

# Non-Independent Randomized Rounding and Coloring

Benjamin Doerr<sup>†</sup>

September 30, 2003

## Abstract

We propose an advanced randomized coloring algorithm for the problem of balanced colorings of hypergraphs (discrepancy problem). Instead of independently coloring the vertices with a random color, we try to use structural information about the hypergraph in the design of the random experiment by imposing suitable dependencies. This yields colorings having smaller discrepancy. We also obtain more information about the coloring, or, conversely, we may enforce the random coloring to have special properties. There are some algorithmic advantages as well.

We apply our approach to hypergraphs of  $d$ -dimensional boxes and to finite geometries. Among others results, we gain a factor  $2^{d/2}$  decrease in the discrepancy of the boxes, and reduce the number of random bits needed to generate good colorings for the geometries down to  $O(\sqrt{n})$  (from  $n$ ). The latter also speeds up the corresponding derandomization by a factor of  $\sqrt{n}$ .

**Key words:** randomized algorithms, hypergraph coloring, discrepancy, randomized rounding, integer linear programming.

## 1 Introduction and Results

### 1.1 The Discrepancy Problem

The *combinatorial discrepancy problem* is to partition the vertex set of a given hypergraph into two classes in a balanced manner, i.e., in such a way that each

---

\*Mathematisches Seminar, Bereich II, Christian-Albrechts-Universität zu Kiel, Christian-Albrechts-Platz 4, D-24098 Kiel, Germany, [bed@numerik.uni-kiel.de](mailto:bed@numerik.uni-kiel.de)

<sup>†</sup>partially supported (associate member) by the graduate school 'Effiziente Algorithmen und Multiskalenmethoden', Deutsche Forschungsgemeinschaft

hyperedge contains roughly the same number of vertices in each of the two partition classes.

More precisely, a *hypergraph* is a pair  $\mathcal{H} = (X, \mathcal{E})$ , where  $X$  is finite set and  $\mathcal{E} \subseteq 2^X$ . The elements of  $X$  are called *vertices*, those of  $\mathcal{E}$  (*hyper*)*edges*. A partition of  $X$  into two classes is usually represented by a coloring  $\chi : X \rightarrow C$  for some two-element set  $C$ . The partition then is formed by the color classes  $\chi^{-1}(i), i \in C$ . It turns out to be useful to select  $-1$  and  $+1$  as colors. For a coloring  $\chi : X \rightarrow \{-1, +1\}$  and a hyperedge  $E \in \mathcal{E}$  the expression

$$\chi(E) := \sum_{x \in E} \chi(x)$$

then counts how many of the  $+1$ -vertices of  $E$  cannot be matched by vertices colored  $-1$ . Thus  $|\chi(E)|$  is a measure for the imbalance of  $\chi$  with respect to  $E$ . As it is our aim to color all hyperedges simultaneously in a balanced manner, we define the discrepancy of  $\chi$  with respect to  $\mathcal{H}$  by

$$\text{disc}(\mathcal{H}, \chi) := \max_{E \in \mathcal{E}} |\chi(E)|.$$

The discrepancy problem originates from the field of number theory (e. g. van der Waerden [vdW27] or Roth [Rot64]), but due to a wide range of applications and connections it has received an increased attention in computer science and applied mathematics during the last twenty years. For reasons of brevity we just mention uniformly distributed sets and numerical integration, computational geometry, communication complexity and image processing. We refer to the books Matoušek [Mat99], Chazelle [Cha00], Kushilevitz and Nisan [KN97] and the paper Asano, Katoh, Obokata and Tokuyama [AKOT02] respectively.

## 1.2 Discrepancies and Integer Linear Programs

The discrepancy problem can be formulated as integer linear program. Since we believe that our methods can be extended to this more general context, let us briefly sketch the connection. Let  $X = \{1, \dots, n\} =: [n]$  and  $\mathcal{E} = \{E_1, \dots, E_m\}$ . Then the following integer linear program (here given as a 0,1 ILP) solves the discrepancy problem for  $\mathcal{H}$ :

**minimize**  $2\lambda$   
**subject to**

$$\begin{aligned} \sum_{i \in E_j} x_i - \frac{1}{2}|E_j| &\leq \lambda, & j = 1, \dots, m, \\ - \sum_{i \in E_j} x_i + \frac{1}{2}|E_j| &\leq \lambda, & j = 1, \dots, m, \\ x_i &\in \{0, 1\}, & i = 1, \dots, n, \\ \lambda &\geq 0. \end{aligned}$$

The problem in using the linear relaxation of this ILP is that there always exists the trivial solution  $x = (x_1, \dots, x_n) = \frac{1}{2}\mathbf{1}_n$ . Therefore, a fruitful connection between solutions of the  $[0, 1]$ -relaxation and the original problem is not to be expected.

On the other hand, randomized rounding strategies for this trivial solution yield random colorings and, vice versa, generators of random colorings can be interpreted as randomized rounding strategies. Thus both problems are strongly connected. It also turns out that the tools used and the difficulties occurring in both the discrepancy problem and randomized rounding problems are very similar. Thus we believe that the methods presented in this paper might have a broader application than just the discrepancy problem. Note that due to work of Beck and Spencer [BS84] and Lovász, Spencer and Vesztergombi [LSV86], arbitrary rounding problems can be reduced to combinatorial discrepancy problems.

