# MULTI–COLOR DISCREPANCIES

BENJAMIN DOERR AND ANAND SRIVASTAV

Abstract. In this article we introduce combinatorial multi-color discrepancies and generalize several classical results from 2–color discrepancy theory to $c$ colors ($c \geq 2$). We give a recursive method that constructs $c$–colorings from approximations of 2–color discrepancies. This method works for a large class of theorems like the 'six standard deviations' theorem of Spencer, the Beck–Fiala theorem, the results of Matoušek, Welzl and Wernisch and Matoušek for bounded VC–dimension and Matoušek's and Spencer's upper bound for the arithmetic progressions. In particular, the $c$–color discrepancy of an arbitrary hypergraph ($n$ vertices, $m$ hyperedges) is $\mathcal{O}(\sqrt{\frac{n}{c}\log m})$. If $m = \mathcal{O}(n)$, then this bound improves to $\mathcal{O}(\sqrt{\frac{n}{c}\log c})$.

On the other hand there are examples showing that discrepancy in $c$ colors can not be bounded in terms of two-color discrepancies in general, even if $c$ is a power of 2. For the linear discrepancy version of the Beck–Fiala theorem the recursive approach also fails.

Here we extend the method of floating colors via tensor products of matrices to multi-colorings and prove multi-color versions of the the Beck–Fiala theorem and the Barany–Grunberg theorem. Using properties of the tensor product we derive a lower bound for the $c$–color discrepancy of general hypergraphs. For the hypergraph of arithmetic progressions in $\{1, \ldots, n\}$ this yields a lower bound of $\frac{1}{25\sqrt{c}}\sqrt[4]{n}$ for the discrepancy in $c$ colors. The recursive method shows an upper bound of $\mathcal{O}(c^{-0.16}\sqrt[4]{n})$.

## 1. Introduction

Combinatorial discrepancy theory deals with the problem of partitioning the vertices of a hypergraph (set-system) in such a way that all hyperedges are split into roughly equal-sized parts. Discrepancy measures the deviation of an optimal partition to an ideal one, that is one where all edges contain the same number of vertices in each class of the partition.

Usually one represents the partition by a coloring, that is a mapping from the vertices into some set such that the classes of equal images form the partition classes. In this language, most results known so far only deal with two colors. Recent results from communication complexity (e. g. [BHK98]) further motivate the study of multi-color discrepancies.

*Date*: June 26, 2004.

1991 *Mathematics Subject Classification.* Primary 11K38, 05C15. Secondary 05C65.

*Key words and phrases.* discrepancy theory, coloring of hypergraphs, arithmetic progressions.

The first author was supported by the graduate school 'Effiziente Algorithmen und Multiskalenmethoden', Deutsche Forschungsgemeinschaft.

In this article we give two methods to construct multi-colorings with low discrepancy. If the given hypergraph fulfills some hereditary property, we can construct multi-colorings from 2–colorings having low discrepancy. This works for a large class of hypergraphs: For hypergraphs $\mathcal{H}$ with $n$ vertices and $n$ hyperedges we obtain an upper bound of $\mathcal{O}(\sqrt{\frac{n}{c}})$ for the discrepancy in $c$ colors. This extends Spencer's famous 'six standard deviations' result to arbitrary numbers of colors. Along the same lines we receive another extension of Spencer's result: We find that in this situation the discrepancy with respect to a given weight $(p, 1-p), p \in [0,1]$ is $\mathcal{O}(\sqrt{pn})$.

We also apply the recursive method to prove multi-color versions of the Beck–Fiala theorem [BF81], the results of Matoušek, Welzl and Wernisch [MWW84] and Matoušek [Mat95] for bounded VC–dimension, and an upper bound of $\mathcal{O}(c^{-0.16}n^{1/4})$ for the multi-color discrepancy of the hypergraph of arithmetic progressions, extending the bound of Matoušek and Spencer [MS96].

The recursive method fails for the linear discrepancy version of the Beck–Fiala theorem and the theorem of Barany-Grunberg. Here we extend the method of floating colors to multi-colorings and use this to prove similar results for any number of colors.

Finally, introducing a tensor product representation for multi-color discrepancy, we extend a lower bound result for the discrepancy function due to Lovász and Sós to the multi-color case: We show that the $c$–color discrepancy is at least $\sqrt{\frac{n(c-1)}{mc^2}\lambda_{\min}(A^\top A)}$, where $A$ is the incidence matrix of the hypergraph and $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue. Applying this to the hypergraph of arithmetic progressions $\mathcal{H}_n$ we get a lower bound for the $c$–color discrepancy of $\Omega(c^{-0.5}n^{1/4})$. Spencer's examples of hypergraphs constructed from Hadamard matrices have $n$ vertices and $n$ hyperedges and show a $c$–color discrepancy of $\Omega(\sqrt{\frac{n}{c}})$. Thus our bounds for the arithmetic progressions and the 'six standard deviations' situation are rather close. In particular, they are sharp for fixed numbers of colors (apart from constants).

## 2. Preliminaries

2.1. **Two-Color Discrepancy.** Let $\mathcal{H} = (X, \mathcal{E})$ denote a finite *hypergraph*, i. e. $X$ is a finite set (called *vertices* or *nodes*) and $\mathcal{E}$ is a family of subsets of $X$ (called *hyperedges*). A partition into two classes can be represented by a *coloring* $\chi : X \to \{-1, +1\}$. We call $-1$ and $+1$ *colors*. The color-classes $\chi^{-1}(-1)$ and $\chi^{-1}(+1)$ form the partition. The imbalance of a hyperedge $E \in \mathcal{E}$ can be expressed as $\chi(E) := \sum_{x \in E} \chi(x)$. The *discrepancy* of $\mathcal{H}$ is defined by

$$\text{disc}(\mathcal{H}) = \min_{\chi: X \to \{-1,+1\}} \max_{E \in \mathcal{E}} |\chi(E)| \tag{1}$$

This concept may be generalized to matrices in a natural way. Let $A = (a_{ij})$ be any $m \times n$–matrix and set

$$\operatorname{disc}(A) := \min_{\chi \in \{-1,1\}^n} \|A\chi\|_\infty. \tag{2}$$

Let $X = \{x_1, \dots, x_n\}$, $\mathcal{E} = \{E_1, \dots, E_m\}$ and define a matrix $A = (a_{ij})$ by $a_{ij} = 1$ if $x_j \in E_i$ and $a_{ij} = 0$ else. $A$ is called the incidence matrix of $\mathcal{H}$. We have $\operatorname{disc}(A) = \operatorname{disc}(\mathcal{H})$, so discrepancy of matrices is indeed a more general concept (if, of course, we restrict ourselves to zero–one–matrices both concepts are equivalent). Sometimes even for hypergraphs the matrix notion is more convenient.

There are two related notions: The *linear discrepancy* of an arbitrary matrix $A$ is defined by

$$\operatorname{lindisc}(A) := \max_{p \in [-1,1]^n} \min_{\chi \in \{-1,1\}^n} \|A(p-\chi)\|_\infty. \tag{3}$$

Linear discrepancy can be regarded as a measure of how well a fractional solution can be rounded to an integer solution. Some authors define linear discrepancy by

$$\max_{p \in [0,1]^n} \min_{\chi \in \{0,1\}^n} \|A(p-\chi)\|_\infty \tag{4}$$

Both versions differ only by the constant factor 2. A special case of the second version is the weighted discrepancy, which refers to the problem of splitting the edges in an arbitrary ratio $p : 1-p$ (instead of $0.5 : 0.5$ as in the discrepancy problem). We define the weighted discrepancy in 2 colors by

$$\operatorname{wd}(A,2) := \max_{p \in [0,1]} \min_{\chi \in \{0,1\}^n} \|A(\mathbf{1}_n p - \chi)\|_\infty, \tag{5}$$

where $\mathbf{1}_n = (1, \dots, 1)^\top \in \mathbb{R}^n$.

Finally the *hereditary discrepancy* is

$$\operatorname{herdisc}(A) := \max_{J \subseteq [n]} \operatorname{disc}((a_{ij})_{i \in [m], j \in J}), \tag{6}$$

where $[n] := \{1, \dots, n\}$.

All notions can be translated to hypergraphs in a natural way, so $\operatorname{herdisc}(\mathcal{H})$ is the maximum discrepancy of all induced subgraphs.

To shorten notation we will write $A_0 \leq A$ to indicate that the matrix $A_0$ consists of some columns of the matrix $A$. Similarly for hypergraphs we will write $\mathcal{H}_0 \leq \mathcal{H}$ if $\mathcal{H}_0$ is an induced subgraph of $\mathcal{H}$.

**Proposition 2.1.** *The following relations between the different notions hold:*

(i) $\operatorname{disc}(A) \leq \operatorname{herdisc}(A)$.
(ii) $\operatorname{disc}(A) \leq \operatorname{lindisc}(A)$.
(iii) $\operatorname{wd}(A,2) \leq \frac{1}{2} \operatorname{lindisc}(A)$.

(iv) $\operatorname{lindisc}(A) \leq 2 \operatorname{herdisc}(A)$.

(i) to (iii) are trivial, while the relation between the linear and the hereditary discrepancy is not. (iv) was discovered by [BS84] and [LSV86]. It is not completely clear how sharp this inequality is. The first author recently gave a slight improvement in [Doe00a]. For our purposes (iv) is sufficient. Matoušek investigated the opposite problem of bounding the hereditary discrepancy in terms of the linear discrepancy in [Mat00].

An excellent survey of classical and recent results in discrepancy theory is the article of Beck and Sós [BS95]. For a more thorough treatment we refer to the books Beck and Chen [BC87] and Matoušek's [Mat99]. Discrepancy theory is put into a broader context in the new book of Chazelle [Cha00].

2.2. **Multi-Color Discrepancy.** Up to now little is known about the discrepancy problem where we ask for a partition into more than two classes. We found the following two results:

**Theorem 2.2** (Beck, Fiala [BF81]). *Given $n \geq 5$ finite sets, it is possible to partition their union into $r$ parts $A_1, \ldots, A_r$ for any positive integer $r$ in such a way, that, for each $i, j$ and $k$,*

$$\big| |E_i \cap A_j| - |E_i \cap A_k| \big| < 24\sqrt{2n \log(2n)}.$$

**Theorem 2.3** (Beck, Sós [BS95]). *Let $\mathcal{H}$ be any hypergraph such that the incidence matrix of $\mathcal{H}$ is unimodular. Then for any number $c \in \mathbb{N}$ there is a $c$–partition $X = X_1 \dot\cup \ldots \dot\cup X_c$ of $X$ such that for any edge $E \in \mathcal{E}$ and $i \in [c]$*

$$\left| |E \cap X_i| - \frac{|E|}{c} \right| < 1.$$

Let us introduce some notation concerning $c$–color discrepancies. A $c$–coloring of $\mathcal{H}$ is simply a mapping $\chi : X \to M$, where $M$ is any set of cardinality $c$. For convenience, normally one has $M = [c] := \{1, \ldots, c\}$. Sometimes a different set $M$ will be of advantage. Note that in applications to communication complexity $M$ can be a finite abelian group [BHK98]. The basic idea of measuring the deviation from the average motivates the definitions of the *discrepancy of an edge $E \in \mathcal{E}$ in color $i \in M$ with respect to $\chi$* by

$$\operatorname{disc}_{\chi,i}(E) := \left| |\chi^{-1}(i) \cap E| - \frac{|E|}{c} \right|, \tag{7}$$

the *discrepancy of $\mathcal{H}$ with respect to $\chi$* by

$$\operatorname{disc}(\mathcal{H}, \chi, c) := \max_{i \in M, E \in \mathcal{E}} \operatorname{disc}_{\chi,i}(E) \tag{8}$$

and the *discrepancy of $\mathcal{H}$ in $c$ colors* by

$$\operatorname{disc}(\mathcal{H}, c) := \min_{\chi : X \to [c]} \operatorname{disc}(\mathcal{H}, \chi, c). \tag{9}$$

Immediately we see

**Remark 2.4.** $\operatorname{disc}(\mathcal{H}, 2) = \frac{1}{2}\operatorname{disc}(\mathcal{H})$.

To add some further motivation to what follows let us give an example which shows that a hypergraph may have very different discrepancies in different numbers of colors.

**Example:** Let $k \in \mathbb{N}$ and $n = 4k$. Set $\mathcal{H}_n = ([n], \{X \subseteq [n] \,|\, |X \cap [\frac{n}{2}]| = |X \setminus [\frac{n}{2}]|\})$. Obviously, $\mathcal{H}_n$ has 2–color discrepancy zero, but $\operatorname{disc}(\mathcal{H}_n, 4) = \frac{1}{8}n$.

*Proof.* Let $\chi : [n] \to [4]$ be any 4–coloring. Let $i \in [4]$ be a color such that $|\chi^{-1}(i)| \leq \frac{1}{4}n$. Then there are sets $E_1 \subseteq [\frac{n}{2}]$, $E_2 \subseteq [n] \setminus [\frac{n}{2}]$ such that $|E_j| = \frac{1}{4}n$ and $\chi^{-1}(i) \cap E_j = \emptyset$. Thus $E_1 \cup E_2$ is an edge in $\mathcal{H}$ and has discrepancy $\frac{1}{8}n$ in color $i$. On the other hand $\chi : x \mapsto \lceil \frac{4x}{n} \rceil$ is a coloring having discrepancy $\frac{1}{8}n$. □

In fact, such examples exist for nearly any two numbers of colors. Unless $c_1$ divides $c_2$, there are hypergraphs $\mathcal{H}_n$ on $n$ vertices having discrepancy $\Theta(n)$ in $c_1$ colors and zero discrepancy in $c_2$ colors. This has been investigated in [Doe02b].

In the above notion we cannot express discrepancies simply by sums of colors. As this is very practical sometimes and a step towards the matrix concept, we describe the color $i \in [c]$ by a vector $m^{(i)} \in \mathbb{R}^c$ defined by

$$m_j^{(i)} := \begin{cases} \frac{c-1}{c} & \text{if } i = j \\ -\frac{1}{c} & \text{otherwise.} \end{cases}$$

Then

$$\operatorname{disc}(\mathcal{H}, \chi, c) = \max_{E \in \mathcal{E}} \left\| \sum_{x \in E} m^{(\chi(x))} \right\|_\infty . \tag{10}$$

Set $M_c := \{m^{(i)} | i \in [c]\}$. Apparently, we have

$$\operatorname{disc}(\mathcal{H}, c) = \min_{\chi : X \to M_c} \max_{E \in \mathcal{E}} \left\| \sum_{x \in E} \chi(x) \right\|_\infty . \tag{11}$$

As for 2 colors, the notion of multi-color discrepancy has a natural extension to matrices. Let $A \in \mathbb{R}^{m \times n}$ be any matrix. Let $\overline{A}$ be the matrix which results from replacing every $a_{ij}$ in $A$ by $a_{ij} I_c$, where $I_c$ shall denote the identity matrix of dimension $c$. Identifying a $\chi : [n] \to M_c$ by a $cn$–dimensional vector in the natural way, we get

$$\operatorname{disc}(A, c) := \min_{\chi : [n] \to M_c} \left\| \overline{A} \chi \right\|_\infty . \tag{12}$$

The other notions of discrepancy transform to the multi-color case in a similar way: Set $\overline{M}_c = \{\sum_{i\in[c]} \lambda_i m^{(i)} | \lambda \in [0,1]^c, \sum_{i\in[c]} \lambda_i = 1\}$, the convex hull of $M_c$. For $p \in \overline{M}_c$ set $\overline{p} : [n] \to \overline{M}_c; i \mapsto p$. We define the *(weighted) discrepancy of $A$ with respect to the weight $p$ and the coloring $\chi$* by

$$\mathrm{wd}(A, \chi, p) := \left\| \overline{A}(\overline{p} - \chi) \right\|_\infty , \tag{13}$$

the *(weighted) discrepancy of $A$ with respect to the weight $p$* by

$$\mathrm{wd}(A, c, p) := \min_{\chi:[n]\to M_c} \mathrm{wd}(A, \chi, p) \tag{14}$$

and the *weighted discrepancy of $A$* by

$$\mathrm{wd}(A, c) := \max_{p\in\overline{M}_c} \mathrm{wd}(A, c, p). \tag{15}$$

There is an equivalent way to define weighted discrepancy which puts more emphasis on the aspect of weights: Denote by $E_c$ the standard basis of $\mathbb{R}^c$ and by $\overline{E}_c$ its convex hull, that is all $p \in [0,1]^c$ such that $\|p\|_1 = 1$. We have

$$\mathrm{wd}(A, c) = \max_{p\in\overline{E}_c} \min_{\chi:[n]\to E_c} \left\| \overline{A}(\overline{p} - \chi) \right\|_\infty . \tag{16}$$

Note that this is an extension of the definition of $\mathrm{wd}(\mathcal{H}, 2)$ in equation (5). For hypergraphs this translates to

$$\mathrm{wd}(\mathcal{H}, c) = \min_{\chi:X\to M_c} \max_{j\in[c],E\in\mathcal{E}} \left| |E \cap \chi^{-1}(j)| - p_j|E| \right| .$$

The linear discrepancy in $c$ colors can be defined by

$$\mathrm{lindisc}(A, c) := \max_{p:[n]\to\overline{M}_c} \min_{\chi:[n]\to M_c} \left\| \overline{A}(p - \chi) \right\|_\infty . \tag{17}$$

Finally the hereditary discrepancy in $c$ colors is

$$\mathrm{herdisc}(A, c) := \max_{A_0\le A} \mathrm{disc}(A_0, c). \tag{18}$$

All these notions shall be defined for hypergraphs as well by taking the incidence matrix of the hypergraph. E. g., for a hypergraph $\mathcal{H}$ with incidence matrix $A$ we have $\mathrm{lindisc}(\mathcal{H}) := \mathrm{lindisc}(A)$. Like in Remark 2.4, these other discrepancy notions are identical with the usual notions up to the constant factor of 2. When citing 2–color results we will use the conventional notation which has no parameter $c$ in it, e. g. $\mathrm{herdisc}(\mathcal{H})$, so there is no danger of confusion.

