Discrete Tomography: a joint Contribution by Optimization, Equivariance Analysis and Learning

Öznur Yaşar* and Gerhard-Wilhelm Weber †
* Department of Mathematics and Statistics, Memorial University of Newfoundland, St. John's, NL, Canada, oznur@math.mun.ca
† Institute of Applied Mathematics, METU, Ankara, Turkey, gweber@metu.edu.tr

Abstract Optimization theory is a key technology for inverse problems of reconstruction with applications in science, technology and economy. Discrete tomography is a modern research field which deals with finite objects from VLSI chip design or medical imaging. This paper focusses on the utilization of modern optimization methods to approximately resolve the NP-hard reconstruction problem of discrete tomography. Our new approaches and introductions are based on modeling and algorithms from coding theory and optimal experimental design. Here, we combine continuous and discrete optimization with exploiting geometrical symmetries, or more generally, equivariances, in a framework of statistical learning.
Keywords Inverse Problem, Discrete Tomography, Optimization, Statistical Learning, VLSI Chip Design and Medical Imaging.

1 Introduction

The starting point of *Computerized Tomography (CT)* might be the need to construct the density distribution within the human body by means of X-ray projections. Let us consider the problem of locating a tumor. We often need an estimate of the location on the basis of noninvasively available data to plan the treatment or an operation. In our case, the available information consists of the projections. The algorithms of CT reconstruct the volume data at a resolution limited by the number of projections. However, it is possible to reconstruct the data at much higher resolutions, if it may take a limited number of values from a discrete set. The accuracy of localization depends on the resolution, while a larger number of projections costs higher doses of ionizing radiation. Thus, *Discrete Tomography (DT)* advances CT whenever it is applicable. In DT, we try to solve such problems in an ideal, at least approximate way, and develop algorithms.

Taking the tomography of a 3-dimensional object, e.g., a human brain, means observing the 2-dimensional slices, whereas in two dimensions the projections are obtained by rays, e.g., X-rays [26, 45]. Tomography investigates the biology of the brain and the technical devices, and it looks for improving numerical and optimal methods. For reference refer to Aster et al. [3].

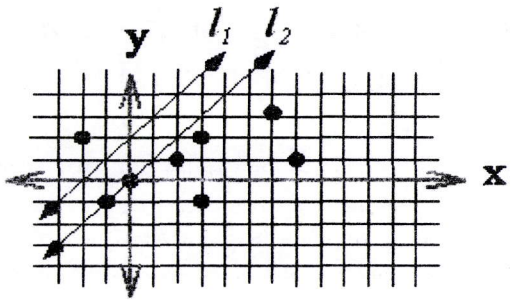

Fig. 1: *Illustration of the main problem of DT.*

We assume that we are given a domain which can be discrete or continuous and a function F with a discrete range. Our aim is to reconstruct F from weighted sums which are the projections of F in the chosen horizontal and vertical directions in the image (cf. Figure 1). In these projections and their records, our measurements and data consist. This is an *inverse problem* in which we want to reconstruct a lattice set from its X-rays or projections. Batenburg [4] presents an improvement of a reconstruction algorithm in order to minimize the time complexity where the optimized version is 50 times faster than the existing approach [30].

When the object to be reconstructed is assumed to be continuous, real analysis is used to develop the theory. By the necessity of numerical evaluations, those *Continuous Tomography* problems often become *discretized*; in many cases, then, a linear system of equations is the resulting problem [3]. As the name *Discrete Tomography* suggests, DT is a part of discrete mathematics [21, 27]. For further information on the theory, algorithms and applications of DT, we refer to the book of Herman and Kuba [35] and, related with our approaches, [71, 73, 74]. In recent years, various applications of DT became reported. DT has been applied to diverse areas such as medical sciences, image processing, electron microscopy, scheduling, statistical data security, game theory and material sciences. For instance, as a first result on medical applications, Reiber et al. [60] reconstructed the right coronary artery from two cineangiograms. For various medical applications of DT such as enhancement of tomographic images, reconstruction of human organs, e.g., blood vessels, we refer to the survey paper of Kuba et al. [47].

In this paper, we first give the notation and basic problems of DT. In Section 3, we present algorithms coming from optimization and statistical learning, supported by the recognition and exploitation of geometrical-algebraical invariances. In Section 4, we provide a short survey of some important applications. We will finish the paper with open problems and an outlook. Since in this paper various methods and traditions come together, we emphasize their concerted interplay whereas, however,

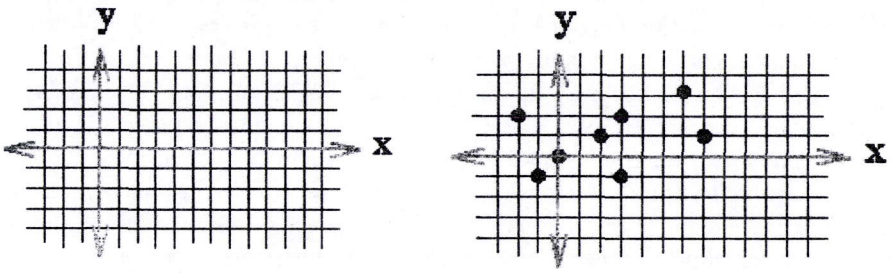

Fig. 2: *A lattice is given on the left hand side; a simple lattice set is shown on the right hand side.*

not all technical aspects can be recalled and presented. This will be done in detail in future works.

2 Preliminaries

Let us introduce the basic notation and definitions which is similar to Shepp et al. [15]. Here, $\mathbb{Z}$ stands for the set of integers, and $\mathbb{N}_0$ denotes the set of natural numbers including 0. *Lattice sets* are discrete sets $F \subset \mathbb{Z}^d$ which are finite subsets of integer vectors (cf. Figure 2). The embedding set $\mathbb{Z}^d$, called a *lattice*, or a rectangular subset of it containing F can be considered as a regular grid of points or positions. The latter ones are also called *cells*. If we are in medical applications of DT, then, this notion has a natural meaning. *Lattice directions* are nonzero vectors v in the lattice $\mathbb{Z}^d$, but over the field $\mathbb{Q}$ of rational numbers, which implies $v \in \mathbb{Q}^d$. A finite sequence of distinct lattice directions will be denoted by D, hence, for some $q \in \mathbb{N}$, $q \geq 2$,

$$D = (v_1, v_2, \ldots, v_q). \tag{1}$$

A *lattice line* l is parallel to a vector $v_k \in D$ and, furthermore, it has nonempty intersection with the lattice: $l \cap \mathbb{Z}^d \neq \emptyset$. For a visualization of lattice lines see l_1 and l_2 in Figure 1. The set of all lattice lines which are parallel to $v_k \in D$ is denoted by L^k, and $\mathcal{E}$ will be the class of finite sets in $\mathbb{Z}^d$. The collection of the set of lattice lines determined by D is represented by

$$L = (L^1, L^2, \ldots, L^q), \quad q \geq 2. \tag{2}$$

Based on these definitions and notations, a lattice line l may, e.g., correspond to an X-ray, and it will be chosen from the set L^k if l is parallel to some chosen direction v_k. In that context, a lattice set represents an atom cluster embedded into a grid (lattice) of possible atom positions.