### 1.3 Algorithmic Aspects of Randomized Coloring and Randomized Rounding

Discrepancy is an  $NP$ -hard problem. For a very restricted class of discrepancy problems that are already  $NP$ -complete, we refer to Asano, Matsui and Tokuyama [AMT00]. Efficient algorithms finding an optimal coloring therefore are not to be expected. Indeed, very little is known about the algorithmic aspect of discrepancy. For some restrictions of the problem a nice solution exist, e. g., for hypergraphs having vertex degree at most  $t$ . Beck and Fiala [BF81] gave a polynomial time algorithm that computes colorings having discrepancy less than  $2t$ .

A common algorithmic approach for the general case (and in fact the only one known to us) are random colorings obtained by independently choosing a random color for each vertex. Via a Chernoff-bound analysis (see e.g. Alon and Spencer [AS00]) these colorings can be shown to have discrepancy  $O(\sqrt{n \log m})$  with high probability, where as above  $n$  shall always denote the number of vertices and  $m$  the number of hyperedges.

**Theorem 1.** *A random coloring obtained by independently choosing a random color for each vertex has discrepancy*

$$\text{disc}(\mathcal{H}, \chi) \leq \sqrt{2n \ln(4m)}$$

*with probability at least  $\frac{1}{2}$ .*

Note that this yields a randomized algorithm computing a coloring of the claimed discrepancy by repeatedly generating and testing a random coloring until the discrepancy guarantee of the theorem is satisfied. This algorithm has

expected run-time  $O(n(R+m))$ , where  $R$  is the complexity of generating a random bit. Random constructions show that (at least for suitable values of  $n$  and  $m$ ) there are hypergraphs having discrepancy  $\Omega(\sqrt{n \log \frac{m}{n}})$ . Thus the random colorings of Theorem 1 are almost optimal in the general case.

Via the transfer sketched in the previous subsection, all of the above holds for general rounding problems as well. In particular, no randomized rounding strategy for a linear problem of  $n$  variables and  $m$  constraints can guarantee a violation in the constraints of less than  $\Omega(\sqrt{n \log \frac{m}{n}})$ .

A central problem with random colorings (and randomized rounding) is therefore how to take into account the structure of the hypergraph (or the ILP). One way to deal with this is to use random colorings as above, but to tighten the analysis using the structural information. Exploiting such dependencies, namely positive correlation among the ‘good’ events, Srinivasan [Sri99] showed improved approximation guarantees for packing and covering problems.

A second approach is to use a different kind of random colorings, i.e., to design the random experiment in a way that it exploits the structure of the hypergraph directly. This is what we do in this paper.

## 1.4 Our Results

We analyze a way of generating random colorings not by independently coloring the vertices, but by imposing some dependencies. This allows us to exploit structural information about the hypergraph. We apply this approach on two examples, namely hypergraphs of higher-dimensional boxes and finite affine and projective planes.

Our approach proves to be effective in several ways. Firstly, it allows to generate random colorings having smaller discrepancy. Being a fairly general class of hypergraphs that have some common structure, we analyze hypergraphs of  $d$ -dimensional boxes. Our randomized colorings beat independent random colorings in terms of discrepancy by a factor of roughly  $2^{d/2}$ .

A second advantage is that we also obtain more information about the random coloring. For example, we may prescribe that our colorings should be *fair*, that is, have equal-sized color classes. This can be useful in applications. The recursive method to construct balanced multi-colorings of [DS03] for example requires fair colorings. A nice feature from the technical point of view is that we get these fair colorings without extra difficulties. Usually, working with fair colorings is more difficult, since the hypergeometric distribution is harder to analyze than the binomial one (cf. Uhlmann [Uhl66] and Chvátal [Chv79]).

A third point concerns the complexity of generating the colorings. Due to the dependencies the number of random bits needed to generate our random

colorings is smaller than for ordinary random colorings. For the hypergraphs of  $d$ -dimensional boxes we reduce the number of random bits by a factor of  $2^d$ . For the geometries, this effect is even stronger. There we reduce the number of random bits to  $O(\sqrt{n})$ , where  $n$  shall denote the number of vertices (or ‘points’ in the language of geometry). This is important if generating random bits is costly, but also admits faster derandomizations.

Finally, computing the discrepancy of our random colorings can be done faster compared to ordinary random colorings. The reason is that (depending on the hypergraph, of course) the number of relevant hyperedges, i.e., those for which  $\chi(E)$  has to be computed, is reduced. Since one approach to obtain a low-discrepancy coloring is by repeatedly generating a random coloring and then computing its discrepancy until a satisfactory solution is found, this fact also speeds up computing low-discrepancy colorings.

## 2 Structured Randomized Coloring

As mentioned in the introduction, our aim is to generate random colorings that do not independently color the vertices, but on the contrary use suitable dependencies that reflect the structure of the hypergraph. To do so, we partition the vertex set into classes. For each class, we fix a coloring that has low discrepancy in the hypergraph induced by these vertices. We obtain a coloring for all vertices by independently at random choosing for each class the associated coloring or its negative.

Thus within each class, we have perfect dependence. For vertices lying in different classes, their colors are mutually independent. The problem of this very general approach is to catch the structure of the hypergraph through the partition and the colorings for the partition classes. We show examples of how to do so in the next sections and proceed by fixing the general framework.

