\|Ux\|_\infty = \|x\|_1$ and hence any coloring $x \in \{-1,1\}$ must have discrepancy $\|x\|_1 = n$. Using the previous logic, the volume lower bound yields by standard estimates

$$\begin{aligned}\mathrm{volLB}(U, B_\infty^m) &\geq \mathrm{vol}_n(\{x \in \mathbb{R}^n : \|x\|_1 \leq 1\})^{-1/n} \\ &= \mathrm{vol}_n(B_1^n)^{-1/n} = (n!/2^n)^{1/n} \geq n/(2e),\end{aligned}$$

which is essentially tight.

A. From Volume to Coloring

The above example gives hope that the volume lower bound can circumvent a dependency on the facet complexity of the norm. Our first main result shows that indeed this is the case:

Theorem 2 (Tightness of the Volume Lower Bound)**.** *For any $U \in \mathbb{R}^{n \times N}$ and symmetric convex body K in $\mathbb{R}^n$, we have that*

$$\mathrm{volLB}^{\mathrm{h}}(U,K) \leq \mathrm{hd}(U,K) \leq O(\log(n)\, \mathrm{volLB}^{\mathrm{h}}(U,K)).$$

Furthermore, there exists a randomized polynomial time algorithm that computes a coloring of U with K-discrepancy $O(\log n \, \mathrm{volLB}^{\mathrm{h}}(U,K))$, given a membership oracle for K.

We note that the above immediately implies the corresponding approximate tightness of the volume lower bound for vector balancing. The above bound can also be shown to be tight. In particular, the counterexample to the 3-permutations conjecture from [28], which has ℓ_∞ discrepancy $\Omega(\log n)$, can be shown to have volume lower bound $O(1)$. The computations for this are somewhat technical, so we defer a detailed discussion to the full version. As mentioned previously, an interesting property of the volume lower bound is its invariance under taking convex hulls, namely $\mathrm{volLB}^{\mathrm{h}}(\mathrm{conv}\{\pm U\}, K) = \mathrm{volLB}^{\mathrm{h}}(U,K)$. In combination with Theorem 2, this establishes the claimed inequality (1). This invariance is proved in the full version by showing that the volume lower bound is convex in each u_i, and hence maximized at extreme points.

Our proof of Theorem 2 is algorithmic, and relies on iterated applications of Rothvoss's partial coloring algorithm. We now explain our high level strategy as well as the differences with respect to prior approaches.

For simplicity of the presentation, we shall assume that $U = (e_1, \dots, e_n) \in \mathbb{R}^{n \times n}$ and that the volume lower bound $\mathrm{volLB}^{\mathrm{h}}((e_i)_{i=1}^n, K) = 1$. This can be (approximately) achieved by applying a standard reduction to the case where U is non-singular, so $N \leq n$, "folding" U into K, and appropriately guessing the volume lower bound.

For any subset $S \subseteq [n]$, let $K_S := \{x \in K : x_i = 0 \ \forall i \in [n] \setminus S\}$ denote the coordinate section of K induced by S. Since the vectors of U now correspond to the coordinate basis, it is direct to verify that

$$\mathrm{volLB}^{\mathrm{h}}((e_i)_{i=1}^n, K) = \max_{S \subseteq [n], k := |S|} \mathrm{vol}_k(K_S)^{-1/k}.$$

In particular, the assumption $\mathrm{volLB}^{\mathrm{h}}((e_i)_{i=1}^n, K) = 1$ implies that

$$\mathrm{vol}_{|S|}(K_S) \geq 1, \ \forall S \subseteq [n]. \tag{5}$$

Under this condition, our goal can now be stated as finding a coloring $x \in \{-1,1\}^n \in O(\log n)K$.

When K is a symmetric polytope with m facets, the algorithm by Bansal [14] uses a "sticky" random walk inside $[0,1]^n$ with increments computed via an SDP. The SDP is used to guarantee that the variance in the normal direction of any facet is at bounded by $s\, \mathrm{hd}((e_i)_{i=1}^n, K)^2$, while the variance in the direction of any (active) coordinate directions is at least s, for a small parameter s. As this only gives probabilistic error guarantees for each constraint in isolation, a union bound is used to get a global guarantee, incurring the $O(\sqrt{\log m})$ dependence in the final bound.

