3–dimensional Single Active Layer Routing

András Recski and Dávid Szeszlér

Department of Computer Science and Information Theory
Budapest University of Technology and Economics, H-1521 Budapest, Hungary
{recski, szeszler}@cs.bme.hu

Abstract. Suppose that the terminals to be interconnected are situated in a rectangular area of length n and width w and the routing should be realized in a box of size $w' \times n' \times h$ over this rectangle (single active layer routing) where $w' = cw$ and $n \leq n' \leq n + 1$. We prove that it is always possible with height $h = O(n)$ and in time $t = O(n)$ for a fixed w and both estimates are best possible (as far as the order of magnitude of n is concerned). The more theoretical case when the terminals are situated in two opposite parallel planes of the box (the 3–dimensional analogue of channel routing) is also studied.

1 Introduction

Traditionally, the detailed routing phase of the design of VLSI (Very Large Scale Integrated) circuits was considered as **a 2-dimensional problem,** gradually extended to $2, 3, \ldots$ layers. Even within this problem single row routing and channel routing (where the terminals to be interconnected are on one side, or on two opposite sides, respectively, of a rectangle) are the better understood subproblems, where the inputs are essentially one-dimensional (one or two lists of terminals, of length n, also called the *length* of the channel). Here the main aim is to realize the routing, and since its 'horizontal' size is given, its 'vertical' size, or *width*, w should be minimized. An important quantity is the *density* of the problem: the maximum number d so that there exists a vertical straight line cutting d nets into two.

In case of the 2-layer Manhattan model

- (1) single row routing is always possible in $O(n)$ time with $w = d$, see [12, 17],
- (2) channel routing is **NP**-hard [18, 26],
- (3) but this latter becomes always possible in $O(nd)$ time with $w = O(d)$ if we may extend the length of the channel by introducing additional columns [3, 6, 14]. In fact, the number of these additional columns may be as large as $O(\sqrt{n})$, see [3, pp. 212–213].

In spite of (2), there are plenty of practically effective algorithms available, which can solve 'difficult' problems (of length 150...200) with width around 20 [17, 10]. In case of switchbox routing (where the terminals are on all the four sides of a rectangle – a real 2-dimensional problem) instances with length 23 and width 15...16 are already 'difficult' problems on 2 layers [5, 8, 15, 22].

If more than two layers are permitted, both channel and switchbox routing become easier, see for example [7] and [4, 25], respectively.

As technology permits more and more layers, a 'real' **3-dimensional** approach becomes reasonable. There are plenty of deep results in this area, see [1, 2, 9, 11, 13, 19, 20, 21, 23, 24], for example. Most of them embed certain 'universal-purpose' graphs (like n-permuters, n-rearrangeable permutation networks, shuffle-exchange graphs) into the 3-dimensional grid, ensuring that *pairs* of terminals can be connected, moreover, in some papers along *edge-disjoint* paths. Our result below is of much simpler structure but it allows *multiterminal nets* as well, and ensures *vertex disjoint* paths (or Steiner-trees) for the interconnections of the terminals within each net.

Throughout, except in the last section, we restrict ourselves to the *single active layer case* (all the terminals are on a single plane and the third dimension (above this plane, with *height* h) is for interconnections only). The terminals occupy certain gridpoints of an $n \times w$ rectangle. Henceforth we will use 'vertical direction' to refer to the direction of h (that is, to the direction perpendicular to this $n \times w$ rectangle) and not for the direction of w.

One can easily see even in small instances like 4×1 that a routing is usually impossible unless either the length n or the width w may be extended by introducing extra rows or columns between rows and columns of the original grid (compare with (3) above). If it is allowed to introduce both extra rows and extra columns then there is a trivial upper bound of $h = O(wn)$ for the height of the routing; see Lemma 1 below.

Our main result is that if w is fixed and n becomes large, a routing of height $O(n)$ can be attained not only in the aforementioned trivial way, but even if the new length n' satisfies $n \leq n' \leq n + 1$ and only w is extended to $w' = cw$, where c is a suitably chosen constant (we will show that $c \geq 8$ suffices). In view of Lemma 2 this linear bound is best possible. Moreover, our algorithm realizes this in $O(n)$ time, which is also essentially best possible.

