EuroComb 2005 DMTCS proc. **AE**, 2005, 323–328

Discrepancy of Products of Hypergraphs

Benjamin Doerr[1], Michael Gnewuch[2†] and Nils Hebbinghaus[3‡]

[1]*Max-Planck-Institut für Informatik, Stuhlsatzenhausweg 85, D-66123 Saarbrücken, e-mail: doerr@mpi-sb.mpg.de*
[2]*Max-Planck-Institut für Mathematik in den Naturwissenschaften, Inselstraße 22, D-04103 Leipzig, e-mail: gnewuch@mis.mpg.de*
[3]*Institut für Informatik und Praktische Mathematik, Christian-Albrechts-Universität Kiel, Christian-Albrechts-Platz 4, D-24118 Kiel, e-mail: nhe@numerik.uni-kiel.de*

For a hypergraph $\mathcal{H} = (V, \mathcal{E})$, its d–fold symmetric product is $\Delta^d\mathcal{H} = (V^d, \{E^d | E \in \mathcal{E}\})$. We give several upper and lower bounds for the c-color discrepancy of such products. In particular, we show that the bound $\operatorname{disc}(\Delta^d\mathcal{H}, 2) \le \operatorname{disc}(\mathcal{H}, 2)$ proven for all d in [B. Doerr, A. Srivastav, and P. Wehr, Discrepancy of Cartesian products of arithmetic progressions, Electron. J. Combin. 11(2004), Research Paper 5, 16 pp.] cannot be extended to more than $c = 2$ colors. In fact, for any c and d such that c does not divide $d!$, there are hypergraphs having arbitrary large discrepancy and $\operatorname{disc}(\Delta^d\mathcal{H}, c) = \Omega_d(\operatorname{disc}(\mathcal{H}, c)^d)$. Apart from constant factors (depending on c and d), in these cases the symmetric product behaves no better than the general direct product $\mathcal{H}^d$, which satisfies $\operatorname{disc}(\mathcal{H}^d, c) = O_{c,d}(\operatorname{disc}(\mathcal{H}, c)^d)$.

Keywords: discrepancy, hypergraphs, Ramsey theory

Introduction

We investigate the discrepancy of certain products of hypergraphs. In [2], Srivastav, Wehr and the first author noted the following. For a hypergraph $\mathcal{H} = (V, \mathcal{E})$ define the d–fold direct product and the d–fold symmetric product by

$$\begin{aligned} \mathcal{H}^d &:= (V^d, \{E_1 \times \cdots \times E_d \,|\, E_i \in \mathcal{E}\}), \\ \Delta^d\mathcal{H} &:= (V^d, \{E^d \,|\, E \in \mathcal{E}\}). \end{aligned}$$

Then for the (two-color) discrepancy

$$\operatorname{disc}(\mathcal{H}) := \min_{\chi: V \to \{-1,1\}} \max_{E \in \mathcal{E}} \left| \sum_{v \in E} \chi(v) \right|,$$

†Supported by the Deutsche Forschungsgemeinschaft, Grant SR7/10-1.
‡Supported by the Deutsche Forschungsgemeinschaft, Graduiertenkolleg 357.

we have

$$\operatorname{disc}(\mathcal{H}^d) \leq \operatorname{disc}(\mathcal{H})^d \quad \text{and} \quad \operatorname{disc}(\Delta^d \mathcal{H}) \leq \operatorname{disc}(\mathcal{H}).$$

In this paper, we show that the situation is more complicated for discrepancies in more than two colors. In particular, it depends highly on the dimension d and the number of colors, whether the discrepancy of symmetric products is more like the discrepancy of the original hypergraph or the d-th power thereof. Let us make this precise:

For $c \in \mathbb{N}_{\geq 2}$, a *c–coloring* of a hypergraph $\mathcal{H} = (V, \mathcal{E})$ is a mapping $\chi : V \to [c]$, where $[r] := \{n \in \mathbb{N} \mid n \leq r\}$ for any $r \in \mathbb{R}$. The discrepancy problem asks for balanced colorings of hypergraphs in the sense that each hyperedge shall contain the same number of vertices in each color. The *discrepancy of* χ and the *c–color discrepancy of* $\mathcal{H}$ are defined by

$$\begin{aligned}\operatorname{disc}(\mathcal{H}, \chi) &:= \max_{E \in \mathcal{E}} \max_{i \in [c]} \left| |\chi^{-1}(i) \cap E| - \tfrac{1}{c}|E| \right|, \\ \operatorname{disc}(\mathcal{H}, c) &:= \min_{\chi: V \to [c]} \operatorname{disc}(\mathcal{H}, \chi).\end{aligned}$$

