View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Springer - Publisher Connector

Math. Program., Ser. B (2012) 134:5–22
DOI 10.1007/s10107-012-0546-7

FULL LENGTH PAPER

Semidefinite optimization in discrepancy theory

Nikhil Bansal

Received: 13 October 2011 / Accepted: 16 April 2012 / Published online: 17 May 2012
© The Author(s) 2012. This article is published with open access at Springerlink.com

Abstract Recently, there have been several new developments in discrepancy theory based on connections to semidefinite programming. This connection has been useful in several ways. It gives efficient polynomial time algorithms for several problems for which only non-constructive results were previously known. It also leads to several new structural results in discrepancy itself, such as tightness of the so-called determinant lower bound, improved bounds on the discrepancy of the union of set systems and so on. We will give a brief survey of these results, focussing on the main ideas and the techniques involved.

Keywords Semidefinite optimization · Discrepancy theory · Rounding error · Algorithms

Mathematics Subject Classification 90 · 05

1 Introduction

The field of combinatorial discrepancy deals with the following type of question. There is a set-system $(V, \mathcal{S})$ specified by the elements $V = \{1, \ldots, n\}$ and a collection of subsets $\mathcal{S} = \{S_1, \ldots, S_m\}$ of V. Find a red-blue coloring of V such that each set in $\mathcal{S}$ is colored as evenly as possible.

Formally, let us use -1 and $+1$ to denote the colors red and blue. Then, given a set-system $(V, \mathcal{S})$, the discrepancy of a coloring $\mathcal{X} : V \to \{-1, +1\}^n$ is defined as $\mathrm{disc}_{\mathcal{X}}(\mathcal{S}) = \max_{j \in [m]} |\mathcal{X}(S_j)|$, where $\mathcal{X}(S) = \sum_{i \in S} \mathcal{X}(i)$ for any $S \subseteq V$.

N. Bansal (✉)
Eindhoven University of Technology, Eindhoven, The Netherlands
e-mail: n.bansal@tue.nl

The discrepancy of the system $(V, \mathcal{S})$ is the minimum discrepancy over all possible colorings, i.e.

$$\mathrm{disc}(\mathcal{S}) = \min_{\mathcal{X}} \max_{j \in [m]} |\mathcal{X}(S_j)|.$$

Another very useful concept is that of *hereditary* discrepancy of a set system $(V, \mathcal{S})$, which is defined as the maximum value of discrepancy over all restrictions W of V. Specifically, given $W \subseteq V$, let $\mathcal{S}_{|W}$ denote the collection $\{S \cap W : S \in \mathcal{S}\}$. Then, the hereditary discrepancy of $(V, \mathcal{S})$ is defined as

$$\mathrm{herdisc}(\mathcal{S}) = \max_{W \subseteq V} \mathrm{disc}(\mathcal{S}_{|W}).$$

We will identify a set-system $(V, \mathcal{S})$ with its incidence matrix A, and interchangeably use $\mathrm{disc}(A)$ and $\mathrm{herdisc}(A)$ to denote $\mathrm{disc}(\mathcal{S})$ and $\mathrm{herdisc}(\mathcal{S})$.

The concept of discrepancy is intimately related to several fundamental questions in mathematics and theoretical computer science. Roughly speaking, the discrepancy and hereditary discrepancy of a set system are useful measures of its inherent complexity, and understanding them can give several important insights. In computer science for example, discrepancy has a range of applications in topics such as probabilistic and approximation algorithms, computational geometry, numerical integration, derandomization, communication complexity, machine learning, data structures and so on. It is far beyond the scope of this article to discuss these connections and applications here and we refer the interested reader to several excellent books [10,12,21] on the subject. We state below one application that arises often in combinatorial optimization and approximation algorithms.

1.1 Application to rounding

Suppose we are given a fractional solution $x \in \mathbb{R}^n$, to a linear system $Ax = b$ on n variables and m constraints. How well can we round x to an integral solution $\tilde{x}$ such that the error in each of the m equations is as low as possible, i.e. find $\tilde{x} \in \mathbb{Z}^n$ that minimizes $|A(x - \tilde{x})|_\infty$.

It turns out that the answer to this question is closely related to the hereditary discrepancy of the matrix[1] A. In particular, Lovász et al. [18] showed that,

Theorem 1 ([18]) *There is a rounding $\tilde{x}$ of x (i.e. $\tilde{x} \in \mathbb{Z}^n$ with $|\tilde{x} - x|_\infty < 1$) such that $|A(x - \tilde{x})|_\infty \leq \mathrm{herdisc}(A)$*

This bound has been further improved to $(1 - 1/2m)\mathrm{herdisc}(A)$ by Doerr [11].

It is useful to contrast this with the commonly used randomized rounding technique. Recall that in randomized rounding, each x_i is rounded independently to either $\lfloor x_i \rfloor$ with probability $\lceil x_i \rceil - x_i$ or to $\lceil x_i \rceil$ otherwise. If A is a $\{0, 1\}$ matrix, then a standard application of probabilistic tail bounds gives an error bound of $O((n \log m)^{1/2})$ for

[1] While we only defined hereditary discrepancy for incidence matrices of set systems, the definition extends to arbitrary matrices in the natural way.

randomized rounding. This follows as for each row i, the expected value of $A_i\tilde{x}$ is b_i and by standard probabilistic tail bounds $\Pr[|A_i\tilde{x} - b_i| \geq c\sqrt{n}] = \exp(-\Omega(c^2))$. Choosing $c = \Omega((\log m)^{1/2})$ and taking the union bound over the m rows gives the claimed bound. Simple examples also show that randomized rounding cannot do better in general.

Interestingly, it turns out that Theorem 1 (together with the known results for discrepancy) always implies a better guarantee. Let us first consider the case when A is the incidence matrix of a general set system. In his celebrated "six standard deviations suffice" result, Spencer [26] showed that

Theorem 2 ([26]) *Any set system with $m = n$ sets has discrepancy is at most $6\sqrt{n}$. More generally, for $m > n$, the discrepancy is $O((n \log(m/n))^{1/2})$.*

Together with Theorem 1, this implies a rounding with error $O((n \log(m/n))^{1/2})$ (vs. $O((n \log m)^{1/2})$ for randomized rounding). Spencer's result is tight up to constant factors, and there exist set systems for which one cannot do better. We shall study Spencer's result in more detail later in Sect. 3. While Theorem 2 guarantees the existence of a good coloring, no efficient (polynomial time) algorithm was known until recently to find such a coloring. In fact, as far as polynomial time algorithms are concerned, nothing better than randomized rounding was known for rounding a general matrix A. This was also true for Theorem 1, i.e. while a good rounding is guaranteed to exist, in most cases it was unclear how to do anything better than randomized rounding.