Throughout this paper we are going to think of w as fixed and try to obtain bounds for the height as a function of n only.

2 Definitions and Main Results

The vertices of a given (planar) grid of size $n \times w$ are called *terminals*. A *net* N is a set of terminals. A *single active layer routing problem* is a set $\mathcal{N} = \{N_1, N_2, \ldots, N_t\}$ of pairwise disjoint nets. n and w are the *length* and the *width* of the routing problem, respectively.

By a *spacing of s_w in direction w* we are going to mean that we introduce $s_w - 1$ pieces of extra columns between every two consecutive columns (and also to the right hand side of the rightmost column) of the original grid. This way the width of the grid is extended to $w' = s_w \cdot w$. A *spacing of s_n in direction n* is defined analogously.

A *solution with a given spacing s_w and s_n* of a routing problem $\mathcal{N} = \{N_1, N_2, \ldots, N_t\}$ is a set $\mathcal{T} = \{T_1, T_2, \ldots, T_t\}$ of pairwise vertex–disjoint

Steiner–trees in the cubic grid of size $(w \cdot s_w) \times (n \cdot s_n) \times h$ (above the original planar grid containing the terminals) such that the terminal set of T_i is N_i for every $1 \leq i \leq t$. h is called the *height* of the routing.

Since w is fixed, one can also imagine the input as w *rows* of terminals (each of length n) or as a set of $\binom{w}{2}$ channel routing problems, each with length n and with a given density. Let D be the maximum of these densities. Clearly, $D \leq n$.

Theorem 1. *If $s_w \geq 8$ then for any fixed value of w and for any n a single active layer routing problem can always be solved in time $t = O(n)$ and with height $h = O(n)$ such that the length n is preserved or increased by at most one. Both linear bounds are best possible.*

Our algorithm gives $t = O(w^3 n)$ and $h = O(wD)$.

3 Straightforward Bounds

Lemma 1. *If $s_w \geq 2$ and $s_n \geq 2$ then every routing problem can be solved with height $h \leq \frac{w}{2}n$.*

Proof. We assign a separate layer to each net. For every terminal we introduce a vertical wire segment to connect the terminal with the layer of its net. The interconnection of the terminals of each net can now be performed trivially on its layer using the extra rows and columns guaranteed by the spacing in both directions.

Since 1–terminal nets can be disregarded, the number of nets is at most $\frac{1}{2}nw$ thus $h \leq \frac{w}{2}n$ follows immediately.

Lemma 2. *For any given n there exists a routing problem that cannot be solved with height h smaller than $\frac{n}{2s_w}$.*

Proof. Let, for simplicity, the width and the length be even, let $w = 2a$ and $n = 2b$. Consider the following example (the idea is very similar to those in [4, 16]). Suppose that each net consists of two terminals in central-symmetric position as shown in Figure 1.

The number of nets is an. Since each net is cut into two by the central vertical line e, any routing with width $w' = s_w w$ and height h must satisfy $w'h \geq an$. Therefore $h \geq (w/2w')n$, hence $h \geq \frac{n}{2s_w}$.

Since in the above example $D = n$, this also proves the lower bound $h = \Omega(D)$.

The straightforward lower bound for the time is the length of the input, that is, $t = \Omega(wn)$.

