BOUNDS FOR DISCRETE TOMOGRAPHY SOLUTIONS

BIRGIT VAN DALEN, LAJOS HAJDU, ROB TIJDEMAN

ABSTRACT. Discrete tomography concerns the reconstruction of (e.g. binary) images when the line sums in certain directions are given. First we give an explicit expression for the projection vector from the origin onto the solution hyperplane in the case of only row and column sums of a finite subset of $\mathbf{Z}^2$. Next we give an upper bound for the distance from any given real solution to the nearest integer solution. This enables us to estimate the stability of a solution. Thereafter we give an upper bound for the difference between the corresponding line sums of a given real solution and the nearest approximating integer solution where the bound depends only on the number of considered directions. Finally we generalize the first mentioned result to the continuous case.

1. INTRODUCTION

The name *discrete tomography* is due to Larry Shepp, who organized the first meeting devoted to the topic, in 1994. The interest in the subject arose from the study of atoms' positions in a crystal (see e.g. [25]), but the developed theory has not only applications in electron microscopy [20], [2], but e.g. also in medical imaging [19], [24] and in nuclear science [21], [22]. The basic problem is to reconstruct a function $f : A \to \{0, 1\}$ where A is a finite subset of $\mathbb{Z}^l$, if the sums of the function values along the lines in a finite number of directions are given.

The first results on such problems are much older. There is a vast literature on the special case where a given finite subset of $\mathbf{Z}^2$ has unknown values 0's and 1's and the problem is to find the values from the

1991 *Mathematics Subject Classification.* 94A08, 15A06.

Key words and phrases. Discrete tomography, upper bounds, approximate solutions, stability, projection vector.

Research supported in part by the Hungarian Academy of Sciences, by the OTKA grants K67580, K75566, and by the TÁMOP 4.2.1/B-09/1/KONV-2010-0007 project. This project is implemented through the New Hungary Development Plan, co-financed by the European Social Fund and the European Regional Development Fund.

given row and column sums. In 1957 Ryser [23] and Gale [15] independently derived necessary and sufficient conditions for the existence of a solution. Ryser also provided a polynomial time algorithm for finding such a solution. However, the problem is usually highly underdetermined and a large number of solutions may exist [26]. Therefore the quest is often to find a solution of a certain type. For some classes of highly-structured images, such as hv-convex polynominoes (the 1's in each row and column are contiguous) polynomial time reconstruction algorithms have been developed (see e.g. [3], [10], [11]). Batenburg [4] developed an evolutionary algorithm for finding the reconstruction which maximises an evolution function and showed that the algorithm can be successfully applied to a wide range of evolution functions. In the present paper we do not present an algorithm, but give an upper bound for the euclidean length of the shortest solution and for the euclidean distance from a given model real solution to the nearest binary solution.

Hajdu and Tijdeman [16] observed that the solution set of 0-1 solutions is precisely the set of shortest vector solutions in the set of functions $f : A \to \mathbb{Z}$ with the given line sums. They also showed that the solutions $f : A \to \mathbb{Z}$ with the given line sums form a multidimensional grid on a linear manifold which consists of all the solutions $f : A \to \mathbb{R}$ with the given line sums. Moreover, they determined the dimension of this manifold and indicated how to find a set of generators of the grid. Later they used their analysis to develop an algorithm to actually construct solutions $f : A \to \{0, 1\}$ in [17], whereafter Batenburg [5], [6] constructed much faster algorithms.

In many cases it is not necessary to know the exact original. It then suffices to have a solution or almost-solution which is guaranteed to be similar to the original. If all the solutions are similar, then they will also be similar to the original and we say that the solution set is stable. Alpers, Gritzmann and Thorens [1] showed that a small change in the data can lead to a dramatic change in the image. This research was generalized and extended by Van Dalen [12], [13], [14] in case of subsets of $\mathbf{Z}^2$ with unknown values 0's and 1's and given row and column sums. Her estimates depend on only few parameters. In this paper we use the distance estimates for the solutions to derive stability results which are applicable in the general case. They involve more complicated computations with more parameters, but, at least in the given example, yield better results. Consequences of the present approach in more complicated examples involving more than two directions can be found in [7].

The division of the paper is as follows. Section 2 contains notation and some general results from the literature. In Section 3 we present an explicit expression for the projection vector in a simple discrete case and in Section 7 we do so in a simple continuous case. Section 4 deals with distance and stability results in the binary case with several directions, and Section 5 with distance results in the case of integer values in two directions. In Section 6 we do not consider exact solutions, but approximate solutions where the line sums have only small errors. Here a result of Beck and Fiala [9] is crucial.

