

# On the use of graphs in discrete tomography

**Dominique de Werra · Marie-Christine Costa ·  
Christophe Picouleau · Bernard Ries**

**Abstract** In this tutorial paper, we consider the basic image reconstruction problem which stems from discrete tomography. We derive a graph theoretical model and we explore some variations and extensions of this model. This allows us to establish connections with scheduling and timetabling applications. The complexity status of these problems is studied and we exhibit some polynomially solvable cases. We show how various classical techniques of operations research like matching, 2-SAT, network flows are applied to derive some of these results.

## 1 Introduction

Discrete tomography has had an explosive development as can be seen from the number of papers and conferences dealing with this topic (see the Introduction in [Herman and Kuba \(2007\)](#)). The interest in this emerging area was mainly due to its applications in medicine and in various domains of science and technology. Discrete tomography has been defined as the area of mathematics dealing with the retrieval of information about a geometric object (or image) from data about its sections or projections or both

(see [Gardner 2006](#)). In mathematical terms an image may be viewed as an unknown function  $f$  whose range is known to be a given discrete set (usually integer numbers corresponding to colors) or a continuous set. The problems have to deal with the determination of function  $f$  (or of an approximation of  $f$ ) from sums over subsets of its domain (or from integrals over subspaces of its domain in the case in which  $f$  is a continuous function defined on a continuous domain). In discrete tomography,  $f$  is usually defined over a discrete set (in our case an  $m$  by  $n$  array consisting of  $mn$  entries) and we are given horizontal and vertical projections (number of entries in each row and in each column where  $f$  has a given value). As mentioned in [Herman and Kuba \(1999b\)](#) the name *discrete tomography* was invented by L. Shepp in 1994 at the first meeting on this topic.

Image reconstruction problems occur in many different contexts. An application to generating grain maps of polycrystals is described in [Alpers et al. \(2007\)](#). Industrial nondestructive testing also uses image reconstruction techniques from discrete tomography (see [Baumann et al. 2007](#)). Applications of discrete tomography to agriculture are described in [Alfandari et al. \(2008\)](#), [Déroche \(2003\)](#), and [Martinis et al. \(2004\)](#) with an extensive use of graph theoretical concepts and mathematical programming. In [Di Gesù and Kuba \(2005\)](#) several problems of combinatorial image analysis are presented; they also belong to discrete tomography. Many more applications are given in the two books by [Herman and Kuba \(1999a, 2007\)](#) which also contain an introduction to discrete tomography and the presentation of topics not treated in this tutorial.

The main focus of this paper is on the use of graph models to derive solutions to the basic problem of image reconstruction. It is not meant to be a survey but rather a tutorial that will hopefully motivate the reader to apprehend the fundamental problem in discrete tomography and to have an overview on various related problems in combinatorial optimization. We shall concentrate on the reconstruction of two-dimensional matrices from two orthogonal projections. We deliberately leave aside the case of objects of higher dimension as well as the reconstruction from more than two projections or with connectivity and convexity constraints. For a three dimensional use of discrete techniques see [Déroche \(1986\)](#). For didactic purposes we include some proofs: they are given either for original results or for showing how various classical techniques of operations research (network flows, 2-SAT, matchings, etc.) may be applied to solve some of these problems.

In the next section we shall state the basic *image reconstruction problem* in discrete tomography. Relations with other problems in scheduling and in timetabling will be established. In Sect. 3 we propose several elementary models using graph theory. Section 4 presents an extension with non orthogonal projections. In Sect. 5 we discuss a problem with preassigned colors. It is a situation where the border between easy and difficult image reconstruction problems can be precisely located. We shall also introduce some extensions and variations of the basic problem which are polynomially solvable. Section 6 will present some extensions in which the projection constraints are generalized and several polynomially solvable cases are exhibited. This includes in particular a metro station refurbishment problem. Finally Sect. 7 will concentrate on cases in which requirements for unions of (two) colors are present. A brief conclusion with perspectives is sketched in the last section.

In addition, basic definitions and results on complexity theory can be found in [Garey and Johnson \(1979\)](#) while graph theoretical terms and elementary concepts are given in [Berge \(1983\)](#).

## 2 The basic image reconstruction problem from projections

We shall assume throughout this paper that we have two-dimensional images to be reconstructed from two projections (say horizontal and vertical projections). We shall be mainly concerned with the existence of solutions to our reconstruction problems but we shall not discuss the important question of uniqueness for simplicity purposes.

In this section, we shall describe the basic problem and introduce notations that will be used in the sequel.

We are given an  $(m \times n)$  array  $A = (a_{ij})$  where each entry may contain a pixel having one of the colors  $1, 2, \dots, k$ . An image is thus entirely characterized by the value  $a_{ij} \in \{1, 2, \dots, k\}$  of each entry  $(i, j)$ . To such an image we may associate its horizontal projections  $h_i^s$  that are the numbers of pixels with color  $s$  in row  $i$  ( $1 \leq s \leq k, 1 \leq i \leq m$ ). In an analogous way we have the vertical projections  $v_j^s$  that are the numbers of pixels with color  $s$  in column  $j$  ( $1 \leq s \leq k, 1 \leq j \leq n$ ). Let us call  $\mathcal{H}$  (resp.  $\mathcal{V}$ ) the collection of values  $h_i^s$  (resp.  $v_j^s$ ). Then the basic image reconstruction problem, denoted by  $RP(m, n, k, \mathcal{H}, \mathcal{V})$ , is the following:

Given the values  $m, n, k, \mathcal{H}, \mathcal{V}$ , find an assignment of a color in  $\{1, 2, \dots, k\}$  to each entry of  $A$  so that in each row  $i$  (resp. column  $j$ ) there are exactly  $h_i^s$  (resp.  $v_j^s$ ) pixels with color  $s$  (for all  $i \leq m, j \leq n, s \leq k$ ).

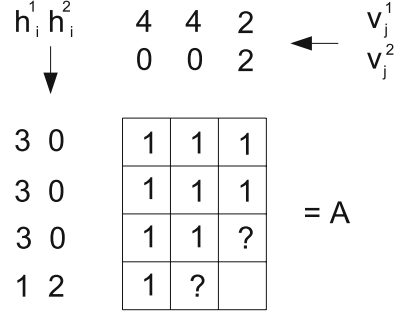
Clearly the values  $h_i^s$  and  $v_j^s$  must satisfy the following conditions.

$$\begin{aligned} \sum_{s=1}^k h_i^s &= n \quad (i = 1, \dots, m) \\ \sum_{s=1}^k v_j^s &= m \quad (j = 1, \dots, n) \\ \sum_{i=1}^m h_i^s &= \sum_{j=1}^n v_j^s \quad (s = 1, \dots, k) \end{aligned} \tag{2.1}$$

We shall assume throughout the text that these conditions do hold. These conditions are clearly necessary for the reconstruction problem to have a solution. They are not sufficient however, as shown by the following example. For  $RP(m = 4, n = 3, k = 2, \mathcal{H}, \mathcal{V})$  given in Fig. 1, one has to find an assignment of color 1 (color 2 will then be given to all entries that are not colored yet). As can be seen easily no such assignment can be found although (2.1) holds.

It is known, however, that for the case  $k = 2$  one can find whether or not there is an image corresponding to the projections  $h_i^s, v_j^s$  satisfying (2.1) as we shall see after the following definitions.

**Fig. 1** An example of  $RP(m, n, k, \mathcal{H}, \mathcal{V})$  without solution



An assignment of color  $s$  in array  $A$  is characterized by the subset of entries  $(i, j)$  with  $a_{ij} = s$ . It will be *feasible* if for each row  $i$  (resp. column  $j$ ), the number of entries  $(i, j)$  with  $a_{ij} = s$  in row  $i$  (resp. column  $j$ ) is  $h_i^s$  (resp.  $v_j^s$ ).

**Proposition 2.1** (Ryser 1957) *Given the values  $h_i^s$  ( $i = 1, \dots, m$ ) and  $v_j^s$  ( $j = 1, \dots, n$ ) there exists a feasible assignment of color  $s$  if and only if the following condition holds: let  $r_1 \geq r_2 \geq \dots \geq r_m$  be the values  $h_i^s$  and  $s_1 \geq s_2 \geq \dots \geq s_n$  be the values  $v_j^s$  rearranged in nonincreasing order; then*

$$\sum_{i=1}^m \min\{t, r_i\} \geq \sum_{j=1}^n s_j \quad (t = 1, \dots, n) \quad (2.2)$$

with equality for  $t = n$ .

Conditions (2.2) are called *Ryser conditions*. One can verify that in the example of Figure 1, they do not hold for color 1 since  $(r_1, r_2, r_3, r_4) = (3, 3, 3, 1)$  and  $(s_1, s_2, s_3) = (4, 4, 2)$ . For  $t = 2$ , we have indeed  $\min\{2, 3\} + \min\{2, 3\} + \min\{2, 3\} + \min\{2, 1\} = 7 < 4 + 4 = s_1 + s_2$ . Notice that the Ryser conditions can be checked in time  $O(\max\{m \log n, n \log m\})$ . For  $k = 2$  colors, there is a solution to  $RP(m, n, k = 2, \mathcal{H}, \mathcal{V})$  if and only if one can find a feasible assignment for color 1, since color 2 will simply be assigned to the remaining entries of  $A$ . This will provide a solution to the problem since (2.1) holds. So we have the following result.

**Theorem 2.2**  $RP(m, n, k = 2, \mathcal{H}, \mathcal{V})$  has a solution if and only if the Ryser conditions hold for color 1. In this case, a solution can be constructed in time  $O(nm)$ .

We shall show in the next section how a solution can be obtained by using a graph theoretical formulation of  $RP(m, n, k = 2, \mathcal{H}, \mathcal{V})$ . For the moment, we just mention the greedy algorithm that constructs such a solution or shows that Ryser conditions are violated (see Herman and Kuba 1999b).