If the matrix A has more structure and better bounds can be shown on the hereditary discrepancy, then of course Theorem 1 implies better guarantees. One particularly interesting case is that of bounded degree set systems, where each element lies in a bounded number of sets. Beck and Fiala [6] showed that

Theorem 3 ([6]) *If A is a set system where each element lies in at most t sets, then* $\mathrm{disc}(A) \leq 2t - 1$.

A slightly improved bound of $2t - 3$ is also known [7]. However, it is believed that this bound can be improved and a well-known conjecture of Beck and Fiala is that the discrepancy of such systems is $O(\sqrt{t})$. If a dependence on n is allowed, bounds of $O(\sqrt{t} \log n)$ due to Srinivasan [27] and $O(\sqrt{t \log n})$ due to Banaszczyk [2] are also known.

We remark here that Theorem 1 also has several other useful consequences. For example, together with the result that totally unimodular (TU) matrices have hereditary discrepancy at most 1 [15], this implies the well known result that if A is TU and b is integral, then the system $Ax = b$ has as integral solution.

1.2 Bounding discrepancy

Upper bounds: Given the various applications, there has been a lot of work on bounding the discrepancy of various types of set systems, and several techniques have been developed for this purpose. One of the most powerful and widely used tools here is the so-called Partial coloring method due to [5] and its refinements [26] based on the

so-called entropy method (we will discuss this in Sect. 3). For example, both Theorem 2 above and the $O(\sqrt{t}\log n)$ bound [27] for bounded-degree systems are based on the entropy method.

It turns out however that both the partial coloring method and the entropy method seem inherently non-constructive. In particular, they are based on a clever application of the pigeonhole principle to the space of all 2^n possible colorings (we recommend the unfamiliar reader to see for e.g. [1]), and only prove the existence of a low discrepancy coloring, without giving any algorithmic insight on how to find them efficiently. This is problematic for applications such as Theorem 1 above.

Note that Theorems 2 and 3 above only give worst-case bounds on discrepancy for specific classes of set systems. A question that arises naturally is whether the discrepancy of a given set system can be computed or approximated efficiently. Not much was known here until the recent work of Charikar et al. [9], who showed that discrepancy is very hard to approximate in general. Among other results, they show that it is NP-Hard to distinguish whether the discrepancy of a set system on $m = O(n)$ sets is 0 or $\Omega(\sqrt{n})$.

Lower bounds: A variety of techniques have also been developed for proving lower bounds on discrepancy. Many of these are based on deep results and connections to various areas of mathematics (see for e.g. [10,21]). One of the strongest known results in this direction is the following *determinant lower bound* due to Lovász et al. [18]. For a real matrix A, define $\mathrm{detlb}(A) := \max_k \max_B |\det B|^{1/k}$, where the maximum is over all $k \times k$ submatrices B of A. For a set system $\mathcal{S}$, we will also use $\mathrm{detlb}(\mathcal{S})$ to denote $\mathrm{detlb}(A)$ where A is the incidence matrix of $\mathcal{S}$.

Theorem 4 ([18]) *For every set system* $\mathcal{S}$, $\mathrm{herdisc}(\mathcal{S}) \geq \frac{1}{2}\mathrm{detlb}(\mathcal{S})$.

Lovász et al. [18] conjectured that herdisc might also be upper bounded by some function of detlb, i.e., $\mathrm{herdisc}(\mathcal{S}) \leq f(\mathrm{detlb}(\mathcal{S}))$ for all $\mathcal{S}$. But, this was refuted by Hoffman who gave an elegant example (see for e.g. [21]) of a set system where $\mathrm{detlb}(\mathcal{S}) = O(1)$ and $\mathrm{disc}(\mathcal{S}) \approx (\log n)/(\log\log n)$. Recently, Pálvölgyi [24] gave an improved example with an $\Omega(\log n)$ gap for the same. However, until recently, there was no non-trivial result upper bound on the gap between the determinant lower bound and the hereditary discrepancy.

1.3 Some recent results

Recently, several important advances have been made in discrepancy theory. First, many previous results that were based on the entropy method and hence non-constructive, can now be made algorithmic. In particular, Bansal [3] showed the following algorithmic version of Spencer's result.

Theorem 5 [3] *For any set system* $\mathcal{S}$ *on* $m = O(n)$ *sets,*[2] *there is a randomized polynomial time algorithm to find a coloring with discrepancy* $O(\sqrt{n})$.

[2] For $m \gg n$, the algorithm of [3] gives the bound $O(\sqrt{n}\log(m/n))$, which has a worse dependence on (m/n) than in Spencer's result.

This result is based on a new and general method for rounding SDPs and several other results based on the entropy method (such as Srinivasan's result [27]) can also be similarly made constructive. This method for rounding SDPs also gives an algorithm to find a low discrepancy coloring for systems with low hereditary discrepancy. In particular,

Theorem 6 ([3]) *For any set system $\mathcal{S}$ on m sets and n elements, there is a randomized polynomial time algorithm to find a coloring with discrepancy $O((\log m \log n)^{1/2} \cdot \mathrm{herdisc}(\mathcal{S}))$.*

Theorem 6 implies the following algorithmic version of Theorem 1.

Corollary 1 *Given any fractional solution $x \in \mathbb{R}^n$ to the linear system $Ax = b$ on m equations, there is a polynomial time algorithm to round x to an integral solution $\tilde{x} \in \mathbb{Z}^n$ such that $|A(x - \tilde{x})|_\infty = O((\log m \log n)^{1/2} \cdot \mathrm{herdisc}(A))$.*

In another surprising result based on this connection, Matoušek [19] gave the first non-trivial upper bound on the gap between the determinant lower bound and hereditary discrepancy.

Theorem 7 ([19]) *For any set system $(V, \mathcal{S})$, $\mathrm{herdisc}(S) \leq O(\log n \sqrt{\log m}) \cdot \mathrm{detlb}(S)$.*

This result is remarkably tight, as shown by the examples of Hoffman and Pálvölgyi mentioned previously. Among other things, this result also implies the following new structural results about discrepancy [19]: for any two set systems $(V, \mathcal{S}_1)$ and $(V, \mathcal{S}_2)$,

$$\mathrm{herdisc}(\mathcal{S}_1 \cup \mathcal{S}_2) \leq \max(\mathrm{herdisc}(\mathcal{S}_1), \mathrm{herdisc}(\mathcal{S}_2)) \cdot O(\log n \sqrt{\log m}).$$

Previously such a result was known only for the special case when $\mathcal{S}_2$ consists of a single set [17]! A further extension of this result to union of t sets can be found in [19].

In addition to these results, there have been several other recent developments. Eisenbrand et al. [13] made a surprising connection between discrepancy and the well known bin-packing problem in combinatorial optimization. In particular, they show that an $O(1)$ discrepancy bound for the three permutation problem, which was a long standing conjecture of Beck, would imply an $O(1)$ additive integrality gap for a natural linear programming formulation for bin packing when all items have size more than $1/4$. Interestingly, the three permutation conjecture itself was recently disproved by Newman and Nikolov [23]. Recently, Rothvoss [25] showed how the entropy method can also be used in the design of approximation algorithms. His technique for example, gives an alternate discrepancy based proof of the celebrated $O(\log^2 n)$ additive guarantee of Karmarkar and Karp for bin-packing. Even though it does not improve the bound for general bin-packing, his technique seems interesting as it gives improved results for several cases where previous techniques do not seem to help.

