Towards a Constructive Version of Banaszczyk's Vector Balancing Theorem

Daniel Dadush*[1], Shashwat Garg†[2], Shachar Lovett‡[3], and Aleksandar Nikolov[4]

1 **Centrum Wiskunde & Informatica, Amsterdam.**
`dadush@cwi.nl`
2 **Department of Mathematics and Computer Science, Eindhoven University of Technology.**
`s.garg@tue.nl`
3 **University of California, San Diego.**
`slovett@cs.ucsd.edu`
4 **University of Toronto, Toronto.**
`anikolov@cs.toronto.edu`

Abstract

An important theorem of Banaszczyk (Random Structures & Algorithms '98) states that for any sequence of vectors of ℓ_2 norm at most $1/5$ and any convex body K of Gaussian measure $1/2$ in $\mathbb{R}^n$, there exists a signed combination of these vectors which lands inside K. A major open problem is to devise a constructive version of Banaszczyk's vector balancing theorem, i.e. to find an efficient algorithm which constructs the signed combination.

We make progress towards this goal along several fronts. As our first contribution, we show an equivalence between Banaszczyk's theorem and the existence of $O(1)$-subgaussian distributions over signed combinations. For the case of symmetric convex bodies, our equivalence implies the existence of a *universal* signing algorithm (i.e. independent of the body), which simply samples from the subgaussian sign distribution and checks to see if the associated combination lands inside the body. For asymmetric convex bodies, we provide a novel *recentering procedure*, which allows us to reduce to the case where the body is symmetric.

As our second main contribution, we show that the above framework can be efficiently implemented when the vectors have length $O(1/\sqrt{\log n})$, recovering Banaszczyk's results under this stronger assumption. More precisely, we use random walk techniques to produce the required $O(1)$-subgaussian signing distributions when the vectors have length $O(1/\sqrt{\log n})$, and use a stochastic gradient ascent method to implement the recentering procedure for asymmetric bodies.

1998 ACM Subject Classification G.3 Probability and Statistics – Probablistic Algorithms

Keywords and phrases Discrepancy, Vector Balancing, Convex Geometry

Digital Object Identifier 10.4230/LIPIcs.APPROX-RANDOM.2016.

1 Introduction

Given a family of sets $S_1, \dots, S_m$ over a universe $U = [n]$, the goal of combinatorial discrepancy minimization is to find a bi-coloring $\chi : U \to \{-1, 1\}$ such that the discrepancy,

* Supported by the NWO Veni grant 639.071.510.
† Supported by the Netherlands Organisation for Scientific Research (NWO) under project no. 022.005.025.
‡ Supported by an NSF CAREER award 1350481 and a Sloan fellowship

i.e. the maximum imbalance, $\max_{j\in[m]} |\sum_{i\in S_j} \chi(i)|$ is made as small as possible. Discrepancy theory, where discrepancy minimization plays a major role, has a rich history of applications in computer science as well as mathematics, and we refer the reader to [19, 10, 11] for a general exposition.

A beautiful question regards the discrepancy of sparse set systems, i.e. set systems in which each element appears in at most t sets. A classical theorem of Beck and Fiala [8] gives an upper bound of $2t-1$ in this setting. They also conjectured an $O(\sqrt{t})$ bound, which if true would be tight. An improved Beck-Fiala bound of $2t - \log^* t$ was given by Bukh [9], where $\log^* t$ is the iterated logarithm function in base 2. Recently, it was shown by Ezra and Lovett [14] that a bound of $O(\sqrt{t \log t})$ holds with high probability when $m \geq n$ and each element is assigned to t sets uniformly at random. The best general bounds having sublinear dependence in t currently depend on n or m. Srinivasan [25] used Beck's *partial coloring method* [7] to give a bound of $O(\sqrt{t} \log \min\{n, m\})$. Using techniques from convex geometry, Banaszczyk [2] proved a general result on vector balancing (stated below) which implies an $O(\sqrt{t \log \min\{n, m\}})$ bound.

The proofs of both Srinivasan's and Banaszczyk's bounds were non-constructive, that is, they provided no efficient algorithm to construct the guaranteed colorings, short of exhaustive enumeration. In the last 6 years, tremendous progress has been made on the question of matching classical discrepancy bounds algorithmically. Currently, essentially all discrepancy bounds proved using the partial coloring method, including Srinivasan's, have been made constructive [4, 18, 15, 22, 13]. Constructive versions of Banaszczyk's result have, however, proven elusive until very recently. In recent work [5], the first and second named authors jointly with Bansal gave a constructive algorithm for recovering Banaszczyk's bound in the Beck-Fiala setting as well as the more general Komlós setting. However, finding a constructive version of Banaszczyk's more general vector balancing theorem, which has further applications in approximating hereditary discrepancy, remains an open problem. This theorem is stated as follows:

▶ **Theorem 1** (Banaszczyk [2]). *Let $v_1, \dots, v_n \in \mathbb{R}^m$ satisfy $\|v_i\|_2 \leq 1/5$. Then for any convex body $K \subseteq \mathbb{R}^m$ of Gaussian measure at least $1/2$, there exists $\chi \in \{-1, 1\}^n$ such that $\sum_{i=1}^n \chi_i v_i \in K$.*

The lower bound $1/2$ on the Gaussian measure of K is easily seen to be tight. In particular, if all the vectors are equal to 0, we must have that $0 \in K$. If we allow Gaussian measure $< 1/2$, then $K = \{x \in \mathbb{R}^n : x_1 \geq \varepsilon\}$, for $\varepsilon > 0$ small enough, is a clear counterexample. On the other hand, it is not hard to see that if K has Gaussian measure $1/2$ then $0 \in K$. Otherwise, there exists a halfspace H containing K but not 0, where H clearly has Gaussian measure less than $1/2$.