2. Notation and general results

We shall use the settings as in the paper [16]. Let m and n be positive integers and

$$A := \{(i,j) \in \mathbb{Z}^2 : 0 \leq i < m, 0 \leq j < n\}.$$

Let $S = \{(a_d, b_d)\}_{d=1}^k$ be a set of directions, with $a_d, b_d \in \mathbb{Z}$, $\gcd(a_d, b_d) = 1$, $a_d \geq 0$ for all $d = 1, \ldots, k$ and $b_d = 1$ if $a_d = 0$. Write $M = \sum\limits_{d=1}^k a_d$ and $N = \sum\limits_{d=1}^k |b_d|$. Assume that $M < m$ and $N < n$. We denote by $l_{d,t}$ the line sums of an $f : A \to \mathbb{R}$ along the d-th direction corresponding to the integer parameter t; i.e. along the line $a_d y = b_d x + t$. It follows from the results of [16] and from elementary linear algebra that the problem admits a unique solution $f_0 : A \to \mathbb{R}$ (i.e. having the prescribed line sums) of minimal euclidean length. We state and prove some theorems concerning this f_0 which we often consider as a point in $\mathbb{R}^{mn}$. Our method of proof provides an easy access to the explicit construction of this f_0, as well.

As one can easily check, for a given d we have $na_d + m|b_d| - a_d|b_d|$ line sums in the direction (a_d, b_d) (see also [16]). Hence altogether we have $mN + nM - \sum$ line sums, with the notation $\sum = \sum\limits_{d=1}^k a_d|b_d|$. Further, write L for the line sum column vector, which is defined as the $(mN + nM - \sum)$-tuple containing the line sums in lexicographically increasing order in (d,t). Similarly, write, for any $f : A \to \mathbb{R}$, $\vec{f}$ for the column vector which is defined as the mn-tuple formed by $f(i,j)$ arranged in lexicographically increasing order in (i,j). We often identify f and $\vec{f}$ and, for example, say that f is the shortest solution (in the euclidean sense) if we mean that $\vec{f}$ is the shortest vector such that the corresponding f is a solution.

We refer to the situation as described above as *the general case*, to the special case that $k = 2$ and the directions are $(1,0),(0,1)$, i.e. we have only row and column sums, as *the simple case*. The first result describes the structure of the solution set over the reals.

Lemma 2.1 ([16])**.** *The solution set H of the problem with zero line sums L, is a linear subspace of $\mathbb{R}^{mn}$ of dimension $(m-M)(n-N)$. Further, having m, n, S, a basis of H can be explicitly constructed.*

Note that in the above lemma the statement concerning the dimension of H is equivalent to saying that there are exactly $MN - \sum$ linearly independent dependencies among the line sums. The following result guarantees the existence of a rational solution, if the line sums are integers. We denote by f_0 the shortest solution corresponding to the line sums given by L.

Lemma 2.2 ([7])**.** *For any m, n, S there exists a rational matrix $\Lambda = (\lambda_{ij})_{\substack{i=1,\dots,mn \\ j=1,\dots,mN+nM-\sum}}$ such that for any line sum vector L we have*

$$\vec{f_0} = \Lambda \cdot L.$$

In particular, if the line sums are rational, then f_0 has rational entries and has line sum vector L.

The following result will be used in Sections 5 and 6.

Lemma 2.3 ([16])**.** *Let m, n, S be as above. If the line sums are in $\mathbb{Z}$ and there exists a solution $f : A \to \mathbb{R}$, then there exist infinitely many solutions $f : A \to \mathbb{Z}$.*

3. Explicit expression for projection in the simple case

In the particular case $S = \{(1,0),(0,1)\}$, i.e. having only row and column sums, we give the explicit form of f_0. For this we simplify our notation. Let c_i $(i = 0, \dots, m-1)$ and r_j $(j = 0, \dots, n-1)$ denote the column sums and row sums, respectively. Further, write $D = \sum_{j=0}^{n-1} r_j$. Note that we have $D = \sum_{i=0}^{m-1} c_i$.

Theorem 3.1. *For any $(i,j) \in A$ we have*

$$f_0(i,j) = \frac{c_i}{n} + \frac{r_j}{m} - \frac{D}{mn}.$$

Proof. To prove the statement, we need to check two properties: that f_0 is a solution, and that f_0 is orthogonal to H.

We start with the first property. Obviously, for any $i = 0, \dots, m-1$ we have

$$\sum_{j=0}^{n-1} f_0(i,j) = \sum_{j=0}^{n-1} \left(\frac{c_i}{n} + \frac{r_j}{m} - \frac{D}{mn} \right) = c_i - \frac{D}{m} + \frac{1}{m} \sum_{j=0}^{n-1} r_j = c_i.$$

Similarly, for any $j = 0, \dots, n-1$,

$$\sum_{i=0}^{m-1} f_0(i,j) = \sum_{i=0}^{m-1} \left(\frac{c_i}{n} + \frac{r_j}{m} - \frac{D}{mn} \right) = r_j - \frac{D}{n} + \frac{1}{n} \sum_{i=0}^{m-1} c_i = r_j,$$

which confirms the first property.

