Reconstructing hv-convex multi-coloured polyominoes

Adam Bains* Therese Biedl*

Technical report CS-2009-33

Abstract

In this paper, we consider the problem of reconstructing polyominoes from information about the thickness in vertical and horizontal directions. We focus on the case where there are multiple disjoint polyominoes (of different colours) that are hv-convex, i.e., any intersection with a horizontal or vertical line is contiguous. We show that reconstruction of such polyominoes is polynomial if the number of colours is constant, but NP-hard for an unbounded number of colours.

1 Introduction

The field of discrete tomography concerns reconstruction of objects given information about the thickness of the object in various projections. See the books by Herrman and Kuba [5, 6] for an extensive overview of this exciting field with many applications in medical imaging.

One special case is when the object to be reconstructed is a binary matrix with m rows and n columns, and the given information are the row and column-sums of the matrix. Testing whether such a matrix exists and finding it can be done easily with flow-methods. However, of more interest is the case when the object is supposed to be a *polyomino*, i.e., from every *black cell* (an entry of the matrix that is 1) to every other black cell, there exists a path along black cells that are adjacent horizontally or vertically.

Reconstructing polyominoes is NP-hard, even if all black cells within each row are contiguous (the polyomino is *h-convex*) or all black cells within each column are contiguous (the polyomino is *v-convex*.) Surprisingly, if the polyomino must *hv-convex* (i.e., both h-convex and v-convex,) reconstructing it from row- and column-sums becomes polynomial. See Chapter 7 of [5] for references and an overview of these results.

We study here reconstruction of objects that are the union of multiple disjoint objects, each of which has a different colour. This has applications in the reconstruction of polyatomic crystals: the number of atoms of each kind in a projection can be determined using a high-resolution transmission electron microscope. See [7, 9] for details. This problem also appears in a recreational puzzle called "Color Pic-a-pix"; see www.conceptispuzzles.com. The general problem (with no restriction on the shape of the objects) was proved to be NP-hard even for 3

*David R. Cheriton School of Computer Science, University of Waterloo, Waterloo, ON N2L 3G1, Canada, e-mail {abains,biedl}@uwaterloo.ca. Supported by NSERC.

colours [3], and very recently even for 2 colours [4]. Since the NP-hard cases for single-colour transfer, the only case that could possibly be polynomial is the case of multiple hv-convex polyominoes.

We resolve this case here, and hence study the following problem: Given C colours $\{1, \ldots, C\}$, and C sets of density-vectors (h_i^c) and (v_j^c) for $i = 1, \ldots, m$, $j = 1, \ldots, n$ and $c = 1, \ldots, C$, do there exist C binary matrices $(x_{i,j}^c)$ such that for each c the matrix $(x_{i,j}^c)$ is an hv-convex polyomino with row-sums and column-sums (h_i^c) and (v_j^c), and such that for any i, j we have $\sum_c x_{i,j}^c \leq 1$? We call this the *C-colour hv-convex polyomino reconstruction problem.*

We show that this problem is polynominal if the number of colours is a constant, but becomes NP-hard if the number of colours is unbounded.

2 Few colours

The natural approach to reconstruct C hv-convex polyominoes is to take one of the existing algorithms to reconstruct a single hv-convex polyomino and modify it so that it handles multiple polyominoes and ensures that they are disjoint. This can in fact be done easily with the algorithm by Barcucci et al. [1], and yields an algorithm for the C-colour reconstruction that takes time $O(C^2 m^{2C+2} n^{2C+2})$. We will not give the details of this, since with a different approach the time complexity can be improved significantly.

There are faster algorithms for single-colour hv-convex polyomino reconstruction, and we tried to generalize the currently fastest known, which is by Chrobak and Dürr [2] and takes $O(\min\{m, n\}^2 mn)$ time. We did not succeed to generalize this algorithm to multiple colours. The main difficulty is that this algorithm stores the computed polyomino implicitly (by storing the "blank area" around it), and hence there is no easy way to add a constraint to ensure that multiple polyominoes are disjoint.

In this paper, we first develop a different algorithm for single-coloured hv-convex polyomino reconstruction, which matches the run-time of Chrobak and Dürr. We then show that it can be generalized to multiple colours easily, yielding a run-time of $O(C^2 \min\{m, n\}^{2C} mn)$.

2.1 Single-colour reconstruction

We first explore the single-colour hv-convex polyomino reconstruction. So assume we are given vectors (h_i) and (v_j) and we want to find a binary matrix $(x_{i,j})$ that is an hv-convex polyomino and has row-sums (h_i) and column-sums (v_j). Note that necessarily $\sum_i h_i = \sum_j v_j$, since otherwise no solution can exist. We also assume $h_i > 0$ and $v_j > 0$. As before, we say that cell (i, j) is *black* if and only if $x_{i,j} = 1$ and *white* otherwise.