Banaszczyk's theorem gives the best known bound for the notorious Komlós conjecture [24], a generalization of the Beck-Fiala conjecture, which states that for any sequence of vectors $v_1, \dots, v_n \in \mathbb{R}^m$ of ℓ_2 norm at most 1, there exists $\chi \in \{-1, 1\}^n$ such that $\|\sum_{i=1}^n \chi_i v_i\|_\infty$ is a constant independent of m and n. In this context, Banaszczyk's theorem gives a bound of $O(\sqrt{\log m})$, because an $O(\sqrt{\log m})$ scaling of the unit ball of ℓ_∞^m has Gaussian measure $1/2$. Banaszczyk's theorem together with estimates on the Gaussian measure of slices of the ℓ_∞^m ball due to Barthe, Guedon, Mendelson, and Naor [6] give a bound of $O(\sqrt{\log d})$, where $d \leq \min\{m, n\}$ is the dimension of the span of $v_1, \dots, v_n$. A well-known reduction (see e.g. Lecture 9 in [24]), shows that this bound for the Komlós problem implies an $O(\sqrt{t \log \min\{m, n\}})$ bound in the Beck-Fiala setting.

While the above results only deal with the case of K being a cube, Banaszczyk's theorem has also been applied to other cases. It was used in [3] to give the best known bound on the Steinitz conjecture. In this problem, the input is a set of vectors $v_1, \ldots, v_n$ in $\mathbb{R}^m$ of norm at most one and summing to 0. The aim is to find a permutation $\pi : [n] \to [n]$ to minimise the maximum sum prefix of the vectors rearranged according to π i.e. to minimize $\max_{k\in[n]} \|\sum_{i=1}^{k} v_{\pi(i)}\|$. The Steinitz conjecture is that this bound should always be $O(\sqrt{m})$, irrespective of the number of vectors, and using the vector balancing theorem Banaszczyk proved a bound of $O(\sqrt{m} + \sqrt{\log n})$ for ℓ_2 norm.

More recently, Banaszczyk's theorem was applied to more general symmetric polytopes in Nikolov and Talwar's approximation algorithm [21] for a hereditary notion of discrepancy. Hereditary discrepancy is defined as the maximum discrepancy of any restriction of the set system to a subset of the universe. In [21] it was shown that an efficiently computable quantity, denoted γ_2, bounds hereditary discrepancy from above and from below for any given set system, up to polylogarithmic factors. For the upper bound they used Banaszczyk's theorem for a natural polytope associated with the set system. However, since there is no known algorithmic version of Banaszczyk's theorem for a general body, it is not known how to efficiently compute colorings that achieve the discrepancy upper bounds in terms of γ_2. The recent work on algorithmic bounds in the Komlós setting does not address this more general problem.

Banaszczyk's proof of Theorem 1 follows an ingenious induction argument, which folds the effect of choosing the sign of v_n into the body K. The first observation is that finding a point of the set $\sum_{i=1}^{n} \{-v_i, v_i\}$ inside K is equivalent to finding a point of $\sum_{i=1}^{n-1} \{-v_i, v_i\}$ in $K - v_n \cup K + v_n$. Inducting on this set is not immediately possible because it may no longer be convex. Instead, Banaszczyk shows that a convex subset $K * v_n$ of $(K - v_n) \cup (K + v_n)$ has Gaussian measure at least that of K, as long as K has measure at least $1/2$, which allows him to induct on $K * v_n$. In the base case, he needs to show that a convex body of Gaussian measure at least $1/2$ must contain the origin, but this fact follows easily from the hyperplane separation theorem, as indicated above. While extremely elegant, Banaszczyk's proof can be seen as relatively mysterious as it does not seem to provide any tangible insights as to what the colorings look like.

1.1 Our Results

As our main contribution, we help demystify Banaszczyk's theorem, by showing that it is equivalent, up to a constant factor in the length of the vectors, to the existence of certain subgaussian coloring distributions. Using this equivalence, as our second main contribution, we give an efficient algorithm that recovers Banaszczyk's theorem up to a $O(\sqrt{\log \min\{m, n\}})$ factor for all convex bodies. This improves upon the best previous algorithms of Rothvoss [22], Eldan and Singh [13], which only recover the theorem for symmetric convex bodies up to a $O(\log \min\{m, n\})$ factor.

As a major consequence of our equivalence, we show that for any sequence $v_1, \ldots, v_n \in \mathbb{R}^m$ of short enough vectors there exists a probability distribution $\chi \in \{-1, 1\}^n$ over colorings such that, for *any symmetric convex body* $K \subseteq \mathbb{R}^m$ of Gaussian measure at least $1/2$, the random variable $\sum_{i=1}^{n} \chi_i v_i$ lands inside K with probability at least $1/2$. Importantly, if such a distribution can be efficiently sampled, we immediately get a *universal sampler* for constructing Banaszczyk colorings for all symmetric convex bodies (we remark that the recent work of [5] constructs a more restricted form of such distributions). Using random walk techniques, we show how to implement an approximate version of this sampler efficiently, which guarantees the same conclusion when the vectors are of length $O(1/\sqrt{\log \min\{m, n\}})$. We provide more details on these results in Section 2

To extend our results to asymmetric convex bodies, we develop a novel *recentering* *procedure* and a corresponding efficient implementation which allows us to reduce the asymmetric setting to the symmetric one. After this reduction, a slight extension of the aforementioned sampler again yields the desired colorings. We note that our recentering procedure in fact depends on the target convex body, and hence our algorithms are no longer universal in this setting. We provide more details on these results in Section 3.

Interestingly, we additionally show that this procedure can be extended to yield a completely different coloring algorithm, i.e. not using the sampler, achieving the same $O(\sqrt{\log \min\{m,n\}})$ approximation factor. Surprisingly, the coloring outputted by this procedure is deterministic (its implementation however is not) and has a natural analytic description, which may be of independent interest.