To prove the second property we check orthogonality for arbitrary $\vec{h} \in H$. Since for any $\vec{h} \in H$ with the corresponding $h : A \to \mathbb{R}$

$$\sum_{j=0}^{n-1} h(i,j) = 0 \quad (i = 0, \dots, m-1)$$

and

$$\sum_{i=0}^{m-1} h(i,j) = 0 \quad (j = 0, \dots, n-1),$$

we have

$$\sum_{i=0}^{m-1} \sum_{j=0}^{n-1} f_0(i,j) h(i,j) = \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} \left(\frac{c_i}{n} + \frac{r_j}{m} - \frac{D}{mn} \right) h(i,j) =$$

$$= \sum_{i=0}^{m-1} \left(\frac{c_i}{n} - \frac{D}{mn} \right) \sum_{j=0}^{n-1} h(i,j) + \sum_{j=0}^{n-1} \frac{r_j}{m} \sum_{i=0}^{m-1} h(i,j) = 0,$$

and the theorem follows. □

4. The pixel values are in $\{0, 1\}$

In this section we consider the general binary case $f : A \to \{0, 1\}$. Write $D := \sum_t l_{d,t}$ for the sum of the 'total line sums' in an arbitrarily chosen direction $(a_d, |b_d|)$. Observe that D is independent of the choice of d. We present a method to give lower bounds for the number of correct pixel values in an approximate solution and to give upper bounds for the number of entries where two solutions can be different. Such results have been obtained in the simple case by Van Dalen [13] by a completely different method. She calls an m by n array F_0 uniquely determined if there is no other m by n array with entries 0 and 1 and the same row and column sums. Let $\alpha(F)$ be half of the sum of the absolute differences between the row sums of F and the row sums of a uniquely determined array F_0 with the same row sums as F. Then Van Dalen derived upper bounds for the number of places where two solutions F_1 and F_2 can differ in terms of $m, n, D = D(F_1)$ and $\alpha(F_1)$. We shall derive some estimates and then compare our results with those in [13].

Theorem 4.1. *For any solution $g : A \to \{0,1\}$ we have*

$$|\vec{g} - \vec{f_0}| = \sqrt{D - |\vec{f_0}|^2}.$$

In other words, such solutions $\vec{g}$ are on a hypersphere in the linear solution manifold, with center $\vec{f_0}$ and radius $\sqrt{D - |\vec{f_0}|^2}$.

Proof. Observe that if g is any binary solution then we have $|\vec{g}| = \sqrt{D}$. This means that such solutions are situated on a hypersphere with the origin as center, and of radius $\sqrt{D}$. According to Lemma 3.1 the solutions $\vec{g}$ are located on a linear manifold of dimension $(m - M)(n - N)$ orthogonal to $\vec{f_0}$. The intersection of this manifold and the hypersphere is a hypersphere (of the appropriate dimension) having $\vec{f_0}$ as center. By the theorem of Pythagoras we get that the radius of this hypersphere is $\sqrt{D - |\vec{f_0}|^2}$, and the theorem follows. □

Put $\langle x \rangle = \min(|x|, |x-1|)$ and $E = \sum_{i=1}^{m} \sum_{j=1}^{n} \langle f_0(i,j) \rangle^2$. Then the euclidean distance between $\vec{f_0}$ and the nearest integer vector with entries in $\{0,1\}$ is exactly $\sqrt{E}$. Hence we have the following consequence of Theorem 4.1.

Corollary 4.1. *If $E + |\vec{f_0}|^2 > D$, then there is no solution $g : A \to \{0,1\}$.*
If $E + |\vec{f_0}|^2 = D$, then the only solutions $g : A \to \{0,1\}$ are obtained by rounding $f_0(i,j)$ to the integer 0 or 1 which is nearest to it.

Note that the only entries where the rounding is not unique are those with value 1/2. If such entries do not exist, the solution is unique.

If $D - E - |f_0|^2 > 0$, but not too large, we still may conclude that a certain fraction of the rounded values agrees with any solution $g : A \to \{0,1\}$. (In most cases we cannot tell which rounded values are correct and it may even be impossible to do so.) Suppose the rounded value $F(i,j) \in \{0,1\}$ of $f_0(i,j)$ is not the right value. If $x := f_0(i,j) \geq 1/2$, then we have to replace $\langle f_0(i,j) \rangle^2 = (x-1)^2$ by x^2. Hence the contribution increases by $2x - 1 = 2f_0(i,j) - 1$. Similarly, if $f_0(i,j) \leq$

1/2, then the contribution increases by $1 - 2x = 1 - 2f_0(i, j)$. Order the values $|2f_0(i, j) - 1|$ in nondecreasing order, $b_1, b_2, \ldots, b_{mn}$, say. According to Theorem 4.1 $D - |f_0|^2$ equals E plus the sum of the values b_i which correspond to wrong values in F. Let s be the value with

$$(1) \qquad b_1 + \cdots + b_s \leq D - E - |\vec{f_0}|^2 < b_1 + \cdots + b_{s+1}.$$

Then at most s pixels can have wrong values. Therefore at least $mn - s$ pixels have right values. Similarly, let t be the value with

$$(2) \qquad b_1 + \cdots + b_t \leq 2(D - E - |\vec{f_0}|^2) < b_1 + \cdots + b_{t+1}.$$

Then for any two solutions at most t corresponding pairs of pixels can have different values. Therefore at least $mn - t$ such pixels have the same values. So we have derived the following result.

