Deadlock Prevention by Acyclic Orientations *

Jean-Claude Bermond † **Miriam Di Ianni** ‡ **Michele Flammini** §
Stephane Perennes ¶

Abstract

Deadlock prevention for routing messages has a central role in communication networks, since it directly influences the correctness of parallel and distributed systems. In this paper we extend some of the computational results presented in [10] on acyclic orientations for the determination of optimal deadlock free routing schemes. In this context, minimizing the number of buffers needed to prevent deadlocks for a set of communication requests is related to finding an acyclic orientation of the network which minimizes the maximum number of changes of orientations on the dipaths realizing the communication requests. The corresponding value is called the rank of the set of dipaths.

We first show that the problem of minimizing the rank is NP-hard if all shortest paths between the couples of nodes wishing to communicate have to be represented and even not approximable if only one shortest path between each couple has to be represented. This last result holds even if we allow an error which is any sublinear function in the number of couples to be connected.

We then improve some of the known lower and upper bounds on the rank of all possible shortest dipaths between any couple of vertices for particular topologies, such as grids and hypercubes, and we find tight results for tori.

Keywords: computational and structural complexity, parallel algorithms, routing and communication in interconnection networks

Contact Author: Michele Flammini
Univ. of L'Aquila - Dipartimento di Matematica Pura ed Applicata
via Vetoio loc.Coppito, I-67100 L'Aquila, Italy
tel.: +39-862-433730, fax: +39-862-433180, e-mail: flammini@univaq.it

1 Introduction

Massively parallel computers with thousands of processors are considered the most promising technology to gain computational power. Large-scale multiprocessors are usually organized as

*Work supported by the EU ESPRIT Long Term Research Project ALCOM-IT under contract N. 20244, by the French action RUMEUR of the GDR PRS and by the Italian MURST 40% project "Algoritmi, Modelli di Calcolo e Strutture Informative".

†Project SLOOP I3S-CNRS/INRIA/Université de Nice–Sophia Antipolis, 930 route des Colles, F–06903 Sophia-Antipolis Cedex, France. E-mail: bermond@unice.fr

‡Dipartimento di Scienze dell'Informazione, University of Rome "La Sapienza", via Salaria 113, I-00198 Rome, Italy. E-mail: diianni@dsi.uniroma1.it

§Dipartimento di Matematica Pura ed Applicata, University of L'Aquila, via Vetoio loc.Coppito, I-67100 L'Aquila, Italy. E-mail: flammini@univaq.it

¶Project SLOOP I3S-CNRS/INRIA/Université de Nice–Sophia Antipolis, 930 route des Colles, F–06903 Sophia-Antipolis Cedex, France. E-mail: sp@unice.fr

ensembles of nodes, each having its own processor, local memory, and other supporting devices. The way nodes are connected to one another varies among machines. In a direct network architecture, each node has a point-to-point, or direct, connection to some number of other nodes. Since they do not physically share memory, nodes must communicate by passing messages through an interconnection network. Neighboring nodes may send messages to one another directly, while nodes that are not directly connected must rely on other nodes in the network to relay messages from source to destination. This is accomplished by a routing function that selects, for each pair of nodes u and v, the set of edges incident on u that can be used to forward messages to v.

Many different models can be defined for message routing in a communication network depending on the way each message moves through the network and it is buffered along the path. In this paper, we consider the *packet routing* (or *packet switching*) and the *wormhole routing* models (see [23, 9, 13, 20]). In the former model each message consists of a single entity (*packet*) which moves through the network and a set of *buffers* is assigned to each vertex. A buffer is the basic storage unit able to contain a single packet. Every time a packet is received by a vertex v, it is first stored in one buffer of that vertex and then, if v is not its destination, it is forwarded to an adjacent vertex according to the routing function. When the next vertex on the path has not free space, the packet cannot be forwarded. In the wormhole routing model each message consists of a sequence of *flits* of equal size, which is usually called *worm* (different worms may have a different number of flits). At each vertex a set of buffers is assigned to each outgoing arc, each capable of containing a single flit. Only the first flit of the worm is subject to the routing selection, while the others follow it along the path in a pipeline fashion. A path can be viewed as a sequence of arcs and if e_1 and e_2 are two successive arcs which must be crossed by a worm, a flit of the worm residing in one buffer of e_1 cannot traverse e_1 if e_2 has not free space.

Since only a finite amount of buffers is assigned to each vertex or arc and messages are allowed to request buffers while holding others, both models are exposed to the occurrence of situations in which no message can be delivered because of cyclic waitings. Such situations are called *deadlocks* and they influence not only the efficiency of the routing strategy but also its correctness.

A possible approach for deadlock handling is deadlock detection and resolution [2, 6, 5, 7, 16, 21, 22], in which the routing algorithm does not take care of deadlocks that are solved by a flow control procedure whenever they occur. In general,this approach requires global control. Since the routing mechanism is a basic mean of communication among the nodes of a distributed system, in which assuming global control is not too realistic, in this paper we consider a different deadlock handling approach, i.e., deadlock prevention in which the routing function is properly designed in order to avoid the occurrence of deadlocks.

Several techniques have been developed to design deadlock-free routing functions in which deadlocks are avoided by ordering the buffers and allowing each message to use them in a monotonically increasing fashion ([17, 15, 19, 9, 1, 12, 11, 18, 3, 4, 8, 13] among the others). As a consequence of the monotone usage of the buffers, resource dependencies are modeled by a directed acyclic graph (DAG) and this insures deadlock prevention. DAG-based methods can be used with slight modifications both for packet and wormhole routing.

In this paper a similar idea is considered, in which the ordering in the set of buffers of each vertex (arc) is based on the concept of acyclic orientations of a graph (see for example [15, 24]). Informally, an acyclic orientation of a graph G is a directed acyclic graph $\vec{G}$ obtained by orienting the edges of G. The buffers contained in each vertex (arc) are then partitioned into a suitable number of classes and a packet (flit) using buffers of class i moves to buffers of class $i + 1$ every time two consecutive traversed links cause a change of orientation in the graph representing the communication network N, i.e. exactly one link corresponds to an edge of its orientation $\overset{\curvearrowleft}{\rightarrow} N$). Such rule guarantees the acyclicity of the resource dependencies graph. The method thus defined was introduced in [10], but it is formally equivalent to the *"Peaks and Valleys"* scheme presented in [15]. However, in this last paper, the author did not give results for specific network topologies.

A more general definition can be found in [24], together with some results on ring networks.

For any DAG-based deadlock prevention technique, and for every network, there exists a lower bound on the number of buffers which have to be maintained at each vertex or arc to allow deadlock-free routing. Of course, a deadlock prevention technique is as better as the buffer requirement is smaller. For this reason, we are interested in optimal deadlock-free routing algorithms, where "optimal" here means that the buffer requirement is minimal among all the algorithms preventing deadlocks by acyclic orientations.

In this paper, we first show (section 4) that approximating the number of buffers yielded by an optimum acyclic orientation is NP-hard even within a $f(k)$ factor of the optimum solution, where k is the number of source-destination couples wishing to communicate and f is any sublinear function. Moreover, we prove the hardness of finding an optimal acyclic orientation when all shortest paths between each couple of vertices wishing to communicate have to be represented.

We then turn our attention to fixed topologies (section 5), by proving some lower and upper bounds on the buffer requirements for butterflies, tori, grids and hypercubes. There is still a little gap left between lower and upper bounds for grids and hypercubes, while the results for tori are tight.

Finally, in section 2 we give the basic notation and definitions we use throughout the paper, in section 3 we show some preliminary results and in section 6 we discuss some conclusive remarks and we address some open problems.

2 Definitions

A communication network can be modeled as a digraph $G = (V, E)$ in which vertices represent processors and arcs communication links between pairs of processors. Since we consider networks in which two processors can communicate in both directions, in the rest of this paper we always refer to symmetric digraphs.

A source vertex u wishing to send a message to a destination vertex v must choose an outgoing link onto which forwarding the message. This is accomplished by using a routing function $f_u : V \rightarrow 2^E$ that selects for each destination vertex v the set of links incident on u that can be used to forward messages to v. The set $F_G = \{f_u \mid u \in V\}$ will be called the *routing function* for G. Due to efficiency requirements, F_G is usually designed in order to route messages along shortest paths (*minimal* routing). According to the degree of freedom left to the messages to choose their paths, a routing function is called *oblivious* if exactly one shortest path can be used to send a message from node u to v ($|f_u(v)| = 1$ for any (u, v), *adaptive* if a message can choose among several shortest paths, *fully adaptive* if it allows messages to use *any* shortest path. Notice that, it would be better to design an adaptive routing function, since an oblivious one, although very simple, is less fault tolerant and too congestion sensitive, and a fully adaptive routing function gives messages much more freedom degrees than required in practical situations. Nevertheless, in order to study the dependancy of the buffer requirement on the number of paths covered by the routing function, we shall mainly refer to the oblivious and fully adaptive cases.