Recently, Chandrasekharan and Vempala [8] have shown how ideas from discrepancy, and in particular a variant of Theorem 5, can be useful in integer programming.

Main ideas: The algorithmic results mentioned above are based on rounding the natural SDP relaxation for discrepancy. However, there are some important new ideas. First, instead of using a single SDP solution and rounding it (as is typically done in other problems), these algorithms produce the solution gradually over time by considering several SDPs. Indeed, the hardness result of Charikar et al. [9] rules out the possibility of finding a good coloring based on rounding a single SDP relaxation. To see this, recall that Charikar et al. [9] show that it is NP-Hard to distinguish whether the discrepancy of a set system is 0 or $\Omega(\sqrt{n})$ is NP-Hard. This implies that there exist instances with discrepancy $\Omega(\sqrt{n})$ for which the SDP relaxation returns the value 0. So, any rounding procedure applied to this SDP solution must lose an additive $\Omega(\sqrt{n})$ term.

The second interesting aspect of the algorithm is that the non-constructive entropy method itself is used crucially in its design. In particular, the various SDPs used over time by the algorithm are determined by the entropy method.

The proof of Theorem 7 is based on SDP duality, and a clever connection between the sub-determinants of the incidence matrix and a dual SDP solution via the use of eigenvalues.

1.4 Organization

The rest of the article is organized as follows. First in Sect. 2, we prove Theorem 6. This is a relatively simple result and illustrates most of the algorithmic ideas involved. Then in Sect. 3, we show how these algorithmic ideas can be refined and combined with the entropy method to obtain Theorem 5. However, instead of describing the optimum $O(\sqrt{n})$ bound, we will show a (slightly) weaker bound of $O((n \log\log\log n)^{1/2})$. This weaker bound illustrates all the ideas involved without getting into the tedious calculations required for the $O(\sqrt{n})$ bound. Finally, in Sect. 4, we will prove Matoušek's result on the determinant lower bound.

2 Systems with low hereditary discrepancy

In this section we prove Theorem 6. We will describe the algorithm completely, but only sketch the analysis. The interested reader can find the details in the recent book by Gärtner and Matoušek [14] or in [3].

Before we begin, let us quickly recall some basic facts about semi-definite programming. A semi-definite program (SDP) can be viewed as an arbitrary linear program on variables of the form $Y_{ij} = \langle w_i, w_j \rangle$ where w_i, for $1 \le i \le n$, are vectors and $\langle , \rangle$ denotes the standard inner product. Any such SDP can be solved to arbitrary accuracy in time polynomial in the input size, provided the dimension m of the vectors w_i is allowed to be arbitrarily large (in general m can be as large as n). For more details about semidefinite programming (see [28]).

The natural SDP relaxation for the problem of finding a ± 1 coloring with discrepancy at most λ is the following.

$$\left\| \sum_{i \in S_j} v_i \right\|_2^2 \leq \lambda^2 \quad \text{for each set} \quad S_j \tag{1}$$

$$\|v_i\|_2^2 = 1 \tag{2}$$

Here, as usual, $\|v\|_2 = (\langle v, v \rangle)^{1/2}$ denotes the length of v. The first constraint says that the discrepancy of each set S_j must be at most λ, and the second constraint says that each v_i must be a unit vector. Clearly, this is a valid relaxation as any ± 1 coloring with discrepancy λ is a feasible solution. We will call any feasible solution to this SDP, a *vector-coloring* for $\mathcal{S}$, and the smallest value of λ for which it is feasible as the vector discrepancy of $\mathcal{S}$, denoted by $\text{vecdisc}(\mathcal{S})$.

Before we describe our algorithm, let us try the most natural approach that does not quite work, and see what we can use from it. First, we can assume that the algorithm knows the value of λ, as it can try all values $0, 1, \ldots, n$ and pick the smallest λ for which the SDP is feasible. Now let us consider some vector-coloring v_i obtained by solving this SDP. In this solution, the vectors v_i are nicely correlated such that for any set S_j, the vector $\sum_{i \in S_j} v_i$ has length at most λ. Our goal then is to convert these vectors v_i into the numbers ± 1 without increasing $\sum_{i \in S_j} v_j$ too much. So we can first try to convert v_i into real numbers (that are hopefully not too far from ± 1) without substantially violating the sums $\sum_{i \in S_j} v_j$. This seems particularly promising as the correlations among the v_i's are preserved upon taking projections. In particular, for an arbitrary vector $g \in \mathbb{R}^n$, the real numbers $y_i = \langle g, v_i \rangle$ satisfy

$$\sum_{i \in S_j} y_i = \sum_{i \in S_j} \langle g, v_i \rangle = \left\langle g, \sum_{i \in S_j} v_i \right\rangle \leq \|g\|_2 \cdot \left\| \sum_{i \in S_j} v_i \right\|_2,$$

implying that $\sum_{i \in S_j} y_i$ is likely to be small if $\| \sum_{i \in S_j} v_i \|$ is small.

To this end, let us take g as a random Gaussian vector in $\mathbb{R}^n$, i.e. each coordinate of g is chosen independently from the standard Gaussian distribution $N(0, 1)$, with mean 0 and variance 1. The crucial property of such a g is the following:

Lemma 1 *For a random Gaussian $g \in \mathbb{R}^n$, and an arbitrary vector $v \in \mathbb{R}^n$, the random variable $\langle g, v \rangle$ is (a real number) distributed as $N(0, \|v\|_2^2)$, i.e. as a Gaussian with mean 0, and variance $\|v\|_2^2$.*

Proof Recall that if g_1 and g_2 are two independent $N(0, 1)$ random variables, then the sum $ag_1 + bg_2$ is distributed as $N(0, a^2 + b^2)$. As $\langle g, v \rangle = \sum_i g(i)v(i)$, where $g(i)$ and $v(i)$ denote the ith coordinate of g and v, and as the $g(i)'s$ are independent, $\langle g, v \rangle$ is distributed as $N(0, \sum_i v(i)^2) = N(0, ||v||_2^2)$.

In our situation, choosing a random Gaussian g and setting $y_i = \langle g, v_i \rangle$, Lemma 1 implies that (i) Each y_i is distributed as $N(0, 1)$. This follows as $\|v_i\|_2^2 = 1$, and (ii) For each j, the discrepancy $\sum_{i \in S_j} y_i$ is distributed as $N(0, \leq \lambda^2)$, i.e. as a gaussian with mean 0 and variance at most λ^2. This follows as $\| \sum_{i \in S_j} v_i \|^2 \leq \lambda^2$.

