Approximation of Multi–Color Discrepancy

Benjamin Doerr[1]* and Anand Srivastav[2]

[1] Mathematisches Seminar, Christian–Albrechts–Universität zu Kiel,
Ludewig–Meyn–Str. 4, D–24098 Kiel, Germany,
`bed@numerik.uni-kiel.de`,
WWW home page: `http://www.numerik.uni-kiel.de/~bed/`
[2] Mathematisches Seminar, Christian–Albrechts–Universität zu Kiel,
Ludewig–Meyn–Str. 4, D–24098 Kiel, Germany,
`asr@numerik.uni-kiel.de`,
WWW home page: `http://www.numerik.uni-kiel.de/~asr/`

Abstract. In this article we introduce (combinatorial) multi–color discrepancy and generalize some classical results from 2–color discrepancy theory to c colors. We give a recursive method that constructs c–colorings from approximations to the 2–color discrepancy. This method works for a large class of theorems like the six–standard–deviation theorem of Spencer, the Beck–Fiala theorem and the results of Matoušek, Welzl and Wernisch for bounded VC–dimension. On the other hand there are examples showing that discrepancy in c colors can not be bounded in terms of two–color discrepancy 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 to multi–colorings and prove multi–color versions of the the Beck–Fiala theorem and the Barany–Grunberg theorem.

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 about equal parts by the partition classes. Discrepancy measures the deviation of an optimal partition to an ideal one, that is one where all edges contains the same number of vertices in any partition class. As discrepancy is a NP–hard problem, efficient methods for constructing a good coloring can only approximate the discrepancy.

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. [BHK]) motivate the study of multi–color discrepancy.

---

* supported by the graduate school ‘Effiziente Algorithmen und Multiskalenmethoden’, Deutsche Forschungsgemeinschaft

1.1 Basic Definitions of Multi–Color Discrepancy

Let $\mathcal{H} = (V, \mathcal{E})$ denote a finite hypergraph, i. e. V is a finite set and $\mathcal{E} \subseteq 2^V$. A c–coloring of $\mathcal{H}$ is a mapping $\chi : X \to M$, where M is any set of cardinality c. For convenience, we may take $M = [c] := \{1, \ldots, c\}$. Sometimes — as we will see in this paper — a different set M will be of advantage. We define the *discrepancy of an edge $E \in \mathcal{E}$ for color $i \in M$ with respect to χ* by

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

This measures the deviation of the number of i–colored points in X from the average $\frac{|E|}{c}$. The *discrepancy of $\mathcal{H}$ with respect to χ* is

$$\mathrm{disc}(\mathcal{H}, \chi) := \max_{i \in M, E \in \mathcal{E}} \mathrm{disc}_{\chi,i}(E),$$

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

$$\mathrm{disc}(\mathcal{H}, c) := \min_{\chi : X \to [c]} \mathrm{disc}(\mathcal{H}, \chi). \tag{1}$$

For a definition of multi–color discrepancy of a matrix, the representation of the colors by suitable vectors is useful. For a color $i \in [c]$ let $m^{(i)} \in \mathrm{I\!R}^c$ be a vector with components defined by

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

for $j = 1, \ldots, c$. Put $M_c := \{m^{(i)} | i \in [c]\}$. For a c–coloring $\chi : X \to M_c$ we define

$$\mathrm{disc}(\mathcal{H}, \chi) = \max_{E \in \mathcal{E}} \left\| \sum_{x \in E} \chi(x) \right\|_\infty ,$$

and

$$\mathrm{disc}(\mathcal{H}, c) := \min_{\chi : X \to M_c} \mathrm{disc}(\mathcal{H}, \chi). \tag{2}$$

It is straightforward to see that the definitions in (1) and (2) are equivalent.