4 Two Simple Steps

1. Suppose at first that $w = 1$. Then what we have is essentially a single row routing problem with density d. Each net determines an interval of length at most

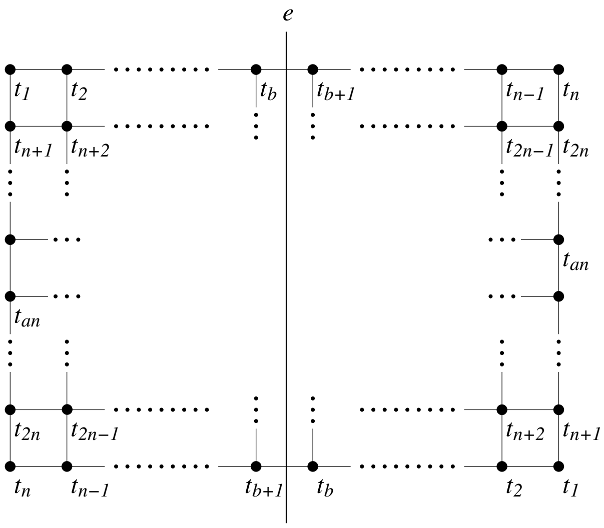

Fig. 1.

n and these intervals can be packed in a vertex-disjoint way into d parallel lines, usually called *tracks*, using the 'left-edge algorithm' [12, 17]. Using the classical 2-layer Manhattan model, we can arrange the tracks in a horizontal plane, as shown in the top of Figure 2, thus realizing a routing with $w' = d$ and $h = 2$. However, alternatively these tracks can occupy either a vertical plane, leading to $w' = 2$ and $h = d$, or two vertical planes, leading to $w' = 3$ and $h = \lceil d/2 \rceil$, see the middle and the bottom drawing of Figure 2, respectively. (Theoretically one can pack the tracks to more vertical planes and thus ensure $h = \lceil 3d/(2w') \rceil$ for larger values of w' as well but it does not seem to be interesting.) Throughout in Figures 2, 3 and 6 continuous lines denote wires while dotted lines are for the indication of coplanarity only.

Similarly, if $w = 2$ then we have a channel routing problem with density d and using the same linear time algorithm we can always realize a routing with $w' = d + 1$ and $h = 3$ or with $w' = 3$ and $h = d + 1$, see Figure 3. This method is also well-known, it dates back at least to [7].

The right hand side of Figure 3 shows the essential idea of our algorithm: immediately at the level of the terminals (when leaving the 'single active layer') we increase the width from w to $w' = s_w w$ in order to make enough space 'between the rows of the terminals' for the vertical plane(-s) containing the interconnecting wires. The same process is illustrated for $w = 6$ in Figure 5 below.

Later we shall need the following observation: if a larger horizontal area is available, we can decrease the height to $h = 2\lceil d/(w' - 2) \rceil + 2$. For this we can arrange the tracks in $\lceil d/(w' - 2) \rceil$ parallel horizontal planes – however, we need an 'empty' plane between two consecutive planes of tracks to ensure that the