Definition 2.1 *An acyclic orientation of a digraph $G = (V, E)$ is an acyclic digraph $\vec{G} = (V, \vec{E})$ such that $\vec{E} \subseteq E$.*

Definition 2.2 *Let $\vec{G} = (V, \vec{E})$ be an acyclic orientation of $G = (V, E)$. We say that two consecutive arcs (u, v) and (v, w) cause a change of orientation if exactly one of the two arcs belongs to $\vec{E}$.*

Definition 2.3 *Let $\vec{G} = (V, \vec{E})$ be an acyclic orientation of $G = (V, E)$. Given a dipath $P = \langle u_1, u_2, \ldots, u_h \rangle$ in G, let c be the number of changes of orientation caused by all the pairs of consecutive arcs along P.*

We define the rank $r(P,\vec{G})$ *of* P *with respect to* $\vec{G}$ *as* $r(P,\vec{G}) = c+1$ *if* $(u_1,u_2) \in\vec{E}$ *and* $r(P,\vec{G}) = c+2$ *if* $(u_1,u_2) \notin\vec{E}$.

Given a set $\mathcal{P}$ *of dipaths in* G, *the rank of* $\mathcal{P}$ *with respect to* $\vec{G}$ *is defined as* $r(\mathcal{P},\vec{G}) = \max_{P\in\mathcal{P}} r(P,\vec{G})$.

Finally, the rank of $\mathcal{P}$ *is* $r_G(\mathcal{P}) = \min_{\vec{G}} r(\mathcal{P},\vec{G})$.

Informally, if a dipath P has rank r, then P can be expressed as the concatenation of r directed subpaths $P_1,\ldots,P_r$ such that for each i, $1\le i\le r$, P_i is a dipath in $\vec{G}$ if i is odd and P_i is a dipath in the opposite orientation of $\vec{G}$ if i is even.

For the sake of brevity, if a set of dipaths $\mathcal{P}$ includes *all* shortest dipaths connecting any couple of vertices in the network, then we denote $r(\mathcal{P},\vec{G})$ and $r_G(\mathcal{P})$ respectively as $r(\vec{G})$ and r_G. Similarly, if a set of dipaths $\mathcal{P}$ includes *exactly one* shortest dipath connecting any couple of vertices in the network, then we denote $r(\mathcal{P},\vec{G})$ and $r_G(\mathcal{P})$ respectively as $r^o(\vec{G})$ and r^o_G.

In packet (wormhole) routing let us denote as s_u ($s_{u,v}$) the number of buffers assigned by the routing scheme to vertex u (arc (u,v)). Then, the importance of acyclic orientations is stated by the following classical theorem (see [15] for a formally equivalent theorem and [24] for a more general statement).

Theorem 2.1 *Given a network* G, *an acyclic orientation* $\vec{G}$ *of* G *and a set of dipaths* $\mathcal{P}$ *there exists a deadlock free packet (wormhole) routing scheme for* G *which routes messages along the dipaths in* $\mathcal{P}$ *and is such that for each vertex* u *(each pair of arcs* $(u,v),(v,u)$*)* $s_u \le r(\mathcal{P},\vec{G})$ *(*$s_{u,v}+s_{v,u} \le r(\mathcal{P},\vec{G})$*).*

Notice that in wormhole routing the previous theorem bounds per each pair of adjacent vertices u and v the sum of number of buffers assigned to (u,v) and (v,u).

3 Preliminary results

In this section we provide some preliminary results about acyclic orientations.

Theorem 3.1 *For any network* G, *there exist a set of paths* $\mathcal{P}$ *including at least one path between each pair of nodes and an acyclic orientation* $\vec{G}$ *such that* $r(\mathcal{P},\vec{G}) = 2$.

Proof. Consider the spanning tree T rooted at any chosen node u, and let $\mathcal{P}$ be the set of all directed paths induced by the edges of the tree. Consider the orientation $\vec{G}$ in which the edges of G are directed from the leaves of T towards u. For any pair of nodes s and d, $\vec{G}$ covers the portion of the path from s to d arriving at their nearest common ancestor and its reversal covers the remaining portion. □

The above result simply states that there always exists a deadlock free routing scheme based on acyclic orientations for a connected graph using at most 2 buffers at each node (edge). This upper bound is tight, as shown in the following theorem:

Theorem 3.2 *Given a network* G *and a set of paths* $\mathcal{P}$ *including at least one path for any pair of nodes,* $r(\mathcal{P},\vec{G}) \ge 2$ *for any acyclic orientation* $\vec{G}$.

Proof. Let $\vec{G}$ be an acyclic orientation for G. Given any pair of nodes (s,d), by hypothesis there exists a directed path $p_1 \in \mathcal{P}$ from s to d. and a directed path $p_2 \in \mathcal{P}$ from d to s both covered by $\vec{G}$. If $r(\mathcal{P}, \vec{G}) = 1$, then p_1 and p_2 induce a cycle in $\vec{G}$, thus contradicting the definition of acyclic orientation. □

Since we are interested only in shortest paths, we can improve the above lower bound as follows.

Lemma 3.3 *Given a ring network R_4 of 4 nodes, it does not exist any acyclic orientation $\vec{(} R_4)$ for R_4 such that $r(\vec{(} R_4)) < 3$.*

Proof. Let $V = \{v_0, v_1, v_2, v_3\}$, $E = \{(v_i, v_{(i+1)\, mod\, 4})\}$. By the acyclicity of $\vec{(} R_4)$, there must exist a node v_i such that arcs $(v_{(i+1)\, mod\, 4}, v_i)$ and $(v_{(i-1)\, mod\, 4}, v_i)$ are included in $\vec{(} R_4)$. The assertion follows by observing that in this case one more change of orientation is required for the shortest path $< v_{(i+1)\, mod\, 4}, v_i, v_{(i-1)\, mod\, 4} >$. □

Similarly, since in any ring of 5 nodes there exists a unique shortest path connecting nodes $v_{(i+1)\, mod\, n}$ and $v_{(i-1)\, mod\, n}$, the following lemma is easily proved.

Lemma 3.4 *Given a ring network R_5 of 5 nodes, it does not exist any acyclic orientation $\vec{(} R_5)$ such that $r^o(\vec{(} R_5)) < 3$.*

As a consequence of the above two lemmata, the following theorems hold.

Theorem 3.5 *Given a graph G containing at least one cordless cycle of length at least 4, it does not exist any acyclic orientation $\vec{G}$ such that $r(\vec{G}) < 3$.*

Proof. It directly follows from lemma 3.3 by observing that $r(\vec{G})$ is at least as much as the minimum size of a similar acyclic orientation for the subgraph induced by nodes in a cordless cycle of length 4. □

Theorem 3.6 *Given a graph G containing at least one cordless cycle of length greater or equal to 5, it does not exist any acyclic orientation $\vec{G}$ such that size less than $r^o(\vec{G}) < 3$.*

Proof. It s a direct consequence of lemma 3.4. Indeed, $r^o(\vec{G})$ is at least as much as the minimum rank of an acyclic orientation for the subgraph induced by nodes in a cordless cycle of length 5. □

As far as upper bounds are concerned, the following theorem holds.

Theorem 3.7 *[24] For every ring network there exists an acyclic orientation $\vec{(} R)$ such that $r^o(\vec{(} R)) = 3$.*

4 Finding minimal acyclic orientations.

In many applications not all pairs of vertices need to exchange messages with each other. Thus, it is worthwhile to specify a set $R = \{(s_1, t_1), \ldots, (s_k, t_k)\} \subseteq V^2$ of *communication requests* denoting the couples of vertices wishing to communicate.

Given a network G, a set of communication requests R, a set $\mathcal{P}$ of dipaths connecting all pairs in R and an integer $k > 0$, we now consider the problem of deciding if $r_G(\mathcal{P}) \leq k$.

Unfortunately, it turns out that the problem of minimizing $r_G(\mathcal{P})$ is NP-hard. We start proving such assertion from the fully adaptive case.

Theorem 4.1 *Given a graph G, a set of communication requests R and the set of dipaths $\mathcal{P}$ containing all shortest dipaths connecting each couple in R, it is* NP*-hard to decide if* $r_G(\mathcal{P}) \leq 5$.

Proof. Consider the 3-SAT problem: given a boolean function f in conjunctive normal form in which each clause contains exactly three literals, decide if there exists a truth assignment satisfying f. We will provide a polynomial-time reduction which associates to an instance of 3-SAT a network G and a set of communication requests R such that there exists a truth assignment for f if and only if $r_G(P_R) \leq 5$, where P_R is the set containing all shortest dipaths between each couple in R. Then the assertion will follow from the NP-completeness of 3-SAT [14].

We say that an orientation $\vec{G}$ is *acceptable* for $< G, R >$ if $r(P_R, \vec{G}) \leq 5$. Notice that if $\vec{G}$ is acceptable then for any request $(s_i, t_i) \in R$ the dipath from s_i to t_i can have at most 4 changes of orientation.

Let $f = c_1 \wedge \ldots \wedge c_m$ be a formula in conjunctive normal form defined on the set of variables $X = \{x_1 \ldots x_n\}$ such that each clause contains three literals. The corresponding network G is constructed as follows.

We associate to each variable x_i 10 columns grouped two by two: the network is built from a set of $10n$ *columns* divided into 5 blocks of $2n$ columns each. We will denote these columns as $C_b(i)$ with $1 \leq i \leq n$, $1 \leq b \leq 5$ and $C \in \{P, Q\}$. $P_b(i)$ (resp. $Q_b(i)$) is the column of type P (resp. Q) belonging to block b and associated to variable x_i. Columns of type P will be said *constrained* and the ones of type Q *free*. C will denote a generic column. *Columns* are parallel vertical dipaths of length L (which will be specified later) and the set of vertices of a column C is $\{C.x \mid 0 \leq x < L\}$. For a given vertex $v = C.x$, we say that x and C are respectively the *coordinate* and the column of v. The edges of column C join vertices $C.x$ and $C.(x+1)$ for $0 \leq x < L-1$. For reasons that will be explained in a few lines, columns are divided into $40n+1$ horizontal slices of thickness S. More formally, the slice s ($0 \leq s < 40n+1$) is the subgraph induced by the vertices of coordinate $x \in [sS, (s+1)S-1]$ and so $L = (40n+1)S$. We will denote by *atom* $A_{s,b}$ the subset of the vertices in slice s and block b.

