

Lattice Approximation and Linear Discrepancy of Totally Unimodular Matrices — Extended Abstract —

Benjamin Doerr ^{*†}

Abstract

This paper shows that the lattice approximation problem for totally unimodular matrices $A \in \mathbb{R}^{m \times n}$ can be solved efficiently and optimally via a linear programming approach. The complexity of our algorithm is $\mathcal{O}(\log m)$ times the complexity of finding an extremal point of a polytope in \mathbb{R}^n described by $2(m+n)$ linear constraints.

We also consider the worst-case approximability. This quantity is usually called linear discrepancy $\text{lindisc}(A)$. For any totally unimodular $m \times n$ matrix A we show

$$\text{lindisc}(A) \leq \min\left\{1 - \frac{1}{n+1}, 1 - \frac{1}{m}\right\}.$$

This bound is sharp. It proves Spencer's conjecture $\text{lindisc}(A) \leq (1 - \frac{1}{n+1}) \text{herdisc}(A)$ for totally unimodular matrices. This seems to be the first time that linear programming is successfully used for a discrepancy problem.

1 Introduction and Results

1.1 Lattice Approximation Problem, Linear Discrepancy and Integer Linear Programs. Let $A \in \mathbb{R}^{m \times n}$ be any real matrix and $b := Ap, p \in \mathbb{R}^n$, a point of the vector space generated by the columns of A (the image of the linear mapping $p \mapsto Ap$). The lattice approximation problem is to find a point $Az, z \in \mathbb{Z}^n$, of the lattice $A\mathbb{Z}^n := \{Az | z \in \mathbb{Z}^n\}$ which is closest to b , i. e. such that $\|Az - b\|_\infty$ is minimal. Usually (and so do we) one also requires $\|p - z\|_\infty \leq 1$ to hold, that is, z evolves from p by some rounding procedure. It is easily seen that any integral part $\lfloor p \rfloor$ just carries over to z , therefore we may restrict ourselves to the case that p is in the unit cube $[0, 1]^n$. The lattice approximation problem then is to find a $0, 1$ vector z which minimizes $\|A(p - z)\|_\infty$.

For given A and p the approximation error of an optimal approximation is also called *linear discrepancy of A with respect to p*

$$\text{lindisc}(A, p) := \min_{z \in \{0, 1\}^n} \|A(p - z)\|_\infty.$$

The worst case inapproximability that can occur with the lattice generated by A is the *linear discrepancy* of A

$$\begin{aligned} \text{lindisc}(A) &:= \max_{p \in [0, 1]^n} \text{lindisc}(A, p) \\ &= \max_{p \in [0, 1]^n} \min_{z \in \{0, 1\}^n} \|A(p - z)\|_\infty. \end{aligned}$$

These notions arise from the close connection to discrepancy theory, the theory of uniform distributions and balanced partitions. Discrepancies occur in several branches of discrete mathematics and computer science, e. g. in discrete geometry and numerical integration. For reasons of space we briefly outline just the partitioning aspect and refer to the chapter of Beck and Sós [BS95] or the recent books of Matoušek [Mat99] and Chazelle [Cha00] for a deeper insight. The *discrepancy* of a matrix A is defined by

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

Discrepancy is a measure of how well the columns of A can be partitioned into two classes such that for each row the two sums of its entries in each class are nearly equal. If A is the incidence matrix of a hypergraph, the discrepancy problem turns into a purely combinatorial problem: Partitioning the vertex set into two classes such that each hyperedge is evenly split by the partition.

Immediately we see

$$\text{disc}(A) = 2 \text{lindisc}(A, \frac{1}{2}\mathbf{1}_n) \leq 2 \text{lindisc}(A).^1$$

Therefore we may view the linear discrepancy as a generalization of the discrepancy, where we have weights assigned to the columns describing the ratio in which (in average) we would like this column to belong to each of the two partition classes. Discrepancies also yield an upper bound for the linear discrepancy, see § 1.2.

A third problem in this context are *integer linear programs*. One approach is to solve their linear relaxation and then try to round the solution to a suitable

^{*}Mathematisches Seminar, Christian-Albrechts-Universität zu Kiel, Ludwig-Meyn-Str. 4, D-24098 Kiel, Germany. Email: bed@numerik.uni-kiel.de

[†]supported by the graduate school 'Effiziente Algorithmen und Multiskalenmethoden', Deutsche Forschungsgemeinschaft

¹Sometimes the linear discrepancy is defined by $\max_p \min_{\chi \in \{-1, 1\}^n} \|A(p - \chi)\|_\infty$. This makes the linear discrepancy a direct generalization of the discrepancy, but puts less emphasis on the connection to the rounding problem. This notion is larger than the one we use by a factor of 2.

integer one. The linear discrepancy bounds the extent to which an optimal rounding violates the constraints. This approach was used by Raghavan [Rag88] for several packing problems. See also Chapter 14 of [Chv83] for more on the concept of approximate solutions.