Based on a discretization prepared and on a slicewise reduction of the reconstruction to two dimensions, we can refine our tomographical problem by a ***feasibility problem*** posed in the planar way of dimension 2. Then, any two distinct lattice directions can be considered. Since a suitable linear transformation can be found which transforms any two dimensional lattice to a lattice with directions $(1,0)$ and $(0,1)$, we will consider this latter case. In applications studied by us below, however, the iterative choice of the lattice directions is part of the challenge and art, an element of a learning process [34]. There are more directions than these standard ones (provided for simplicity) possible, namely, e.g., the diagonal-wise and counterdiagonal-wise ones, and, moreover, a refinement by partial and local instead of global measurements.

At this stage, our aim is to find a binary vector which satisfies a matrix equation given by

$$Px = b, \tag{3}$$

where $P \in \{0,1\}^{M \times N}$ and $b \in \mathbb{N}_0^M$. Namely, if the smallest rectangular box containing the finite set to be reconstructed has dimension $n_1 \times n_2$, then, $M = n_1 + n_2$ and $N = n_1 \cdot n_2$. Hence, M is the number of lattice lines, in this case being parallel to the directions $(1,0)$ and $(0,1)$ on which there is at least one element from our discrete set, and N is the total number of points or positions considered in our reconstruction problem.

According to the projections taken we define the matrix P, sometimes referred to as a *view matrix* [28]. The vector b consists of the raywise recorded experimental data. Considering that the surface or tissue to become reconstructed is discretized by cells, this matrix will consist of rows whose components are 1 for any cell where the ray (represented by a row of P) goes through, and 0 if the ray does not meet that cell. Here, the binary values come from the provided orthogonality of the axes of directions and from the distance of a unit 1 between neighbouring lattice points along of these directions. In cases of nonorthogonality or, more generally, coming from an underlying and discretized continuous problem, P may also have nonbinary values of distance or values of further physical, biological or chemical dimensions [3].

We could as well extend our feasibility problem by imposing the maximization of $f(x)$, defined as the sum of the components x_i ($i \in \{1,2,\ldots,N\}$) with respect to eq. 3, i.e.,

$$(\mathcal{P}_1) \qquad \text{maximize} \quad f(x) := \sum_{j=1}^{N} x_j, \quad \text{subject to } Px = b \text{ and } x \in \{0,1\}^N.$$

A solution of this ***optimization problem*** implies feasibility, but it additionally aims at a maximal density of atoms or of other considered discrete objects, in other words, at a high distribution or a large support. For example, the value associated with each cell, by relaxation of the discrete case, can correspond to the amount of

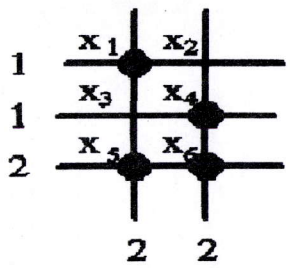

Fig. 3: *An instance illustrating the problem as a linear programming problem. Projection on two lattice lines in the directions* $(1,0)$ *and* $(0,1)$.

radiation passed onto the next cell. Hence, by maximizing the sum $f(x)$ in $(\mathcal{P}_1)$ we reversely allocate the X-ray measurement data b_i more widely, towards portions for any single cell under a threshold. Herewith, we minimize the radiation emerging from that object. Especially, in corresponding medical applications, this is very important for protecting health under X-ray application.

Other objective functions rather than $f(x)$ are possible and they can be selected depending on the studied application from science, technology, ecology, social science or medicine. For example, we could also *minimize* the objective function $f(x)$ from $(\mathcal{P}_1)$. For this minimization problem, a more general interpretation can be given by the theory of inverse problems: There is a *trade-off* between the interest in a high (re-)solution quality on the one side, and not too high a problem complexity, i.e., a minimal number of nonvanishing parameters x_j on the other side [3]. That second goal expresses itself in norm-minimizing of x, e.g., in the ℓ^2- (or Euclidean) norm, or, as in our case, the ℓ^1- (or sum-) norm. Please note the nonnegativity implied by the binary constraints.

For our optimization problem $(\mathcal{P}_1)$, there are polynomial-time interior point methods (cf. [36, 66]). In *Section 3*, we give detailed interpretation and new insights from the viewpoint of optimization.

Example 1. *Consider the lattice set given in Figure 3. It is contained in a* (3×2)*-rectangle, hence,* $M = 5$ *and* $N = 6$. *Here, we know the projections along* 3 *horizontal and* 2 *vertical directions. For this instance we have the following system of equations*

$$\begin{aligned}
x_1 + x_2 &= 1\\
x_3 + x_4 &= 1\\
x_5 + x_6 &= 2\\
x_1 + x_3 + x_5 &= 2\\
x_2 + x_4 + x_6 &= 2.
\end{aligned}$$

Here, P, x and b are

$$P = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ lot\ 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix}, \quad b = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \\ 2 \end{pmatrix}. \quad \blacksquare$$

The projection of a lattice set in direction v_k is $p_F^{(k)} : L^{(k)} \to \mathbb{N}_0$ such that

$$p_F^{(k)}(l) = |F \cap l| = \sum_{x \in l} f(x), \tag{4}$$

where f is the characteristic function of F, i.e., $f(x) = 1$ if $x \in F$, and $f(x) = 0$ if $x \notin F$.

Two lattice sets F and F' are said to be *tomographically equivalent* with respect to the directions D, if the following equality is satisfied:

$$p_F^{(k)} = p_{F'}^{(k)}, \quad k = 1, 2, \ldots, q. \tag{5}$$

Now, let us state the three main problems which discrete tomography is concerned with:

Consistency$(\mathcal{E}, L)$

Given: Functions $p^{(k)} : L^{(k)} \to \mathbb{N}_0$, $k = 1, 2, \ldots, q$, with finite support (i.e., for every k: $p^{(k)}(l) \neq 0$ for finitely many $l \in L^{(k)}$ only).

Question: Does there exist an $F \in \mathcal{E}$ such that $p_F^{(k)} = p^k$ for $k = 1, 2, \ldots, q$?

Uniqueness $(\mathcal{E}, L)$

Given: An $F \in \mathcal{E}$.

Question: Does there exist a different $F' \in \mathcal{E}$ such that F and F' are disjoint and tomographically equivalent with respect to the directions of D?

If a set is nonunique, then it cannot by its projections be distinguished from some other set in $\mathbb{Z}^d$. We note that utilizing available a priori information on the reconstructed object may help to distinguish the most likely solution. The successive discovering of such a knowledge will later on be part of our algorithmic approach.

Reconstruction $(\mathcal{E}, L)$

Given: Functions $p^{(k)} : L^{(k)} \to \mathbb{N}_0$, $k = 1, 2, \ldots, q$, with finite support.

Task: Construct a finite set $F \in \mathcal{E}$ such that $p_F^{(k)} = p^k$ for $k = 1, 2, \ldots, q$.