First we set  $v'_j = v_j^1$  for each column  $j$  ( $v'_j$  is the number of remaining cells of color 1 to still be assigned). Then we fill the rows of  $A$  in any order. When we handle row  $i$ , we insert a 1 in the  $h_i^1$  entries located in the columns  $j$  with the largest values  $v'_j$  and we update the  $v'_j$  accordingly.

It is clear that if we cannot insert a missing 1 in some row, it means that the Ryser conditions are not satisfied.

Generally if a solution of  $RP(m, n, k = 2, \mathcal{H}, \mathcal{V})$  exists, it is not necessarily unique. In Ryser (1957), Ryser proved that a solution is unique if and only if in (2.2) we have equality for all values of  $t$ , i.e.,  $\sum_{i=1}^m \min\{t, r_i\} = \sum_{j=1}^n s_j$ , ( $t = 1, \dots, n$ ).

Let us now examine the connections with a classical timetabling problem that is formulated in the following way. We are given a collection of  $m$  classes  $c_1, \dots, c_m$  and a collection of  $k$  teachers  $t_1, \dots, t_k$  together with the  $m \times k$  requirement matrix  $R$  whose entry  $r_{is}$  is the number of 1-h lectures that teacher  $t_s$  has to give to class  $c_i$ . In addition we have a set  $H = \{1, \dots, n\}$  of periods (of 1 h) in which lectures will have to take place. For each teacher  $t_s$  a set  $T_s \subseteq H$  is given. It represents the periods where  $t_s$  is available to give lectures.

For the moment let us assume that the number  $|T_s|$  of periods in which teacher  $t_s$  is available is equal to the number  $\sum_{i=1}^m r_{is}$  of lectures which  $t_s$  has to give (we can introduce lectures to be given to fictitious classes otherwise). All classes are moreover assumed to be available at all periods in  $H$ .

A timetable will be an assignment of each lecture in  $R$  to some period such that no teacher is involved in two lectures at a time and no class is involved in two lectures at a time. Furthermore a lecture of  $t_s$  can be scheduled only at periods in  $T_s$ . This problem, called  $TIMETABLE(m, k, n, \tau)$  where  $\tau$  is the collection of subsets  $T_s$ , can be reformulated as follows.

We consider an  $(m \times n)$  array  $A$  whose rows correspond to the  $m$  classes  $c_i$  and whose columns are associated to the  $n$  periods. For each teacher  $t_s$  and each class  $c_i$  we set  $h_i^s = r_{is}$  (number of lectures of  $t_s$  to  $c_i$ ). Also for each teacher  $t_s$  and each period  $j$  we set  $v_j^s = 1$  if  $j \in T_s$  or  $v_j^s = 0$  otherwise.

A feasible timetable is an assignment of each teacher  $t_s$  to some entries  $(i, j)$  in  $A$  to be chosen in such a way that:

- (a)  $t_s$  is assigned to exactly  $h_i^s = r_{is}$  entries in row  $i$ ;
- (b)  $t_s$  is assigned to  $v_j^s \in \{0, 1\}$  entries in column  $j$ .

So  $t_s$  will be in entry  $(i, j)$  of  $A$  (or equivalently  $a_{ij} = s$ ) if and only if at period  $j$ ,  $c_i$  gets a lecture of  $t_s$ .

Notice that here  $\sum_{s=1}^k h_i^s = \sum_{s=1}^k r_{is} = n$  means that each class  $c_i$  has to have one lecture at each one of the  $n$  periods and  $\sum_{s=1}^k v_j^s = m$  means that at each period  $j$  exactly  $m$  teachers are available to give lectures to the  $m$  classes.

Clearly a timetable will exist for  $TIMETABLE(m, k, n, \tau)$  if and only if the corresponding problem  $RP(m, n, k, \mathcal{H}, \mathcal{V})$  constructed above has a solution.

**Remark 2.3** What happens if we do not have  $\sum_{s=1}^k r_{is} = n$ ? Then we must also have  $\sum_{s=1}^k v_j^s < m$  for some periods  $j$ . So we introduce a new multi-teacher  $t_{k+1}$  with

- 1.  $h_i^{k+1} = r_{i,k+1} = n - \sum_{s=1}^k r_{is} \geq 0$ ;
- 2.  $v_j^{k+1} = m - \sum_{s=1}^k v_j^s \geq 0$ .

This multi-teacher gives  $h_i^{k+1}$  lectures of “idle time” to each class  $c_i$ . It may give  $v_j^{k+1}$  simultaneous lectures in each period  $j$ . Clearly we now have a  $TIMETABLE$  problem

(with  $k + 1$  teachers instead of  $k$ ) that satisfies condition (2.1). It has a solution if and only if the initial timetabling problem has one.

In terms of image reconstruction, the timetabling problem is thus a special case in which all colors  $s$  except the last one satisfy  $v_j^s \in \{0, 1\}$ . It is known to be  $NP$ -complete as soon as  $n \geq 3$  where  $n$  is the number of periods (Even et al. 1976). A polynomially solvable case with two teachers will be mentioned in Sect. 5.

We shall examine later some extensions of  $RP(m, n, k, \mathcal{H}, \mathcal{V})$  but we shall first give some formulations in terms of integer linear programming and of graphs.

### 3 Multicommodity flow and graph theoretical formulations

We intend to give some formulations of  $RP(m, n, k, \mathcal{H}, \mathcal{V})$  that are based on graphs and on networks.

#### 3.1 A multicommodity flow problem

Given an  $(m \times n)$  array  $A$  we can associate with it a network  $N = (R \cup C, U)$  constructed as follows: Each row  $i$  of  $A$  becomes a vertex  $r_i$  and  $R = \{r_i | i = 1, \dots, m\}$ . Each column  $j$  becomes a vertex  $c_j$  and  $C = \{c_j | j = 1, \dots, n\}$ . The set of arcs is  $U = \{(r_i, c_j) | i = 1, \dots, m; j = 1, \dots, n\}$ . Now we consider that each vertex  $r_i$  in  $R$  has a supply of value  $h_i^s$  for each commodity  $s$  ( $1 \leq s \leq k$ ) and each vertex  $c_j$  in  $C$  has a demand of value  $v_j^s$  for each commodity  $s$ . Furthermore each arc  $(i, j)$  has a capacity  $c(i, j) = 1$ .

A feasible multicommodity flow in  $N$  is a function  $f^s(i, j) \geq 0$  defined for each arc  $(r_i, c_j)$  and each commodity  $s$  that satisfies:

$$\sum_{j=1}^n f^s(i, j) = h_i^s \quad (i = 1, \dots, m; s = 1, \dots, k) \quad (3.1)$$

$$\sum_{i=1}^m f^s(i, j) = v_j^s \quad (j = 1, \dots, n; s = 1, \dots, k) \quad (3.2)$$

$$\sum_{s=1}^k f^s(i, j) \leq c(i, j) = 1 \quad (i = 1, \dots, m; j = 1, \dots, n) \quad (3.3)$$

Then it is immediate to observe that  $RP(m, n, k, \mathcal{H}, \mathcal{V})$  has a solution if and only if there exists in  $N$  a feasible multicommodity flow that is integral:  $f^s(i, j)$  will be 0 or 1 on each arc  $(i, j)$  for each commodity  $s$  and we shall have  $f^s(i, j) = 1$  if and only if  $a_{ij} = s$ , i.e., entry  $(i, j)$  of  $A$  gets a pixel of color  $s$ . We recall that finding a feasible multicommodity flow is easy by linear programming. The difficulty is to have one with integral values.

### 3.2 Models based on graphs

Let us now introduce two models based on graphs. The first one associates with every row  $i$  (resp. column  $j$ ) of  $A$  a vertex  $r_i$  (resp.  $c_j$ ) as in the multicommodity flow formulation. Let  $G = (R \cup C, E)$  be the graph obtained after introducing into  $E$  all edges  $[r_i, c_j]$  for  $i = 1, \dots, m$  and  $j = 1, \dots, n$ .  $G$  is the complete bipartite graph  $K_{m,n}$  where  $m = |R|$  and  $n = |C|$ . Each edge  $[r_i, c_j]$  corresponds to an entry  $(i, j)$  in  $A$ . If all values  $h_i^s$  and  $v_j^s$  are given then  $RP(m, n, k, \mathcal{H}, \mathcal{V})$  amounts to finding a partition of the edge set  $E$  satisfying some conditions.

#### Model bipartite

Given a complete bipartite graph  $G = (R \cup C, E)$  and values  $h_i^s, v_j^s$  for  $1 \leq i \leq m$ ,  $1 \leq j \leq n$  and  $1 \leq s \leq k$  find a partition  $E^1, \dots, E^k$  of  $E$  such that for  $1 \leq s \leq k$ :

$$e^s(r_i) = h_i^s \quad (i = 1, \dots, m) \quad (3.4)$$

$$e^s(c_j) = v_j^s \quad (j = 1, \dots, n) \quad (3.5)$$

where  $e^s(v)$  is the number of edges in  $E^s$  that are incident to vertex  $v$ .

In other words one has to partition  $E$  into subsets  $E^s$  having specified degrees. It is in fact an edge coloring problem in  $G = K_{m,n}$ .

The above model suggests that one could also try to formulate  $RP(m, n, k, \mathcal{H}, \mathcal{V})$  as a vertex coloring problem. This can indeed be done if we construct another type of graph.