We say that $\vec{G}$ is *uniform* on column C and slice s if C is uniformly oriented downward or upward in the slice s, that is, either the dipath from $C.(sS)$ to $C.((s+1)S-1)$ belongs to $\vec{G}$ or the dipath from $C.((s+1)S-1)$ to $C.(sS)$ belongs to $\vec{G}$. Similarly, $\vec{G}$ is *uniform* on slice s (resp. on atom $A_{s,b}$) if $\vec{G}$ is uniform on each column of slice s (resp. of atom $A_{s,b}$). We will say that $\vec{G}$ is *strongly uniform downward* (resp. *upward*) on atom $A_{s,b}$ if $\vec{G}$ is uniform on $A_{s,b}$ and furthermore all columns of type P (the constrained ones) are oriented downward (resp. upward). Again, this means that all the dipaths from $P_b(i).(sS)$ to $P_b(i).((s+1)S-1)$ belong to $\vec{G}$ (resp. all the dipaths from $P_b(i).((s+1)S-1)$ to $P_b(i).(sS)$ belong to $\vec{G}$), for $1 \leq i \leq n$.

Now, we put in the set of requests the pairs formed by the initial and the terminal vertices of each column (all the couples $(C.0, C.(L-1))$). Then, if $\vec{G}$ is acceptable for this set of request, $\vec{G}$ is uniform on at least one slice. Indeed, since the unique shortest dipath from $C.0$ to $C.(L-1)$ is the column C itself, we know that on each column there are at most 4 changes of orientation. Since the total number of columns is $10n$ and each column may contribute to the non uniformity of at most 4 slices, the maximum number of non uniform slices is $40n$. Hence, there must exists a uniform slice, because the total number of slices is $40n+1$.

We now add some new edges and requests so that any acceptable orientation $\vec{G}$ has to be strongly uniform on an atom $A_{s,b}$. In each slice we perform the same construction as described in the following. We refer to the coordinate of a node in slice s by its offset from the coordinate of the initial vertex of the slice $s_0 = sS$, so in what follows the vertex $C.(sS + x)$ will be simply denoted as $C.x$.

The edges needed to complete the following dipath are added (see figure 1).

$$
\begin{array}{rlllllll}
T = & < P_1(1).0, & P_1(1).1, & P_1(2).0, & P_1(2).1, & P_1(3).0, & \ldots, & P_1(n).1, \\
& P_2(1).0, & P_2(1).1, & P_2(2).0, & P_2(2).1, & P_2(3).0, & \ldots, & P_2(n).1, \\
& P_3(1).0, & P_3(1).1, & P_3(2).0, & P_3(2).1, & P_3(3).0, & \ldots, & P_3(n).1, \\
& P_4(1).0, & P_4(1).1, & P_4(2).0, & P_4(2).1, & P_4(3).0, & \ldots, & P_4(n).1, \\
& P_5(1).0, & P_5(1).1, & P_5(2).0, & P_5(2).1, & P_5(3).0, & \ldots, & P_5(n).1 >
\end{array}
$$

The communication request $(P_1(1).0, P_5(n).1)$ is added to the set of requests.

Let us consider now a slice s_0, such that $\vec{G}$ is uniform on s_0. Since the dipath from $P_1(1).0$ to $P_5(n).1$ has at most four orientation changes, $\vec{G}$ is necessarily such that for some b_0, $1 \leq b_0 \leq 5$, all columns $Q_{b_0}(i)$, $1 \leq i \leq n$, have the same orientation in slice s_0. Thus $\vec{G}$ is strongly uniform downward or upward in the atom A_{s_0,b_0}.

The remaining and the key part of our construction is devoted to show the requests (and the shortest dipaths) associated to the clauses of f in such a way that f is satisfiable if and only if there exists an acceptable orientation $\vec{G}$ for a strongly uniform block.

To this aim, we add edges and requests for each atom and each clause in the same way. We first split the vertices of a column C in slice s as follows. On each column and for each slice s we reserve h_0 vertices (namely vertices $C.x$ with $x \in [sS, sS + h_0 - 1]$) for the path T defined above, h vertices per clause c_k (namely vertices $C.x$ with $x \in [sS + h_0 + kh, sS + h_0 + (k+1)h - 1]$) and h_0 vertices (namely vertices $C.x$ with $x \in [sS + h_0 + mh, sS + h_0 + mh + h_0 - 1]$) at the end of the atom to separate it from the next one. Thus, $S = mh + 2h_0$. The two parameters h_0 and h will be adjusted later in such a way that the dipaths that we consider in the proof are unique shortest dipaths. In order to have simpler notations, for the clause c_k in a generic atom $A_{s,b}$ we will denote the vertex $C_b(i).sS + kh + h_0 + x$ by $C_b(i).x$.

Let $c_k = l_{j_1} \vee l_{j_2} \vee l_{j_3}$, with $j_1 < j_2 < j_3$, where l_{j_u} is either x_{j_u} or $\overline{x_{j_u}}$. The node E is defined as $Q(j_3).1$ if $l_{j_3} = x_3$, or as $Q(j_3).0$ if $l_{j_3} = \overline{x_3}$. The communication request $(P(j_1).0, E)$ is added to the set of requests and the edges necessary to build the following dipaths are added to G (see also figure 2):

- $< P(j_1).0,\ P(j_1).1,\ Q(j_1).0,\ Q(j_1).1,\ P(j_2).0 >$ if $l_{j_1} = x_{j_1}$
  $< P(j_1).0,\ P(j_1).1,\ Q(j_1).1,\ Q(j_1).0,\ P(j_2).0 >$ if $l_{j_1} = \overline{x_{j_1}}$
- $< P(j_2).0,\ P(j_2).1,\ Q(j_2).0,\ Q(j_2).1,\ P(j_3).0 >$ if $l_{j_2} = x_{j_2}$
  $< P(j_2).0,\ P(j_2).1,\ Q(j_2).1,\ Q(j_2).0,\ P(j_3).0 >$ if $l_{j_2} = \overline{x_{j_2}}$
- $< P(j_3).0,\ P(j_3).1,\ Q(j_3).0,\ Q(j_3).1 >$ if $l_{j_3} = x_{j_3}$
  $< P(j_3).0,\ P(j_3).1,\ Q(j_3).1,\ Q(j_3).0 >$ if $l_{j_3} = \overline{x_{j_3}}$

Consider now an acceptable orientation $\vec{G}$ for the set of communication requests $\{(C.0, C.(L-1))\}$ for every column C and a strongly uniform atom $A_{s_0 b_0}$ for $\vec{G}$. It is possible to associate to $\vec{G}$ a truth assignment for X as follows. In slice s_0, all columns $P_{b_0}(i)$ are oriented downward (resp. upward) if $\vec{G}$ is strongly uniform downward (resp. upward) in the atom, and each column $Q_{b_0}(i)$ can be independently oriented downward or upward. If the orientations of $P_{b_0}(i)$ and $Q_{b_0}(i)$ are identical (resp. opposite) we will associate to x_i the value true (resp. false).

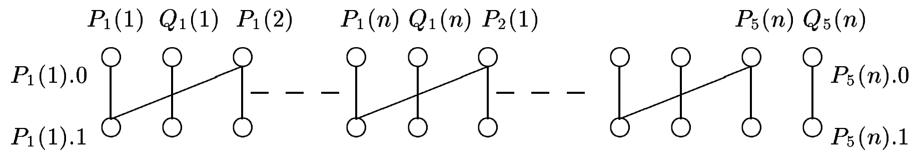

Figure 1: The dipath constructed in each slice

Notice that, if the truth assignment associated to the strongly uniform atom A_{s_0,b_0} is such that the clause c_k is false, then the dipath from $P(j_1).0$ to E when $\overrightarrow{G}$ is strongly uniform downward or upward in the atom has to use the orientations $\overrightarrow{G}\overleftarrow{G}\overrightarrow{G}\overleftarrow{G}\overrightarrow{G}\overleftarrow{G}$, where $\overleftarrow{G}$ is the reversal of $\overrightarrow{G}$. Hence, such a dipath has at least 5 changes of orientation, i.e. rank at least 6, and $\overrightarrow{G}$ cannot be acceptable. Thus, if $\overrightarrow{G}$ is acceptable then f is satisfiable.

To complete the proof we must provide an acceptable orientation $\overrightarrow{G}$ when f is satisfiable. To this aim, we choose a truth assignment for variables x_i satisfying f and we define $\overrightarrow{G}$ as follows: constrained columns are directed downward, free columns are directed according to the truth assignment (as previously shown), horizontal arcs are directed from left to right. More formally:

all columns of $P_b(i)$ such that $1 \leq b \leq 5$ and $1 \leq i \leq n$ are directed downward (that is arc $(P_b(i).x, P_b(i).(x+1))$ is in $\overrightarrow{G}$).

if x_i is true all columns $Q_b(i)$ such that $1 \leq i \leq n$ are directed downward $((Q_b(i).x, Q_b(i).(x+1)) \in \overrightarrow{G})$, otherwise they are directed upward $((Q_b(i).(x+1), Q_b(i).x) \in \overrightarrow{G})$.

if there is an edge between two vertices $C_b(i).x$ and $C_{b'}(i').x'$ with $b < b'$ or $b = b'$ and $i < i'$, then the arc $(C_b(i).x, C_{b'}(i').x')$ belongs to $\overrightarrow{G}$.

if there is an edge between two vertices $P_b(i).x$ and $Q_b(i).x'$, then the arc $(P_b(i).x, Q_b(i).x')$ belongs to $\overrightarrow{G}$.