To avoid the "union bound", we instead use Rothvoss's partial coloring algorithm, which simply samples a random Gaussian vector $X \in \mathbb{R}^n$ and computes the closest point in Euclidean distance x to X in $K \cap [-1,1]^n$ as the candidate partial coloring. As long as K has "large enough" Gaussian measure, Rothvoss shows that x has at least a constant fraction of its components at ± 1. While this method can in better leverage the geometry of K than Bansal's method (in particular, it does not need an explicit description of K), it is apriori unclear why Gaussian measure should be large enough in the present context.

Our main technical result is that if all the coordinate sections of K have volume at least 1 (i.e. condition (5)), then there indeed exists a section of K of dimension close to n, whose Gaussian measure is "large" after an appropriate scaling. Specifically, we show that for any $\delta \in (0,1)$, there exists a subspace H of dimension $(1-\delta)n$ such that the Gaussian measure of $2^{O(1/\delta)}(K \cap H)$ is at least $2^{-\delta n}$ (the exact statement is given in the full version). We sketch the ideas in the next subsection.

The existence of a large section of K with Gaussian measure which is not too small in fact suffices to run Rothvoss's partial coloring algorithm. Conveniently, one does not need to know the section explicitly, as its existence is only used in the analysis of the algorithm. Since condition 5 is hereditary, we can now find partial colorings of K-discrepancy $O(1)$ on any subset of coordinates. Thus, applying $O(\log n)$ partial coloring phases in the standard way yields the desired full coloring.

A useful restatement of the above is that Rothvoss's algorithm can always find partial colorings with discrepancy $O(1)$ times the volume lower bound.

B. Finding a section with large Gaussian measure.

We now sketch how to find a section of K of large Gaussian measure under the assumption that $\mathrm{vol}_{|S|}(K_S) \geq 1, \forall S \subseteq [n]$. The main tool we require is the M-ellipsoid from convex geometry [29]. The M-ellipsoid E of K is an ellipsoid which approximates K well from the perspective of covering, that is $2^{O(n)}$ translates of E suffice to cover K and vice versa.

The main idea is to use the volumetric assumption to show that the longest $(1-\delta)n$ axes of E, for $\delta \in (0,1)$ of our choice, have length at least $\sqrt{n}2^{-O(1/\delta)}$, and then use the subspace generated by these axes for the section of K we use. On this subspace H, we have that a $2^{O(1/\delta)}$ scaling of $E \cap H$ contains the $\sqrt{n}$ ball, and thus by the covering estimate $2^{O(n)}$ translates of $2^{O(1/\delta)}(K \cap H)$ cover the $\sqrt{n}$ ball. Since the $\sqrt{n}$ ball on H has Gaussian measure at least $1/2$, the prior covering estimate indeed implies that $2^{O(1/\delta)}(K \cap H)$ has Gaussian measure $2^{-O(n)}$, noting that shifting $2^{O(1/\delta)}(K \cap H)$ away from the origin only reduces Gaussian measure. Using an M-ellipsoid with appropriate regularity properties (as in [30]), one can scale $K \cap H$ by another $2^{O(1/\delta)}$ factor so that the preceding argument yields Gaussian measure at least $2^{-\delta n}$.

We now explain why the $(1-\delta)n$ longest axes of E are indeed long enough. By the covering estimates, for any $S \subseteq [n]$, $|S| = \delta n$, the sections E_S and K_S satisfy

$$\mathrm{vol}_{\delta n}(E_S)^{1/\delta n} \geq 2^{-O(1/\delta)} \mathrm{vol}_{\delta n}(K_S)^{1/\delta n} \geq 2^{-O(1/\delta)},$$

where the last inequality is by assumption. Using a form of the restricted invertibility principle for determinants (see e.g. the full version of [31]), one can show that if all coordinate sections of E of dimension δn have large volume, then so does every section of E of the same dimension. Precisely, one gets that

$$\min_{\dim(W)=\delta n} \mathrm{vol}_{\delta n}(E \cap W)^{1/\delta n} \geq \binom{n}{\delta n}^{-1/\delta n} \min_{|S|=\delta n} \mathrm{vol}_{\delta n}(E_S)^{1/\delta n} \geq 2^{O(-1/\delta)}.$$

In particular, the above implies that the geometric average of the lengths of the *shortest* δn axes of E (corresponding to the minimum volume section of E), must be at least $\sqrt{n}2^{-O(1/\delta)}$ since the ball of volume 1 in dimension δn has radius $\Omega(\sqrt{\delta n})$. But then, the longest $(1-\delta)n$ axes all have have length $\sqrt{n}2^{-O(1/\delta)}$. This completes the proof sketch.