Theorem 4.2. (a) *Let s be defined as above. For any solution $g : A \to \{0, 1\}$ we have $g(i, j) = F(i, j)$ for at least $mn - s$ pairs (i, j) $(i = 1, \ldots, m; j = 1, \ldots, n)$.*
(b) *Let t be defined as above. For any two solutions $g_1, g_2 : A \to \{0, 1\}$ we have $g_1(i, j) = g_2(i, j)$ for at least $mn - t$ pairs (i, j) $(i = 1, \ldots, m; j = 1, \ldots, n)$.*

This result can be slightly improved if there are line sums over the $F(i, j)$ which do not agree with the corresponding sum over the $f(i, j)$. Consider some direction with such line sums. We know that some values of F have to be wrong and can therefore increase the value of E by securing that among $b_1, ..., b_s$ and $b_1, ..., b_t$ in (1) and (2), respectively, there are not too many representatives from each row. This may yield lower value of s and t, hence better results.

Example 4.1. Let $m = 6, n = 5$. Let the row sums be given by $5, 4, 3, 2, 1$ and the column sums by $4, 4, 3, 2, 1, 1$, respectively. Then $D = 15$ and, according to Theorem 3.1, $30f_0$ is given by

34	34	28	22	16	16
29	29	23	17	11	11
24	24	18	12	6	6
19	19	13	7	1	1
14	14	8	2	−4	−4

Thus $|f_0|^2 = 166/15$. Further we get the following table F by rounding.

1	1	1	1	1	1
1	1	1	1	0	0
1	1	1	0	0	0
1	1	0	0	0	0
0	0	0	0	0	0

A calculation gives $E = 13/5$, hence $D - |\vec{f_0}|^2 - E = 4/3$. From this we compute $s = 9, t = 13$. Thus, by Theorem 4.2, we conclude that every solution differs at most at 9 places from F and that any two solutions differ from each other at most at 13 places. In fact, the solutions below, due to Van Dalen [13], show that the actual numbers s and t can be as large as 8 and 10, respectively.

1	1	1	1	0	1
1	1	1	0	1	0
1	1	0	1	0	0
1	0	1	0	0	0
0	1	0	0	0	0

1	1	1	1	1	0
1	1	1	1	0	0
1	1	1	0	0	0
1	1	0	0	0	0
0	0	0	0	0	1

Theorem 2 of [13] implies that two solutions can differ at most at 20 places. This estimate is worse than the 13 places which follows from the new estimate. The present estimation can be applied for any set of directions. The estimate of Van Dalen depends only on few parameters, but holds only for the simple case.

Since in the example F does not satisfy the required row sums, an improvement of the value of E is possible. The column sums are correct, but the top row sum is 1 too high and the bottom row sum 1 too low. Hence elements have to be pushed downwards from top to bottom either in one stroke or in steps as in the above examples from [13]. In this special example (the vertical difference of consecutive elements in the first table is constant) it does not matter in which way the lowering is carried out, provided that the final line sums are correct. It follows that $s = 8$ (as in the above examples from [13]) is optimal.

The example can be generalized as follows. Let n, q be positive integers and set $m = (n+1)q$. Let the row sums be given by $c_j = n - 1$ for $1 \leq j \leq q, c_{lq+j} = n - l$ for $1 \leq l \leq n, 1 \leq j \leq q$, and $c_j = 1$ for $nq + 1 \leq j \leq (n+1)q$. In [13] Van Dalen showed that $t \leq 2q\sqrt{4nq(n+1)+1} - 2q$ and that $t = 2nq$ can be reached. After an elaborate computation the argument in the present paper yields the estimates $s < 2nq$ and $t < 4nq - 2q$. Therefore we have obtained an improvement by roughly a factor $\sqrt{q}$ compared to the estimate for t from

[13]. Moreover, the present bound for t cannot be further improved by a factor 2.

5. The pixel values in the simple case are in $\mathbb{Z}$

In this section we consider the case that we have only row and column sums for functions $f : A \to \mathbb{Z}$ (but not necessarily in $\{0, 1\}$). We give an upper bound for the solution closest to some given function $t : A \to \mathbb{R}$ in the euclidean sense and further an upper bound for the euclidean length of the shortest integer solution which cannot be improved by a factor > 2. The function t can be considered as a prescribed model for the ideal integer solution f satisfying the row and column sums.

Theorem 5.1. *Let $t : A \to \mathbb{R}$ have integer row and column sums. There exists a function $f : A \to \mathbb{Z}$ with the same row and column sums and satisfying*

$$|\vec{f} - \vec{t}| \leq \sqrt{mn}.$$

In the proof we shall use the following lemma.