Such an orientation is clearly acyclic (any dipath in $\overrightarrow{G}$ either stays on a column and goes upward or downward, or it goes strictly from left to right). Since all the clauses are true under the chosen truth assignment and consequently each of them contains at least one true literal, one can check that the dipaths associated to clauses have rank at most 5. All the other requests are fulfilled with no change of orientation, thus we have constructed an acceptable orientation for the graph G.

In order to complete the proof it suffices to observe that by choosing $h_0 \geq 2n$ and $h \geq 10$ all the considered dipaths are the (unique) shortest ones. This leads to $S = 10m+4n$, $L = (40n+1)(10m+4n)$ and to a total number of vertices in the graph equal to $10nL = 10n(40n+1)(10m+4n)$.

□

Concerning the adaptive and the oblivious cases, things are even worse. In fact, consider the possibility of devising polynomial time algorithms able to find approximate solutions, that is, solutions whose sizes have constant approximation error with respect to the optimal ones. The formal definition of the approximation error of a minimization problem Π is defined as $\frac{m(S_A)}{m(S^*)}$, where $m(S^*)$ is the size of an optimum solution S^* and $m(S_A)$ is the approximate solution computed by some algorithm A. A problem is said to be ϵ-approximable if a polynomial time algorithm A exists such that the approximation error is never greater than ϵ.

The technique used in the previous theorem can be exploited to prove that, in the oblivious and adaptive cases, it is NP-hard even to approximate $r_G(\mathcal{P})$.

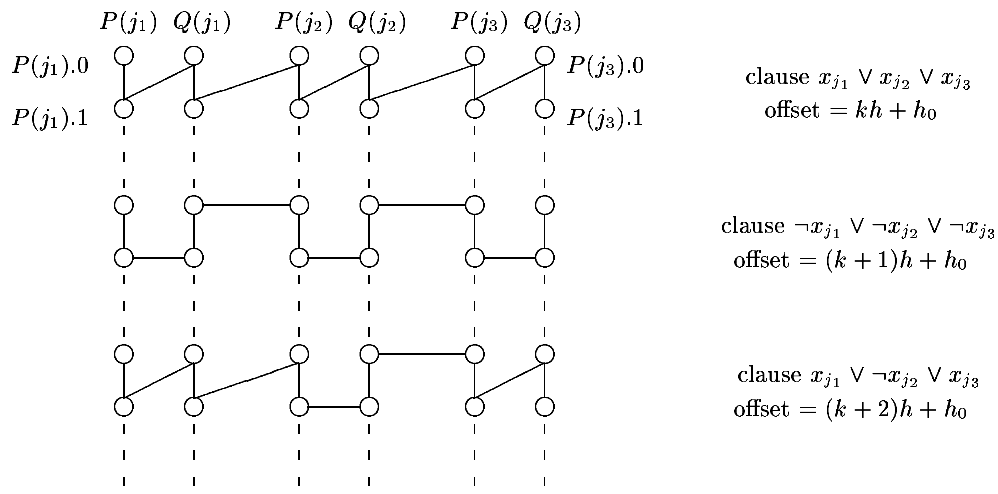

Figure 2: Sample of 3 out of 8 possible clauses.

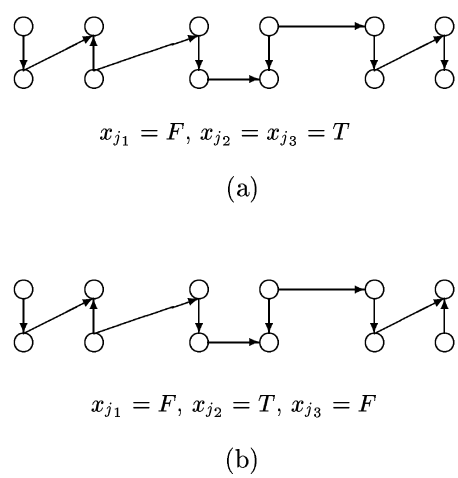

Figure 3: Two possible truth assignments to clause $C = x_{i1} \vee \neg x_{i2} \vee x_{i3}$: (a) satisfying c, (b) not satisfying C.

Theorem 4.2 *Given a graph G and a set of communication requests R in G, it is* NP*-hard to approximate the*

$$\min\{r_G(\mathcal{P}) : \mathcal{P} \text{ includes exactly one shortest dipath for each couple in } R\}$$

within an error in $\mathbf{O}(k^\epsilon)$ *for any* $\epsilon < 1$*, where* $k = |R|$*.*

Proof. We still use a polynomial-time reduction from the 3-SAT problem. In this case, the reduction builds a network G_f and a set of communication requests R_f corresponding to a boolean formula f such that there is a large gap between the number of changes sufficient for a set of shortest dipaths in G_f connecting each couple in R_f when f is satisfiable and the number of changes necessary for any set of shortest dipaths when f is not satisfiable.

The construction of G_f is very similar to the one shown in theorem 4.1: we still have a number S of slices and a number B of blocks, each block b including n pairs of columns $P_b(i)$ and $Q_b(i)$, $1 \leq i \leq n$. This time, the thickness of each slice is $70m + 40n$. Again, we have the communication requests $(P_b(i).0, P_b(i).(L-1))$ and $(Q_b(i).0, Q_b(i).(L-1))$. As in the previous theorem, we have that every acceptable orientation makes at least one atom strongly uniform, where *acceptable* now means that it produces no change of orientation in at least one shortest dipath connecting every pair of communication request.

The only differences with the network of the previous theorem are:

- the number of slices $S = n^2$,
- the number of blocks is $B = n^2$, and
- for each clause c_j, n^2 communication requests $(E_j^1, F_j^1), (E_j^2, F_j^2), \ldots, (E_j^{n^2}, F_j^{n^2})$ are included in R (one for each slice) and any shortest dipath connecting E_j^h to F_j^h must use one out of 7 shortest dipaths in each block b.

More precisely, let $c_k = l_{j_1} \vee l_{j_2} \vee l_{j_3}$, with $j_1 < j_2 < j_3$, where l_{j_u} is either x_{j_u} or $\overline{x_{j_u}}$. Instead of formally describing the 7 shortest dipaths that must be used in block b to connect E_j^h to F_j^h, they are shown in figure 4. Each of such "subpath" corresponds to a truth assignment satisfying f.

Needless to say, the previous construction can be performed in polynomial time. We now claim that if f is satisfiable an acceptable orientation for the network exists, otherwise every acyclic orientation requires at least $n^2 - 1$ changes for any set of shortest dipaths connecting the couples in R_f.

Suppose first that f is satisfiable and let $\mathcal{A}$ be a truth assignment satisfying f. The same orientation $\vec{G}$ described in the proof of theorem 4.1 with rank at most 5 can now be used: all horizontal edges are oriented from left to right, all columns P downwards, column $Q_b(i)$ is oriented downwards if x_i has received by $\mathcal{A}$ the value true, upwards otherwise. In this case, since each clause c_j is satisfied, there exists one shortest dipath connecting E_j^h to F_j^h that uses only edges in $\vec{G}$: in every block the subpath corresponding to a truth value assigned by $\mathcal{A}$ to the literals in c_j is chosen . Trivially, no change of orientation is now necessary.

Conversely, we now prove that if f is not satisfiable then every shortest dipath connecting any communication request requires a non constant number of changes. If f is not satisfiable, every truth assignment $\mathcal{A}'$ is unable to satisfy at least one clause. Notice that, if we orient the edges of $\vec{G}$ according to a truth assignment $\mathcal{A}'$ (as explained above) that does not satisfy clause c_j, then each of the seven subpaths to be used in a block to connect E_j^h to F_j^h requires at least two changes. Thus, an acyclic orientation corresponding to a truth assignment yields at least $B = n^2 - 1$ changes. On the other hand, trying to have a lower the number of changes needed to the shortest dipaths from E_j^h to F_j^h while keeping the horizontal edges oriented from left to right corresponds to increasing the number of changes necessary to the columns: if a dipath from E_j^h to