Each entry  $(i, j)$  of  $A$  is associated with a vertex  $(i, j)$ . We link every vertex  $(i, j)$  to the vertices corresponding to the entries  $(i, j+1), (i-1, j), (i, j-1), (i+1, j)$  if they exist in  $A$ . We obtain a graph  $\hat{G} = (V, E)$  with  $|V| = mn$  and  $|E| = 2mn - (m+n)$ . Such a graph is called an  $(m \times n)$  grid graph. Let  $H_i = \{(i, j) | 1 \leq j \leq n\}$  and  $V_j = \{(i, j) | 1 \leq i \leq m\}$ . Notice that the vertices in  $H_i$  (resp.  $V_j$ ) form a chain. We can formulate  $RP(m, n, k, \mathcal{H}, \mathcal{V})$  in the following way.

#### Model grid

Given an  $(m \times n)$  grid graph  $G = (V, E)$ , a collection of maximal chains  $H_i$  ( $i = 1, \dots, m$ ) and  $V_j$  ( $j = 1, \dots, n$ ), find a partition of  $V$  into  $k$  subsets  $W^1, \dots, W^k$  such that for  $1 \leq s \leq k$ :

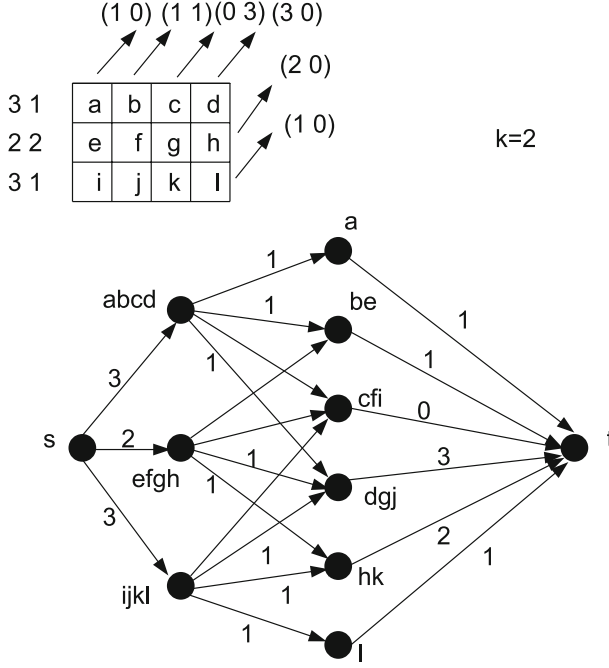
$$|W^s \cap H_i| = h_i^s \quad (i = 1, \dots, m) \quad (3.6)$$

$$|W^s \cap V_j| = v_j^s \quad (j = 1, \dots, n) \quad (3.7)$$

Both models will give rise to some extensions and variations that will be discussed in the next sections.

### 4 Network flow models for some extensions of $RP(m, n, k = 2, \mathcal{H}, \mathcal{V})$

We intend to consider here an extension of the basic problem where the projection conditions are replaced by more general requirements. It will turn out that for  $k = 2$



**Fig. 2** Non-orthogonal projections

we can still solve the problem in polynomial time. For the case  $k = 2$  network flow models are often used to find the assignment of color 1.

One should observe that the projections given by  $\mathcal{H}, \mathcal{V}$  (horizontal and vertical) in  $RP(m, n, k = 2, \mathcal{H}, \mathcal{V})$  may be considered in a more general form. In fact we may have any two families  $\mathcal{T}, \mathcal{U}$  of lines (a line being a subset of entries of  $A$ ) such that each one of  $\mathcal{T}, \mathcal{U}$  is a partition of the entries of  $A$ . If  $\mathcal{T}, \mathcal{U}$  correspond to projections in some given directions, then for any  $T_i \in \mathcal{T}$  and any  $U_j \in \mathcal{U}$  we shall have  $|T_i \cap U_j| \leq 1$ .

For instance we may take  $\mathcal{T} = \mathcal{H}$  (the rows of  $A$ ) and for  $\mathcal{U}$  we take the lines parallel to a “diagonal” of  $A$ . An example of such a problem is given in Fig. 2 where a network flow model is exhibited with a flow giving a feasible assignment of color 1.

Let us consider another extension of  $RP(m, n, k = 2, \mathcal{H}, \mathcal{V})$ : in addition to the usual data we are given disjoint subsets  $S_*^1, S_*^2$  of entries with the requirement that, for  $s = 1, 2$ , all entries in  $S_*^s$  have color  $s$ .

For this problem with preassignments  $\mathcal{H}, \mathcal{V}$  always contain the updated values after coloring the preassigned entries. This image reconstruction problem with preassignments is denoted by  $RP(m, n, k = 2, \mathcal{H}, \mathcal{V}, S_*)$ . Deleting in the network the arcs  $(i, j)$  corresponding to the entries in  $S_*^s, s = 1, 2$  and searching a maximum flow in the remaining network will give us a solution to  $RP(m, n, k = 2, \mathcal{H}, \mathcal{V}, S_*)$ .

Since in practice, the goal of tomography is to obtain the reconstruction of an unknown original image and not just to find any solution that has the given projections, additional prior knowledge must be used. In [Batenburg \(2007\)](#), the weighted reconstruction problem is introduced: with each entry  $(i, j)$  we associate a positive





We shall assume that the preassignments satisfy some obvious necessary conditions: for  $1 \leq s \leq k$ ,  $S_*^s$  has at most  $h_i^s$  entries in row  $i$  ( $1 \leq i \leq m$ ) and at most  $v_j^s$  entries in column  $j$  ( $1 \leq j \leq n$ ).

For the reconstruction problem with preassignments, the boundary between easy and difficult problems is easily located, as shown by the next results.

**Proposition 5.2**  *$RP(m, n, k = 3, \mathcal{H}, \mathcal{V}, S_*)$  is NP-complete.*

*Proof* In order to derive this result we shall use a reduction from the following problem:

Given a bipartite graph  $\widehat{G} = (V, E)$  and two integers  $p, q$  does there exist in  $\widehat{G}$  two disjoint matchings  $M_p, M_q$  such that  $|M_p| = p, |M_q| = q$ ?

This problem called 2-MATCH in [Gabow et al. \(1985\)](#) has been shown to be NP-complete (even if  $M_p$  is a perfect matching).

Assume we have a bipartite graph  $\widehat{G} = (R^- \cup C^-, E)$ . We call  $R^-$  and  $C^-$  the two sets of vertices and let  $m = |R^-|, n = |C^-|$ .

We introduce into  $\widehat{G}$  all edges that were missing between  $R^-$  and  $C^-$ . Let  $S_*^3$  be the set of these edges. These are the edges that will be constrained to have color 3 in the Model Bipartite corresponding to our problem.

We introduce two vertices  $c_{n+1}, c_{n+2}$  linked to all vertices in  $R^-$  and two vertices  $r_{m+1}, r_{m+2}$  linked to all vertices in  $C^-$ .

We now set  $R = R^- \cup \{r_{m+1}, r_{m+2}\}$   $C = C^- \cup \{c_{n+1}, c_{n+2}\}$  and we link  $r_{m+1}, r_{m+2}$  to  $c_{n+1}, c_{n+2}$ : these four edges are introduced into  $S_*^3$ .

The new graph  $G$  is now the complete bipartite graph  $K_{m+2, n+2}$ .

We set

$$\begin{aligned}
 h_i^1 &= 1 \quad h_i^2 = 1 \quad (i = 1, \dots, m) \\
 h_{m+1}^1 &= n - p \quad h_{m+1}^2 = 0 \\
 h_{m+2}^1 &= 0 \quad h_{m+2}^2 = n - q \\
 v_j^1 &= 1 \quad v_j^2 = 1 \quad (j = 1, \dots, n) \\
 v_{n+1}^1 &= m - p \quad v_{n+1}^2 = 0 \\
 v_{n+2}^1 &= 0 \quad v_{n+2}^2 = m - q
 \end{aligned} \tag{5.1}$$

For every vertex  $r_i$  we set  $h_i^3 = n + 2 - h_i^1 - h_i^2 \geq 0$  and for every  $c_j$ , we set  $v_j^3 = m + 2 - v_j^1 - v_j^2 \geq 0$ . Furthermore we set  $S_*^1 = S_*^2 = \emptyset$ .

We claim that there is a solution to 2-MATCH in  $\widehat{G}$  if and only if the problem  $RP(m + 2, n + 2, k = 3, \mathcal{H}, \mathcal{V}, S_*)$  defined above has a solution.

Assume 2-MATCH has a solution consisting of matchings  $M_p, M_q$ : we give color 1 to the entries  $(i, j)$  of  $A$  corresponding to the edges  $[r_i, c_j]$  of  $M_p$  (there are  $p$  such edges, with  $i \leq m, j \leq n$ ) and to the  $m - p$  entries  $(i, n + 1)$  for which  $r_i$  is not adjacent to an edge of  $M_p$ . Also the  $n - p$  entries  $(m + 1, j)$  for which  $c_j$  is not adjacent to an edge of  $M_p$  are given color 1. The color 2 is given to the entries  $(i, j)$  corresponding to edges of  $M_q$ . In addition we give also color 2 to the  $m - q$  entries  $(i, n + 2)$  for which  $r_i$  is not adjacent to an edge of  $M_q$  and also to the  $n - q$  entries  $(m + 2, j)$  for which  $c_j$  is not adjacent to an edge of  $M_q$ . Then all remaining entries

of  $A$  get color 3. This assignment of colors satisfies (3.4) and (3.5) with the values given in (5.1) as can be verified easily.

Conversely if we have a solution of the problem  $RP(m+2, n+2, k=3, \mathcal{H}, \mathcal{V}, S_*)$ , then the entries  $(i, j)$  with color 1 that have  $i \leq m$  and  $j \leq n$  define a matching  $M$  in  $\widehat{G}$ . Its cardinality  $|M|$  is  $p$  (the  $m-p$  entries  $(i, n+1)$  and the  $n-p$  entries  $(m+1, j)$  with color 1 correspond to vertices  $r_i$  and  $c_j$  of  $\widehat{G}$  not adjacent to edges of  $M$ ). So we have  $M_p$ .