A *foot* of a polyomino is the intersection of the polyomino with the leftmost/rightmost column or top/bottom row, and each of the four feet is named after the meridial directions. For our algorithm, assume that $m \leq n$ after possible rotation. Try every possible *west-foot and east-foot*, i.e., all possible indices $w, e \in \{1, \ldots, m\}$. [1] We say that a polyomino *respects these feet* if cells $(w, 1)$ and (e, n) are black. We now show how to find a polyomino that

[1] Using feet and the "spine" S_{WE} is inspired by [1].

respects these feet, if one exists, in $O(nm)$ time using 2-SAT. We give the algorithm only for the case where $w \leq e$; the other case is similar.

Barcucci et al. [1] used as one of their main ingredients that in any row between w and e, they can find cells that are guaranteed to be in any polyomino respecting the feet. More precisely, let S_{WE} be the set of all cells (i,j) with $w \leq i \leq e$, $\sum_{k=1}^{j} v_k \geq \sum_{k=1}^{i-1} h_k$, and $\sum_{k=1}^{i} h_k \geq \sum_{k=1}^{j-1} v_k$. Set S_W to be the first h_w cells in row w and let S_E to be the last h_e cells in row e. Exactly as in [1], one can show that any cell in $S_W \cup S_E \cup S_{WE}$ must be black in any polyomino that respects the west and east foot. Furthermore, the cells in $S_W \cup S_E \cup S_{WE}$ form a polyomino that contains at least one cell in every column.

Now define a 2-SAT instance. We have two variables $x_{i,j}$ and $R_{i,j}$ for every cell, where $x_{i,j} = TRUE$ means that cell (i,j) is black, and $R_{i,j} = TRUE$ means that cell (i,j) is to the right of the polyomino, i.e., it and all cells to its right are white.[2] For every column j, let α_j be such that (α_j, j) is in $S_{WE} \cup S_W \cup S_E$; recall that at least one such cell must exist for any j, and it must be black in any polyomino respecting the feet.

Add the following clauses for all $j = 1, \ldots, n$:[3]

$$\begin{array}{rl} x_{i,j} = TRUE & \text{for } i = \alpha_j \\ x_{i,j} = FALSE & \text{for } i \notin [\alpha_j - v_j + 1, \alpha_j + v_j - 1] \\ x_{i,j} \Leftrightarrow \overline{x_{i+v_j,j}} & \text{for } i \in [\alpha_j - v_j + 1, \alpha_j - 1] \\ x_{i,j} \Rightarrow x_{i+1,j} & \text{for } i \in [\alpha_j - v_j, \alpha_j - 2] \\ x_{i,j} \Rightarrow x_{i-1,j} & \text{for } i \in [\alpha_j + 2, \alpha_j + v_j] \end{array}$$

One can easily verify that these clauses ensure that $x_{i,j}$ is true for exactly v_j cells in column j, and these cells are contiguous. For each i and j, also add the clauses

$$R_{i,j} \Rightarrow R_{i,j+1} \text{ and } R_{i,j} \Rightarrow \overline{x_{i,j}} \text{ and } x_{i,j} \Rightarrow R_{i,j+h_i},$$

which ensures that $R_{i,j}$ describes indeed the white region to the right of the polyomino in row i, and row i contains at most h_i cells for which $x_{i,j}$ is true.

If this 2-SAT instance has a solution, then define a cell to be black if and only if $x_{i,j}$ is true. Then the total number of cells that are black is exactly $\sum_j v_j$ and at most $\sum_i h_i$, which implies equality since $\sum_i h_i = \sum_j v_j$. Since $S_W \cup S_E \cup S_{WE}$ was connected and every column is contiguous, the resulting polyomino is also connected, and hence the desired reconstruction.

Computing the set S_{WE}, building the 2-SAT instance, and solving it can be done in $O(mn)$ time. Trying this for the $O(\min\{m,n\}^2)$ possible foot configurations yields the answer to the reconstruction problem in $O(\min\{m,n\}^2 mn)$ time, which matches the run-time by Chrobak and Dürr [2].

2.2 Fast multi-coloured reconstruction

Our single-colour algorithm easily generalizes to multiple colours. For each colour c, find the leftmost/rightmost column with $v_j^c > 0$. Choose an east and west foot for the polyomino

[2] This is loosely inspired by the variables for white corner regions used by Chrobak and Dürr [2].