C. The Discrepancy of Partial Colorings

Our analysis of Rothvoss's algorithm opens up the tantalizing possibility that it may indeed be optimal for finding partial colorings in a hereditary sense. More precisely, we conjecture that if when run on an instance U with norm ball K, the algorithm almost always produces partial colorings with K-discrepancy at least D, then there exists a subset of S of the columns of U such that every partial coloring of U_S has discrepancy $\Omega(D)$. The starting point for this conjecture is our upper bound of $O(\mathrm{volLB}^{\mathrm{h}}(U, K))$, on the discrepancy of the partial colorings the algorithm computes. It remains to show that the volume lower bound is also a lower bound on the hereditary discrepancy of *partial* colorings. We now provide a purely geometric conjecture, which would imply the above "hereditary optimality" for Rothvoss's algorithm.

As in the last subsection, we may assume that $U = (e_1, \dots, e_n)$ is the standard basis of $\mathbb{R}^n$ and that $\mathrm{volLB}((e_i)_{i=1}^n, K) = 1$. To prove the conjecture, it suffices to show that there exists some subset $S \subseteq [n]$ of coordinates, such that all partial colorings have K-discrepancy $\Omega(1)$. For concreteness, let us ask for partial colorings which color at least $|S|/2$ coordinates (the precise constant will not matter). For $x \in [-1,1]^n$, define $\mathrm{bounds}(x) = \{i \in [n] : x_i \in \{-1, 1\}\}$. With this notation, our goal is to find $S \subseteq [n]$, such that $\forall x \in [-1,1]^S$, $|\mathrm{bounds}(x)| \geq |S|/2$, $\|\sum_{i \in S} x_i e_i\|_K \geq \Omega(1)$.

We explain the candidate geometric obstruction to low discrepancy partial colorings, which is a natural generalization of the so-called spectral lower bound for ℓ_2 discrepancy. Assume that for some subset $S \subseteq [n]$, we have that

$$K_S \subseteq c\sqrt{|S|}B_2^S, \tag{6}$$

where $B_2^S := (B_2^n)_S$, for some constant $c > 0$. Since any partial coloring $x \in [-1,1]^S$, $|\mathrm{bounds}(x)| \geq |S|/2$, clearly has $\|x\|_2 \geq \sqrt{|S|/2}$, we must have that

$$\frac{1}{c\sqrt{2}} \leq \Big\|\sum_{i \in S} x_i e_i\Big\|_{c\sqrt{|S|}B_2^S} \leq \Big\|\sum_{i \in S} x_i e_i\Big\|_{K_S}. \tag{7}$$

In particular, every partial coloring on S has discrepancy at least $\frac{1}{c\sqrt{2}} = \Omega(1)$, as desired.

Given the above, we may now reduce the conjecture to the following natural geometric question:

Conjecture 3 (Restricted Invertibility for Convex Bodies)**.** *There exists an absolute constant $c \geq 1$, such that for any $n \in \mathbb{N}$ and symmetric convex body $K \subseteq \mathbb{R}^n$ of volume* *at most* 1*, there exists* $S \subseteq [n]$*,* $S \neq \emptyset$*, such that* $K_S \subseteq c\sqrt{|S|}B_2^S$.

To see that this indeed implies the required statement, note that if $\mathrm{volLB}((e_i)_{i=1}^n, K) = 1$, then by definition there exists $A \subseteq [n]$, $|A| \geq 1$, such that $\mathrm{vol}_{|A|}(K_A) \leq 1$. Now applying the above conjecture to K_A yields the desired result.