Let  $\mathcal{H} = (X, \mathcal{E})$  be a hypergraph and  $\mathcal{P} = \{P_1, \dots, P_r\}$  be a partition of its vertex set. For each class  $P_i$ , let  $\chi_{P_i} : P_i \rightarrow \{-1, +1\}$  be a coloring. The family  $(\chi_{P_i})_{i \in [r]}$  is *admissible* for  $\mathcal{H}$  if  $|\chi_{P_i}(E \cap P_i)| \leq 1$  holds for all  $E \in \mathcal{E}$  and all  $i \in [r]$ . For a hyperedge  $E \in \mathcal{E}$  set

$$I(\mathcal{P}, E) := \{i \in [r] \mid \chi_{P_i}(E \cap P_i) \neq 0\}.$$

Assume that the family  $\chi_{P_i}$  is admissible for  $\mathcal{H} = (X, \mathcal{E})$ . Then  $\chi_{P_i}(E \cap P_i) \in \{-1, +1\}$  for all  $i \in I(\mathcal{P}, E)$ . We generate a random coloring like this: For each  $i \in [r]$  we ‘flip a coin’, i.e., independently and uniformly choose a random sign  $\varepsilon_i \in \{-1, +1\}$ . Let  $\chi : X \rightarrow \{-1, +1\}$  denote the union of the  $\varepsilon_i \chi_{P_i}$ , that is, we have  $\chi(x) = \varepsilon_i \chi_{P_i}(x)$  for all  $x \in P_i$ . We call  $\chi$  a *structured random coloring* with respect to the colorings  $\chi_{P_i}, i \in [r]$ . We have the following:

**Lemma 1.** *Let  $\chi$  be a structured random coloring with respect to the  $\chi_{P_i}, i \in [r]$ . For any hyperedge  $E \in \mathcal{E}$  we have*

$$P(|\chi(E)| > \lambda) < 2e^{-\frac{\lambda^2}{2|I(\mathcal{P}, E)|}}.$$

*Proof.* For each  $i \in I(\mathcal{P}, E)$  define a random variable  $Z_i = \chi_{P_i}(E \cap P_i) = \sum_{x \in E \cap P_i} \chi_{P_i}(x)$ . Set  $Z = \sum_{i \in I(\mathcal{P}, E)} Z_i$ . Note that  $Z = \chi(E)$ . Since the  $Z_i$  are mutually independent  $-1, 1$  random variables, we may apply the Chernoff bound (cf. [AS00], Corollary A.1.2) and get

$$P(|Z| > \lambda) < 2e^{-\frac{\lambda^2}{2|I(\mathcal{P}, E)|}}.$$

□

Comparing Lemma 1 with the analogous estimate for ordinary random colorings

$$P(|\chi(E)| > \lambda) < 2e^{-\frac{\lambda^2}{2|E|}},$$

we see that in our version we replaced the cardinality  $|E|$  of the hyperedge by the possibly smaller number of  $P_i$  such that  $\chi_{P_i}(E \cap P_i) \neq 0$ . We thus reduced the relevant size of the hyperedges.

There is a second way structured random colorings can improve discrepancy bounds, namely by reducing the number of relevant hyperedges. Set

$$E_{\mathcal{P}} := \bigcup \{(E \cap P_i) \mid \chi_{P_i}(E \cap P_i) \neq 0\}$$

for all  $E \in \mathcal{E}$  and  $\mathcal{E}_{\mathcal{P}} := \{E_{\mathcal{P}} \mid E \in \mathcal{E}\}$ . From the definition of structured random colorings it is clear that any structured random coloring  $\chi$  with respect to the  $\chi_P, P \in \mathcal{P}$  fulfills  $\chi(E) = \chi(E_{\mathcal{P}})$ . In particular, we have

$$\text{disc}(\mathcal{H}, \chi) = \text{disc}((X, \mathcal{E}_{\mathcal{P}}), \chi).$$

Depending on the partition  $\mathcal{P}$  and the colorings  $\chi_{P_i}$ , the mapping  $E \mapsto E_{\mathcal{P}}$  is not injective, and hence  $|\mathcal{E}_{\mathcal{P}}| < |\mathcal{E}|$ . In this case we only need to consider the smaller number  $|\mathcal{E}_{\mathcal{P}}|$  of hyperedges. Since the discrepancy bound depends on the number of hyperedges just logarithmically, this effect is less important than the reduction of the relevant sizes of the hyperedges. It can however be useful, as it makes the computation of  $\text{disc}(\mathcal{H}, \chi)$  easier.