Before we continue with a more detailed description of our results, we begin with some terminology and a well-known reduction. Given a set of vectors $v_1, \dots, v_n \in \mathbb{R}^m$, we shall call a property *hereditary* if it holds for all subsets of the vectors. We note that Banaszczyk's vector balancing bounds restricted to a set of vectors are hereditary, since a bound on the maximum ℓ_2 norm of the vectors is hereditary. We shall say that a property of colorings holds in the *linear setting*, if when given any shift $t \in \sum_{i=1}^n [-v_i, v_i] \stackrel{\text{def}}{=} \{\sum_{i=1}^n \lambda_i v_i : \lambda \in [-1,1]^n\}$, one can find a coloring (or distribution on colorings) $\chi \in \{-1,1\}^n$ such that $\sum_{i=1}^n \chi_i v_i - t$ satisfies the property. It is well-known that Banaszczyk's theorem also extends by standard arguments to the linear setting after reducing the ℓ_2 norm bound from $1/5$ to $1/10$ (a factor 2 drop). This follows, for example, from the general inequality between hereditary and linear discrepancy proved by Lovasz, Spencer, and Vesztergombi [16].

All the results in this work will in fact hold in the linear setting. When treating the linear setting, it is well known that one can always reduce to the case where the vectors $v_1, \dots, v_n$ are linearly independent, and in our setting, when $m = n$. In particular, assume we are given some shift $t \in \sum_{i=1}^n [-v_i, v_i]$ and that $v_1, \dots, v_n$ are *not* linearly independent. Then, using a standard linear algebraic technique, we can find a "fractional coloring" $x \in [-1,1]^n$ such that $\sum_{i=1}^n x_i v_i = t$, and the vectors $(v_i : i \in A_x)$ are linearly independent, where $A_x \stackrel{\text{def}}{=} \{i : x_i \in (-1,1)\}$ is the set of fractional coordinates (see Lecture 5 in [24], or Chapter 4 in [19]). We can think of this as a reduction to coloring the linearly independent vectors indexed by A_x. Specifically, given x as above, define the lifting function $L_x : [-1,1]^{A_x} \to [-1,1]^n$ by

$$L_x(z)_i = \begin{cases} z_i & : i \in A_x \\ x_i & : i \in [n] \setminus A_x \end{cases}, \ \forall i \in [n] \ . \tag{1}$$

This map takes any coloring $\chi \in \{-1,1\}^{A_x}$ and "lifts" it to a full coloring $L_x(\chi) \in \{-1,1\}^n$. It also satisfies the property that $L_x(\chi) - t = \sum_{i \in A_x} \chi_i v_i - \sum_{i \in A_x} x_i v_i$. So, if we can find a coloring $\chi \in \{-1,1\}^{A_x}$ such that $\sum_{i \in A_x} \chi_i v_i - \sum_{i \in A_x} x_i v_i \in K$, then we would have $L_x(\chi) - t \in K$ as well. Moreover, if we define W as the span of $(v_i : i \in A_x)$, then $\sum_{i \in A_x} \chi_i v_i - \sum_{i \in A_x} x_i v_i \in K$ if and only if $\sum_{i \in A_x} \chi_i v_i - \sum_{i \in A_x} x_i v_i \in K \cap W$, so we can replace K with $K \cap W$, and work entirely inside W. For convex bodies K with Gaussian measure at least $1/2$, the central section $K \cap W$ has Gaussian measure that is at least as large, so we have reduced the problem to the case of $|A_x|$ linearly independent vectors in an $|A_x|$-dimensional space (details are given in the full version of this paper.) We shall thus, for simplicity, state all our results in the setting where the vectors $v_1, \dots, v_n$ are in $\mathbb{R}^n$ and are linearly independent.

2 Symmetric Convex Bodies and Subgaussian Distributions

In this section, we detail the equivalence of Banaszczyk's theorem restricted to symmetric convex bodies with the existence of certain subgaussian distributions. We begin with the main theorem of this section, which we note holds in a more general setting than Banaszczyk's result.

▶ **Theorem 2** (Main Equivalence). *Let $T \subseteq \mathbb{R}^n$ be a finite set. Then, the following parameters are equivalent up to a universal constant factor independent of T and n:*

1. *The minimum $s_b > 0$ such that for any symmetric convex body $K \subseteq \mathbb{R}^n$ of Gaussian measure at least $1/2$, we have that $T \cap s_b K \neq \emptyset$.*
2. *The minimum $s_g > 0$ such that there exists an s_g-subgaussian random variable Y supported on T.*

We recall that a random vector $Y \in \mathbb{R}^n$ is s-subgaussian, or subgaussian with parameter s, if for any unit vector $\theta \in S^{n-1}$ and $t \geq 0$, $\Pr[|\langle Y, \theta \rangle| \geq t] \leq 2e^{-(t/s)^2/2}$. In words, Y is subgaussian if all its 1-dimensional marginals satisfy the same tail bound as the 1-dimensional Gaussian of mean 0 and standard deviation s.

To apply the above to discrepancy, we set $T = \sum_{i=1}^n \{-v_i, v_i\}$, i.e. all signed combinations of the vectors $v_1, \dots, v_n \in \mathbb{R}^n$. In this context, Banaszczyk's theorem directly implies that $s_b \leq 5 \max_{i \in [n]} \|v_i\|_2$, and hence by our equivalence that $s_g = O(1) \max_{i \in [n]} \|v_i\|_2$. Furthermore, the above extends to the linear setting letting $T = \sum_{i=1}^n \{-v_i, v_i\} - t$, for $t \in \sum_{i=1}^n [-v_i, v_i]$, because, as mentioned above, Banaszczyk's theorem extends to this setting as well.