Two natural relaxations of the conjectures are to ask (1) does it hold for ellipsoids and (2) does it hold for general sections instead of coordinate sections? Our main evidence for this conjecture is that indeed both these statements are true. We note that (1) indeed implies the optimality of Rothvoss's partial coloring algorithm for ℓ_2 discrepancy. Our results here are slightly stronger than (1)+(2), as we in some sense manage to get "halfway there" with coordinates sections, by working with the M-ellipsoid, and only for the last step do we need to resort to general sections. We note that the above conjecture is closely related to the Bourgain-Tzafriri restricted invertibility principle [32], and indeed our proof for ellipsoids reduces to it.

D. A Factorization Approach for Vector Balancing.

While Theorem 2 gives an efficient and approximately optimal method of balancing a given set of vectors, it does not give an efficiently computable tight upper bound on the vector balancing constant or on hereditary discrepancy. Even though we proved that, after an appropriate scaling, the volume lower bound also gives an upper bound on the vector balancing constant, we are not aware of an efficient algorithm for computing the volume lower bound, which is itself a maximum over an exponential number of terms. To address this shortcoming, we study a different approach to vector balancing which relies on applying Banaszczyk's theorem in an optimal way in order to get an efficiently computable, and nearly tight, upper bound on both vector balancing constants and hereditary discrepancy.

Recall that Banaszczyk's vector balancing theorem states that if a body K has Gaussian measure at least $1/2$, then $\mathrm{vb}(B_2^n, K) \leq 5$. In order to apply the theorem to bodies K of small Gaussian measure, we can use rescaling. In particular, if r is the smallest number such that the Gaussian measure of rK is $\frac{1}{2}$, then the theorem tells us that $\mathrm{vb}(B_2^n, K) \leq 5r$. A natural way to use this upper bound for bodies C different from B_2^n is to find a mapping of C into B_2^n, and then use the theorem as above. As an illustration of this idea, let us see how we can get nearly tight bounds on $\mathrm{vb}(B_p^n, B_q^n)$ (the ℓ_p and ℓ_q balls) by applying Banaszczyk's theorem. Let us take an arbitrary sequence of points $u_1, \ldots, u_N \in B_p^n$, and rescale them to define new points $v_i := u_i / \max\{1, n^{1/2-1/p}\}$. The rescaled points $v_1, \ldots, v_N$ lie in B_2^n and we can apply Banaszczyk's theorem to them and the convex body $K := L\sqrt{q}n^{1/q}B_q^n$, which has Gaussian measure at least $\frac{1}{2}$ as long as we choose L to be a large enough constant. We get that there exist signs $\varepsilon_1, \ldots, \varepsilon_N \in \{-1, 1\}$ such that

$$\left\|\sum_{i=1}^N \varepsilon_i v_i\right\|_K \leq 5 \iff \left\|\sum_{i=1}^N \varepsilon_i u_i\right\|_q \leq 5L\sqrt{q}\max\{n^{1/q}, n^{1/q+1/2-1/p}\}.$$

In other words, we have that

$$\mathrm{vb}(B_p^n, B_q^n) \leq 5L\sqrt{q}\max\{n^{1/q}, n^{1/q+1/2-1/p}\}.$$

The volume lower bound (Lemmas 1) can be used to show that this bound is tight up to the $O(\sqrt{q})$ factor. Indeed one can show that B_p^n contains n vectors $u_1, \ldots u_n$ such that the matrix $U := (u_1, \ldots, u_n)$ has determinant $\det(U) \geq e^{-1}\max\{1, n^{1/2-1/p}\}$ (see [33] or [31]). By standard estimates, $\mathrm{vol}(B_q^n)^{1/n} \geq cn^{1/q}$ for an absolute constant $c > 0$. Plugging these estimates into Lemma 1 shows $\mathrm{vb}(B_p, B_q) \geq c' \max\{n^{1/q}, n^{1/q+1/2-1/p}\}$ for a constant $c' > 0$.