This observation together with Lemma 1 yields

**Theorem 2.** *Let  $s_0 := \max_{E \in \mathcal{E}} |I(\mathcal{P}, E)|$  and  $m_0 := |\mathcal{E}_{\mathcal{P}}|$ . Then a structured random coloring with respect to the  $\chi_P, P \in \mathcal{P}$  has discrepancy*

$$\text{disc}(\mathcal{H}, \chi) \leq \sqrt{2s_0 \ln(4m_0)}$$

*with probability at least  $\frac{1}{2}$ .*

*Proof.* Let  $\lambda = \sqrt{2s_0 \ln(4m_0)}$ . Then

$$\begin{aligned}
P(\text{disc}(\mathcal{H}, \chi) > \lambda) &= P(\text{disc}((X, \mathcal{E}_{\mathcal{P}}), \chi)) \\
&= P(\exists E \in \mathcal{E}_{\mathcal{P}} : |\chi(E)| > \lambda) \\
&\leq \sum_{E \in \mathcal{E}_{\mathcal{P}}} P(|\chi(E)| > \lambda) \\
&< \sum_{E \in \mathcal{E}_{\mathcal{P}}} 2e^{-\frac{\lambda^2}{2|I(\mathcal{P}, E)|}} \\
&\leq m_0 2e^{-\frac{\lambda^2}{2s_0}} = \frac{1}{2}.
\end{aligned}$$

□

There are few more points to add concerning structured random colorings. One is that we may get information about  $\chi$  through properties of the colorings  $\chi_P, P \in \mathcal{P}$ . For example, if each  $\chi_P, P \in \mathcal{P}$  has equal-sized color classes, then this also holds for  $\chi$ . Conversely of course, we may enforce certain properties on  $\chi$  by choosing suitable colorings  $\chi_P, P \in \mathcal{P}$ .

### Algorithmic Aspects

From the definition of structured random colorings it is clear that to generate a structured random coloring with respect to  $\chi_P, P \in \mathcal{P}$ , we need only  $|\mathcal{P}|$  random bits instead of  $n$  random bits needed for independent random colorings. This can be essential, when generating random bits is considered to be costly.

We may also completely remove the random element by using derandomization techniques. The method of conditional probabilities together with so-called pessimistic estimators can be applied also to our setting, as follows from Raghawan [Rag88].

Since the time complexity of these derandomization methods is proportional to the number of single random experiments conducted, the reduction of the number of random bits needed yields a similar speed-up for the derandomizations. We give more details about derandomizations in Section 4.

## 3 Higher-Dimensional Boxes

In this section we apply the method described above to hypergraphs of higher-dimensional boxes. They display some regularity that can be exploited, but still are fairly general.

We say that a hypergraph  $\mathcal{H} = (X, \mathcal{E})$  is a *hypergraph of  $d$ -dimensional boxes* for some  $d \in \mathbb{N}$ , if there is a decomposition  $X = X_1 \times \cdots \times X_d$  such that each hyperedge  $E \in \mathcal{E}$  has a representation  $E = E_1 \times \cdots \times E_d$  such that  $E_i \subseteq X_i$  holds for all  $i \in [d]$ . Let us agree to call any such set  $E$  a *box*.

For an arbitrary number  $r$  we denote by  $\lceil r \rceil_2$  the smallest even integer not being less than  $r$ .

### 3.1 Discrepancy Bound

For the discrepancy of box-hypergraphs, we show

**Theorem 3.** *Let  $\mathcal{H} = (X, \mathcal{E})$  be a hypergraph of  $d$ -dimensional boxes. Let  $X = X_1 \times \cdots \times X_d$  be a corresponding decomposition. Set  $n := |X|$ ,  $n_i := |X_i|$  for  $i \in [d]$ ,  $\tilde{n} := \prod_{i \in [d]} \lceil n_i \rceil_2$  and  $m := |\mathcal{E}|$ . Then there are structured random colorings  $\chi : X \rightarrow \{-1, 1\}$  having discrepancy at most*

$$\text{disc}(\mathcal{H}, \chi) \leq 2^{-\frac{d-1}{2}} \sqrt{\tilde{n} \ln(4m)}$$

*with probability at least  $\frac{1}{2}$ . Generating these structured random colorings needs  $2^{-d\tilde{n}}$  random bits.*

Note that Theorem 1 using ordinary random colorings only proves a bound of  $\sqrt{2n_1 \cdots n_d \ln(4m)} \leq \sqrt{2\tilde{n} \ln(4m)}$ . This is worse by a factor of  $2^{d/2}$  (in the case of even  $n_i$ ).

*Proof.* Without loss of generality we may assume that  $X_i = [n_i]$ . We first consider the case that all  $n_i, i \in [d]$  are even.

Set  $\mathcal{P} := \{\{2x_1 - 1, 2x_1\} \times \cdots \times \{2x_d - 1, 2x_d\} \mid \forall i \in [d] : x_i \in [\frac{n_i}{2}]\}$ , that is, we partition the vertex set into small cubes of size  $2^d$  in a canonical way.

The coloring corresponding to each small cube shall be such that adjacent (in the Hamming distance sense) corners always receive opposite colors. More formally, for a given cube  $P \in \mathcal{P}$  we define a coloring  $\chi_P : P \rightarrow \{-1, 1\}$  by

$$\chi_P(x) = 1 \iff \sum_{i \in [d]} x_i \text{ is even}$$

for all  $x = (x_1, \dots, x_d)$ .

Let  $E \in \mathcal{E}$  and  $P \in \mathcal{P}$ . As both  $E$  and  $P$  are boxes, so is  $E \cap P$ . From the definition of  $\chi_P$  we see that any subbox  $S$  of  $P$  such that  $|S| \neq 1$  fulfills  $\chi(S) = 0$ . Hence  $|\chi_P(E \cap P)| \leq 1$  for all  $E \in \mathcal{E}$  and  $P \in \mathcal{P}$ , i.e., the family  $(\chi_P)$  is admissible for  $\mathcal{H}$ . Let  $\chi$  be a structured random coloring with respect to the  $\chi_P, P \in \mathcal{P}$ .