Lemma 5.1. *Let be given a $d-$dimensional parallelopiped whose edges have lengths $l_1, l_2, \ldots, l_d$. Then the distance from any interior point of the parallelopiped to the nearest vertex is at most $\sqrt{l_1^2 + l_2^2 + \cdots + l_d^2}/2$.*

Note that the bound is the best possible and that it is attained by the centre point of a hyperblock with side lengths $l_1, l_2, \ldots, l_d$.

Proof. By induction on d. We indicate the directions corresponding to the sides of lengths $l_1, l_2, \ldots, l_d$ by $1, 2, \ldots, d$, respectively.

For $d = 2$: consider a point P inside the parallelogram. Project P parallel to direction 2 onto both parallel sides into direction 1. Then the sum of the distances from P to the two projections equals l_2. So at least one of the distances is at most $l_2/2$. Let Q be a point on one of the sides into direction 1 that has minimal distance to P. Then the distance from P to Q is at most $l_2/2$. If Q is a vertex of the parallelogram, then the distance from P to the nearest vertex is at most $\sqrt{l_1^2 + l_2^2}/2$ indeed. Otherwise Q is a point on a side in the direction 1 and the line segment PQ is orthogonal to direction 1. Observe that the distance from Q to the nearest vertex on that side is at most $l_1/2$. Now apply Pythagoras' theorem to find that the distance from P to this vertex is at most $\sqrt{l_1^2 + l_2^2}/2$.

For larger d: project P parallel to direction d onto both parallel $(d-1)$-dimensional parallelopipeds containing the edges into the directions $1, \ldots, d-1$. Since the sum of the distances from P to the two projections equals l_d, one of these distances is at most $l_d/2$. Let Q be a point in one of both parallel parallelopipeds of dimension $d-1$ with minimal distance to P. For $i = 1, \ldots, d-1$ consider the intersection of the parallelopiped under consideration containing Q and the line through Q into the direction i. If Q is one of both end points of this line segment, then we consider in the sequel the subparallelopipiped through Q of one dimension lower by projecting parallel to direction i. If Q is an interior point of this line segment, then the line segment PQ is orthogonal to the direction i. Let I be the set of i's of the latter type. We then are left with a parallelopiped of dimension $|I|$ containing Q such that PQ is orthogonal to direction i for $i \in I$. By the induction hypothesis, applied to this parallelopiped of dimension $|I|$, the distance from Q to the nearest vertex is at most $\sqrt{\sum_{i \in I} d_i^2}/2$. By Pythagoras' theorem the distance from P to the nearest vertex is at most

$$\frac{1}{2}\sqrt{l_d^2 + \sum_{i \in I} l_i^2} \leq \frac{1}{2}\sqrt{\sum_{i=1}^{d} l_i^2}$$

. □

Proof of Theorem 5.1. According to Lemma 3.1 there exist infinitely many integer solutions which are located in a linear manifold of dimension $(m-1)(n-1)$. More precisely, according to [16], they form a lattice in this manifold generated by the $(m-1)(n-1)$ vectors in the mn-dimensional space corresponding with the m by n rectangles $A^{i,j}$ $(i = 1, \ldots, m-1; j = 1, \ldots, n-1)$ given by $A^{i,j}(i,j) = 1, A^{i,j}(i, j+1) = -1, A^{i,j}(i+1, j) = -1, A^{i,j}(i+1, j+1) = 1$ and $A^{i,j}(k,l) = 0$ for other values of k, l. Each of these vectors has euclidean length 2. The $(m-1)(n-1)$ vectors generate a (closed) parallelopiped such that the union of all the parallelopideds which arise by shifting over a lattice vector cover the mn-dimensional real space. Therefore the vector $\vec{f_0}$ is in one of these shifted parallelopipeds. It follows from Lemma 5.1 that the distance to its nearest lattice point is at most $\sqrt{(m-1)(n-1)}$. Since f_0 is the shortest rational solution and the shortest integer solution is the solution nearest to f_0, this proves the theorem. □

If there exists a solution f with all entries in $\{0, 1\}$, then this solution is the shortest among all integer solutions, since

$$D = \sum_{i=1}^{m}\sum_{j=1}^{n} f(i,j) \leq \sum_{i=1}^{m}\sum_{j=1}^{n} f(i,j)^2$$

and equality holds if and only if $f : A \to \{0, 1\}$. If there does not exist such a solution, it is natural to ask for the shortest solution among all the integer solutions in the euclidean sense. Here we give an upper bound for the euclidean length of the shortest integer solution.

Corollary 5.1. *Suppose $f : A \to \mathbb{Z}$ has integer row and column sums. Let $\vec{f_0}$ be the projection vector. Let $g : A \to \mathbb{Z}$ be the shortest integer solution in the euclidean sense satisfying the row and column sums. Then*

$$|\vec{g} - \vec{f_0}| \leq \sqrt{mn}$$

Moreover, this result cannot be improved by a factor greater than 2.