This seems quite close to what we would like. As $y_i \sim N(0,1)$, we have that $y_i/(c(\log n)^{1/2}) \in [-1,1]$ with high probability (for some large enough constant c), and that the y_i's are typically bounded away from 0. Moreover, for any j, the discrepancy $|\sum_{i \in S_j} y_i| = O(\lambda(\log n)^{1/2})$ with high probability. Perhaps, one could now hope to round these y_i's to ± 1 without increasing the discrepancy substantially. Unfortunately it turns out that this cannot be done (recall that the hardness result of [9] implies that there are instances with discrepancy $\Omega(\sqrt{n})$ but vector-discrepancy 0, which rules out the existence of the desired rounding procedure).

So, we adopt a different approach. Instead of trying to round the y_i's into ± 1 and incurring large discrepancy, we will obtain a ± 1 solution by combining several different collections of correlated $y_i's$. This is also where we will really use the upper bound of λ on the hereditary discrepancy.

Main idea: As mentioned above, instead of trying to obtain a coloring using a single SDP solution, we will gradually produce a solution by using several SDPs over time. At time 0, we start with the "empty" coloring $x^0 = (0, \ldots, 0)$ where each element is colored 0. We slowly modify it over time as follows: suppose x^{t-1} denotes the coloring of elements at time $t-1$, we obtain the coloring x^t by adding a small perturbation vector u^t to x^{t-1}, i.e. $x^t(i) = x^{t-1}(i) + u^t(i)$ for each element i. As the perturbations are added, the color of the elements will evolve over time. Whenever an element's color reaches -1 or $+1$, we freeze that element's color and it is not longer updated. It remains to specify how to generate the updates u^t. This is done exactly using the Gaussian rounding mentioned above. In particular, at time t, we consider the SDP given by (1)–(2) (but only restricted to elements i that are still unfrozen). As the hereditary discrepancy is at most λ, no matter which variables are unfrozen at time t, the SDP is always feasible. We take such a solution v^t, scale it by a small constant γ (it suffices to choose $\gamma = 1/n$) and apply the Gaussian rounding, i.e. $u_i^t = \gamma \langle g, v_i^t \rangle$.

Formal Algorithm: We now state the algorithm formally.

1. Let x^t denote the coloring at time t. Let $\gamma = 1/n$ and $\ell = 8 \log n/\gamma^2$. We initialize, $x^0(i) = 0$ for all $i \in [n]$. The F^t denote the set of frozen variables by time t, where we initialize $F^0 = \emptyset$.
2. For each time step $t = 1, 2, \ldots, \ell$ repeat the following steps:
   (a) Find a feasible solution to the following semidefinite program:

$$\left\| \sum_{i \in S_j} v_i \right\|_2^2 \leq \lambda^2 \quad \text{for each set} \quad S_j$$
$$\|v_i\|_2^2 = 1 \quad \forall i \notin F^{t-1}$$
$$\|v_i\|_2^2 = 0 \quad \forall i \in F^{t-1}$$

   (b) Pick a random Gaussian $g^t \in \mathbb{R}^n$.
   (c) For each $i \in [n]$, update $x^t(i) = x^{t-1}(i) + \gamma \langle g^t, v_i^t \rangle$.
   (d) Set $F^t = F^{t-1}$. For each i, freeze i if $|x^t(i)| > 1$, and update $F^t = F^t \cup \{i\}$.

3. After time $t = \ell$, if some $|x^\ell(i)| < 1$ (i.e. some element is still not frozen), return fail.
4. For each i, set $x_i^\ell = -1$ if $x_i^\ell < -1$, and set $x_i^\ell = 1$ if $x_i^\ell > 1$. Output the final coloring x_ℓ.

Remark We note that the SDP in step 2(a) changes only when the set of frozen variables F changes. Moreover, the SDP is always feasible irrespective of the set F as $\text{herdisc}(\mathcal{S}) \leq \lambda$.

Analysis:

Theorem 8 *The algorithm described above produces a coloring with discrepancy* $O(\lambda(\log m \log n)^{1/2})$ *with probability at least* $1/2$.

Proof (*Sketch*) Let us consider an element i and see how its color $x^t(i)$ evolves over time until it crosses ± 1 and is frozen. Starting from 0, at each step the update $u^t(i) = \gamma\langle v_i^t, g^t\rangle$ is added. As $\|v_i^t\| = 1$ we have that u_i^t is distributed as $N(0, \gamma^2)$. Moreover, as g^t is chosen independently at each time step, $u^t(i)$ is independent[3] of all previous increments $u^{t'}(i)$, where $t' < t$. As the increments are $N(0, \gamma^2)$, with constant probability $x^t(i)$ will reach ± 1 in $O(1/\gamma^2)$ steps, and standard probabilistic tail bounds imply that it will reach ± 1 by time $\ell = O(\log n/\gamma^2)$ with probability at least $1 - 1/n^2$. So, with probability at least $1 - 1/n$, all elements will be frozen by time ℓ.

Let us now see how the discrepancy $x^t(S_j) = \sum_{i \in S_j} x^t(i)$ of a set S_j evolves over time. Clearly, it is 0 initially at $t = 0$. At each step t, it is updated by $u^t(S_j) = \sum_{i \in S_j} \gamma\langle g^t, v_i^t\rangle$. Let us denote $\lambda_t = \|\sum_{i \in S_j} v_j^t\|_2$. The SDP constraint ensures that $\lambda_t \leq \lambda$. Now, λ_t may depend on the previous choices $g^{t'}$ for $t' \leq t-1$ (as these choices affect the SDP at time t), but as g^t is chosen independently at time t, conditioned on the previous random choices $g^{t'}$, the update $u^t(S_j) \sim N(0, \gamma^2\lambda_t^2)$. Thus, $x^t(S_j)$ forms a martingale with increments $N(0, \leq \gamma^2\lambda^2)$. Again, standard martingale tail bound arguments imply that after ℓ steps,

$$\Pr[x^\ell(S_j) > c(\log m)^{1/2} \cdot \sqrt{\ell} \cdot \gamma\lambda] \leq 1/m^2$$

for some suitable constant c. Thus, with probability at least $1 - 1/m$

$$\text{disc}(\mathcal{S}) = O((\log m)^{1/2} \cdot \sqrt{\ell} \cdot \gamma\lambda) = O(\lambda \cdot (\log m \log n)^{1/2}).$$

Finally, we note that truncating $x^\ell(i)$ to ± 1 in the last step introduces very low error. Indeed, since $|x_i^t| < 1$ holds just before it is frozen, it must hold that $|x_i^t| < 1 + \gamma \cdot O((\log n)^{1/2})$ with high probability when it is frozen. Thus, the truncation can add an additional discrepancy of at most $n \cdot \gamma \cdot O((\log n)^{1/2}) = O((\log n)^{1/2})$ to any set.