These notions were introduced in [1] extending the discrepancy problem for hypergraphs to arbitrary numbers of colors. Note that $\operatorname{disc}(\mathcal{H}) = 2\operatorname{disc}(\mathcal{H}, 2)$ holds for all $\mathcal{H}$. In this more general setting, the product bound proven in [2] is

$$\operatorname{disc}(\mathcal{H}^d, c) \leq c^{d-1} \operatorname{disc}(\mathcal{H}, c)^d. \tag{1}$$

However, as we show in this paper the relation $\operatorname{disc}(\Delta^d \mathcal{H}, c) = O(\operatorname{disc}(\mathcal{H}, c))$ does not hold in general. We give a characterization of those values of c and d, for which it is fulfilled for every hypergraph $\mathcal{H}$. In particular, we present for all c, d, k such that c does not divide $d!$ a hypergraph $\mathcal{H}$ having $\operatorname{disc}(\mathcal{H}, c) \geq k$ and $\operatorname{disc}(\Delta^d \mathcal{H}, c) = \Omega_d(k^d)$. In the light of (1), this is largest possible apart from factors depending on c and d only.

On the other hand, there are further situations where this worst case does not occur. We state some results of this type in the last section.

Symmetric Direct Products Having Large Discrepancy

Let $S(d, l)$, $d, l \in \mathbb{N}$, denote the Stirling numbers of the second kind. For $c \in \mathbb{N}$ and $\lambda \in N_0$ we write $c \mid \lambda$ if there exists an $m \in \mathbb{N}_0$ with $mc = \lambda$.

Theorem 1 *Let* $c, d \in \mathbb{N}$.

If $c \mid k!\, S(d, k)$ *for all* $k \in \{2, \ldots, d\}$, *then every hypergraph* $\mathcal{H}$ *satisfies*

$$\operatorname{disc}(\Delta^d \mathcal{H}, c) \leq \operatorname{disc}(\mathcal{H}, c)\,. \tag{2}$$

If $c \nmid k!\, S(d, k)$ *for some* $k \in \{2, \ldots, d\}$, *then there exists a hypergraph* $\mathcal{K}$ *such that*

$$\operatorname{disc}(\Delta^d \mathcal{K}, c) \geq \frac{1}{3\, k!} \operatorname{disc}(\mathcal{K}, c)^k\,, \tag{3}$$

and $\mathcal{K}$ *can be chosen to have arbitrary large discrepancy* $\operatorname{disc}(\mathcal{K}, c)$.

We state some simple consequences of Theorem 1:

Corollary 2 *(a) Let $d \geq 3$ be an odd number. Then $\operatorname{disc}(\Delta^d\mathcal{H}, 3) \leq \operatorname{disc}(\mathcal{H}, 3)$ holds for any hypergraph $\mathcal{H}$.*

(b) Let $d \geq 2$ be an even number and $c = 3l$, $l \in \mathbb{N}$. There exists a hypergraph $\mathcal{H}$ with arbitrary large discrepancy that fulfills $\operatorname{disc}(\Delta^d\mathcal{H}, c) \geq \frac{1}{6}\operatorname{disc}(\mathcal{H}, c)^2$.

Proof: Obviously $3|\, k!$ for all $k \geq 3$. Since $S(d, 2) = 2^{d-1} - 1$, we have $3|\, S(d, 2)$ if and only if d is odd. □

Corollary 3 *Let $l \in \mathbb{N}$ and $c = 4l$. For all $d \geq 2$ there exists a hypergraph $\mathcal{H}$ with arbitrary large discrepancy such that $\operatorname{disc}(\Delta^d\mathcal{H}, c) \geq \frac{1}{6}\operatorname{disc}(\mathcal{H}, c)^2$.*

Proof: As $S(d, 2) = 2^{d-1} - 1$ is an odd number, we have $4 \nmid 2!\, S(d, 2)$. □

Our proof of Theorem 1 uses the following lemma.

Lemma 4 *Let $c, d \in \mathbb{N}$. For all $m \in \mathbb{N}$ there exists an $n \in \mathbb{N}$ having the following property: For each c-coloring $\chi : [n]^d \to [c]$ we find a subset $T \subseteq [n]$ with $|T| = m$ such that for all $l \in [d]$ each l-dimensional simplex in T^d is monochromatic with respect to χ.*

Hereby an l-dimensional simplex in T^d is of the following form: Fix a partition $\{J_1, \ldots, J_l\}$ of $[d]$ and define vectors $f^{(i)}$ by $f^{(i)}_k = 1$ if $k \in J_i$ and 0 else. Then

$$S := \Big\{ \sum_{i=1}^{l} \alpha_i f^{(i)} \,|\, \alpha_1, \ldots, \alpha_l \in T\,, \alpha_1 < \ldots < \alpha_l \Big\}$$

is an l-dimensional simplex in T^d. The proof of Lemma 4 is based on an argument from Ramsey theory (see, e.g., [3, Section 1.2]).