In the same way we get  $M_q$  by taking the edges  $[r_i, c_j]$  corresponding to entries  $(i, j)$  of  $A$  with color 2 and  $i \leq m, j \leq n$ .

Since all entries in  $S_*^3$  have color 3, we have used edges in  $\widehat{G}$  to construct  $M_p$  and  $M_q$ . So we have a solution of 2-MATCH.  $\square$

**Remark 5.3** In fact 2-MATCH remains  $NP$ -complete if the maximum degree  $\Delta(G)$  of the bipartite graph  $G$  is 3. So we may strengthen Proposition 5.2 by stating that  $RP(m, n, k=3, \mathcal{H}, \mathcal{V}, S_*)$  is  $NP$ -complete even if in each line (row or column) of  $A$  there are at most three entries that are not in  $S_*$ .

Before exhibiting a solvable case with preassignments  $S_*$  we need to state a preliminary result. We recall the necessary and sufficient conditions of Hall (see Berge 1983) for a matching saturating all vertices in  $X$  to exist in a bipartite graph  $G = (X, Y, E)$ : for every subset  $A \subseteq X$  of vertices, the set  $N(A)$  of neighbors of vertices in  $A$  satisfies  $|N(A)| \geq |A|$ . Clearly these conditions do hold if  $\min_{x \in X} d_G(x) \geq \max_{y \in Y} d_G(y)$ . They will be used in the next Lemma.

**Lemma 5.4** Assume  $G = (R^- \cup C^-, E)$  is a bipartite graph with  $d_G(r_i) = r$  for each  $r_i$  in  $R^-$  and  $d_G(v_j) = t$  for each  $v_j$  in  $C^-$ . Then 2-MATCH is polynomially solvable in  $G$ .

*Proof* If  $r = t = 1$ , these matchings exist if and only if  $p + q \leq m$ . Assume that  $r \geq 2$  and  $t = 1$ . We can find a maximum matching  $M_1$  with  $|M_1| = m$  (we have  $m \leq n$  since  $mr = nt$  and the conditions of Hall are satisfied in  $G$ ). In this case, one can find disjoint matchings  $M_p, M_q$  if  $\max(p, q) \leq m$ .

Suppose now that  $r = t \geq 2$ . Clearly if  $r = t$  we have  $m = n$  and since  $G$  is regular, we can find two disjoint maximum matchings  $M, M'$  with  $|M| = |M'| = m$ . There is a solution if and only if  $\max(p, q) \leq m$ .

Assume now that  $r > t \geq 2$ . As before we can find a maximum matching  $M$  in  $G$  with  $|M| = m$ . Then after removal of  $M$ , all degrees in  $R^-$  are  $r - 1 \geq t$ . Since all degrees in  $C^-$  are at most  $t$ , the conditions of Hall do still hold and another maximum matching  $M'$  can be found with  $|M'| = m$  and we continue as above.  $\square$

**Proposition 5.5**  $RP(m+2, n+2, k=3, \mathcal{H}, \mathcal{V}, S_*)$  can be solved in polynomial time if we have the following:

- (a)  $h_i^1 = h_i^2 = 1 \ i \leq m, v_j^1 = v_j^2 = 1 \ j \leq n$ ;
- (b)  $h_{m+1}^1 = n - p, h_{m+1}^2 = 0, v_{n+1}^1 = m - p, v_{n+1}^2 = 0,$   
 $h_{m+2}^1 = 0, h_{m+2}^2 = n - q, v_{n+2}^1 = 0, v_{n+2}^2 = m - q$ ;

- (c)  $S_* = (S_*^1, S_*^2, S_*^3)$  where  $S_*^3$  contains entries  $(m+1, n+1)$ ,  $(m+1, n+2)$ ,  $(m+2, n+1)$ ,  $(m+2, n+2)$  and in addition  $S_*^1 \cup S_*^2 \cup S_*^3$  contains  $n-r$  entries in each row  $i \leq m$  and  $m-t$  entries in each column  $j \leq n$  (where  $r \leq n$  and  $t \leq m$  are given integers).

*Proof* This follows by observing that this reconstruction problem will have a solution if and only if one can find a solution to 2-MATCH in the bipartite graph associated to the rows  $i$  with  $i \leq m$  and the columns  $j$  with  $j \leq n$  in the array  $A$ . According to Lemma 5.4, 2-MATCH in this case can be solved in polynomial time.  $\square$

We shall next examine some special cases of  $RP(m, n, k = 3, \mathcal{H}, \mathcal{V})$  for which solutions can be obtained in polynomial time. Among the special cases let us mention the following result.

**Proposition 5.6** (Costa et al. 2005) *Assume that in a reconstruction problem  $RP(m, n, k = 3, \mathcal{H}, \mathcal{V})$  the number  $q$  of multiple occurrences of color 1 in the lines (rows and columns) is bounded by a constant  $t$ , i.e.,  $q = \sum_{i: h_i^1 > 1} h_i^1 + \sum_{j: v_j^1 > 1} v_j^1 \leq t$ . Then  $RP(m, n, k = 3, \mathcal{H}, \mathcal{V})$  can be solved in polynomial time.*

Let us now say that a color  $s$  is *semi-unary* in  $A$  if either  $h_i^s \leq 1$  for all rows  $i \leq m$  or  $v_j^s \leq 1$  for all columns  $j \leq n$ . With simple alternating chain arguments one can show the following.

**Proposition 5.7** (Costa et al. 2006a) *Assume that in a reconstruction problem  $RP(m, n, k = 3, \mathcal{H}, \mathcal{V})$  colors 1 and 2 are semi-unary. Then the problem can be solved in polynomial time.*

To refer to the timetabling application of Sect. 2 this would correspond to a timetabling problem with two teachers and one multi-teacher.

Finally another special case with  $k = 3$  has been exhibited: it assumes the existence of *binary colors*, i.e., colors  $s$  for which  $h_i^s, v_j^s \in \{0, 2\}$  for every row  $i$  and every column  $j$ .

**Proposition 5.8** (Costa et al. 2006a) *If in  $RP(m, n, k = 3, \mathcal{H}, \mathcal{V})$  colors 1 and 2 are binary, the problem can be solved in polynomial time.*

We skip the proof consisting in the examination of a long collection of cases in which the problem has no solution.

However we shall discuss another case in which the problem can be solved in polynomial time by giving a sufficient condition for a solution to exist. The condition involves the largest degree  $p$  in  $E^{12} = E^1 \cup E^2$ . We shall assume that  $p \geq 2$  since the case  $p = 1$  is trivial.

**Proposition 5.9** (Bentz et al. 2007) *In a complete bipartite graph  $G = K_{X,Y}$  let  $p = \max_{i \in X, j \in Y} \{h_i^1 + h_i^2, v_j^1 + v_j^2\} \geq 2$ . Then  $RP(m, n, k = 3, \mathcal{H}, \mathcal{V})$  has a solution, i.e., there exists a partition  $E^1, E^2, E^3$  of  $E$  satisfying (3.4) and (3.5) if  $|E^{12}| \geq 2p(p-2) + 3$ .*

*Proof* From [Ryser \(1957\)](#), we know how to separately construct  $E^1$  and  $E^2$ . If there are no cycles of length 2, then we are done.  $E^1$  and  $E^2$  are disjoint and the remaining (uncolored) edges will necessarily belong to  $E^3$ .

Otherwise we have at least one cycle of length 2,  $[x, y]_1, [x, y]_2$ , where  $x \in X, y \in Y, [x, y]_1 \in E^1$  and  $[x, y]_2 \in E^2$ . If we can find an edge  $[z, t] \in E^{12}$  ( $z \in X, t \in Y$ ) of color  $s$  such that  $[x, t], [z, y] \notin E^{12}$ , then by replacing  $[x, y]_s$  and  $[z, t]$  by  $[x, t]$  and  $[z, y]$  that get color  $s$ , conditions (3.4) and (3.5) are still satisfied and we have at least one less cycle of length 2. By repeating this procedure while there is a cycle of length 2, we shall finish by getting 2 disjoint edge sets  $E^1, E^2$  satisfying (3.4) and (3.5) and thus we get a solution of our problem.

Let us now show in which case we can always find an edge  $[z, t] \in E^{12}$  such that  $[x, t], [z, y] \notin E^{12}$ . Such an edge will be called a ‘good’ edge. Notice that  $x$  and  $y$  are considered as linked by two edges. Clearly all edges having as endvertices  $x$  or  $y$  are not ‘good’. We have at most  $2(p-1)$  such edges. Furthermore, all edges  $[u, v] \in E^{12}$  such that  $[x, v]$  or  $[u, y]$  belongs to  $E^{12}$  are not ‘good’. There are at most  $2(p-2)(p-1)$  such edges. Clearly every other edge in  $E^{12}$  not belonging to these two sets will be a ‘good’ edge. Thus if we require that  $E^{12}$  contains at least  $2(p-1) + 2(p-2)(p-1) + 1 = 2p(p-2) + 3$  edges, there will always be a ‘good’ edge and hence all cycles of length 2 can be replaced by two disjoint edges.  $\square$

More solvable cases of  $RP(m, n, k = 3, \mathcal{H}, \mathcal{V})$  where one specifies the structure of  $E^1, E^2$  (for instance each  $E^s$  is a Hamiltonian cycle on  $V(E^s)$ ,  $s = 1, 2$ ) can be found in [Bentz et al. \(2007\)](#).