[3] Strictly speaking, this is not true as the previous choices of $g^{t'}$ determine whether i is frozen by time t or not, and hence determine whether $u^t(i) = 0$ or not. More precisely, $x^t(i)$ forms martingale with $N(0, \gamma^2)$ increments, until it is absorbed at ± 1.

2.1 An extension

In a nutshell, the above algorithm works by gradually producing a coloring over time, and at each time step the low hereditary discrepancy ensures that regardless of which variables are frozen, there is a low discrepancy solution to the SDP relaxation.

It turns out that this algorithm can be modified to work with a weaker condition, instead of a bound on the hereditary discrepancy. To motivate this condition, let us consider the Beck-Fiala setting. Here, while it is conjectured that the hereditary discrepancy is $O(\sqrt{t})$, the best bound that we currently know of is only $O((t \log n)^{1/2})$ [2], and hence Theorem 6 only guarantees a $O(\sqrt{t} \log^{3/2} n)$ discrepancy coloring. It turns out that one can do better, and find a coloring with discrepancy $O(\sqrt{t} \log n)$ by using the fact that there always exists a partial coloring with discrepancy $O(\sqrt{t})$.[4]

Formally, let us define the *partial hereditary discrepancy* of a system as the smallest number λ such that given any sub-system, there is a partial coloring (say that colors at least half the elements in that sub-system) with discrepancy λ. The algorithm above can modified to obtain a coloring with discrepancy at most $O((\log n \log m)^{1/2})$ times the partial hereditary discrepancy as follows. Consider the algorithm that is defined exactly as before, except that instead of the SDP (1)–(2) above, one considers the following SDP.

$$\left\| \sum_{i \in S_j} v_i \right\|_2^2 \leq \lambda^2 \quad \text{for each set} \quad S_j \tag{3}$$

$$\sum_{i \notin F} \|v_i\|_2^2 \geq |F^c|/2 \tag{4}$$

$$\|v_i\|_2^2 \leq 1 \qquad \forall i \in F^c \tag{5}$$

$$\|v_i\|_2^2 = 0 \qquad \forall i \in F \tag{6}$$

Here F^c denotes the complement of F (i.e. unfrozen variables) and λ is the partial hereditary discrepancy. Note the difference here is that constraint (2) is replaced by the constraints (4) and (5), which only require that at least half of the unfrozen variables be colored.

While the algorithm is same as before, the analysis needs some more care. The problem is that the unfrozen variables do not necessarily satisfy $\|v_i\|_2 = 1$ anymore, but only the weaker condition (4). So, a priori it is possible that some variable always has $|v_i| \approx 0$ and hence never makes progress towards reaching ± 1. To get around this, one needs a more careful "energy increment" argument to show that after every $1/\gamma^2$ time steps, a constant fraction of the variables do reach ± 1 in expectation. One can then show that all elements are eventually colored ± 1 in $O(\log n/\gamma^2)$ time steps with reasonable probability. The details can be found in [3].

[4] In fact this is how Srinivasan's bound of $O(\sqrt{t} \log n)$ is obtained: by iteratively finding a partial coloring over the remaining uncolored elements. As at least half the elements are colored in each step, there are $O(\log n)$ steps each of which adds discrepancy $O(\sqrt{t})$.

3 Algorithm version of Spencer's result

Here we consider Theorem 5, which is much more subtle. Actually, to keep the presentation clean, we will only describe a weaker algorithm, that given an arbitrary set system on $m = O(n)$, finds a coloring efficiently with discrepancy $O((n \log\log\log n)^{1/2})$. Note that is already substantially better than randomized rounding.

The reason why this setting is more tricky is the following. First, unlike previously, it is not at all clear whether semidefinite programming can help. In particular, let us again consider the SDP given by (1)–(2). For Theorem 5, the natural thing[5] is to set $\lambda = O(\sqrt{n})$ and hope to round the SDP solution (over time). However, when $\lambda = \sqrt{n}$, it turns out that this SDP always has the trivial solution $v_i = e_i$ for $i \in [n]$, where e_i denotes the unit vector in the ith direction. This is feasible as the v_i's are unit vectors and hence (1) holds, and the orthogonality of the v_i's implies that $\| \sum_{i \in S_j} v_i \|_2 = (|S_j|)^{1/2} \leq n^{1/2}$ and hence (2) holds. Thus, the SDP does not reveal any useful information. A second problem is that in the previous algorithm, the discrepancy for each set performs a random walk and hence some sets are likely to deviate from their expected discrepancy by a factor of $\Omega(\sqrt{\log n})$. As we must set $\lambda = \Omega(\sqrt{n})$ in general, this would suggest that we cannot do better than $(n \log n)^{1/2}$ (which is what randomized rounding would give us anyways).

Main idea: To get around the above problems, the main idea is to let the discrepancy bound λ_S (in the SDP) depend upon the set S. As we shall see, the entropy method says that there always exists a partial coloring even if the discrepancy bounds λ_S are chosen non-uniformly, provided that these λ_S's are $\Omega(\sqrt{n})$ in a certain "average" sense.

So the algorithm works as follows: it starts with the all-zero coloring, and considers the partial coloring SDP [(3)–(6)] requiring that λ be $O(\sqrt{n})$ for each set (so initially, this SDP could just return the trivial all-orthogonal solution mentioned above). As previously, it projects the SDP solution on to a random gaussian vector, scales it by γ, and then adds it to the current coloring. As this process continues for $T = O(1/\gamma^2)$ steps, the discrepancy of each set does a random walk which is expect to stay $O(\sqrt{n})$. However, the random walk for some sets will get unlucky and they will start to incur a discrepancy of much larger than the desired $O(\sqrt{n})$. At this point, we reduce the λ_S for such sets to something much smaller than $\sqrt{n}$, thereby ensuring that their random walk has low variance henceforth and will not cause problems any more. Of course, we cannot reduce λ_S for all sets, so the key point is to show that λ_S is only reduced for a few sets, and that λ_S can still be kept $O(\sqrt{n})$ on "average" (in the sense needed by the entropy method), so that the SDPs remain feasible.

Before describing the algorithm formally, we first describe the entropy method, and also see how Spencer's original result [26] follows from it.

[5] One cannot set $\lambda = o(\sqrt{n})$ in general, as the Hadamard set system (see [21]) has vector discrepancy $\Omega(\sqrt{n})$.

3.1 The entropy method

The following result is known as the partial coloring lemma. Its proof is based on a clever pigeonhole principle argument, often referred to as the entropy method, and seems inherently non-constructive. The form we present below is from [20]. More details can be found in [1].