For each  $E \in \mathcal{E}$  we have  $|I(\mathcal{P}, E)| \leq |\mathcal{P}| = 2^{-d}n$ . Applying Theorem 2 with  $s_0 = 2^{-d}n$ , we get the bound

$$\text{disc}(\mathcal{H}, \chi) \leq 2^{-\frac{d-1}{2}} \sqrt{n \ln(4m)},$$

which finishes the proof in the case that all  $n_i, i = 1, \dots, d$ , are even.

For the general case we consider the hypergraph  $\mathcal{H}_1 = ([n_1]_2 \times \dots \times [n_d]_2, \mathcal{E})$ . Since  $\mathcal{H}$  is a subhypergraph of  $\mathcal{H}_1$ , any coloring  $\chi_1$  for  $\mathcal{H}_1$  by restriction yields a coloring  $\chi = (\chi_1)|_X$  for  $\mathcal{H}$ . The claim follows from  $\text{disc}(\mathcal{H}, \chi) \leq \text{disc}(\mathcal{H}_1, \chi_1)$  and applying the case of even cardinality sets to  $\mathcal{H}_1$ .

□

In the meantime, the colorings proposed in the proof above were also used to prove relatively sharp bounds for the  $L_p$ -discrepancy of the hypergraph of all boxes in the  $[n]^d$  grid, see the result [ADLS02] of Alon, Luczak, Schoen and the author.

## 3.2 Structural Information

### 3.2.1 Geometric Boxes

Apart from this improved discrepancy bound, we also gained some information about the coloring itself. For example, all geometric boxes are colored very nicely. We call a box  $B \subseteq X$  a *geometric box*, if it can be represented in the form  $B = I_1 \times \dots \times I_d$  for some intervals  $I_i \subseteq [n_i], i \in [d]$ . As can be seen easily, these boxes fulfill  $|\chi(B)| \leq 2^d$  for any structured random coloring  $\chi$  with respect to  $\chi_P, P \in \mathcal{P}$ .

### 3.2.2 Fairness

The structured random colorings used above are *fair*, that is, they are perfectly balanced on the whole vertex set. We have  $\chi(X) = 0$ , if  $|X|$  is even, and  $\chi(X) \in \{-1, 1\}$ , if  $|X|$  is odd (note that any odd cardinality set  $S$  cannot have discrepancy  $\chi(S) = 0$ , no matter what the coloring  $\chi$  is like).

Fair colorings are important in recursive algorithms and divide-and-conquer procedures. The relation between combinatorial discrepancies and  $\varepsilon$ -approximations (and thus also the “transfer principle” connecting geometric and combinatorial discrepancies) rely on the concept of fair colorings. We refer to the first chapter of Matoušek [Mat99] for the details. Another example for the use of fair colorings is the recursive method to construct balanced multi-colorings from 2-color discrepancy information (cf. [DS03]).

### 3.2.3 Relevant Hyperedges

The structural knowledge about the random coloring can also be used to reduce the number of relevant hyperedges. To show this, we examine a special class of box hypergraphs: The hypergraph of *all*  $d$ -dimensional boxes in  $[n_0]^d$  for some  $n_0 \in \mathbb{N}$  is  $\mathcal{H}_{n_0}^d := ([n_0]^d, \{S_1 \times \cdots \times S_d \mid S_i \subseteq [n_0]\})$ . Independent random colorings  $\chi$  (Theorem 1) fulfill

$$\begin{aligned} \text{disc}(\mathcal{H}_{n_0}^d, \chi) &\leq \sqrt{2n_0^d \ln(4 \cdot 2^{n_0 d})} \\ &\approx 1.18 n_0^{\frac{d+1}{2}} \sqrt{d} (1 + o(1)) \end{aligned}$$

with probability at least  $\frac{1}{2}$ . In the following theorem we improve this bound and also show that less than  $3^{n_0 d/2}$  of the  $2^{n_0 d}$  hyperedges are relevant. For convenience let us concentrate on the case that  $n_0$  is even. The general result can be obtained from similar reasoning as in the proof of Theorem 3.

**Theorem 4.** *Let  $n_0, d \in \mathbb{N}$ ,  $n_0$  even,  $d \geq 2$  and  $n := n_0^d$ . There are structured random colorings  $\chi$  for  $\mathcal{H}_{n_0}^d$  that have*

$$\text{disc}(\mathcal{H}_{n_0}^d, \chi) \leq 1.05 \cdot 2^{-d/2} n_0^{\frac{d+1}{2}} \sqrt{d}$$

*with probability at least  $\frac{1}{2}$ . Generating these colorings needs  $2^{-d} n$  random bits. To compute their discrepancy, only  $2^{-d} 3^{n_0 d/2}$  hyperedges have to be regarded.*

*Proof.* Set  $\mathcal{P} := \{\{2x_1 - 1, 2x_1\} \times \cdots \times \{2x_d - 1, 2x_d\} \mid x_1, \dots, x_d \in [\frac{n_0}{2}]\}$  and define  $\chi_P, P \in \mathcal{P}$  as in the proof of Theorem 3. Let  $\chi$  be a random coloring with respect to  $\chi_P, P \in \mathcal{P}$ . As above we have  $|I(\mathcal{P}, E)| \leq 2^{-d} n_0^d$ .