2.3. **Tensor Products.** As it will be convenient to substitute matrices into each another in the way described above, let us analyze this briefly: For any two matrices $A_k \in \mathbb{C}^{m_k \times n_k}$, $k = 1, 2$, the tensor (or Kronecker) product $A_1 \otimes A_2$ is the matrix $B = (b_{ij}) \in \mathbb{C}^{m_1 m_2 \times n_1 n_2}$ such that

$$b_{(i_1-1)m_1+i_2,(j_1-1)n_1+j_2} = a_{i_1 j_1} a_{i_2 j_2}$$

for all $i_k \in [m_k]$, $j_k \in [n_k]$, $k = 1, 2$. Hence $B$ is produced by replacing every entry $a_{ij}$ of $A_1$ by $a_{ij} A_2$.

**Lemma 2.5.** *The following laws hold for the tensor product:*

(i) *Associativity: All matrices $A, B, C$ fulfill $(A \otimes B) \otimes C = A \otimes (B \otimes C)$.*
(ii) *Distributivity with $+$: For all matrices $A, B, C$ such that $A + B$ is defined we have $(A + B) \otimes C = A \otimes C + B \otimes C$ and $C \otimes (A + B) = C \otimes A + C \otimes B$.*
(iii) *'Mixed Product Rule': $(AB) \otimes (CD) = (A \otimes C)(B \otimes D)$ for all matrices $A, B, C, D$ such that $AB$ and $CD$ are defined.*
(iv) *$\otimes$ is compatible with inversion: $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ for all non-singular matrices $A$ and $B$.*
(v) *The (complex) eigenvalues of $A \otimes B$ are exactly the products of an eigenvalue of $A$ and one of $B$.*
(vi) $\operatorname{rank}(A \otimes B) = \operatorname{rank}(A) \operatorname{rank}(B)$.
(vii) $\det(A \otimes B) = (\det A)^{n_B} (\det B)^{n_A}$ *for all matrices $A \in \mathbb{C}^{n_A \times n_A}$ and $B \in \mathbb{C}^{n_B \times n_B}$.*

Some patience and knowledge of linear algebra is enough to prove Lemma 2.5. Most books on multilinear algebra will contain these results in chapters concerning tensor products of linear mappings and matrices. An elementary approach can be found in [Gra81].

In the tensor product notation e. g. equation (12) transforms to

$$\operatorname{disc}(\mathcal{H}, c) := \min_{\chi : X \to M_c} \|(A \otimes I_c)\chi\|_\infty . \tag{19}$$

2.4. **The Basic Probabilistic Bound.** An elementary probabilistic approach is to consider a random coloring. We color each vertex independently with a random color. Using the so-called Chernoff-bound, we prove that with positive probability our random coloring is balanced to a certain extent. Let $\mathcal{H} = (V, \mathcal{E})$ be any hypergraph. Set $m := |\mathcal{E}|$ and $s = \max_{E \in \mathcal{E}} |E|$. For the 2–color case it is well known that $\operatorname{disc}(\mathcal{H}) \leq \sqrt{2s \ln(2m)}$ holds ([AS92, Theorem 12.1.1]). In $c$ colors we have:

**Proposition 2.6.** $\operatorname{disc}(\mathcal{H}, c) \leq \sqrt{\frac{1}{2} s \ln(2mc)}$.

*Proof.* Define a random $c$–coloring $\chi$ by independently picking a random color uniformly distributed from $[c]$ for every vertex $v \in V$. Define random variables $X_{i,v}$ by

$$X_{i,v} := \begin{cases} \frac{c-1}{c} & \text{if } \chi(v) = i \\ -\frac{1}{c} & \text{else} \end{cases}$$

for all $v \in V$, $i \in [c]$. Set $X_{i,E} := \sum_{v \in E} X_{i,v}$ for all $E \in \mathcal{E}$, $i \in [c]$. From [AS92, Theorem A.1.4] we know

$$P(|X_{i,E}| > \alpha) < 2e^{-2\alpha^2/|E|}$$

for any real $\alpha > 0$.

For $\alpha = \sqrt{\frac{1}{2} s \ln(2mc)}$ we have

$$\begin{aligned} P(\forall i \in [c], E \in \mathcal{E} : |X_{i,E}| \leq \alpha) &= 1 - P(\exists i \in [c], E \in \mathcal{E} : |X_{i,E}| > \alpha) \\ &\geq 1 - \sum_{i \in [c], E \in \mathcal{E}} P(|X_{i,E}| > \alpha) \\ &> 1 - \sum_{i \in [c], E \in \mathcal{E}} 2e^{-2\alpha^2/|E|} \\ &\geq 1 - cm\, 2e^{-2\alpha^2/s} = 0 \end{aligned}$$

by choice of $\alpha$. Hence with positive probability our random $\chi$ has discrepancy not greater than $\sqrt{\frac{1}{2} s \ln(2mc)}$, thus such a coloring exists. □

## 3. Recursive Coloring

For some 2–color discrepancy results the proofs seem to rely heavily on the fact that only two colors are used. One example is Spencer's $\mathcal{O}(\sqrt{n})$ bound for hypergraphs having $n$ vertices and edges. A key step in the proof is to construct a low discrepancy partial coloring $\chi := \frac{1}{2}(\chi_1 - \chi_2)$ from two colorings $\chi_1, \chi_2$ with $\chi_1(E) \approx \chi_2(E)$ for all $E \in \mathcal{E}$. It is not clear to us how this idea can be generalized to $c$ colors.

As the partial coloring method has been a major break-through in 2–color discrepancy theory, it is desirable to have a similar method for $c$ colors as well. What we do in this section is not partial coloring, i. e. enlarging the partition classes by successively coloring points, but recursive 2–coloring, i. e. successively enlarging the number of partition classes. The basic idea is to find a suitable 2–coloring of $X$ with color classes $X_1, X_2$ and then to iterate this process on the subhypergraphs induced by $X_1$ and $X_2$. If the weighted 2–color discrepancies of suitable induced subhypergraphs are bounded, such a recursive method can be analyzed, even if $c$ is not a power of 2. This will lead to a generalization of the 'six standard deviations' theorem of Spencer [Spe85], the discrepancy bound of Beck–Fiala [BF81] and the bounds using the primal and dual shatter function of Matoušek, Welzl and Wernisch [MWW84] and Matoušek [Mat95].

At the end of this long section we will show the limits of the recursive approach. For example, for the linear discrepancy in $c$ colors recursive methods are rather weak, and we need other methods, which will be introduced in the next section.

3.1. **The Recursive Method.** The following lemma analyses a single step in the recursion. It shows that an imbalance inflicted in the first step of the recursion is evenly split up in the remainder of the partitioning process.

We call a function $p : C \to [0,1]$ a *weight* of the set $C$ of colors if $\sum_{i \in C} p_i = 1$. For $D \subseteq C$ set $p(D) := \|p_{|D}\|_1 = \sum_{i \in D} p_i$.

**Lemma 3.1.** *Let $C$ be a set of colors with $c = |C|$ and let $\{C_1, C_2\}$ be a partition of $C$. Let $p$ be a weight of $C$. Set $q_j = p(C_j)$, $j \in [2]$. Let $\chi_0 : X \to [2]$ be a 2–coloring of $X$. Set $X_j := \chi_0^{-1}(j)$, $j \in [2]$. Let $\chi_j : X_j \to C_j$ be any colorings. Set $\chi := \chi_1 \cup \chi_2$. For all $E \in \mathcal{E}$, $j \in [2]$ and $i \in C_j$ the discrepancy of $E$ with respect to the color $i$, the coloring $\chi$ and the weight $p$ is*

$$\big||E \cap \chi^{-1}(i)| - p_i|E|\big| \le \tfrac{p_i}{q_j}\,||E \cap X_j| - q_j|E|| + ||E \cap X_j \cap \chi_j^{-1}(i)| - \tfrac{p_i}{q_j}|E \cap X_j||.$$

*In particular*

$$\mathrm{wd}(\mathcal{H}, c, p) \le \max_{j \in [2], i \in C_j} \left( \tfrac{p_i}{q_j}\,\mathrm{wd}(\mathcal{H}, 2, (q_1, q_2)) + \max_{\mathcal{H}_0 \le \mathcal{H}} \mathrm{wd}(\mathcal{H}_0, |C_j|, \tfrac{1}{q_j} p_{|C_j}) \right).$$

*Proof.* Let $j \in [2]$, $i \in C_j$, $E \in \mathcal{E}$. Then

$$\begin{aligned} &\big||E \cap \chi^{-1}(i)| - p_i|E|\big| \\ &\quad = \big||E \cap X_j \cap \chi_j^{-1}(i)| - p_i|E|\big| \\ &\quad \le \left||E \cap X_j \cap \chi_j^{-1}(i)| - \tfrac{p_i}{q_j}|E \cap X_j|\right| + \left|\tfrac{p_i}{q_j}|E \cap X_j| - p_i|E|\right|. \end{aligned}$$

If the $\chi_j$, $j = 0,1,2$ are chosen such that $||E \cap X_j| - q_j|E|| \le \mathrm{wd}(\mathcal{H}, 2, (q_1, q_2))$ and $\left||E \cap X_j \cap \chi_j^{-1}(i)| - \frac{p_i}{q_j}|E \cap X_j|\right| \le \mathrm{wd}(\mathcal{H}_{|X_j}, |C_j|, \frac{1}{q_j} p_{|C_j})$ for all $E \in \mathcal{E}$, $j \in [2]$, $i \in C_j$, then the second claim follows from the first. □

As this section is quite lengthy, here is a short overview of what is going to come. We first analyze recursive coloring assuming that we have a uniform bound on the weighted 2–color discrepancies of the induced subhypergraphs. We derive a first result for the weighted $c$–color discrepancy and then improve it in the case of equi-weighted discrepancy. Finally we replace the uniformity assumption with the assumption that subhypergraphs on $n_0$ vertices have weighted discrepancy $\mathcal{O}(n_0^\alpha)$ for some $\alpha \in ]0,1[$. With this stronger precondition we get a number of beautiful results, among them a near tight $c$–color analogue of Spencer's 'six standard deviations' theorem.

3.2. **Weighted Discrepancy.** In the following two subsections we analyze the case that all induced subgraphs have a common bound on all weighted discrepancies in two colors. This is an important case for two reasons: Firstly, the proof of some results on two-color discrepancy provides some information about the weighted discrepancy of the induced subgraphs (e. g. in the Beck–Fiala setting). Secondly, the linear discrepancy and thus also the weighted discrepancies of all subgraphs are bounded by the hereditary discrepancy: From Proposition 2.1 we get

**Remark 3.2.** *For all induced subhypergraphs $\mathcal{H}_0$ of $\mathcal{H}$ we have* $\operatorname{wd}(\mathcal{H}_0, 2) \leq \operatorname{herdisc}(\mathcal{H})$.

Hence a bound on the hereditary discrepancy is also sufficient to apply the main theorems of these two subsection. Bounds on the hereditary discrepancy are often encountered in situations where the partial coloring method is used in the 2–color case — for the simple reason that the uncolored points induce a subhypergraph which has to be colored in the next iteration.

It will be convenient to represent the iterated partitioning of the set of colors $C$ by a binary tree. We call a binary rooted tree $T = (V_T, E_T)$ a *partition tree* for $C$, if the following conditions are satisfied: The root of $T$ is $C$, all nodes are subsets of $C$, all leaves are singletons of $C$ and each two son nodes form a partition of their common father node. For every color $i \in C$ there is a unique path $C = C_0^{(i)} \supset C_1^{(i)} \supset \ldots \supset C_{k(i)}^{(i)} = \{i\}$ in the partition tree. We write $h(T)$ for the height of $T$, that is the length of a longest path connecting a leaf and the root.

For a color $i \in C$ set $v(T,p,i) := \sum_{l=1}^{k(i)} \frac{p_i}{p(C_l^{(i)})}$ and $v(T,p) = \max_{i \in C} v(T,p,i)$. As the next theorem shows, these constants reflect the influence of the partition tree chosen for the recursive coloring process. In Lemma 3.4 and 3.6 we will give partition trees for which these values (and hence the resulting discrepancy) is small.

**Theorem 3.3.** *Let* $\operatorname{wd}(\mathcal{H}_0, 2) \leq K$ *for all induced subgraphs* $\mathcal{H}_0$ *of* $\mathcal{H}$*. Let* $C$ *be a set of colors with* $c = |C|$ *and let* $p$ *be a weight of* $C$*. Let* $T = (V_T, E_T)$ *be a partition tree of* $C$*. Then there is a coloring* $\chi : X \to C$ *such that for all colors* $i \in C$ *and all* $E \in \mathcal{E}$ *we have*

$$\big||E \cap \chi^{-1}(i)| - p_i|E|\big| \leq K v(T,p,i).$$

*In particular,* $\operatorname{wd}(\mathcal{H}, p, c) \leq K v(T,p)$.