The existence of the *universal sampler* claimed in the previous section is in fact the proof that $s_b = O(s_g)$ in the above Theorem. In particular, it follows directly from the following lemma.

▶ **Lemma 3.** *Let $Y \in \mathbb{R}^n$ be an s-subgaussian random variable. There exists an absolute constant $c > 0$, such for any symmetric convex body $K \subseteq \mathbb{R}^n$ of Gaussian measure at least $1/2$, $\Pr[Y \in s \cdot cK] \geq 1/2$.*

Here, if Y is the s_g-subgaussian distribution supported on $\sum_{i=1}^n \{-v_i, v_i\} - t$ as above, we simply let χ denote the random variable such that $Y = \sum_{i=1}^n \chi_i v_i - t$. That χ now yields the desired universal distribution on colorings is exactly the statement of the lemma.

As a consequence of the above, we see that to recover Banaszczyk's theorem for symmetric convex bodies, it suffices to be able to efficiently sample from an $O(1)$-subgaussian distribution over sets of the type $\sum_{i=1}^n \{-v_i, v_i\} - t$, for $t \in \sum_{i=1}^n [-v_i, v_i]$, when $v_1, \dots, v_n \in \mathbb{R}^n$ are linearly independent and have ℓ_2 norm at most 1. Here we rely on homogeneity, that is, if Y is an s-subgaussian random variable supported on $\sum_{i=1}^n \{-v_i, v_i\} - t$ then αY is αs-subgaussian on $\sum_{i=1}^n \{-\alpha v_i, \alpha v_i\} - \alpha t$, for $\alpha > 0$.

The proof of Lemma 3 follows relatively directly from well-known convex geometric estimates combined with Talagrand's majorizing measure theorem [26] (see also [27]), which gives a powerful characterization of the supremum of any Gaussian process.

Unfortunately, Lemma 3 does not hold for asymmetric convex bodies. In particular, if $Y = -e_1$, the negated first standard basis vector, and $K = \{x \in \mathbb{R}^n : x_1 \geq 0\}$, the conclusion is clearly false no matter how much we scale K, even though Y is $O(1)$-subgaussian and K has Gaussian measure $1/2$. One may perhaps hope that the conclusion still holds if we ask for either Y or $-Y$ to be in $s \cdot cK$ in the asymmetric setting, though we do not know how to prove this. We note however that this only makes sense when the support of Y is symmetric, which does not necessarily hold in the linear discrepancy setting.

We now describe the high level idea of the proof for the reverse direction, namely, that $s_g = O(s_b)$. For this purpose, we show that the existence of a $O(s_b)$-subgaussian distribution on T can be expressed as a two player zero-sum game, i.e. the first player chooses a distribution on T and the second player tries to find a non-subgaussian direction. Here the value of the game will be small if and only if the $O(s_b)$-subgaussian distribution exists. To bound the value of the game, we show that an appropriate "convexification" of the space of subgaussianity tests for the second player can be associated with symmetric convex bodies of Gaussian measure at least 1/2. From here, we use von Neumann's minimax principle to switch the first and second player, and deduce that the value of the game is bounded using the definition of s_b.

2.1 The Random Walk Sampler

From the algorithmic perspective, it turns out that subgaussianity is a very natural property in the context of random walk approaches to discrepancy minimization. Our results can thus be seen as a good justification for the random walk approaches to making Banaszczyk's theorem constructive.

At a high level, in such approaches one runs a random walk over the coordinates of a "fractional coloring" $\chi \in [-1,1]^n$ until all the coordinates hit either 1 or -1. The steps of such a walk usually come from Gaussian increments (though not necessarily spherical), which try to balance the competing goals of keeping discrepancy low and moving the fractional coloring χ closer to $\{-1,1\}^n$. Since a sum of small centered Gaussian increments is subgaussian with the appropriate parameter, it is natural to hope that the output of a correctly implemented random walk is subgaussian. Our main result in this setting is that this is indeed possible to a limited extent, with the main caveat being that the walk's output will not be "subgaussian enough" to fully recover Banaszczyk's theorem.

▶ **Theorem 4.** *Let $v_1, \ldots, v_n \in \mathbb{R}^n$ be vectors of ℓ_2 norm at most 1 and let $t \in \sum_{i=1}^n [-v_i, v_i]$. Then, there is an expected polynomial time algorithm which outputs a random coloring $\chi \in \{-1,1\}^n$ such that the random variable $\sum_{i=1}^n \chi_i v_i - t$ is $O(\sqrt{\log n})$-subgaussian.*

To achieve the above sampler, we guide our random walk using solutions to the so-called vector Kómlos program, whose feasibility was first given by Nikolov [20], and show subgaussianity using well-known martingale concentration bounds. Interestingly, the random walk's analysis does not rely on phases, and is instead based on a simple relation between the walk's convergence time and the subgaussian parameter. As an added bonus, we also give a new and simple constructive proof of the feasibility of the vector Kómlos program which avoids the use of an SDP solver.

Given the results of the previous section, the above random walk is a universal sampler for constructing the following colorings.

▶ **Corollary 5.** *Let $v_1, \ldots, v_n \in \mathbb{R}^n$ be vectors of ℓ_2 norm at most 1, let $t \in \sum_{i=1}^n [-v_i, v_i]$, and let $K \subseteq \mathbb{R}^n$ be a symmetric convex body of Gaussian measure 1/2 (given by a membership oracle). Then, there is an expected polynomial time algorithm which outputs a coloring $\chi \in \{-1,1\}^n$ such that $\sum_{i=1}^n \chi_i v_i - t \in O(\sqrt{\log n})K$.*

As mentioned previously, the best previous algorithms in this setting are due to Rothvoss [22], Eldan and Singh [13], which find a signed combination inside $O(\log n)K$. Furthermore, these algorithms are not universal, i.e. they heavily depend on the body K. We note that these algorithms are in fact tailored to find *partial colorings* inside a symmetric convex body K of

Gaussian measure at least 2^{-cn}, for $c > 0$ small enough, a setting in which our sampler does not provide any guarantees.