## 6 An extension of $RP(m, n, k, \mathcal{H}, \mathcal{V})$

The first extension of our basic problem that we shall consider in this section is based on Model Grid where  $G$  is a grid graph and each  $H_i$  (resp. each  $V_j$ ) is a maximal horizontal (resp. vertical) chain in  $G$ . We shall relax the statement that  $G$  is a grid graph and consider more generally that a family  $\mathcal{P}^* = (P_i | i \in I)$  of chains  $P_i$  in  $G$  is given. Each  $P_i$  will sometimes for commodity be identified with its vertex set  $V(P_i)$  whenever no confusion arises. We shall also associate with each  $P_i$  in  $\mathcal{P}^*$  a vector  $h(P_i) = (h_i^1, \dots, h_i^k)$  such that  $\sum_{s=1}^k h_i^s = |P_i|$  for each  $i \in I$ , where  $|P_i|$  is the number of vertices in  $P_i$ . We now have to find a partition  $W^1, \dots, W^k$  of  $V(G)$  such that  $|P_i \cap W^s| = h_i^s$  for all  $i \in I$  and all  $s \leq k$ . This problem will be called  $RP(G, k, \mathcal{P}^*, \mathcal{H})$  where  $\mathcal{H}$  is the collection of values  $h_i^s$  associated to the chains  $P_i$  in  $\mathcal{P}^*$ .

As an application of  $RP(G, k, \mathcal{P}^*, \mathcal{H})$  let us mention the following problem consisting in scheduling the refurbishment of the stations in a city subway network. The network is represented by a graph  $G = (V, E)$  where the vertices are the stations. Each metro line is associated with a chain  $P_i$  in  $\mathcal{P}^*$ . Assuming that the renovation operation of every single station takes one month, we want to schedule these operations while taking into account the following requirements: in month  $s$ , the number of stations in metro line  $P_i$  that will be closed for renovation is  $h_i^s$ . The problem of assigning a date (month) for the renovation of every station with the above constraints is precisely  $RP(G, k, \mathcal{P}^*, \mathcal{H})$  if the whole refurbishment has to take place within a period of  $k$

months. In some cases, it is desired to avoid closing two consecutive stations along the same metro line. The assignment of dates is then a proper  $k$ -coloring of the underlying graph  $G$  (adjacent vertices get different colors) and the problem is denoted by  $RP^*(G, k, \mathcal{P}^*, \mathcal{H})$ .

In addition to the above mentioned application, our problem may be viewed in a different context related to constraint satisfaction in logic. Essentially we are given a collection of  $n$  Boolean variables as well as a collection of clauses  $P_i$  (each one of them involves a subset of the Boolean variables). It is required to find an assignment of values True or False to each Boolean variable in such a way that in each clause  $P_i$  the number of variables with value False is exactly (or at most) a given number  $h_i^F$ . Notice that here we have a number  $k$  of colors that is  $k = 2$ . The general  $k$ -coloring case would then correspond to  $k$ -valued logical variables.

We shall be giving some special cases where a solution can be found in polynomial time.

We need some more definitions and notations for  $\mathcal{P}^*$ . For a family  $\mathcal{P}^* = (P_i \mid i = 1, \dots, p)$  of subsets  $P_i$  of a set  $V$ , we call *cover index* of  $\mathcal{P}^*$  and denote by  $c(\mathcal{P}^*)$  the maximum number of members of  $\mathcal{P}$  that may cover a single element of  $V$  (i.e., that have a non-empty intersection).

For instance in the basic reconstruction problem of discrete tomography we have  $c(\mathcal{P}) = 2$ .

A family  $\mathcal{P}^* = (P_i \mid i = 1, \dots, p)$  of subsets  $P_i$  of a set  $V$  is called *nested* if for any  $P_i, P_f \in \mathcal{P}^*$ , we have either  $P_i \subseteq P_f$  or  $P_f \subseteq P_i$  or  $P_i \cap P_f = \emptyset$ .

Consider now a partition of  $\mathcal{P}^*$  into nested families. One can look for a partition into the smallest possible number of nested families. This number, denoted by  $\text{Nest}(\mathcal{P}^*)$ , is called the *nesticity* of  $\mathcal{P}^*$ .

**Lemma 6.1** (Hansen and de Werra 1997) *One can determine in polynomial time if for a family  $\mathcal{P}^*$  we have  $\text{Nest}(\mathcal{P}^*) \leq 2$ .*

*Proof* Assign a vertex to each  $P_i \in \mathcal{P}^*$  and link by an edge  $P_i$  and  $P_f$  whenever  $P_i \cap P_f \neq \emptyset$ ,  $P_i \not\subseteq P_f$  and  $P_f \not\subseteq P_i$ . The resulting graph is bipartite if and only if  $\text{Nest}(\mathcal{P}^*) \leq 2$ .  $\square$

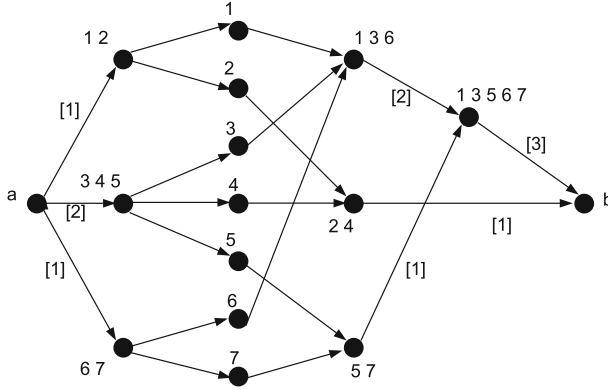
Observe that  $c(\mathcal{P}^*)$  and  $\text{Nest}(\mathcal{P}^*)$  are unrelated: we may have  $c(\mathcal{P}^*) > \text{Nest}(\mathcal{P}^*)$  or  $c(\mathcal{P}^*) < \text{Nest}(\mathcal{P}^*)$ . In fact, for  $\mathcal{P}^* = (\{a, b\}, \{a, c\}, \{b, c\})$ , we have  $c(\mathcal{P}^*) = 2$ ,  $\text{Nest}(\mathcal{P}^*) = 3$  and for  $\mathcal{P}' = (\{a, b, c\}, \{a, b\})$ , we have  $c(\mathcal{P}') = 2$ ,  $\text{Nest}(\mathcal{P}') = 1$ .

We recall that the basic image reconstruction problem in discrete tomography is polynomially solvable for  $k = 2$  when the  $P_i$ 's are the rows and the columns of the associated grid graph  $G$ . Remember that in this special case we have  $c(\mathcal{P}^*) = 2$ .

More generally, we can state:

**Proposition 6.2** (Bentz et al. 2008)  *$RP(G, k = 2, \mathcal{P}^*, \mathcal{H})$  is polynomially solvable if  $c(\mathcal{P}^*) = 2$ .*

*Proof* We construct a multigraph  $G'$  as follows: assign a vertex  $P_i$  to each chain  $P_i$  in  $\mathcal{P}^*$ . Each vertex of  $G$ , that is in  $P_i$  and in  $P_f$  is represented by an edge in  $G'$  between  $P_i$  and  $P_f$ . Each vertex that is covered by a unique  $P_i$  is associated to an edge in



**Fig. 4** The network associated with a family  $\mathcal{P}^*$  with  $\text{Nest}(\mathcal{P}^*) = 2$

$G'$  between vertex  $P_i$  and a new vertex  $P'_i$ . So there is a one-to-one correspondence between the vertices of  $G$  and the edges of  $G'$ .

Then a solution, if there is one, will correspond to a subset  $F$  of edges of  $G'$  such that for each vertex  $P_i$ ,  $F$  has  $h_i^1$  edges adjacent to  $P_i$  (there is no restriction for the vertices  $P'_i$ ).

In  $G'$ , the edges of  $F$  will give  $W^1$  in  $G$  and the edges not in  $F$  will correspond to  $W^2$  in  $G$ . There are polynomial algorithms (see [Lovasz and Plummer 1986](#)) to construct such subsets  $F$  if they exist or to decide that there is no solution.  $\square$

One can derive the following from results in [Hansen and de Werra \(1997\)](#).

**Proposition 6.3** ([Bentz et al. 2008](#))  $RP(G, k = 2, \mathcal{P}^*, \mathcal{H})$  is polynomially solvable if  $\text{Nest}(\mathcal{P}^*) = 2$ .

*Proof* Starting from the inclusion tree of each one of the two nested families covering  $\mathcal{P}^*$ , one can build as shown in [Hansen and de Werra \(1997\)](#) a network flow model where a compatible integral flow will define the subset  $W^1 \subseteq V$  and  $W^2 = V - W^1$  will be obtained immediately.

Assume  $\mathcal{P}^*$  can be decomposed into nested subfamilies  $A$  and  $B$ . We represent both families by the inclusion tree of their subsets  $P_i$ . A source  $a$  (resp. a sink  $b$ ) is linked to all maximal (inclusionwise) subsets of  $A$  (resp.  $B$ ). We link each  $l \in V$  to the unique minimal subset  $A_r$  of  $A$  (resp.  $B_s$  of  $B$ ) that contains  $l$  by an arc  $(A_r, l)$  (resp.  $(l, B_s)$ ). The network is obtained by orienting all remaining edges from  $a$  to  $b$ . The arc entering (resp. leaving) each  $P_i$  in  $A$  (resp.  $B$ ) has a capacity and a lower bound of flow equal to  $h_i^1$ . The arcs adjacent to the vertices corresponding to the elements of  $V$  have capacity 1 and a lower bound of flow equal to 0.

In Fig. 4 an example is given for a set  $V = \{1, 2, \dots, 7\}$  and a family  $\mathcal{P}^*$  with  $\text{Nest}(\mathcal{P}^*) = 2$ . Here  $A = (\{1, 2\}, \{3, 4, 5\}, \{6, 7\})$  and  $B = (\{1, 3, 6\}, \{2, 4\}, \{5, 7\}, \{1, 3, 5, 6, 7\})$ . The values  $h_i^1$  are shown in brackets.