*Proof.* We use induction on the height $h(T)$ of $T$. For $h(T) = 0$ we have just one color and both sides of the inequality become zero. Let $T$ be of height $h(T) > 0$ and assume that the theorem is true for all partition tree of height strictly less than $h(T)$. Let $C_1$ and $C_2$ be the sons of $C$ in $T$. Set $q_j := p(C_j) = \sum_{k \in C_j} p_k$, $j = 1, 2$. By assumption there is a 2–coloring $\chi_0 : X \to [2]$ such that

$$\big||E \cap \chi_0^{-1}(j)| - q_j|E|\big| \leq \operatorname{wd}(\mathcal{H}, 2, (q_1, q_2)) \leq K \tag{20}$$

holds for all $j \in [2]$ and $E \in \mathcal{E}$. Put $X_j := \chi_0^{-1}(j)$, $j = 1, 2$. Denote by $T_j$ the subtree having $C_j$ as its root. Then the hypergraph $\mathcal{H}_{|X_j}$ together with the set of colors $C_j$, the weight $\frac{1}{q_j}p_{|C_j}$ and the partition tree $T_j$ fulfills the assumption of this theorem. By induction there are colorings $\chi_j : X_j \to C_j$, $j \in 1,2$ such that

$$\left| |E \cap X_j \cap \chi_j^{-1}(i)| - \tfrac{1}{q_j} p_i |E \cap X_j| \right| \leq K v(T_j, \tfrac{1}{q_j} p_{|C_j}, i) \leq K \sum_{l=2}^{k(i)} \frac{\frac{p_i}{q_j}}{\frac{1}{q_j} p(C_l^{(i)})} \tag{21}$$

for all $i \in C_j$. Set

$$\chi = \chi_1 \cup \chi_2 : x \mapsto \begin{cases} \chi_1(x) & \text{if } x \in X_1 \\ \chi_2(x) & \text{otherwise.} \end{cases}$$

Let $j \in [2]$ and $i \in C_j$. Then $C_1^{(i)} = C_j$ and $q_j = p(C_1^{(i)})$. Let $E \in \mathcal{E}$. From (20), (21) and Lemma 3.1 we get

$$\begin{aligned} \left| |E \cap \chi^{-1}(i)| - p_i |E| \right| &\leq \left| |E \cap X_j \cap \chi_j^{-1}(i)| - \tfrac{p_i}{q_j} |E \cap X_j| \right| + \tfrac{p_i}{q_j} \left| |E \cap X_j| - q_j |E| \right| \\ &\overset{(20),(21)}{\leq} \sum_{l=2}^{k(i)} \frac{\frac{p_i}{q_j}}{\frac{1}{q_j} p(C_l^{(i)})} K + \tfrac{p_i}{q_j} K \\ &= K \sum_{l=1}^{k(i)} \frac{p_i}{p(C_l^{(i)})} = K v(T, p, i). \end{aligned}$$

Hence $\chi$ satisfies the claim. □

In the following corollary we give a first upper bound on the constant $v(T, p, i)$. An improvement in the case of equi-weighted discrepancy will be discussed in more detail in Subsection 3.3.

**Lemma 3.4.** *In the situation of Theorem 3.3 there is a partition tree $T$ such that*

$$v(T, p, i) < 4$$

*for all $i \in [c]$. Thus* $\mathrm{wd}(\mathcal{H}, p, c) < 4K$.

*Proof.* Recursively we construct a partition tree $T$ for $C$ with $v(T, p) \leq 4$. We start with the tree consisting of the unique node $C$. For a leaf $C_0$ of cardinality greater than 1 let us define sons by the following rule: If there is a color $i \in C_0$ with weight $p_i \geq \frac{1}{2} p(C_0)$, then the sons of $C_0$ shall be $\{i\}$ and $C_0 \setminus \{i\}$. Otherwise partition $C_0$ in any way $(C_1, C_2)$ such that $p(C_j) \in [\frac{1}{3} p(C_0), \frac{2}{3} p(C_0)]$. Repeat this process until all leaves are singletons. The resulting tree $T$ is a partition tree for $C$. All father-son pairs $(C_0, C_1)$ in the resulting tree fulfill $\frac{2}{3} p(C_0) \geq p(C_1)$ or $|C_1| = 1$ and $p(C_0) > p(C_1)$. In the notation of Theorem 3.3 we have $p(C_{k(i)}^{(i)}) = p_i$, $p(C_{k(i)-1}^{(i)}) \geq p_i$ and $p(C_{k(i)-1-l}^{(i)}) \geq \left(\frac{3}{2}\right)^l p_i$ for all $l \in [k(i) - 1]$. Now

$$v(T, p) \leq \max_{i \in C} \sum_{l=1}^{k(i)} \frac{p_i}{p(C_l^{(i)})} \leq \max_{i \in C} p_i \left( \frac{1}{p_i} + \sum_{l=0}^{k(i)-2} \frac{1}{\left(\frac{3}{2}\right)^l p_i} \right) \leq 1 + \sum_{l=0}^{h(T)-2} \left(\tfrac{2}{3}\right)^l < 4,$$

and Theorem 3.3 gives the bound $\mathrm{wd}(\mathcal{H}, p, c) \le 4K$. □

3.3. **Equi-Weighted Discrepancy.** In this subsection we consider the case of equi-weighted discrepancy in $c$ colors. Hence our assumptions are identical with the ones from the preceding subsection except that we always have $p = \frac{1}{c}\mathbf{1}_c$. In this case only the size of the color sets is important, as all colors are equivalent. Therefore the following simpler structure can be investigated:

A *partition tree* for a positive integer $n$ is a binary tree $T = (V_T, E_T)$ together with a labeling $l : V_T \to [n]$ such that the following conditions are satisfied:

- The root $r$ is labeled $l(r) = n$.
- For every non-leaf $v$ with sons $s_1$ and $s_2$ we have $l(v) = l(s_1) + l(s_2)$.
- The leaves are labeled 1.

Note that we can not assume $l$ to be injective anymore. For a path $P : r = v_0^{(i)}, v_1^{(i)}, \ldots, v_{k(i)}^{(i)}$ connecting the root $r$ and a leaf $v_{k(i)}^{(i)}$ labeled $i$ we call $v(T, P) = \sum_{l=1}^{k(i)} \frac{1}{l(v_l^{(i)})}$ the value of $P$ and $v(T)$ the maximum $v(T, P)$ over all these paths $P$. Finally $v(n)$ is the minimum $v(T)$ over all partition trees $T$ of $n$.

There is a natural correspondence between partition trees for sets of colors and for positive integers. Let $T = (V_T, E_T)$ denote a partition tree for the set of colors $C$. Define a labeling $l_T : V_T \to [|C|]; v \mapsto |v|$. Then $T$ together with $l_T$ is a partition tree for $|C|$.

Now let $T$ together with $l$ denote a partition tree for a positive integer $c$. Let $C$ be any set of colors such that $|C| = c$. We construct a partition tree $T^*$ for $C$ such that $l_{T^*} = l$. Define $f : V_T \to 2^{[c]}$ recursively: Set $f(r) = C$ for the root $r$ of $T$. For every node $v$ with sons $s_1$ and $s_2$ such that $f(v)$ is already defined choose $f(s_1)$ to be any subset of $f(v)$ of size $l(s_1)$ and $f(s_2) = f(v) \setminus f(s_1)$. Note that $f$ is injective, and by replacing every $v \in V_T$ by $f(v)$ we get a partition tree $T^*$ for $C$. Clearly, $l_{T^*} = l$.

Furthermore, we have

$$v(T^*, \tfrac{1}{c}\mathbf{1}_c) = \max_{i \in C} \frac{1}{c} \sum_{l=1}^{k(i)} \frac{1}{\frac{1}{c}|C_l^{(i)}|} = \max_{i \in C} \sum_{l=1}^{k(i)} \frac{1}{l(v_l^{(i)})} = v(T).$$

**Corollary 3.5.** *Let* $\mathrm{wd}(\mathcal{H}_0, 2) \le K$ *for all induced subgraphs* $\mathcal{H}_0$ *of* $\mathcal{H}$.

*Then* $\mathrm{disc}(\mathcal{H}, c) \le v(c)K$.

*Proof.* Let $T = (V_T, E_T)$ together with $l$ be a partition tree for $c$ such that $v(T) = v(c)$. We build $T^*$ as above and apply Theorem 3.3 on $T^*$ and $p = \frac{1}{c}\mathbf{1}_c$:

$$\mathrm{disc}(\mathcal{H}, c) = \mathrm{wd}(\mathcal{H}, p, c) \leq Kv(T^*, p) = Kv(T) = Kv(c).$$

□

The exact calculation of $v(c)$ seems to be a difficult task. In particular, the optimal partition trees are in general not of minimal height. Put $\lfloor c \rfloor_2 := 2^{\lfloor \log_2 c \rfloor}$ and $\lceil c \rceil_2 := 2^{\lceil \log_2 c \rceil}$. Denote by $n_1(c)$ the number of 1's in the binary expansion of $c$ (e. g. $n_1(9) = 2$). We give a lower bound and an upper bound on $v(c)$. If $c$ is a power of 2, both bounds coincide.

**Lemma 3.6.** *For all $c \in \mathbb{N}$, $c \geq 2$ we have*

$$2 - \tfrac{2}{\lceil c \rceil_2} \leq v(c) \leq 2 + (n_1(c) - 3)\tfrac{1}{\lfloor c \rfloor_2}.$$

*In particular, $v(c) \leq 2.0005$.*

*Proof.* Let $T = (V_T, E_T)$ together with $l$ be any partition tree for $c$. Then there is a path $v_0, \ldots, v_k$ of length $k \geq \log_2 \lceil c \rceil_2$ such that $v_k$ is a leaf and $l(v_{i-1}) \leq 2l(v_i)$ for all $i \in [k]$. Thus $\sum_{i=1}^{k} \frac{1}{l(v_i)} \geq \sum_{i=0}^{k-1} 2^{-i} = 2 - \frac{1}{2^{k-1}} \geq 2 - \frac{1}{\lceil c \rceil_2}$.

For the upper bound we recursively construct a partition tree $T$ for $c$. For a vertex $v$ labeled $\sum_{i \in [k]} a_i 2^k \neq 1$, $a_i \in \{0, 1\}$, we add sons $s_1(v)$ and $s_2(v)$ labeled $l(s_1(v)) = 2^{\min\{i \in [k] | a_i = 1\}}$ and $l(s_2(v)) = l(v) - l(s_1(v))$, if $l(v)$ is not a power of two, and labeled $l(s_1(v)) = l(s_2(v)) = \frac{1}{2}l(v)$ otherwise. Immediately we see that we only need to investigate the path $P : r, s_2(r), s_2(s_2(r)), \ldots$ — if $r$ denotes the root of $T$ —, because the labels of all other paths occur also on this path. Thus $v(P)$ is maximal. The labels of the first $n_1(c)$ vertices of $P$ are greater than or equal to $\lfloor c \rfloor_2$, so their contribution to $v(P)$ is not greater than $(n_1(c) - 1)\frac{1}{\lfloor c \rfloor_2}$. The rest of the vertices are labeled by $\frac{2}{\lfloor c \rfloor_2}, \frac{4}{\lfloor c \rfloor_2}, \ldots$ up to 1. This sums up to $2 - \frac{2}{\lfloor c \rfloor_2}$ and the inequality is proven.

The last assertion is clear for $c \geq c_0 := 2^{15} - 1$, as $(n_1(c) - 3)\frac{1}{\lfloor c \rfloor_2} \leq \frac{\log_2(\lfloor c \rfloor_2) - 2}{\lfloor c \rfloor_2} \leq \frac{\log_2(\lfloor c_0 \rfloor_2) - 2}{\lfloor c_0 \rfloor_2}$. For the remaining small numbers, $v(c)$ can be computed in $\mathcal{O}(c^2)$–time and attains its maximum value for $c = 909$, namely $v(909) \approx 2.000450$. □

Now Corollary 3.5 and Lemma 3.6 yield

**Theorem 3.7.** *Let $\mathrm{wd}(\mathcal{H}_0, 2) \leq K$ for all induced subgraphs $\mathcal{H}_0$ of $\mathcal{H}$. Then $\mathrm{disc}(\mathcal{H}, c) \leq 2.0005K$ holds for any number $c$ of colors.*

We apply Theorem 3.7 on the Beck–Fiala setting and get

**Theorem 3.8.** *For any hypergraph $\mathcal{H}$ we have*

$$\mathrm{disc}(\mathcal{H}, c) < v(c)\,\Delta(\mathcal{H}) \leq 2.0005\,\Delta(\mathcal{H}).$$

*Proof.* The Beck–Fiala theorem states that $\operatorname{lindisc}(\mathcal{H}) < 2\Delta(\mathcal{H})$ holds for any hypergraph $\mathcal{H}$. In particular, we have $\operatorname{wd}(\mathcal{H}_0, 2) \leq \frac{1}{2} \operatorname{lindisc}(\mathcal{H}_0) < \Delta(\mathcal{H}_0) \leq \Delta(\mathcal{H})$ for all induced subhypergraphs $\mathcal{H}_0$ of $\mathcal{H}$. From Corollary 3.5 and Lemma 3.6 we conclude $\operatorname{disc}(\mathcal{H}, c) \leq v(c) \operatorname{wd}(\mathcal{H}, 2) < 2.0005\Delta$. □

A similar result is proven in Section 4. Note that Theorem 3.7 also yields the following bounds:

- For any hypergraph $\mathcal{H} = (V, \mathcal{E})$ with $n := |V| = |\mathcal{E}|$ sufficiently large we have
$$\operatorname{disc}(\mathcal{H}, c) < 12\sqrt{n}.$$
- Let $\mathcal{H} = (V, \mathcal{E})$ be a hypergraph on $n$ points. Let $d > 1$. If $\pi_{\mathcal{H}} = \mathcal{O}(m^d)$, then $\operatorname{disc}(\mathcal{H}, c) = \mathcal{O}(n^{\frac{1}{2} - \frac{1}{2d}})$. If $\pi^*_{\mathcal{H}} = \mathcal{O}(m^d)$, then $\operatorname{disc}(\mathcal{H}, c) = \mathcal{O}(n^{\frac{1}{2} - \frac{1}{2d}} \log n)$. In both cases the implicit constants are independent of $c$.
- The hypergraph $\mathcal{A}_n$ of arithmetic progressions in $[n]$ fulfills
$$\operatorname{disc}(\mathcal{A}_n, c) \leq v(c) C \sqrt[4]{n} \leq 2.0005 C \sqrt[4]{n},$$
where $C$ is the constant of Matoušek and Spencer such that $\operatorname{disc}(\mathcal{A}_n) \leq C\sqrt[4]{n}$.

Using the fact that in these cases the discrepancies of smaller induced subhypergraphs are decreasing, we improve these bounds in the next two subsections.

The following example shows that the recursive approach is nearly optimal in the general case. Let $n = kc$ for some $k \in \mathbb{N}$. Set $\mathcal{H} = \left([n], \binom{[n]}{k}\right) = ([n], \{E \subseteq [n] \,|\, |E| = k\})$. Any $c$–coloring for $\mathcal{H}$ produces a monochromatic hyperedge which has discrepancy $k(1 - \frac{1}{c})$. Hence $\operatorname{disc}(\mathcal{H}, c) \geq k(1 - \frac{1}{c})$. Now let $(p, 1-p)$ be any 2–color weight. Assume without loss of generality that $p \leq \frac{1}{2}$. Put $\chi : [n] \to [2]; i \mapsto 2$. Now each hyperedge has weighted discrepancy $k(1-p)$ with respect to $\chi$ and $(p, 1-p)$. Thus $\operatorname{wd}(\mathcal{H}, c) \leq \frac{1}{2}k$, and of course this holds as well for any induced subhypergraph $\mathcal{H}_0$ of $\mathcal{H}$. This shows
$$\operatorname{disc}(\mathcal{H}, c) \geq 2(1 - \tfrac{1}{c}) \max_{\mathcal{H}_0 \leq \mathcal{H}} \operatorname{wd}(\mathcal{H}_0, 2).$$
In particular, the recursive method yields optimal colorings in this case if $c$ is a power of 2, and it is asymptotically optimal for $c \to \infty$.

3.4. **Refined Recursive Coloring.** In this subsection we extend the recursive approach to make use of the additional assumption that subhypergraphs on fewer vertices have smaller discrepancy. This is a natural assumption as many results are of this type (see Subsection 3.5 where we prove their multi-color analogies).