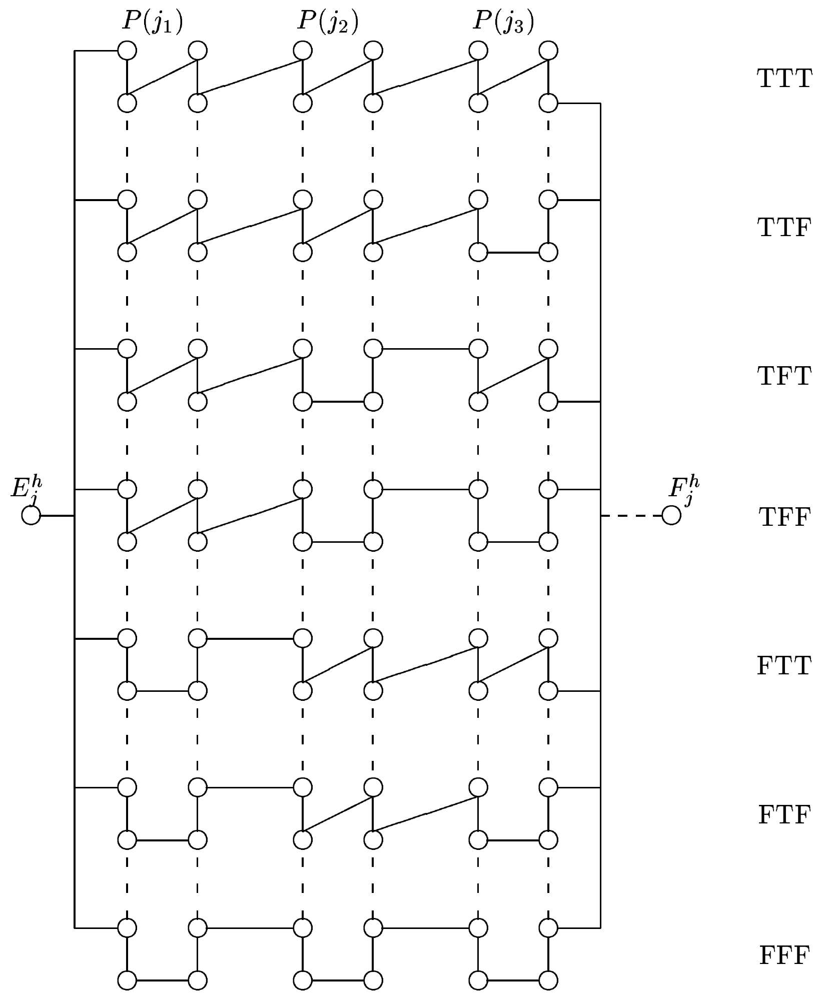

Figure 4: The seven dipaths in block b corresponding to the seven possible truth assignments satisfying clause $C_j = x_{i1} \vee \neg x_{i2} \vee x_{i3}$.

F_j^h uses $n^2 - l$ changes then l columns involved in that dipath need at least 2 changes. Thus, if all pairs $(E_j^1, F_j^1), (E_j^2, F_j^2), \ldots, (E_j^{n^2}, F_j^{n^2})$ should be connected by using at most $n^2 - l$ changes then at least one of the involved 6 columns would require

$$\geq \lceil \frac{l}{6} n^2 \rceil$$

changes. This means that if the horizontal edges are all oriented from left to right every set of shortest dipaths $\mathcal{P}$ has rank at least

$$r_G(\mathcal{P}) = \min_k \{\max\{n^2 - k, \lceil \frac{k}{6} n^2 \rceil\}\} = n^2.$$

Observe now that orienting some horizontal edge from right to left could help only if the columns are oriented according to a truth assignment $\mathcal{A}'$ and *all* horizontal edges corresponding to a false literal in clause c_j are oriented from right to left. However, in this case

- either some literal in some other clause c_l is true under $\mathcal{A}'$ and orienting its horizontal edges from left to right (so that one of the dipaths from $E^h - l$ to F_l^h uses a constant number of changes) would induce a cycle in $\vec{G}$
- or all horizontal edges are oriented from right to left. In this last case, it is sufficient to invert the orientation of every column to obtain the symmetric (and equal) situation of the horizontal edges oriented from left to right.

Hence, changing the orientation of the horizontal edges does not help in keeping "small" the number of changes and the above claim is proved.

Suppose now an $g(k)$-approximation algorithm T exists for the minimum acyclic orientation problem, with $g(k)$ some sublinear function in the number of communication requests k. Thus, if $r^*(G,k) = \min\{r_G(\mathcal{P}) : \mathcal{P}$ includes exactly one shortest dipath for each couple in $R\}$ and $r^T(G,k)$ denotes the minimum number of changes used by the acyclic orientation found by T for any set of shortest dipaths connecting the couples in R_f, the following relation holds:

$$\frac{r^T(G,k)}{r^*(G,k)} \leq g(k).$$

We now show how it is possible, by using T and the reduction above, to decide if a boolean formula is satisfiable in polynomial time. Consider a boolean formula f: transform f into a pair (G_f, R_f) and apply to it algorithm T. If f is satisfiable then T finds for (G_f, R_f) an acyclic orientation $\vec{G}$ such that $r^T(G_f,k) \leq g(k) r^*(G_f,k)$, where k is the size of R_f; conversely, if f is not satisfiable then T finds for (G_f, R_f) an acyclic orientation such that $r^T(G_f,k) \geq r^*(G_f,k) \geq n^2$, n being the number of boolean variables used by f. Since $k = \mathbf{O}(n^5)$, then $g(n^5) < n^2$ whenever $g(k) \leq k^{\frac{2}{5}}$, that is, a $k^{\frac{2}{5}}$-approximation algorithm for the minimum acyclic orientation problem is also a polynomial time algorithm deciding satisfiability.

Finally, notice that if in the reduction above we use $\mathbf{O}(n^h)$ instead of n^2 slices and blocks, the number of communication requests becomes $\mathbf{O}(n^{h+3})$ and the rank in the case of a no instance becomes $\mathbf{O}(n^h)$. This implies that a $g(k)$-approximation algorithm cannot exist for any $g(k) \leq k^{\frac{h}{h+3}}$. Since the previous assertion holds for *any* $h > 0$, the theorem is completely proved. □

Of course the non approximability result for the adaptive case follows by generalization.

The previous results motivate us to look for minimal schemes for some classes of graph which are widely used in distributed and parallel systems.

5 Bounds for fixed topologies

In this section, we consider only the case where $R = V \times V$ and we provide some bounds both on r_G and on r^o_G when G is a butterfly, a torus, a grid, an hypercube, all of them being classical interconnection networks.

Before starting the analysis for fixed topologies, we need some preliminary steps.

Recall that a coloring of a graph G is a function $\mathcal{C}_G : V \to |V|$, such that for each pair of nodes u, v, $u \neq v$, if $(u, v) \in E$ then $\mathcal{C}_G(u) \neq \mathcal{C}_G(v)$. The size of $\mathcal{C}_G$ is $max\{\mathcal{C}_G(u) \mid u \in V\}$ and the color of u is $\mathcal{C}_G(u)$. A coloring of size d simply partitions the set of nodes into d independent subsets $S_1, \dots, S_d$.

Given a coloring of a graph G, any path p in G can be expressed as the concatenation $p_1 \circ \ldots \circ p_h$ of h directed paths $p_1, \dots, p_h$ such that, for each odd i, edges in p_i are "positive" (that is, they are traversed from a node having color i to a node having color $j > i$), while, for each even i, edges in p_i are "negative". We define the *depth* of p as the minimum h such that $p = p_1 \circ \ldots \circ p_h$.

Notice that the depth of p merely corresponds to the number of alternances between positive and negative edges in p, starting with an eventually empty succession of positive edges. Given a set of paths $\mathcal{P}$, the depth $\mathcal{D}(\mathcal{C}_G, \mathcal{P})$ of a coloring $\mathcal{C}_G$ w.r.t $\mathcal{P}$ is the maximum depth of a path in $\mathcal{P}$.

There is an important relation between the depth of a coloring and the rank of an acyclic orientation, as stated in the following theorem.

Theorem 5.1 *Given a graph G and a set of paths $\mathcal{P}$, there exists aa acyclic orientation $\vec{G}$ for $\mathcal{P}$ with $r(\mathcal{P}, \vec{G}) = k$ if and only if there exists a coloring $\mathcal{C}_G$ such that $\mathcal{D}(\mathcal{C}_G, \mathcal{P}) = k$.*

Proof. Assume first that an acyclic orientation $\vec{G}$ such that $r(\mathcal{P}, \vec{G}) = k$ exists. Since $\vec{G}$ is acyclic, we can construct a topological order ord of the nodes in $\vec{G}$ satisfying the property according to which if there is a directed path in $\vec{G}$ from u to v then $ord(u) < ord(v)$. Consider the coloring $\mathcal{C}_G$ such that $\mathcal{C}_G(u) = ord(u)$. It is easy to see that h changes of orientations are sufficient for any path $p \in \mathcal{P}$ if and only if $\mathcal{D}(\mathcal{C}_G, \mathcal{P}) \leq h$, thus $\mathcal{D}(\mathcal{C}_G, \mathcal{P}) = k$.

Assume now that a coloring $\mathcal{C}_G$ with $\mathcal{D}(\mathcal{C}_G, \mathcal{P}) = k$ exists. Let $\vec{G}$ be the acyclic orientation of G obtained by orienting all the edges $(u, v) \in E$ from u to v if and only if $\mathcal{C}_G(u) \leq \mathcal{C}_G(v)$. Given $p \in \mathcal{P}$ of depth h, let $p_1, \dots, p_h$ the directed subpaths (induced by positive or negative edges in $\mathcal{C}_G$) such that $p = p_1 \circ \ldots \circ p_h$. Clearly if i is odd (even) then edges in p_i belong to $\vec{G}$ (its reversal) and the acyclic orientation of size h obtained by alternating $\vec{G}$ and its reversal covers p. The claim follows by observing that this holds for any $p \in P$. □

Thus, given a graph G and a set of paths $\mathcal{P}$, the problem of finding a minimum rank acyclic orientation for $\mathcal{P}$ can be reduced to the problem of finding a coloring $\mathcal{C}_G$ having minimum $\mathcal{D}(\mathcal{C}_G, \mathcal{P})$.