[3] To ease notation we allow here indices of variables to be outside $[1 \ldots m] \times [1 \ldots n]$; any clause containing such variables can be omitted.

of colour c in these columns and build the 2-SAT for colour c, using variables $x^c_{i,j}$ and $R^c_{i,j}$. Combine all these 2-SAT instances for all C colours into one 2-SAT instance with $O(Cnm)$ variables and $O(Cnm)$ clauses.

To ensure that no two colours use a same cell, add the exclusion clause

$$x^{c_1}_{i,j} \Rightarrow \overline{x^{c_2}_{i,j}} \quad \text{for all colours } c_1 \neq c_2.$$

This adds $O(C^2mn)$ clauses. Hence the time to build and solve the 2-SAT instance for one fixed set of foot configurations is $O(C^2mn)$.

Each colour has at most m^2 possible foot configurations (assuming $m \leq n$ after possible configurations), so the total number of combinations of foot configurations is $\min\{m,n\}^{2C}$. Running 2-SAT for each of them hence computes disjoint hv-convex polyominoes, if they exist, in $O(\min\{m,n\}^{2C}mnC^2)$ time.

Theorem 1 *The C-colour reconstruction problem for hv-convex polyominoes can be solved in time $O(\min\{m,n\}^{2C}mnC^2)$.*

2.3 Variants

The variables $R_{i,j}$ for the "white region to the right" turn out to be useful for two variants of the C-colour reconstruction problem. The first variant concerns ordered reconstruction. As defined, a C-colour reconstruction problem gives the row/column-sums as an unordered set (one for each colour.) A variant would be to give an ordered set, such that all cells of colour c_1 must be to the left of all cells of colour c_2, etc.

We can solve this variant easily. Whenever colour c_1 must be left of colour c_2 in row i, add the clause $x^{c_2}_{i,j} \Rightarrow R^{c_1}_{i,j}$ for all j. So if the cell has colour c_2, then it must be in the right region with respect to colour c_1, so all cells of colour c_1 must be to its left as desired. Similarly we can add clauses for columns, after defining another variable (say $B^c_{i,j}$) that expresses that a cell is below all cells of colour c.

If the order of colours is total, and we add this clause only for consecutive colours, then we can even drop the exclusion clauses (which are then always satisfied), which reduces the run time of our algorithm to $O(\min\{m,n\}^{2C}mnC)$.

A second variant concerns the shape of the union of all coloured polyominoes. The current setup does not put any restriction on this shape. But if we wanted the union to be h-convex, we could simply add the clause $x^{c_1}_{i,j} \Rightarrow R^{c_2}_{i,j+h_i}$, for any i, j, c_1, c_2, where $h_i = \sum_c h^c_i$. Row i then can contain at most h_i cells of any colour, which ensures that coloured cells are contiguous. Similarly (again after adding variables $B^c_{i,j}$) we can ensure v-convexity of the union.

3 NP-hardness results

The previous section gave a polynomial algorithm for multicolour reconstruction for a constant number C of colours. The dependency on C is exponential, and as such it is of no surprise that if C is not constant, the problem becomes NP-complete. We prove this now.

Surely C-colour reconstruction is in NP. To see that it is NP-hard, we use a reduction from 3-SAT. We use three gadgets: a transmitter gadget, a splitter gadget, and a crosser gadget (see Figure 1.) Each occurrence of a gadget has its own unique colour (and 3 colours for a splitter gadget.)

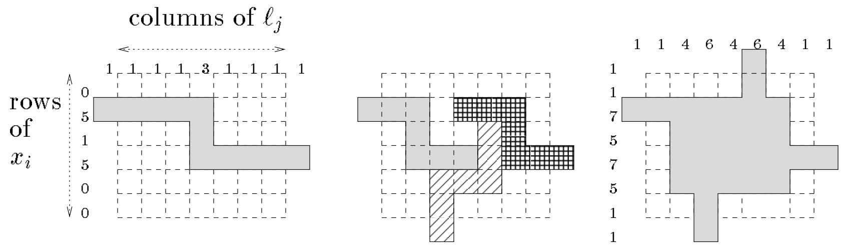

Figure 1: The transmitter gadget, the splitter gadget, and the crosser gadget. We omit the row-/column-sums for the splitter gadget to avoid confusion among the three colours.