Roughly speaking we show that if the 2–color discrepancy of the subhypergraphs on $n_0$ vertices is bounded by $\mathcal{O}(n_0^\alpha)$, then the $c$–color discrepancy is bounded by $\mathcal{O}((\frac{n}{c})^\alpha)$. It seems a little surprising that this bound is achievable by a recursive approach, as the first step in the recursion will find a 2–coloring for the whole hypergraph with discrepancy guarantee $\mathcal{O}(n^\alpha)$ only. We still get the $\mathcal{O}((\frac{n}{c})^\alpha)$–discrepancy for the final coloring due to the fact that imbalances inflicted in earlier rounds of the recursion are split up in a balanced manner by later steps (cf. Lemma 3.1). It turns out that this effect even exceeds the effect of decreasing discrepancy of smaller subhypergraphs. Crucial therefore is the last step of the recursion where colorings for hypergraphs on roughly $\frac{2n}{c}$ vertices are looked for.

There are two points though that need further attention: Firstly, like in the case where we only assumed a uniform bound on the discrepancies of the induced subhypergraphs, this simple approach only works if the number of colors is a power of 2. This is the reason why we have to use weighted discrepancies again.

A second point is that to use the assumption of decreasing discrepancies we need to make sure that the vertex sets considered actually become smaller. Unfortunately, in general we do not know the size of the color classes generated by a low discrepancy coloring. If the whole vertex set is a hyperedge, we know at least that the sizes of the color classes deviate from the aimed at value by at most the discrepancy guarantee. This is not too bad if the discrepancy is relatively small, but even then keeping track of these deviations during the recursion is tedious. Better bounds seem achievable by the cleaner approach of only investigating *fair* colorings, that is, those which have discrepancy less than one on the set of all vertices.

To ease notation let us agree the following. Let $p \in [0,1]^c$ be a $c$–color weight and $\mathcal{H} = (X, \mathcal{E})$ a hypergraph. We say that *$\chi$ is a fair $p$–coloring of $\mathcal{H}$ having discrepancy at most $d_i$ in color $i \in [c]$* to denote that

- $\chi$ is a $c$–coloring of $\mathcal{H}$,
- $\chi$ is fair with respect to $p$, that is, for all $i \in [c]$ we have $|\,|\chi^{-1}(i)| - p_i|X|\,| \le 1$,
- the discrepancy of $\mathcal{H}$ with respect to $\chi$ and $p$ in color $i \in [c]$ is at most $d_i$.

One remark that eases work with the fractional parts: Let us call a weight $p \in [0,1]^c$ *integral* with respect to $\mathcal{H}$ (or $\mathcal{H}$–integral for short) if all $p_i, i \in [c]$ are multiples of $\frac{1}{|X|}$. From the definition it is clear that a fair coloring $\chi$ with respect to an integral weight $p$ fulfills $|\chi^{-1}(i)| = p_i|X|$ for all colors $i \in [c]$. Suppose that we know that for a given hypergraph and for all integral weights $p$ there is a fair $p$–coloring that has discrepancy at most $k$. Then there are fair colorings having discrepancy at most $k+1$ for any weight: For an arbitrary weight $p$ there is an integral weight $p'$ such that $|p_i - p'_i| < \frac{1}{|X|}$ holds for all $i \in [c]$. Therefore, a fair coloring with respect to $p'$ is also fair with respect to $p$, and its discrepancy with respect to $p$ is larger (if at all) than the one with respect to $p'$ by less than one. For these reasons we may restrict ourselves to the more convenient case that all weights are integral.

Using the following recoloring argument we can transform arbitrary colorings into fair colorings.

**Lemma 3.9.** *Let $\mathcal{H} = (X, \mathcal{E})$ be a hypergraph such that $X \in \mathcal{E}$. Let $p$ be a 2–color weight. Then any 2–coloring $\chi$ of $\mathcal{H}$ can be modified in $\mathcal{O}(|X|)$ time into a fair $p$–coloring $\overline{\chi}$ such that*

$$\mathrm{wd}(\mathcal{H}, \overline{\chi}, p) \leq 2\,\mathrm{wd}(\mathcal{H}, \chi, p).$$

*Proof.* Let $\chi$ be a coloring such that $\mathrm{wd}(\mathcal{H}, \chi, p) = \mathrm{wd}(\mathcal{H}, c, p)$. Set $x := q|X| - |\chi^{-1}(1)|$. Since $X$ is an edge in $\mathcal{H}$, $|x| \leq \mathrm{wd}(\mathcal{H}, c, p)$. Let $\overline{\chi}$ denote a coloring arising from $\chi$ by changing the color of $\lfloor |x| \rfloor$ points in such a way that $|q|X| - |\overline{\chi}^{-1}(1)|| < 1$. Now $\overline{\chi}$ is a fair coloring with respect to the weight $(q, 1-q)$. For an edge $E \in \mathcal{E}$ we compute

$$\begin{aligned} &|\, q|E| - |\overline{\chi}^{-1}(1) \cap E|| \\ &\leq \ |\, q|E| - |\chi^{-1}(1) \cap E|| + ||\chi^{-1}(1) \cap E| - |\overline{\chi}^{-1}(1) \cap E|| \\ &\leq \ |\, q|E| - |\chi^{-1}(1) \cap E|| + \lfloor |x| \rfloor \\ &\leq \ 2\,\mathrm{wd}(\mathcal{H}, c, p). \end{aligned}$$

□

Lemma 3.9 requires the whole vertex set to be a hyperedge. Fortunately, most discrepancy results are relatively robust concerning the addition of a single hyperedge. In these cases we may just replace the hypergraph under consideration by the one obtained from adding $X$ as additional edge.

To analyze our recursive algorithm we need the following constants. Let $\alpha \in ]0,1[$. For each $p \in ]0,1[$ define $v_\alpha(p)$ by

$$v_\alpha(p) = \max\left\{ \sum_{i=1}^{k} \prod_{j=1}^{i} q_j^\alpha \prod_{j=i+1}^{k} q_j \,\middle|\, k \in \mathbb{N}, q_1, \ldots, q_{k-1} \in [0, \tfrac{2}{3}], q_k \in [0,1], \prod_{j=1}^{k} q_j = p \right\}.$$

Set $c_\alpha := \frac{2}{2^{1-\alpha}-1}\left(1 + \frac{1}{1-\left(\frac{2}{3}\right)^{(1-\alpha)}}\right)$. Then we have

**Lemma 3.10.** *Let $\alpha \in ]0,1[$.*

(i) *Let $0 < p < q \leq \frac{2}{3}$. Then $q^\alpha v_\alpha(\frac{p}{q}) + q^\alpha \frac{p}{q} \leq v_\alpha(p)$.*
(ii) *For all $p \in [0,1]$, $\frac{2}{2^{1-\alpha}-1} v_\alpha(p) \leq c_\alpha p^\alpha$.*

*Proof.* Let $k \in \mathbb{N}, q_1, \ldots, q_{k-1} \in [0, \frac{2}{3}], q_k \in [0,1]$ such that $\prod_{j=1}^k q_j = \frac{p}{q}$ and $v_\alpha(\frac{p}{q}) = \sum_{i=1}^k \prod_{j=1}^i q_j^\alpha \prod_{j=i+1}^k q_j$. With $q_0 := q$ we have

$$\begin{aligned} q^\alpha v_\alpha(\tfrac{p}{q}) + q^\alpha \tfrac{p}{q} &= q_0^\alpha \sum_{i=1}^k \prod_{j=1}^i q_j^\alpha \prod_{j=i+1}^k q_j + q_0^\alpha \prod_{j=1}^k q_j \\ &= \sum_{i=0}^k \prod_{j=0}^i q_j^\alpha \prod_{j=i+1}^k q_j \le v_\alpha(p), \end{aligned}$$

since $\prod_{j=0}^k q_j = q \prod_{j=1}^k q_j = p$. This is $(i)$.

Let $k \in \mathbb{N}, q_1, \ldots, q_{k-1} \in [0, \frac{2}{3}], q_k \in [0,1]$ such that $\prod_{j=1}^k q_j = p$ and $v_\alpha(p) = \sum_{i=1}^k \prod_{j=1}^i q_j^\alpha \prod_{j=i+1}^k q_j$. For $i \in [k]$ set $x_i := \prod_{j=1}^i q_j^\alpha \prod_{j=i+1}^k q_j$. Then $x_k = p^\alpha$ and $x_{k-1} \le x_k$. For $i \in [k-2]$ we have

$$\frac{x_i}{x_{i+1}} = \frac{q_{i+1}}{q_{i+1}^\alpha} = q_{i+1}^{1-\alpha} \le \left(\tfrac{2}{3}\right)^{1-\alpha},$$

and hence $x_{k-1-i} \le \left(\frac{2}{3}\right)^{(1-\alpha)i} x_k$. Thus

$$\tfrac{2}{2^{1-\alpha}-1} v_\alpha(p) = \tfrac{2}{2^{1-\alpha}-1} \sum_{i=1}^k x_i \le \tfrac{2}{2^{1-\alpha}-1} \left(1 + \sum_{i=0}^{k-2} \left(\tfrac{2}{3}\right)^{(1-\alpha)i}\right) x_k < c_\alpha p^\alpha.$$

□

Here is the precise setting we investigate in this section:

**Assumption** (Decreasing-Discrepancies-Assumption)**.** *Let $\mathcal{H} = (X, \mathcal{E})$ be a hypergraph. Set $n := |X|$. Let $p_0, \alpha \in ]0,1[$ and $D > 0$. For all $X_0 \subseteq X$ such that $|X_0| \ge p_0|X|$ and all $q \in [0,1]$ such that $(q, 1-q)$ is $\mathcal{H}_{|X_0}$–integral there is a fair $(q, 1-q)$–coloring $\chi$ of $\mathcal{H}_{|X_0}$ having discrepancy at most $D|X_0|^\alpha$.*

In addition to what we already explained there is one further detail involved in our assumption. As we do recursive partitioning, we never need a discrepancy result concerning induced subhypergraphs on fewer than roughly $\frac{n}{c}$ vertices (in the equi-weighted case). This observation will be useful in some applications, e. g. in the case $|\mathcal{E}| = |X|$.

Concerning the computational complexity there are two possible measures. We can count how many 2–colorings have to be computed, or how often a 2–coloring for a vertex has to be found. The latter is useful if the complexity of computing the 2–colorings is proportional to the number of vertices of the induced subhypergraph as in Theorem 3.14.

**Theorem 3.11.** *Suppose that the Decreasing-Discrepancies-Assumption holds. Then for each $\mathcal{H}$–integral weight $p \in [0,1]^c$ there is a fair $p$–coloring $\chi$ of $\mathcal{H}$ such that the discrepancy* *is at most* $\frac{2}{2^{1-\alpha}-1} D v_\alpha(p_i) n^\alpha \leq D c_\alpha(p_i n)^\alpha$ *in all those colors* $i \in [c]$ *such that* $p_i \geq p_0$. *Such colorings can be obtained by computing at most* $(c-1)\left\lceil \log_2(\frac{1}{p_0})\right\rceil$ *colorings as in the Decreasing-Discrepancies-Assumption. At most* $3n \log_{1.5}(\frac{1}{p_0})$ *times a color for a vertex has to be computed.*

For the proof we first show a stronger bound for the 2–color discrepancy with respect to a weight $(q, 1-q)$, if $q$ is small.

**Lemma 3.12.** *Suppose that the Decreasing-Discrepancies-Assumption holds. Then for each* $k \in \mathbb{N}$ *such that* $p = (2^{-k}, 1-2^{-k})$ *is an* $\mathcal{H}$*–integral weight and* $2^{-k} \geq p_0$*, a fair* $p$*–coloring* $\chi$ *having discrepancy at most*

$$\mathrm{wd}(\mathcal{H}, \chi, p) \leq \sum_{i=0}^{k-1} 2^{-k+1+i} 2^{-\alpha i}\, D n^\alpha$$

*can be computed from* $k$ *colorings as in the Decreasing-Discrepancies-Assumption. This requires* $\sum_{i=0}^{k-1} 2^{-i} n \leq 2n$ *times computing a color for a vertex.*

*Proof.* We proceed by induction. For $k = 1$, there is nothing to show. Let $k > 1$. Let $\chi_0 : X \to [2]$ be a fair $(0.5, 0.5)$–coloring having discrepancy at most $Dn^\alpha$. Set $X_1 := \chi_0^{-1}(1)$. Let $\chi_1 : X_1 \to [2]$ be a fair $(2^{-k+1}, 1-2^{-k+1})$–coloring. Note that $(2^{-k+1}, 1-2^{-k+1})$ is integral for $\mathcal{H}_{|X_1}$. By induction we may assume that $\chi_1$ has discrepancy at most $\sum_{i=0}^{k-2} 2^{-k+2+i} 2^{-\alpha i} D(\frac{n}{2})^\alpha$. Define a coloring $\chi : X \to [2]$ by $\chi(x) = 1$ if and only if $\chi_0(x) = 1$ and $\chi_1(x) = 1$. Then $\chi$ is a fair $(2^{-k}, 1-2^{-k})$–coloring. Using a similar argument as in Lemma 3.1, we compute the discrepancy of an edge $E \in \mathcal{E}$ with respect to $(2^{-k}, 1-2^{-k})$ in color 1:

$$\begin{aligned}
&\big||E \cap \chi^{-1}(1)| - 2^{-k}|E|\big| \\
&\quad= \big||E \cap \chi_0^{-1}(1) \cap \chi_1^{-1}(1)| - 2^{-k}|E|\big| \\
&\quad\leq \big||E \cap \chi_0^{-1}(1) \cap \chi_1^{-1}(1)| - 2^{-k+1}|E \cap \chi_0^{-1}(1)|\big| + \big|2^{-k+1}|E \cap \chi_0^{-1}(1)| - 2^{-k}|E|\big| \\
&\quad\leq \big||(E \cap X_1) \cap \chi_1^{-1}(1)| - 2^{-k+1}|E \cap X_1|\big| + 2^{-k+1}\big||E \cap \chi_0^{-1}(1)| - 0.5|E|\big| \\
&\quad\leq \sum_{i=0}^{k-2} 2^{-k+2+i} 2^{-\alpha i} D(\tfrac{n}{2})^\alpha + 2^{-k+1} D n^\alpha \\
&\quad= \sum_{i=0}^{k-1} 2^{-k+1+i} 2^{-\alpha i} D n^\alpha.
\end{aligned}$$

As 2–colorings have the same discrepancy in both colors, this proves Lemma 3.12. □

From our assumptions on $\mathcal{H}$ it is clear that the assertion of Lemma 3.12 also holds for any induced subhypergraph $\mathcal{H}_{|X_0}$ of $\mathcal{H}$ as long as $2^{-k}|X_0| \geq p_0|X|$. We use this fact to extend Lemma 3.12 to arbitrary weights.