Let A be the incidence matrix of $\mathcal{H}$ and let $\overline{A}$ be the matrix which results from replacing every a_{ij} in A by $a_{ij} I_c$, where I_c denotes the unit matrix of dimension c. By identifying a $\chi : X \to M_c$ with a $c|X|$–dimensional vector in the natural way, we get

$$\mathrm{disc}(\mathcal{H}, c) = \min_{\chi : X \to M_c} \left\| \overline{A} \chi \right\|_\infty .$$

This motivates the matrix notion of multi–color discrepancy for arbitrary matrices $A \in \mathbb{R}^{m \times n}$:

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

Let $\overline{M}_c$ be the the convex hull of M_c in IR^c. For $p \in \overline{M}_c$ set $\overline{p} : X \to \mathbb{R}^c; x \mapsto p$. We define the *weighted discrepancy of $\mathcal{H}$ with weight p* by

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

and the *weighted discrepancy of $\mathcal{H}$* by

$$\operatorname{wd}(\mathcal{H}, c) := \max_{p \in \overline{M}_c} \operatorname{wd}(\mathcal{H}, c, p).$$

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

$$\operatorname{wd}(\mathcal{H}, c) = \max_{p \in \overline{E}_c} \min_{\chi : X \to E_c} \left\| \overline{A}(\overline{p} - \chi) \right\|_\infty .$$

We define the *linear discrepancy* of a matrix $A \in \mathbb{R}^{m \times n}$ in c colors by

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

The linear discrepancy of a hypergraph is of course the linear discrepancy of its incidence matrix. Let us write $A_0 \leq A$ to indicate that the matrix A_0 consists of some columns of the matrix A. For hypergraphs we write $\mathcal{H}_0 \leq \mathcal{H}$ if $\mathcal{H}_0$ is an induced subgraph of $\mathcal{H}$. Finally, the hereditary discrepancy in c colors is defined by

$$\operatorname{herdisc}(\mathcal{H}, c) := \min_{\mathcal{H}_0 \leq \mathcal{H}} \operatorname{disc}(\mathcal{H}_0, c).$$

Replace $\mathcal{H}$ by A for the matrix version. It is easy to show that the different notions of c–color discrepancy in the case $c = 2$ are identical with the usual notions of discrepancy (cf. the survey of Beck and Sós [BSó]) up to the constant factor of $1/2$. When citing 2–color results we will use the conventional notation (which has no parameter c in it, e. g. $\operatorname{herdisc}(\mathcal{H})$, so there is no danger of confusion).

1.2 Results

In this paper we give two methods for approximating the discrepancy in c colors, a recursive approach and the extension of the floating–color technique to multi–colorings. The recursive approach uses 2–color discrepancy theory. It turns out that under some not too strong assumptions we have an upper bound for the discrepancy in any number of colors that is roughly twice the bound for 2 colors. We show $\operatorname{disc}(\mathcal{H}, c) \leq 2.0005\Delta(\mathcal{H})$[1] (Beck–Fiala theorem for c colors), prove $\operatorname{disc}(\mathcal{H}, c) \leq 12\sqrt{n}$ for $|V| = |\mathcal{E}| = n$ (Spencer's bound generalized to c colors)

---

[1] $\Delta(\mathcal{H})$ is the degree of the hypergraph.

and derive bounds for c–color discrepancy of hypergraphs with bounded VC–dimension extending theorems of Matoušek, Welzl and Wernisch [MWW].

We give an example which shows the limit of the recursive approach: There are hypergraphs having 2–discrepancy zero, but arbitrarily large c–discrepancy even if c is a power of 2. Furthermore, the recursive method does not apply to the stronger linear discrepancy version of the Beck–Fiala theorem. For this situation and in the setting of the Barany–Grunberg theorem we present an approximation algorithm using vector–colorings and an extension of the floating–color technique to multi–colorings.

2 Recursive Coloring

As a warming up exercise let us fix the bound for discrepancy in c colors by the basic probabilistic method.