It should be clear from the definitions that if the reconstruction problem is solvable, then the consistency problem can be solved as well. In the following, we come to *complexity* results [22] for our reconstruction problems by lattice lines. For closer information refer to [20, 42], and we begin with a small table. Here, **PTA** stands for the existence of a corresponding *polynomial-time algorithm.*

	Consistency	Uniqueness	Reconstruction
$q = 2, d = 2$	PTA	PTA	PTA
$q \geq 3$	NP Complete	NP Complete	NP Hard

One of the results stated by Gritzmann and Gardner [75] says that any discrete subset of $\mathbb{Z}^2$ is unique with respect to D if the cardinality of D exceeds 6.

The complexity of the problem is not in general developped yet for r-dimensional X-rays when r is greater than 1. Only a few results are known. For instance, when $r = 2$, $d = 3$, and L consists of the 3 coordinate planes, the problems still remain open.

If we have $q = 2$ lattice directions only, then, due to *Ryser's Theorem* [64], by checking whether a known submatrix is contained in a binary matrix or not, we can say whether it is unique or not. On the other hand, it is known that one cannot check uniqueness with the same procedure if there are $q \geq 2$ directions.

Because of the described problem complexities, one looks for new algorithmical methods. In the following, we give a contribution to this by modern optimization combined with geometrical-algebraical properties of structure and in a process of learning, of information gaining and using.

3 Optimization in Discrete Tomography

For the sake of simplicity, we shall in the following concentrate on referring to the applied background from *microchip* design. This preference in motivation will also well fit to our methods from experimental *design* which will be later on presented.

3.1 Motivation and Preparation by Optimization in VLSI Chip Design

An *integrated circuit (IC)* is a tiny semiconductor chip, on which a complex of electronic components and their interconnections are fabricated with a set of pattern defining masks. The technique of fabricating an IC with a set of masks is somewhat analogous to that of creating a photograph with a negative.

The evolution of IC technology is measured by the number of components integrated on a single chip. The IC industry has gone through the milestones of *small scale integration (SSI)*, *medium scale integration (MSI)*, *large scale integration (LSI)*, and *very large scale integration (VLSI)*. Known as *Moore's Law*, the maximum number of transistors on a chip approximately doubles every eighteen months. The creation of a VLSI circuit involves a large number of activities such as modeling, simulation and logic design,[69]. *Optimization* plays a very important role in all of these tasks.

Here, DT deals with the reconstruction of an atom cluster which is faced with in *quality control* of a microchip's boundary layer, [48, 55]. Since the problem of reconstruction is NP-hard, we look for *approximative* algorithms using optimization in a refined way, coupled by structure analysis and statistical learning.

In fact, the entire production process of microchips concludes with a guarantee control of whether the outer layer on a microchip, consisting, e.g., of silicon, is not too rough. Therefore, we investigate the layer's surface structure. In other words, we wish homogeneous boundary layers. Detecting the *holes* and all kinds of *nonconvexities* can be defined in the discrete terms of the grid and of the lattice set of atoms, and it asks to reconstruct the atom cluster and its structure. For this purpose, we want to find out where the atoms are located. This is our *inverse problem*. Finding the concrete layer is our *reconstruction* problem or, in other words, our problem on *estimation* or *approximation* of parameters. By means of measurements, of discovering the geometrical and algebraical structure, and by a combined process of statistical learning and optimization, we try to solve this problem efficiently.

Our learning consists in a stepwise approximation of the atom cluster: Firstly and, then, in every iteration, we may embed the possible cluster into a slighly bigger (neighbouring) discrete set, which also lies in the underlying grid and may be a bit easier to investigate. If fact, there may be small holes or cuts (i.e., paths without atoms) become closed by us in order that we can expect more symmetry-like structures. Such a *preprocessing* by a set extension can be made by an insertion of further atoms which will be taken off again at the end of our solution algorithm. Later on, we will come back to this small aspect, but for the moment disregard it. It is enough for us to consider the atom cluster lying in a sufficiently large discrete rectangle which is a subset of the grid.

Now, using a series of X-ray measurements, we discover and systematically exploit our lattice set's *equivariant* (e.g., symmetrical) properties. This can be done *globally*, but also *locally* or *partially* for a finite number of subsets which entirely cover our neighbouring rectangle, and by minimizing an appropriate objective function. Such discrete subsets or ***windows*** may, e.g., be a smaller (discrete) squares, discs, but also (less locally, but partially), frames (differences between concentric squares) and rings or, since our reflections also apply in higher dimensional spaces, cubes, balls, frames of higher dimensions, or tori. Herewith, they may have very different connectivity properties, or, more generally, their algebraical topological may differ very much [68]. As being given in these examples, the subsets should reveal geometrical-algebraical symmetries which can be represented as a closedness under motions, i.e., group orbits, to which we will come next. Herewith, the subsets become the place where those parts of the atom cluster which are lying in them, just like in a subspace, can appropriately be investigated under the qualitative criterion of equivariance. Let us mention that our approach on window covering (or partitioning) with a variety of maturities of reduced dimensions being possible within each window, can be considered as a special and discrete version of *exhaustion* principle [13, 25, 63, 70] applied on our problem from DT.

In fact that objective function represents an expected error or a variance of our estimation, and it should be *invariant* with respect to those motions, like rotations, permutations and sign-changes, which are describing the equivarince struc-

tures. This will below become closer defined by means of group elements and regular transformations. Then, finally, these equivariances, if verified by the X-ray measurements, help us to simplify the representation of the atom (sub-) cluster or of its estimate (iterate) and, herewith, globally, partially or locally to dimensionally reduce the complexity of our reconstruction problem. Besides of the probabilistic meaning of such a reduction [18], there can also a geometrical interpretation be given: Globally, or in the partial or local windows of our subsets, we look at less and, possibly, other coordinates (e.g., polar ones), or at lower dimensional spaces where the group orbits pass through. In terms of corresponding continuous models, we may think of transversal sections [38] traversed by trajectories (say, solving a system of differential equations [2, 56]). Later on, when the reconstruction problem is approximately resolved, we follow the obits or discrete trajectories back, assign the problem result from the points on the section to all the other points on the corresponding orbit. This *backward assignment* will orbitwise be done by the same values, or up to a regular change of coordinates of the measurement vector, e.g., by a permutation of the components or another transformation related with the group of orbits.

In the course of this approach, we make use of statistical learning with its algorithms by *training* (i.e., first determination of model parameters) and *testing* (error analysis and successive minimization). In this context, the most famous optimization algorithms apply for minimizing *least squares* of errors, they find a *maximum likelihood* of the unknown parameters, or they are called *EM-algorithms* (expectation-maximization). In this sense, we will present various approaches to solve variants and generalizations of the problem $(\mathcal{P}_1)$.

3.2 Coding Theory Applied in Discrete Tomography

We consider a sufficiently large rectangular subset of the lattice with the unknown atom cluster included (to be studied by rays, mainly, X-rays), as a *word* being close to one or another element of a finite linear space $\mathcal{C}$. This word can be easily found by aligning the columns (or rows) of the binary matrix given by the rectangular set and the lattice directions. This initial linear code incorporates any preinformation which we have about the possible location of the atom cluster. At first, we only have a vague initial guess about the atom distribution. We try to find the codeword which is closest to our guess by *decoding*. In order to decode, in tomographical terms: to reconstruct, we can use the *Optimal Decoding Rule* or *Maximum Likelihood Decoding Rule* [50].