**Lemma 3.13.** *Suppose that the Decreasing-Discrepancies-Assumption holds. For each $\mathcal{H}$–integral weight $(q, 1-q)$, $p_0 \leq q \leq \frac{1}{2}$, there is a fair $(q, 1-q)$–coloring $\chi$ having discrepancy at most*

$$\operatorname{wd}(\mathcal{H}, \chi, p) \leq \frac{2}{2^{1-\alpha}-1} D\,(qn)^{\alpha}.$$

*A coloring of this kind can be computed by $\left\lceil \log_2(\frac{1}{q}) \right\rceil$ times computing a coloring as in the Decreasing-Discrepancies-Assumption. This requires at most $3n$ times computing a color for a vertex.*

*Proof.* Let $k \in \mathbb{N}_0$ be maximal subject to the condition that $q' = 2^k q \leq 1$. Since $(q, 1-q)$ is $\mathcal{H}$–integral, so is $(q', 1-q')$. According to the Decreasing-Discrepancies-Assumption there is a fair $(q', 1-q')$–coloring $\chi_0 : X \to [2]$ having discrepancy at most $Dn^{\alpha}$. From $|\chi_0^{-1}(1)| = q'|X|$ we have $\frac{q}{q'}|\chi_0^{-1}(1)| = q|X| \in \mathbb{N}_0$. Hence $(\frac{q}{q'}, 1-\frac{q}{q'})$ is $(\mathcal{H}_{|\chi_0^{-1}(1)})$–integral. By Lemma 3.12 we may compute a fair $(\frac{q}{q'}, 1-\frac{q}{q'})$–coloring $\chi_1 : \chi_0^{-1}(1) \to [2]$ that has discrepancy at most $\sum_{i=0}^{k-1} 2^{-k+1+i}2^{-\alpha i}D(q'n)^{\alpha}$. Define a coloring $\chi : X \to [2]$ by $\chi(x) = 1$ if and only if $\chi_0(x) = 1$ and $\chi_1(x) = 1$. Then $\chi$ is a fair $(q, 1-q)$–coloring. For an edge $E \in \mathcal{E}$ we compute its discrepancy in color 1:

$$\begin{aligned}
&\big||E \cap \chi^{-1}(1)| - q|E|\big| \\
&= \big||E \cap \chi_0^{-1}(1) \cap \chi_1^{-1}(1)| - q|E|\big| \\
&\leq \left||E \cap \chi_0^{-1}(1) \cap \chi_1^{-1}(1)| - \tfrac{q}{q'}|E \cap \chi_0^{-1}(1)|\right| + \left|\tfrac{q}{q'}|E \cap \chi_0^{-1}(1)| - q|E|\right| \\
&= \big||E \cap \chi_0^{-1}(1) \cap \chi_1^{-1}(1)| - 2^{-k}|E \cap \chi_0^{-1}(1)|\big| + 2^{-k}\,\big||E \cap \chi_0^{-1}(1)| - q'|E|\big| \\
&\leq \sum_{i=0}^{k-1} 2^{-k+1+i}2^{-\alpha i}D(q'n)^{\alpha} + 2^{-k}Dn^{\alpha} \\
&< \sum_{i=0}^{k-1} 2^{-k+1+i}2^{-\alpha i}D(q'n)^{\alpha} + 2^{-k}2^{\alpha}D(q'n)^{\alpha} \\
&< 2q'^{\alpha}\frac{2^{-\alpha k}}{2^{1-\alpha}-1}Dn^{\alpha} = \frac{2}{2^{1-\alpha}-1}\,D(qn)^{\alpha}.
\end{aligned}$$

Note that if $q' = 1$, then we may compute $\chi$ directly using Lemma 3.12. Therefore the computation of $\chi$ requires $\left\lceil \log_2(\frac{1}{q}) \right\rceil$ times computing a coloring assured by the Decreasing-Discrepancies-Assumption. Computing $\chi_0$ means computing a color for $n$ vertices. By Lemma 3.12, $\chi_1$ can be computed by at most $2q'n$ times computing a color for a vertex.

To get $\chi$ we therefore computed at most $3n$ times a color for a vertex. This proves Lemma 3.13. $\square$

*Proof of Theorem 3.11.* To make the recursion work properly we need to fix a set $C$ of colors at the beginning. A weight then is a vector $p = (p_i)_{i \in C}$ indexed by the colors, or, more formally, a function $p : C \to [0,1]$, such that $\|p\|_1 = \sum_{i \in C} p_i = 1$. To avoid trivial cases we shall always assume that no color $i \in C$ has the weight $p_i = 0$.

We analyze the following recursive algorithm:

**Input:** A hypergraph $\mathcal{H} = (X, \mathcal{E})$ fulfilling the Decreasing-Discrepancies-Assumption, a set $C$ of at least 2 colors and an $\mathcal{H}$–integral weight function $p : C \to [0,1]$.
**Output:** A coloring $\chi : X \to C$ as in Theorem 3.11.
**1:** Choose a partition $\{C_1, C_2\}$ of the set of colors $C$ such that $\|p_{|C_1}\|_1, \|p_{|C_2}\|_1 \le \frac{2}{3}$ or $C_1$ contains a single color with weight at least $\frac{1}{3}$. Set $(q_1, q_2) := (\|p_{|C_1}\|_1, \|p_{|C_2}\|_1)$.
**2:** Following Lemma 3.13, compute a fair $(q_1, q_2)$–coloring $\chi_0 : X \to [2]$ that has discrepancy at most $\frac{2}{2^{1-\alpha}-1} D(q_i n)^\alpha$ in color $i = 1, 2$ if $q_i \ge p_0$. Set $X_i := \chi^{-1}(i)$ for $i = 1, 2$.
**3:** For $i = 1, 2$ do
**if:** $|C_i| > 1$,
**then:** by recursion compute a fair $\frac{1}{q_i} p_{|C_i}$–coloring $\chi_i : X_i \to C_i$ for $\mathcal{H}_{|X_i}$ having discrepancy at most $\frac{2}{2^{1-\alpha}-1} D v_\alpha(\frac{p_j}{q_i})(q_i n)^\alpha$ in each color $j \in C_i$ such that $p_j \ge p_0$
**else:** if $C_i = \{j\}$ for some $j \in C$, choose $\chi_i : X_i \to \{j\}$ as the constant mapping.
**4:** Return $\chi : X \to C$ defined by $\chi(x) := \chi_1(x)$, if $x \in X_1$, and $\chi(x) := \chi_2(x)$, if $x \in X_2$, for all $x \in X$.

We prove that our algorithm produces a coloring as claimed in Theorem 3.11 and also fulfills the complexity statements. Suppose by induction that this holds for sets of less than $c$ colors. We analyze the algorithm being started on an input as above with $|C| = c$.

We first show correctness. For Step 1 note that both $C_1$ and $C_2$ are non-empty and that $q_2 \le \frac{2}{3}$ holds. Therefore by Lemma 3.13 and induction the colorings $\chi_i$, $i = 0, 1, 2$ can be computed as desired in Step 2 and 3. Let $E \in \mathcal{E}$, $i \in [2]$ and $j \in C_i$ such that $p_j \ge p_0$. If $|C_i| > 1$, then

$$\begin{aligned}
&\big||E \cap \chi^{-1}(j)| - p_j|E|\big| \\
&= \big||E \cap \chi_0^{-1}(i) \cap \chi_i^{-1}(j)| - p_j|E|\big| \\
&\le \left||E \cap \chi_0^{-1}(i) \cap \chi_i^{-1}(j)| - \tfrac{p_j}{q_i}|E \cap \chi_0^{-1}(i)|\right| + \left|\tfrac{p_j}{q_i}|E \cap \chi_0^{-1}(i)| - p_j|E|\right| \\
&\le \left||(E \cap X_i) \cap \chi_i^{-1}(j)| - \tfrac{p_j}{q_i}|E \cap X_i|\right| + \tfrac{p_j}{q_i}\big||E \cap \chi_0^{-1}(i)| - q_i|E|\big| \\
&\le \tfrac{2}{2^{1-\alpha}-1} D v_\alpha(\tfrac{p_j}{q_i})(q_i n)^\alpha + \tfrac{p_j}{q_i}\tfrac{2}{2^{1-\alpha}-1} D(q_i n)^\alpha \\
&\le \tfrac{2}{2^{1-\alpha}-1} D v_\alpha(p_j) n^\alpha
\end{aligned}$$

by Lemma 3.10 (i). On the other hand, if $C_i$ contains a single color $j$, then $p_j = q_i$ and

$$\begin{aligned}\big||E \cap \chi^{-1}(j)| - p_j|E|\big| &= \big||E \cap \chi_0^{-1}(i)| - q_i|E|\big| \\ &\leq \tfrac{2}{2^{1-\alpha}-1} D(q_i n)^\alpha \\ &\leq \tfrac{2}{2^{1-\alpha}-1} D v_\alpha(p_j) n^\alpha.\end{aligned}$$

This is the correctness statement.

Concerning the complexity note that the computation of $\chi_0$ takes at most $\left\lceil \log_2(\frac{1}{p_0}) \right\rceil$ and (by induction) the one of the $\chi_i$ takes at most $(|C_i| - 1) \left\lceil \log_2(\frac{q_i}{p_0}) \right\rceil$ colorings as in the Decreasing-Discrepancies-Assumption. These are not more than $(c-1) \left\lceil \log_2(\frac{1}{p_0}) \right\rceil$ colorings altogether.

By Lemma 3.13 we compute at most $3n$ times a color for a vertex in Step 2. If $|C_i| > 1$ for both $i = 1, 2$, then $q_i \leq \frac{2}{3}$ and computing $\chi_i$ involves at most $3q_i n \log_{1.5}(\frac{q_i}{p_0}) \leq 3q_i n \log_{1.5}(\frac{2}{3p_0})$ times computing a color for a vertex. Altogether this makes at most $3n + 3q_1 n \log_{1.5}(\frac{q_1}{p_0}) + 3q_2 n \log_{1.5}(\frac{q_2}{p_0}) \leq 3n(1 + \log_{1.5}(\frac{2}{3p_0})) = 3n \log_{1.5}(\frac{1}{p_0})$ times computing a color for a vertex. If $|C_i| = 1$ then there is nothing to do to get $\chi_i$ and the respective term just vanishes in the calculation above. □

### 3.5. **Applications of the Refined Recursive Coloring Approach.** 

We are now ready to prove $c$–color versions of a series of discrepancy results.

3.5.1. *General Hypergraphs.* Let $\mathcal{H} = (X, \mathcal{E})$ denote an arbitrary hypergraph. Set $n := |X|$ and $m := |\mathcal{E}|$ for convenience. The approach of Proposition 2.6 shows that a random coloring generated by coloring each vertex independently with each color with probability $\frac{1}{c}$ has discrepancy at most $\sqrt{\frac{1}{2} n \ln(4mc)}$ with probability at least $\frac{1}{2}$. This yields a randomized algorithm computing such a coloring by repeatedly generating and testing such a random coloring until its discrepancy is at most $\sqrt{\frac{1}{2} n \ln(4mc)}$.

In this subsection we show that via the recursive approach of Theorem 3.11 a better bound can be achieved. In particular, the discrepancy tends to decrease for larger numbers of colors.

**Theorem 3.14.** *Let $p$ denote an $\mathcal{H}$–integral $c$–color weight. Set $p_0 := \min\{p_i | i \in [c]\}$. Then a $c$–coloring $\chi$ having discrepancy at most $45\sqrt{p_i n \ln(4m)}$ in color $i \in [c]$ can be computed in expected time $\mathcal{O}(nm \log(\frac{1}{p_0}))$. In particular, a $c$–coloring $\chi$ such that*

$$\operatorname{disc}(\mathcal{H}, \chi, c) \leq 45\sqrt{\tfrac{n}{c} \ln(4m)} + 1$$

*can be computed in expected time $\mathcal{O}(nm \log c)$.*

*Proof.* There is little to do for $m = 1$, so let us assume that $m \geq 2$. We show that the colorings required by the Decreasing-Discrepancies-Assumption can be computed in expected time $\mathcal{O}(|X_0|m)$. Denote by $\overline{\mathcal{H}}$ the hypergraph obtained from $\mathcal{H}$ by adding the whole vertex set as an additional hyperedge. Let $X_0 \subseteq X$ and $(q, 1-q)$ be a 2–color weight. Let $\chi : X_0 \to [2]$ be a random coloring independently coloring the vertices with probabilities $P(\chi(x) = 1) = q$ and $P(\chi(x) = 2) = 1 - q$ for all $x \in X_0$. A standard application of the Chernoff inequality (cf. [AS92]) shows that

$$(*) \qquad \mathrm{wd}(\overline{\mathcal{H}}_{|X_0}, \chi, (q, 1-q)) \leq \sqrt{\tfrac{1}{2}|X_0| \ln(4m)}$$

holds with probability at least $\frac{m-1}{2m}$. Hence by repeatedly generating and testing these random colorings until $(*)$ holds we obtain a randomized algorithm computing such a coloring with expected running time $\mathcal{O}(nm)$. By Lemma 3.9 we get a fair $(q, 1-q)$–coloring for $\mathcal{H}_{|X_0}$ having discrepancy at most $\sqrt{2|X_0| \ln(4m)}$. Hence for $\alpha = \frac{1}{2}$, $D = \sqrt{2\ln(4m)}$ and arbitrary $p_0$ the colorings required in the Decreasing-Discrepancies-Assumption can be computed in expected time $\mathcal{O}(|X_0|m)$.

Therefore we may apply Theorem 3.11 with $p_0 = \min\{p_i | i \in [c]\}$. The discrepancy bounds follow from $c_\alpha \leq 31.15$. Computing such a coloring involves $\mathcal{O}(\log(\frac{1}{p_0})n)$ times computing a color for a vertex. As this can be done in expected time $\mathcal{O}(m)$, we have the claimed bound of $\mathcal{O}(nm\log(\frac{1}{p_0}))$. $\square$

Some remarks concerning the theorem and its proof above. For the complexity guarantee we assumed that the complexity contribution of computing the 2–colorings dominates the remaining operations of the recursive algorithm given in the proof of Theorem 3.11. This is justified by the fact that we may assume $c \leq n$ since integrality ensures $p_i \geq \frac{1}{|X|}$ for all colors $i \in C$.

A second point is that the constant of 45 could be improved by a more careful way of generating the random 2–colorings. In particular by taking a random fair coloring we could avoid the extra factor of 2 inflicted by Lemma 3.9. This though requires an analysis of the hypergeometric distribution, which is considerably more difficult that ours.

Finally let us remark that the construction of the 2–colorings can be derandomized with standard derandomization techniques like an algorithmic version of the Chernoff-Hoeffding inequality (cf. [SS96] or [Sri01]). Thus the colorings in Theorem 3.14 can be computed by a deterministic algorithm as well.

3.5.2. *Six Standard Deviations.* The celebrated 'six standard deviations' result due to Spencer [Spe85] states that there is a constant $K$ such that for all hypergraphs $\mathcal{H} = (X, \mathcal{E})$ having $n$ vertices and $m \geq n$ edges

$$\mathrm{disc}(\mathcal{H}) \leq K\sqrt{n \ln(\tfrac{2m}{n})}$$

holds.

The interesting case is of course the one where $m = \mathcal{O}(n)$ and thus $\operatorname{disc}(\mathcal{H}) = \mathcal{O}(\sqrt{n})$. For $m$ significantly larger than $n$ this result is outnumbered (due to the implicit constants) by a simple random coloring. The title "Six Standard Deviations Suffice" of this paper comes from the fact that for $n = m$ large enough, $\operatorname{disc}(\mathcal{H}) \le 6\sqrt{n}$ holds.

Using the relation between discrepancies respecting a particular weight and hereditary discrepancy (Remark 3.2) and the recoloring argument (Lemma 3.9), we derive from Spencer's result

**Lemma 3.15.** *For any $X_0 \subseteq X$ and $\mathcal{H}_{|X_0}$–integral weight $p = (q, 1-q)$ there is a fair $p$–coloring $\chi$ of $\mathcal{H}_{|X_0}$ that has $\operatorname{wd}(\mathcal{H}_{|X_0}, \chi, p) \le 2K\sqrt{|X_0| \ln(\frac{2m+2}{|X_0|})}$.*

*Proof.* Let $X_0 \subseteq X$. Then any induced subgraph of $\mathcal{H}_{|X_0}$ has discrepancy at most $K\sqrt{|X_0| \ln(\frac{2m}{|X_0|})}$, simply because Spencer's bound is monotone in the number of vertices. From Remark 3.2, we have $\operatorname{wd}(\mathcal{H}_{|X_0}, 2, (q, 1-q)) \le \operatorname{herdisc}(\mathcal{H}_{|X_0}) \le K\sqrt{|X_0| \ln(\frac{2m}{|X_0|})}$.