We now recall prior work on random walk based discrepancy minimization. The random walk approach was pioneered by Bansal [4], who used a semidefinite program to guide the walk and gave the first efficient algorithm matching the classic $O(\sqrt{n})$ bound of Spencer [23] for the combinatorial discrepancy of set systems satisfying $m = O(n)$. Later, Lovett and Meka [18] provided a greatly simplified walk, removing the need for the semidefinite program, which recovered the full power of Beck's entropy method for constructing partial colorings. Harvey, Schwartz, and Singh [15] defined another random walk based algorithm, which, unlike previous work and similarly to our algorithm, doesn't explicitly use phases or produce partial colorings. The random walks of [18] and [15] both depend on the convex body K; the walk in [18] is only well-defined in a polytope, while the one in [15] remains well-defined in any convex body, although the analysis still applies only to the polyhedral setting. Most directly related to this paper is the recent work [5], which gives a walk that can be viewed as a randomized variant of the original $2t - 1$ Beck-Fiala proof. This walk induces a distribution $\chi \in \{-1, 1\}^n$ on colorings for which *each coordinate* of the output $\sum_{i=1}^n \chi_i v_i$ is $O(1)$-subgaussian. From the discrepancy perspective, this gives a sampler which finds colorings inside any axis parallel box of Gaussian measure at least $1/2$ (and their rotations, though not in a universal manner), matching Banaszczyk's result for this class of convex bodies.

3 Asymmetric Convex Bodies

In this section, we explain how our techniques extend to the asymmetric setting. The main difficulty in the asymmetric setting is that one cannot hope to increase the Gaussian mass of an asymmetric convex body by simply scaling it. In particular, if we take $K \subseteq \mathbb{R}^n$ to be a halfspace through the origin, e.g. $\{x \in \mathbb{R}^n : x_1 \geq 0\}$, then K has Gaussian measure exactly $1/2$ but $sK = K$ for all $s > 0$. At a technical level, the lack of any measure increase under scaling breaks the proof of Lemma 3, which is crucial for showing that subgaussian coloring distributions produce combinations that land inside K.

The main idea to circumvent this problem will be to reduce to a setting where the mass of K is "symmetrically distributed" about the origin, in particular, when the barycenter of K under the induced Gaussian measure is at the origin. For such a body K, we show that a constant factor scaling of $K \cap -K$ also has Gaussian mass at least $1/2$, yielding a direct reduction to the symmetric setting.

To achieve this reduction, we will use a novel *recentering procedure*, which will both carefully fix certain coordinates of the coloring as well as shift the body K to make its mass more "symmetrically distributed". The guarantees of this procedure are stated below:

▶ **Theorem 6** (Recentering Procedure). *Let $v_1, \ldots, v_n \in \mathbb{R}^n$ be linearly independent, $t \in \sum_{i=1}^n [-v_i, v_i]$, and $K \subseteq \mathbb{R}^n$ be a convex body of Gaussian measure at least $1/2$. Then, there exists a fractional coloring $x \in [-1, 1]^n$, such that for $p = \sum_{i=1}^n x_i v_i - t$, $A_x = \{i \in [n] : x_i \in (-1, 1)\}$ and $W = \mathrm{span}(v_i : i \in A_x)$, the following holds:*

1. *$p \in K$.*
2. *The Gaussian measure of $(K - p) \cap W$ on W is at least the Gaussian measure of K.*
3. *The barycenter of $(K - p) \cap W$ is at the origin, i.e. $\int_{(K-p)\cap W} y e^{-\|y\|^2/2} dy = 0$.*

By convention, if the procedure returns a full coloring $x \in \{-1, 1\}^n$ (in which case, since $p \in K$, we are done), we shall treat conditions 2 and 3 as satisfied, even though $W = \{0\}$. At a high level, the recentering procedure allows us to reduce the initial vector balancing problem to one in a possibly lower dimension with respect to "well-centered" convex body of no smaller Gaussian measure, and in particular, of Gaussian measure at least $1/2$. Interestingly, as mentioned earlier in the introduction, the recentering procedure can also be extended to yield a full coloring algorithm. We explain the high level details of its implementation together with this extension in the next subsection.

To explain how to use the fractional coloring x from Theorem 6 to get a useful reduction, recall the lifting function $L_x : [-1,1]^{A_x} \to [-1,1]^n$ defined in (1). We reduce the initial vector balancing problem to the problem of finding a coloring $\chi \in \{-1,1\}^{A_x}$ such that $\sum_{i \in A_x} \chi_i v_i - \sum_{i \in A_x} x_i v_i \in (K-p) \cap W$ (note that $\sum_{i \in A_x} \chi_i v_i - \sum_{i \in A_x} x_i v_i \in W$ by construction). Then we can lift this coloring to $L_x(\chi)$, which satisfies

$$\sum_{i \in A_x} \chi_i v_i - \sum_{i \in A_x} x_i v_i \in (K-p) \cap W \Leftrightarrow \sum_{i=1}^{n} L_x(\chi)_i v_i - t \in K.$$

From here, the guarantee that $K' \stackrel{\text{def}}{=} (K-p) \cap W$ has Gaussian measure at least $1/2$ and barycenter at the origin allows a direct reduction to the symmetric setting. Namely, we can replace K' by the symmetric convex body $K' \cap -K'$ without losing "too much" of the Gaussian measure of K'. This is formalized by the following extension of Lemma 3, which directly implies a reduction to subgaussian sampling as in section 2.