Lemma 2 (Partial Coloring via the Entropy Method) *Let $\mathcal{S}$ be a set system on an n-point set V, and let a number $\Delta_S > 0$ be given for each set $S \in \mathcal{S}$. Suppose Δ_S satisfy the condition*

$$\sum_{S \in \mathcal{S}} g\left(\frac{\Delta_S}{\sqrt{|S|}}\right) \leq \frac{n}{5} \tag{7}$$

where

$$g(\lambda) = \begin{cases} K e^{-\lambda^2/9} & \text{if} \quad \lambda > 0.1 \\ K \ln(\lambda^{-1}) & \text{if} \quad \lambda \leq 0.1 \end{cases}$$

and K is some absolute constant. Then there is a partial coloring $\mathcal{X}$ that assigns ± 1 to at least $n/2$ variables (and 0 to the rest), and satisfies $|\mathcal{X}(S)| \leq \Delta_S$ for each $S \in \mathcal{S}$.

Proof of Spencer's (non-constructive) result [26]*:* The coloring is constructed in phases. In phase i, for $i = 0, \ldots, \log n$, the number of uncolored elements left is at most $n_i \leq n/2^i$. In phase i, we apply Lemma 2 to these n_i elements with $\Delta_S^i = c(n_i \log(2m/n_i))^{1/2}$ and verify that (7) holds when c is a large enough constant. This gives a partial coloring on at least $n_i/2$ elements, with discrepancy for any set S at most Δ_S^i. Summing up over the phases, the overall discrepancy for any set is at most

$$\sum_i \Delta_S^i = \sum_i c\left(n2^{-i} \log\left(\frac{2m}{n2^{-i}}\right)\right)^{1/2} = O((n \log(2m/n))^{1/2}).$$

For our purposes below, we only need the following corollary of the above lemma.

Corollary 2 *Let $\mathcal{S}$ be any set system on $m = O(n)$ sets, and $\mathcal{S}' \subset \mathcal{S}$ be any subcollection of $O(n/(\log\log n)^2)$ sets. Then there exists a partial coloring where each set in $\mathcal{S}'$ has discrepancy $O(\sqrt{n}/\log n)$ and $O(\sqrt{n})$ otherwise.*

Proof This is a simple computation. For each set in $\mathcal{S}'$, the contribution to the left hand side of (7) is $g(\log n) \log O(\log\log n)$. Now, as $|\mathcal{S}'| = O(n/\log\log^2 n)$, the overall entropy contribution due to such sets is $o(n)$. On the other hand, choosing $\lambda_S = c\sqrt{n}$ for a large enough constant c, ensures that the contribution of any set S to (7) is $\ll 1$. As $m = O(n)$, this ensures that (7) holds and the partial coloring with the desired properties exists.

3.2 Algorithmic subroutine and analysis

We now describe our algorithm. We only describe the first phase when the number of uncolored variables reduces from n to $n/2$. This is the hardest phase and conveys all the ideas we need. The problem only gets easier in subsequent phases when fewer uncolored variables remain (e.g. when only $n/\log n$ variables remain, randomized rounding only adds $O(\sqrt{n})$ discrepancy).

The algorithm in the first phase works as follows: we start with the all 0 coloring, and consider the following partial coloring SDP.

$$\left\| \sum_{i \in S} v_i \right\|_2^2 \leq \lambda_S^2 \quad \text{for each set} \quad S = S_1, \ldots, S_m \tag{8}$$

$$\sum_{i \notin F} \|v_i\|_2^2 \geq |F^c|/2 \tag{9}$$

$$\|v_i\|_2^2 \leq 1 \quad \forall i \notin F \tag{10}$$

$$\|v_i\|_2^2 = 0 \quad \forall i \in F \tag{11}$$

Initially we set each $\lambda_S = cn^{1/2}$ for c large enough c such that (7) is satisfied easily with some slack. As previously, for each time step t, we obtain $u_i^t = \gamma \langle g^t, v_i^t \rangle$ for $i = 1, \ldots, n$ and add it to the coloring thus far. We repeat this for $O(1/\gamma^2)$ steps, at which point we expect half the colors to reach ± 1.

During these steps, if the discrepancy $|x^t(S)|$ exceeds $2(n \log\log\log n)^{1/2}$ for some set S, we label S *dangerous* and set $\lambda_S = n^{1/2}/\log n$ for it in the SDP. This ensures that its discrepancy increment $u^t(S)$ will have standard deviation $O(\gamma \cdot (n^{1/2}/\log n))$ henceforth, making S extremely unlikely to incur an additional $\Omega(n^{1/2})$ discrepancy over the remaining $O(1/\gamma^2)$ steps.

Thus, at the end of the algorithm, each set has $O((n \log\log\log n)^{1/2})$ discrepancy with high probability. So, it only remains to show that the SDP never becomes infeasible. We show that this happens with, say, probability at least 1/2. Indeed, since the discrepancy of any set forms a martingale with Gaussian increments with standard deviation $O(\sqrt{n})$, the probability of a set ever becoming dangerous is $O(\exp(-2\log\log\log n) = O(\log\log n)^{-2}$. So, with probability at least, say 3/4, the number of dangerous sets does not exceed $O(n(\log\log n)^{-2})$. Let us condition on this event. As we set $\lambda_S = n^{1/2}/\log n$ for the dangerous sets and $O(\sqrt{n}$ for others, Corollary 2 implies that the SDP does not become infeasible. This gives the desired result.

We remark that the $O(n^{1/2})$ bound in Theorem 5 is simply obtained by refining this idea, where we choose multiple danger levels and set the bounds λ_S appropriately for each danger level.

4 Matoušek's lower bound

We now prove Theorem 7, which says that the determinant lower bound is within poly-logarithmic factors of the hereditary discrepancy.

Main idea: The proof uses several interesting ideas. First, one observes that for any set system $(V, \mathcal{S})$, the proof of Theorem 6 implies that there must be some subset J of elements such that the vector discrepancy of the system $\mathcal{S}_{|J}$ satisfies

$$\text{vecdisc}(\mathcal{S}_{|J}) = \Omega(\text{herdisc}(\mathcal{S})/(\log m \log n)^{1/2}) \tag{12}$$

To see this, suppose that (12) does not hold for any J. Then, the algorithm for Theorem 6 can be used to color any system $\mathcal{S}_{|J'}$, for $J' \subseteq V$ with discrepancy strictly less then $\text{herdisc}(\mathcal{S})$, contradicting the definition of hereditary discrepancy.

Now, the key idea is to show that if $\text{vecdisc}(\mathcal{S}_{|J})$ is large, then $\text{detlb}(\mathcal{S}_{|J})$ is also large (Theorem 9 below has the formal statement). This is done by considering the dual for the SDP (1)–(2), which gives a witness for the large vector discrepancy of $\mathcal{S}_{|J}$. This witness is then used to show that the incidence matrix of $\mathcal{S}_{|J}$ has a sub-matrix with large eigenvalues, which in turn implies that it has a large determinant.

Details: As $\text{vecdisc}(S)$ is an optimum solution to a semidefinite program, one can use the duality of semidefinite programming to lower bound $\text{vecdisc}(S)$ in the following way.

