Efficient Collective Communication in Optical Networks*

Jean-Claude Bermond†
I3S, CNRS
Université de Nice, Sophia Antipolis
06903 Sophia Antipolis Cedex, France

Luisa Gargano‡
Dipartimento di Informatica ed Applicazioni
Università di Salerno
84081 Baronissi (SA), Italy

Stephan Perennes†
I3S, CNRS
Université de Nice, Sophia Antipolis
06903 Sophia Antipolis Cedex, France

Adele A. Rescigno‡
Dipartimento di Informatica ed Applicazioni
Università di Salerno
84081 Baronissi (SA), Italy

Ugo Vaccaro‡
Dipartimento di Informatica ed Applicazioni
Università di Salerno
84081 Baronissi (SA), Italy

Abstract

This paper studies the problems of One–to–All and All–to–All Communication in optical networks. In such networks the vast bandwidth available is utilized through *wavelength division multiplexing*: a single physical optical link can carry several logical signals, provided that they are transmitted on different wavelengths. In this paper we consider both *single–hop* and *multi-hop* optical networks. In single–hop networks the information, once transmitted as light, reaches its destination without being converted to electronic form in between, thus reaching high speed communication. In multihop networks a packet may have to be routed through a few intermediate nodes before reaching its final destination. In both models we give efficient One–to–All and All–to–All Communication algorithms, in terms of time and number of wavelengths. We consider both networks with arbitrary topologies and particular networks of practical interest. Several of our algorithms exhibit optimal performances.

*An extended abstract of this work was presented at *ICALP '96*, Paderborn, Germany.

†Work partially supported by the French GDR/PRC Project PRS and by Galileo Project.

‡Work partially supported by the Italian Ministry of the University and Scientific Research, Project: Algoritmi, Modelli di Calcolo e Strutture Informative, and by Galileo Project.

1 Introduction

Motivations. Optical networks offer the possibility of interconnecting hundreds to thousands of users, covering local to wide area and providing capacities exceeding those of traditional technologies by several orders of magnitude. Optical–fiber transmission systems also achieve very low bit error rate compared to their copper–wire predecessors, typically 10^{-9} compared to 10^{-5}. Optics is thus emerging as a key technology in state–of–the–art communication networks and is expecting to dominate many applications. The most popular approach to realize these high–capacity networks is to divide the optical spectrum into many different channels, each channel corresponding to a different wavelength. This approach, called *wavelength–division multiplexing* (WDM) [12] allows multiple data streams to be transferred concurrently along the same fiber–optic, with different streams assigned separate wavelengths.

The major applications for such networks are video conferencing, scientific visualisation and real-time medical imaging, high–speed super-computing and distributed computing [21, 44, 48]. We refer to the books of Green [21] and McAulay [35] for a presentation of the physical theory and applications of this emerging technology.

In order to state the new algorithmic issues and challenges concerning communication in optical networks, we need first to describe the most accepted models of optical networks architectures.

The Optical Model. In WDM optical networks, the bandwidth available in optical fiber is utilised by partitioning it into several channels, each at a different wavelength. Each wavelength can carry a separate stream of data. In general, such a network consists of routing nodes interconnected by point–to–point unidirectional optic fiber links. Each link can support a certain number of wavelengths. The routing nodes in the network are capable of routing a wavelength coming in on an input port to one or more output ports, independently of the other wavelengths. The same wavelength on two input ports *cannot* be routed to the same output port. WDM lightwave networks can be classified into two categories: *switchless* (also called *broadcast–and–select* or *non–reconfigurable*) and *switched* (also called *reconfigurable*). Each of these in turn can be classified as either *single–hop* (also called *all-optical*) or *multihop* [44]. In switchless networks, the transmission from each station is broadcast to all stations in the network. At the receiver, the desired signal is then extracted from all the signals. These networks are practically important since the whole network can be constructed out of passive optical components, hence it is reliable and easy to operate. However, switchless networks suffer of severe limitations that make problematic their extension to wide area networks. Indeed it has been proven in [1] that switchless networks require a large number of wavelengths to support even simple traffic patterns. Other drawbacks of switchless networks are discussed in [44]. Therefore, optical switches are required to build large networks.

A switched optical network consists of nodes interconnected by point–to–point optic communication lines. Each of the fiber–optic links supports a given number of wavelengths. The nodes can be terminals, switches, or both. Terminals send and receive signals. Switches direct their input signals to one or more of the output links. Each link is bidirectional and actually consists of a pair of unidirectional links [44].

In this paper we consider switched networks with *generalised switches*, as done in [1, 3, 11, 28, 36, 43]. In this kind of networks, signals for different requests may travel on the same communication link into a node v (on different wavelengths) and then exit v along different links. Thus the photonic switch can differentiate between several wavelengths coming along a communication link and direct each of them to a different output of the switch. The only constraint is that no two paths in the network sharing the same optical link have the same wavelength assignment. In switched networks it is possible to "reuse wavelengths" [44], thus obtaining a drastic reduction on the number of required wavelengths with respect to switchless networks [1]. We remark that optical switches *do not* modulate the wavelengths of the signals passing through them; rather, they direct the incoming waves to one or more of their outputs.

Single–hop networks (or all-optical networks) are networks where the information, once transmitted as light, reaches its final destination directly without being converted to electronic form in between. Maintaining the signal in optic form allows to reach high speed in these networks since there is no overhead due to conversions to and from the electronic form. However, engineering reasons [44] suggest that in some situations the multihop approach can be preferable. In these networks, a packet from a terminal node may have to be routed through a few terminal nodes before reaching its final destination. At each terminal node, the packet is converted from light to electronic form and retransmitted on another wavelength. See [37, 38] for more on these questions. In the present paper we consider both switched single–hop and switched multihop networks.

A last word on the model: from our description it follows that a single node could multicast a message along a path in the network in such a way that all nodes along the path get the message. While this is technically feasible, it introduces a phenomenon, called *splitting loss* in the engineering literature, that is considered incompatible with the construction of wide area optical networks. More precisely, if the transmitting station transmits with a power P, then each node in the path will get only a fraction of it, this fraction can be a low as P/n^2 if n is the number of nodes in the path [44]. Wideband optical amplifiers could be used to partially recover from this loss [14] but still they suffer of several limitations [32, 22]. The usual approach to overcome this problem, cfr. [2] is to provide for each source–destination pair a separate path–wavelength pair assignment. In the peresent paper we follow this last point of view.

Our results. In this paper we study the problem of designing efficient algorithms for collective communication in switched optical networks.

Collective communication among the processors is one of the most important issues in multiprocessor systems. The need for collective communication arises in many problems of parallel and distributed computing including many scientific computations [10, 15, 19] and database management [49]. Due to the considerable practical relevance in parallel and distributed computation and the related interesting theoretical issues, collective communication problems have been extensively studied in the literature (see the surveys [23, 27, 20]). In this paper we will consider the design of efficient algorithms for two widely used collective communication operations: *One–to–All Communication* and *All–to–All Communication*. Formally, the One–to–All and All–to–All Communication processes can be described as follows.

One–to–All Communication: One terminal node v, called the source, has a block of data $B(v)$. The goal is to disseminate this block so that each other terminal node in the network gets $B(v)$.

All–to–All Communication: Each terminal node v in the network has a block of data $B(v)$. The goal is to disseminate these blocks so that each terminal node gets all the blocks $B(u)$, for each terminal u in the network.