Now let us bound the number of hyperedges that are relevant for the discrepancy of  $\chi$  with respect to  $\mathcal{H}$ . We first compute  $|\mathcal{E}_{\mathcal{P}}|$ . Let  $E = S_1 \times \cdots \times S_d$ . Assume that for some  $i \in [d]$  and  $x \in [\frac{n_0}{2}]$  we have  $\{2x - 1, 2x\} \subseteq S_i$ . Then no box  $P = \{2x_1 - 1, 2x_1\} \times \cdots \times \{2x_d - 1, 2x_d\}$  such that  $x_i = x$  intersects  $E$  in exactly one vertex. From some elementary properties of boxes and the definition of  $\chi_P$  we derive  $\chi_P(E \cap P) = 0$ . Thus  $E_{\mathcal{P}} = (S_1 \times \cdots \times (S_i \setminus \{2x - 1, 2x\}) \times \cdots \times S_d)_P$ . By induction we see that  $\pi : E \mapsto E_{\mathcal{P}}$  is a projection of  $\mathcal{E}$  onto  $\mathcal{E}_{\mathcal{P}}$ . Therefore we need to count its fixed points only to get  $|\mathcal{E}_{\mathcal{P}}|$ . We just exhibited that a necessary (and sufficient) condition for a hyperedge  $E = S_1 \times \cdots \times S_d$  to be a fixed point is

$$\forall i \in [d] \forall x \in [\frac{n_0}{2}] : |S_i \cap \{2x - 1, 2x\}| \leq 1.$$

For each  $i \in [d], x \in [\frac{n_0}{2}]$  we therefore have exactly three possibilities:  $S_i \cap \{2x - 1, 2x\}$  is empty or  $\{2x - 1\}$  or  $\{2x\}$ . This makes  $|\mathcal{E}_{\mathcal{P}}| = 3^{n_0 d/2}$  fixed points.

Still, not all hyperedges in  $\mathcal{E}_{\mathcal{P}}$  are relevant. From the structure of  $\chi$  we derive a further reduction: Note that for all  $i \in [d]$ ,

$$\gamma_i : \mathcal{E} \rightarrow \mathcal{E}; S_1 \times \cdots \times S_i \times \cdots \times S_d \mapsto S_1 \times \cdots \times ([n_0] \setminus S_i) \times \cdots \times S_d$$

is a fixed-point-free bijection of  $\mathcal{E}$  that leaves the set  $\mathcal{E}_{\mathcal{P}}$  of reduced hyperedges invariant and preserves discrepancy: We have

$$\chi(E) = -\chi(\gamma_i(E))$$

for all hyperedges  $E \in \mathcal{E}$ . In particular, the group  $\langle \gamma_1, \dots, \gamma_d \rangle \simeq \mathbb{Z}_2^d$  acts on  $\mathcal{E}$  and  $\mathcal{E}_{\mathcal{P}}$  in such a way that all orbits have length  $2^d$ . As all elements of an orbit have the same discrepancy with respect to  $\chi$ , it is enough to consider just one representative from each orbit. Let  $\mathcal{E}_0 \subseteq \mathcal{E}$  be system of representatives of the orbits in  $\mathcal{E}_{\mathcal{P}}$ , that is,  $\mathcal{E}_0$  contains exactly one element of each orbit in  $\mathcal{E}_{\mathcal{P}}$ . Since  $|\mathcal{E}_0| = 2^{-d}|\mathcal{E}_{\mathcal{P}}|$ , we reduced the number of relevant hyperedges by another factor of  $2^d$ . From Theorem 2 we finally get (with probability at least  $\frac{1}{2}$ )

$$\begin{aligned} \text{disc}(\mathcal{H}, \chi) &= \text{disc}((X, \mathcal{E}_0), \chi) \\ &\leq \sqrt{2 \cdot 2^{-d} n_0^d \ln(4 \cdot 2^{-d} 3^{n_0 d/2})} \\ &\leq 1.05 \cdot 2^{-d/2} n_0^{\frac{d+1}{2}} \sqrt{d}. \end{aligned}$$

□

We should remark that the size reduction yields a change in the order of magnitude in terms of  $d$ , namely the additional  $2^{-d/2}$  factor, whereas counting the relevant edges (less than  $(7/8)^n$  of the total number of edges already for  $d = 2$ ) only improves the constant by about 12%. Recall however, that reducing the number of relevant hyperedges does reduce the complexity of checking whether a structured random coloring fulfills the discrepancy bound of the theorem or not. The current best lower bound for the discrepancy of  $\mathcal{H}_{n_0}^d$  of  $8^{-d/2} n^{(d+1)/2}$  can be found in [ADLS02].

## 4 Finite Geometries

In this section we provide another example where structured randomization can be applied, namely finite affine and projective planes. This will not improve the discrepancy guarantee significantly, but reduce the number of random bits needed from  $n$  to  $\sqrt{n}$ . We also show how this can be used to gain a  $\sqrt{n}$ -factor speed-up for a derandomized coloring algorithm.