Assume we are given an instance of 3-SAT with variables $x_1, \ldots, x_N$ and clauses $c_1, \ldots, c_M$, where each clause contains exactly three literals. We dedicate 6 rows to every variable and 7 columns for each of literal of a clause. At the place where the rows for variable x_i meets the columns of literal ℓ_j, we place a splitter gadget if ℓ_j uses x_i (i.e., if $\ell_j = x_i$ or $\ell_j = \overline{x_i}$), a crosser gadget if ℓ_j uses some x_h with $h < i$, and a transmitter gadget otherwise. See Figure 2.

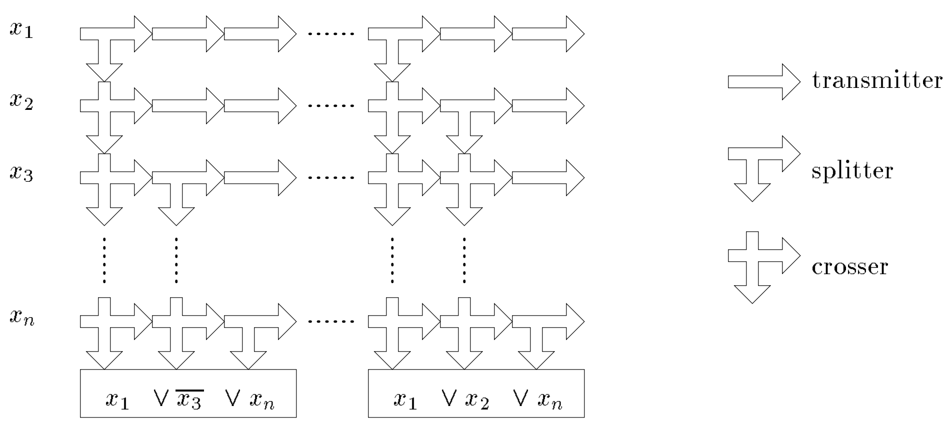

Figure 2: The overall layout.

Each gadget has some "legs" sticking out, i.e., places where it enters the region of adjacent gadgets. A transmitter gadget and a splitter gadget have two possible realizations: either the upper leg (in row 2) sticks out to the left and the lower leg (in row 4) sticks out to the right, or vice versa. For the splitter gadget, this also determines whether the leg to the bottom sticks out in column 3 or 5. A crosser gadget has four legs sticking out (in row 2 and 4 and columns 3 and 5), and four possible realizations. See Figure 3 for some alternate layouts.

Note that all gadgets within one row must either all have their upper legs to the left or all to the right; otherwise they would overlap. Given a realization of this coloured polyomino

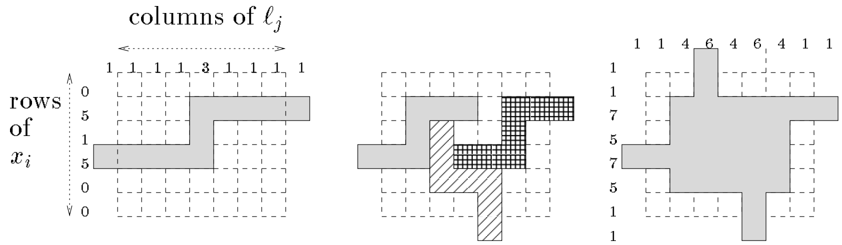

Figure 3: Alternative layouts of the transmitter gadget, splitter gadget, and crosser gadget.

instances, set variable x_i to be true if all upper legs in gadgets in the rows assigned to x_i stick out to the left. If $\ell_j = x_i$ or $\ell_j = \overline{x_i}$, then this also means that in the columns of ℓ_j, the leg of the splitter gadget (in the rows of x_i) sticks out to the bottom in column 3, and this transmits along all crosser gadgets below the rows of x_i.

Finally we define 8 different clause gadgets (depending on the negation status of the literals in the clause), and place for each clause the corresponding gadget in the columns of the clause, below all other gadgets. Figure 4 shows the clause-gadget for $\overline{x_{j_1}} \vee \overline{x_{j_2}} \vee x_{j_3}$ and $x_{j_1} \vee x_{j_2} \vee \overline{x_{j_3}}$. Generally, the column-sum vector for this gadget is

$$(\underbrace{0\ldots0}_{a_1}4\underbrace{5\ldots5}_{6-a_1}\underbrace{5\ldots5}_{a_2}7\underbrace{5\ldots5}_{6-a_2}\underbrace{5\ldots5}_{a_3}3\underbrace{0\ldots0}_{6-a_3}),$$

where $a_i = 4$ if the ith literal of the clause is positive and $a_i = 2$ otherwise. The row-sum vector is $(1, 1, W, W, W, W, W, 1, 1)$, where $W = 14 + a_3 - a_1$.