It is easy to see that, unlike the example above, in general simply rescaling C and K and applying Banaszczyk's theorem to the rescaled bodies may not give a tight bound on $\mathrm{vb}(C, K)$. However, we will show that we can get such tight bounds if we expand the class of transformations we allow on C and K from simple rescaling to arbitrary linear transformations. It turns out that the most convenient language for this approach is that of linear operators between normed spaces. We can generalize the notion of a vector balancing constant between a pair of convex bodies to arbitrary linear operators $U : X \to Y$ between two n-dimensional normed spaces X, with norm $\|\cdot\|_X$, and Y, with norm $\|\cdot\|_Y$, as follows

$$\mathrm{vb}(U) = \sup \min_{\varepsilon_1,\ldots,\varepsilon_N \in \{-1,1\}} \left\|\sum_{i=1}^N \varepsilon_i U(x_i)\right\|_Y, \quad (8)$$

where $B_X = \{x : \|x\|_X \leq 1\}$ is the unit ball of X, and the supremum is over positive integers N and sequences $x_1, \ldots, x_N \in B_X$. This definition is indeed a generalization of the geometric one. If C and K are two centrally symmetric convex bodies in $\mathbb{R}^n$, and we define the corresponding normed spaces $X_C = (\mathbb{R}^n, \|\cdot\|_C)$ and $X_K = (\mathbb{R}^n, \|\cdot\|_K)$, then the vector balancing constant $\mathrm{vb}(I)$ of the formal identity operator $I : X_C \to X_K$ recovers $\mathrm{vb}(C, K)$. However, the more abstract setting makes it plain that a simple rescaling is not the right approach to applying Banaszczyk's theorem to arbitrary norms: if X is an arbitrary norm, then X and B_2^n may not be defined on the same vector space, and rescaling B_X so that it is a subset of B_2^n does not even make sense. Instead, when dealing with general norms, it becomes very natural to embed B_X into B_2^n via a linear map $T : X \to \ell_2^n$ so that $T(B_X) \subseteq B_2^n$. Our approach is based on this idea, and, in particular, on choosing such a map T optimally.

To formalize the above, we use the ℓ-norm, which has been extensively studied in the theory of operator ideals, and in asymptotic convex geometry (see e.g. [34]–[36]). For a linear operator $S : \ell_2^n \to Y$ into an n-dimensional normed space Y with norm $\|\cdot\|_Y$, the ℓ-norm of S is defined as

$$\ell(S) := \left(\int \|S(x)\|_Y^2 d\gamma_n(x) \right)^{1/2},$$

where γ_n is the standard Gaussian measure on $\mathbb{R}^n$. I.e., if Z is a standard Gaussian random variable in $\mathbb{R}^n$, then $\ell(S) = (\mathbb{E}\|S(Z)\|_Y^2)^{1/2}$. It is direct to verify that $\ell(\cdot)$ is a norm on the space of linear operators from ℓ_2^n to Y, for any normed space Y as above. The reason the ℓ-norm is useful to us is the fact that the smallest r for which the set $K = \{x \in \mathbb{R}^n : \|Sx\|_Y \le r\}$ has Gaussian measure at least 1/2 is approximately $\ell(S)$, due to the concentration of measure phenomenon.

We now define our main tool: a factorization constant λ, which, for any two n-dimensional normed spaces X and Y and an operator $U : X \to Y$ is defined by

$$\lambda(U) := \inf\{\ell(S)\|T\| : T : X \to \ell_2^n,\ S : \ell_2^n \to Y, U = ST\}.$$

In other words, $\lambda(U)$ is the minimum of $\ell(S)\|T\|$ over all ways to factor U through ℓ_2^n as $U = ST$. Here $\|T\|$ is the operator norm, equal to $\max\{\|Tx\|_2/\|x\|_X\}$. This definition captures an optimal application of Banaszczyk's theorem. Using the theorem, it is not hard to show that $\mathrm{vb}(U) \le C\lambda(U)$ for an absolute constant C. Our main result is showing $\mathrm{vb}(U)$ and $\lambda(U)$ are in fact equal up to a factor which is polynomial in $\log n$. To prove this, we formulate $\lambda(U)$ as a convex minimization problem. Such a formulation is important both for our structural results, which rely on Lagrange duality, and also for giving an algorithm to compute $\lambda(U)$ efficiently, and, therefore, approximate $\mathrm{vb}(U)$ efficiently, which turns out to be sufficient to approximate hereditary discrepancy in arbitrary norms.