▶ **Lemma 7.** *Let $Y \in \mathbb{R}^n$ be an s-subgaussian random variable. There exists an absolute constant $c > 0$, such for any convex body $K \subseteq \mathbb{R}^n$ of Gaussian measure at least $1/2$ and barycenter at the origin,* $\Pr[Y \in s \cdot c(K \cap -K)] \geq 1/2$.

In particular, if there exists a distribution over colorings $\chi \in \{-1,1\}^{A_x}$ such that $\sum_{i \in A_x} \chi_i v_i - \sum_{i \in A_x} x_i v_i$ as above is $1/c$-subgaussian, Lemma 7 implies that the random signed combination lands inside K' with probability at least $1/2$. Thus, the asymmetric setting can be effectively reduced to the symmetric one, as claimed.

Crucially, the recentering procedure in Theorem 6 can be implemented in probabilistic polynomial time if one relaxes the barycenter condition from being exactly 0 to having "small" norm. Furthermore, the estimate in Lemma 7 will be robust to such perturbations. Thus, to constructively recover the colorings in the asymmetric setting, it will still suffice to be able to generate good subgaussian coloring distributions.

Combining the sampler from Theorem 4 together with the recentering procedure, we constructively recover Banaszczyk's theorem for general convex bodies up to a $O(\sqrt{\log n})$ factor.

▶ **Theorem 8** (Weak Constructive Banaszczyk)**.** *There exists a probabilistic polynomial time algorithm which, on input a linearly independent set of vectors $v_1, \ldots, v_n \in \mathbb{R}^n$ of ℓ_2 norm at most $c/\sqrt{\log n}$, $c > 0$ small enough, $t \in \sum_{i=1}^n [-v_i, v_i]$, and a (not necessarily symmetric) convex body $K \subseteq \mathbb{R}^n$ of Gaussian measure at least $1/2$ (given by a membership oracle), computes a coloring $\chi \in \{-1,1\}^n$ such that with high probability $\sum_{i=1}^n \chi_i v_i - t \in K$.*

As far as we are aware, the above theorem gives the first algorithm to recover Banaszczyk's result for asymmetric convex bodies under any non-trivial restriction. In this context, we note that the algorithm of Eldan and Singh [13] finds "relaxed" partial colorings, i.e. where the fractional coordinates of the coloring are allowed to fall outside $[-1,1]$, and the resulting vector lies inside an n-dimensional convex body of Gaussian measure at least 2^{-cn}. However, it is unclear how one could use such partial colorings to recover the above result, even with a larger approximation factor.

3.1 The Recentering Procedure

In this section, we describe the details of the recentering procedure as well as its extension to full colorings, which produces deterministic colorings matching the guarantees of Theorem 8. We provide only its abstract instantiation here, leaving a detailed description of its implementation to the full version of the paper.

Before we begin, we give a more geometric view of the vector balancing problem and the recentering procedure, which help clarify the exposition. Let $v_1, \dots, v_n \in \mathbb{R}^n$ be linearly independent vectors and $t \in \sum_{i=1}^n [-v_i, v_i]$. Given the target body $K \subseteq \mathbb{R}^n$ of Gaussian measure at least $1/2$, we can restate the vector balancing problem geometrically as that of finding a vertex of the parallelepiped $P = \sum_{i=1}^n [-v_i, v_i] - t$ lying inside K. Here, the choice of t ensures that $0 \in P$. Note that this condition is necessary, since otherwise there exists a halfspace separating P from 0 having Gaussian measure at least $1/2$.

Recall now that in the linear setting, and using this geometric language, Banaszczyk's theorem implies that if P contains the origin, and $\max_{i\in[n]} \|v_i\|_2 \leq 1/10$ (which we do not need to assume for the validity of the recentering procedure), then any convex body of Gaussian measure at least $1/2$ contains a vertex of P. Thus, for our given target body K, we should make our situation better by replacing P and K by $P - q$ and $K - q$ respectively, if $q \in P$ is a shift such that $K - q$ has higher Gaussian measure than K. In particular, given the symmetry of Gaussian measure, one would intuitively expect that if the Gaussian mass of K is not symmetrically distributed around 0, there should be a shift of K which increases its Gaussian measure.

In the current language, fixing a color $\chi_i \in \{-1, 1\}$ for vector v_i, corresponds to restricting ourselves to finding a vertex in the facet $F = \chi_i v_i + \sum_{j\neq i}[-v_j, v_j] - t$ of P lying inside K. Again intuitively, restricting to a facet of P should improve our situation if the Gaussian measure of the corresponding slice of K in the lower dimension is larger than that of K. To make this formal, note that when inducting on a facet F of P (which is an $n-1$ dimensional parallelepiped), we must choose a center $q \in F$ to serve as the new origin in the lower dimensional space. Precisely, this can be expressed as inducting on the parallelepiped $F - q$ and shifted slice $(K - q) \cap \mathrm{span}(F - q)$ of K, using the $n - 1$ dimensional Gaussian measure on $\mathrm{span}(F - q)$.

With the above viewpoint, one can restate the goal of the recentering procedure as that of finding a point $q \in P \cap K$, such that smallest face F of P containing q, satisfies that $(K - q) \cap \mathrm{span}(F - q)$ has its barycenter at the origin and Gaussian measure no smaller than that of K. Recall that as long as $(K - q) \cap \mathrm{span}(F - q)$ has Gaussian measure at least $1/2$, we are guaranteed that $0 \in K - q \Rightarrow q \in K$. With this geometry in mind, we implement the recentering procedure as follows:

Compute $q \in P$ so that the Gaussian mass of $K - q$ is maximized. If q is on the boundary of P, letting F denote a facet of P containing q, induct on $F - q$ and the slice $(K - q) \cap \mathrm{span}(F - q)$ as above. If q is in the interior of P, replace P and K by $P - q$ and $K - q$, and terminate.