Lemma 3 *For any set system $(V, \mathcal{S})$, we have* $\text{vecdisc}(\mathcal{S}) \geq D$ *if and only if there are nonnegative reals $w_1, \ldots, w_m$ with $\sum_{i=1}^m w_i \leq 1$ and reals $z_1, \ldots, z_n$ with $\sum_{j=1}^n z_j \geq D^2$ such that for all $\mathbf{x} \in \mathbb{R}^n$,*

$$\sum_{i=1}^m w_i \left(\sum_{j \in S_i} x_j \right)^2 \geq \sum_{j=1}^n z_j x_j^2. \tag{13}$$

Computing the dual formally needs some work (and we refer the reader to [19] for details). However, this dual has a rather intuitive interpretation. Suppose one could construct a convex combination w_i of the set discrepancy constraints $(\sum_{j \in S_i} x_j)^2$ such that this sum always exceeds $\sum_{j=1}^n z_j x_j^2$ no matter what real values are assigned to x_j's. Or, in other words, $\sum_{i=1}^m w_i (\sum_{j \in S_i} x_j)^2 - \sum_{j=1}^n z_j x_j^2$ is a positive quadratic form. Then, indeed $\sum_{j=1}^n z_j x_j^2 = \sum_j z_j$ for any ± 1 assignment to the x_j's and hence this certifies that $\sum_{j=1}^n z_j$ is a lower bound on the discrepancy. The duality says that this is also a lower bound on the vector discrepancy, and moreover, if the vector discrepancy is D, then there always exists a choice of witnesses w_i's and z_j's of this form.

We are now ready to prove the following theorem, which directly implies Theorem 7.

Theorem 9 *Let $\mathcal{A} = \{A_1, \ldots, A_m\}$ be a set system on $[n]$ with* $\text{vecdisc}(\mathcal{A}) = D$. *Then* $\text{detlb}(\mathcal{A}) = \Omega(D/\sqrt{\log n})$.

Before proving Theorem 9, we first show how it implies Theorem 7. Applying Theorem 9 to $A = \mathcal{S}_{|J}$ (on the universe J, which satisfies $|J| \leq n$) implies that $\text{detlb}(\mathcal{S}_{|J}) = \Omega(\text{vecdisc}(\mathcal{S}_{|J})/\sqrt{\log n})$. Together with (12), this gives

$$\operatorname{detlb}(\mathcal{S}) \geq \operatorname{detlb}(\mathcal{S}_{|J}) \geq \frac{1}{\sqrt{\log n}} \cdot \Omega\left(\frac{\operatorname{herdisc}(\mathcal{S})}{(\log m \log n)^{1/2}}\right),$$

where the first inequality follows as any sub set system of $\mathcal{S}_{|J}$ is also a sub set system of $\mathcal{S}$.

Proof (*Theorem 9*) Let us consider the dual formulation of vector discrepancy from Lemma 3. For more convenient notation, let us write the nonnegative weight w_i as β_i^2. Moreover, let $L \subseteq [n]$ consist of the indices j with $z_j > 0$. We will only consider the inequality (13) in Lemma 3 with $x_j = 0$ whenever $j \notin L$. Writing $z_j = \gamma_j^2$ for $j \in L$, we obtain

$$\sum_{i=1}^{m} \beta_i^2 \left(\sum_{j \in A_i \cap L} x_j\right)^2 \geq \sum_{j \in L} \gamma_j^2 x_j^2 \tag{14}$$

for all $x \in \mathbb{R}^L$, where $\|\beta\| \leq 1$ and $\|\boldsymbol{\gamma}\| \geq D$.

Next, we select $K \subseteq L$ with $\|\gamma[K]\| = \Omega(D/\sqrt{\log n})$ and such that all entries of $\gamma[K]$ are within a factor of 2 of each other (here $\gamma[K]$ denotes the vector γ restricted to coordinates in K). Such a subset K exists for the following reason: let $\gamma_{\max} = \max_j |\gamma_j|$. For $i = 0, 1, 2, \ldots$, let $K_i = \{j : |\gamma_j| \in (2^{-i-1}\gamma_{\max}, 2^{-i}\gamma_{\max}]\}$. The contribution to $\|\gamma\|$ of the components of γ with indices in K_i for $i \geq 2\log n$, say, is negligible, and so there exists some i_0 for which $\sum_{j \in K_{i_0}} \boldsymbol{\gamma}_j^2 = \Omega(\|\boldsymbol{\gamma}\|^2/\log n)$.

Let us denote $k = |K|$ and $\tilde{D} = \frac{1}{2}\|\boldsymbol{\gamma}[K]\|$. As $\sum_{j \in K} \gamma_j^2 = 4\tilde{D}^2$, and all these γ_j are within twice of each other, we have that $\gamma_j \geq \tilde{D}/\sqrt{k}$ for all $j \in K$. So, restricting (14) to vectors x with $x_j = 0$ for $j \notin K$, we have that

$$\sum_{i=1}^{m} \beta_i^2 \left(\sum_{j \in A_i \cap K} x_j\right)^2 \geq \frac{\tilde{D}^2}{k} \sum_{j \in K} x_j^2. \tag{15}$$

Let $C = A[*, K]$ be the $m \times k$ incidence matrix of the system $A|_K$ and let $\check{C}$ be the $m \times k$ matrix obtained from C by multiplying the ith row by β_i. Then (15) can be rewritten as

$$x^T \check{C}^T \check{C} x = \|\check{C}x\|^2 \geq \frac{\tilde{D}^2}{k}\|x\|^2 \quad \text{for all} \quad x \in R^k.$$

This, by the usual variational characterization of eigenvalues, tells us that the smallest eigenvalue of the $k \times k$ matrix $\check{C}^T\check{C}$ is at least $\tilde{D}^2/k$. Now, as the determinant is the product of eigenvalues, this implies that $\det(\check{C}^T\check{C}) \geq (\tilde{D}^2/k)^k$. By the Binet–Cauchy formula we obtain

$$\det(\check{C}^T\check{C}) = \sum_{I} \det(\check{C}[I, *])^2, \tag{16}$$

where the summation is over all k-element subsets $I \subseteq [m]$ and $\check{C}[I,*]$ consists of the rows of $\check{C}$ whose indices lie in I. Setting $M := \max_I |\det(C[I,*])|$ and noting that $\det(\check{C}[I,*]) = \det(C[I,*]) \prod_{i \in I} \beta_i$, we can bound the right-hand side of (16) as

$$\sum_I \det(\check{C}[I,*])^2 = \sum_I \det(C[I,*])^2 \prod_{i \in I} \beta_i^2 \leq M^2 \sum_I \prod_{i \in I} \beta_i^2$$
$$\leq M^2 \frac{\left(\sum_{i=1}^m \beta_i^2\right)^k}{k!} \leq \frac{M^2}{k!},$$

where the second inequality follows as every term $\prod_{i \in I} \beta_i^2$ occurs $k!$ times in the multinomial expansion of $(\beta_1^2 + \cdots + \beta_m^2)^k$. Letting $B := C[I,*]$ for an I maximizing $|\det C[I,*]|$, we have

$$\det(B)^2 \geq k! \det(\check{C}^T \check{C}) \geq k!(\tilde{D}^2/k)^k \geq (k/e)^k (\tilde{D}^2/k)^k = \Omega(D/\sqrt{\log n})^{2k}.$$

So the $k \times k$ matrix B witnesses $\mathrm{detlb}(\mathcal{F}) = \Omega(D/\sqrt{\log n})$, and the lemma is proved.