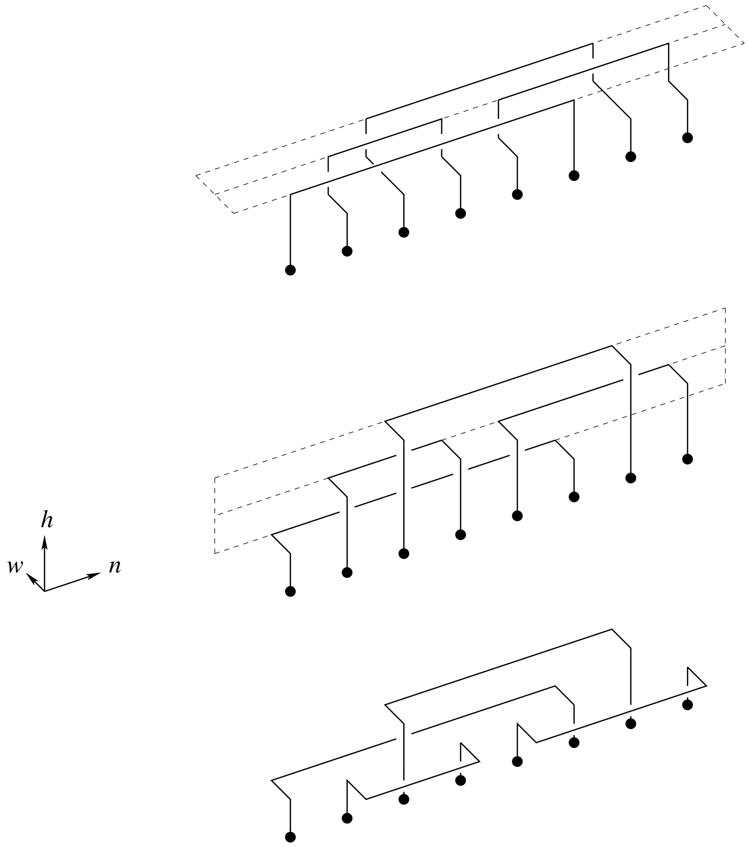

Fig. 2.

endpoints of the intervals can always reach the terminals, even if they are 'in the wrong side', as terminals t_2 and t_3 in Figure 3. Therefore the basic quantity hw' can be upper bounded by essentially $3d/2$ for $w = 1$ but only by $2d$ for $w = 2$.

2. Let us turn now to the general problem with width w. Since the terminals occupy certain gridpoints of an $n \times w$ rectangle, we consider them as a collection of w parallel *rows,* each of length n. We wish to solve $\binom{w}{2}$ channel routing problems one after the other. Figure 4 illustrates this for $w = 6$. At first (going from the bottom of the figure to the top) we solve those $w - 1$ channel routings where the rows are adjacent (first 'floor'), then those $w - 2$ ones where the distance of two rows is two (second 'floor') etc.

The $w - 1$ channel routings at the first floor (actually the $w - 1$ vertical planes containing the wire segments of these channel routings) do not interfere with one another (these are illustrated by the $w - 1$ empty dots in the bottom horizontal line of the figure).

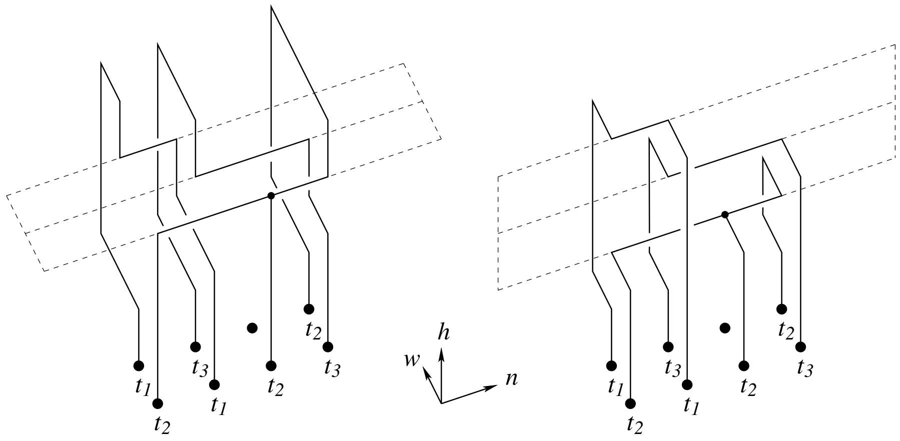

Fig. 3.

On the other hand, the $w-2$ channel routings at the second floor do not have this property hence this floor will have two levels, one for the row pairs 1 and 3, 3 and 5, etc and one for the row pairs 2 and 4, 4 and 6 etc. In general, floor f contains l_f levels where $l_f = f$ if $1 \leq f \leq \lfloor w/2 \rfloor$ and $l_f = w - f$ if $\lfloor w/2 \rfloor < f \leq w-1$. Solid dots illustrate the rows, the vertical lines in the figure show that the terminals within a row may appear at different floors. Empty dots indicate the areas where these two rows can be interconnected. Hence such an empty dot may indicate a contribution of at most D to the final height (compare with the right hand side of Figure 3).

Observe that there are two empty dots between two solid ones in the second floor, three empty dots between two solid ones in the third floor etc. Hence the total height requirement is *not* $D \times \sum_{f=1}^{w-1} l_f = O(w^2 D)$ but only $2D \times \sum_{f=1}^{w-1} \frac{1}{f} l_f = O(wD)$. The extra constant 2 is due to the necessary empty planes between the consecutive planes of tracks, as explained in the last remark in Step **1** above, concerning the empty plane between the consecutive planes of tracks.