The most immediate way to formulate $\lambda(U)$ as an optimization problem is to minimize $\ell(UT^{-1})$ over operators $T : X \to \ell_2^n$ and subject to the constraint $\|T\| \le 1$. Unfortunately, this optimization problem is not convex in T: the value of the objective function is finite for any nonzero T, but infinite for $0 = \frac{1}{2}(T + (-T))$, for example. The key observation that allows us to circumvent this issue is that the objective function is completely determined by the operator $A := T^*T$, and is in fact convex in A. Here T^* is the dual operator of T. We use $f(A)$ to denote this objective function, i.e. to denote $\ell(UT^{-1})$ where T is an operator such that $T^*T = A$. In the full version we prove that this function is well-defined and convex. Then, our convex formulation of $\lambda(U)$ is

$$\begin{aligned} &\inf f(A) \\ &\text{s.t.} \\ &A : X \to X^*, \|A\| \le 1 \\ &A \succ 0. \end{aligned}$$

Above, X^* is the dual space of X, and $\|A\|$ is the operator norm. The first constraint is equivalent to the constraint $\|T\| \le 1$ where $U = ST$ is the factorization in the definition of $\lambda(U)$. The last constraint says that A should be positive definite, which is important so that A can be written as T^*T and $f(A)$ is well-defined.

We utilize this convex formulation and Lagrange duality to derive a dual formulation of $\lambda(U)$ as a supremum over "dual certificates". Such a formulation is useful in approximately characterizing $\mathrm{vb}(U)$ in terms of $\lambda(U)$ because it reduces our task to relating the dual certificates to the terms in the volume lower bound (2). If we can show that every dual certificate bounds from below one of the terms of the volume lower bound (up to factors polynomial in $\log n$), then we can conclude that $\lambda(U)$ also bounds the volume lower bound from below, and therefore $\mathrm{vb}(U)$ as well.

Before we can give the dual formulation, we need to introduce the dual norm ℓ^* of the ℓ-norm, defined via trace duality: for any linear operator $R : Y \to \ell_2^n$, let

$$\ell^*(R) := \sup\{\mathrm{tr}(RS) : S : \ell_2^n \to Y, \ell(S) \le 1\}.$$

The norms ℓ and ℓ^* form a dual pair, and in particular we have

$$\ell(S) = \sup\{\mathrm{tr}(RS) : R : Y \to \ell_2^n, \ell^*(R) \le 1\}.$$

For a finite dimensional space Y, both suprema above are achieved.

The derivation of our dual formulation uses standard tools, but is quite technical due to the complicated nature of the function $f(A)$. We give the formulation for norms X such that $B_X = \mathrm{conv}\{\pm x_1, \ldots, \pm x_m\}$. This is without loss of generality since every symmetric convex body can be approximated by a symmetric polytope (but has implications for the complexity of our algorithms). The dual formulation is as follows:

$$\begin{aligned} &\sup \mathrm{tr}((RU(\sum_{i=1}^m p_i x_i \otimes x_i)U^*R^*)^{1/3})^{3/2} \\ \text{s.t.}\quad &R : Y \to \ell_2^n, \ell^*(R) \le 1 \\ &\sum_{i=1}^m p_i = 1, \quad p_1, \ldots, p_m \ge 0. \end{aligned}$$

Above $x_i \otimes x_i$ is the rank-1 operator from the dual space X^* to X, given by $(x_i \otimes x_i)(x^*) = \langle x^*, x_i \rangle x_i$.

We relate the volume lower bound to this dual via deep inequalities between the ℓ^* and the ℓ norms (K-convexity), and between the ℓ norm and packing and covering numbers (Sudakov's minoration). Our second main result is the theorem below.

Theorem 4. *There exists a constant C such that for any two n-dimensional normed spaces X and Y, and any linear operator $U : X \to Y$ between them, we have*

$$\frac{1}{C} \leq \frac{\lambda(U)}{\mathrm{vb}(U)} \leq C(1 + \log n)^{5/2}.$$

Moreover, for any vectors $u_1, \ldots, u_N$ and convex body K in $\mathbb{R}^n$ we can define a norm X on $\mathbb{R}^n$ so that for the space Y with unit ball K and the identity map $I : X \to Y$,

$$\frac{\lambda(I)}{C(1 + \log n)^{5/2}} \leq \mathrm{hd}((u_i)_{i=1}^N, K) \leq \mathrm{vb}(I) \leq C\lambda(I).$$

Finally, $\lambda(U)$ is computable in polynomial time given appropriate access to X and Y. [1]

Acknowledgment

Daniel Dadush is supported by NWO Veni grant 639.071.510. Aleksandar Nikolov is supported by an NSERC Discovery Grant (RGPIN-2016-06333).