Although our work seems to be the first to address the problem of collective communication in switched optical networks, there is a substantial body of literature that has considered related problems. Optical routing in arbitrary networks has been recently considered in [1, 3, 36, 43]. The above papers contain also efficient algorithms for routing in networks of practical interest. Routing in hypercube based networks has been considered by [3, 39, 43]. Lower bounds on the number of wavelengths necessary for routing permutations have been given in [39, 4, 42]. All–to–All Communication in broadcast–and–select optical networks has been considered in [1]. Other work related to ours is contained in [17, 25, 18, 26, 27]. In these papers the problem of designing efficient One–to–All and All–to–All Communication algorithms in traditional networks has been considered under the assumption that data exchange can take place through arc–disjoint paths in the network.

In this paper we consider both single-hop and multihop networks. In case of single-hop networks we design One–to–All and All–to–All communication algorithms that do not need buffering at intermediate nodes. The algorithms have to guarantee that there is a path between each pair of nodes requiring communication and no link will carry two different signals on the same wavelength. For our purposes, a wavelength will be an integer in the interval $[1, W]$. Generally, we wish to minimise the quantity W, since the cost of switching and amplification devices depends on the number of wavelengths they handle. For single–hop networks we obtain:

- *Optimal* One–to–All Communication algorithms (Theorems 3.1 and 3.2);
- upper and lower bounds on the minimum number of wavelengths necessary to perform All–to–All Communication in arbitrary graphs in terms of the edge–expansion factor of the graph, *optimal* All–to–All Communication algorithms for rings and hypercubes (Theorems 3.4 and 3.5, respectively).

For multihop networks we derive non–trivial tradeoffs between the number of wavelengths and the number of hops necessary to complete the One–to–All Communication process. We obtain, among several results:

- Asymptotically tight bounds for bounded degree networks (Corollary 4.1);
- tight bounds for hypercubes (Theorem 4.3), and meshes and tori (Theorem 4.4).

Some of our results generalise previously known ones; indeed the results of [17] and [25] can be seen as particular cases of our results, when only *one* wavelength is available.

2 Notations and Definitions

The optical network will be represented as a graph $G = (V(G), E(G))$, where each undirected edge represents a pair of point–to–point unidirectional optical fiber links connecting a pair of nodes [44, 36]. We will use the term edge and link interchangeably, however the term link will always be associated with the direction in which an edge is used, in particular, our algorithms will assign *different* wavelengths to all the signals crossing the same link, i.e., crossing an edge in the same direction.

The number of vertices of G will be always denoted by n. Given $v \in V(G)$, we denote with $d(v)$ the *degree* of v, with $d_{\max}$ and $d_{\min}$ we denote the maximum and minimum degree of G, respectively.

The communication processes are accomplished by a set of calls; a call consists of the transmission of a message from some node x to some destination node y along a path from x to y in G. Each call requires one round and is assigned a fixed wavelength. A node can be involved in an arbitrary number of calls during each round, but we require that if two calls share a link during the same round then they must be assigned different wavelengths.

Given a network G, a node $x \in V(G)$, and an integer t, we denote by $\mathtt{wO}(G, x, t)$ the minimum possible number of *wavelengths* necessary to complete the One–to–All Communication process in G in at most t rounds, when x is the source of the process; we set $\mathtt{wO}(G, t) = \max_{x \in V(G)} \mathtt{wO}(G, x, t)$.

Analogously, with $\mathtt{wA}(G,t)$ we shall denote the minimum possible number of *wavelengths* necessary to complete the All–to–All Communication process in G in at most t rounds.

Given G, a node $x \in V(G)$, and an integer w, we denote by $\mathtt{tO}(G,x,w)$ the minimum possible number of *rounds* necessary to complete the One–to–All Communication process in G using up to w wavelengths per round, when x is the source of the process; we set $\mathtt{tO}(G,w) = \max_{x\in V(G)} \mathtt{tO}(G,x,w)$. We denote by $\mathtt{tA}(G,w)$ the minimum possible number of *rounds* necessary to complete the All–to–All Communication process using up to w wavelengths per round.

The *edge–expansion* $\beta(G)$ of G [31], (also called *isoperimetric number* in [47] and *conductance* in [33]) is the minimum over all subsets of nodes $S \subset V(G)$ of size $|S| \leq n/2$, of the ratio of the number of edges having exactly one endpoint in S to the size of S.

A graph G is k–edge–connected if k is the minimum number of edges to be removed in order to disconnect G, G is *maximally edge–connected* if its edge–connectivity equals its minimum degree.

A *routing* for a graph G is a set of $n(n-1)$ paths $R = \{R_{x,y} \mid x, y \in V(G),\ x \neq y\}$, where $R_{x,y}$ is a path in G from x to y. Given a routing R for the graph G, the load of an edge $e \in E(G)$, denoted by $\mathtt{load}(R,e)$, is the number of paths of R going through e in either direction. The *edge–forwarding index* of G [24], denoted by $\pi(G)$, is the minimum over all routings R for G of the maximum over all the edges of G of the load posed by the routing R on the edge, that is, $\pi(G) = \min_R \max_{e\in E(G)} \mathtt{load}(R,e)$. It is known that [47]

$$\pi(G) \geq \frac{n}{\beta(G)}. \tag{1}$$

Unless otherwise specified, all logarithms in this paper are in base 2.

3 Single–Hop Networks

In this section we study the number of wavelengths necessary to realize the One–to–All and All–to–All Communication processes in single–hop (all-optical) networks.

In the single–hop model it is sufficient to study the number of wavelengths necessary when only *one* communication round is used. Indeed, any one–round algorithm that uses w wavelengths can also be executed in t rounds using $\lceil w/t \rceil$ wavelengths per round, that is,

$$\mathtt{wA}(G,t) \leq \left\lceil \frac{\mathtt{wA}(G,1)}{t} \right\rceil, \qquad \mathtt{wO}(G,t) \leq \left\lceil \frac{\mathtt{wO}(G,1)}{t} \right\rceil. \tag{2}$$

On the other hand, the assumption of a single–hop system implies that if we have a realization of a process in t rounds using up to w wavelengths per round, we can easily obtain a new realization using wt wavelengths and one round. Therefore, in the sequel of this section we will focus on one–round algorithms; we will write $\mathtt{wO}(G)$ and $\mathtt{wA}(G)$ to denote $\mathtt{wO}(G,1)$ and $\mathtt{wA}(G,1)$, respectively.

3.1 One–to–All Communication

The problem here is to set up $n-1$ (light)paths from the source of the One–to–All Communication process to any other node in the network. Given a graph G and a node $v \in V(G)$, when v is the source of the process there must exist at least $(n-1)/d(v)$ paths out of the $n-1$ paths originated at v that share the same edge incident on v. Therefore,

$$\mathtt{wO}(G) \geq \left\lceil \frac{n-1}{d_{\min}(G)} \right\rceil . \tag{3}$$

On the other hand, if G is k–edge–connected, for any source v and any subset of k nodes it is possible to choose k edge–disjoint paths to connect v to these nodes (see [8], Corollary 3, p. 167); it follows that a same wavelength can be assigned to these paths. Therefore,

$$\mathtt{wO}(G) \leq \left\lceil \frac{n-1}{k} \right\rceil . \tag{4}$$

From (3) and (4) we get

Theorem 3.1 *If G is maximally edge–connected then*

$$\mathtt{wO}(G) = \left\lceil \frac{n-1}{d_{\min}(G)} \right\rceil .$$

The above theorem gives the exact value of the number of wavelengths necessary to perform One–to–All Communication in one round in various classes of important networks. By Mader's theorem [34], Theorem 3.1 gives the exact value of $\mathtt{wO}(G)$ for the wide class of vertex–transitive graphs. For other classes of graphs G for which the edge connectivity is equal to $d_{\min}$ and, therefore, for which $\mathtt{wO}(G) = \left\lceil \frac{n-1}{d_{\min}(G)} \right\rceil$, see the survey paper [9].