Theorem 1. *Let $\mathcal{H} = (V, \mathcal{E})$ be any hypergraph. Set $m := |\mathcal{E}|$ and $s = \max_{E \in \mathcal{E}} |E|$. Then $\operatorname{disc}(\mathcal{H}, c) \le \sqrt{\frac{1}{2} s \ln(2mc)}$. With probability greater than $\frac{1}{2}$ we can find a c–coloring χ with discrepancy at most $\sqrt{\frac{1}{2} s \ln(4mc)}$ (this leads to a randomized algorithm with arbitrarily small error probability).*

Proof. Define a random c–coloring χ 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 $v \in V$, $i \in [c]$. For fixed i these are independent random variables. Set $X_{i,E} := \sum_{v \in E} X_{i,v}$ for all $E \in \mathcal{E}$, $i \in [c]$. From [ASE, Theorem A.4] we know $P(|X_{i,E}| > \alpha) < 2e^{-2\alpha^2/|E|}$ for all real $\alpha > 0$. With $\alpha = \sqrt{\frac{1}{2} s \ln(2mc)}$ it is easy to see that the random coloring χ fulfills $\operatorname{disc}(\mathcal{H}, \chi) \le \sqrt{\frac{1}{2} s \ln(2mc)}$ with non–zero probability. Choosing $\alpha = \sqrt{\frac{1}{2} s \ln(4mc)}$ we get this probability below $\frac{1}{2}$.

For 2–colors we get $\operatorname{disc}(\mathcal{H}) \le \sqrt{2s \ln(4m)}$, while the best bound in the classical approach is $\operatorname{disc}(\mathcal{H}) \le \sqrt{2s \ln(2m)}$. Note that with the method of conditional probabilities the existence result of Theorem 1 can be derandomized.

The basic idea of recursive coloring 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 2–color discrepancy of all induced subhypergraphs is uniformly bounded, such a recursive method can be analyzed, even for n not a power of 2. This will lead to a generalization of the six–standard–deviation theorem of Spencer, the discrepancy bound of Beck–Fiala and the bound using the primal and dual shatter function of Matoušek, Welzl and Wernisch. At the end of this section we will show the limits of the recursive approach. For example, for the linear discrepancy in c–colors recursive methods fail, and we need other methods, which will be introduced in the next section.

The following lemma analyses a single step in the recursion.

Lemma 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, i. e. $p \in [0,1]^c$ such that $\|p\|_1 = 1$. Denote by $p_{|C_j}$ the vector taking the components of p with indices corresponding to the colors in the set C_j, and set $q_j = \|p_{|C_j}\|_1$, $j \in [2]$. Let χ_0 be a 2–coloring of X, set $X_1 := \chi_0^{-1}(1)$, $X_2 := \chi_0^{-1}(-1)$. 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 χ and the weight p is*

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

In particular $\mathrm{wd}(\mathcal{H}, c, p) \le \frac{p_i}{q_j}\,\mathrm{wd}(\mathcal{H}, 2, (q_1, q_2)) + \max_{\mathcal{H}_0 \le \mathcal{H}} \mathrm{wd}(\mathcal{H}_0, |C_j|, \frac{1}{q_j} p_{|C_j})$.

The proof is straightforward and will appear in the full version. We now investigate the case where we all induced subgraphs have a common bound on the weighted discrepancy in two colors. This is an important case for two reasons: Firstly, the proof of some results on two–color discrepancy actually gives an information about the weighted discrepancy of the induces subgraphs (e. g. in the Beck–Fiala setting). Secondly, the linear discrepancy (and thus also the weighted discrepancies of all subgraphs, note $\mathrm{wd}(\mathcal{H}, c) \le \frac{1}{2}\,\mathrm{lindisc}(\mathcal{H})$) are bounded by the hereditary discrepancy: From [BSp,LSV] we know $\mathrm{lindisc}(\mathcal{H}) \le 2\,\mathrm{herdisc}(\mathcal{H})$ (for a recent improvement see also [D]).

We represent the iterated partitioning of C by a binary tree. We call a binary tree $T = (V_T, E_T)$ a *partition tree* for C, if the following conditions are satisfied: the nodes are subsets of C, the root is C, all leaves are singletons of C and the two son nodes form a partition of their common father node. For every color $i \in C$ there is a unique path of type $C = C_0^{(i)} \subset C_1^{(i)} \subset \ldots \subset C_{k(i)}^{(i)} = \{i\}$ in the partition tree.