We now explain why the above achieves the desired result. First, if the maximizer q is in a face F of P, then a standard convex geometric argument reveals that the Gaussian measure of $(K - q) \cap \mathrm{span}(F - q)$ is no smaller than that of $K - q$, and in particular, no smaller than that of K. Thus, in this case, the recentering procedure fixes a color for "free". In the second case, if q is in the interior of P, then a variational argument gives that the barycenter of $K - q$ under the induced Gaussian measure must be at the origin, namely, $\int_{K-q} xe^{-x^2/2}dx = 0$.

To conclude this section, we explain how to extend the recentering procedure to directly produce a deterministic coloring satisfying Theorem 8. For this purpose, we shall assume that $v_1, \ldots, v_n$ have length at most $c/\sqrt{\log n}$, for a small enough constant $c > 0$. To begin, we run the recentering procedure as above, which returns P and K, with K having its barycenter at the origin. We now replace P, K by a joint scaling $\alpha P, \alpha K$, for $\alpha > 0$ a large enough constant, so that αK has Gaussian mass at least $3/4$. At this point, we run the original recentering procedure again with the following modification: every time we get to the situation where K has its barycenter at the origin, induct on the facet of P closest to the origin. More precisely, in this situation, compute a point p on the boundary of P closest to the origin, and, letting F denote the facet containing p, induct on $F - p$ and $(K - p) \cap \mathrm{span}(F - p)$. At the end, return the final found vertex.

Notice that, as claimed, the coloring (i.e. vertex) returned by the algorithm is indeed deterministic. The reason the above algorithm works is the following. While we cannot guarantee, as in the original recentering procedure, that the Gaussian mass of $(K - p) \cap \mathrm{span}(F - p)$ does not decrease, we can instead show that it decreases only *very slowly*. In particular, we use the bound of $O(1/\sqrt{\log n})$ on the length of the vectors $v_1, \ldots, v_n$ to show that every time we induct, the Gaussian mass drops by at most a $1 - c/n$ factor. More generally, if the vectors had length at most $d > 0$, for d small enough, the drop would be of the order $1 - ce^{-1/(cd)^2}$, for some constant $c > 0$. Since we "massage" K to have Gaussian mass at least $3/4$ before applying the modified recentering algorithm, this indeed allows to induct n times while keeping the Gaussian mass above $1/2$, which guarantees that the final vertex is in K. To derive the bound on the rate of decrease of Gaussian mass, we prove a new inequality on the Gaussian mass of sections of a convex body near the barycenter, which may be of independent interest. The new inequality is stated below:

▶ **Theorem 9.** *Let $K \subseteq \mathbb{R}^n$ be a convex body with Gaussian measure $\gamma_n(K) = \alpha \geq 3/5$ such that its barycenter, b satisfies $\|b\|_2 \leq \eta$, for $\eta > 0$ small enough. For $\theta \in S^{n-1}$ and $a \in \mathbb{R}$, let $K_a^\theta = (K - a\theta) \cap \{x \in \mathbb{R}^n : \langle \theta, x \rangle = 0\}$. Then, there exists universal constants $a_0, c > 0$, such that for $|a| \leq a_0$, we have that*

$$\gamma_{n-1}(K_a^\theta) \geq (\alpha - c\eta)(1 - \frac{e^{-\frac{1}{100a^2}}}{4\sqrt{2\pi}}).$$

As a final remark, we note that unlike the subgaussian sampler, the recentering procedure is not scale invariant. Namely, if we jointly scale P and K by some factor α, the output of the recentering procedure will not be an α-scaling of the output on the original K and P, as Gaussian measure is not homogeneous under scalings. Thus, one must take care to appropriately normalize P and K before applying the recentering procedure to achieve the desired results.

We now give the high level overview of our recentering step implementation. The first crucial observation in this context, is that the task of finding $t \in P$ maximizing the Gaussian measure of $K - t$ is in fact a *convex program.* More precisely, the objective function (Gaussian measure of $K - t$) is a logconcave function of t and the feasible region P is convex. Hence, one can hope to apply standard convex optimization techniques to find the desired maximizer.

It turns out however, that one can significantly simplify the required task by noting that the recentering strategy does not in fact necessarily need an exact maximizer, or even a maximizer in P. To see this, note that if p is a shift such that $K - p$ has larger Gaussian measure than K, then by logconcavity the shifts $K - \alpha p$, $0 < \alpha \leq 1$, also have larger Gaussian measure. Thus, if a we find a shift $p \notin P$ with larger Gaussian measure, letting αp be the intersection point with the boundary ∂P, we can induct on the facet of $P - \alpha p$ containing 0 and the corresponding slice of $K - \alpha p$ just as before. Given this, we can essentially "ignore" the constraint $p \in P$ and the treat the optimization problem as unconstrained.

This last observation will allow us to use the following simple gradient ascent strategy. Precisely, we simply take steps in the direction of the gradient until either we pass through a facet of P or the gradient becomes "too small". As alluded to previously, the gradient will exactly equal a fixed scaling of the barycenter of $K - p$, p the current shift, under the induced Gaussian measure. Thus, once the gradient is small, the barycenter will be very close to the origin, which will be good enough for our purposes. The last nontrivial technical detail is how to efficiently estimate the barycenter, where we note that the barycenter is the expectation of a random point inside $K - p$. For this purpose, we simply take an average of random samples from $K - p$, where we generate the samples using standard random walk samplers for logconcave distributions over convex bodies [1, 17, 12].