In case of an arbitrary set of source–destination pairs of nodes requiring communication, the problem of determining the minimum possible number of wavelengths necessary to set up paths between each source–destination pair is NP–complete already for trees and cycles [16]. In contrast, we show that the computation of $\mathtt{wO}(G)$ can be done in polynomial time by computing at most $\log |V|$ maximum flows on a graph with $O(|V|^2)$ nodes and $O(|V||E|)$ edges.

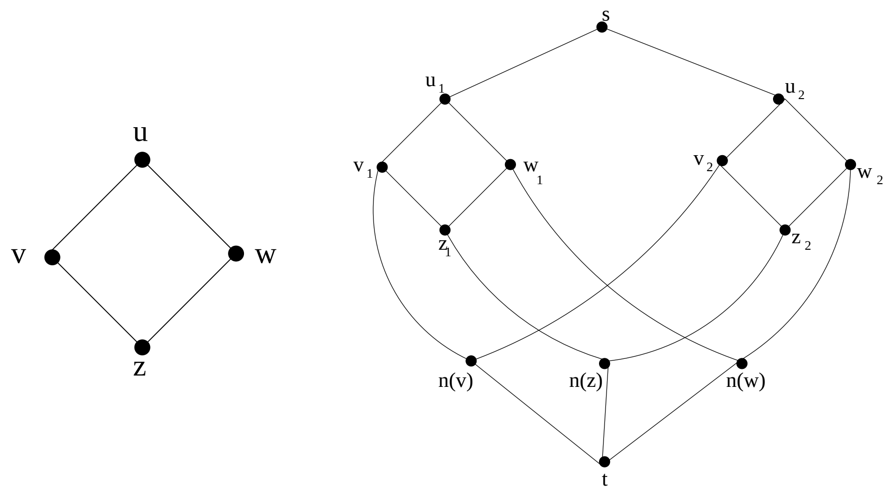
Given $G = (V, E)$, let u be the source of the One–to–All Communication process and w be an integer greater than 0. Construct w copies of G: $G_1 = (V_1, E_1), \ldots, G_w = (V_w, E_w)$; for any $v \in V$, let $v_1, \ldots, v_w$ be the copies of v in $G_1 = (V_1, E_1), \ldots, G_w = (V_w, E_w)$, respectively. Moreover, for any vertex $v \in V - \{u\}$ let $n(v)$ be a new vertex. Define the flow network $G'_w = (V', E')$ as follows:

$$\begin{aligned}
V' &= \left(\bigcup_{i=1}^{w} V_i \right) \cup \{s, t\} \cup \left(\bigcup_{v \in V - \{u\}} n(v) \right) \\
E' &= \bigcup_{i=1}^{w} \{(s, u_i)\} \cup \left\{ \bigcup_{i=1}^{w} E_i \right\} \cup \left(\bigcup_{v \in V - \{u\}} \bigcup_{i=1}^{w} \{(v_i, n(v))\} \right) \cup \left(\bigcup_{v \in V - \{u\}} \{(n(v), t)\} \right) .
\end{aligned}$$

Vertex s is the source and vertex t is the sink of the flow network G'_w. For any $e \in E'$ we set the capacity $c(e)$ of e equal to ∞ if $e = (s, u_i)$, for $i = 1, \ldots, w$, and $c(e) = 1$ otherwise. As an example, when G is the cycle on four vertices in figure 1(a), the flow network G'_2 is represented in figure 1(b).

The way we have constructed the graph $G'_w = (V', E')$ directly implies the following result which, in turn, implies a polynomial time algorithm to compute $\mathtt{wO}(G)$.

Theorem 3.2 *For any graph $G = (V, E)$ we have that $\mathtt{wO}(G) \leq w$ if and only if there is a flow of value $|V| - 1$ in the associated flow network $G'_w = (V', E')$.*



(a) (b)

Figure 1: An example of a network G(a) and its corresponding flow network G'_2(b)

3.2 All–to–All Communication

In this section we study the minimum possible number of wavelengths necessary to perform All–to–All Communication in single–hop networks in exactly one round.

We first notice that, since each node in the graph G has to send its block of information to each other node, the number of paths crossing an edge in either direction cannot be less than the edge–forwarding index of G; since at least half of them cross the same link, we get

Lemma 3.1 *For each graph G,*

$$\mathtt{wA}(G) \geq \pi(G)/2.$$

Minimising the number of wavelengths is in general not the same problem as that of realizing a routing that minimises the number of paths sharing the same link. Indeed, our problem is made much harder due to the further requirement of wavelengths assignment on the paths. In order to get equality in Lemma 3.1 one should find a routing R achieving the bound $\pi(G)$ for which the associated *conflict graph*, that is, the graph with a node for each path in R and an edge between any two paths sharing a link, is $\pi(G)/2$–vertex colorable.

It is possible to put in relation the minimum possible number of wavelengths necessary to perform All–to–All Communication in G in one round with the edge–expansion of G. From Lemma 3.1 and (1) we get the lower bound

$$\mathtt{wA}(G) = n/2\beta(G).$$

An h–relation is a set of communication requests in which each node of G appears at most h time as source of the communication and at most h time as recipient of the communication. Since All–to–All communication corresponds to an $(n-1)$–relation, from [3] we get

Theorem 3.3 [3] *In any bounded degree graph G on n nodes*

$$\mathtt{wA}(G) = O\left(\frac{n \log^2 n}{\beta^2(G)}\right).$$

For each graph G with small edge–expansion $\beta(G)$ the above bound can be very weak; for instance, if G is a rectangular mesh Theorem 3.3 gives $\mathtt{wA}(G) = O(n^3 \log n)$, which is meaningless. In such cases the following bound that follows from (4) for any k–edge connected graph can be more useful

$$\mathtt{wA}(G) \le \left\lceil \frac{n(n-1)}{k} \right\rceil.$$

We show now that for some classes of important networks the lower bound on $\mathtt{wA}(G)$ given in Lemma 3.1 can be efficiently reached.

In case of the path P_n on n nodes it is not hard to prove that the shortest path routing gives a set of paths that can be coloured with an optimal number of colours $\pi(P_n)/2 = \frac{1}{2}\left\lfloor \frac{n^2}{2} \right\rfloor$, so that all paths sharing a link have different colours. In [2] the value $\mathtt{wA}(\cdot)$has been determined for the ring C_n on n nodes, n multiple of 4. In the next theorem we determine $\mathtt{wA}(C_n)$ for any n.

Theorem 3.4 *Let C_n be the ring on n nodes. Then*

$$\mathtt{wA}(C_n) = \left\lceil \frac{\pi(C_n)}{2} \right\rceil = \left\lceil \frac{1}{2} \left\lfloor \frac{n^2}{4} \right\rfloor \right\rceil.$$

Proof. It is known that $\pi(C_n) = \left\lfloor \frac{n^2}{4} \right\rfloor$ [24]. Therefore, from Lemma 3.1 we have $\mathtt{wA}(C_n) \ge \left\lceil \frac{1}{2} \left\lfloor \frac{n^2}{4} \right\rfloor \right\rceil$. We give a routing which attains this bound and we show how to colour the paths of the routing with $\left\lceil \frac{1}{2} \left\lfloor \frac{n^2}{4} \right\rfloor \right\rceil$ colours so that for any edge of C_n all the paths crossing each of the corresponding links have different colours. Let us denote by $\{0, 1, \ldots, n-1\}$ the vertex set of C_n and by $\oplus$ and $\ominus$ the addition and the subtraction modulus n, respectively. For any pair of nodes $x, y \in V(C_n)$, the shortest path from x to y in C_n is unique if either n is odd or n is even and

$y \neq x \oplus n/2$, otherwise we have two shortest paths from x to $x \oplus n/2$. For our purpose, we choose the path $x, x \oplus 1, \ldots, x \oplus n/2$ if x is even and the path $x, x \ominus 1, \ldots, x \ominus n/2$ if x is odd, as the shortest path from x to $x \oplus n/2 = x \ominus n/2$ in C_n.