Theorem 2. *Let $\mathrm{wd}(\mathcal{H}_0, 2) \le K$ 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, i. e. $p \in [0,1]^c$ such that $\sum_{i=1}^{c} p_i = 1$. 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*

$$\big||E \cap \chi^{-1}(i)| - p_i|E|\big| \le p_i \sum_{l=1}^{k(i)} \frac{1}{\|p_{|C_l^{(i)}}\|_1} K.$$

Proof. By induction on the height of T. For $h(T) = 0$ we have just one color and both sides of the inequality become zero. So let T be of height greater than zero and assume that the theorem is true for smaller heights. Let C_1 and C_2 be the sons of C in T. Set $q_j := \sum_{k \in C_j} p_k$, $j = 1, 2$. By the assumption of the theorem there is a 2–coloring $\chi_0 : X \to [2]$ such that $\big||E \cap \chi_0^{-1}(j)| - q_j|E \cap X_j|\big| \le K$ for all colors $j \in [2]$ and edges $E \in \mathcal{E}$. Put $X_j := \chi_0^{-1}(j)$, $j = 1, 2$. Denote by T_j the subtree having C_j as its root. Then $\mathcal{H}_{|X_j}$, the set of colors C_j with weight $\frac{1}{q_j}p_{|C_j}$ together with the T_j fulfill the assumption of this theorem. Hence by induction hypothesis there are $\chi_j : X_j \to C_j$, $j \in 1, 2$ such that

$$\left| |E \cap X_j \cap \chi_j^{-1}(i)| - \frac{1}{q_j} p_i |E \cap X_j| \right| \le \frac{p_i}{q_j} \sum_{l=2}^{k(i)} \frac{1}{\frac{1}{q_j}\|p_{|C_l^{(i)}}\|_1} K$$

for all $E \in \mathcal{E}$, $j \in [2]$, $i \in C_j$. This and lemma 1 prove that $\chi = \chi_1 \cup \chi_2$ is as desired.

The c–color discrepancy problem is simply the case where all weights are equal. In this case only the size of the partitioning sets is of importance. Hence the following simpler structure can be investigated:

A *partition tree* for a positive integer n is binary tree $T = (V_T, E_T)$ together with a labeling $V_T \to [n]$ such that the following conditions are satisfied: the root is labeled n, all leaves are labeled 1 and for all non–leaf nodes the labels of the two sons sum up to the label of the node itself.

For every path $P : r = v_0, v_1, v_2, \dots, v_k$ connecting the root r and a leaf v_k we call $v_T(P) = \sum_{i=1}^{k} \frac{1}{l(v_i)}$ 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.

Theorem 3. *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 $v(T) = v(c)$. From (V_T, E_T) we build a partition tree of $[c]$ with same tree structure. 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 $[c]$. Set $p = \frac{1}{c}\mathbf{1}_c$. For every path P connecting the root with a leaf $\{i\}$ in T^* we get that the bound in theorem 2 for the discrepancy in color i is equal to $n_T(P)K$, hence the discrepancy is bounded by $v(T)K = v(c)K$.

What remains is the calculation of $v(c)$. Set $\lfloor n \rfloor_2 = 2^{\lfloor \log_2 n \rfloor}$, the largest power of 2 that is not larger than n, and $\lceil n \rceil_2 = 2^{\lceil \log_2 n \rceil}$, the smallest one not smaller than n. Denote by $n_1(n)$ the number of 1s in the binary representation of n (e.g. $n_1(9) = 2$). We give a lower and an upper bound for $v(n)$. If n is a power of 2, both bounds coincide.

Lemma 2. *For all* $n \in \mathbb{N}$, $n \ge 2$ *we have*

(i) $v(n) \ge 2 - \frac{1}{\lceil n \rceil_2}$.
(ii) $v(n) \le 2 + (n_1(n) - 2)\frac{1}{\lfloor n \rfloor_2} \le 2 + (\log_2(\lfloor n \rfloor_2) - 1)\frac{1}{\lfloor n \rfloor_2}$.
(iii) $v(n) \le 2.0005$.