### 4.1 Discrepancy of Affine and Projective Planes

An *affine plane* is a hypergraph  $\mathcal{A} = (P, \mathcal{L})$  such that the following three axioms are fulfilled. Common language in geometry calls the elements of  $P$  *points*, those of  $\mathcal{L}$  *lines*. Two lines are said to be *parallel*, if they are equal or their intersection is empty.

- (i) Each two points are connected by exactly one line, i. e., for all  $p_1, p_2 \in P$  there exists a unique line  $L \in \mathcal{L}$  such that  $p_1, p_2 \in L$ .
- (ii) For each line  $L \in \mathcal{L}$  and each point  $p \in P \setminus L$  not on this line, there is exactly one line  $L'$  containing  $p$  and intersecting  $L$  trivially.
- (iii) There are three points such that no line contains all three.

It can be shown that for each affine plane  $\mathcal{A} = (P, \mathcal{L})$  there is a number  $o$  such that

- $|P| = o^2$ ,
- $|\mathcal{L}| = o(o + 1)$ ,
- each point is contained in exactly  $o + 1$  lines,
- $|L| = o$  for all line  $L \in \mathcal{L}$ ,
- each class of pairwise parallel lines has cardinality  $o$ .

We call  $o$  the *order* of  $\mathcal{A}$ .

The classical example of an affine plane is constructed over a two-dimensional vector space. Let  $V$  be a two-dimensional vector space over a field  $K$ . A one-dimensional *affine subspace*  $A$  is a translate of a one-dimensional subvectorspace  $U$ , hence,  $A = x + U$  for some  $x \in V$ . Simple calculations show that

$$(V, \{A \mid A \text{ one-dimensional affine subspace of } V\})$$

is an affine plane. There are further examples non-isomorphic to any of this type. For more information about this and other issues concerning finite geometries, we refer to the exhaustive treatment of Dembowski [Dem68].

In the following we deal with affine planes only. We shortly sketch why our results apply to projective ones as well. Each affine plane of order  $o$  can be transformed into a projective one by adding (suitable)  $o + 1$  more points and one more line. Conversely, deleting any line (and its points) from a projective plane yields an affine one. These two transformations are inverses of each other. Since no two of these extra or deleted points are contained in the same line of the affine plane, the discrepancy of the two geometries can differ by at most one.

The discrepancy of these finite geometries theoretically is well understood. A beautiful eigenvalue argument attributed to Lovász and Sós in [BS95] shows a lower bound of  $\Omega(\sqrt{o})$ . Deep results of Spencer [Spe89] and in a more general way Matoušek [Mat95] prove that colorings of discrepancy  $O(\sqrt{o})$  indeed exist. Unfortunately, these results use the partial coloring method of Beck [Bec81],

which highly depends on the pigeon principle. Thus both cannot be transformed into an efficient algorithm.

The best algorithmic solution known so far is randomized coloring using  $|P|$  random bits or a derandomization thereof. This yields colorings having discrepancy  $O(\sqrt{o \log o})$ . The objective of the following theorem is to show that there are structured randomized colorings satisfying the same discrepancy guarantee, but which can be generated by  $\sqrt{|P|}$  random bits only.

**Theorem 5.** *For each finite affine geometry of order  $o$ , there are structured random colorings that have discrepancy  $2\sqrt{o \ln o}$  with probability at least a half. They can be generated with  $o$  random bits.*

*Proof.* Let  $L_0 \in \mathcal{L}$  be any line. Denote by  $\mathcal{P} := \{L' \in \mathcal{L} \mid L_0 \cap L' = \emptyset\} \cup \{L_0\}$  the parallel class of  $L_0$ . For each  $L \in \mathcal{P}$  let  $\chi_L : L \rightarrow \{-1, +1\}$  be any coloring of  $L$  such that  $|\chi_L(L)| \leq 2\sqrt{o \ln o}$ . Of course, there is nothing wrong with taking fair colorings  $\chi_L, L \in \mathcal{P}$ . Let  $L' \in \mathcal{L} \setminus \mathcal{P}$ . By definition of an affine plane, we have  $|\chi_L(L \cap L')| = 1$ . Thus the partial colorings  $\chi_L, L \in \mathcal{P}$  are admissible for the hypergraph  $(P, \mathcal{L} \setminus \mathcal{P})$ .

Let  $\chi$  be a structured random coloring with respect to  $\chi_L, L \in \mathcal{P}$ . Applying Theorem 2 to  $\mathcal{H} = (P, \mathcal{L} \setminus \mathcal{P})$  with  $s_0 = o$  and  $m_0 = o^2$ , we get  $|\chi(L)| \leq 2\sqrt{o \ln o}$  for all  $L \in \mathcal{L} \setminus \mathcal{P}$  with probability at least a half. The lines in  $\mathcal{P}$  fulfill  $|\chi(L)| \leq 2\sqrt{o \ln o}$  by choice of the  $\chi_L, L \in \mathcal{P}$ . Hence  $\chi$  satisfies the discrepancy claim. Obviously, we needed  $|\mathcal{P}| = o$  random bits to generate  $\chi$ .  $\square$

In the proof of Theorem 5 we slightly reduced the number of relevant hyperedges from  $o(o+1)$  to  $o^2$ . Thus also our discrepancy guarantee is slightly superior to the one obtained by independent randomized coloring. Nevertheless, we chose this example to demonstrate the effect of reducing the number of random bits, which is more impressive than the discrepancy improvement.