Since already the discrepancy problem for $0, 1$ matrices is NP-hard and difficult to approximate, there is no hope for a general solution to the lattice approximation problem. For some special cases, better results are available: For sparse matrices, Beck and Fiala [BF81] gave a solution with approximation error less than $\|A\|_1$, where $\|\cdot\|_1$ denotes the operator norm $\|A\|_1 := \max_{\|x\|_1=1} \|Ax\|_1 = \max\{\sum_{i \in [m]} a_{ij} \mid j \in [n]\}$ induced by the 1-norm. This result though does not give any comparison to the approximation error $\text{lindisc}(A, p)$ of an optimal approximation.

In this paper we present a complete solution for the case of totally unimodular matrices. An $m \times n$ matrix A is called *totally unimodular* if each square submatrix has determinant $-1, 0$ or 1 . In particular, $A \in \{-1, 0, 1\}^{m \times n}$. Totally unimodular matrices arise naturally in many areas (cf. also [Hof79]). For example, incidence matrices of bipartite graph are totally unimodular, as well as network matrices.

The discrepancy problem for totally unimodular matrices is well-understood. Their discrepancy is at most one. By definition, submatrices of totally unimodular matrices are totally unimodular, hence the discrepancy of all submatrices is at most one as well. The beautiful theorem of Ghouila-Houri [GH62] states that also the converse holds:

THEOREM 1.1. (GHOUILA-HOURI, 1962) *A matrix is totally unimodular if and only if each submatrix has discrepancy at most one.*

Defining the *hereditary discrepancy* $\text{herdisc}(\cdot)$ to be the maximal discrepancy among all submatrices, we have that totally unimodular matrices are exactly the ones having hereditary discrepancy at most one. Contrary to the discrepancy and the hereditary discrepancy, for the linear discrepancy of totally unimodular matrices even a sharp upper bound was missing so far, not to mention an algorithmic solution for the corresponding lattice approximation problem. This paper solves these problems via linear programming. To our knowledge, this is the first time that linear programming is used efficiently for discrepancy problems.

1.2 Previous Results. Using the result due to Lovász, Spencer and Vesztergombi [LSV86] that

$$\text{lindisc}(A) \leq \text{herdisc}(A)$$

holds for any matrix A , immediately we have $\text{lindisc}(A) \leq 1$ for a totally unimodular matrix A . Spencer conjectures that even $\text{lindisc}(A) \leq (1 - \frac{1}{n+1}) \text{herdisc}(A)$ holds for any A . This would yield $\text{lindisc}(A) \leq 1 - \frac{1}{n+1}$ for totally unimodular matrices, but Spencer's over 15 years old conjecture seems far from being proven. The current best bound is $\text{lindisc}(A) \leq (1 - \frac{1}{2m}) \text{herdisc}(A)$, cf. [Doe98].

The conjecture is backed up by the fact that Spencer [Spe87] provides an example of a matrix A such that $\text{lindisc}(A) = (1 - \frac{1}{n+1}) \text{herdisc}(A)$: Define $A \in \{0, 1\}^{(n+1) \times n}$ by $a_{ij} = 1$ if and only if $i = j$ or $i = n+1$. Set $p = \frac{1}{n+1} \mathbf{1}_n$. Then $\text{lindisc}(A, p) = 1 - \frac{1}{n+1}$. As A is totally unimodular, this example also shows that $\text{lindisc}(A) \leq 1 - \frac{1}{n+1}$ is the best possible general upper bound for the linear discrepancy of totally unimodular matrices.

For a special class of totally unimodular matrices, Peng and Yan [PY00] used a combinatorial approach. A matrix A is called *strongly unimodular*, if it is totally unimodular and if each matrix obtained from A by replacing a single non-zero entry by zero is also totally unimodular. Peng and Yan show that for a strongly unimodular $0, 1$ matrix A ,

$$\text{lindisc}(A) \leq 1 - 3^{-\frac{n+1}{2}}$$

holds. They use a decomposition lemma due to Crama, Loebl and Poljak [CLP92], which states that such a matrix is the union of incidence matrices of digraphs, if all rows contain an even number of ones. In the same work, Peng and Yan show an upper bound of $1 - \frac{1}{n+1}$ for strongly unimodular $0, 1$ matrices which have at most two non-zeros in every row.