We generate a *training set* and a *test set* of measurements. By training we refine and suitably build up a code where we suppose that the atom distribution should be an element in or approximated by. Afterwards, but *iteratively* coupled with training, by testing we validate or falsify and successively improve the code and, herewith, the approximation of the real atom cluster. Within of this learning process, we step by step ***reduce*** the dimension or complexity of the code (herewith, adapting it to our

increasing structural insight). This implies an alternating sequence of training and test errors being analyzed. This is also a proceeding source for conclusions about the further measurement design which is given, e.g., by the window partitioning and by the choice of lattice directions therein.

Since a linear code is defined as a vector space over some finite field $\mathbb{Z}_p$, where p is some prime number, we use coding theory to find geometrical properties of our atom cluster in a step by step process of measurement and improvement. Since we are dealing with discrete tomography in VLSI chip design, our code is over the field $\mathbb{Z}_2$, hence $p = 2$. Here, 1 means existence and 0 means nonexistence of an atom at a lattice point. Let us assume again that we are working in the dimension $N = n_1 \cdot n_2$ of a, say, smallest rectangle containing the finite (sub-) set to be reconstructed. For measuring the difference between the approximative iterate and the closest codewords, we use the Hamming distance.

Formalization of ML: Let us have the preinformation that our atom distribution is firstly approximated by at least one element c^0 of a code $\mathcal{C} \subset (\mathbb{Z}_2)^N$. Then, for any $u \in (\mathbb{Z}_2)^N$ we ask:

$$(\mathcal{P}_2) \qquad \text{minimize} \qquad f(x) := \sum_{i=1}^{N} (x_i - u_i)^2, \quad x \in \mathcal{C}.$$

In fact, the *Hamming distance* of two words of a binary linear code is defined as the number of places where the words differ [41]. We note, $(x_i - u_i)^2 = x_i - u_i$ in the field $\mathbb{Z}_2$.

As our first kind of equivariance and pattern for following approaches, here, we look for *cyclicity* which implies many simplifications in decoding. We can analyze this code, hence, find out what a kind of an error-correcting code it is or what its minimum distance is. Cyclic linear codes have many special features which make them more easily applicable [41]. We remark that our problem ML can also be refined in a stochastic manner [34] based on distribution assumptions about noise in the X-ray measurement data.

A *cyclic linear code* $\mathcal{C}$ is a cyclic subspace of the vector space $(\mathbb{Z}_p)^N$ in which it is defined. Given a code word here and considering the first component (letter) succeeding the last component in this word, then, any uniform shift to the right in the coordinates does not lead to a leaving from the code. This cyclic permutation of the coordinates can be imagined by (closed) *orbits*. Herewith, cyclicity means a closeness condition in neighbouring structure of the atoms. Let us enumerate the position of the word and state the following correspondence:

$$u = (a_0, a_1, \ldots, a_{N-1}) \in (\mathbb{Z}_q)^N \quad \Leftrightarrow \quad u(x) = a_0 + a_1 x + \ldots + a_{N-1} x^{N-1}.$$

In the sequel, we may use this identification without any special mentioning of it. Furthermore, note that working on $\mathcal{C}$ is equivalent to working in the ring $\mathbb{Z}_p[x]/\rho(x)$ of polynomials modulo $\rho(x) := x^N - 1$. We recall that

$$u(x) = w(x) \mod (g(x)) \quad \Leftrightarrow \quad u(x) - w(x) \text{ is divisible by } g(x),$$

and if $g(x)$ has degree r, then $|\mathbb{Z}_p[x]/g(x)| = p^r$. Let $g(x) = g_0 + g_1 x + \ldots + g_r x^r$ be the *generator polynomial* of our code $\mathcal{C}$ (under the above identification), i.e., $\mathcal{C} = \{a(x)g(x)|a(x) \in \mathbb{Z}_p[x]/\rho(x)\}$.[1] In other words, $\mathcal{C}$ is the code generated by $g(x)$, in symbol: $\mathcal{C} =< g(x) >$. Then, a generator matrix G for $\mathcal{C}$ representing $\mathcal{C}$, by definition, in the way $\mathcal{C} = \{Gc^T|c \in (\mathbb{Z}_p)^N\}$, is

$$G = \begin{pmatrix} g_0 & g_1 & g_2 & \cdots & g_r & 0 & 0 & \cdots & 0 \\ 0 & g_0 & g_1 & g_2 & \cdots & g_r & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & g_0 & g_1 & g_2 & \cdots & g_r \end{pmatrix}. \tag{6}$$

For the multiplication with such as special matrix G we can use *shift registers*. Applying the generator polynomial $g(x)$ instead of general generator matrices G, i.e., scalar polynomial algebra and calculation instead of higher dimensional matrix algebra is much more convenient. Herewith, we have reduced the problem complexity to a great extent. Then, but iteratively coupled, by a suitable decoding algorithm we can find out what the most closest code words, hence, most probable approximate atoms' locations, are. Furthermore, by using $g(x)$ we canonically define the *control polynomial*

$$h(x) := \frac{x^N - 1}{g(x)},$$

where, by definition, $\mathcal{C}$ just consists of the polynomials $c(x) \in \mathbb{Z}_p[x]/\rho(x)$ with $h(x)c(x) = 0$ (modulo $\rho(x)$). By $g(x)$ we can do error-correction more easily than with a canonical *control matrix*. We illustrate this by the following example which we learn from [41].

Example 2. *Let our dimensions be $p = 2$, $r = 3$, $N = 2^r - 1 = 7$ and let our cyclic linear code be a Hamming code, i.e., the columns of our control matrix H are pairwise different and nonvanishing. Here, the generator polynomial which we use is just the minimal polynomial $\mu(x) = x^3 + x + 1$. Referring to a primitive element a (where $\mu(a) = 1$), we can represent our preferred control matrix by*

$$H = (1, a, a^2, \ldots, a^6).$$

In fact, calculating modulo $\mu(x)$ in the field of polynomials $\beta_0 x^2 + \beta_1 x + \beta_2$ which can in the standard way be represented by the columns $(\beta_0, \beta_1, \beta_2)^T$ (and inserting $x = a$), our control matrix takes the form

$$H = \begin{pmatrix} 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 1 \end{pmatrix}.$$

[1] We hope that the traditional notation $g(x)$ for the generator polynomial made by us in this Subsection 3.2 does not lead to any confusion with the symbol g used for orbits outside of this subsection.

This special matrix, obtained from polynomial calculation and just representable by powers of a, is more convenient than the canonical control matrix which results from the form of the corresponding generator matrix. Now, let us think that we have received the word u or, equivalently, polynomial $u(x) = c(x) + x^i$, where $c(x)$ is the correct code word (where being in the code just means $c(a) = 0$) and x^i stands for the error term at the i-th position of the word. Both $c(x)$ and x^i are unknowns. By the equations

$$Hv^T = u(a)\ (= c(a) + a^i) = a^i,$$

however, we identify the error x^i. In fact, if, e.g., $u(x) = x + x^4$, with $c(x)$ and x^i hidden, then, the calculation $u(a) = a + a^4 = a + (a + a^2) = (a + a) + a^2 = a^2$ (in $\mathbb{Z}_2$) allows us detect the error at the place $i = 2$. Herewith, we get as our result of reconstruction by ML principle: $c(x) = x + x^2 + x^4$ and the corresponding code word c. ∎

In DT applications, the dimensions are of course, much larger than in this example, but our basic idea from *de*coding remains the same for a best approximation of the atom distribution by a code word. We recall that, in our approach, this approximation is embedded in a learning processes with, e.g., cyclicity found and exploited. Concerning that structural property, in the next Subsection 3.3 we present and use further examples in the rich, modern and general class of equivariances.