Proof. Apply Theorem 5.1 with $\vec{t} = \vec{f_0}$. This yields a solution $g_0 : A \to \mathbb{Z}$ with $|\vec{g_0} - \vec{f_0}| \leq \sqrt{mn}$. Since the shortest integer solution is the integer solution which is nearest to f_0 in the euclidean sense, we conclude that the shortest integer solution $g : A \to \mathbb{Z}$ satisfies $|\vec{g} - \vec{f_0}| \leq \sqrt{mn}$.

Let m and n be even. Consider $f : A \to \{0, 1\}$ given by $A(i, j) = 1$ if $i+j$ is even and $A(i, j) = 0$ if $i+j$ is odd. Then $\vec{f_0} = (0.5, 0.5, \ldots, 0.5)$. Therefore the distance from $\vec{f_0}$ to the nearest integer vector is $\sqrt{mn}/2$. This proves the second assertion. □

6. Approximate solutions in the general case

Until now we have studied exact solutions. In this section we assume that a function $t : A \to \mathbb{R}$ or its line sums are given and derive upper bounds for the maximal difference between the line sums of t and the 'nearest' function $f : A \to \mathbb{Z}$ where we may also restrict the range of f further. A crucial tool in the proofs will be the following result of Beck and Fiala (see [8], p.244, formula (35)). A small improvement for large $\deg(Z)$ has been obtained by Helm [18].

Lemma 6.1. *Let $X = \{x_1, x_2, \ldots, x_p\}$ be an arbitrary finite set and $Z = \{Y_1, Y_2, \ldots, Y_q\}$ an arbitrary family of subsets of X. Let $\deg(Z)$ be the maximum degree of Z, i.e.*

$$\deg(Z) = \max_{x \in X} \sharp\{Y \in Z : x \in Y\}.$$

Let the linear discrepancy of Z be given by

$$\text{lindis}(Z) = \max_{\alpha_1,\ldots,\alpha_p} \min_{a_1,\ldots,a_p} \max_{Y \in Z} |\sum_{x_i \in Y} (a_i - \alpha_i)|,$$

where $\alpha_1, \ldots, \alpha_p \in [0, 1]$ and $a_1, \ldots, a_p \in \{0, 1\}$. Then

$$\text{lindis}(Z)) < \deg(Z).$$

As an immediate consequence we prove the following theorem.

Theorem 6.1. *Let $t : A \to [0,1]$ and k directions S be given. Then there exists a function $f : A \to \{0,1\}$ such that each difference between corresponding line sums of t and f in the directions S is less than k.*

Proof. Let X be the entries of A and $Y_1, \dots, Y_q$ the subsets of A which determine the line sums. Then $\deg(Z) = k$. By Lemma 6.1 there exist integers $a_1, \dots, a_p \in \{0,1\}$ such that $\text{lindis}(Z) < k$. Observe that $|\sum_{x_i \in Y}(a_i - \alpha_i)|$ is the difference between the line sums of t and f along the line Y. Hence the maximal difference over all line sums is less than k. □

The crucial point of the theorem is that the upper bound is independent of m and n, hence of the size of A. It is likely that the dependence on k can be considerably improved (cf. the remark at the bottom of p. 242 of [8]).

The situation becomes somewhat more complicated if only the line sums are given and not the entries of the model function t. Nevertheless, the following result on grey values is a consequence of Theorem 6.1.

Corollary 6.1. *Let r be a positive integer and $R := \{0, 1, \dots, r\}$. Let $t : A \to [0, r]$ be an unknown function with given line sums in the k directions S. Then there exists a function $f : A \to R$ such that each difference between corresponding line sums of t and f in the directions S is less than k.*

Proof. Subtract the integer part $\lfloor t_{i,j} \rfloor$ from each entry $t_{i,j}$ of t. This yields a function $t^* : A \to [0,1]$ to which we apply Theorem 6.1. Thus there exists a function $f^* : A \to \{0,1\}$ such that the corresponding line sums of t^* and f^* in the directions S are less than k. Add $\lfloor t_{i,j} \rfloor$ to each entry $f^*_{i,j}$ of f^* to obtain a function $f : A \to R$. Then the difference between the corresponding line sums of t and f in the directions S is at most k too. □

At least in principle it is possible for every practical situation to determine whether there exists a real function $t : A \to [0, r]$ satisfying the line sums. First the linear manifold of all real solutions mentioned in Theorem 2.2 can be computed by using the method described in [16]. Subsequently the orthogonal projection $\vec{f_0}$ of the origin onto the solution manifold may be determined. If this vector is within the hyperblock R, then it can serve as function t in the Corollary. Otherwise it has to be checked whether the projection of f_0 in the solution manifold onto the hyperfaces of R has a point in common. If such a point exists, then it serves as a function t. Otherwise such a function t does not exist.

7. The continuous version in the simple case

Finally we prove an analogue of Theorem 3.1.

Let m and n be positive real numbers and let A be a Lebesgue-measurable subset of $T = [0, m] \times [0, n]$. Write $f_A(x, y)$ for the characteristic function of A inside T.