5 Concluding remarks

While all the algorithms that we described were randomized, these results can also be made deterministic [4]. Surprisingly, this turns out to be non-trivial and the entropy method plays a key role here. To derandomize, instead of a picking a random gaussian vector g at each time step, one needs to deterministically pick a vector on which to project the SDP solution. However, it is not apriori clear why a vector that simultaneously guarantees low increase in discrepancy and that the energy of the solution increases sufficiently (to make progress towards reaching a ± 1 coloring) should exist, even though these properties hold in expectation for a random gaussian vector. The fix is to add some additional constraints to the SDP formulation that help ensure that such a vector will exist. However, one now needs to argue that the resulting SDP with these extra constraints is still feasible, and the entropy method plays a key role here.

Even though we now have a constructive algorithm for Spencer's result, recall that this algorithm crucially relies on the non-constructive entropy method. In particular, it does not give a different non-counting based proof of Spencer's result. This is quite different from, say, Moser's constructive algorithm [22] for Lovász Local Lemma, which can be viewed as giving an alternate algorithmic proof of the Local Lemma. It would be interesting to find a truly constructive proof of Spencer's result.

While we now have algorithmic results for many discrepancy problems, several related problems are still open. In particular, while the best known non-constructive bound for the Beck-Fiala problem is $O((t \log n)^{1/2})$ [2], the best known constructive bound is only $O(\sqrt{t} \log n)$. One reason for this is that the result of [2] is based on techniques from convex-geometry and does not use the partial coloring lemma. To this end, a more combinatorial proof of the $O((t \log n)^{1/2})$ result would also be very interesting. Of course, it is an outstanding open question to obtain an $O(\sqrt{t})$ bound. Another open question is to improve the $O(\lambda(\log n \log m)^{1/2})$ guarantee in Theorem 6.

It may be possible to remove the factor of $O(\log^{1/2} n)$. In particular, we do not know of any instance where the discrepancy exceeds the vector hereditary discrepancy by a factor of $\Omega(\log^{1/2} m)$.

Finally, it would be extremely interesting to see if similar techniques can be helpful for other problems where the existence of good solutions is always guaranteed using some non-constructive argument, but it is unclear how to find them. Some well known examples of such problems include Minkowski's first theorem on the existence of short vectors in lattices, and the existence of perfect matchings in hypergraphs [16].

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

References

1. Alon, N., Spencer, J.H.: The Probabilistic Method. 2nd edn. Wiley, New York (2000)
2. Banaszczyk, W.: Balancing vectors and gaussian measures of n-dimensional convex bodies. Random Struct. Algorithms **12**(4), 351–360 (1998)
3. Bansal, N.: Constructive algorithms for discrepancy minimization. In: Foundations of Computer Science (FOCS), pp. 3–10 (2010)
4. Bansal, N., Spencer, J.: Deterministic discrepancy minimization. In: ESA, pp. 408–420 (2011)
5. Beck, J.: Roth's estimate on the discrepancy of integer sequences is nearly sharp. Combinatorica **1**, 319–325 (1981)
6. Beck, J., Fiala, T.: Integer-making theorems. Discret. Appl. Math. **3**, 1–8 (1981)
7. Bednarchak, D., Helm, M.: A note on the beck-fiala theorem. Combinatorica **17**, 147–149 (1997)
8. Chandrasekaran, K., Vempala, S.: A discrepancy based approach to integer programming. Arxiv 1111.4649 (2011)
9. Charikar, M., Newman, A., Nikolov, A.: Tight hardness results for minimizing discrepancy. In: Proceedings of the 22nd Annual ACM-SIAM Symposium on Discrete Algorithms (SODA) (2011)
10. Chazelle, B.: The Discrepancy Method: Randomness and Complexity. Cambridge University Press, Cambridge (2001)
11. Doerr, B.: Linear and hereditary discrepancy. Combin. Probab. Comput. **9**(4), 349–354 (2000)
12. Drmota, M., Tichy, R.F.: Sequences, Discrepancies and Applications. Springer, Berlin (1997)
13. Eisenbrand, F., Pálvölgyi, D., Rothvoss, T.: Bin packing via discrepancy of permutations. In: Symposium on Discrete Algorithms (SODA), pp. 476–481 (2011)
14. Gärtner, B., Matoušek, J.: Approximation Algorithms and Semidefinite Programming. Springer, Berlin (2012)
15. Ghouila-Houri, A.: Caracterisation des matrices totalement unimodularies. Comptes Rendus Hebdomadaires Des Seances de l'Academie Des Sciences **254**, 1191–1194 (1962)
16. Haxell, P.: A condition for matchability in hypergraphs. Graphs Combin. **11**, 245–248 (1995)
17. Kim, J.H., Matoušek, J., Vu, V.: Discrepancy after adding a single set. Combinatorica **25**(4), 499–501 (2005)
18. Lovász, L., Spencer, J., Vesztergombi, K.: Discrepancy of set-systems and matrices. Eur. J. Combin. **7**, 151–160 (1986)
19. Matoušek, J.: The determinant bound for discrepancy is almost tight. Manuscript, Arxiv 1101:0767
20. Matoušek, J.: An lp version of the beck-fiala conjecture. Eur. J. Combin. **19**, 175–182 (1998)
21. Matoušek, J.: Geometric Discrepancy: An Illustrated Guide. Springer, Berlin (2010)
22. Moser, R.A.: A constructive proof of the lovász local lemma. In: Symposium on Theory of Computering (STOC), pp. 343–350 (2009)
23. Newman, A., Nikolov, A.: A counterexample to beck's conjecture on the discrepancy of three permutations. Arxiv 1104.2922 (2011)
24. Pálvölgyi, D.: Indecomposable coverings with concave polygons. Discret. Comput. Geom. **44**, 577–588 (2010)

25. Rothvoss, T.: The entropy rounding method in approximation algorithms. In: Symposium on Discrete Algorithms (SODA), pp. 356–372 (2012)
26. Spencer, J.: Six standard deviations suffice. Trans. Am. Math. Soc. **289**(2), 679–706 (1985)
27. Srinivasan, A.: Improving the discrepancy bound for sparse matrices: better approximations for sparse lattice approximation problems. In: Symposium on Discrete Algorithms (SODA), pp. 692–701 (1997)
28. Vandenberghe, L., Boyd, S.: Semidefinite programming. SIAM Rev. **38**, 49–95 (1996)