## 4.2 Derandomizations

Having mentioned already that our approach may also be used to obtain more efficient derandomizations, let us now make this point more clear. The method of conditional probabilities roughly speaking works like this: Our random colorings emerge from a series of independent random experiments ('coin flips'). The probability  $p$  that the resulting coloring is 'good', i.e., has discrepancy at most a specific value, is positive. Let us assume that we conduct the random experiments in some order. Among all outcomes of the first random experiment, there has to be one such that with positive probability (of at least  $p$ ) the coloring resulting from prescribing this first outcome and then conducting the remaining random experiments is good. Hence, instead of conducting the first random experiment, we can also compute these conditional probabilities

for all outcomes and then choose the best result. Proceeding along the series of random experiments, we sequentially determine values for the (ex-)random variables ending up with a good coloring.

One problem with this approach is that the conditional probabilities cannot be computed efficiently in many cases. We can overcome this by using so-called *pessimistic estimators*, which are nothing else than upper bounds for the conditional probabilities that behave similar and can be computed efficiently. Raghawan [Rag88] has shown that pessimistic estimators yielding (and in fact mimicking) the bounds obtained by Chernoff-bounds exist for all randomized rounding problems. It is clear that our approach via structural randomization can be interpreted as a randomized roundings approach for a modified linear problem having  $|\mathcal{P}|$  decision variables only. Hence pessimistic estimators derandomizing our approach exist.

From the exposition above it is clear that the time complexity of derandomizations arising from the method of conditional probabilities and pessimistic estimators is proportional to the number of random experiments conducted. Hence any reduction in the number of random bits needed reduces the run time of the derandomization in the same order.

## 5 Summary and Conclusion

In this paper we presented a new way of generating random colorings for the discrepancy problem of hypergraphs. This allows to use structural information about the hypergraph and thus

- improves discrepancy guarantees,
- allows to prescribe additional properties regarding the coloring, e.g. fairness,
- reduces the number of random bits needed to generate the coloring, and thus speeds up the corresponding derandomizations,
- reduces the number of relevant hyperedges, and thus reduces the complexity of computing the discrepancy of the random coloring.

Since generating random colorings for a discrepancy problem is equivalent to generating random roundings for the trivial solution of the corresponding ILP-relaxation, we hope that these ideas can be applied to a broader range of ILPs as well.

## References

- [ADLS02] N. Alon, B. Doerr, T. Łuczak, and T. Schoen. On the discrepancy of combinatorial rectangles. *Random Structures & Algorithms*, 21:205–215, 2002.
- [AKOT02] T. Asano, N. Katoh, K. Obokata, and T. Tokuyama. Matrix rounding under the  $L_p$ -discrepancy measure and its application to digital halftoning. In *Proceedings of the 13th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 896–904, 2002.
- [AMT00] T. Asano, T. Matsui, and T. Tokuyama. Optimal roundings of sequences and matrices. *Nordic J. Comput.*, 7:241–256, 2000.
- [AS00] N. Alon and J. H. Spencer. *The Probabilistic Method*. John Wiley & Sons, Inc., 2nd edition, 2000.
- [Bec81] J. Beck. Roth’s estimate of the discrepancy of integer sequences is nearly sharp. *Combinatorica*, 1:319–325, 1981.
- [BF81] J. Beck and T. Fiala. “Integer making” theorems. *Discrete Applied Mathematics*, 3:1–8, 1981.
- [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*, pages 1405–1446. Elsevier, 1995.
- [Cha00] B. Chazelle. *The Discrepancy Method*. Cambridge University Press, 2000.
- [Chv79] V. Chvátal. The tail of the hypergeometric distribution. *Discrete Math.*, 25:285–287, 1979.
- [Dem68] P. Dembowski. *Finite Geometries*. Springer-Verlag, Berlin, 1968.
- [DS03] B. Doerr and A. Srivastav. Multicolour discrepancies. *Combinatorics, Probability and Computing*, 12:365–399, 2003.
- [KN97] E. Kushilevitz and N. Nisan. *Communication Complexity*. Cambridge University Press, 1997.
- [LSV86] L. Lovász, J. Spencer, and K. Vesztegombi. 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.

- [Rag88] P. Raghavan. Probabilistic construction of deterministic algorithms: Approximating packing integer programs. *J. Comput. Syst. Sci.*, 37:130–143, 1988.
- [Rot64] K. F. Roth. Remark concerning integer sequences. *Acta Arithmetica*, 9:257–260, 1964.
- [Spe89] J. Spencer. Coloring the projective plane. *Discrete Math.*, 73:213–220, 1989.
- [Sri99] A. Srinivasan. Improved approximation guarantees for packing and covering integer programs. *SIAM J. Comput.*, 29:648–670, 1999.
- [Uhl66] W. Uhlmann. Vergleich der hypergeometrischen mit der Binomial-Verteilung. *Metrika*, 10:145–158, 1966.
- [vdW27] B. L. van der Waerden. Beweis einer Baudetschen Vermutung. *Nieuw Arch. Wsk.*, 15:212–216, 1927.