In the introduction we mentioned that the width of the input must be extended to $w' = s_w w$. Figure 4 might give the wrong impression that $s_w = 2$ suffices. However, as we shall see in the next section, the realization of the 'crossings' in the figure requires much more space, leading to $s_w = 8$.

5 The Real Routing

For future reference we are going to introduce the following terminology. By a *w–plane* we are going to mean a plane that is perpendicular to the width of the routing, that is, to the 'vector' w of Figures 2, 3 and 6. Analogously, *h–planes* and *n–planes* are planes perpendicular to the height and the length of the routing (or to the vectors h and n), respectively. Similarly, by a *w–wire segment* we are going to mean a wire segment that is parallel to the width of the routing or

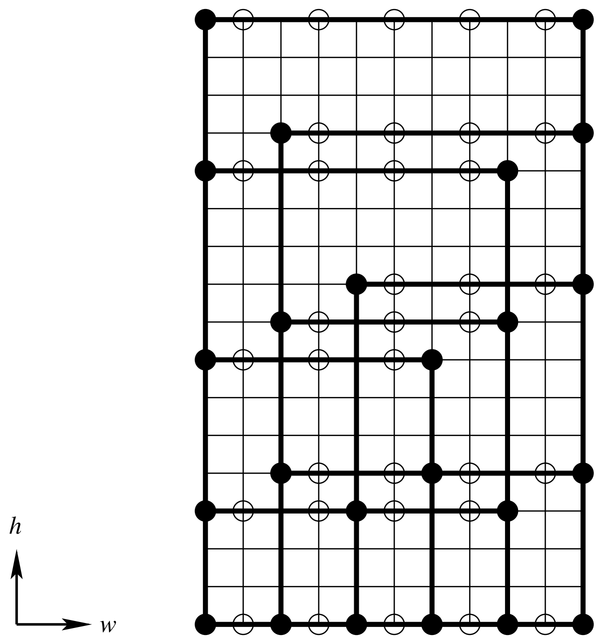

Fig. 4.

to the vector w. *h–wire segments* and *n–wire segments* are defined in the same way. (Note that for example an h–plane is a horizontal plane, while an h–wire segment is a vertical wire segment.)

In the previous section Figure 4 illustrated the order how the $\binom{w}{2}$ channel routing problems are routed one above the other. (Of course it is possible that the terminals in row i and those in row k do not share any net and therefore a whole level within a floor is missing.)

However, Figure 4 may alternatively be considered as a 'cross-section' of the routing by an n–plane of size $w' \times h$ (and then there are n copies of these cross-sections, one behind the other). In this sense an empty dot in a particular level indicates a whole w–plane of length n and of height at most $2\lceil D/f \rceil$ containing several wires one above the other for that channel routing. Hence a horizontal line in Figure 4 between a solid and an empty dot indicates a wire segment going towards this plane – but it may go for a wire segment running in this plane or it may wish to avoid it and go for a wire segment in one of the other parallel planes (that is, towards one of the further empty dots).

It is very important to realize, therefore, that there are two types of 'crossings' which have to be avoided if we wish to realize the final 3-dimensional routing along the n–planes one after the other:

– Type 1 — A vertical line, connecting two solid dots, and a horizontal line, connecting two empty dots, may cross each other in Figure 4.

– Type 2 — A horizontal line which passes through an empty dot in Figure 4 may, in fact, not use that particular n–wire (which is perpendicular to the actual n–plane).

The basic point is that crossings of Type 2 can be avoided by a detour within the actual n–plane (increasing the height by one and the width by two) but crossings of Type 1 can be avoided outside the n–plane only. Therefore we must use the adjacent n–plane as well. But what happens if in this latter n–plane there is another vertical line interconnecting two solid dots (or in the real routing: a h–wire segment coming from a terminal) blocking the detour?