1.3 Our Contribution. In this paper we do not follow the approach via the hereditary discrepancy, nor do we use any structure theory for totally unimodular matrices. Instead, we consider suitable linear programs and apply the theorem of Hoffman and Kruskal. This yields two types of results: A theoretical one bounding the approximation error of an optimal solution and characterizing the critical cases, and a practical one, namely an efficient algorithm solving the lattice approximation problem optimally.

We show that there is always a solution to the lattice approximation problem with error at most $1 - g(Ap)$, where $g(Ap)$ is a number of at least $\frac{1}{m}$. Using some basic linear algebra, we refine this result to

THEOREM 1.2. *Let $A \in \mathbb{R}^{m \times n}$ be a totally unimodular matrix and $p \in [0, 1]^n$. Then there is an $z \in \{0, 1\}^n$ such that*

$$\|A(p - z)\|_\infty \leq \min\{1 - \frac{1}{n+1}, 1 - \frac{1}{m}\}.$$

In particular, $\text{lindisc}(A) \leq 1 - \frac{1}{n+1}$.

This result is sharp, as the example due to Spencer proves. Theorem 1.2 shows that Spencer's conjecture $\text{lindisc}(A) \leq (1 - \frac{1}{n+1}) \text{herdisc}(A)$ holds for totally unimodular matrices. As a side product, our approach yields a characterization of all totally unimodular matrices such that $\text{lindisc}(A) = 1 - \frac{1}{n+1}$:

THEOREM 1.3. *Let A be an $m \times n$ totally unimodular matrix. Then $\text{lindisc}(A) = 1 - \frac{1}{n+1}$ holds if and only if there is a collection of $n+1$ rows of A such that each n thereof are linearly independent. If $\text{lindisc}(A, p) = 1 - \frac{1}{n+1}$ for some $p \in [0, 1]^n$, then $p_i \in \{\frac{1}{n+1}, \dots, \frac{n}{n+1}\}$ for all $i \in [n]$.*

Thus all such 'extreme' matrices contain a matrix resembling Spencer's example and possibly some additional rows which have no influence on the linear discrepancy.

This is the theoretical analysis. As mentioned we are also able to solve the lattice approximation problem optimally. This is

THEOREM 1.4. *There is an algorithm that computes for any totally unimodular $A \in \mathbb{R}^{m \times n}$ and $p \in [0, 1]^n$ an optimal solution x for the lattice approximation problem, i. e. an $x \in \{0, 1\}^n$ such that $\|A(p - x)\|_\infty = \text{lindisc}(A, p)$. The complexity of this algorithm is $\mathcal{O}(\log m)$ times the complexity of finding an extremal point of a polytope in \mathbb{R}^n described by $2(m+n)$ linear constraints or proving its emptiness.*

2 Definitions and Notation

For a real number $r \in \mathbb{R}$ write $\lfloor r \rfloor := \max\{z \in \mathbb{Z} \mid z \leq r\}$ for the largest integer not greater than r , and $\lceil r \rceil := \min\{z \in \mathbb{Z} \mid z \geq r\}$ for the smallest integer not being less than r . Set $\{r\} := r - \lfloor r \rfloor$, the fractional part of r .

Let $b \in \mathbb{R}^m$. We assume the above notation lifted to vectors in the natural way, e. g. $\lfloor b \rfloor := (\lfloor b_i \rfloor)_{i \in [m]}$. Part of our strategy will be to round those components of b to the nearest integer which are already very close to an integer. For $d \in [0, \frac{1}{2}]$ we define $I^-(b, d) := \{i \in [m] \mid \{b_i\} < d\}$, the set of indices such that b_i is less than d above the nearest integer (and hence a candidate for being rounded down), and $I^+(b, d) := \{i \in [m] \mid 1 - \{b_i\} < d\}$, the set of indices such that b_i is less than d below the nearest integer. Set $I(b, d) := I^-(b, d) \cup I^+(b, d)$. Let $r(b, d) \in \mathbb{R}^m$ denote the vector resulting from rounding the components with index in $I(b, d)$ to the nearest integer, i. e. for all $i \in [m]$ we have

$$r(b, d)_i = \begin{cases} \lfloor b_i \rfloor & \text{if } i \in I^-(b, d) \\ \lceil b_i \rceil & \text{if } i \in I^+(b, d) \\ b_i & \text{else.} \end{cases}$$

The total error of this rounding is described by

$$\begin{aligned} e(b, d) &:= \|r(b, d) - b\|_1 \\ &= \sum_{i \in I^-(b, d)} (b_i - \lfloor b_i \rfloor) + \sum_{i \in I^+(b, d)} (\lceil b_i \rceil - b_i). \end{aligned}$$

Let $g(b)$ denote the maximum value of $d \in [0, \frac{1}{2}]$ such that $e(b, d) < 1$ (the maximum exists, since $d \mapsto e(b, d)$ is left-continuous). For a matrix $A \in \mathbb{R}^{m \times n}$ set $g(A) := \max_{p \in [0, 1]^n} g(Ap)$.