Related to Lemma 4 is a result of Gravier, Maffray, Renault and Trotignon [4]. They have shown that for any $m \in \mathbb{N}$ there is an $n \in \mathbb{N}$ such that any collection of n different sets contains an induced subsystem on m points such that one of the following holds: (a) each vertex forms a singleton, (b) for each vertex there is a set containing all m points except this one, or (c) by sufficiently ordering the points $p_1, \ldots, p_m$ we have that all sets $\{p_1, \ldots, p_\ell\}$, $\ell \in [m]$, are contained in the system.[§]

In our language, this means that any $0, 1$ matrix having n distinct rows contains a $m \times m$ submatrix that can be transformed through row and column permutations into a matrix that is (a) a diagonal matrix, (b) the inverse of a diagonal matrix, or (c) a triangular matrix.

[§] To be precise, the authors also have the empty set contained in cases (a) and (c) and the whole set in case (b). It is obvious that by altering m by one, one can transform one result into the other.

Hence this result is very close to the assertion of Lemma 4 for dimension $d = 2$ and $c = 2$ colors. It is stronger in the sense that not only monochromatic simplices are guaranteed, but also a restriction to 3 of the 8 possible color combinations for the 3 simplices is given. Of course, this stems from the facts that (a) column and row permutations are allowed, (b) not a submatrix with index set T^2 is provided but only one of type $S \times T$, and (c) the assumption of having different sets ensures sufficiently many entries in both colors.

Further Upper Bounds

Besides the first part of Theorem 1, there are more ways to obtain upper bounds.

Theorem 5 *Let $\mathcal{H} = (V, \mathcal{E})$ be a hypergraph. Let p be a prime number, $q \in \mathbb{N}$ and $c = p^q$. Furthermore, let $d \geq c$ and $s = d - (p-1)p^{q-1}$. Then* $\operatorname{disc}(\Delta^d \mathcal{H}, c) \leq \operatorname{disc}(\Delta^s \mathcal{H}, c)$.

As a corollary, we state a less general (but also less technical) version of Theorem 5:

Corollary 6 *Let $\mathcal{H} = (V, \mathcal{E})$ be a hypergraph. If c is a prime number, $q \in \mathbb{N}$ and $d = c^q$, then* $\operatorname{disc}(\Delta^d \mathcal{H}, c) \leq \operatorname{disc}(\mathcal{H}, c)$.

Proof: Use $c^q = 1 + (c-1) \sum_{j=0}^{q-1} c^j$ and Theorem 5 (repeatedly). □

The following result is an extension of the first statement of Theorem 1.

Theorem 7 *Let c, $d \in \mathbb{N}$, and let $d' \in \{2.\ldots, d\}$. If $c \,|\, k!\, S(d', k)$ for all $k \in \{2, \ldots, d'\}$, then*

$$\operatorname{disc}(\Delta^d \mathcal{H}, c) \leq \operatorname{disc}(\Delta^{d-d'+1} \mathcal{H}, c) \tag{4}$$

holds for every hypergraph $\mathcal{H}$.

Remark 8 *The condition in Theorem 7 is only sufficient but not necessary for the validity of (4), as the following example shows: Let $c = 4$, $d \geq c$ and $d' = 3$. According to Theorem 5, we get for each hypergraph $\mathcal{H}$ that* $\operatorname{disc}(\Delta^d \mathcal{H}, c) \leq \operatorname{disc}(\Delta^{d-2} \mathcal{H}, c) = \operatorname{disc}(\Delta^{d-d'+1} \mathcal{H}, c)$. *But we have* $2!\, S(d', 2) = 6 = 3!\, S(d', 3)$ *and* $4 \nmid 6$. *This example indicates also that the proof methods of Theorem 5 and Theorem 7 are different.*

References

[1] B. Doerr and A. Srivastav, Multi-Color Discrepancies, Comb. Probab. Comput. 12(2003), 365-399.

[2] B. Doerr, A. Srivastav, and P. Wehr, Discrepancy of Cartesian products of arithmetic progressions, Electron. J. Combin. 11(2004), Research Paper 5, 16 pp.

[3] R. L. Graham, B. L. Rothschild, and J. H. Spencer, Ramsey Theory, Second Edition, Wiley, New York, USA, 1990.

[4] S. Gravier, F. Maffray, J. Renault, and N. Trotignon, Ramsey-type results on singletons, co-singletons and monotone sequences in large collections of sets, European J. Combin. 25 (2004), 719-734.