For any bounded Lebesgue-measurable function $f : T \to \mathbb{R}$, let $c_f(x) : [0, m] \to \mathbb{R}$ and $r_f(y) : [0, n] \to \mathbb{R}$ be the column integrals and row integrals of $f(x, y)$, that is

$$c_f(x) = \int_0^n f(x, y)dy \quad (x \in [0, m])$$

and

$$r_f(y) = \int_0^m f(x, y)dx \quad (y \in [0, n])),$$

respectively, where integration is always meant in the Lebesgue sense. Since $f(x, y)$ is bounded, by the theorem of Fubini we know that these functions exist. Note that the same is true for $c_{f_A}(x)$ and $r_{f_A}(y)$.

Let L denote the set of bounded Lebesgue-integrable functions $T \to \mathbb{R}$ having column integrals $c_{f_A}(x)$ $(x \in [0, m])$ and row integrals $r_{f_A}(y)$ $(y \in [0, n])$. Further, write H for the set of bounded Lebesgue-integrable functions $T \to \mathbb{R}$ having vanishing row integrals and column integrals. Observe that H is a closed linear subspace of the linear space $\mathcal{L}$ of bounded integrable functions $T \to \mathbb{R}$. Further, for any $g_1, g_2 \in L$ we obviously have $g_1 - g_2 \in H$. In other words, $L = g + H$ with any $g \in L$.

We recall that the well known inner product in $\mathcal{L}$ is given by

$$\langle f(x, y), g(x, y)\rangle = \int\int_T f(x, y)g(x, y)dxdy$$

for $f, g \in \mathcal{L}$. The following theorem describes the shortest element in L, with respect the usual norm

$$||f(x, y)|| = \sqrt{\langle f, f\rangle} = \left(\int\int_T f^2(x, y)dxdy\right)^{1/2}$$

for $f \in \mathcal{L}$.

Theorem 7.1. *The shortest element in L exists, and is given by*

$$f_0(x, y) = \frac{c_{f_A}(x)}{n} + \frac{r_{f_A}(y)}{m} - \frac{\lambda(A)}{mn} \qquad ((x, y) \in T),$$

where $\lambda(A)$ is the Lebesgue-measure of A.

Proof of Theorem 7.1. Since H is a closed linear subspace of $\mathcal{L}$ and L is just a shift of H, L has a shortest element $f_0(x, y)$ indeed. This $f_0(x, y)$ is uniquely determined by the following two properties:

- $f_0(x, y)$ has column integrals $c_{f_A}(x)$ $(x \in [0, m])$ and row integrals $r_{f_A}(y)$ $(y \in [0, n])$,
- $f_0(x, y)$ is orthogonal to H, i.e. $\langle f_0(x, y), h(x, y)\rangle = 0$ for every $h(x, y) \in H$.

We prove that the choice for $f_0(x, y)$ in the statement meets these requirements. To prove the first property, observe that

$$\int_0^n r_{f_A}(y)dy = \int\int_T f_A(x, y)dxdy = \lambda(A).$$

Thus for any $x \in [0, m]$ we have

$$\int_0^n \left(\frac{c_{f_A}(x)}{n} + \frac{r_{f_A}(y)}{m} - \frac{\lambda(A)}{mn}\right) dy = c_{f_A}(x) + \frac{\lambda(A)}{m} - n\frac{\lambda(A)}{mn} = c_{f_A}(x).$$

Similarly, by

$$\int_0^m c_{f_A}(x)dx = \int\int_T f_A(x, y)dxdy = \lambda(A)$$

for any $y \in [0, n]$

$$\int_0^m \left(\frac{c_{f_A}(x)}{n} + \frac{r_{f_A}(y)}{m} - \frac{\lambda(A)}{mn}\right) dx = \frac{\lambda(A)}{n} + r_{f_A}(y) - m\frac{\lambda(A)}{mn} = r_{f_A}(y),$$

which proves the first property.

In order to check the second property, take an arbitrary $h(x, y) \in H$. Then for any $x \in [0, m]$ and $y \in [0, n]$ we have

$$c_h(x) = \int_0^n h(x, y)dy = 0$$

and

$$r_h(y) = \int_0^m h(x, y)dx = 0,$$

respectively. Hence

$$\left\langle h(x, y), \frac{c_{f_A}(x)}{n} + \frac{r_{f_A}(y)}{m} - \frac{\lambda(A)}{mn}\right\rangle =$$

$$= \int_0^n \int_0^m h(x,y) \left(\frac{c_{f_A}(x)}{n} + \frac{r_{f_A}(y)}{m} - \frac{\lambda(A)}{mn} \right) dxdy =$$

$$\int_0^m \frac{c_{f_A}(x)}{n} \left(\int_0^n h(x,y)dy \right) dx +$$

$$+ \int_0^n \left(\frac{r_{f_A}(y)}{m} - \frac{\lambda(A)}{mn} \right) \left(\int_0^m h(x,y)dx \right) dy = 0$$

as the inner integrals are 0. This proves the second property, and the theorem follows. □

References

[1] A. Alpers, P. Gritzmann, L. Thorens, Stability and instability in discrete tomography, in G. Bertrand *et al.* (Eds.), *Digital and Image Geometry*, LNCS 2243 (2001), 175-186.