The authors would like to thank the American Institute of Mathematics for hosting a workshop on Hereditary Discrepancy and Factorization Norms, where work on this project began.

References

[1] J. Matoušek, *Geometric Discrepancy (An Illustrated Guide)*. Springer, 1999.

[2] B. Chazelle, *The Discrepancy Method*. Cambridge University Press, 1991.

[3] J. Spencer, "Six standard deviations suffice," *Trans. Amer. Math. Soc.*, vol. 289, pp. 679–706, 1985.

[4] E. D. Gluskin, "Extremal properties of orthogonal parallelepipeds and their applications to the geometry of banach spaces," *Mathematics of the USSR-Sbornik*, vol. 64, no. 1, p. 85, 1989.

[5] I. Bárány and V. Grinberg, "On some combinatorial questions in finite-dimensional spaces," *Linear Algebra and its Applications*, vol. 41, pp. 1–9, 1981.

[6] J. Beck and T. Fiala, "Integer-making theorems," *Discrete Applied Mathematics*, vol. 3, no. 1, pp. 1–8, 1981.

[7] J. Spencer, *Ten lectures on the probabilistic method*, 2nd ed., ser. CBMS-NSF Regional Conference Series in Applied Mathematics. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994, vol. 64. [Online]. Available: http://dx.doi.org/10.1137/1.9781611970074

[8] W. Banaszczyk, "Balancing vectors and gaussian measures of n-dimensional convex bodies," *Random Structures & Algorithms*, vol. 12, no. 4, pp. 351–360, 1998.

[1]See the full version of the paper for the necessary assumptions.

[9] A. Ghouila-Houri, "Caractérisation des matrices totalement unimodulaires," *C. R. Acad. Sci. Paris*, vol. 254, pp. 1192–1194, 1962.

[10] L. Lovász, J. Spencer, and K. Vesztergombi, "Discrepancy of set-systems and matrices," *European Journal of Combinatorics*, vol. 7, no. 2, pp. 151–160, 1986.

[11] R. Hoberg and T. Rothvoss, "A logarithmic additive integrality gap for bin packing," in *Proceedings of the Twenty-Eighth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, Philadelphia, PA, 2017, pp. 2616–2625. [Online]. Available: https://doi.org/10.1137/1.9781611974782.172

[12] T. Rothvoss, "The entropy rounding method in approximation algorithms," in *Symposium on Discrete Algorithms (SODA)*, 2012, pp. 356–372.

[13] N. Bansal and V. Nagarajan, "Approximation-friendly discrepancy rounding," in *IPCO*, 2016, pp. 375–386.

[14] N. Bansal, "Constructive algorithms for discrepancy minimization," in *51st Annual IEEE Symposium on Foundations of Computer Science (FOCS), 2010*. IEEE, 2010, pp. 3–10.

[15] S. Lovett and R. Meka, "Constructive discrepancy minimization by walking on the edges," *Arxiv preprint arXiv:1203.5747*, 2012.

[16] A. A. Giannopoulos, "On some vector balancing problems," *Studia Mathematica*, vol. 122, no. 3, pp. 225–234, 1997.

[17] T. Rothvoß, "Constructive discrepancy minimization for convex sets," in *55th IEEE Annual Symposium on Foundations of Computer Science, FOCS 2014, Philadelphia, PA, USA, October 18-21, 2014*. IEEE Computer Society, 2014, pp. 140–145. [Online]. Available: http://dx.doi.org/10.1109/FOCS.2014.23

[18] R. Eldan and M. Singh, "Efficient algorithms for discrepancy minimization in convex sets," *CoRR*, vol. abs/1409.2913, 2014.

[19] N. Bansal, D. Dadush, and S. Garg, "An algorithm for Komlós conjecture matching Banaszczyk's bound," in *Proceedings of the IEEE 57th Annual Symposium on Foundations of Computer Science — FOCS 2016*. IEEE, 2016, pp. 788–799.