There is a one-to-one correspondence between the feasible integral flows from  $a$  to  $b$  and the subset  $W^1$  of vertices in a coloring  $(W^1, W^2)$  satisfying the requirements.  $\square$

**Proposition 6.4** (Bentz et al. 2008) *Let  $G$  be an arbitrary graph and  $\mathcal{P}^*$  a family of chains  $P_i$  such that any  $P_i$  has at most two vertices belonging to some other chains of  $\mathcal{P}^*$ . Then  $RP(G, k = 2, \mathcal{P}^*, \mathcal{H})$  can be solved in polynomial time.*

*Proof* We shall transform the problem into a 2-SAT problem which is known to be polynomially solvable (see Aspvall et al. 1979).

We associate a binary variable  $x$  to every vertex of  $G$  that belongs to at least two chains  $P_i$ . Notice that we may assume that  $\min\{h_i^1, h_i^2\} \geq 1$ ,  $i \leq p$ , otherwise there is only one color occurring in  $P_i$  and the problem can be reduced. We first remove all vertices that belong to exactly one  $P_i$  (these will be considered later). Now each  $P_i$  contains one or two vertices. For each  $P_i$  that has exactly two vertices, say  $x, y$ , that belong to other chains, we write a clause  $c_i$  as follows. If  $h_i^1 = 2, h_i^2 = 1$ , we set  $c_i = x \vee y$  (this means that at least one of the vertices  $x, y$  must have color one) and if  $h_i^1 = 2, h_i^2 = 2$ , we set  $c_i = \bar{x} \vee \bar{y}$  (at least one of  $x, y$  must have color 2). If  $\min\{h_i^1, h_i^2\} \geq 2$ , we do nothing (since  $x$  and  $y$  can get any color). Finally when  $h_i^1 = h_i^2 = 1$ , we introduce a constraint  $x = \bar{y}$  (because  $x$  and  $y$  must get different colors). For any  $P_i$  that has exactly one vertex belonging to more than one chain in  $\mathcal{P}^*$ , we do nothing since by assumption ( $\min\{h_i^1, h_i^2\} \geq 1$ ) this vertex can have any color. We define  $\mathcal{C} = \bigwedge$  and using the equality constraints  $x = \bar{y}$  we may substitute variable  $y$  to variable  $x$ . We are left with a 2-SAT instance. It has a solution if and only if  $RP(G, k = 2, \mathcal{P}^*, \mathcal{H})$  has a solution.

From a solution of 2-SAT, we derive a partition  $W^1, W^2$  of the vertices associated to the binary variables. The bicoloring  $W^1, W^2$  of the vertices of  $G$  belonging to more than one chain of  $\mathcal{P}^*$  is given by  $W^1 = \{v \mid v \text{ is true}\}$ ,  $W^2 = \{v \mid v \text{ is false}\}$ . For each  $P_i$  it is possible to assign color 1 or 2 to the yet uncolored vertices so that the number of occurrences of color  $s$  is  $h_i^s$  (for  $s = 1, 2$ ). This will provide the required coloring of  $G$ .

Conversely if  $RP(G, k = 2, \mathcal{P}^*, \mathcal{H})$  has a solution, then by setting  $x = \text{true}$  (resp.  $x = \text{false}$ ) for all variables corresponding to the vertices  $x$  that are in more than one chain and have color 1 (resp. color 2), we shall satisfy all clauses in  $\mathcal{C}$  (as well as the equality constraints).  $\square$

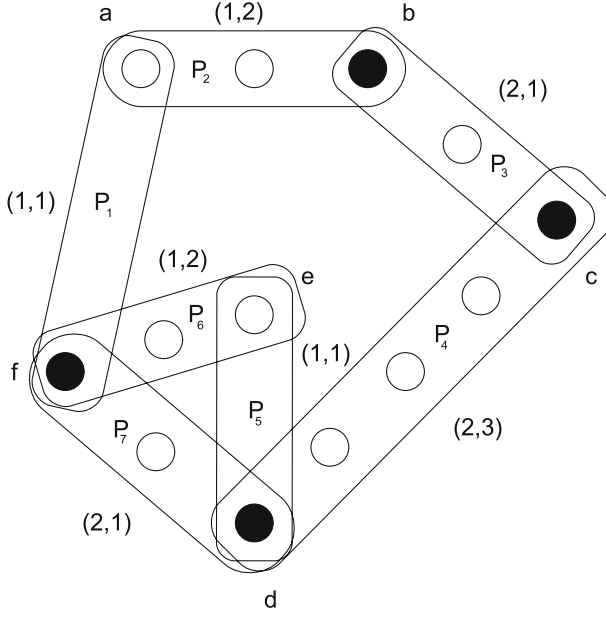
An example of this construction is given in Fig. 5 where we introduce the following clauses:  $a = \bar{f}, \bar{a} \vee \bar{b}, b \vee c, e = \bar{d}, \bar{e} \vee \bar{f}, d \vee f$ . Thus we get  $\mathcal{C} = (\bar{a} \vee \bar{b}) \wedge (b \vee c) \wedge (\bar{e} \vee \bar{f}) \wedge (d \vee f) = (\bar{a} \vee \bar{b}) \wedge (b \vee c) \wedge (d \vee a) \wedge (d \vee \bar{a})$  and a solution is given by  $W^1 = \{b, c, d, f\}$  and  $W^2 = \{a, e\}$ .

**Remark 6.5** The above results are in some sense best possible since it has been shown in Bentz et al. (2008) that  $RP(G, k = 2, \mathcal{P}^*, \mathcal{H})$  is  $NP$ -complete when  $\text{Nest}(\mathcal{P}^*) = 3$  or  $c(\mathcal{P}^*) = 3$ .

We shall finally mention another situation with an arbitrary number of colors where  $RP(G, k, \mathcal{P}^*, \mathcal{H})$  is polynomially solvable.

**Proposition 6.6** (Bentz et al. 2008) *For any graph  $G$  and any  $\mathcal{P}^*$  such that every  $|P_i| \leq 2$ ,  $RP(G, k, \mathcal{P}^*, \mathcal{H})$  can be solved in polynomial time.*





**Fig. 5** Transformation into 2-SAT

*Proof* Consider  $RP(G, k, \mathcal{P}^*, \mathcal{H})$ . Eliminate all  $P_i$ 's such that  $h_i^s = |P_i| \leq 2$  for some color  $s \leq k$  (these have a unique coloring) and apply the reductions implied by these eliminations.

Consider a pair  $P_i, P_j$  with  $|P_i \cap P_j| = 1$ . For each  $P_l$  let  $\Pi_l$  be the set of colors  $s$  with  $h_l^s > 0$ . If  $\Pi_i \cap \Pi_j = \emptyset$ , there is no solution. If  $|\Pi_i \cap \Pi_j| = 1$ , then assign this color to the vertex in  $P_i \cap P_j$  and the rest of  $P_i, P_j$  is also determined. We apply these reductions until either we get a contradiction or we have a collection of connected components  $C_1, \dots, C_r$  where in each connected component all  $P_i$ 's have the same set  $\Pi_i$  of possible colors (remember that  $|P_i| = 2$  and  $|\Pi_i| = 2$ ). Then our problem has a solution if and only if every connected component is bipartite.  $\square$

One should also mention that when  $|P_i| = 3, h_i^1 = h_i^2 = h_i^3 = 1$  for each  $P_i$  in  $\mathcal{P}^*$ ,  $RP(G, k = 3, \mathcal{P}^*, \mathcal{H})$  is  $NP$ -complete even if  $c(\mathcal{P}^*) = 2$  (see [Bentz et al. 2008](#)).

This can be seen by using a transformation from edge three-coloring of a three-regular graph  $G'$  (this problem is known to be  $NP$ -complete ([Holyer 1981](#))): we construct a graph  $G$  and a family  $\mathcal{P}^*$  of chains in  $G$ . We shall associate a chain  $P_i$  in  $G$  to each vertex  $w_i$  of  $G'$ . Each edge  $[w_i, w_f]$  of  $G'$  is associated with a vertex  $v_{if} \equiv v_{fi}$  of  $V(G)$ .  $P_i$  will be a chain in  $G$  containing the three vertices corresponding to the three edges adjacent to  $w_i$  in  $G'$ . If in  $G'$  vertex  $w_i$  is adjacent to  $w_r, w_u, w_t$  ( $r < u < t$ ) then in  $G$ ,  $P_i = \{v_{ir}, v_{iu}, v_{it}\}$  and the corresponding chain will be formed by edges  $[v_{ir}, v_{iu}], [v_{iu}, v_{it}]$ . We set  $h_i^s = 1$  for  $i = 1, \dots, p$  and  $s = 1, 2, 3$ . Then there is an edge three-coloring of  $G'$  if and only if there is a partition  $W^1, W^2, W^3$  of  $V(G)$  such that for each  $P_i, |P_i \cap W^s| = 1 = h_i^s$  for any  $i, s$ .

## 7 Constraints on unions of colors

We shall now examine in Model Bipartite (edge partitioning in the complete bipartite graph  $G = (R \cup C, E) = K_{m,n}$ ) a situation where in a partition  $E^1, E^2, \dots, E^k$  of  $E$  we have additional requirements besides (3.4) and (3.5).

In general no requirement is imposed on the structure of the graphs generated by  $E^s$  or by  $E^{st} \equiv E^s \cup E^t$  besides satisfying (3.4) and (3.5). Here we shall first examine some variations where the union of some subsets  $E^s$  has to satisfy additional constraints. We shall focus on these subsets and we shall not care about the other subsets corresponding to the remaining colors.

Let us observe that from the first two constraints of (2.1) we see that there are indeed  $k - 1$  independent colors. The last one, say color  $k$ , will be the ground color (the number of its occurrences in each row and in each column is entirely determined by the occurrences of the first colors  $1, \dots, k - 1$ ). Here we shall assume  $k = 3$ , so we shall have to determine disjoint sets  $E^1, E^2$  and  $E^3 = E - (E^1 \cup E^2)$  will be automatically determined and it will satisfy (3.4) and (3.5).