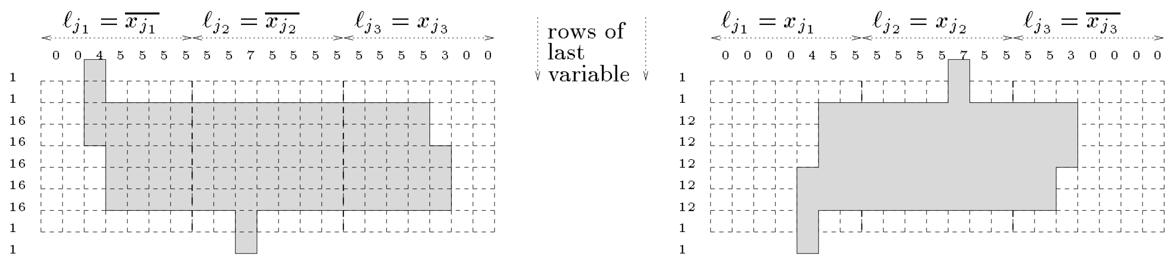

Figure 4: Two of the eight possible clause-gadgets.

The clause gadget has $W + 1$ non-zero columns and 5 rows with density W; this implies that all non-zero columns except the first and last have black cells in these rows. Therefore, the leg that sticks out at the top of the clause-gadget must occur in one of the *special columns*, which are the columns with density 4, 7, or 3. On the other hand, for each of these three special columns, there indeed exists a reconstruction of the clause-gadget where the leg is in this column; Figure 4 shows two of them.

The clause gadget has been configured such that the special columns coincide exactly with the column where the gadget in the rows above has a leg sticking out if this literal is false. So if the coloured polyomino instance can be realized, then to avoid overlap for every clause there must be at least one true literal, so 3-SAT has a solution. Similarly one shows that any 3-SAT solution gives a realization, which finishes the reduction.

Clearly the reduction is polynomial: it uses $O(N)$ rows, $O(M)$ columns, and $O(NM)$ colours.

Theorem 2 *C-colour hv-convex reconstruction is NP-complete if C is part of the input.*

4 Conclusion and open problems

In this note, we showed that reconstruction of multiple disjoint hv-convex polyominoes from their projections is polynomial if the number of colours is constant, but NP-hard if the number of colours is unbounded.

The main remaining open problem is to determine the dependency on the number of colours C. Is it possible to separate the exponentiality in C from the size of the grid, in other words, is the problem fixed-parameter tractable in the number of colours? (See for example Niedermeier [8] for an introduction to fixed-parameter tractability.) We suspect that this is the case, but have not been able to prove it; the main obstacle is that foot positions of different colours need not be in the same column, and we have not been able to bound the number of relevant foot positions in terms of C only.

Finally, are there faster algorithms for single-coloured hv-convex polyomino reconstruction? And can they be generalized to multiple colours?

References

[1] E. Barcucci, A. Del Lungo, M. Nivat, and R. Pinzani. Median of polyominoes: a property for the reconstruction. *International Journal of Imaging Systems and Technology*, 9(2-3):69–77, 1998.

[2] M. Chrobak and C. Dürr. Reconstructing hv-convex polyominoes from orthogonal projections. *Information Processing Letters*, 69:283–289, 1999.

[3] M. Chrobak and C. Dürr. Reconstructing polyatomic structures from discrete x-rays: NP-completeness proof for three atoms. *Theoretical Computer Science*, 259(1-2):81–96, 2001.

[4] C. Dürr, F. Guinez, and M. Matamala. Reconstructing 3-colored grids from horizontal and vertical projections is np-hard, 2009. arXiv:0904.3169.

[5] G.T. Herman and A. Kuba, editors. *Discrete Tomography: Foundations, Algorithms, and Applications*. Birkhäuser, 1999.

[6] G.T. Herman and A. Kuba, editors. *Advances in Discrete Tomography and Its Applications*. Birkhäuser, 2007.

[7] C. Kisielowski, P. Schwander, F.H. Baumann, M. Seibt, Y. Kim, and A. Ourmazd. An approach to quantitative high-resolution transmission electron microscopy of crystalline materials. *Ultramicroscopy*, 58(2):131–155, 1995.

[8] Rolf Niedermeier. *Invitation to Fixed-Parameter Algorithms.* Oxford University Press, 2006.

[9] P. Schwander, C. Kisielowski, M. Seibt, F.H. Baumann, Y. Kim, and A. Ourmazd. Mapping projected potential, interfacial roughness, and composition in general crystalline solids by quantitative transmission electron microscopy. *Physical Review Letters*, 71(25):4150–4153, 1993.