It remains to show the existence of a fair coloring. Let $\overline{\mathcal{H}}$ denote the hypergraph arising from $\mathcal{H}$ by adding the set $X$ as an additional edge (unless of course $X \in \mathcal{E}$ already holds). Then $\overline{\mathcal{H}}_{|X_0}$ has at most $m+1$ edges, and from the previous paragraph we know $\operatorname{wd}(\overline{\mathcal{H}}_{|X_0}, 2, (q, 1-q)) \le K\sqrt{|X_0| \ln(\frac{2m+2}{|X_0|})}$. Lemma 3.9 now yields the claim. □

Lemma 3.15 and Theorem 3.11 yield

**Theorem 3.16.** *Let $\mathcal{H} = (X, \mathcal{E})$ denote a hypergraph having $n$ vertices and $m \ge n$ edges and $p \in [0,1]^c$ an integral weight. Set $p_0 := \min_{i \in [c]} p_i$. Then there is a fair $p$–coloring having discrepancy at most $63K\sqrt{p_i n \ln(\frac{2m+2}{p_0 n})}$ in color $i$.*

*In particular, in the case $|X| = |\mathcal{E}| = n$ we have*

$$\operatorname{disc}(\mathcal{H}, c) \le \mathcal{O}\left(\sqrt{\tfrac{n}{c} \ln c}\right).$$

*Proof.* By Lemma 3.15 we may apply Theorem 3.11 with $\alpha = \frac{1}{2}$, $D = 2K\sqrt{\ln(\frac{2m+2}{p_0 n})}$ and $p_0$. This yields a fair $p$–coloring having discrepancy at most $Dc_\alpha\sqrt{p_i n}$ in color $i \in [c]$. The claim follows from $c_\alpha \le 31.15$. □

This is quite close to the optimum. Theorem 5.2 shows a class of hypergraphs such that $|X| = |\mathcal{E}| = n$ and $\operatorname{disc}(\mathcal{H}, c) = \Omega(\sqrt{\frac{n}{c}})$. Again we should remark that we did not try to optimize the constant.

The following corollary on 2 color discrepancies with respect to a given weight seems worth mentioning. Already from combining Lemma 3.13 and Lemma 3.15 we derive:

**Corollary 3.17.** *Let $\mathcal{H} = (X, \mathcal{E})$ denote a hypergraph such that $|X| = |\mathcal{E}| =: n$ and $(q, 1-q)$ an integral 2–color weight. Assume $q \leq \frac{1}{2}$. Then the weighted discrepancy* $\mathrm{wd}(\mathcal{H}, 2, (q, 1-q))$ *is at most* $10K\sqrt{qn\ln(\frac{3}{q})}$.

3.5.3. *Arithmetic Progressions.* A third classical example is the hypergraph of arithmetic progressions on the first $n$ numbers. This is probably the most famous of the few non–trivial examples where discrepancy is well–understood. For $a, d, l \in \mathbb{N}$ denote by $A_{adl} := \{a + id \,|\, 0 \leq i \leq l-1\}$ the arithmetic progression with starting point $a$, difference $d$ and length $l$. Denote by $\mathcal{E}_n$ the set of all arithmetic progressions in $[n]$, that is $\mathcal{E}_n = \{A_{adl} \cap [n] | a, d, l \in [n]\}$. Set $\mathcal{A}_n = ([n], \mathcal{E}_n)$.

Roth [Rot64] proved the celebrated lower bound $\mathrm{disc}(\mathcal{A}_n) = \Omega(n^{1/4})$. Roth himself believed that this bound was too small and that the discrepancy actually should be close to $n^{1/2}$. This was disproved by Sárközy [Sár74], who showed an upper bound of $\mathcal{O}(n^{1/3+\varepsilon})$. Inventing the partial coloring method, Beck [Bec81] showed a nearly tight bound of $\mathcal{O}(n^{1/4}(\log n)^{5/2})$. Finally Matoušek and Spencer [MS96] solved the discrepancy problem for $\mathcal{A}_n$ by proving the asymptotically tight upper bound $\mathcal{O}(n^{1/4})$.

This bound holds in any fixed number of colors. Moreover, we prove that the discrepancy decreases for larger numbers of colors.

**Theorem 3.18.** *For an absolute constant $C'$ the following holds: Let $p \in [0,1]^c$ be a weight. Then there is a fair coloring of $\mathcal{A}_n$ with respect to $p$ having discrepancy at most $C' p_i^{0.16} n^{0.25}$ in each color $i$ such that $p_i \geq n^{0.25}$. In particular,*

$$\mathrm{disc}(\mathcal{A}_n, c) = \mathcal{O}(c^{-0.16} n^{0.25})$$

*holds for $c \leq n^{0.25}$ colors.*

*Proof.* From Lemma 5.3 of [MS96] we learn than an induced subgraph $\mathcal{H}_0 = (\mathcal{A}_n)_{|X_0}$ of $\mathcal{A}_n$ on $|X_0| = \rho n \geq n^{0.25}$ vertices has discrepancy at most $C_1 \rho^{0.16} n^{0.25}$ for some absolute constant $C_1$. We first show that $\mathrm{herdisc}(\mathcal{H}_0) \leq 2C_1 \rho^{0.16} n^{0.25}$:

Let $\mathcal{H}_1 = (X_1, \mathcal{E}_1)$ be an induced subhypergraph of $\mathcal{H}_0$. If $|X_1| \geq n^{0.25}$ we are done by the Lemma of Matoušek and Spencer. Let us therefore assume $|X_1| < n^{0.25}$. We show that $(\mathcal{H}_1)_{|[\frac{n}{2}]}$ and $(\mathcal{H}_1)_{|[n]\setminus[\frac{n}{2}]}$ have discrepancy at most $C_1 \rho^{0.16} n^{0.25}$ and conclude $\mathrm{disc}(\mathcal{H}_1) \leq 2C_1 \rho^{0.16} n^{0.25}$. Consider the hypergraph $\mathcal{H}_2 := \mathcal{H}_{(X_1 \cap [\frac{n}{2}]) \cup \{n - n^{0.25} + |X_1 \cap [\frac{n}{2}]|+1, \ldots, n\}}$. This hypergraph has exactly $n^{0.25} \leq \rho n$ vertices and thus discrepancy at most $C_1 \rho^{0.16} n^{0.25}$. As every edge of $(\mathcal{H}_1)_{|[\frac{n}{2}]}$ is also an edge of $\mathcal{H}_2$, we conclude $\mathrm{disc}((\mathcal{H}_1)_{|[\frac{n}{2}]}) \leq C_1 \rho^{0.16} n^{0.25}$. A similar argument shows $\mathrm{disc}((\mathcal{H}_1)_{|[n]\setminus[\frac{n}{2}]}) \leq C_1 \rho^{0.16} n^{0.25}$.

Thus $\operatorname{herdisc}(\mathcal{H}_0) \leq 2C_1\rho^{0.16}n^{0.25}$. The relation between the linear and hereditary discrepancy yields that all weighted discrepancies of $\mathcal{H}_0$ are bounded by $2C_1\rho^{0.16}n^{0.25}$. As $[n]$ is an arithmetic progression, we may apply Lemma 3.9 and conclude that twice this discrepancy may be achieved by a fair coloring respecting the underlying weight.

Thus we may apply Theorem 3.11 with $D = 4C_1n^{0.09}$, $\alpha = 0.16$ and $p_0 = n^{0.25}$, which proves our claim. □

Theorem 3.18 has a nice corollary in two colors extending Matoušek's and Spencer's bound to a more general counterpart of Roth's theorem. Though often only the (ordinary) discrepancy result is cited, Roth's famous paper actually shows a lower bound for the weighted discrepancy in two colors: For all $p \in [0,1]$,

$$\operatorname{disc}(\mathcal{A}_n, (p, 1-p), 2) = \Omega(p^{1/2}n^{1/4}).$$

Theorem 3.18 applied to $c = 2$ shows an upper bound of

$$\operatorname{disc}(\mathcal{A}_n, (p, 1-p), 2) = \mathcal{O}(p^{0.16}n^{1/4}).$$

3.5.4. *Bounded Shatter Functions.* The recursive approach also generalizes results of Matoušek, Welzl and Wernisch [MWW84] and Matoušek [Mat95] connecting discrepancy with the primal shatter function $\pi_{\mathcal{H}}$ and dual shatter function $\pi^*_{\mathcal{H}}$ of a hypergraph. Note that this also yields a discrepancy bound in terms of the $VC$–dimension $\dim(\mathcal{H})$ of $\mathcal{H}$: Already Vapnik and Chervonenkis [VC71] showed $\pi_{\mathcal{H}} \in \mathcal{O}(n^{\dim(\mathcal{H})})$.

**Theorem 3.19.** *Let $\mathcal{H} = (X, \mathcal{E})$ be a hypergraph on $n$ points. Let $d > 1$. If $\pi_{\mathcal{H}} = \mathcal{O}(m^d)$, then $\operatorname{disc}(\mathcal{H}, c) = \mathcal{O}((\frac{n}{c})^{\frac{1}{2}-\frac{1}{2d}})$. If $\pi^*_{\mathcal{H}} = \mathcal{O}(m^d)$, then $\operatorname{disc}(\mathcal{H}, c) = \mathcal{O}((\frac{n}{c})^{\frac{1}{2}-\frac{1}{2d}} \log n)$. In both cases the implicit constants are independent of $c$.*

*Proof.* Clearly the assumptions on the shatter functions are hereditary in the sense that a shatter function of an induced subhypergraph is less or equal the one of the whole hypergraph. They are also very robust: Adding the whole vertex set as additional edge changes the primal shatter function by at most 1, and does not change the dual shatter function. Without loss of generality we may therefore assume $X \in \mathcal{E}$. The remainder of the proof is standard — bound the weighted discrepancies of the induced subhypergraphs using Remark 3.2, buy fairness at the price of a factor of 2 (Lemma 3.9) and apply Theorem 3.11. □

3.6. **Summary: Recursive Coloring.** We have seen that a recursive approach is very effective in situations where we can bound the weighted discrepancies of induced subhypergraphs. We always get a uniform bound from the hereditary discrepancy of $\mathcal{H}$ (Remark 3.2) and often find a bound decreasing for smaller induced subhypergraphs.

There many are situations where the recursive approach is the only result we have. We do not have a direct proof for a result like Theorem 3.16 or Theorem 3.18. We feel that the original proof relies heavily on the fact that only two colors are considered.

Surprisingly, the recursive approach and direct methods are sometimes nearly equally effective. An example is the (equi-weighted) multi–color discrepancy in the case of bounded degree (as in the theorem of Beck and Fiala). The direct approach of Section 4 yields $\operatorname{disc}(\mathcal{H}, c) \le 2\Delta(\mathcal{H})$, the recursive one gives $\operatorname{disc}(\mathcal{H}, c) \le v(c)\Delta(\mathcal{H})$ for constants $v(c) \in [2(1-\frac{1}{c}), 2.0005]$. Both ways are constructive. For $c$ tending to infinity both methods give the same bound.

On the other hand the recursive approach is limited: We can get results on weighted discrepancies, but we do not get a nice bound on the linear discrepancy, e. g. in the Beck–Fiala setting. A second point to keep in mind is that to apply recursion, we need a two–color result on the hereditary discrepancy, even in the case that $c$ is a power of 2. See the example in Section 2.

Recently, the first author showed a converse result [Doe02a]: If the hereditary discrepancy in $c$ colors is bounded, one may construct $c_2$ colorings (and in particular 2–colorings) with low discrepancy. Together with Remark 3.2 and Corollary 3.5 this shows that the hereditary discrepancy in two numbers of colors deviates at most by a constant factor (depending on the numbers of colors, but not on the hypergraph).

## 4. Vector-Coloring

In this section we extend the Beck–Fiala theorem and the Barany–Grunberg theorem to any number of colors. In the 2–color case both are proved using 'floating colors', i. e. colors initially floating in $[-1, 1]$ are successively changed to colors in $\{-1, 1\}$. Linear algebra is the key tool there. For the $c$–color case we need vector colors and tensor products as well. In the Beck–Fiala situation we will derive a bound independent on the number of colors (and twice the bound of the original result), whereas in the Barany–Grunberg case our bound is $(c-1)$ times the original bound (and thus coincides with the original result in the case $c = 2$).

4.1. **Beck–Fiala Theorem.** Denote by $\Delta(\mathcal{H}) := \max_{x \in X} |\{E \in \mathcal{E} | x \in E\}|$ the maximum degree of the hypergraph $\mathcal{H}$. This is one of the few parameters of a hypergraph which give a good bound on the discrepancy. The Beck–Fiala theorem states that $\operatorname{disc}(\mathcal{H}) < 2\Delta(\mathcal{H})$ for any hypergraph $\mathcal{H}$ (cf. [BF81]).

It is quite easy to see from the proof that this bound can be improved to $2\Delta(\mathcal{H}) - 2$. With more effort Martin Helm [Hel99] further improved the bound to $2\Delta(\mathcal{H}) - 3$. The much stronger conjecture of Beck and Fiala is:

**Conjecture.** $\operatorname{disc}(\mathcal{H}) \in \mathcal{O}(\sqrt{\Delta(\mathcal{H})})$.

This conjecture remains far from being proven. Good results in this direction are Srinivasan [Sri97] and Banaszczyk [Ban98], though neither succeeds in avoiding a logarithmic dependence of the number of vertices.

Beck and Fiala actually proved a more general result. For any matrix $A = (a_{ij}) \in \mathbb{R}^{m\times n}$ denote by $\|A\|_1 := \max_{j\in[n]} \sum_{i\in[m]} |a_{ij}|$ the operator norm induced by the 1–norm.

**Theorem 4.1** ([BF81])**.** *For any matrix $A \in \mathbb{R}^{m\times n}$,*

$$\operatorname{lindisc}(A) < 2\|A\|_1.$$

For $c$ colors we prove

**Theorem 4.2.** *For any matrix $A \in \mathbb{R}^{m\times n}$,*

$$\operatorname{lindisc}(A, c) < 2\|A\|_1.$$

Note that the linear discrepancy versions of the Beck–Fiala theorem do not allow the improvements cited above. They rely on the fact that the incidence matrix of a hypergraph is a 0–1–matrix. A multi-color version of the Beck–Fiala theorem for hypergraph exploiting this fact was recently given by Biedl et al. [BČC+02].

The following very elementary remark plays a crucial role in the proofs of both the generalized Beck–Fiala theorem and the Barany–Grunberg theorem.

**Lemma 4.3.** *Let $x \in \overline{M}_c$. Assume that there is a $j' \in [c]$ such that $x_{j'} \notin \{-\frac{1}{c}, \frac{c-1}{c}\}$. Then there is at least a second index $j''$ (different from $j'$) such that $x_{j''} \notin \{-\frac{1}{c}, \frac{c-1}{c}\}$.*

*Proof.* By assumption we have $cx_{j'} \notin \mathbb{Z}$. As $c\sum_{j\in[c]} x_j = 0 \in \mathbb{Z}$ by definition of $\overline{M}_c$, there exists a $j'' \in [c]$, $j' \neq j''$, such that $cx_{j''} \notin \mathbb{Z}$. In particular, $x_{j''} \notin \{-\frac{1}{c}, \frac{c-1}{c}\}$. □

*Proof of Theorem 4.2.* Set $\Delta := \|A\|_1$ and $\overline{A} = (\overline{a}_{ij}) := A \otimes I_c$. Note that $\Delta = \|\overline{A}\|_1$. Let $p : [n] \to \overline{M}_c$. Set $\chi = p$. Successively we will change $\chi$ to a mapping $[n] \to M_c$. Again we regard $p$ and $\chi$ as $cn$–dimensional vectors.

Put $J := \{j \in [cn] | \chi_j \notin \{-\frac{1}{c}, \frac{c-1}{c}\}\}$, and call the columns from $J$ floating (the others fixed). Set $I := \{i \in [cm] | \sum_{j\in J} |\overline{a}_{ij}| > 2\Delta\}$, and call the rows from $I$ active (the others ignored). We will ensure that during the rounding process the following conditions are fulfilled (this is clear for the start, because $\chi = p$):

(i) $(\overline{A}(p - \chi))_{|I} = 0$, i. e. all active rows have discrepancy zero, and
(ii) all colors are in $\overline{M}_c$, in particular we have $\sum_{k=0}^{c-1} \chi_{cj-k} = 0$ for all $j \in [n]$.

Note that (ii) is the crucial difference to the 2–color case, where we only need a condition of type (i). This increases the number of equations investigated below and is the reason why the multi-color bound is twice the 2–color bound.