Proof. (i): For the proof of the lower bound let $T = (V_T, E_T)$ together with l be any partition tree of n. Then there is a path $v_0, \dots, v_k$ of length $k \geq \log_2 \lceil n \rceil_2$ such that v_k is a leaf and $l(v_{i-1}) \leq 2l(v_i)$ for all $i \in [k]$. Thus

$$\sum_{i=1}^{k} \frac{1}{l(v_i)} \geq \sum_{i=0}^{k-1} 2^{-i} = 2 - \frac{1}{2^{k-1}} \geq 2 - \frac{1}{\lceil n \rceil_2}.$$

(ii): For the proof of the upper bound we give a strategy how to construct a partition tree of n. We do so recursively: 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 $l(s_2(v)) = l(v) - l(s_1(v))$. Immediately we see that we only need to investigate the path $P : r, s_2(r), s_2(s_2(r)), \dots$ — 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 $n_1(n)$ vertices of P are greater than or equal to $\lfloor n \rfloor_2$, so their contribution to $v(P)$ is not greater than $n_1(n) \frac{1}{\lfloor n \rfloor_2}$. The rest of the vertices are labeled by $\frac{2}{\lfloor n \rfloor_2}, \frac{4}{\lfloor n \rfloor_2}, \dots$ up to 1. This sums up to $2 - \frac{2}{\lfloor n \rfloor_2}$ and the the first inequality is proven. For the second inequality note that $n_1(n) - 1 \leq \log_2(\lfloor n \rfloor_2)$.

(iii): For $n \geq n_0 := 2^{17} - 1$ the last inequality follows from (ii), as $(n_1(n) - 2)\frac{1}{\lfloor n \rfloor_2} \leq \frac{\log_2(\lfloor n \rfloor_2) - 1}{\lfloor n \rfloor_2} \leq \frac{\log_2(\lfloor n_0 \rfloor_2) - 1}{\lfloor n_0 \rfloor_2}$. For the remaining small numbers, $v(n)$ can be computed in $\mathcal{O}(n^2)$–time and attains its maximum value for $n = 909$, namely $v(909) \approx 2.000450$.

From this we derive a c–color version of the Beck–Fiala theorem:

Theorem 4. *For any hypergraph $\mathcal{H}$ we have*

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

The complexity to construct a c–coloring respecting this bound is less than $2\log_2 \lfloor c \rfloor_2$ times the complexity for the 2–color case.

Proof (sketch). The bound is a direct consequence of theorem 3 and Lemma 2 and the Beck–Fiala theorem for 2 colors. For the complexity estimation let C be the constant such that the construction of a 2–coloring as in the theorem of Beck–Fiala has complexity bounded by Cn^4, where n shall denote the number of vertices of $\mathcal{H}$. Then the partition tree of Lemma 2 (ii) yields the bound of $n_1(c) - 1 + \log_2 \lfloor c \rfloor_2 \leq 2\log_2 \lfloor c \rfloor_2$.

The generalization of the 'Six Standard Deviation' theorem of Spencer [Sp1] is:

Theorem 5. *For any hypergraph $\mathcal{H} = (V, \mathcal{E})$ such that $n := |V| = |\mathcal{E}|$ we have*

$$\operatorname{disc}(\mathcal{H}, c) < 5.32 v(c)\sqrt{n} \leq 12\sqrt{n}.$$

Proof. Spencer's proof shows that $\operatorname{herdisc}(\mathcal{H}) \leq 5.32\sqrt{n}$.

Without proof we may remark that this bound is tight (apart from the constant).

The recursive approach also generalizes a result of Matoušek, Welzl and Wernisch [MWW] connecting discrepancy with the primal shatter function $\pi_{\mathcal{H}}$ and dual shatter function $\pi^*_{\mathcal{H}}$ of a hypergraph (the shatter functions are closely related to the VC–dimension of $\mathcal{H}$).

Theorem 6. *Let* $\mathcal{H} = (V, \mathcal{E})$ *be a hypergraph on* n *points. Let* $d > 1$. *If* $\pi_{\mathcal{H}} = \mathcal{O}(m^d)$, *then* $\operatorname{disc}(\mathcal{H}, c) = \mathcal{O}(n^{\frac{1}{2}-\frac{1}{2d}}(\log n)^{1+\frac{1}{2d}})$. *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 can be chosen independent of* c.