Conclusion and Open Problems

In conclusion, we have shown a tight connection between the existence of subgaussian coloring distributions and Banaszczyk's vector balancing theorem. Furthermore, we make use of this connection to constructively recover a weaker version of this theorem. The main open problem we leave is thus to fully recover Banaszczyk's result. As explained above, this reduces to finding a distribution on colorings such that the output random signed combination is $O(1)$-subgaussian, when the input vectors have ℓ_2 norm at most 1. We believe this approach is both attractive and feasible, especially given the recent work [5], which builds a distribution on colorings for which each coordinate of the output random signed combination is $O(1)$-subgaussian.

Acknowledgments

We would like to thank the American Institute for Mathematics for hosting a recent workshop on discrepancy theory, where some of this work was done.

References

1. David Applegate and Ravi Kannan. Sampling and integration of near log-concave functions. In *Proceedings of the 23rd annual ACM symposium on Theory of Computing*, pages 156–163, 1991.
2. Wojciech Banaszczyk. Balancing vectors and Gaussian measures of n-dimensional convex bodies. *Random Structures Algorithms*, 12(4):351–360, 1998.
3. Wojciech Banaszczyk. On series of signed vectors and their rearrangements. *Random Struct. Algorithms*, 40(3):301–316, 2012. URL: `http://dx.doi.org/10.1002/rsa.20373`, `doi:10.1002/rsa.20373`.
4. Nikhil Bansal. Constructive algorithms for discrepancy minimization. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science FOCS 2010*, pages 3–10. IEEE Computer Soc., Los Alamitos, CA, 2010.
5. Nikhil Bansal, Daniel Dadush, and Shashwat Garg. An algorithm for Komlós conjecture matching Banaszczyk's bound. To appears in FOCS, 2016.
6. Franck Barthe, Olivier Guédon, Shahar Mendelson, and Assaf Naor. A probabilistic approach to the geometry of the l_p^n-ball. *Ann. Probab.*, 33(2):480–513, 2005.
7. József Beck. Roth's estimate of the discrepancy of integer sequences is nearly sharp. *Combinatorica*, 1(4):319–325, 1981.

8 József Beck and Tibor Fiala. "Integer-making" theorems. *Discrete Appl. Math.*, 3(1):1–8, 1981.

9 Boris Bukh. An improvement of the Beck-Fiala theorem. *CoRR*, abs/1306.6081, 2013.

10 Bernard Chazelle. *The discrepancy method*. Cambridge University Press, Cambridge, 2000. Randomness and complexity.

11 William Chen, Anand Srivastav, Giancarlo Travaglini, et al. *A Panorama of Discrepancy Theory*, volume 2107. Springer, 2014.

12 Ben Cousins and Santosh Vempala. A cubic algorithm for computing Gaussian volume. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 1215–1228. ACM, New York, 2014.

13 Ronen Eldan and Mohit Singh. Efficient algorithms for discrepancy minimization in convex sets. CoRR, abs/1409.2913, 2014.

14 Esther Ezra and Shachar Lovett. On the Beck-Fiala conjecture for random set systems. *Electronic Colloquium on Computational Complexity (ECCC)*, 22:190, 2015.

15 Nicholas J. A. Harvey, Roy Schwartz, and Mohit Singh. Discrepancy without partial colorings. In *Approximation, randomization, and combinatorial optimization*, volume 28 of *LIPIcs. Leibniz Int. Proc. Inform.*, pages 258–273. Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2014.

16 L. Lovász, J. Spencer, and K. Vesztergombi. Discrepancy of set-systems and matrices. *European J. Combin.*, 7(2):151–160, 1986. URL: `http://dx.doi.org/10.1016/S0195-6698(86)80041-5`, `doi:10.1016/S0195-6698(86)80041-5`.

17 László Lovász and Santosh Vempala. Fast algorithms for logconcave functions: Sampling, rounding, integration and optimization. In *Proceedings of the 47th Annual IEEE Symposium on Foundations of Computer Science*, pages 57–68, 2006.

18 Shachar Lovett and Raghu Meka. Constructive discrepancy minimization by walking on the edges. *SIAM J. Comput.*, 44(5):1573–1582, 2015. Preliminary version in FOCS 2012.

19 Jiří Matoušek. *Geometric discrepancy*, volume 18 of *Algorithms and Combinatorics*. Springer-Verlag, Berlin, 1999. An illustrated guide.

20 Aleksandar Nikolov. The Komlós conjecture holds for vector colorings. arXiv preprint arXiv:1301.4039, 2013.

21 Aleksandar Nikolov and Kunal Talwar. Approximating hereditary discrepancy via small width ellipsoids. In *Symposium on Discrete Algorithms, SODA*, pages 324–336, 2015.

22 Thomas Rothvoss. Constructive discrepancy minimization for convex sets. In *55th Annual IEEE Symposium on Foundations of Computer Science—FOCS 2014*, pages 140–145. IEEE Computer Soc., Los Alamitos, CA, 2014.

23 Joel Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289(2):679–706, 1985.

24 Joel Spencer. *Ten lectures on the probabilistic method*, volume 52. SIAM, 1987.

25 Aravind Srinivasan. Improving the discrepancy bound for sparse matrices: better approximations for sparse lattice approximation problems. In *Proceedings of the Eighth Annual ACM-SIAM Symposium on Discrete Algorithms (New Orleans, LA, 1997)*, pages 692–701. ACM, New York, 1997.

26 Michel Talagrand. Regularity of Gaussian processes. *Acta Math.*, 159(1-2):99–149, 1987. URL: `http://dx.doi.org/10.1007/BF02392556`, `doi:10.1007/BF02392556`.

27 Michel Talagrand. *Upper and lower bounds for stochastic processes*, volume 60 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer, Heidelberg, 2014. Modern methods and classical problems. URL: `http://dx.doi.org/10.1007/978-3-642-54075-2`, `doi:10.1007/978-3-642-54075-2`.