LEMMA 2.1. *Let $b \in \mathbb{R}^m$ and $d \in [0, \frac{1}{2}]$. Then*

- (i) $e(b, d) < |I(b, d)| d \leq md$,
- (ii) $g(b) \geq \frac{1}{m}$.

In particular $g(A) \geq \frac{1}{m}$ holds for any $m \times n$ matrix A .

Proof. We have

$$\begin{aligned} e(b, d) &= \sum_{i \in I^-(b, d)} (b_i - \lfloor b_i \rfloor) + \sum_{i \in I^+(b, d)} (\lceil b_i \rceil - b_i) \\ &< \sum_{i \in I^-(b, d)} d + \sum_{i \in I^+(b, d)} d \\ &= |I(b, d)| d \leq md. \end{aligned}$$

In particular $e(b, \frac{1}{m}) < 1$. Thus $g(b) \geq \frac{1}{m}$ by definition.

These bounds are sharp. The vector $(d - \varepsilon)\mathbf{1}_m, \varepsilon > 0$ shows that (i) does not allow any further improvement, and $b = \frac{1}{m}\mathbf{1}_m$ is an example for $g(b) = \frac{1}{m}$.

3 Solving the Lattice Approximation Problem

In this section we present an algorithm that solves the lattice approximation problem for totally unimodular matrices efficiently and optimally.

3.1 Polyhedra. Our proof is self-contained apart from the well-known theorem of Hoffman and Kruskal [HK56]. This states that the set of feasible solutions of a linear program is an integral polyhedron, if the constraint matrix is totally unimodular and the right-side vector is integral. Hence in this case the existence of optimal solutions implies that there are also integral optimal solutions. All polyhedra in this work will be bounded and thus compact (of course everything is finite-dimensional). Hence the existence of optimal solutions is ensured if the polyhedron is non-empty.

Let A be a totally unimodular $m \times n$ matrix and $p \in [0, 1]^n$. Set $b = Ap$. For all $d \in [0, \frac{1}{2}]$ define

$$P_d := \{x \in [0, 1]^n \mid \lfloor \text{rd}(b, d) \rfloor \leq Ax \leq \lceil \text{rd}(b, d) \rceil\}.$$

P_d is an integral polyhedron by [HK56]. We first observe

LEMMA 3.1. For all $d \in [0, \frac{1}{2}[$, $x \in P_0$ we have

$$x \in P_d \iff \|b - Ax\|_\infty \leq 1 - d.$$

Proof. As a compact polyhedron is the convex hull of its extremal points we may assume x to be an extremal point of P_d . Thus x is integral. Let $i \in [m]$. Note first that $\lfloor b_i \rfloor \leq \lfloor \text{rd}(b, d)_i \rfloor$ and $\lceil \text{rd}(b, d)_i \rceil \leq \lceil b_i \rceil$. Thus $(Ax)_i$ can take at most two values, namely $\lfloor b_i \rfloor$ and $\lceil b_i \rceil$.

If $\{b_i\} \notin [0, d[\cup]1 - d, 1]$, then it does not matter which of these values is taken as $|b_i - (Ax)_i| \leq 1 - d$ holds in both cases. This is different if $\{b_i\} \in [0, d[\cup]1 - d, 1]$. Taking the wrong value would yield an approximation error of more than $1 - d$. Fortunately, $\text{rd}(b, d)_i$ is integral if $\{b_i\} \in [0, d[\cup]1 - d, 1]$ by definition. Moreover, $\text{rd}(b, d)_i$ equals the closer of the values $\lfloor b_i \rfloor$ and $\lceil b_i \rceil$. Thus we have $(Ax)_i = \text{rd}(b, d)_i$ and $|b_i - (Ax)_i| \leq \frac{1}{2} \leq 1 - d$.

The second implication is proved similarly.

3.2 The algorithm. We claim that the following algorithm solves the lattice approximation problem for totally unimodular matrices:

- (i) Set $D := \{d \in [0, \frac{1}{2}[\mid \exists i \in [m] : \{b_i\} \in \{d, 1 - d\}\} \cup \{\frac{1}{2}\}$.
- (ii) Using a binary search strategy determine the largest $d \in D$ such that $P_d \neq \emptyset$.
- (iii) Find an extremal point x of this P_d .