We avoid this problem in the following way: Since $w' > w$, we must, in any case, start the routing in each n–plane by 'expanding' the w terminals into larger distances. This increases the height by $w/2$ and can be performed like in either of the ways shown in Figure 5. Now, if we use the two kinds of expansions alternatively then the terminals within a single row will form a zigzag pattern and hence two h–wires that 'should' block each other in two consecutive n–planes, will actually be shifted by two units away.

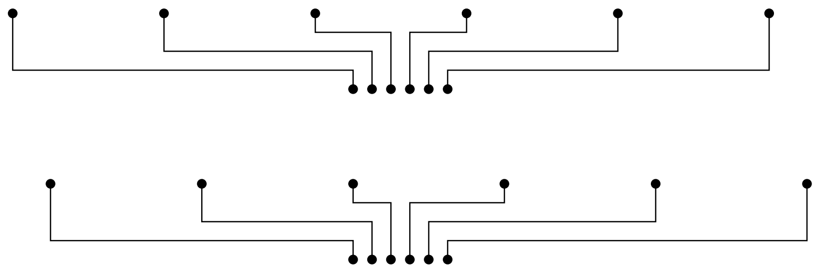

Fig. 5.

This way we have ensured that a Type 1 detour will not be blocked by an h–wire segment coming from a terminal. However, one difficulty remains to be solved: a Type 1 detour can still cross a w–wire segment that goes in the same h–plane one unit behind (connecting two w–planes within a level). In order to handle this, let us number the n–planes from 'front' to 'back' with the numbers $1, 2, \ldots, n$ and the h–planes form 'bottom' to 'top' with the numbers $1, 2, \ldots, h$. Now let us declare the following rule: if a w–wire segment goes in an n–plane numbered with an even number, then it must go in an h–plane also numbered with an even number; similarly, if a w–wire segment goes in an n–plane numbered with an odd number, then it must go in an h–plane also numbered with an odd number. Since the height of a level is always the double of the number of tracks in it, the above rule can obviously be fulfilled. (This way some of the Type 2 detours will become unnecessary: if a w–wire segment reaches a w–plane (containing some of the tracks within a level) between two consecutive tracks, then the w–wire segment can cross the w–plane without meeting the tracks, there is no need for a Type 2 detour.)

Figure 6 illustrates most of these situations in a single drawing. Recall that continuous lines are wires, dotted lines are for the indication of coplanarity only. Observe that $t_1, t_2, \ldots$ are terminals within a single row (illustrating the aforementioned zigzag pattern), while the terminals $t'_1, t'_2, \ldots$ form the next row. It might be instructive to recall that the detour between A and B is of Type 1 while that between C and D is of Type 2.