In the following we assign colours only to the paths $x, x \oplus 1, \ldots, x \oplus \ell$ (denoted by $x \overset{n}{\leadsto} x \oplus \ell$) for any x and ℓ, where

$$\ell \leq \begin{cases} \lfloor n/2 \rfloor & \text{if } n \text{ is odd, or } n \text{ is even and } x \text{ is even,} \\ n/2 - 1 & \text{if } n \text{ is even and } x \text{ is odd.} \end{cases}$$

Indeed, it is possible to use the same colours for the remaining paths which use the links in the reverse direction. For example, for each x and ℓ we can assign to the path $x, x \ominus 1, \ldots, x \ominus \ell$ the same colour assigned to $(x \oplus 1) \overset{n}{\leadsto} (x \oplus 1) \oplus \ell$. To prove the theorem we proceed by induction on the length n of the cycle.

Let $n = 3$. We have just to colour the paths $x \overset{3}{\leadsto} x \oplus 1$, for $x \in \{0, 1, 2\}$. Trivially, one colour suffices (see Figure 2a).

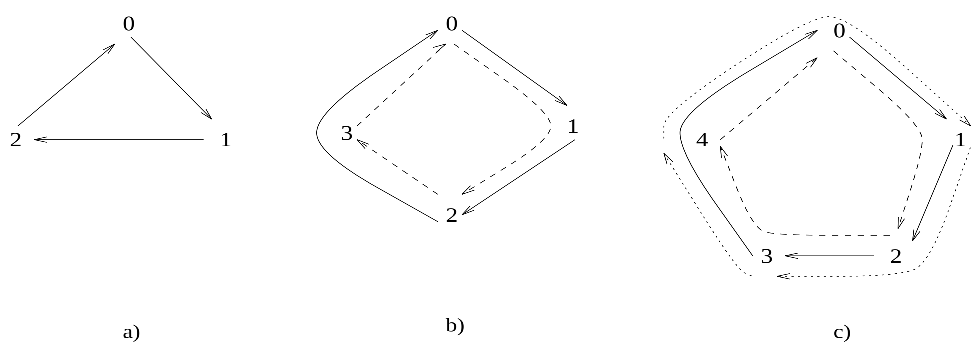


Figure 2

Let n be odd. Suppose by induction that we are able to colour optimally the paths of C_n using the colours $0, 1, \ldots, \mathtt{wA}(C_n) - 1 = \frac{n^2-1}{8} - 1$. Denote by $\mathtt{c}(i \overset{n}{\leadsto} j)$ the colour given in C_n to the path $i \overset{n}{\leadsto} j$. In the following we will colour the paths of C_{n+1} and C_{n+2}.

Case 1. Let us consider the cycle C_{n+2}. We colour the paths of C_{n+2} with colours $\{0, 1, \ldots, \frac{n^2-1}{8} + \frac{n+1}{2} - 1\}$; thus proving that $\mathtt{wA}(C_{n+2}) = \mathtt{wA}(C_n) + \frac{n+1}{2} = \frac{(n+2)^2-1}{8}$. Denote by $\oplus'$ the addition modulus $n+2$. For any $i \in V(C_{n+2})$ and $j = i \oplus' \ell$ with $\ell \leq (n+1)/2$, the path $i, i \oplus' 1, \ldots, j$ will be denoted by $i \leadsto j$. We denote by $\mathtt{c}'(i \leadsto j)$ the colour given to the paths in C_{n+2}.

1) Consider node $i \leq (n-1)/2$ and the path $i \leadsto i \oplus' \ell$, for any $\ell \leq (n+1)/2$. If $\ell \leq (n-1)/2$ then the path $i \leadsto i \oplus' \ell$ in C_{n+2} is made of the same nodes of the path $i \overset{n}{\leadsto} i \oplus \ell$ in C_n.

Therefore, we assign

$$\mathsf{c}'(i \rightsquigarrow i \oplus' \ell) = \begin{cases} \mathsf{c}(i \overset{n}{\rightsquigarrow} i \oplus \ell) & \text{if } \ell \leq (n-1)/2, \\ \frac{n^2-1}{8} + i & \text{if } \ell = (n+1)/2. \end{cases}$$

2) Consider node $k_i = (n+1)/2 + i$, with $i = 0, \ldots, (n-3)/2$. For each $\ell \leq (n+1)/2$ we assign

$$\mathsf{c}'(k_i \rightsquigarrow k_i \oplus' \ell) = \begin{cases} \mathsf{c}(k_i \overset{n}{\rightsquigarrow} i) & \text{if } k_i \oplus' \ell = n, \\ \frac{n^2-1}{8} + i & \text{if } k_i \oplus' \ell = n+1, \\ \mathsf{c}(k_i \overset{n}{\rightsquigarrow} k_i \oplus \ell) & \text{otherwise.} \end{cases}$$

3) Consider node n. For each $i = n+1, 0, \ldots, (n-3)/2$, we assign

$$\mathsf{c}'(n \rightsquigarrow i) = \begin{cases} \frac{n^2-1}{8} + \frac{n-1}{2} & \text{if } i = n+1, \\ \mathsf{c}(k_i \overset{n}{\rightsquigarrow} i) & \text{otherwise.} \end{cases}$$

4) Consider node $n+1$. For each $i = 0, \ldots, (n-1)/2$, we assign

$$\mathsf{c}'((n+1) \rightsquigarrow i) = \frac{n^2-1}{8} + i.$$

The colouring of the paths of C_3 and the corresponding colouring of the paths of C_5 are shown in Figure 2a) and 2c).

We now check that for any colour $\mathsf{c} \in \{0, 1, \ldots, \frac{n^2-1}{8} + \frac{n+1}{2} - 1\}$, any link is crossed by at most one path of colour c.

Let c be such that $0 \leq \mathsf{c} \leq \frac{n^2-1}{8} - 1$ and let $\mathcal{P}$ be the set of paths coloured c in C_n. Notice that the paths in $\mathcal{P}$ are originated at i, for $i \leq \frac{n-1}{2}$, and at k_i, for $i \leq \frac{n-3}{2}$; furthermore, only the paths originated at k_i include node 0. Then, we can distinguish the following cases

- By 1) we have that if $i \rightsquigarrow i \oplus' \ell$ is coloured c in C_{n+2} then the path $i \overset{n}{\rightsquigarrow} i \oplus \ell$ is in $\mathcal{P}$.
- By 2) and 3) we have that if $k_i \rightsquigarrow n$ and $n \rightsquigarrow i$ are coloured c in C_{n+2} then $k_i \overset{n}{\rightsquigarrow} i$ is in $\mathcal{P}$.
- By 2) we have that if $k_i \rightsquigarrow k_i \oplus' \ell = (k_i, \ldots, n, n+1, 0, \ldots, k_i \oplus' \ell)$ is coloured c in C_{n+2} then the path $k_i \overset{n}{\rightsquigarrow} k_i \oplus \ell$ is in $\mathcal{P}$.

Since, by the inductive hypothesis, any link is crossed by at most one path of colour c in C_n we have that any link is crossed by at most one path of colour c in C_{n+2}.

Let $\mathsf{c}_i = \frac{n^2-1}{8} + i$, for $i = 0, \ldots, \frac{n-1}{2}$. The paths of colour c_i in C_{n+2} are

$i \rightsquigarrow i \oplus' \frac{n+1}{2}$ by 1) $\quad i \oplus' \frac{n+1}{2} \rightsquigarrow n+1$ by 2), $\quad$ and, in case $i = n+1$, also $n+1 \rightsquigarrow i$ by 3).