3.3 Correctness. Let d, x be the output of the algorithm. As P_d is integral, $x \in \{0, 1\}^n$. From Lemma 3.1 we have $\|b - Ax\|_\infty \leq 1 - d$. Let $y \in \{0, 1\}^n$ such that $\text{lindisc}(A, p) = \|b - Ay\|_\infty$, that is, y is an optimal approximation. Assume first that $l := \text{lindisc}(A, p) > \frac{1}{2}$. Then $P_{1-l} \neq \emptyset$, as $y \in P_{1-l}$ by Lemma 3.1. Since $\|b - Ay\|_\infty = l$, there is an $i \in [m]$ such that $|b_i - (Ay)_i| = l$. As Ay is integral, $b_i \in \{l, 1 - l\}$ and $1 - l \in D$. From the maximality of d we deduce $1 - l \leq d$. From $\|b - Ax\|_\infty \leq 1 - d \leq l$ and the optimality of y we conclude $\|b - Ax\|_\infty = \text{lindisc}(A, p)$.

Now let us consider the case that $\text{lindisc}(A, p) \leq \frac{1}{2}$. Then for each $i \in [m]$ we have

$$(3.1) \quad \begin{aligned} (Ay)_i = \lfloor b_i \rfloor &\iff \{b_i\} < \frac{1}{2} \\ (Ay)_i = \lceil b_i \rceil &\iff \{b_i\} > \frac{1}{2}. \end{aligned}$$

In particular, $\|b - Ay\|_\infty = \max_{i \in [m]} \min\{\{b_i\}, 1 - \{b_i\}\}$. We also find that $y \in P_{\frac{1}{2}}$. Thus $d = \frac{1}{2}$ and (3.1) holds as well with y replaced by x . Therefore we also have $\|b - Ax\|_\infty = \max_{i \in [m]} \min\{\{b_i\}, 1 - \{b_i\}\}$. Thus $\|b - Ax\|_\infty = \text{lindisc}(A, p)$ again.

This proves that x is an optimal solution of the lattice approximation problem corresponding to A and p .

3.4 Complexity. It remains to show that our algorithm is efficient. We do not want discuss any linear programming theory here and simply assume that linear programs can be solved efficiently. See any book on linear programming for a discussion of that problem. Nor do we want to discuss any problems concerned with exact number representations and complexities of elementary calculations. We therefore assume that all elementary calculations can be done in constant time with perfect accuracy.

Computing D requires m steps, namely checking whether $\{b_i\}$ or $1 - \{b_i\}$ should be included in D for each $i \in [m]$. This also shows $|D| \leq m + 1$. For the binary search we need to sort the elements of D by their size which has worst-case complexity $\mathcal{O}(m \log m)$ (cf. e. g. [CLR90]). Finally, up to $\lceil \log_2(m + 1) \rceil$ times a linear system of $2(m + n)$ constraints has to be solved to decide emptiness of the corresponding polytope and to compute the extremal point of the final P_d . Note that if the linear systems are solved using the simplex method any solution already is an extremal point of the polytope. Summarizing we see that solving the linear systems dominates the other steps of the algorithm in terms of complexity. This finally proves Theorem 1.4.

4 Approximability, Linear Discrepancy

The next lemma analyses the linear discrepancy problem for fixed A and $p \in [0, 1]^n$, that is, analyzes how bad an optimal approximation can be in the worst case. The proof is independent of the preceding section.

LEMMA 4.1. Let $A \in \mathbb{R}^{m \times n}$ be a totally unimodular matrix and $p \in [0, 1]^n$. Then

$$\text{lindisc}(A, p) \leq 1 - g(Ap).$$

In particular, $\text{lindisc}(A) \leq 1 - \frac{1}{m}$.

In the language of Section 3, Lemma 4.1 claims that $P_{g(Ap)} \neq \emptyset$, but this approach is misleading. Instead we consider the polytope P_0 together with a suitable objective function.

Proof. Set $b := Ap$. Let

$$P := \{x \in [0, 1]^n \mid \lfloor b \rfloor \leq Ax \leq \lceil b \rceil\}.$$

As A is totally unimodular, P is an integral polyhedron (this is [HK56]). Define $f : P \rightarrow \mathbb{R}$ by

$$(4.2) \quad \begin{aligned} f(x) &= \sum_{i \in I^-(b, g(b))} ((Ax)_i - \lfloor b_i \rfloor) \\ &+ \sum_{i \in I^+(b, g(b))} (\lceil b_i \rceil - (Ax)_i). \end{aligned}$$

for all $x \in P$. Thus $f(x)$ is the total error inflicted by rounding Ax in that way that was used to get $r(b, g(b))$ from b .² By definition, f is non-negative. We first show that for all $x \in P \cap \mathbb{Z}^n$

$$(4.3) \quad \forall i \in [m] \begin{cases} \{b_i\} < g(b) \Rightarrow (Ax)_i = \lfloor b_i \rfloor \\ \{b_i\} > 1 - g(b) \Rightarrow (Ax)_i = \lceil b_i \rceil \end{cases}$$

is equivalent to $f(x) < 1$.