3.3 Optimal Experimental Design Applied in DT

Here, the basic idea is to look for *equivariances* in the atom distribution or of an approximate of it. By these, we continue our reflections initiated with cyclicity in Subsection 3.2. They will simplify the problem and reduce the dimension. These generalized symmetry-like properties could be worked out by our X-ray measurements in a way of trial and error. However, we will do it more systematically by remaining in a framework of statistical learning. Then, with that structural information at hand, we ask for an optimal approximation of the atom distribution, where the optimization function is, e.g., likelihood of the estimation (to be maximized), its bias and/or variance (to be minimized) [34], and it should be compatible with the equivariance structure which we found. This compatibility will be called *invariance*. Both invariance and equivariance information will be represented and utilized by *optimal experimental design* theory from statistics [18]. Before we introduce them closer, please note that these geometrical-algebraical methods of Gaffke and Heiligers have already successfully been applied for the elasticity of crystals in material science [6].

Let us regard a *linear regression model* $y(x) = \theta^T p(x) \quad (x \in \mathcal{X})$, where $p = (p_1, p_2, \ldots, p_n)^T$ is an $\mathbb{R}^n$-valued function on the experimental region $\mathcal{X} \subset \mathbb{R}^q$ with compact range $p(\mathcal{X})$, and $\theta = (\theta_1, \theta_2, \ldots, \theta_n)^T \in \mathbb{R}^n$ is an unknown parameter vector. For example, the basis functions $p_j(x)$ may be polynomials or basic trigonometric functions. By an approximate design $\mathcal{E}$ we mean a discrete probability measure on

$\mathcal{X}$ with finite support, i.e., it assigns nonnegative weights $\mathcal{E}(x_i)$ to finitely many points $x_i \in \mathcal{X}$ $(i = 1, 2, \ldots, N)$ only and such that $\sum_{i=1}^{N} \mathcal{E}(x_i) = 1$. The moment matrix, $M(\mathcal{E})$, which reflects the statistical properties of the design $\mathcal{E}$, is a positive semidefinite $(n \times n)$-matrix. The components of $M(\mathcal{E})$ are the covariances $Cov(p_i, p_j)$ $(i, j = 1, 2, \ldots, n)$.

To decide about the optimality of a design developed in this way, we need a real-valued convex criterion on the convex cone $\mathcal{C}$ of symmetric $(n \times n)$-matrices. This is related with cone or semidefinite programming; we refer to [5]. Our feasible set, being a convex cone, can be viewed as the infinite intersection of hyperplanes, yielding a semi-infinite optimization problem [36]. Let us call a design *optimal* [17] if it is an optimal solution of

$$(\mathcal{P}_2) \qquad \min \quad f(M(\mathcal{E})), \quad \text{subject to} \quad M(\mathcal{E}) \in \mathcal{C}.$$

An offer of various optimality criteria, f, is provided, the most common ones being *Kiefer's Φ_p-criteria* or *integrated variance* or, more generally spoken, an *error criterion*. Mostly, these are not so complicated functions, defined, e.g., by the eigenvalues or the trace of C or of C^{-1}, where C is a symmetric positive definite $(n \times n)$-matrix. Herewith, by such an approximate criterion or objective function, we aim at a highest probability of atom cluster reconstruction, at a smallest prediction error and standard deviation of measurement error or noise, etc., or at their additive combinations [34]. Higher complexity and computational expenses arise for this minimization task, when we increase the dimension. However, we can turn from a global to a partial or local model (cf. Subsection 3.1) and, in those windows, we can use special structural properties in order to reduce the dimension. To this end we will look for *optimal designs*, globally or, at least, in the partial or local sense of subdesigns, and approximatively. In this sense, the following denotation and terminology can also be understood in the relative sense of partial or local models and, furthermore, they can from time to time be adapted to the growing insights we gain in the course of measurement and learning.

Let all $(n \times n)$-moment matrices C $(C = M(\mathcal{E}))$ be lying in some convex cone $\mathcal{C}$ with semidefinite elements and the given measurement space $\mathcal{X}$. As in Subsection 3.2, we can again by a suitable alignment of columns or rows directly turn from matrices to vectors, e.g., for applying optimization methods below. Furthermore, f is some suitable criterion on $\mathcal{C}$ to be minimized, $F(x)$ stands for the measurements of atom distribution along of a considered lattice direction and passing through a lattice point x. Furthermore, $\mathcal{Q}$ is a compact group of regular matrices $\mathcal{Q} \in \mathbb{R}^{n \times n}$ and $\mathcal{G}$ is a group of bijective $g : \mathcal{X} \to \mathcal{X}$. These functions can by their image or trace be imagined as *orbits*. For example, $\mathcal{G}$ is the group of orthogonal transformations (e.g., rotations), of permutations or the group of sign changes, or groups jointly generated by such groups. These group elements and their orbits by which we model symmetry-like structures of equivariance, can be of different discrete densities. In combination with the X-ray measurements, these orbits and their densities help

to describe structure and homogenity of the unknown atom cluster. In cases of high density and, for this, a more global lattice section provided, the orbits may be regarded as discrete approximations of continuous trajectories. We remark that in cases of rotations, the orbits pointwise follow a moving and detecting *radar beam*; this model is well-known in the discrete mathematics of oriented matroids [7, 8]. In the partial or local windows of our research, or just globally, we may for our research think of *generalized* radar beams working.

The concerted interactions of all these functions and the matrices, called **invariance** and equivariance, leads to dimensional reduction [17, 18]. Finally, semidefinite and nonlinear integer programming problems have approximatively to be resolved. Before we mention some basic ideas on these optimization methods, let us state the central conditions which, in fact, are fulfilled for a number of constellations of our examples on functions and matrices:

Invariance means

$$f(QCQ^T) = f(C) \quad \forall\, Q \in \mathcal{Q},\ C \in \mathcal{C},$$

where f is one of those appropriate criteria.

Equivariance means

$$F(g(x)) = Q_g F(x) \quad \forall\, g \in \mathcal{G},\ x \in \mathcal{X}.$$

Under these assumptions and the fact that our set is compact and convex, *Jensen's inequality* [62] implies that $(\mathcal{P}_2)$ has an optimal solution. In the paper [17] on optimal experimental design, an iterative algorithm for solving $(\mathcal{P}_2)$ approximately is given. We adapt and apply it suitably for our discrete tomography problems. In order to achieve a better convergence rate, the following quadratic approximation is constructed:

$$f(m) \approx f(m_\nu) + < \mathrm{g}_\nu, m - m_\nu > + \frac{1}{2} < \mathrm{H}_\nu(m - m_\nu), m - m_\nu > \quad (m \in \mathcal{C}).$$

Here, g_ν is the gradient, H_ν is the Hessian of f at m_ν, the iteration point at the ν-th iteration, respectively, and $< \cdot, \cdot >$ stands for scalar product. Sometimes, we only have approximates g_ν, H_ν for the gradient and Hessian in hand. The line search method and the step length can be chosen most efficiently. In order to minimize our quadratic approximation to the objective function f, *Higgins-Polak method* [37] is used [17] (for further optimization methods cf., e.g., [5, 51, 52]).