Let us assume that the rounding process is at a stage where $J$ and $I$ are as above and (i) and (ii) hold. If there is no floating color, i. e. $J = \emptyset$, then all $\chi_j$, $j \in [cn]$, are in $\{-\frac{1}{c}, \frac{c-1}{c}\}$ and $\chi$ has the desired form.

Hence assume that there are still floating colors. We consider the system of equations

$$\sum_{k=0}^{c-1} \chi_{cj-k} = 0,\ j \in [n] \text{ such that } c(j-1)+k \in J \text{ for some } k \in [c]. \tag{22}$$

By Lemma 4.3, in every equation of (22) there are at least two floating variables $\chi_{j'}, \chi_{j''}$, i. e. $j', j'' \in J$. Thus (22) is a system of at most $\frac{1}{2}|J|$ equations.

We have

$$|J|\,\Delta \geq \sum_{j \in J} \sum_{i \in I} |\overline{a}_{ij}| = \sum_{i \in I} \sum_{j \in J} |\overline{a}_{ij}| > |I|\,2\Delta,$$

hence $|J| > 2|I|$. We conclude that the system

$$\begin{aligned} \overline{A}_{|I \times J}\ \ \chi_{|J} &= 0 \\ \sum_{k=0}^{c-1} \chi_{cj-k} &= 0,\ j \in [n] \text{ such that } c(j-1)+k \in J \text{ for some } k \in [c] \end{aligned} \tag{23}$$

consists of at most $|I| + \frac{1}{2}|J| < |J|$ equations and hence is under-determined (taking just the $\chi_j, j \in J$ as variables). Thus there is a non-trivial solution $x \in \mathbb{R}^{|J|}$ for (23). We extend $x$ to $x_E \in \mathbb{R}^{cn}$ by

$$(x_E)_j := \begin{cases} x_j & \text{if } j \in J \\ 0 & \text{else} \end{cases}.$$

By (ii) and the definition of $J$, all variables $\chi_j, j \in J$ are in $]-\frac{1}{c}, \frac{c-1}{c}[$. Thus there is a $\lambda > 0$ such that at least one component of $\chi + \lambda x_E$ becomes fixed and all colors are still in $\overline{M}_c$, i. e. $\chi + \lambda x_E \in \overline{M}_c^n$. Note that $\chi + \lambda x_E$ also fulfills (i) since $(\overline{A} x_E)_{|I} = 0$. Set $\chi := \chi + \lambda x_E$. Since (i), (ii) are fulfilled for this new $\chi$, we can continue this rounding process until all $\chi_j$, $j \in [cn]$ are in $\{-\frac{1}{c}, \frac{c-1}{c}\}$.

We show $\|\overline{A}(p-\chi)\|_\infty < 2\Delta$. Let $i \in [cm]$. Denote by $\chi^{(0)}$ and $J^{(0)}$ the values of $\chi$ and $J$ when the row $i$ first became ignored. We have $\chi_j^{(0)} = \chi_j$ for all $j \notin J^{(0)}$ and $|\chi_j^{(0)} - \chi_j| < 1$

for all $j \in J^{(0)}$. Note that $\sum_{j \in J^{(0)}} |\overline{a}_{ij}| \leq 2\Delta$, since $i$ is ignored. Thus

$$|(\overline{A}(p-\chi))_i| = |(\overline{A}(p-\chi^{(0)}))_i + (\overline{A}(\chi^{(0)}-\chi))_i| = |0 + \sum_{j \in J^{(0)}} \overline{a}_{ij}(\chi_j^{(0)} - \chi_j)| < 2\Delta.$$

□

For the $c$–color discrepancy we have

**Corollary 4.4.** $\operatorname{disc}(\mathcal{H}, c) < 2\Delta(\mathcal{H})$.

Note that this result is very similar to Theorem 3.8 of Section 3.

4.2. **Theorem of Barany–Grunberg.** For the remainder of this section let $\|\cdot\|$ denote any norm on $\mathbb{R}^d$. The theorem of Barany–Grunberg states

**Theorem 4.5** ([BG81])**.** *Let $\|\cdot\|$ be any norm on $\mathbb{R}^n$ and $v_1, v_2, \ldots, v_k$ be a finite sequence of vectors in $\mathbb{R}^n$ of arbitrary length such that $\|v_i\| \leq 1$ for all $i = 1, \ldots, k$. Then there are signs $\varepsilon_i \in \{-1, +1\}, i = 1, \ldots, k$ such that for all $l \in [k]$ we have*

$$\left\| \sum_{i=1}^{l} \varepsilon_i v_i \right\| < 2n.$$

This seems to be similar to the Beck–Fiala theorem, but has a slightly different flavor: Here partial sums are considered, and we may choose any norm for the input and the discrepancy. The Beck–Fiala theorem formulated in terms of a vector sequence states that for any vectors $v_1, \ldots, v_k$ of $\|\cdot\|_1$–norm at most one there are signs $\varepsilon_i \in \{-1, +1\}, i = 1, \ldots, k$ such that $\left\|\sum_{i=1}^{k} \varepsilon_i v_i\right\|_\infty < 2$. Thus neither theorem is a special case of the other.

The signs $-1$ and $+1$ are a convenient way to represent a partition. From this point of view the theorem of Barany–Grunberg states that under the given assumptions there is a 2–partition $(I_1, I_2)$ of the set $X = \{v_1, \ldots, v_k\}$ such that for any subset $X_0 = \{v_1, \ldots, v_l\}$

$$\left\| \sum_{v \in I_j \cap X_0} v - \tfrac{1}{2} \sum_{v \in X_0} v \right\| < n$$

holds for both $j = 1, 2$. This motivates the following definition:

**Definition** (Discrepancy of sets and vector sequences)**.** Let $X$ be a finite set of vectors in $\mathbb{R}^n$ and $\mathcal{P} = \{I_1, \ldots, I_c\}$ a $c$–partition of $X$. Let $\|\cdot\|$ be any norm on $\mathbb{R}^n$. We define the *discrepancy of the set* $X$ w. r. t. $\mathcal{P}$ and $\|\cdot\|$ by

$$\operatorname{disc}(\mathcal{P}, \|\cdot\|) := \max_{j \in [c]} \left\| \sum_{v \in I_j} v - \tfrac{1}{c} \sum_{v \in X} v \right\|.$$

Given a subset $X_0 \subseteq X$ set $\mathcal{P}_{|X_0} := \{I_1 \cap X_0, \dots, I_c \cap X_0\}$. Let $v_1, v_2, \dots, v_k$ be a finite sequence of vectors and $\mathcal{P} = \{I_1, \dots, I_c\}$ be a $c$–partition of $\{v_1, v_2, \dots, v_k\}$. We define the *discrepancy of the sequence* $v_1, v_2, \dots, v_k$ w. r. t. $\mathcal{P}$ and $\|\cdot\|$ by

$$\operatorname{disc}((v_l)_{l\in[k]}, \mathcal{P}, \|\cdot\|) := \max_{l\in[k]} \operatorname{disc}(\mathcal{P}_{|\{v_1,\dots,v_l\}}, \|\cdot\|).$$

In this notation the Barany–Grunberg theorem states that there is a 2–partition $\mathcal{P} = \{I_1, I_2\}$ such that $\operatorname{disc}((v_l)_{l\in[k]}, \mathcal{P}, \|\cdot\|) < n$. We define a norm $\|\cdot\|_c$ on $\mathbb{R}^{cn}$ by

$$\|w\|_c := \max_{j\in[c]} \|w_{|\{j,j+c,\dots,j+(n-1)c\}}\|,$$

where we write $w_{|\{j,j+c,\dots,j+(n-1)c\}}$ for the $n$–dimensional vector $(w_j, w_{j+c}, \dots, w_{(n-1)c})^\top$. Then

**Lemma 4.6.** *Let $X \subseteq \mathbb{R}^n$ be a finite set of vectors and $\mathcal{P} = \{I_1, \dots, I_c\}$ be any $c$–partition of $X$. Let $\chi : X \to [c]$ be the corresponding coloring (i. e. for all $v \in X, l \in [c]$ we have $\chi(v) = l$ if and only if $v \in I_l$). Then the discrepancy of $X$ w. r. t. $\mathcal{P}$ and $\|\cdot\|$ is*

$$\operatorname{disc}(\mathcal{P}, \|\cdot\|) = \left\| \sum_{v\in X} v \otimes m^{(\chi(v))} \right\|_c .$$

*Proof.* Remember that by definition $m_j^{(\chi(v))} = \begin{cases} 1 - \frac{1}{c} & \text{if } \chi(v) = j \\ -\frac{1}{c} & \text{otherwise.} \end{cases}$. Thus

$$(v \otimes m^{(\chi(v))})|_{\{j,j+c,\dots,j+(n-1)c\}} = m_j^{(\chi(v))} v \tag{24}$$

and

$$\sum_{v\in X} m_j^{\chi(v)} v = \sum_{\substack{v\in X\\ \chi(v)=j}} (1 - \tfrac{1}{c}) v - \sum_{\substack{v\in X\\ \chi(v)\neq j}} \tfrac{1}{c} v = \sum_{\substack{v\in X\\ \chi(v)=j}} v - \tfrac{1}{c} \sum_{v\in X} v. \tag{25}$$

So

$$\begin{aligned}
\left\| \sum_{v\in X} v \otimes m^{(\chi(v))} \right\|_c &= \max_{j\in[c]} \left\| \left( \sum_{v\in X} v \otimes m^{(\chi(v))} \right)\Bigg|_{\{j,j+c,\dots,j+(n-1)c\}} \right\| \\
&\overset{(24)}{=} \max_{j\in[c]} \left\| \sum_{v\in X} m_j^{(\chi(v))} v \right\| \\
&\overset{(25)}{=} \max_{j\in[c]} \left\| \sum_{v\in I_j} v - \tfrac{1}{c} \sum_{v\in X} v \right\| \\
&= \operatorname{disc}(\mathcal{P}, \|\cdot\|).
\end{aligned}$$

□

We are now ready to prove the following multi-color version:

**Theorem 4.7.** *Let* $\|\cdot\|$ *be any norm on* $\mathbb{R}^n$ *and* $v_1, v_2, \dots, v_k$ *be a finite sequence of vectors in* $\mathbb{R}^n$ *such that* $\|v_i\| \le 1$ *for all* $i = 1, \dots, k$*. Then there is a c–partition* $\mathcal{P} = \{I_1, \dots, I_c\}$ *of* $\{v_1, v_2, \dots, v_k\}$ *such that*

$$\mathrm{disc}((v_l)_{l\in[k]}, \mathcal{P}, \|\cdot\|) < (c-1)n.$$

*Proof.* We may assume $k > n$. By Lemma 4.6 it suffices to show the existence of a coloring $\chi : [k] \to M_c$ such that $\left\|\sum_{i\in[l]} v_i \otimes \chi^{(i)}\right\|_c < (c-1)n$ for all $l \in [k]$.

As in the proof of the Barany-Grunberg theorem we give an algorithmic construction of $\chi$. At the beginning define $A := [n]$ and $\chi_j^{(i)} := 0$ for all $i \in [k]$, $j \in [c]$. Let us call those $\chi_j^{(i)}$ where $i \in A$ and $\chi_j^{(i)} \notin \{\frac{c-1}{c}, -\frac{1}{c}\}$ *variables* and the corresponding color vector $\chi^{(i)}$ *active.* Hence at the beginning we have $cn$ variables and $n$ active color vectors. Furthermore all color vectors $\chi^{(i)}, i \in [k]$ are in $\overline{M_c}$ and we have $\sum_{i\in A} v_i \otimes \chi^{(i)} = 0$.

We repeat the following rounding process: Set $A_0 := \{i \in [k] | \exists j \in [c] : \chi_j^{(i)} \notin \{\frac{c-1}{c}, -\frac{1}{c}\}\}$, the set of indices of active color vectors. We try to find a nontrivial solution to the system of equations

$$\begin{aligned} \sum_{i\in A} v_i \otimes \chi^{(i)} &= 0 \\ \sum_{j\in[c]} \chi_j^{(i)} &= 0 \quad \text{for all } i \in A_0. \end{aligned} \tag{26}$$

Let $n'$ be the number of variables and $m'$ the rank of the system (26). By Lemma 4.3, each active vector contains at least two variables, so $n' \ge 2|A_0|$. On the other hand, $m' \le (c-1)n + |A_0|$, since $\sum_{j\in[c]} \chi_j^{(i)} = 0$ for all $i \in [k]$ holds at any stage of the rounding process.

If there is no nontrivial solution to (26), then there are at most $m'$ variables. From $2|A_0| \le n' \le m' \le (c-1)n + |A_0|$ we conclude $|A_0| \le (c-1)n$.

If there are still vectors that have not been active, i. e. $A \ne [k]$, we increase the number of active vectors by setting $A := A \cup \{\max(A) + 1\}$ and continue the rounding process considering the updated system (26). If $A = [k]$ we terminate the rounding process by changing the remaining variables to $\frac{c-1}{c}$ or $-\frac{1}{c}$ in any way such that all $\chi^{(i)}$ are in $M_c$.

If there is a nontrivial solution to (26), then we can change $\chi$ in the way that some variables become $\frac{c-1}{c}$ or $-\frac{1}{c}$ and all variables stay in $[-\frac{1}{c}, \frac{c-1}{c}]$ in the same fashion as in the proof of

Beck–Fiala. Note that the conditions $\chi^{(i)} \in \overline{M_c}$ for all $i \in [k]$ and $\sum_{i\in A} v_i \otimes \chi^{(i)} = 0$ are still satisfied. Hence we can continue the rounding process.

For the analysis let $l \in [k]$. Denote by $\widetilde{\chi}^{(1)}, \dots, \widetilde{\chi}^{(k)}$ the value of the color vectors at that stage of the rounding process when $A = [l]$ and no nontrivial solution to (26) can be found. Denote by $\widetilde{A_0}$ the value of $A_0$ at this stage. Let $\chi_f^{(1)}, \dots, \chi_f^{(k)}$ denote the final values of the color vectors. From above we know $|\widetilde{A_0}| \le (c-1)n$. Since $\chi^{(i)} \in \overline{M_c}$ we have $\|\widetilde{\chi}^{(i)} - \chi_f^{(i)}\|_\infty < 1$ for all $i \in [l]$. Furthermore $\widetilde{\chi}^{(i)} = \chi_f^{(i)}$ holds if $i \notin \widetilde{A}_0$, since an inactive vector never becomes active again. By (26) we also have the equation $\sum_{i\in[l]} v_i \otimes \widetilde{\chi}^{(i)} = 0$.

Now

$$\begin{aligned}
\left\| \sum_{i\in[l]} v_i \otimes \chi^{(i)} \right\|_c &\le \underbrace{\left\| \sum_{i\in[l]} v_i \otimes \widetilde{\chi}^{(i)} \right\|_c}_{=0 \text{ by } (26)} + \left\| \sum_{i\in[l]} v_i \otimes (\chi^{(i)} - \widetilde{\chi}^{(i)}) \right\|_c \\
&= \left\| \sum_{i\in\widetilde{A}_0} v_i \otimes (\chi^{(i)} - \widetilde{\chi}^{(i)}) \right\|_c \\
&= \max_{j\in[c]} \left\| \sum_{i\in\widetilde{A}_0} (v_i \otimes (\chi^{(i)} - \widetilde{\chi}^{(i)}))|_{\{j,j+c,\dots,j+(n-1)c\}} \right\| \\
&= \max_{j\in[c]} \left\| \sum_{i\in\widetilde{A}_0} v_i (\chi^{(i)} - \widetilde{\chi}^{(i)})_j \right\| \\
&< \sum_{i\in\widetilde{A}_0} \|v_i\| \\
&\le (c-1)n.
\end{aligned}$$

□

## 5. Lower Bounds

In this section we give a general lower bound and analyze two prominent examples: Hypergraphs arising from Hadamard matrices and arithmetic progressions. We start with the $c$–color version of a result attributed to Lovász and Sós in [BS95].