Since these paths are edge-disjoint, any link is crossed by at most one path of colour c_i in C_{n+2}.

Case 2. Let us consider the cycle C_{n+1}. Denote by $\oplus'$ the addition modulus $n+1$; furthermore, for any $i \in V(C_{n+1})$ and $j = i \oplus' \ell$ with

$$\ell \leq \begin{cases} \frac{n-1}{2} & \text{if } i \text{ is odd,} \\ \frac{n+1}{2} & \text{if } i \text{ is even,} \end{cases} \tag{5}$$

the path $i, i \oplus' 1, \ldots, j$ in C_{n+1} will be denoted by $i \rightsquigarrow j$. We have

$$\pi(C_{n+1}) = \left\lceil (n+1)^2/8 \right\rceil = (n^2-1)/8 + \lceil n/4 \rceil = \pi(C_n) + \lceil n/4 \rceil . \tag{6}$$

We will optimally colour the paths of C_{n+1} using the colours $0, 1, \ldots, \frac{n^2-1}{8} + \lceil \frac{n}{4} \rceil - 1$ (cfr. (6)). We denote by $\mathsf{c}'(i \rightsquigarrow j)$ the colour given to the paths in C_{n+1}.

1) Consider node $i \leq (n-1)/2$ and the path $i \rightsquigarrow i \oplus' \ell$, for ℓ as in (5). If $\ell \leq (n-1)/2$ then the path $i \rightsquigarrow i \oplus' \ell$ in C_{n+1} coincides with the path $i \overset{n}{\rightsquigarrow} i \oplus \ell$ in C_n and we assign

$$\mathsf{c}'(i \rightsquigarrow i \oplus' \ell) = \begin{cases} \mathsf{c}(i \overset{n}{\rightsquigarrow} i \oplus \ell) & \text{if } \ell \leq (n-1)/2 \\ \frac{n^2-1}{8} + i/2 & \text{if } i \text{ is even and } \ell = (n+1)/2. \end{cases}$$

2) Consider node $k_i = (n+1)/2 + i$, with $i = 0, \ldots, (n-3)/2$. For each ℓ as in (5) we assign

$$\mathsf{c}'(k_i \rightsquigarrow k_i \oplus' \ell) = \begin{cases} \frac{n^2-1}{8} + \left\lfloor \frac{i}{2} \right\rfloor & \text{if } k_i \oplus' \ell = n \text{ and } k_i \text{ is even,} \\ \mathsf{c}(k_i \overset{n}{\rightsquigarrow} i) & \text{if } k_i \oplus' \ell = n \text{ and } k_i \text{ is odd,} \\ \mathsf{c}(k_i \overset{n}{\rightsquigarrow} k_i \oplus \ell) & \text{otherwise.} \end{cases}$$

3) Consider node n. For each $i = 0, \ldots, (n-1)/2 - 1$, we assign

$$\mathsf{c}'(n \rightsquigarrow i) = \begin{cases} \frac{n^2-1}{8} + \frac{i}{2} & \text{if } i \text{ is even,} \\ \mathsf{c}(k_i \overset{n}{\rightsquigarrow} i) & \text{if } i \text{ is odd and } n \equiv 3 \pmod 4 , \\ \mathsf{c}(k_{i+1} \overset{n}{\rightsquigarrow} i+1) & \text{if } i \text{ is odd and } i \leq \frac{n-1}{2} - 3 \text{ and } n \equiv 1 \pmod 4 , \\ \frac{n^2-1}{8} + \frac{n-1}{4} & \text{if } i = \frac{n-1}{2} - 1 \text{ and } n \equiv 1 \pmod 4). \end{cases}$$

The colouring of the paths of C_4 from that of the paths of C_3 is shown in Figure 2(b).

We now check that for any $\mathsf{c} = 0, \ldots, \frac{n^2-1}{8} + \lceil \frac{n}{4} \rceil - 1$ any link is crossed by at most one path of colour c.

Consider first $\mathsf{c} \leq \frac{n^2-1}{8} - 1$ and let $\mathcal{P}$ be the set of paths coloured c in C_n. Notice that the paths in $\mathcal{P}$ are originated at i, for $i \leq \frac{n-1}{2}$, and at k_i, for $i \leq \frac{n-3}{2}$; furthermore, only the paths originated at k_i include node 0. We distinguish the following cases

- By 1) we have that if $i \rightsquigarrow i \oplus' \ell$ is coloured c in C_{n+1} then the path $i \overset{n}{\rightsquigarrow} i \oplus \ell$ is in $\mathcal{P}$.
- Let $n \equiv 1 \pmod 4$. By 2) and 3) we have that if $k_i \rightsquigarrow n$ and $n \rightsquigarrow i-1$ are coloured c in C_{n+1}, that is k_i is odd (and then i is even), then the path $k_i \overset{n}{\rightsquigarrow} i$ is in $\mathcal{P}$, for $i > 0$. Furthermore, by 2) we have that if $k_0 \rightsquigarrow n$ is coloured c in C_{n+1} then the path $k_0 \overset{n}{\rightsquigarrow} 0$ is in $\mathcal{P}$.

Let $n \equiv 3 \pmod 4$. By 2) and 3) we have that if $k_i \rightsquigarrow n$ and $n \rightsquigarrow i$ are coloured c in C_{n+1}, that is k_i is odd (and then i is odd), then $k_i \overset{n}{\rightsquigarrow} i$ is in $\mathcal{P}$.

- By 2) we have that if $k_i \rightsquigarrow k_i \oplus' \ell = k_i, k_i \oplus' 1, \ldots, n, 0, \ldots, k_i \oplus' \ell$ is coloured c in C_{n+1} then the path $k_i \overset{n}{\rightsquigarrow} k_i \oplus \ell$ is in $\mathcal{P}$.

Since, by the inductive hypothesis, any link is crossed by at most one path of colour c in C_n we have that any link is crossed by at most one path of colour c in C_{n+1}.

Let $\mathsf{c}_i = \frac{n^2-1}{8} + i$, for $i = 0, \ldots, \lceil \frac{n}{4} \rceil - 1$. If $n \equiv 1 \pmod 4$ then paths of colour c_i in C_{n+1} are

$2i \rightsquigarrow 2i \oplus' \frac{n+1}{2}$ for $i \leq \frac{n-1}{4}$ (by 1)),

$(2i+1) \oplus' \frac{n+1}{2} \rightsquigarrow n$ for $i < \frac{n-1}{4}$ (by 2)),

$n \rightsquigarrow 2i$ for $i < \frac{n-1}{4}$, and $n \rightsquigarrow \frac{n-1}{2} - 1$ for $i = \frac{n-1}{4}$ (by 3)).

If $n \equiv 3 \pmod 4$ then paths of colour c_i in C_{n+1} are

$2i \rightsquigarrow 2i \oplus' \frac{n+1}{2}$ (by 1)), $2i \oplus' \frac{n+1}{2} \rightsquigarrow n$ (by 2)), $n \rightsquigarrow 2i$ (by 3)).

Therefore, any link is crossed by at most one path of colour c_i in C_{n+1}. □

In the next theorem we determine $\mathtt{wA}(\cdot)$ for the d–dimensional hypercube.

Theorem 3.5 *Let H_d be the d–dimensional hypercube. We have*

$$\mathtt{wA}(H_d) = \pi(H_d)/2 = 2^{d-1}.$$

Proof. It is known that $\pi(H_d) = 2^d$ [24]. Therefore, from Lemma 3.1 we have $\mathtt{wA}(H_d) \geq 2^{d-1}$. We give a routing which attains this bound and we show how to colour the paths of the routing so that for any link all the 2^{d-1} paths crossing that link have different colours.