Suppose that $f(x) < 1$ for $x \in P \cap \mathbb{Z}^n$. As A and x are integral, $f(x) \leq 0$, and since f is non-negative, $f(x) = 0$. Since all parts of the sum in (4.2) are non-negative, they are all zero. Hence $(Ax)_i = r(b, g(b))_i$ for all $i \in I(b, g(b))$. This is (4.3). On the other hand, if (4.3) is fulfilled, we have $f(x) = 0$ by (4.2).

Consider the linear optimization problem

$$\min_{x \in P} f(x).$$

p is a feasible solution and $f(p) = e(Ap, g(b)) = e(b, g(b)) < 1$. Hence there is an optimal solution x^* such that $f(x^*) < 1$. As P is integral, we may assume $x^* \in \mathbb{Z}^n$.

Let us compute $\|A(p - x^*)\|_\infty = \|b - Ax^*\|_\infty$. Let $i \in [m]$. If $i \in I^-(b, g(b))$, then $(Ax^*)_i = \lfloor b_i \rfloor$ by (4.3). Hence $|b_i - (Ax^*)_i| < g(b)$. Similarly for $i \in I^+(b, g(b))$. Thus we may assume $i \in [m] \setminus I(b, g(b))$, i. e. $b_i \in [\lfloor b_i \rfloor + g(b), \lceil b_i \rceil - g(b)]$. As $(Ax^*)_i \in \{\lfloor b_i \rfloor, \lceil b_i \rceil\}$ due to $x^* \in P$, we conclude $|b_i - (Ax^*)_i| \leq 1 - g(b)$.

The second claim follows from Lemma 2.1.

Lemma 4.1 is sharp in the worst-case, as this example due to Spencer shows: Set $m := n + 1$. Let $A \in \{0, 1\}^{m \times n}$ denote the $m \times n$ matrix with

$$a_{ij} = \begin{cases} 1 & \text{if } i = j \text{ or } i = n + 1 \\ 0 & \text{else} \end{cases}.$$

Set $p = \frac{1}{m} \mathbf{1}_n$. It is easy to see that any $z \in \{0, 1\}^n$ fulfills $\|A(p - z)\|_\infty \geq 1 - \frac{1}{m}$: If any $z_j, j \in [n]$ equals 1, then $(A(p - z))_j = p_j - z_j = \frac{1}{m} - 1$. Otherwise we have $(A(p - z))_{n+1} = n \frac{1}{m} = 1 - \frac{1}{m}$.

5 The Refinement

In this section we refine the result of the previous one and finally prove Theorem 1.2. Before doing so let us remark that a weaker bound in terms of n follows from a purely combinatorial argument. If $A \in \mathbb{R}^{m \times n}$ is totally unimodular and any two rows of A are linearly independent, then

$$(5.4) \quad m \leq \binom{n}{2} + \binom{n}{1}$$

²To be precise: We do not really round Ax . We do round p to a 0, 1 vector x . If $x \in P$, then $\|Ax - b\|_\infty < 1$, i. e. Ax looks like a rounding of b .

holds. We would show (5.4) using the connection between the VC-dimension and the primal shatter function of hypergraphs, but may-be more direct ways are possible as well. Unfortunately, (5.4) is sharp. To prove Theorem 1.2, we therefore need a different approach.

Proof. [of Theorem 1.2] Let $p \in [0, 1]^n$ and $b := Ap$. Denote the rows of A by a_1, \dots, a_m . Set

$$I := I(b, \frac{1}{n+1}) = \{i \in [m] \mid |b_i - \text{rd}(b_i, \frac{1}{2})| < \frac{1}{n+1}\}.$$

Let us call these rows ‘critical’ for the moment, because they are the ones where a rounding error of more than $1 - \frac{1}{n+1}$ can occur when using the approach of the previous section.

We proceed by showing that it is enough to consider at most n critical rows. Let $I_0 \subseteq I$ be chosen such that $\{a_i \mid i \in I_0\}$ is a basis for the vector space generated by all critical rows. In particular, $|I_0| \leq n$. Let A_0 denote the matrix obtained from A by deleting all critical rows except $a_i, i \in I_0$. Then $A_0 p = b_{|[m] \setminus I_0 \cup I_0} =: b_0$. From Lemma 2.1 and $I(b_0, \frac{1}{n+1}) = I_0$ we conclude $e(b_0, \frac{1}{n+1}) < \frac{n}{n+1}$. Hence $g(A_0 p) \geq \frac{1}{n+1}$. By Lemma 4.1 there is a $z \in \{0, 1\}^n$ such that $\|A_0(p - z)\|_\infty < 1 - \frac{1}{n+1}$. In particular, for all $i \in I_0$ we have

$$(5.5) \quad |a_i \cdot (p - z)| < \frac{1}{n+1},$$

where \cdot denotes the usual inner product on \mathbb{R}^n .