**Theorem 5.1.** *Let* $A \in \mathbb{R}^{m\times n}$*. Then* $\operatorname{disc}(A, c) \ge \sqrt{\frac{n(c-1)}{mc^2} \lambda_{\min}(A^\top A)}$.

*Proof.* Let $\chi : [n] \to M_c$ be an optimal coloring with respect to $c$–color discrepancy. Then

$$\begin{aligned}
\operatorname{disc}(A,c) &= \|(A\otimes I_c)\chi\|_\infty \\
&\geq \frac{1}{\sqrt{cm}}\|(A\otimes I_c)\chi\|_2 \\
&\geq \frac{1}{\sqrt{cm}}\|\chi\|_2\sqrt{\lambda_{\min}((A\otimes I_c)^\top(A\otimes I_c))} \\
&\overset{\text{Lemma 2.5(iii)}}{=} \frac{1}{\sqrt{cm}}\sqrt{\frac{n(c-1)}{c}}\sqrt{\lambda_{\min}((A^\top A)\otimes I_c)} \\
&\overset{\text{Lemma 2.5(v)}}{=} \sqrt{\frac{n(c-1)}{mc^2}}\sqrt{\lambda_{\min}(A^\top A)}.
\end{aligned}$$

□

5.1. **Hadamard Matrices.** Hypergraphs corresponding to Hadamard matrices show that Spencer's 'six standard deviations' result is best possible apart from constant factors. The following theorem extends this result to $c$ colors:

**Theorem 5.2.** *There is a universal constant $K > 0$ such that for an infinite sequence of $n \in \mathbb{N}$ there is a hypergraph with $n$ vertices and $n$ edges and discrepancy at least $K\sqrt{\frac{n}{c}}$.*

*Proof.* Let $n \in \mathbb{N}$ be such that there exists a Hadamard matrix $H$ of dimension $n$, i. e. $H \in \{+1,-1\}^{n\times n}$ and all rows of $H$ are pairwise orthogonal. By multiplying some rows by $-1$ we may assume that all entries of the first column $v_1$ are 1. Let $v_2,\dots,v_n$ denote the remaining columns. Set $A = \frac{1}{2}(H+J)$, where $J$ is the $n\times n$ matrix consisting of 1s only. $A$ is the incidence matrix of a hypergraph $\mathcal{H}$ of $n$ edges on $n$ vertices. We show that $\mathcal{H}$ has the desired discrepancy.

Let $\chi : [n] \to M_c$ be any coloring. Let $i \in [c]$ be such that

$$|\chi^{-1}(m^{(i)}) \setminus \{1\}| \geq \tfrac{n-1}{c}. \tag{27}$$

For all $j \in [c]$ set $\chi_j : [n] \to \{-\frac{1}{c}, \frac{c-1}{c}\}; k \mapsto \chi(k)_j$. Let $a_1,\dots,a_n$ be the row vectors of $A$ and for $x, y \in \mathbb{R}^n$ let $x\cdot y$ be the usual inner product in $\mathbb{R}^n$. Then

$$\begin{aligned}
\operatorname{disc}_\chi(\mathcal{H},c) &= \|(A\otimes I_c)\chi\|_\infty \\
&= \|(a_1\cdot\chi_1,\dots,a_1\cdot\chi_c,\dots,a_n\cdot\chi_1,\dots,a_n\cdot\chi_c)^\top\|_\infty \\
&\geq \|(a_1\cdot\chi_i,\dots,a_n\cdot\chi_i)^\top\|_\infty \\
&= \|A\chi_i\|_\infty \\
&\geq \tfrac{1}{\sqrt{n}}\|A\chi_i\|_2.
\end{aligned}$$

By (27) we have

$$|\{k \in [n]\setminus\{1\} | \chi_i(k)| = \tfrac{c-1}{c}\} \geq \tfrac{n-1}{c}. \tag{28}$$

By definition of $A$ there is a $\lambda \in \mathbb{R}$ such that $A\chi_i = \sum_{k=2}^n \frac{1}{2}\chi_i(k)v_k + \lambda v_1$. Since the $v_1, \dots, v_n$ are pairwise orthogonal, we have

$$\begin{aligned}
\|A\chi_i\|_2 &= \sqrt{\sum_{k=2}^{n} \chi_i(k)^2 \|\tfrac{1}{2}v_k\|_2^2 + \lambda^2\|v_1\|_2^2} \\
&\geq \sqrt{\sum_{k=2}^{n} \chi_i(k)^2 \|\tfrac{1}{2}v_k\|_2^2} \\
&= \tfrac{1}{2}\sqrt{n}\sqrt{\sum_{k=2}^{n} \chi_i(k)^2} \\
&\geq \tfrac{1}{2}\sqrt{n}\sqrt{\tfrac{n-1}{c}\left(\tfrac{c-1}{c}\right)^2 + \tfrac{(n-1)(c-1)}{c}\left(-\tfrac{1}{c}\right)^2} \quad \text{(by (28))} \\
&= \tfrac{1}{2}\sqrt{n}\sqrt{\tfrac{(n-1)(c-1)}{c^2}}.
\end{aligned}$$

Hence $\operatorname{disc}(\mathcal{H}, c) \geq \frac{1}{2}\sqrt{\frac{(n-1)(c-1)}{c^2}}$. □

5.2. **Arithmetic Progressions.** In this section we give a lower bound for the $c$–color discrepancy of the arithmetic progressions. We refer to Subsection 3.5.3 for an introduction to this problem.

**Theorem 5.3.** *The hypergraph of arithmetic progressions fulfills*

$$\operatorname{disc}(\mathcal{H}_n, c) \geq 0.04\frac{1}{\sqrt{c}}\sqrt[4]{n}.$$

*Proof.* For the lower bound we will follow the approach of [BS95]. Set $k = \lfloor\sqrt{\frac{1}{6}n}\rfloor$. Let $\mathcal{E}$ be the set of arithmetic progressions of length $k$ and difference less than $6k$ computed modulo $n$ (hence our arithmetic progressions may be over-wrapped from $n$ to 1 at most once). Every arithmetic progression of $\mathcal{E}$ is a union of at most two arithmetic progressions from $\mathcal{E}_n$, so the discrepancy of $\mathcal{H}_n$ is at least half the discrepancy of $\mathcal{E}$.

Recall that a matrix is called circulant if the $i$-th row can be obtained from the first by shifting it $i-1$ times to the right. Let us enumerate the arithmetic progressions in $\mathcal{E}$ in a way that if $i$ is not divisible by $n$, then $E_{i+1} = E_i + 1$ (always computed modulo $n$), i. e. $E_{i+1}$ is $E_i$ shifted right by one. Thus the incidence matrix $A = (a_{ij}) \in \{0,1\}^{6kn\times n}$ defined by $a_{ij} = 1$ if and only if $j \in E_i$ consists of $6k$ circulant sub-matrices. As sum and product of two circulant matrices is circulant again, $A^\top A$ is circulant. The eigenvectors of circulant matrices are known to be of the form $(1, \varepsilon, \varepsilon^2, \dots, \varepsilon^{n-1})^\top$, where $\varepsilon$ is an $n$th root of unity. Thus we find that the minimum eigenvalue $\lambda_{\min}(A^\top A)$ of $A^\top A$ is greater than $\frac{1}{4}k^2$.

Using Theorem 5.1 we have $\mathrm{disc}(([n],\mathcal{E}_n),c)^2 \geq \frac{n(c-1)}{6knc^2}\frac{1}{4}k^2 = \frac{(c-1)k}{24c^2}$. Hence

$$\begin{aligned}\mathrm{disc}(\mathcal{H}_n,c) &\geq 0.5\,\mathrm{disc}(([n],\mathcal{E}),c)\\ &\geq \sqrt{\frac{c-1}{96\sqrt{6}c^2}}\sqrt[4]{n}\\ &\geq 0.0652\sqrt{\frac{c-1}{c^2}}\sqrt[4]{n}\\ &\geq 0.04\sqrt{\frac{1}{c}}\sqrt[4]{n}.\end{aligned}$$

□

We may remark that the lower bound can also be proved using harmonic analysis as in [Weh97, DSW98].

## 6. Open Problems

From this paper several open problems arise, some of which we would like to emphasize here. For the hypergraph of arithmetic progressions and the 'six standard deviations' situation we gave upper and lower bounds for the $c$–color discrepancy. For fixed number $c$ of colors, these bounds are optimal apart from constant factors. Concerning the influence of the number of colors, we showed that these discrepancies decrease with larger numbers of colors. Still, our bounds display a multiplicative gap of $\mathcal{O}(c^{0.34})$ and $\mathcal{O}(\sqrt{\log c})$ respectively. Reducing these gaps seems to be a nice problem. For the arithmetic progressions this should as well yield more information about the weighted discrepancy in two colors, reducing the gap between Roth's bound and ours.

Another interesting question is whether there is a direct proof for Spencer's 'six standard deviations' result for the multi-color discrepancy problem or, more generally, a suitable generalization of the partial coloring method. A negative answer would suggest that discrepancy in 2 colors is a rather special situation compared to discrepancy in arbitrary numbers of colors.

A problem field investigated only a little in this paper is the one of linear discrepancies. The following observation suggests that they might behave differently in more than 2 colors. Consider a totally unimodular $m \times n$ matrix $A$. Various proofs show that $\mathrm{lindisc}(A,2) \leq 1$ holds. The sharp bound of $1-\frac{1}{n+1}$ was recently proven in [Doe01]. For more than 2 colors, it is not difficult to find a totally unimodular matrix such that even $\mathrm{lindisc}(A,c) > 1$ holds (e. g. in [Doe00b] an example was given that fulfills $\mathrm{lindisc}(A,3) \geq 1+\frac{1}{9}$).

Recently, Hebbinghaus, Schoen and Srivastav [HSS02] introduced the notion of positive $c$–color discrepancy and proved for two colors a tight discrepancy bound for hyperplanes in the $r$–dimensional vector space $\mathbb{F}_2^r$. Whether a tight bound also exists for the multicolor discrepancy, is an open problem.

## References


[AS92] N. Alon and J. H. Spencer. *The probabilistic method.* John Wiley & Sons Inc., New York, 1992.

[Ban98] W. Banaszczyk. Balancing vectors and Gaussian measures of $n$-dimensional convex bodies. *Random Structures & Algorithms*, 12:351–360, 1998.

[Bec81] J. Beck. Roth's estimate of the discrepancy of integer sequences is nearly sharp. *Combinatorica*, 1:319–325, 1981.

[BC87] J. Beck and W. L. Chen. *Irregularities of distribution*, volume 89 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 1987.

[BČC+02] T. Biedl, E. Čenek, T. Chan, E. Demaine, M. Demaine, R. Fleischer, and M. Wang. Balanced $k$-colorings. *Discrete Mathematics*, 254:19–32, 2002.

[BF81] J. Beck and T. Fiala. "Integer making" theorems. *Discrete Applied Mathematics*, 3:1–8, 1981.

[BG81] I. Barany and V. S. Grunberg. On some combinatorial questions in finite dimensional spaces. *Linear Algebra Appl.*, 41:1–9, 1981.

[BHK98] L. Babai, T. P. Hayes, and P. G. Kimmel. The cost of the missing bit: Communication complexity with help. In *Proceedings of the 30th STOC*, pages 673–682, 1998.

[BS84] J. Beck and J. Spencer. Integral approximation sequences. *Math. Programming*, 30:88–98, 1984.

[BS95] J. Beck and V. T. Sós. Discrepancy theory. In R. Graham, M. Grötschel, and L. Lovász, editors, *Handbook of Combinatorics.* 1995.

[Cha00] B. Chazelle. *The Discrepancy Method.* Princeton University, 2000.

[Doe00a] B. Doerr. Linear and hereditary discrepancy. *Combinatorics, Probability and Computing*, 9:349–354, 2000.

[Doe00b] B. Doerr. *Multi-Color Discrepancies.* 2000. Dissertation, Christian-Albrechts-Universität zu Kiel.

[Doe01] B. Doerr. Linear discrepancy of totally unimodular matrices. In *Proceedings of the 12th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2001.

[Doe02a] B. Doerr. Balanced coloring: Equally easy for all numbers of colors? In H. Alt and A. Ferreira, editors, *STACS 2002*, volume 2285 of *Lecture Notes in Computer Science*, pages 112–120, Berlin–Heidelberg, 2002. Springer Verlag.

[Doe02b] B. Doerr. Discrepancy in different numbers of colors. *Discrete Mathematics*, 250:63–70, 2002.

[DSW98] B. Doerr, A. Srivastav, and P. Wehr. Discrepancies of cartesian products of arithmetic progressions. Technical Report 98–42, Berichtsreihe des Mathematischen Seminars der Universität Kiel, 1998.

[Gra81] A. Graham. *Kronecker Products and Matrix Calculus.* Ellis Horwood Ltd., 1981.

[HSS02] N. Hebbinghaus, T. Schoen, A. Srivastav. Tight discrepancy of hyperplanes in finite vector spaces. Technical Report, Mathematisches Seminar, Universität zu Kiel, 2002.

[Hel99] M. Helm. On the Beck-Fiala theorem. *Discrete Math.*, 207:73–87, 1999.

[LSV86] L. Lovász, J. Spencer, and K. Vesztergombi. Discrepancies of set-systems and matrices. *Europ. J. Combin.*, 7:151–160, 1986.

[Mat95] J. Matoušek. Tight upper bound for the discrepancy of half-spaces. *Discr. & Comput. Geom.*, 13:593–601, 1995.

[Mat99] J. Matoušek. *Geometric Discrepancy.* Springer-Verlag, Berlin, 1999.

[Mat00] J. Matoušek. On the linear and hereditary discrepancies. *Europ. J. Combin.*, 21:519–521, 2000.

[MS96] J. Matoušek and J. Spencer. Discrepancy in arithmetic progressions. *J. Amer. Math. Soc.*, 9:195–204, 1996.

[MWW84] J. Matoušek, E. Welzl, and L. Wernisch. Discrepancy and approximations for bounded VC–dimension. *Combinatorica*, 13:455–466, 1984.

[Rot64] K. F. Roth. Remark concerning integer sequences. *Acta Arithmetica*, 9:257–260, 1964.

[Sár74] A. Sárközy. In P. Erdős and J. Spencer, editors, *Probabilistic Methods in Combinatorics.* Akadémia Kiadó, Budapest, 1974.

[Spe85] J. Spencer. Six standard deviations suffice. *Trans. Amer. Math. Soc.*, 289:679–706, 1985.

[Sri97] A. 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, New York, 1997. ACM.

[Sri01] A. Srivastav. Derandomization in combinatorial optimization. In: *Handbook of Randomized Computing*, Chapter 18, pp. 731–842, S. Rajasekaran, P.M. Pardalos, J. H. Reif, and J. D. P. Rolim, editors, Kluver Academic Publishers, 2001.

[SS96] A. Srivastav and P. Stangier. Algorithmic Chernoff-Hoeffding inequalities in integer programming. *Random Structures & Algorithms*, 8:27–58, 1996.

[VC71] V. N. Vapnik and A. Ya. Chervonenkis. On the uniform convergence of relative frequencies of events to their probabilities. *Theory Prob. Appl.*, 16:264–280, 1971.

[Weh97] P. Wehr (Knieper). *The Discrepancy of Arithmetic Progressions.* 1997. Dissertation, Institut für Informatik, Humboldt-Universität zu Berlin.

Benjamin Doerr, Mathematisches Seminar, Christian–Albrechts–Universität zu Kiel, Ludewig–Meyn–Str. 4, D–24098 Kiel, Germany

*E-mail address*: `bed@numerik.uni-kiel.de`

Anand Srivastav, Mathematisches Seminar, Christian–Albrechts–Universität zu Kiel, Ludewig–Meyn–Str. 4, D–24098 Kiel, Germany

*E-mail address*: `asr@numerik.uni-kiel.de`