Summary of the recursive method: We see that the recursive method is very effective in situations where we can bound the weighted discrepancy of the induced subgraphs. This is always the case if we know the hereditary discrepancy of $\mathcal{H}$. There are cases where the recursive approach is the only result we have. We do not have a direct proof for a result like theorem 5. We feel that the original proof relies heavily on the fact that only two colors are considered. On the other hand the recursive approach clearly has its limitations: We can get results on weighted discrepancy, but we do not get any on linear discrepancy (e. g. in the Beck–Fiala setting). To apply recursion, we need a 2–color result for the weighted discrepancy, even in the case that c is a power of two. The following example illustrates this limitation.

Example: Let $n \in \mathbb{N}$. Set $\mathcal{H}_n = ([2n], \{X \subseteq [2n] \mid |X \cap [n]| = |X \setminus [n]|\})$. Obviously, $\operatorname{disc}(\mathcal{H}_n) = 0$. On the other hand it is not difficult to show $\operatorname{disc}(\mathcal{H}_n, 4) = \frac{1}{8}n$ for all $n \in \mathbb{N}$.

3 Floating Vector–Colors

In this section we give analogous results to the Beck–Fiala theorem and the Barany–Grunberg theorem. In the 2–color case both are proved using a rounding strategy. We show how this strategy can be extended to the multi–color case. The key in both cases is the representation of the colors by the vectors $m^{(i)}$ defined in section 1.

3.1 Beck–Fiala Theorem for Linear Discrepancy

The maximum degree $\Delta(\mathcal{H}) := \max_{x \in X} |\{E \in \mathcal{E} | x \in E\}|$ is one of the few parameters of a hypergraph which give a good bound on the discrepancy. The Beck–Fiala theorem [BF] states $\operatorname{disc}(\mathcal{H}) < 2\Delta(\mathcal{H})$. Actually Beck and Fiala proved a stronger 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 on $\mathbb{R}^n$. Then $\operatorname{lindisc}(A) < 2||A||_1$. We were not able to generalize this theorem to c–colors by the recursive method, and this might also not be possible. The difficulty is that in the phase of coloring a subhypergraph in the recursive method some information on the weights $p_j(v)$, j a color, v a vertex, is not available anymore.

The method of floating colors though can be extended to multi–colorings. We have

Theorem 7. *For any matrix* A*,* $\mathrm{lindisc}(A,c) < 2\|A\|_1$*. The problem of computing a* $\chi : [n] \to M_c$ *for a given* $p : [n] \to \overline{M}_c$ *such that* $\|\overline{A}(p-\chi)\|_\infty < 2\|A\|_1$ *has time–complexity* $\mathcal{O}(c^4)$ *times the complexity of the* 2*–color problem.*

Proof. Set $\Delta := \|A\|_1$ and $\overline{A} = (\overline{a}_{ij})$ the matrix resulting from A by replacing every entry a_{ij} by $a_{ij} I_c$ as introduced in section 1. Note that $\Delta = \|\overline{A}\|_1$. Let $p : [n] \to \overline{M}_c$. Set $\chi = p$. Successively we will change χ to a $\chi : [n] \to M_c$.

Set $J := \{j \in [cn] | \chi_j \notin \{-\frac{1}{c}, \frac{c-1}{c}\}\}$ and call these columns floating (the others fixed). Set $I := \{i \in [cm] | \sum_{j\in J} |\overline{a}_{ij}| > 2\Delta\}$ and call these rows 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 $\sum_{k=0}^{c-1} \chi_{cj-k} = 0$ 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 will increase the number of equations investigated below, and is thus the reason why the multi–color bound is off the classical result by a factor of 2.