Finally, let us remark one of the key properties of orientations related to the traversability of a cycle. Let C be a 4-cycle consisting of the arcs e_0, e_1, e_2, e_3. As any orientation $\vec{G}$ is acyclic, in the subgraph induced by the cycle C there is at least one sink and one source. So, if we consider any four dipaths of length 2: P_0, P_1, P_2, P_3, where P_i contains arcs e_i and $e_{(i+1)\,mod\,4}$, at least two of them have one change of orientation in the cycle.

5.1 Butterfly

Consider a butterfly B_d of dimension d, with $V = \{v_{i,\alpha} \mid 1 \leq i \leq d+1, \alpha \in \{0,1\}^d\}$, and $E = \{(v_{i,\alpha}, v_{i+1,\alpha}) \mid 1 \leq i \leq d\} \cup \{(v_{i,\alpha}, v_{i+1,\alpha(i)}) \mid 1 \leq i \leq d\}$, where $\alpha(i)$ is the string differing from α only on the i-th bit.

Since B_d contains at least a cycle of length 4, from theorem 3.5 it results $r_{B_d} \geq 3$. The following theorem provides an upper bound.

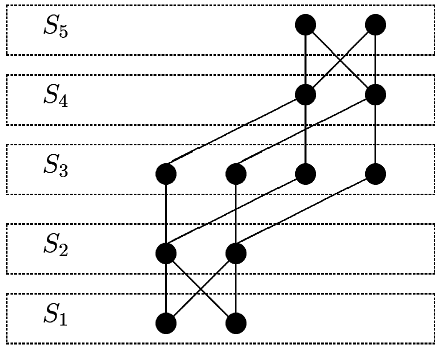

Figure 5: Coloring of a butterfly

Theorem 5.2 $r_{B_d} \leq 4$.

Proof. It follows directly from theorem 5.1 by observing that the coloring $\mathcal{C}_{B_d}(v_{i,\alpha}) = i$ has depth $\mathcal{D}(\mathcal{C}_{B_d}) = 4$.

As an example, a coloring matching the lower bound for 2-dimensional butterflies can be seen in figure 5.

□

5.2 Torus

We now consider torus networks. The vertices of $T_{p\times q}$ will be denoted as (i, j) with $i \in Z_p$, $j \in Z_q$. Vertex (i, j) is joined to vertices $(i+1, j)$ and $(i-1, j)$ by horizontal arcs and to vertices $(i, j+1)$ and $(i, j-1)$ by vertical arcs.

Let's start the analysis with oblivious routing. By theorem 3.6, for $n \geq 5$ $\;r^o_{T_{n\times n}} \geq 3$. Such a bound is tight, as proved in the next theorem.

Theorem 5.3 $r^o_{T_{n\times n}} \leq 3$.

Proof. By theorem **??**, given a ring R_n of n nodes, with $V = \{v_i \mid i = 1, \ldots, n\}$ and $E = \{v_i, v_{(i+1)\, mod\, n} \mid i = 1, \ldots, n\}$, it results $r^o_{R_n} \leq r_{R_n} = 3$, thus there exists an acyclic orientation $\overset{(}{\to} R_n)$ such that $r^o(R_n) = 3$.

Consider now the acyclic orientation $\overset{(}{\to} T_{n\times n})$ obtained by orienting all rows and columns of $T_{n\times n}$ according to $\overset{(}{\to} R_n)$ (the subgraph induced by all nodes in a same row or in a same column is a ring of n nodes), i.e. such that $(v_{i,j}, v_{i,((j+1)\, mod\, n)}) \in \overset{(}{\to} T_{n\times n})$ iff $(v_j, v_{((j+1)\, mod\, n)}) \in \overset{(}{\to} R_n)$ and $(v_{i,j}, v_{((i+1)\, mod\, n),j}) \in \overset{(}{\to} T_{n\times n})$ iff $(v_i, v_{(i+1)\, mod\, n}) \in \overset{(}{\to} R_n)$.

To prove the claim it suffices to show that the $r^o(\overset{(}{\to} T_{n\times n})) = 3$. In fact, for each pair of nodes $v_{i,j}$ and $v_{i',j'}$ in $T_{n\times n}$, there always exists a shortest path from $v_{i,j}$ to $v_{i',j'}$ obtained by following first some directed edges of $\overset{(}{\to} T_{n\times n})$ along the i-th row until a certain node $v_{i,j''}$ and then some directed edges of $\overset{(}{\to} T_{n\times n})$ along the j''-th column until a certain node $v_{i'',j''}$, and then doing the same for the reversal orientation, and again the same for $\overset{(}{\to} T_{n\times n})$, since by theorem **??** three changes of orientations are sufficient to reach the right row and the right column. □

Concerning the fully adaptive case, the lower and the upper bounds differ only for a constant additive, as shown in the following two theorems.

Theorem 5.4 *Let $p \geq q$, then $r_{T_{p\times q}} \geq \lfloor \frac{q}{2} \rfloor + 2$.*

Proof. Let $p' = \lfloor \frac{p}{2} \rfloor$, $q' = \lfloor \frac{q}{2} \rfloor$ and $N = pq$ (the number of vertices).

Consider first the case $p' = q'$. Let $\mathcal{P}_s$ be the subset of the set of all shortest dipaths $\mathcal{P}$ constituted by the following $8N$ "staircase dipaths": for each vertex (i, j) we associate 8 shortest dipaths of length the diameter $D = p' + q' = 2q'$ where arcs alternate in directions. Such dipaths are of the form $(e_1, f_1, e_2, f_2, \ldots, e_{q'}, f_{q'})$ where the e_i's are all horizontal (resp. vertical) arcs and all the f_i's vertical (resp. horizontal). These dipaths join vertex (i, j) to vertices $(i \stackrel{-}{+} p', j \stackrel{-}{+} q')$.

Notice that if a dipath from (i, j) to (i', j') belongs to $\mathcal{P}_s$ then the opposite dipath from (i', j') to (i, j) also belongs to $\mathcal{P}_s$.

Due to the symmetry of the torus, each of the 8 dipaths of length 2 of any 4-cycle belongs to the same number $2(2q' - 1)$ of dipaths in $\mathcal{P}_s$. So, for any acyclic orientation $\vec{T}_{p\times q}$, the N cycles of length 4 yield globally a total of $4N(2q' - 1)$ changes of orientation over the $8N$ dipaths in $\mathcal{P}_s$.

Therefore, either one dipath of $\mathcal{P}_s$ has at least $q' + 1$ changes or $4N$ dipaths in $\mathcal{P}_s$ have exactly q' changes of orientation and the remaining $4N$ dipaths of $\mathcal{P}_s$ have $q' - 1$ changes. If there is a dipath P with $q' + 1$ changes, then by definition of rank $r_{T_{p\times q}}(\mathcal{P}) \geq r(\mathcal{P}, \vec{T}_{p\times q}) \geq q' + 2$ and we have proven the lower bound, so let us suppose that the second condition holds.

In this case assume by contradiction that $r_{T_{p\times q}}(\mathcal{P}) \leq q' + 1$. Since there are as many dipaths in $\mathcal{P}_s$ starting with an arc in $\vec{T}_{p\times q}$ than with an arc not in $\vec{T}_{p\times q}$, this means that all the dipaths starting with an arc not in $\vec{T}_{p\times q}$ have $q' - 1$ changes of orientation (otherwise they would have rank $q' + 2$) and all the dipaths starting with an arc in $\vec{T}_{p\times q}$ have q' changes.

In this case, all the dipaths in $\mathcal{P}_s$ should have the last (vertical arc) in $\vec{T}_{p\times q}$ if q' is even and not in $\vec{T}_{p\times q}$ if q' is odd, but this is impossible since for for any i and j there are dipaths in $\mathcal{P}_s$ ending with arc $((i, j), (i, j + 1))$ and dipaths ending with arc $((i, j + 1), (i, j))$.

Suppose now $p' > q'$. We use a similar technique, but now we take the set of dipaths $\mathcal{P}_s$ as the $4N$ shortest dipaths of length $2q' + 1$ $(\leq D)$ starting at any vertex with a horizontal arc and where arcs alternate (so the last one is horizontal). These dipaths join vertex (i, j) to vertices $(i \stackrel{-}{+} (q' + 1)', j \stackrel{-}{+} q')$. The total number of changes yielded by the N 4-cycles is now $2N(2q')$ for the $4N$ dipaths of $\mathcal{P}_s$. Therefore, either one dipath in $\mathcal{P}_s$ has at least $q' + 1$ changes of orientation, or all dipaths of $\mathcal{P}_s$ have q' changes. If there is a dipath with $q' + 1$ changes we have proven the lower bound, otherwise all dipaths in $\mathcal{P}_s$ starting with an arc not in $\vec{T}_{p\times q}$ (one half of the total) have rank at least $q' + 2$. □

Theorem 5.5 *Let $p \geq q$, then $r_{T_{p\times q}} \leq \lceil \frac{q}{2} \rceil + 4$.*

Proof(sketch). It suffices to consider the acyclic orientation such that all vertical arcs are oriented from (i, j) to $(i, j + 1)$ for $0 \leq j \leq n - 2$ and from $(i, 0)$ to $(i, n - 1)$. Horizontal arcs are oriented if j is even from (i, j) to $(i + 1, j)$ for $0 \leq i \leq n - 2$ and from $(0, j)$ to $(n - 1, j)$, while if j is odd from $(i + 1, j)$ to (i, j) for $0 \leq i \leq n - 2$ and from $(n - 1, j)$ to $(0, j)$.