We shall examine in particular the situation where the color classes  $E^1, E^2$  form together a tree or a collection of disjoint chains.

For all these problems we shall also examine the corresponding problem in the case in which instead of having an underlying graph  $G$  that is bipartite (as was  $K_{X,Y}$ ) we have a complete graph  $G = K_X$  on  $|X| = m$  vertices. So we are given for each vertex  $i$  in  $G$  and each color  $s$  a non negative integer  $h_i^s$ . Our problem then consists in finding a partition of the edge set  $E(G)$  of  $G = K_X$  into  $k$  subsets  $E^1, \dots, E^k$  such that for each color  $s$  we have

$$h_i^s \text{ is the number } e^s(i) \text{ of edges of } E^s \text{ adjacent to vertex } i \text{ in } X \quad (7.1)$$

Clearly for a solution to exist the following conditions must hold.

$$\sum_{s=1}^k h_i^s = m - 1 \quad (i = 1, \dots, m) \quad (7.2)$$

$$\sum_{i=1}^m h_i^s \text{ is even} \quad (s = 1, \dots, k) \quad (7.3)$$

For the rest of the section, when working in complete graphs, we assume that conditions (7.2) and (7.3) are satisfied.

The first problem that we consider can be formulated as follows: We assume that  $k = 3$ . Given a complete bipartite graph  $G = K_{X,Y}$  with values  $h_i^1, h_i^2, v_j^1, v_j^2$  (for  $i = 1, \dots, m; j = 1, \dots, n$ ) find two disjoint subsets  $E^1, E^2$  of edges of  $E(G)$  such that (3.4) and (3.5) hold for  $s = 1, 2$  and in addition  $E^{12}$  is a tree.

We assume without loss of generality that every vertex in  $G$  will be adjacent to some edge of  $E^{12}$  (otherwise we simply delete the vertices not adjacent to any edge

of  $E^{12}$  and consider the remaining graph). This assumption can be stated as

$$V(E^1) \cup V(E^2) = V(G) \quad (7.4)$$

Also observe that  $h_i^1 + h_i^2$  (resp.  $v_j^1 + v_j^2$ ) will be the degree of vertex  $i$  in  $X$  (resp. vertex  $j$  in  $Y$ ) in the tree  $E^{12}$ .

To avoid dealing with trivial cases, we shall assume that our problem is not degenerate so that each one of colors 1,2 occurs on at least one edge.

We shall first state two lemmas of [Bentz et al. \(2007\)](#) that will be repeatedly used to construct the required subset  $E^{12}$  of edges by reducing the number of connected components.

**Recoloring Lemma** *Let  $C_1, C_2$  be two connected components of  $E^{12}$  satisfying (3.4) and (3.5) and such that  $C_2$  contains at least one cycle.*

*Assume one can find an edge  $[x_1, y_1]$  in  $C_1$  and an edge  $[x_2, y_2]$  belonging to some cycle  $C$  of  $C_2$  such that  $[x_1, y_1]$  and  $[x_2, y_2]$  have the same color (both are in  $E^1$  or both are in  $E^2$ ).*

*Then by replacing  $[x_1, y_1], [x_2, y_2]$  by  $[x_1, y_2], [x_2, y_1]$  and by giving them the same color as the removed edges we get a single connected component  $C'$  that still satisfies (3.4) and (3.5).*

*Proof* Since  $[x_1, y_2], [x_2, y_1]$  get the same color as  $[x_1, y_1], [x_2, y_2]$ , (3.4) and (3.5) are still satisfied. One verifies that  $x_2$  and  $y_2$  are still connected in  $C - [x_2, y_2]$ . Furthermore in  $C'$ ,  $x_1$  and  $y_1$  are connected by a chain consisting of edge  $[x_1, y_2]$  followed by  $C - [x_2, y_2]$  and by edge  $[x_2, y_1]$ . It follows that there are chains between any two vertices of  $C'$ . Notice furthermore that  $C'$  is still bipartite if  $C_1$  and  $C_2$  were bipartite.  $\square$

**Recycling Lemma** *Assume we have a connected component  $C$  of  $E^{12}$  containing some cycle  $\bar{C}$  and let  $e$  be an edge of  $C$  not contained in any cycle. If there is a chain  $\tilde{C}$  in  $C$  containing  $e$  and starting with some edge  $[x_2, y_2]$  in  $\bar{C}$  and ending with an edge  $[x_1, y_1] \neq e$  in  $C - \bar{C}$  with the same color as  $[x_2, y_2]$ , then one may exchange edges so that (3.4) and (3.5) still hold and  $e$  is on a cycle.*

*Proof* Notice that  $[x_1, y_2], [x_2, y_1]$  are not in  $E^{12}$  (otherwise  $e$  would be in a cycle). Replacing  $[x_1, y_1], [x_2, y_2]$  by  $[x_1, y_2], [x_2, y_1]$  and assigning them the same color as  $[x_1, y_1], [x_2, y_2]$  gives another connected component. It can be checked that there is a cycle  $C'$  (possibly of length 2) containing  $e$  that goes either through  $[x_1, y_2]$  or through  $[x_2, y_1]$ .  $\square$

**Proposition 7.1** ([Bentz et al. 2007](#)) *In  $G = K_{X,Y}$  there exist two disjoint subsets  $E^1, E^2$  of edges such that (3.4) and (3.5) hold and  $E^{12}$  is a tree if and only if:*

- (a)  $\sum_{i \in X} (h_i^1 + h_i^2) = \sum_{j \in Y} (v_j^1 + v_j^2) = (m + n - 1);$
- (b)  $\sum_i h_i^s = \sum_j v_j^s \leq (m_s + n_s - 1)$  for  $s = 1, 2$  where  $m_s$  (resp.  $n_s$ ) is the number of vertices  $i$  in  $X$  (resp.  $j$  in  $Y$ ) with  $h_i^s > 0$  (resp.  $v_j^s > 0$ ) for  $s = 1, 2$ .

The proof uses repeatedly the recoloring and recycling lemmas to construct the required set  $E^{12}$ .

Let us now briefly study the case in which  $E^{12}$  is a collection of vertex disjoint chains in  $K_{X,Y}$ . We are given values  $h_i^s, v_j^s$  that satisfy

$$1 \leq h_i^1 + h_i^2 \leq 2 \quad \text{for each } i \text{ in } X \quad (7.5)$$

$$1 \leq v_j^1 + v_j^2 \leq 2 \quad \text{for each } j \text{ in } Y \quad (7.6)$$

Here  $E^{12}$  will have to consist of a collection of elementary open chains having their endvertices at vertices  $r$  (resp.  $u$ ) with  $h_r^1 + h_r^2 = 1$  (resp.  $v_u^1 + v_u^2 = 1$ ). These will be called odd vertices. Clearly we must have an even positive number of odd vertices for the existence of a solution.

Notice that we exclude cycles in a solution, i.e., we have to show that we only have open chains.

**Proposition 7.2** (Bentz et al. 2007) *In a complete bipartite graph  $K_{X,Y}$  there exist subsets  $E^1, E^2$  of edges of  $E(G)$  satisfying (3.4) and (3.5) and such that  $E^{12}$  is a collection of elementary open chains if and only if:*

- (a) *for each color  $s$ , there is at least one vertex that has to be adjacent to exactly one edge of color  $s$ ;*
- (b) *there exists a vertex  $i \in X$  with  $h_i^1 + h_i^2 = 1$  or a vertex  $j \in Y$  with  $v_j^1 + v_j^2 = 1$ .*

As before, the proof uses the recoloring and recycling lemmas.

As mentioned similar problems can be raised for the case in which  $G$  is simply a complete graph  $K_X$  on a set  $X$  of  $m$  vertices.

Although this has no immediate connection with discrete tomography as before, we mention it for its interest in a graph theoretical context.

We now want to find a partition  $E^1, \dots, E^k$  of the edge set  $E(G)$  of  $K_X$  such that in each  $E^s$  there are exactly  $h_i^s$  edges adjacent to vertex  $i$  ( $i = 1, \dots, |X|$ ) (condition (7.1)).  $E^s$  is usually called a  $b$ -factor. Since we are in  $K_X$ , the conditions of existence are given by the Erdős Gallai theorem (existence of a simple graph with given degrees; see Chap. 6 in Berge (1983)).

Furthermore we assume, as before, that (7.4) holds, i.e., every vertex is adjacent to at least one edge of  $\bigcup_{s=1}^{k-1} E^s$ .

Are such problems easier or more difficult when we consider complete graphs? A partial answer is given in the following.