We have

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

Note further that for every vertex it cannot happen that exactly one color is floating, so

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

is a system of at most $\frac{1}{2}|J|$ equations. Hence the system

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

is under–determined (taking just the $\chi_j, j \in J$ as variables). Thus there is a non–trivial solution $x \in \mathbb{R}^J$. Expand x to $x_E \in \mathbb{R}^{cn}$ by

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

Note that for any such x we can replace χ by $\chi + x_E$ in (i), (ii). Choose $\lambda \in \mathbb{R}$ such that at least one component of $\chi + \lambda x_E$ becomes fixed and all colors are still in $\overline{M}_c$, i. e. $\chi + \lambda x_E \in \overline{M}_c^n$. Set $\chi := \chi + \lambda x_E$. Since (i), (ii) are fulfilled, we

can continue this rounding process until all χ_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 χ and J when the row i first became ignored. Note that $\sum_{j \in J^{(0)}} |\overline{a}_{ij}| < 2\Delta$ by definition of I. Thus
$|(A(p-\chi))_i| = |(A(p-\chi^{(0)}))_i + (A(\chi^{(0)}-\chi))_i| = 0 + \sum_{j \in J^{(0)}} \overline{a}_{ij}(\chi^{(0)}-\chi)(i) < 2\Delta.$

3.2 Theorem of Barany–Grunberg

The theorem of Barany–Grunberg [BG] for 2 colors states:

Theorem 8. *Let* $||\cdot||$ *be any norm on* $\mathbb{R}^n$ *and* $v_1, v_2, \ldots, v_k$ *be a finite sequence of arbitrary length of vectors of norm at most* 1 *in* $\mathbb{R}^n$*. Then there are signs* $\varepsilon_i, i = 1, \ldots, k$ *such that for all* $l \in [k]$ *we have*

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

As in the proof of the Beck–Fiala theorem we describe the colors by vectors from the set M_c. As above let $v_1, v_2, \ldots, v_k$ be a finite sequence of vectors in $\mathbb{R}^n$ and $||\cdot||$ a norm on $\mathbb{R}^n$. A mapping $\chi : [k] \to M_c, i \mapsto \chi^{(i)}$ is called a coloring for these vectors. Since Barany–Grunberg works for any norm, we need to lift our norm to a suitable norm on $\mathbb{R}^{cn}$: Define a norm $||\cdot||_c$ on $\mathbb{R}^{cn}$ by $||v||_c := \max_{j \in [c]} ||v_{|\{j, j+c, \ldots, j+(n-1)c\}}||$.

We need a calculus for substituting vectors into each other. For any two vectors $v \in \mathbb{R}^n$, $w \in \mathbb{R}^m$, we define $v * w$ to be the vector $u \in \mathbb{R}^{nm}$ such that $u_{(i-1)n+j} = v_i w_j$ for all $i \in [n]$, $j \in [m]$. So u is obtained by replacing every entry v_i of v by $v_i w$. The following lemma follows from direct calculations.

Lemma 3. $\left\| \sum_{i \in [k]} v_i * \chi^{(i)} \right\|_c = \max_{j \in [c]} \left\| \sum_{i \in [k], \chi^{(i)} = m_j} v_i - \frac{1}{c} \sum_{i \in [k]} v_i \right\|.$

The latter expression in the lemma measures the maximal deviation (over the colors) of the sum of vectors in this color from the average $\frac{1}{c} \sum_{i \in [k]} v_i$ with respect to the norm $||\cdot||$. This is the c–color analogue of the discrepancy term $\left\| \sum_{i=1}^{l} \varepsilon_i v_i \right\|$ in the Barany–Grunberg theorem.

The multi–color version of Barany–Grunberg is:

Theorem 9. *Let* $||\cdot||$ *be any norm on* $\mathbb{R}^n$ *and* $v_1, v_2, \ldots, v_k$ *be a finite sequence of vectors of norm at most* 1 *in* $\mathbb{R}^n$*. Then there is a* c*–partition* $I_1, \ldots, I_c$ *of* $[k]$ *such that for all* $l \in [k]$ *and* $j \in [c]$ *we have*

$$\left\| \sum_{i \in I_j \cap [l]} v_i - \frac{1}{c} \sum_{i=1}^{l} v_i \right\| < (c-1)n.$$

The complexity for computing this partition is $\mathcal{O}(c^3)$ *times the complexity of the 2–color case.*

Proof. By Lemma 3 it is enough to construct a coloring $\chi : [k] \to M_c$ such that

$$(*) \qquad \left\| \sum_{i \in [l]} v_i * \chi^{(i)} \right\|_c \leq (c-1)n \text{ for all } l \in [k].$$