A path $(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k)$ from node $\mathbf{x}_0$ to $\mathbf{x}_k$ is called *ascending* if for each $i = 1, \ldots, k$ the node $\mathbf{x}_i$ is obtained from $\mathbf{x}_{i-1}$ by complementing the bit in position p_i, with $p_1 < p_2 < \ldots < p_k$; the ascending path from $\mathbf{x}_0$ to $\mathbf{x}_k$ will be denoted by $\mathbf{x}_0 \rightsquigarrow \mathbf{x}_k$. We will consider ascending paths only.

Let us denote by $\oplus$ the componentwise vector addition modulo 2 and by $\mathbf{e}_i \in \{0,1\}^d$ the vector with i-th component equal to 1 and all the remaining components equal to 0. We first assign a colour $\mathsf{c}(\mathbf{x})$ to each $\mathbf{x} \in \{0,1\}^d$ so that that for each $\mathbf{x}, \mathbf{y} \in \{0,1\}^d$

$$\mathsf{c}(\mathbf{x}) = \mathsf{c}(\mathbf{y}) \quad \text{if and only if} \quad \mathbf{y} = \overline{\mathbf{x}}, \tag{7}$$

where $\overline{\mathbf{x}}$ represents the binary complement of $\mathbf{x}$. This requires 2^{d-1} colours. To each path $\mathbf{v} \rightsquigarrow \mathbf{u}$ we assign colour

$$\mathsf{c}(\mathbf{v} \rightsquigarrow \mathbf{u}) = \mathsf{c}(\mathbf{v} \oplus \mathbf{u}). \tag{8}$$

We prove now that each link $(\mathbf{z}, \mathbf{z} \oplus \mathbf{e}_i)$ is crossed by exactly one path of any colour. Since we are considering ascending paths, the link $(\mathbf{z}, \mathbf{z} \oplus \mathbf{e}_i)$ is crossed only by paths $\mathbf{s} \rightsquigarrow \mathbf{x}$ with

$$\mathbf{s} = s_1 \ldots s_{i-1} z_i \ldots z_d, \quad \text{and} \quad \mathbf{x} = z_1 \ldots z_{i-1} \overline{z}_i x_{i+1} \ldots x_d. \tag{9}$$

Let $\mathbf{a} = s_1 \ldots s_{i-1} \overline{z}_i x_{i+1} \ldots x_d$, by (8) and (9) we have

$$\mathsf{c}(\mathbf{s} \rightsquigarrow \mathbf{x}) = \mathsf{c}(\mathbf{s} \oplus \mathbf{x}) = \mathsf{c}(\mathbf{z} \oplus \mathbf{a}). \tag{10}$$

Consider now any other path $\mathbf{s}' \rightsquigarrow \mathbf{x}'$ crossing the link $(\mathbf{z}, \mathbf{z} \oplus \mathbf{e}_i)$. By (9) and (10), we have

$$\mathsf{c}(\mathbf{s}' \rightsquigarrow \mathbf{x}') = \mathsf{c}(\mathbf{s}' \oplus \mathbf{x}') = \mathsf{c}(\mathbf{z} \oplus \mathbf{a}'), \tag{11}$$

where $\mathbf{a}' = s'_1 \dots s'_{i-1} \overline{z}_i x'_{i+1} \dots x'_d$.
It is immediate to see that $\mathbf{a}' \neq \overline{\mathbf{a}}$ and, by (9), that $\mathbf{a} = \mathbf{a}'$ only if $\mathbf{s} = \mathbf{s}'$ and $\mathbf{x} = \mathbf{x}'$. Therefore, by (7), (8), (10) and (11) we get $\mathsf{c}(\mathbf{s} \rightsquigarrow \mathbf{x}) \neq \mathsf{c}(\mathbf{s}' \rightsquigarrow \mathbf{x}')$. □

4 Multihop Networks

In this section we show that by exploiting the capabilities of the multihop optical model, a drastic reduction on the number of wavelengths can be obtained with respect to (2).

In the following, we will be mostly interested in investigating One–to–All Communication algorithms. Indeed, as it is well known, the All–to–All Communication process can be accomplished by first accumulating all blocks at one node and then spreading the resulting message from this node. Since accumulation process corresponds to the inverse process of One–to–All Communication we get the obvious result:

Lemma 4.1 *For each graph G and number of wavelengths w*

$$\mathtt{tO}(G, w) \leq \mathtt{tA}(G, w) \leq 2\, \mathtt{tO}(G, w).$$

4.1 Lower Bounds

Lemma 4.2 *For each graph G on n nodes of minimum degree $d_{\min}$ and maximum degree $d_{\max}$*

$$\mathtt{tO}(G, w) \geq \left\lceil \frac{\log(1 + (n-1) d_{\max}/d_{\min})}{\log(w d_{\max} + 1)} \right\rceil. \tag{12}$$

Proof. Let the source of the One–to–All Communication process be a node x of degree $d(x) = d_{\min}$. Indicate by n_i the maximum number of nodes that can have received the data $B(x)$ within the i-th round; initially we have $n_0 = 1$.
During round $i \geq 1$ node x can send the message $B(x)$ up to $w d_{\min}$ nodes, whereas any node y that has received the message by round $i - 1$ can send it to $w d(y) \leq w d_{\max}$ other nodes. Therefore, we have

$$n_i \leq n_{i-1} + w d_{\min} + (n_{i-1} - 1) w d_{\max} = n_{i-1}(w d_{\max} + 1) - (d_{\max} - d_{\min}) w, \tag{13}$$

By iterating (13), in view of $n_0 = 1$ we get

$$n_i \leq (w d_{\max} + 1)^i \, (d_{\min}/d_{\max}) + 1 - d_{\min}/d_{\max}. \tag{14}$$

Since it is possible to complete the One–to–All Communication in t rounds only if $t \geq \min\{i \mid n_i \geq n\}$, from (14) we get the following inequality

$$n \leq (wd_{\max}+1)^t \frac{d_{\min}}{d_{\max}} + 1 - \frac{d_{\min}}{d_{\max}}$$

that implies (12). $\square$

The following lemma allows to find lower bounds on $\mathtt{tA}(G,w)$ in terms of $\mathtt{tO}(G,w)$ and $\pi(G)$.

Lemma 4.3 *Let G be a graph on n nodes of maximum degree d, and w be a positive integer. The following relation between $\mathtt{tA}(G,w)$, $\mathtt{tO}(G,w)$ and $\pi(G)$ holds:*

$$2(n-1)\frac{(wd+1)^{\mathtt{tA}(G,w)-\mathtt{tO}(G,w)}-1}{wd} + (2\mathtt{tO}(G,w)-\mathtt{tA}(G,w))(wd+1)^{\mathtt{tA}(G,w)-1} \geq \pi(G)/(2w). \quad (15)$$

Proof. Let $t = \mathtt{tA}(G,w)$ be the number of round of the All–to–All Communication process and $\mathtt{t}_0 = \mathtt{tO}(G,w)$. We first notice that, from Lemma 4.1

$$\mathtt{t}_0 \leq t \leq 2\mathtt{t}_0. \quad (16)$$

Fix an edge (x,y) and consider a round i, with $1 \leq i \leq t$. In this round i there are up to w messages that cross the link (x,y) from x to y, say $M_1,\ldots,M_{w'}$, $w' \leq w$, originated in some node x_j and destinate to some y_j. Let b_i be the total number of nodes that will receive at least one block contained in $M_1,\ldots,M_{w'}$, in one of the rounds $i, i+1,\ldots,t$. Obviously, $\sum_{i=1}^t b_i$ represents the load posed by the All–to–All Communication process on the link (x,y), therefore

$$\sum_{i=1}^{t} b_i \geq \pi(G)/2. \quad (17)$$

We now want to upper bound each b_i. We first notice that since node y_j, $1 \leq j \leq w'$, receives message M_j at round i, during the subsequent rounds from $i+1$ to t, node y_j can disseminate these blocks to at most

$$Y_i = \begin{cases} (wd+1)^{t-i} & \text{if } t-i < \mathtt{t}_0 \\ n-1 & \text{if } t-i \geq \mathtt{t}_0. \end{cases} \quad (18)$$

nodes (other than the sender x_j of the message M_j) (cfr. (14), noticing that $d_{\min} \leq d_{\max} = d$). We evaluate now the size (number of blocks) of each M_j. Since x_j sends M_j at round i, then M_j can contain only the blocks known to x_j within i rounds, therefore for each $j = 1,\ldots,w'$, the size of M_j is at most

$$m_i = \begin{cases} (wd+1)^{i-1} & \text{if } i-1 < \mathtt{t}_0 \\ n-1 & \text{if } i-1 \geq \mathtt{t}_0, \end{cases} \quad (19)$$

(cfr. (14) and notice that we do not count in M_j the eventual block of the receiving node y_j).