[2] K.J. Batenburg

[3] E. Barucci, A. Del Lungo, M. Nivat, R. Pinzani, Reconstructing convex polynominoes from horizontal and vertical projections, *Theoret. Comput. Sci.* **155** (1996), 321-347.

[4] K.J. Batenburg, An evolutionary algorithm for discrete tomography, *Discr. Appl. Math.* **151**(2005), 36-54.

[5] K.J. Batenburg, A network flow algorithm for binary image reconstruction from few projections, in A. Kuba, L.G. Nyul, K. Palágyi, *DGCI 2006*, LNCS 4245 (2006), 86-97.

[6] K.J. Batenburg, Network flow algorithms for discrete tomography, in G.T. Herman, A. Kuba (Eds.) *Advances in Discrete Tomography and its Applications* Birkhäuser, 2007, pp. 175-205.

[7] K.J. Batenburg, W. Fortes, L. Hajdu, R. Tijdeman, *Bounds on the difference between reconstructions in binary tomography,* to appear.

[8] J. Beck, W.W.L. Chen, *Irregularities of distribution*, Cambridge University Press, 1987.

[9] J. Beck, T. Fiala, Integer-making theorems, *Discr. Appl. Math.* **3** (1981), 1-8.

[10] S. Brunetti, A. Del Lungo, F. Del Ristoro, A. Kuba, M. Nivat, Reconstruction of 4- and 8-connected convex discrete sets from row and column projections, *Linear Algebra Appl.* **339** (2001), 37-57.

[11] G. Dahl, T. Flatberg, Optimization and reconstruction of hv-convex (0,1)-matrices, in: A. Del Lungo, V. Di Gesy, A. Kuba (Eds.), *Electronic Notes Discr. Math.* **12** (2003), 58-69.

[12] B. Van Dalen, Stability results for uniquely determined sets in two directions, *Discr. Math.* **309** (2009), 3905-3916.

[13] B. Van Dalen, On the difference between solutions of discrete tomography problems, *J. Combinatorics Number Th.* **1** (2009), 15-29.

[14] B. Van Dalen, On the difference between solutions of discrete tomography problems II, *Pure Math. Appl.* **20** (2009), 103-112.

[15] D. Gale, A theorem on flows in networks, *Pacific J. Math.* **7** (1957), 1073-1082.
[16] L. Hajdu and R. Tijdeman, Algebraic aspects of discrete tomography, *J. Reine Angew. Math.* **534** (2001), 119–128.
[17] L. Hajdu and R. Tijdeman, An algorithm for discrete tomography, *Lin. Algebra Appl.* **339** (2001), 147-169.
[18] M. Helm, On the Beck-Fiala theorem, *Discr. Math.* **207** (1999), 73-87.
[19] G.T. Herman, A. Kuba, Discrete tomography in medical imaging, *Proc. IEEE* **91** (2003), 1612-1626.
[20] J.R. Jinschek, K.J. Batenburg, H.A. Calderon, R. Kilaas, V. Radmilovic, C. Kisielowski, 3-D reconstruction of the atomic positions in a simulated gold nanocrystal based on discrete tomography, *Ultramicroscopy* **108** (May 2008), 589-604.
[21] A. Kuba, L. Rodek, Z. Kiss, L. Riskó, A. Nagy, M. Balaskó, Discrete tomography in neutron radiography, *Nucl. Instrum. Methods Phys. Res.* **A**542 (2005), 376-382.
[22] J.C. Palacios, L.C. Longoria, J. Santos, R.T. Perry, *Nucl. Instrum. Methods Phys. Res.* **A**508 (2003), 500-511.
[23] H.J. Ryser, Combinatorial properties of matrices of zeros and ones, *Canad. J. Math.* **9** (1957), 371-377.
[24] H. Slump, J.J. Gerbrands, A network flow approach to reconstruction of the left ventricle from two projections, *Comput. Graph. Image Process.* **18** (1982), 18-36.
[25] P. Schwander, C. Kisielowski, M. Seibt, F.H. Baumann, Y. Kim, A. Ourmazd, Mapping projected potential, interfacial roughness, and composition in general crystalline solids by quantitative transmission electron microscopy, *Phys. Rev. Lett.* **71** (1993), 4150-4153.
[26] B. Wang, F. Zhang, On the precise number of $(0,1)$-matrices in $\mathcal{A}(R,S)$, *Discrete Math.* **187** (1998), 211-220.

B.E. Van Dalen, R. Tijdeman
Mathematisch Instituut, Leiden University
Postbus 9512
2300 RA Leiden
The Netherlands
E-mail address: bevandalen@gmail.com
E-mail address: tijdeman@math.leidenuniv.nl

L. Hajdu
University of Debrecen, Institute of Mathematics
and the Number Theory Research Group
of the Hungarian Academy of Sciences
Debrecen
P.O. 12.
H-4010
Hungary
E-mail address: hajdul@math.unideb.hu