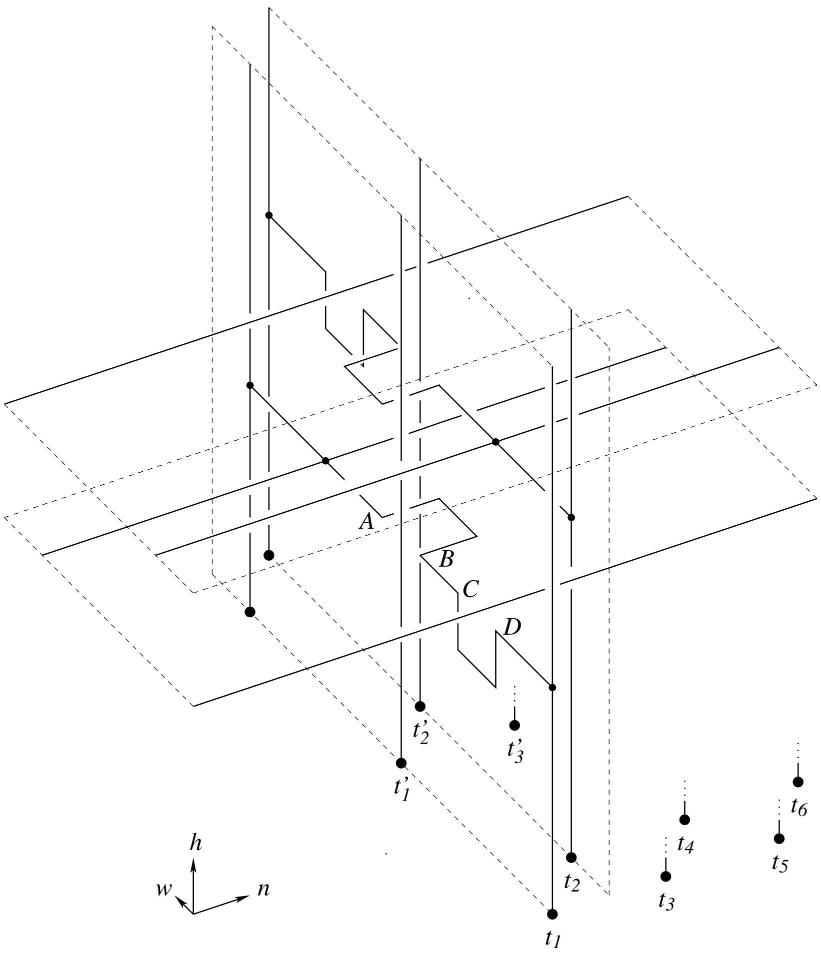

Fig. 6.

Figure 7 shows a part of Figure 6 again in order to explain why s_w needs to be as large as 8.

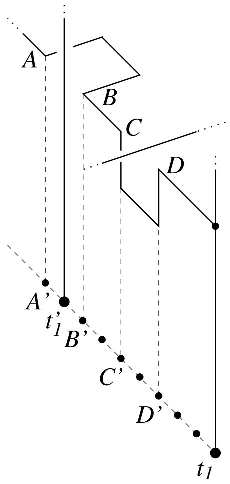

Fig. 7.

6 3–dimensional Channel Routing

Throughout this paper we have dealt with the single active layer routing problem, which is, in a sense, the 3–dimensional analogue of single row routing. However, the 3–dimensional analogue of channel routing may also be of interest not only from a technical point of view (see [11] for example), but also in a theoretical sense: in constrast to the essential difference in complexity between single row routing and channel routing in the 2–layer Manhattan model (see (1) and (2) in the Introduction), there does not seem to be such a difference between their 3–dimensional analogues. The *3–dimensional channel routing problem* is defined as two parallel rectangular planes of size $n \times w$ containing all the terminals to be interconnected (by vertex disjoint Steiner trees) in a box of size $n' \times w' \times h$ between the two parallel planes. As before, we suppose that w is fixed and n can be very large hence we allow $w' = s_w w$ but $n \leq n' \leq n+1$ only.

Theorem 2. *For any fixed value of w and for any n such a 3-dimensional channel can always be routed in $t = O(n)$ time such that the length n is preserved or increased by at most one, the width is extended to $w' = s_w w$ and the required height is $h = O(n)$. Both linear bounds are best possible.*

The lower bound is the same as in Lemma 2. The routing can be performed basically along the same line as explained in Sections 4 and 5 for the single active layer case. Just like the two 'expansions' of Figure 6 are alternating along the direction n, the expansions opposite to each other on the two parallel planes should also be shifted by two units away.

Acknowledgements: Grants No. OTKA 29772 and 30122 of the Hungarian National Science Foundation, grant No. FKFP 409/1997 of the Hungarian Ministry of Education and the NATO Collaborative Linkage Grant PST.CLG.976383 are gratefully acknowledged.

References

[1] Aggarwal, A., M. Klawe, D. Lichtenstein, N. Linial and A. Wigderson, 1991. A lower bound on the area of permutation layouts, *Algorithmica* 6, 241-255.

[2] Aggarwal, A., J. Kleinberg and D. P. Williamson, 2000. Node-disjoint paths on the mesh and a new trade-off in VLSI layout, *SIAM J. Computing* 29, 1321-1333.

[3] Baker S. B., S. N. Bhatt and F. T. Leighton, 1984. An approximation algorithm for Manhattan routing, *Advances in Computing Research* 2, 205–229.