Formulæ (16), (18), and (19) give

$$b_i \leq \sum_{j=1}^{w'} m_i Y_i \leq w \begin{cases} (n-1)(wd+1)^{i-1} & \text{if } i \leq t - \mathtt{t}_0 \\ (wd+1)^{t-1} & \text{if } t - \mathtt{t}_0 < i \leq \mathtt{t}_0 \\ (wd+1)^{t-i}(n-1) & \text{if } \mathtt{t}_0 < i, \end{cases}$$

for each $i = 1, \ldots, t$, and

$$\begin{aligned} \sum_{i=1}^{t} b_i &\leq w \left[\sum_{i=1}^{t-\mathtt{t}_0} (n-1)(wd+1)^{i-1} + \sum_{i=t-\mathtt{t}_0+1}^{\mathtt{t}_0} (wd+1)^{t-1} + \sum_{i=\mathtt{t}_0+1}^{t} (n-1)(wd+1)^{t-i} \right] \\ &= \frac{2(n-1)((wd+1)^{t-\mathtt{t}_0} - 1)}{d} + w(wd+1)^{t-1}(2\mathtt{t}_0 - t). \end{aligned}$$

Therefore, from (17) we get

$$\frac{2(n-1)((wd+1)^{t-\mathtt{t}_0} - 1)}{d} + w(wd+1)^{t-1}(2\mathtt{t}_0 - t) \geq \sum_{i=1}^{t} b_i \geq \pi(G)/2 \tag{20}$$

and the lemma holds □

Assume that $\mathtt{tO}(G, w) \leq \log_{wd+1} cn$, for some constant $c \geq 1$. It is easy to see under this hypothesis that from formula (15) we get

$$\mathtt{tA}(G, w) \geq \mathtt{tO}(G, w) + \log_{wd+1} \frac{d\pi(G)}{n} - O(\log_{wd+1} \mathtt{tO}(G, w)). \tag{21}$$

In case G is the cycle C_n on n vertex, it is immediate to see that $\mathtt{tO}(C_n, w) = \lceil \log_{wd+1} n \rceil$, therefore from above inequality we get

$$\mathtt{tA}(C_n, w) \geq \mathtt{tO}(C_n, w) + \log_{wd+1} n - O(\log_{wd+1} \log_{wd+1} n)$$

which shows that the trivial upper bound on $\mathtt{tA}(G, w)$ given in Lemma 4.1 is almost tight for the cycle C_n. Finally, we also remark that Lemma 4.3 allows to get a lower bound in terms of the edge bisection width of G. In fact, if G has edge bisection width k one can use the relation

$$\pi(G) \geq \frac{n^2}{4k}$$

and proceed as above to get lower bounds on $\mathtt{tA}(G, w)$ in terms of n and k. An analogous bound in terms of vertex bisection width when $w = 1$ has been given in [29].

4.2 Upper Bounds

In order to obtain our general upper bound on the number of rounds to perform One–to–All communication in G with a fixed number of wavelengths, we need the following covering property.

Definition 4.1 *An s–*tree cover *for a graph $G = (V, E)$ is a family $\mathcal{F}$ of edge–disjoint subtrees of G such that:*

1. $\cup_{F \in \mathcal{F}} V(F) = V$;
2. *For each $F, F' \in \mathcal{F}$ it holds $|V(F) \cap V(F')| \leq 1$;*
3. *For each $F \in \mathcal{F}$ it holds $|V(F)| \leq s$.*

The s–tree cover number of G is the minimum size of an s–tree cover for G.

The following result upper bounds the s–tree cover number of a graph; its proof also furnishes an efficient way to determine an s–tree cover which attains the bound. The proof is in Appendix A.

Lemma 4.4 *For each graph G on n nodes and bound s, the s–tree cover number of G is upper bounded by $2n/s$.*

Before giving the upper bound on the time to perform One–to–All Communication in general graphs, we notice the following application of Lemma 4.4 to the function $\mathtt{wO}(\cdot)$.

Theorem 4.1 *For each k–edge connected graph G on n nodes*

$$\left\lceil \frac{\sqrt{1+(n-1)d_{\max}/d_{\min}}-1}{d_{\max}} \right\rceil \leq \mathtt{wO}(G,2) \leq \left\lceil \sqrt{\frac{2n}{k}} \right\rceil .$$

Proof. The lower bound follows from Lemma 4.2. Let $s = \left\lceil \sqrt{2n/k} \right\rceil$, by Lemma 4.4 we can construct an s–tree cover $\mathcal{F} = \{F_1, \ldots, F_p\}$ for G with

$$p \leq 2n/\lceil \sqrt{2n/k} \rceil \quad \text{and} \quad |F_i| \leq s = \left\lceil \sqrt{2n/k} \right\rceil, \text{ for } i = 1, \ldots, p.$$

Since G is k–edge connected, it is possible to find k edge–disjoint paths connecting the source of the One–to–All Communication process to k arbitrary other nodes in the graph. From this we get that in the first round of the One–to–All Communication process it is possible to inform one node in each F_i, for $i = 1, \ldots, p$, using at most

$$\lceil p/k \rceil \leq \lceil \sqrt{2n/k} \rceil$$

wavelengths.
Since no two elements of $\mathcal{F}$ share an edge, in the second round the informed nodes of each tree F_i can independently disseminate the information to all the other nodes of F_i using at most

$$|F_i| - 1 < s = \lceil \sqrt{2n/k} \rceil$$

wavelengths. □

By using Lemma 4.4 we can prove a general upper bound on $\mathtt{tO}(G, w)$ for any $w \geq 2$; in the case $w = 1$ the bound $\mathtt{tO}(G, 1) \leq \lceil \log n \rceil$ has been given in [17].

Theorem 4.2 *For each graph G on n nodes and number of wavelengths $w \geq 2$*

$$\mathtt{tO}(G,w) \leq \lceil \log n / (\log(w+1) - 1) \rceil .$$

Proof. Let $s = \lceil \frac{2n}{w+1} \rceil$. By Lemma 4.4 we can construct for G an s–tree cover $\mathcal{F} = \{F_1, \ldots, F_p\}$; with

$$p \leq \frac{2n}{\lceil 2n/(w+1) \rceil} \leq w+1 \quad \text{and} \quad |F_i| \leq s = \left\lceil \frac{2n}{w+1} \right\rceil, \text{ for } i = 1, \ldots, p.$$

In the first round the source of the process v can inform one node in each F_i, for $i = 1, \ldots, p$, apart the one containing v itself. Since no two trees in $\mathcal{F}$ share an edge the process can proceed independently and recursively in each tree $F_i \in \mathcal{F}$. Therefore, $\mathtt{tO}(G,w) \leq \lceil \log n / (\log(w+1) - 1) \rceil$. □

By Lemma 4.2 and Theorem 4.2 we get

Corollary 4.1 *For each bounded degree graph G on n nodes*

$$\mathtt{tO}(G,w) = \Theta(\log_{w+1} n).$$

We give now a sharper bound on the time to perform One–to–All Communication in the d-dimensional hypercube in terms of the maximum number of wavelengths. In the special case $w = 1$ it is proved in [25] that $\mathtt{tO}(H_d, 1) = \Theta(d/\log d)$.