For this acyclic orientation we can check that any shortest dipath has rank at most $\lceil \frac{q}{2} \rceil + 4$. □

5.3 Grid

Consider an $n \times n$ grid $G_{n\times n}$, with $V = \{v_{i,j} \mid 1 \leq i \leq n, 1 \leq j \leq n\}$, and $E = \{(v_{i,j}, v_{i+1,j}) \mid 1 \leq i \leq n - 1, 1 \leq j \leq n\} \cup \{(v_{i,j}, v_{i,j+1}) \mid 1 \leq i \leq n, 1 \leq j \leq n - 1\}$.

Concerning the oblivious routing case, the following theorem holds.

Theorem 5.6 $r^o_{G_{n\times n}} = 2$.

Proof. By theorem 3.2 $r^o_{G_{n\times n}} \geq 2$.

Consider now the set $\mathcal{P}$ of all shortest paths constituted by a succession of moves along a row of the grid, followed by a succession of moves along a column. By construction $\mathcal{P}$ contains one shortest path between every pair of nodes. Consider now the coloring $\mathcal{C}_{G_{n\times n}}$ such that $\mathcal{C}_{G_{n\times n}}(v_{i,j}) = i + j$. Clearly is such a coloring two adjacent nodes in $G_{n\times n}$ cannot share the same color. The claim follows by theorem 5.1 by observing that $\mathcal{D}(\mathcal{C}_{G_{n\times n}}, \mathcal{P}) = 2$. □

Good bounds can be determined also for the fully adaptive routing case. Let's start with the lower bound.

Theorem 5.7 *Let $p \geq q$, then $r_{G_{p\times q}} \geq \lceil (2-\sqrt{2})q \rceil - 1$.*

Proof. Consider only the $q \times q$ subgrid $G_{q\times q}$ of $G_{p\times q}$ induced by nodes (i,j) such that $0 \leq i \leq q-1$ and $0 \leq j \leq q-1$. Let α be a fixed number such that $\frac{q-1}{2} \leq \alpha \leq q-1$. The sets of shortest dipaths considered will consist of two disjoint sets $\mathcal{P}_1$ and $\mathcal{P}_2$. $\mathcal{P}_1$ contains the 2α dipaths from $(0,0)$ to $(q-1,q-1)$ constituted by a sequence of horizontal (resp. vertical) arcs till a given vertex $(j,0)$ (resp. $(0,j)$), where $1 \leq j \leq \alpha$, then followed by arcs alternating in direction starting with a vertical (resp. horizontal) arc, then by a vertical (resp. horizontal) dipath from $(q-1,q-1-j)$ (resp. $(q-1-j,q-1)$) to $(q-1,q-1)$. We will call such dipaths "almost staircase". $\mathcal{P}_2$ consists of the 2α "almost staircase" shortest dipaths from $(0,q-1)$ to $(q-1,0)$ constituted by an horizontal (resp. vertical) dipath till $(j,q-1)$ (resp. $(0,q-1-j)$) with $1 \leq j \leq \alpha$, then a staircase dipath and finally a vertical (resp. horizontal) one.

Any 4-cycle will be said to be "inner" if it consists of the four vertices $(i,j),(i,j+1),(i+1,j+1)$ and $(i+1,j)$ where $q-\alpha-1 \leq i+j \leq q+\alpha-3$, $i-j \leq \alpha-1$, $j-i \leq \alpha-1$. Hence, the total number of inner cycles is $c = (q-1)^2 - 2(q-\alpha-1)(q-\alpha)$.

Notice that, for each inner cycle there are exactly 2 dipaths of $\mathcal{P}_1$ using respectively the arcs $(i,j)(i+1,j)(i+1,j+1)$, $(i,j)(i,j+1)(i+1,j+1)$ and 2 dipaths of $\mathcal{P}_2$ using the arcs $(i,j+1)(i,j)(i+1,j)$ and $(i,j+1)(i+1,j+1)(i+1,j)$. By the remark on the acyclicity of the orientations, at least two of these dipaths must change orientation inside the cycle. Hence, the c inner cycles yield globally a total of at least $2c$ changes of orientation over all the dipaths of $\mathcal{P}_1 \cup \mathcal{P}_2$.

Since $|\mathcal{P}_1 \cup \mathcal{P}_2| = 4\alpha$, one dipath $P \in \mathcal{P}_1 \cup \mathcal{P}_2$ has at least $\frac{c}{2\alpha} = \frac{1}{2\alpha}(-(q-1)^2 + 2(2q-1)\alpha - 2\alpha^2)$ changes of orientation.

A simple derivation shows that $\frac{c}{2\alpha}$ is the maximum for $\alpha = \sqrt{\frac{q^2-1}{2}}$. For this value of α, it gives $\frac{c}{2\alpha} = 2(q-1) - \sqrt{2(q^2-1)}$. Since we are considering only integers one can show that $\frac{c}{2\alpha} \geq (2-\sqrt{2})q - 2$. So the dipath P has rank at least $\lceil (2-\sqrt{2})q \rceil - 1$. □

We conjecture that this lower bound is asymptotically the right order for the value $r_{G_{q\times q}}$. Till now we have been able to design a simple construction giving $\frac{2q}{3} + o(q)$ orientations and a slightly more complicated one of order $\frac{3q}{5} + o(q)$, and the method should work for q large enough for any fraction $\frac{aq}{b}$ such that $a/b \geq 2 - \sqrt{2}$.

Is this the right theorem?

Theorem 5.8 $r_{G_{m\times n}} \leq min(\lceil \frac{2}{3} \cdot (m+1) \rceil, \lceil \frac{2}{3} \cdot (n+1) \rceil) + 1$.

Proof(sketch). We partition the grid into nine subgrids, as shown in figure 6. Arcs in each subgrid are then oriented in G_1 as described in the following (see also the figure):

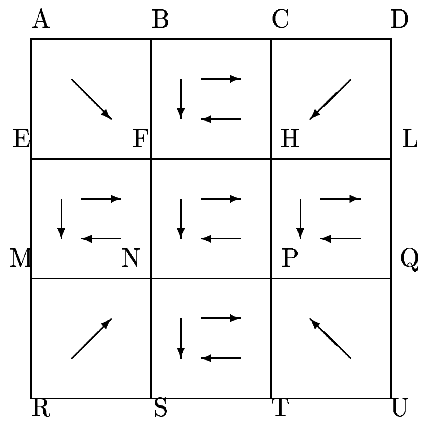

Figure 6: Partitioning of a grid into nine subgrids and orientations of each subgrid.

Subgrid ABFE: columns are oriented from top to bottom, rows from left to right;

Subgrid CDLH: columns are oriented from top to bottom, rows from right to left;

Subgrid PQUT: columns are oriented from bottom to top, rows from right to left;

Subgrid MNSR: columns are oriented from bottom to top, rows from left to right;

All the others subgrids: columns are oriented from top to bottom, odd rows from left to right, even rows from right to left.

It can be seen that the dipaths requiring the largest number of orientations to be covered include one of the following subpath: (E, B), (L, C), (S, M), (T, Q), (M, F, C), (B, H, Q), (L, P, S), (T, N, E). The assert follows by noticing that each of the previous dipaths can be covered by at most $min(\lceil\frac{2}{3}\cdot(m+1)\rceil, \lceil\frac{2}{3}\cdot(n+1)\rceil)$ orientations $\langle G_1, \neg G_1\rangle$ and that the remaining of a dipath including one of them as a subpath can be covered by a single orientation. □

5.4 Hypercubes

Consider an hypercube H_d of dimension d, with $V = \{v_\alpha \mid \alpha \in \{0,1\}^d\}$, and $E = \{(v_\alpha, v_{\alpha(i)}) \mid 1 \leq i \leq d\}$, where $\alpha(i)$ is the string differing from α only on the i-th bit.

In the oblivious routing case, the following theorem holds.

Theorem 5.9 $r^o_{H_d} = 2$.

Proof. By theorem 3.2 $r^o_{H_d} \geq 2$.

In order to prove that also $r^o_{H_d} \leq 2$ holds, observe that given any two nodes v_α and v_β in H_d, there always exists a shortest path from v_α to v_β that goes from v_α to the node $v_{\alpha\cup\beta}$ (where $\alpha\cup\beta$ is the string obtained by the bit-to-bit *or* operation of α and β), and from $v_{\alpha\cup\beta}$ to v_β. Let $\mathcal{P}$ be the set of shortest paths between all pairs source-destination satisfying the above property. Consider now the coloring $\mathcal{C}_{H_d}$ such that $\mathcal{C}_{H_d}(v_\alpha)$ is the number of bits equal to 1 in α. Clearly, in such a definition two adjacent nodes in H_d cannot share the same color (thus satisfying the definition of coloring), and $\mathcal{D}(\mathcal{C}_{H_d}, \mathcal{P}) = 2$. But $\mathcal{P}$ contains at least one shortest path between every pair of nodes, and the result follows directly from theorem 5.1. □

Concerning the fully adaptive case, it is possible to prove the following lower bound.

Theorem 5.10 $r_{H_d} \geq \lceil \frac{d+1}{2} \rceil$.