[20] D. Dadush, S. Garg, S. Lovett, and A. Nikolov, "Towards a constructive version of Banaszczyk's vector balancing theorem," in *Proceedings of the 20th International Workshop on Randomization and Computation — RANDOM 2016*, ser. LIPIcs. Leibniz Int. Proc. Inform., vol. 60. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016, pp. Art. No. 28, 12.

[21] N. Bansal, D. Dadush, S. Garg, and S. Lovett, "The Gram-Schmidt Walk: a cure for the Banaszczyk blues," in *To appear in STOC 2018*, 2018.

[22] M. Charikar, A. Newman, and A. Nikolov, "Tight hardness results for minimizing discrepancy," in *SODA*, 2011, pp. 1607–1614.

[23] K. G. Larsen, "Constructive discrepancy minimization with hereditary l2 guarantees," Arxiv report 1711.02860, 2017.

[24] J. Matoušek, "The determinant bound for discrepancy is almost tight," 2011, http://arxiv.org/abs/1101.0767.

[25] A. Nikolov and K. Talwar, "Approximating hereditary discrepancy via small width ellipsoids," in *Proceedings of the Twenty-Sixth Annual ACM-SIAM Symposium on Discrete Algorithms*. SIAM, Philadelphia, PA, 2015, pp. 324–336. [Online]. Available: https://doi.org/10.1137/1.9781611973730.24

[26] J. Matoušek, A. Nikolov, and K. Talwar, "Factorization norms and hereditary discrepancy," *International Mathematics Research Notices*, p. rny033, 2018. [Online]. Available: http://dx.doi.org/10.1093/imrn/rny033

[27] W. Banaszczyk, "Balancing vectors and convex bodies," *Studia Math.*, vol. 106, no. 1, pp. 93–100, 1993.

[28] A. Newman, O. Neiman, and A. Nikolov, "Beck's three permutations conjecture: a counterexample and some consequences," in *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science—FOCS 2012*. IEEE Computer Soc., Los Alamitos, CA, 2012, pp. 253–262.

[29] V. D. Milman, "Inégalité de Brunn-Minkowski inverse et applications à la théorie locale des espaces normés," *C. R. Acad. Sci. Paris Sér. I Math.*, vol. 302, no. 1, pp. 25–28, 1986.

[30] G. Pisier, "A new approach to several results of V. Milman," *J. Reine Angew. Math.*, vol. 393, pp. 115–131, 1989. [Online]. Available: https://doi.org/10.1515/crll.1989.393.115

[31] A. Nikolov, "Randomized rounding for the largest simplex problem [extended abstract]," in *STOC'15—Proceedings of the 2015 ACM Symposium on Theory of Computing*. ACM, New York, 2015, pp. 861–870.

[32] J. Bourgain and L. Tzafriri, "Invertibility of large submatrices with applications to the geometry of banach spaces and harmonic analysis," *Israel journal of mathematics*, vol. 57, no. 2, pp. 137–224, 1987.

[33] K. Ball, "Volumes of sections of cubes and related problems," in *Geometric aspects of functional analysis (1987–88)*, ser. Lecture Notes in Math. Springer, Berlin, 1989, vol. 1376, pp. 251–260. [Online]. Available: https://doi.org/10.1007/BFb0090058

[34] N. Tomczak-Jaegermann, *Banach-Mazur distances and finite-dimensional operator ideals*, ser. Pitman Monographs and Surveys in Pure and Applied Mathematics. Longman Scientific & Technical, Harlow; copublished in the United States with John Wiley & Sons, Inc., New York, 1989, vol. 38.

[35] G. Pisier, *The volume of convex bodies and Banach space geometry*, ser. Cambridge Tracts in Mathematics. Cambridge University Press, Cambridge, 1989, vol. 94. [Online]. Available: http://dx.doi.org/10.1017/CBO9780511662454

[36] S. Artstein-Avidan, A. Giannopoulos, and V. D. Milman, *Asymptotic geometric analysis. Part I*, ser. Mathematical Surveys and Monographs. American Mathematical Society, Providence, RI, 2015, vol. 202. [Online]. Available: https://doi.org/10.1090/surv/202