We sketch an algorithm for this task: To start with put $A := [n]$ and $\chi_j^{(i)} := 0$ for all $i \in [k]$, $j \in [c]$. We repeat the following rounding process: Let us call those $\chi_j^{(i)}$ where $i \in A$ and $\chi_j^{(i)} \notin \{\frac{c-1}{c}, -\frac{1}{c}\}$ *variables.* We try to find a nontrivial solution of the system of equations

$$\sum_{i \in A} v_i * \chi^{(i)} = 0 \quad \text{for all } j \in [c-1].$$

$$\sum_{j \in [c]} \chi_j^{(i)} = 0 \quad \text{for all } i \in A.$$

If one exists, change χ in the way that one variable becomes $\frac{c-1}{c}$ or $-\frac{1}{c}$ and all variable stay in $[-\frac{1}{c}, \frac{c-1}{c}]$. If not, then increase the number of vectors under consideration, i. e. set $A := A \cup \{\max A + 1\}$, if $A \neq [k]$, and stop, if $A = [k]$. After the rounding process stopped, change the remaining variables to $\frac{c-1}{c}$ or $-\frac{1}{c}$ in such a way that all $\chi^{(i)}$ are in M_c. For the correctness proof we calculate the number of indices in A such that the color of v_i is not completely determined to be at most $(c-1)n$. This together with the properties of $*$ and $\|\cdot\|_c$ shows that χ fulfills the sufficient condition $(*)$. The journal version of this paper will contain the missing details.

Note that this time the number of colors influences our bound in a much stronger way than in the theorems before: In the Beck–Fiala situation, the multi–color bound is off the 2–color one by a factor of 2 for the simple reason that in the 2–color case one can actually ignore one color (the discrepancy in both colors is the same). In the Barany–Grunberg theorem the bound contains a factor of $c-1$ and both the original result (translated to our notion) and our result yield the same bound. This is due to the fact that we use an arbitrary norm. Thus in the analysis we can 'ignore' a vector v_i only from the point on when its color $\chi^{(i)}$ is completely determined. Compare this to the proof of Beck–Fiala, where the fixing of a single χ_k (which is equivalent to a $\chi_j^{(i)}$ here) improves the situation.

4 Conclusion

This paper presents two types of results. Firstly, some important discrepancy theorems can be lifted to c–colorings via 2–color results. Secondly, using the right vector–representation of the colors and an appropriate matrix calculus some results of the 2–color discrepancy theory can be generalized to any number of colors. Nevertheless c–color discrepancy theory is more than a generalization for generalization's sake as the recent applications in communication complexity might indicate. We hope that our paper can spur further research in this part of discrepancy theory.

References

[ASE] N. Alon, J. Spencer, P. Erdős: *The Probabilistic Method,* John Wiley & Sons, Inc., 1992.

[BHK] L. Babai, T. P. Hayes and P. G. Kimmel, *The cost of the Missing Bit: Communication Complexity with Help,* 30th STOC, 1998, 673–682.

[BF] J. Beck and T. Fiala, *"Integer making" Theorems,* Discrete Applied Mathematics **3** (1981), 1–8.

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

[BSó] J. Beck and V. Sós, *Discrepancy Theory* in R. Graham, M. Grötschel and L. Lovász, *Handbook of Combinatorics*, 1995, Chapter 26.

[BSp] J. Beck and J. Spencer, *Integral approximation sequences*, Math. Programming **30** (1984), 88–98.

[D] B. Doerr, *Linear and Hereditary Discrepancy*, 1999, accepted for publication in Combinatorics, Probability and Computing.

[LSV] L. Lovász, J. Spencer and K. Vesztergombi, *Discrepancies of set–systems and matrices*, European J. Combin. **7** (1986), 151–160.

[Sp1] J. Spencer, *Six Standard Deviation Suffice*, Trans. Amer. Math. Soc. **289** (1985), 679–706.

[Sp2] J. Spencer, *Ten Lectures on the Probabilistic Method*, SIAM, Philadelphia, 1987.

[MWW] J. Matoušek, E. Welzl and L. Wernisch, *Discrepancy and approximations for bounded VC–Dimension,* Combinatorica **13** (1984), 455–466.