Proof. Please, check this proof

Given the set $\mathcal{P}$ of all shortest paths between every pair source-destination, consider $\mathcal{P}' \subset \mathcal{P}$ constituted by all shortest paths of length d (i.e. restricted to pairs of opposite nodes at distance d in H_d).

Clearly, since $\mathcal{P}' \subset \mathcal{P}$, $r_{H_d} \geq r_{H_d}(\mathcal{P}')$.

Given any cycle C in H_d of four nodes v_0, v_1, v_2 and v_3, consider the set $S = \{(e_i, e_{(i+1)\,mod\,4}) \mid 0 \leq i \leq 3, e_i = (v_{(i-1)\,mod\,4}, v_i)$ *and* $e_{(i+1)\,mod\,4} = (v_i, v_{(i+1)\,mod\,4})\}$ of adjacent edges along C. In every orientation $\overset{\curvearrowleft}{\rightarrow} H_d)$, any pair $(e_i, e_{(i+1)\,mod\,4}) \in S$ is such that e_i and $e_{(i+1)\,mod\,4}$ either have the same sense (with respect to v_i) or opposite sense. If $p \in \mathcal{P}'$ contains e_i and $e_{(i+1)\,mod\,4}$, and e_i and $e_{(i+1)\,mod\,4}$ have the same sense, then p has to change orientation, that is $p = p' \circ p_1 \circ p_2 \circ p''$, where p_1 ends with e_i and p_2 starts with $e_{(i+1)\,mod\,4}$ or p_1 ends with $e_{(i+1)\,mod\,4}$ and p_2 starts with e_i, so that p_1 and p_2 cannot belong to the same acyclic orientation.

Since the orientations are acyclic, at least two of the four pairs in S must be formed by two edges having the same sense (in the cycle there must be at least one source and one sink), and the ratio between changing pairs and the total number of pairs is at least $\frac{1}{2}$.

By symmetry, every pair of edges in S belongs to the same number of directed paths $p \in \mathcal{P}'$, thus if we denote by $load(C)$ the cardinality of the subset of paths in $\mathcal{P}'$ stepping through two edges of the cycle C, then at least $\frac{load(C)}{2}$ paths in $\mathcal{P}'$ have to change orientation along C.

If we denote by c the number of cycles of length four in the hypercube, since by symmetry every cycle C has the same load $l = load(C)$, by summing up over all the cycles it results that the paths in $\mathcal{P}'$ have to change orientation in total at least $\frac{c \cdot l}{2}$ times. Hence at least one $p \in \mathcal{P}'$ has to change $\frac{c \cdot l}{2 \cdot |\mathcal{P}'|}$ orientations, so that p is covered by at least $\frac{c \cdot l}{2 \cdot |\mathcal{P}'|} + 1$ different orientations.

The theorem follows by observing that $l = \frac{(d-1) \cdot |\mathcal{P}'|}{c}$. In fact, every path $p \in \mathcal{P}'$ increases by one the load of each one of the $d-1$ cycles it shares two edges with, thus the sum of the loads of all cycles is $(d-1) \cdot |\mathcal{P}'|$, and since by symmetry every node has the same load, $l = \frac{(d-1) \cdot |\mathcal{P}'|}{c}$.

Therefore,

$$\frac{c \cdot l}{2 \cdot |\mathcal{P}'|} + 1 \geq \frac{d-1}{2} + 1 = \frac{d+1}{2}.$$

□

Notice that, in the fully adaptive case every acyclic orientation has size at least equal to half of the diameter, while concerning upper bounds the following theorem holds.

Theorem 5.11 $r_{H_d} \leq d + 1$.

Proof. By theorem **??**, every coloring $\mathcal{C}_{H_d}$ has depth $\mathcal{D}(\mathcal{C}_{H_d}) \leq d + 1$, and the result follows directly from theorem 5.1. □

Thus, the lower and the upper bounds differ by a multiplicative factor equal to 1/2. Moreover, it is possible to show that for every acyclic orientation $\overset{\curvearrowleft}{\rightarrow} H_d)$ satisfying the property that all edges in the same dimension are oriented in the same sense $r(\overset{\curvearrowleft}{\rightarrow} H_d)) = d + 1$.

6 Conclusions and open problems

In this paper we have investigated the problem of finding acyclic orientations for communication networks in order to prevent deadlock configurations.

In particular, new results have been presented both from a theoretical computational complexity point of view and from a practical one by providing concrete bounds on deadlock free routing schemes for specific topologies.

One of the main questions left open in this paper is whether or not the problem of minimizing the number of buffers yielded by the acyclic orientations can be approximated in polynomial time when all shortest dipaths between each communication request must be represented.

Concerning the topology dependent results, while tight bounds have been determined for tori, it would be worthwhile to establish the exact order for $q \times q$ grids (we conjecture a value of order $(2-\sqrt{2})q$) and hypercubes of dimension d (we conjecture order d). As the shown results for tori, grids and hypercubes suggest, even for particular cases the task of determining tight bounds is not trivial. Anyway, in all the above cases there exist acyclic orientations of rank at most twice the optimal one.

Finally, it would be worth to extend the known results to more general classes of networks.

References

[1] B. Awerbuch, S. Kutten, and D. Peleg. Efficient deadlock-free routing. In *10th Annual ACM Symposium on Principles of Distributed Computing (PODC)*, pages 177–188, Montreal, Canada, 1991.

[2] F. Belik. An efficient deadlock avoidance technique. *IEEE Trans. on Computers*, 39, 1990.

[3] P.E. Berman, L. Gravano, G.D. Pifarré, and J.L.C. Sanz. Adaptive deadlock and livelock-free routing with all minimal paths in torus networks. In *4th Symposium on Parallel Algorithms and Architectures (SPAA)*, pages 3–12, June 1992.

[4] J.C. Bermond and M. Syska. Routage wormhole et canaux virtuel. In M. Cosnard M. Nivat and Y. Robert, editors, *Algorithmique Parallèle*, pages 149–158. Masson, 1992.

[5] C. Chan and T. Yum. An algorithm for detecting and resolving store-and-forward deadlocks in packet switched networks. In *ICC-86*, pages 114–118, 1986.

[6] K. Mani Chandy, J. Misra, and L. Haas. Distributed deadlock detection. *ACM Transactions in Computer Systems*, 1:144–156, 1983.

[7] J. Cidon, J. Jaffe, and M. Sidi. Global distributed deadlock detection and resolution with finite buffers. In *ICC-86*, pages 124–128, 1986.

[8] Robert Cypher and Luis Gravano. Requirements for deadlock-free, adaptive packet routing. In 11*th Annual ACM Symposium on Principles of Distributed Computing (PODC)*, pages 25–33, Vancouver, Aug 1992.

[9] W. J. Dally and C. L. Seitz. Deadlock-free message routing in multiprocessor interconnection networks. *IEEE Trans. Comp.*, C-36, N.5:547–553, May 1987.

[10] M. Di Ianni, M. Flammini, R. Flammini, and S. Salomone. Systolic acyclic orientations for deadlock prevention. In *2nd Colloquium on Structural Information and Communication Complexity (SIROCCO)*. Carleton University Press, 1995.

[11] J. Duato. Deadlock-free adaptive routing algorithms for multicomputers: evaluation of a new algorithm. In *3rd IEEE Symposium on Parallel and Distributed Processing*, Dec 1991.

[12] J. Duato. On the design of deadlock-free adaptive routing algorithms for multicomputers: theoretical aspects. In *2nd European Conference on Distributed Memory Computing*, LNCS 487, pages 234–243, 1991.

[13] E. Fleury and P. Fraigniaud. Deadlocks in adaptive wormhole routing. Research Report, Laboratoire de l'Informatique du Parallélisme, LIP, École Normale Supérieure de Lyon, 69364 Lyon Cedex 07, France, March 1994.

[14] M.R. Garey and D.S. Johnson. *Computers and Intractability. A Guide to the Theory of NP-completeness.* W.H. Freeman, 1977.

[15] K.D. Gunther. Prevention of deadlock in packet-switched data transport system. *IEEE Trans. on Commun.*, COM-29:512–514, May 1981.

[16] J. Jaffe and M. Sidi. Distributed deadlock resolution in store-and-forward networks. *Algorithmica*, 4:417–436, 1989.

[17] P.M. Merlin and P.J. Schweitzer. Deadlock avoidance in store-and-forward networks: Store and forward deadlock. *IEEE Trans. on Commun.*, COM-28:345–352, March 1980.

[18] G.D. Pifarré, L. Gravano, S.A. Felperin, and J.L.C. Sanz. Fully-adaptive minimal deadlock-free packet routing in hypercube, meshes, and other networks. In *3rd Symposium on Parallel Algorithms and Architectures (SPAA)*, pages 278–290, June 1991.

[19] A.G. Ranade. How to emulate shared memory. In *Foundation of Computer Science*, pages 185–194, 1985.

[20] Jean De Rumeur. *Communication dans les réseaux de processeurs.* Collection Etudes et Recherchers en Informatique. Masson, 1994.

[21] M. Sherman and H. Rudin. Using automated validation techniques to detect lockups in packet switched networks. *IEEE Trans. on Comm.*, COM-30:1762–1767, 1982.

[22] M. Singhal. Deadlock detection in distributed systems. *IEEE Computer*, 22:37–48, 1989.

[23] A.S. Tannenbaum. *Computer Networks.* Englewood Cliffs, Prentice Hall, 1988.

[24] Gerard Tel. *Introduction to Distributed Algorithms.* Cambridge University Press, Cambridge, U.K., 1994.