We end the proof by showing that this z also fulfills $\|A(p - z)\|_\infty \leq 1 - \frac{1}{n+1}$. Let $j \in I \setminus I_0$. As I_0 is a basis for the vector space generated by all critical rows, there are $\lambda_i, i \in I_0$ such that $a_j = \sum_{i \in I_0} \lambda_i a_i$. Since A is totally unimodular, Cramer’s rule implies $\lambda_i \in \{-1, 0, 1\}$ for all $i \in I_0$. Now

$$|a_j \cdot (p - z)| \leq \sum_{i \in I_0} |\lambda_i a_i \cdot (p - z)| < n \frac{1}{n+1}$$

by (5.5).

6 A Characterization

The proof above yields some more information which we now use for a characterization of totally unimodular matrices that have $\text{lindisc}(A) = 1 - \frac{1}{n+1}$.

Proof. [of Theorem 1.3] Let $\text{lindisc}(A) = 1 - \frac{1}{n+1}$. Choose $p \in [0, 1]^n$ such that $\text{lindisc}(A, p) = 1 - \frac{1}{n+1}$. Set $b := Ap$ and

$$I := \{i \in [m] \mid |b_i - \text{rd}(b_i, \frac{1}{2})| \leq \frac{1}{n+1}\}.$$

Note that $I = I(b, d)$ for some $d > \frac{1}{n+1}$. For any $J \subseteq [m]$ define V_J to be the vector space generated

by the rows $a_j, j \in J$. Let I_0 be a minimal subset of I such that $V_{I_0} = V_I$. In particular, the rows $a_i, i \in I_0$ form a basis of V_I . If there is an $i \in I_0$ such that $\{b_i\} \notin \{\frac{1}{n+1}, 1 - \frac{1}{n+1}\}$, or if $|I_0| < n$, then by mimicking the proof above we get a $z \in \{0, 1\}^n$ such that $\|A(p-z)\|_\infty \leq \max\{1-d, \sum_{i \in I_0} |b_i - \text{rd}(b_i, \frac{1}{2})|\} < 1 - \frac{1}{n+1}$. We conclude $|I_0| = n$ and $\{b_i\} \in \{\frac{1}{n+1}, 1 - \frac{1}{n+1}\}$ for all $i \in I_0$, and hence also for all $i \in I$. From Lemma 2.1 we get $|I| \geq n+1$ (otherwise $g(b) \geq d$, and Lemma 4.1 yields a contradiction).

From the fact that A is totally unimodular, we know that each $a_i, i \in I \setminus I_0$, can be expressed in the form $a_i = \sum_{j \in I_0} \lambda_j a_j$ with some $\lambda_j \in \{-1, 0, 1\}, j \in I_0$. Let us assume that for each $i \in I \setminus I_0$ there is such an expression $a_i = \sum_{j \in I_0} \lambda_j a_j$ such that at least one of the $\lambda_j, j \in I_0$ is zero. Then by mimicking the proof of Theorem 1.2 above (using this I_0), we find a $z \in \{0, 1\}^n$ such that $|a_i \cdot (p-z)| = \frac{1}{n+1}$ for all $i \in I_0$ and $|a_i \cdot (p-z)| \leq \frac{n-1}{n+1}$ for all $i \in I \setminus I_0$. This is again a contradiction to our choice of p . Hence there is an $i \in I \setminus I_0$ such that $a_i = \sum_{j \in I_0} \lambda_j a_j$ with some (by the way unique) $\lambda_j \in \{-1, 1\}, j \in I_0$. In particular, any n of the rows with index in $I_0 \cup \{i\}$ are linearly independent.

Let A' and b' denote the restrictions of A and b on the rows with index in I_0 . Then A' is non-singular, and thus p is already determined by $A'p = b'$. As $(n+1)b' \in \{1, n\}^n$ was shown in the first paragraph, we have $(n+1)p \in \mathbb{Z}^n$ (the inverse of a totally unimodular matrix is totally unimodular, and thus integral). Clearly, none of the $p_i, i \in [n]$ is 0 or 1 — otherwise we may just put $z_i = p_i$ reducing the dimension of the problem by one. Hence all $p_i, i \in [n]$ are in $\{\frac{1}{n+1}, \dots, \frac{n}{n+1}\}$ as claimed.