Figure 4, being the result obtained by both Gaffke and Heiligers [17] and Sloane [33], illustrates the outcome of our algorithm. Since [17] asks for optimal designs with rational weights, Gaffke and Heiligers finally apply the group orbits g on their optimal approximate (real) solution points, e.g., they apply rotations, perturbations and sign changes, on their optimal approximate (real) solution points, and a rounding procedure on the weights [59]. This procedure works a bit like a cake-cutting algorithm. All of the indicated points constitute a (symmetrical) exact design with

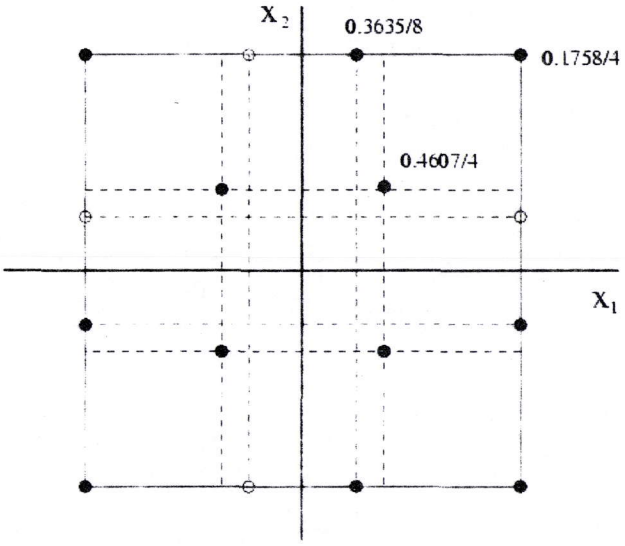

Fig. 4: *An optimally reconstructed (small) atom cluster (design).*

rational weights. The dark points constitute an approximate design with real-valued weights. Concerning our application from VLSI chip design, we consider an atom to be given at a grid point, if the corresponding weight exceeds some small threshold $\varepsilon > 0$ which is prescribed by us. 50] Since usually the set to be reconstructed is *non*symmetric, i.e., not ideally equivariant, we use the following treatment. We can still apply our orbits g, e.g., make rotations, perturbations or sign changes, whereby, we just include those lattice points which are neighbouring but not necessarily expected to be in the atom cluster, by giving them zero weights. Namely, we can add some small neighbouring parts of the lattice to get more symmetries so that we can optimize more easily. This means a *preprocessing* which may, in addition, also be made at the beginning of every iteration step. At the end of the algorithm, we delete these further auxiliary points, at least partially. Moreover, as explained above, we perform a ***backward assignment*** by means of all the equivariance conditions used, including a successive increase of the dimensions. Finally, we are are back in our original problem dimension for all the *windows* which, herewith, such to say, became *closed*. This concludes our algorithm.

We mention that Haase [29] suggests wavelet methods for the investigation for fractal-like structures of a high roughness, e.g., also in VLSI design. Such a roughness is often given in situations of little local symmetries, which we took account of in our previous reflections. One further important way how to take into account

violations of equivariance, e.g., *symmetry breaking*, can be learned from the theory of discontinuous dynamical systems [57, 58]. In fact, let us think that our orbits are close in approximation to solutions of a system of differential equations where, in addition, jumps are allowed as soon as the trajectories transversally hit a manifold [38], i.e., without any tangential effect. This means impulsive guidance of the local or global flow to a different place where, then, the flow goes on. With these jumps or impulses, we can take account of discontinuous parts in the structure of the atom cluster or tissue, which may result from damages in the production processes or, in medical applications, from violations, surgeries or mutations. For related considerations we refer to [1, 24]. On the stage of the objection function, this occurrence of jumps may in future necessitate a further development of *discontinuous optimization* [10, 39, 72].

Our combination of continuous optimization methods with exploiting geometrical-algebraical symmetries or, as we just discussed, broken symmetries, is a new contribution to the inverse problems of DT.

4 Applications

Although the birth of Discrete Tomography is not long a time ago, it has been used in many applications. Mostly, DT is applied for electron microscopy-techniques, for quality control of industrial products in industrial imaging, for reconstruction of the shape of heart chambers from orthogonal biplane cardiac angiograms in medical imaging and for quality control in VLSI chip design. These applications promise to lead to significant improvements in their fields.

The study of DT with *medical applications* is emerging as an important new research area. To recover cross-section images from a number of projections, the object is exposed to X-ray beams from a number of directions. The transmitted rays convey information about the density distribution inside of the body. The problem consists in reconstructing the best approximation of the real cross-section. In order to model the 3 dimensional shape of the left or right heart chambers from the density distributions of orthogonal biplane ventriculograms, Onnasch and Prause [54] introduce a reconstruction method. In [54], techniques of image acquisition and restoration are also presented. We remark that our interest in a small number of X-ray measurements is not only of economical interest, but also for the sake of health of a medical patient. Finally, in molecular biology, atom structures change their geometry in time. This raises a prediction problem [43] and a necessity to take the dynamical aspect into consideration during the course of measurements and statistical learning.

In *industrial imaging*, to obtain shape and dimensional information of industrial parts, CT has been used as an important and powerful tool [12], e.g., in reconstructing radioactive materials. Pointing out that most of the objects are made by one material only, one can use DT by representing the related material with 1 and air

with 0. In [9], the authors explain how DT offers valuable support in industrial imaging and manufacturing. This may be applied to a wide range of materials [65].

5 Conclusion

Discrete Tomography is a promising field of mathematics which is developing rapidly. Having various applications, its use requires appropriate modelling of the problem and efficient implementation of the algorithms. Hence, further research in this field suggests significant improvements. In this paper, we have investigated DT from the new viewpoint of a combination of optimization theory, of analysis and utilization of equivariance structures, and of statistical learning.

To mention an open problem, please consider a grid with a known number of objects of the same size to be put on each row or column. The aim is to cover the whole area with the objects. The problem is solvable in polynomial time if there is only one object, whereas the computational complexity of the two object case is a open problem. For three or more objects the problem is NP-complete. It is also known that it is NP-hard for six or more objects.

Such NP-hard problem complexities ask for approximative algorithms. The utilization of optimization theory in a framework of permanent learning is promising for this purpose, and recommended in this paper. In the future, we want to continue to refine, extend and apply the results and methods which we introduced.

Acknowledgements We would like to express our gratitude to Institute of Applied Mathematics of METU, especially, to Prof. Dr. M.U. Akhmet, Prof. Dr. A. Aytuna, Prof. Dr. B. Karasözen and Dr. H. Öktem for encouragement.