[4] Boros, E., A. Recski and F. Wettl, 1995. Unconstrained multilayer switchbox routing, *Ann. Oper. Res.* 58, 481-491.

[5] Burstein, M. and R. Pelavin, 1983. Hierarchical wire routing, *IEEE Trans. Computer-Aided Design* CAD-2, 223-234.

[6] Burstein, M. and R. Pelavin, 1983. Hierarchical channel router, *Integration, the VLSI Journal* 1, 21-38.

[7] Chan, Wan S., 1983. A new channel routing algorithm, *in:* R. Bryant, ed. *3rd Caltech Conference on VLSI,* Comp. Sci. Press, Rockville, 117-139.

[8] Cohoon, J. P. and P. L. Heck, 1988. BEAVER: a computational-geometry-based tool for switchbox routing, *IEEE Trans. Computer-Aided Design* CAD-7, 684-697.

[9] Cutler, M. and Y. Shiloach, 1978. Permutation layout, *Networks* 8, 253-278.

[10] Deutsch, D, 1976. A dogleg channel router, *Proc. 13rd Design Automation Conf.* 425-433.

[11] Enbody, R.J., G. Lynn and K. H. Tan, 1991. Routing the 3–D chip, *Proc. 28th ACM/IEEE Design Automation Conf.* 132–137.

[12] Gallai T. His unpublished results have been announced in A. Hajnal and J. Surányi, 1958. Über die Auflösung von Graphen in vollständige Teilgraphen, *Ann. Univ. Sci. Budapest Eötvös Sect. Math.* 1, 115-123.

[13] Games, R. A., 1986. Optimal book embeddings of the FFT, Benes, and barrel shifter networks, *Algorithmica* 1, 233-250.

[14] Gao, S. and M. Kaufmann, 1994. Channel routing of multiterminal nets, *J. of the ACM* 41, 791-818.

[15] Grötschel, M., A. Martin and R. Weismantel, 1993. Routing in grid graphs by cutting planes, *in:* G. Rinaldi and L. Wolsey, eds. *Integer programming and combinatorial optimization,* 447-461.

[16] Hambrusch, S. E., 1985. Channel routing in overlap models, *IEEE Trans. Computer-Aided Design of Integrated Circ. Syst.* CAD-4, 23-30.

[17] Hashimoto, A. and J. Stevens, 1971. Wire routing by optimizing channel assignment, *Proc. 8th Design Automation Conf.* 214-224.

[18] LaPaugh A. S., 1980. A polynomial time algorithm for optimal routing around a rectangle, *Proc. 21st FOCS Symp.* 282-293.

[19] Leighton, T., 1983. *Complexity issues in VLSI: Optimal layouts for the shuffle-exchange graph and other networks,* The MIT Press, Cambridge, MA.

[20] Leighton, T. and A. L. Rosenberg, 1986. Three-dimensional circuit layouts, *SIAM J. Computing* 15, 793-813.

[21] Leighton, T., S. Rao and A. Srinivasan, 1999. New algorithmic aspects of the Local Lemma with applications to routing and partitioning, *Proc. Tenth Annual ACM-SIAM Symp. on Discrete Algorithms,* ACM/SIAM, New York and Philadelphia, 643-652.

[22] Luk, W. K., 1985. A greedy switch-box router, *Integration, the VLSI Journal* 3, 129-149.
[23] Rosenberg, A. L., 1983. Three-dimensional VLSI: a case study, *J. ACM.* 30, 397-416.
[24] Shirakawa, I., 1980. Some comments on permutation layout, *Networks* 10, 179-182.
[25] Szeszlér D., 1997. Switchbox routing in the multilayer Manhattan model, *Ann. Univ. Sci. Budapest Eötvös Sect. Math,* 40, 155–164.
[26] Szymanski, T. G., 1985. Dogleg channel routing is **NP**-complete, *IEEE Trans. Computer-Aided Design of Integrated Circ. Syst.* CAD-4, 31-41.