Now let A be such that there are $n+1$ rows each n thereof being linearly independent. Without loss of generality we may assume these to be the rows a_1, \dots, a_{n+1} . As above there are $\lambda_1, \dots, \lambda_n \in \{-1, 1\}$ such that $a_{n+1} = \sum_{i \in [n]} \lambda_i a_i$. Define $b' \in \mathbb{R}^n$ by

$$b'_i := \begin{cases} \frac{1}{n+1} & \text{if } \lambda_i = 1 \\ 1 - \frac{1}{n+1} & \text{else} \end{cases}$$

for all $i \in [n]$. Let A_0 denote the matrix consisting of the rows a_1, \dots, a_n only. As A_0 has full rank, the system $A_0 x = b'$ has a unique solution x . Since A_0 is totally unimodular and $(n+1)b' \in \mathbb{Z}^n$, $(n+1)x$ is also integral. Set $p = \{x\}$ and $b = Ap$. Then $\{b_i\} = \{b'_i\}$ for $i \in [n]$. We claim that any $z \in \{0, 1\}^n$ fulfills $\|A(p-z)\|_\infty \geq 1 - \frac{1}{n+1}$. Let us assume $|a_i \cdot (p-z)| < 1 - \frac{1}{n+1}$ for all $i \in [n]$ (otherwise we are done). Then $\lambda_i a_i \cdot (p-z) = \frac{1}{n+1}$ holds for all $i \in [n]$ by definition of b' . Thus $a_{n+1} \cdot (p-z) = \sum_{i \in [n]} \lambda_i a_i \cdot (p-z) = n \frac{1}{n+1}$. This proves the claim.

It is a trivial consequence of the definition of the linear discrepancy that if a matrix B consists of some rows of the matrix A , then $\text{lindisc}(B) \leq \text{lindisc}(A)$. In the light of Theorem 1.3 it makes sense to call a totally unimodular $m \times n$ matrix critical, if $m = n+1$ and $\text{lindisc}(A) = 1 - \frac{1}{n+1}$. Theorem 1.3 then states that a totally unimodular $m \times n$ matrix has linear discrepancy $1 - \frac{1}{n+1}$ if and only if it contains a critical one. The reasoning above also shows that for critical matrices A , there are just two different p such that $\text{lindisc}(A, p) = 1 - \frac{1}{n+1}$ holds, namely the one constructed, call it $p^{(1)}$, and $p^{(2)} := 1 - p^{(1)}$.

References

- [BF81] J. Beck and T. Fiala. “Integer making” theorems. *Discrete Applied Mathematics*, 3:1–8, 1981.
- [BS95] J. Beck and V. 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.
- [Chv83] V. Chvátal. *Linear Programming*. W. H. Freeman and Company, New York, 1983.
- [CLP92] Y. Crama, M. Loeb, and S. Poljak. A decomposition of strongly unimodular matrices into incidence matrices of digraphs. *Disc. Math.*, 102:143–147, 1992.
- [CLR90] T. H. Cormen, C. E. Leiserson, and R. L. Rivest. *Introduction to algorithms*. MIT Press, Cambridge, MA, 1990.
- [Doe98] B. Doerr. Linear and hereditary discrepancy. Technical Report 98–34, Berichtreihe des Mathematischen Seminars der Universität Kiel, 1998. 7 pages. Available at <http://www.numerik.uni-kiel.de/reports/1998/>. Accepted for publication in *Combinatorics, Probability and Computing*.
- [GH62] A. Ghouila-Houri. Caractérisation des matrices totalement unimodulaires. *C. R. Acad. Sci. Paris*, 254:1192–1194, 1962.
- [HK56] A. J. Hoffman and J. G. Kruskal. Integral boundary points of convex polyhedra. In H. W. Kuhn and A. W. Tucker, editors, *Linear Inequalities and Related Systems*, pages 223–246. 1956.
- [Hof79] A. J. Hoffmann. The role of unimodularity in applying linear inequalities to combinatorial theorems. *Ann. Disc. Math.*, 4:73–84, 1979.
- [LSV86] L. Lovász, J. Spencer, and K. Vesztegombi. Discrepancies of set-systems and matrices. *Europ. J. Combin.*, 7:151–160, 1986.
- [Mat99] J. Matoušek. *Geometric Discrepancy*. Springer-Verlag, Berlin, 1999.
- [PY00] H. Peng and C. H. Yan. On the discrepancy of strongly unimodular matrices. *Discrete Mathematics*, 219:223–233, 2000.

- [Rag88] P. Raghavan. Probabilistic construction of deterministic algorithms: Approximating packing integer programs. *J. Comput. Syst. Sci.*, 37:130–143, 1988.
- [Spe87] J. Spencer. *Ten Lectures on the Probabilistic Method*. SIAM, 1987.