References

[1] Akhmet, M.U., Kirane, M., Tleubergerova M.A., and Weber, G.-W. (2004/2005) Control and optimal response problems for quasilinear impulse integrodifferential equations, to appear in: special issue *Advances of Continuous Optimization* of European Journal of Operational Research.

[2] Amann, H. (1983) *Gewöhnliche Differentialgleichungen*, Walther de Gruyter, Berlin, New York.

[3] Aster, A., Borchers, B., and Thurber, C. (2004) *Parameter Estimation and Inverse Problems*, Academic Press.

[4] Batenburg, K.J. (2003) Analysis and optimization of an algorithm for discrete tomography, preprint submitted to Elsevier Science.

[5] Ben Tal, A., and Nemirovskii, A. (2001) *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*, MPS-SIAM Series on Optimization.

[6] Bertram, A., Böhlke, T., Gaffke, N., Heiligers, B., and Offinger, R. (2000) On the generation of discrete isotropic orientation distribution for linear elastic crystals, J. Elast. 58, pp. 233-248.

[7] Björner, A., Las Vergnas, M., Sturmfels, B., White, N., and Ziegler, G. (1993) *Oriented Matroids*, Cambridge University Press.

[8] Bokowski, J. (1996) *Einführung in die Diskrete Mathematik*, lecture notes, Darnmstadt University of Technology, Department of Mathematics (winter semester 1995/96).

[9] Browne, J.A., and Koshy, M. (1999) C T - Assisted Engineering and Manufacturing, in: *Discrete Tomography - Foundations, Algorithms and Applications*, Birkäuser, pp. 345-360.

[10] Chew, S., and Zheng, Q. (1988) *Integral Global Optimization*, Springer.

[11] Dalinger, E. (2003) Aktuelle Fragestellungen der Diskreten Tomographie und ihre komplexitätstheoretische Behandlung, diploma thesis, Darmstadt University of Technology, Department of Mathematics.

[12] Dusaussoy, N.J., Cao , Q., Yancey, R.N., and Stanley, J.H. (1995) Image processing for CT-assisted reverse engineering and part characterization, Proc. 1995 IEEE Int. Conf. on Image Processing 3, pp. 33-36.

[13] Emerenko, A. (2000) Ahlfors' contribution to the theory of meromorphic functions, in: *Lectures in Memory of Lars Ahlfors*, Israel Math. Conf. Proc. 14, American Mathematical Society, Providence RI, pp. 41-63.

[14] Fabricius, E.D. (1990) *Introduction to VLSI Design*, John Wiley and Sons.

[15] Fishburn, P.C., and Shepp, L.A. (1999) Sets of uniqueness and additivity in integer lattices, in: *Discrete Tomography - Foundations, Algorithms and Applications*, Birkäuser, pp. 35-57.

[16] Fletcher, R. (1987) *Practical Methods of Optimization*, Wiley, second edition.

[17] Gaffke, N., and Heiligers, B. (1995) Algorithms for optimal design with applications to multiple polynomial regression, Metrika 42, pp. 173-190.

[18] Gaffke, N., and Heiligers, B. (1996) Approximative designs for polynomial regression: invariance, admissibility, and optimality, Handbook of Statistics 13, eds.: Ghosh, S., and Rao, C.R., Elsevier Science, pp. 1149-1199.

[19] Garboczi, E.J. (2002) Three-dimensional mathematical analysis of particle shape using x-ray tomography and spherical harmonics: Application to aggregates used in concrete, Cement and Concrete Research 32, 10, pp. 1621-1638.

[20] Gardner, R.J., and Gritzmann, P. (1999) Uniqueness and complexity in discrete tomography, in: *Discrete Tomography - Foundations, Algorithms and Applications*, Birkhäuser, pp. 85-111.

[21] Gardner, R.J., and Gritzmann, P. (2001) Reconstructing polyatomic structures from discrete x-rays: NP-completeness proof for three atoms, with Marek Chrobak, Theoretical Computer Science 259, pp. 81-98.

[22] Garey, M.R., and Johnson, D.S. (1969) *Computers and Intractability - A Guide to the Theory of NP-Completeness*, Freedman and Company, New York.

[23] Gebert, J., Lätsch, M., Pickl, S.W., Weber, G.-W., and Wünschiers, R. (2003) Towards DNA chip analysis and discrete tomography by inverse problems and optimization, Proceedings of 4th MATHMOD Vienna - Fourth International Symposium on Mathematical Modelling, pp. 1298-1302.

[24] Gebert, J., Lätsch, M., Pickl, S.W., Weber G.-W., and Wünschiers, R. (2004) An algorithm to analyze stability of gene-expression pattern, in: special issue *Discrete Mathematics and Data Mining*, guest eds.: Boros, E., Hammer, P.L., and Kogan, A., Discrete Applied Mathematics 144.

[25] Gehring, F., Kra, I., Krantz, S.G., and Osserman, R. (1998) The mathematics of Lars Valerian Ahlfors, Notices of the American Mathematical Society 45, 2, pp. 233-242.

[26] Gencer, N.G., and Tek, M.N. (1999) Electrical conductivity imaging via contactless measurements, IEEE Transactions On Medical Imaging 18, 7, pp. 617-627.

[27] Gritzmann, P., Gardner, R.J., and Prangenberg, D. (1999) On the computational complexity of reconstructing lattice sets from their X-rays, Discrete Math. 202, pp. 45-71.

[28] Groetsch, C.W. (1993) *Inverse Problems in the Mathematical Sciences*, Vieweg, Braunschweig.

[29] Haase, M. (2000) An Exploration of self-similar structures with wavelets, in: *Similarity Methods*, Institut für Statik und Dynamik der Luft und Raumfahrtkonstruktionen, University of Stuttgart, pp. 183-192.

[30] Hajdu, L., and Tijdeman, R. (2001) An algorithm for discrete tomography, Lin. Alg. Appl. 339, pp. 147-169.

[31] Hall, P. (2000) A model for learning human vascular anatomy, in: *Discrete Mathematical Problems with Medical Applications*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science 55, pp. 11-27.

[32] Hammer, P.C. (1963) Problem 2, Proc. Symp. Pure Math. 7 (Convexity), pp. 498-499.

[33] Hardin, R.H., and Sloane, N.J.A. (1993) A new approach to the construction of optimal design, Journal of Statistical Planning and Inference 37, pp. 339-369.

[34] Hastie, T., Tibshirani, R., and Freedman, J. (2001) *The Elements of Statistical Learning - Data Mining, Inference and Prediction*, Springer Series in Statistics.

[35] Herman, G.T., and Kuba, A. (1999) *Discrete Tomography - Foundations, Algorithms and Applications*, Birkhäuser.

[36] Hettich, R., and Zencke, R. (1982) *Numerische Methoden der Approximation und semi-infiniten Optimierung*, Teubner, Stuttgart.

[37] Higgins, J.E., and Polak, E. (1990) Minimizing pseudoconvex compact sets, Journal of Optimization Theory and Applications 65, pp. 1-27.

[38] Hirsch, M.W. (1976) *Differential Topology*, Springer.

[39] Hoffmann, A., and Geletu, A. (2004) On robustness of set valued maps and marginal value functions, preprint, Institute of Mathematics, Ilmenau University of Technology.