**Proposition 7.3** (Bentz et al. 2007) *The degree-constrained edge  $k$ -partitioning problem  $P'$  in a complete graph is at least as difficult as the degree-constrained edge  $k$ -partitioning problem  $P$  in a complete bipartite graph.*

*Proof* We are given a problem  $P$  defined by a complete bipartite graph  $G = K_{X,Y}$  and values  $h_i^s$  ( $i \in X$ ),  $v_j^s$  ( $j \in Y$ ) for  $1 \leq s \leq k$  satisfying (2.1). We construct a complete graph  $G' = K_{X \cup Y}$  on  $X \cup Y$  by introducing in  $G$  a clique on  $X$  and a clique on  $Y$ . Let  $m = |X|$  and  $n = |Y|$ . For each  $i \in X$  we set  $h_i'^1 = h_i^1 + m - 1$ ,  $h_i'^s = h_i^s$  ( $s = 2, \dots, k$ ) and for each  $j \in Y$  we set  $v_j' = v_j^2 + n - 1$ ,  $v_j'^s = v_j^s$  ( $1 \leq s \leq k$ ;

$s \neq 2$ ). This defines a problem  $P'$  on  $G'$ . Clearly if  $P$  has a solution  $S$ , we can derive a solution  $S'$  to  $P'$  by keeping the colors of the edges  $[x_i, y_j]$  of  $G'$ , by giving color 1 to all edges  $[x_r, x_s]$  and color 2 to all edges  $[y_u, y_v]$ . Conversely assume that  $P'$  has a solution  $S'$  in  $G'$ . Then all edges with both ends in  $X$  (resp. in  $Y$ ) have color 1 (resp. color 2): suppose an edge  $[x_i, x_j]$  has some color  $c \neq 1$ ; then  $x_i$  and  $x_j$  are adjacent to  $m - 2$  edges of color 1 (instead of  $m - 1$ ) with both ends in  $X$ ; so the number of edges of color 1 going out of  $X$  will be at least  $\sum_{i \in X} h_i^1 + 2 > \sum_{j \in Y} v_j^1 = \sum_{j \in Y} v_j'^1$  that is at least as large as the number of edges of color 1 that may have one or two ends in  $Y$ . This is impossible. For color 2, the same holds (interchanging the roles of  $X$  and  $Y$ ). Then by keeping the colors of all edges  $[x, y]$  of  $K_{X \cup Y}$ , we get a solution for  $P$  in  $K_{X, Y}$ .  $\square$

From Proposition 7.3 and from the  $NP$ -completeness of the image reconstruction problem for  $k = 4$  (see Chrobak and Dürr 2001), we obtain the following.

**Corollary 7.4** (Bentz et al. 2007) *For any fixed  $k \geq 4$ , the degree-constrained edge  $k$ -partitioning problem in a complete graph is  $NP$ -complete.*

By analogy with Proposition 5.9 we can state the corresponding result for complete graphs.

**Proposition 7.5** (Bentz et al. 2007) *In a complete graph  $G = K_X$ , let  $p = \max_{i \in X} (h_i^1 + h_i^2)$ ,  $p \geq 2$ . There exists a partition  $E^1, E^2, E^3$  of the edge set  $E$  satisfying (7.1) if  $|E^{12}| \geq p^2 - 2p + 3$ .*

The proof is very similar to the proof of Proposition 5.9.

Here we shall deal with the case in which we have  $k = 3$  colors. We can give an analogous statement to Proposition 7.1. We first consider the case in which  $E^{12}$  is a tree.

**Proposition 7.6** (Bentz et al. 2007) *In a complete graph  $G = K_X$  there exist disjoint subsets  $E^1, E^2$  of edges such that (7.1) holds for each vertex  $i$  and for  $s = 1, 2$  and  $E^{12}$  is a tree, if and only if*

- (a)  $\sum_{i \in X} (h_i^1 + h_i^2) = 2(|X| - 1)$
- (b)  $\sum_{i \in X} h_i^s \leq 2(m_s - 1)$  for  $s = 1, 2$  where  $m_s$  is the number of vertices  $i$  with  $h_i^s > 0$ .

The proof follows the same lines as the proof of Proposition 7.1 (except that we do not have to take care about the bipartite character of  $E^{12}$  when connecting different components).

We may as well consider the case in which a Hamiltonian cycle has to be constructed while taking condition (7.1) into account.

**Proposition 7.7** (Bentz et al. 2007) *Given values  $h_i^1, h_i^2$ , satisfying  $h_i^1 + h_i^2 = 2$  for each vertex  $i$  of a complete graph  $G = K_X$ , there are disjoint subsets  $E^1, E^2$  of the edge set  $E(G)$  such that (7.1) holds for each vertex  $i$  and for  $s = 1, 2$  and in addition  $E^{12}$  is a Hamiltonian cycle, if and only if there exists at least one vertex with  $h_i^1 = h_i^2 = 1$ .*

## 8 Conclusion

We have investigated some graph theoretical formulations of the basic image reconstruction problem in discrete tomography.

This has led us to several variations and extensions that have some interest as graph theory problems and that are relevant since they are able to model situations occurring in other contexts like the metro station refurbishment.

Some special cases of the famous three-color problem  $RP(m, n, k = 3, \mathcal{H}, \mathcal{V})$  have been solved with polynomial time algorithms, while the complexity of the basic problem with 3 colors remains open. Besides this we have shown that various techniques of optimization could be used to tackle several variations of the basic problem.

There are many more connections between graph theory and discrete tomography. On the one hand the various requirements occurring in image reconstruction stimulate the development of generalized graph models and graph algorithms (for instance to handle connectivity and convexity constraints). On the other hand the availability of involved graph theoretical algorithms allows their use in more realistic image reconstruction problems. An example is given by an interesting use of network flows in discrete tomography (see [Batenburg 2007](#)). Furthermore the use of graphs like switching graphs may help to get a better insight on the space of feasible solutions to an image reconstruction problem (see [Kaneko and Nagahama 2006](#)).

Most of the problems discussed in this tutorial could also be formulated in complete graphs (not necessarily bipartite) and in addition one might consider oriented versions of these problems. We are thus moving far away from our starting point in discrete tomography but the research avenue is wide open and there are a lot of fascinating topics to explore there.

## References

- Ahuja RK, Magnanti TL, Orlin JB (1993) Network flows. Prentice-Hall, Englewood Cliffs
- Alfandari L, Lemalade JL, Nagih A, Plateau G (2008) A MIP flow model for crop-rotation planning in a context of forest sustainable development (submitted)
- Alpers A, Rodek L, Poulsen HF, Knudsen E, Herman GT (2007) Discrete tomography for generating maps of polycrystals. In: Herman GT, Kuba A (eds). Advances in discrete tomography and its applications. Birkhauser, Boston, pp 271–301
- Aspvall B, Plass MF, Tarjan R (1979) A linear-time algorithm for testing the truth of certain quantified Boolean formulas. Inform Process Lett 8:121–123
- Batenburg KJ (2007) Network flow algorithms for discrete tomography. In: Herman GT, Kuba A (eds) Advances in discrete tomography and its applications. Birkhauser, Boston, pp 175–207
- Baumann J, Kiss Z, Krimmel S, Kuba A, Nagy A, Rodek L, Schillinger B, Stephan J (2007) Discrete tomography methods for nondestructive testing. In: Herman GT, Kuba A (eds) Advances in discrete tomography and its applications. Birkhauser, Boston, pp 303–332
- Bentz C, Costa M-C, de Werra D, Picouleau C, Ries B (2008) On a graph coloring problem arising from discrete tomography. Networks (to appear)
- Bentz C, Costa M-C, de Werra D, Picouleau C, Ries B (2007) Degree-constrained edge partitioning in graphs arising from discrete tomography. Technical Report ORWP 07/03
- Berge C (1983) Graphes. Gauthier-Villars, Paris
- Brocchi S, Frosini A, Picouleau C (2007) Reconstruction of binary matrices under fixed size neighborhood constraints (submitted)
- Chrobak M, Dürr C (2001) Reconstructing polyatomic structures from X-rays: NP-completeness proof for three atoms. Theor Comput Sci 259(1):81–98

- Costa M-C, de Werra D, Picouleau C (2006a) Using graphs for some discrete tomography problems. *Discrete Appl Math* 154:35–46
- Costa M-C, de Werra D, Picouleau C, Ries B (2006b) Graph coloring with cardinality constraints on the neighborhood. Technical report ORWP 06/08
- Costa M-C, de Werra D, Picouleau C, Schindl D (2005) A solvable case of image reconstruction in discrete tomography. *Discrete Appl Math* 148:240–245
- Déroche G (1986) Guy Dupuy: sculpteur discret. *Horizons d'Argonne* 52:104–105
- Déroche G (2003) Tomographie agricole des vallées de l'Aisne et de l'Aire. *Horizons d'Argonne* 80:17–20
- Di Gesù V, Kuba A (eds) (2005) Special issue: IW CIA 2003, Ninth International Workshop on Combinatorial Image Analysis. *Discrete Appl Math* 151:3
- Even S, Itai A, Shamir A (1976) On the complexity of timetable and multicommodity flow problems. *SIAM J Comput* 5:691–703
- Gabow H, Nishizeki T, Kariv O, Leven D, Tereda O (1985) Algorithms for edge-coloring. Technical report 41/85, Tel Aviv University
- Gardner RJ (2006) *Geometric Tomography*, 2nd edn. Cambridge University Press, New York, NY
- Garey M, Johnson DS (1979) *Computer and intractability*. Freeman, San Francisco
- Hansen P, de Werra D (1997) Nesticity, DIMACS Series. *Discrete Math Theor Comput Sci* 37:225–232
- Herman GT, Kuba A (1999b) Discrete tomography: a historical overview. In: Herman GT, Kuba A (eds) *Discrete tomography: foundations, algorithms and applications*. Birkhauser, Boston, pp 3–34
- Herman GT, Kuba A (eds) (1999a) *Discrete tomography: foundations, algorithms and applications*. Birkhauser, Boston
- Herman GT, Kuba A (eds) (2007) *Advances in discrete tomography and its applications*. Birkhauser, Boston
- Holyer I (1981)  $NP$ -completeness of edge-coloring. *SIAM J Comput* 10:718–720
- Kaneko A, Nagahama R (2006) Reconstruction algorithm and switching graph for two-projection tomography with prohibited subregion. In: *Proceedings of the 13th international conference on discrete geometry for computer imagery*, Szeged, Hungary, pp 110–121
- Lovasz L, Plummer M (1986) *Matching theory*. North Holland, USA
- Martinis R, Socco LV, Sambuelli L, Nicolotti G, Schmitt O, Bucur V (2004) Tomographie ultrasonore pour les arbres sur pied. *Ann Forest Sci* 61:157–162
- Ryser HJ (1957) Combinatorial properties of matrices of zeros and ones. *Can J Math* 9:371–377