[40] Honghui, W. (1984) Structure and cardinality of the class $\mathfrak{U}(R,S)$ of (0,1)-matrices, J. Math. Research and Exposition 4, pp. 87-93.

[41] Ihringer, Th. (1999) *Diskrete Mathematik*, Teubner Verlag.

[42] Irwing, R.W., and Jerrum, M. R. (1994) Three-dimensional statistical data security problems, SIAM J. Comput. 23, pp. 170-184.

[43] Lim, K.F., Beliakov, G., and Batten, L.M. (2003) Predicting molecular structures: Application of the cutting angle method, Physical Chemistry Chemical Physics 5, pp. 3884-3890.

[44] Lee, Y.-J., Mangasarian. O.L. and Wolberg, W.H. (2000) Breast cancer survival and chemotherapy: A support vector machine analysis, in: *Discrete Mathematical Problems with Medical Applications*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science 55, pp. 1-10.

[45] Karbeyaz, B.Ü., and Gencer, N.G. (2003) Electrical conductivity imaging via contactless measurements: an experimental study, IEEE Transactions On Medical Imaging 22, 5, pp. 627-635.

[46] Kisielowski, C., Schwander, P., Baumann, F.H., and Siebt, M. (1995) An approach to quantitative high-resolution transmission electron microscopy of crystaline materials, Ultramicroscopy 58, pp. 417-432.

[47] Kuba, A., Herman, G.T., Matej, S., and Todd-Pokropek, A. (2003) Medical applications of discrete tomography, preprint, University of Szeged, Hungary.

[48] Kuba, A., Yasar, O., Weber, G.W., Özgür, O. and Dagli, E. (2005) Discrete Tomography: New Approaches and Results by Optimization Theory and Statistical Learning, keynote lecture, ISPRS 2005 SPATIAL DATA MINING WORKSHOP METU, Ankara, Turkey, November, pp. 24-25.

[49] Minati, L. (2005) Neuroimaging techniques: a conceptual overview of physical principles, contribution and history, to appear in the proceedings of CASYS'05, Seventh International Conference on Computing Anticipatory Systems, Liege, Belgium.

[50] Morelos-Zaragoza, R.H. (2002) *The Art of Error Correcting Coding*, John Wiley and Sons.

[51] Nash, S.G., and Sofer, A. (1996) *Linear and Nonlinear Programming*, McGraw-Hill.

[52] Nocedal, J., and Wright, S.J. (1999) *Numerical Optimization*, Springer Series in Operations Research, Springer.

[53] Nystrom, I., and Smedby, O. (2000) Analysis of magnetic resonance angiography images using skeletonization and distance transforms, in: *Discrete Mathematical Problems with Medical Applications*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science 55, pp. 75-90.

[54] Onnasch, D.G.W., and Prause, G.P.M. (1999) Heart chamber reconstruction from biplane angiography, in: *Discrete Tomography, Foundations, Algorithms and Applications*, Birkhäuser, pp. 385-401.

[55] Özgür, O. (2006) Discrete Tomographic Reconstruction Methods from the Theories of Optimization and Inverse Problems: Application in VLSI Chip Production, MSc thesis, Institute of Applied Mathematics, Ankara, Turkey.

[56] Palis, J., and de Melo, W. (1982) *Geometric Theory of Dynamical Systems: An Introduction*, Springer.

[57] Pavlidis, T. (1965) A new model for simple neural nets and its application in the design of a neural oscillator, Bull. Math. Biophys. 27. pp. 215-229.

[58] Pavlidis, T. (1966) Stability of a class of discontinuous dynamical systems, Information and Control 9, pp. 298-322.

[59] Pukelsheim, F., and Rieder, S. (1992) Efficient rounding of approximate designs, Biometrika 79, pp. 763-770.

[60] Reiber, J.H.C., Gerbrands, J.J., Troost, G.J., Kooijman, C.J., and Slump, C.H. (1983) 3-D reeconstruction of coronararterial segments from two projections, in: *Digital Imaging in Cardiovascular Radiology*, Georg Thieme Verlag: Stuttgart, Germany, pp. 151-163.

[61] Riele, H., and Batenburg, J. (2002) Research project: Discrete Tomography, Centrum voor Wiskunde en Informatica Annual Report, Mathetical Institute in Leiden, Netherlands.

[62] Rockafellar, R.T. (1970) *Convex Analysis*, Princeton University Press.

[63] Rückmann, J.J., and Weber, G.-W. (1998) Semi-infinite optimization: Excisional stability of noncompact feasible sets, Sibirskij Mathematicheskij Zhurnal 39, pp. 129-145, in Russian; also appeared in English in Sibirian Mathematical Journal 39, pp. 113-125.

[64] Ryser, H.J. (1957) Combinatorial properties of matrices of zeros and ones, Canadian Journal of Mathematics 9, pp. 371-377.

[65] Salzberg, P., and Figueroa, R. (1999) Tomography on the 3D-tori and cyrstals, in: *Discrete Tomography - Foundations, Algorithms and Applications*, Birkhäuser, pp. 385-401.

[66] Schwander, P., and Fishburn, P., and Shepp, L. and Vanderbei, R.J. (1997) The discrete Radon transform and its approximate inversion via linear programming, Discrete Applied Mathematics 75, 1, pp. 39-61.

[67] Schwander, P., Kisielowski, C., Baumann, F.H., Siebt, M., Kim, Y. and Ourmazd, A. (1993) Mapping projected potential, interfacial roughness, and composition in general crystaline solids by quantitative transmission electron microscopy, Physical Review Letters 71, pp. 4150-4153.

[68] Spanier, E.H. (1966) *Algebraic Topology*, Mc-Graw Hill, New York.

[69] Vai, M.M. (2001) VLSI Design, IEEE Computer Society Press (California) 3, CRC Press, pp. 33-36.

[70] Weber, G.-W. (1992) *Charakterisierung struktureller Stabilität in der nichtlinearen Optimierung*, Aachener Beiträge zur Mathematik 5, Augustinus publishing house (now: Mainz publishing house) Aachen.

[71] Weber, G.-W. (2002) On the topology of generalized semi-infinite optimization, Journal of Convex Analysis 9, pp. 665-691.

[72] Weber, G.-W. (2003) *Generalized Semi-Infinite Optimization and Related Topics*, Heldermann publishing house, Research and Exposition in Mathematics 29, Lemgo.

[73] Weber, G.-W., and Yaşar, Ö. (2004) Discrete tomography: A modern inverse problem reconsidered by optimization, in: Part 1 of common special issue of Journal of Computational Technologies 9 and The Bulletin of Kazakh National University 42, 3 (Mathematics, Mechanics and Informatics Issue), at the International Conference *Computational Technologies and Mathematical Models for Research, Engineering and Education*, Almaty, pp. 115-121.

[74] Yaşar, Ö., Diner, Ç., Dogan, A., Weber, G.-W., Özbudak, F., and Tiefenbach, A. (2004) On the applied mathematics of discrete tomography, Journal of Computational Technologies 9, 5, pp. 14-32.

[75] Yung Kong, T., and Herman, G.T. (1999) Tomographic equivalence and switching operations, in: *Discrete Tomography - Foundations, Algorithms and Applications*, Birkhäuser, pp. 59-83.