Theorem 4.3 *For each d and number of wavelengths w*

$$\left\lceil \frac{d}{\log(wd+1)} \right\rceil \leq \mathtt{tO}(H_d, w) \leq c(d,w) \frac{d}{\lfloor \log(wd+1) \rfloor} + 2$$

with $c(d,w) \leq 4$ for any $d \geq 3$, and $\lim_{d \to \infty} c(d,w) \leq \begin{cases} 1 & \text{if } \log w = o(2^d), \\ 1 + \frac{\log e}{e} & \text{otherwise.} \end{cases}$

Proof. The lower bound is given in Lemma 4.2. We prove here the upper bound. Given a sequence $\mathbf{a} = a_1 \ldots a_L \in \{0,1\}^L$, for some $1 \leq L \leq d-1$, let us denote by $H(\mathbf{a})$ the subcube of dimension $d - L$ of H_d consisting of all nodes $\mathbf{x} = x_1 \ldots x_{d-L}\mathbf{a}$.

We recall that a path $(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_k)$ from node $\mathbf{x}_0$ to node $\mathbf{x}_k$ is called ascending if for each $i = 1, \ldots, k$ the node $\mathbf{x}_i$ is obtained from $\mathbf{x}_{i-1}$ by complementing the bit in position p_i, with $p_1 < p_2 < \ldots < p_k$. Without loss of generality we assume that the source of the One–to–All Communication process is node $\mathbf{0}$. Let

$$L = \lfloor \log(wd+1) \rfloor, \tag{22}$$

and $A = \{0,1\}^L - \{0^L\}$ be the set of all sequences of length L containing at least one 1. We first establish in H_d paths from $\mathbf{0}$ to a node in each subcube $H(\mathbf{a})$, for $\mathbf{a} \in A$, so that any link is crossed by no more than w paths. The paths are assigned as follows:

i) Select in A pairwise disjoint subsets $A_1, \ldots, A_L$ such that

$$A_i \subset \{\mathbf{v} = v_1 \ldots v_L \; | \mathbf{v} \in \{0,1\}^L \text{ and } \; v_i = 1\} \quad \text{and} \quad |A_i| = w, \;\; \text{for each } i = 1, \ldots, L;$$

if $2^L - 1 \leq wL$ then $A_1 = \ldots = A_L = \emptyset$.
For each $\mathbf{v} \in A_i$, for $i = 1, \ldots, L$, the path $P(\mathbf{v})$ from $\mathbf{0}$ to $\mathbf{0}^{d-L}\mathbf{v}$ is obtained as follows: if $v_1 = \ldots = v_{i-1} = 0$ then $P(\mathbf{v})$ is the ascending path from $\mathbf{0}$ to $\mathbf{0}^{d-L}\mathbf{v}$, otherwise $P(\mathbf{v})$ is formed by the ascending path from $\mathbf{0}$ to $\mathbf{0}^{d-L+i-1}v_i \ldots v_L$ followed by the ascending path from $\mathbf{0}^{d-L+i-1}v_i \ldots v_L$ to the destination node $\mathbf{0}^{d-L}\mathbf{a} = \mathbf{0}^{d-L}v_1 \ldots v_L$.

ii) Consider now the set of sequences $B = A - \cup_{i=1}^{L} A_i = \{\mathbf{b}_1, \ldots, \mathbf{b}_{2^L-1-wL}\}$. By (22), we can assign to each $\mathbf{b} \in B$ an integer $f(\mathbf{b}) \leq d - L$ so that no more than w element of B have the same value of f. Let $\mathbf{0}^{d-L}\mathbf{b} \oplus \mathbf{e}_{f(\mathbf{b})}$ be the node obtained from $\mathbf{0}^{d-L}\mathbf{b}$ by complementing the bit in position $f(\mathbf{b})$. The path $P(\mathbf{b})$ is formed by the link $(\mathbf{0}, \mathbf{e}_{f(\mathbf{b})})$ followed by the ascending path from $\mathbf{e}_{f(\mathbf{b})}$ to the end node $\mathbf{e}_{f(\mathbf{b})} \oplus \mathbf{0}^{d-L}\mathbf{b}$.

The above set of paths $P(\mathbf{v})$, for $\mathbf{v} \in A$, establishes in H_d paths from $\mathbf{0}$ to one node in each subcube $H(\mathbf{v})$ so that any link is crossed by no more than w paths. Therefore, in the first round the source $\mathbf{0}$ can send out the information along the paths $P(\mathbf{v})$, for $\mathbf{v} \in A$, and inform one node in each $(d-L)$–dimensional subcube $H(\mathbf{a})$, $\mathbf{a} \in \{0,1\}^L$; in $H(\mathbf{0})$ the informed node is the source $\mathbf{0}$.
In the subsequent rounds each node can iterate the process independently in the $(d-L)$–dimensional subcube to which it belongs.

The above reasoning implies that in one round the given communication algorithms reduces the dimension of the problem from d to $d - \lfloor \log(wd+1) \rfloor$, that is,

$$\mathtt{tO}(H_d, w) \leq 1 + \mathtt{tO}(H_{d-\lfloor \log(wd+1) \rfloor}, w). \tag{23}$$

We show now that (23) gives the desired upper bound on $\mathtt{tO}(H_d, w)$. Let us first notice that $\mathtt{tO}(H_d, w) = 1$ whenever $w \geq (2^d - 1)/d$. Let then

$$w = (2^{\alpha d} - 1)/d \tag{24}$$

for some $0 \leq \alpha < 1$; this implies $\lfloor \log(wd+1) \rfloor = \lfloor \alpha d \rfloor$.
Define Δ as the maximum integer such that $w \geq (2^\Delta - 1)/\Delta$. By (23) we have

$$\mathtt{tO}(H_d, w) \leq \left\lceil \frac{(wd+1) - 2^{\lfloor \alpha d \rfloor}}{\lfloor \alpha d \rfloor w} \right\rceil + \sum_{i=\Delta}^{\lfloor \alpha d \rfloor - 1} \left\lceil \frac{2^i}{wi} \right\rceil + 1. \tag{25}$$

Therefore,

$$\mathtt{tO}(H_d, w) \leq \left(\frac{(wd+1) - 2^{\lfloor \alpha d \rfloor}}{\lfloor \alpha d \rfloor w} + 1 - \frac{1}{\lfloor \alpha d \rfloor w} \right) + \left(\sum_{i=\Delta}^{\lfloor \alpha d \rfloor - 1} \frac{2^i}{wi} + \lfloor \alpha d \rfloor - \Delta - \sum_{i=\Delta}^{\lfloor \alpha d \rfloor - 1} \frac{1}{iw} \right) + 1. \tag{26}$$

Since $\sum_{i=\Delta}^{\lfloor\alpha d\rfloor-1}\frac{2^i}{wi}\le 2^{\lfloor\alpha d\rfloor}/(w(\lfloor\alpha d\rfloor-2))$ we get

$$\begin{aligned}
\mathtt{tO}(H_d,w) &\le \frac{(wd+1)-2^{\lfloor\alpha d\rfloor}}{\lfloor\alpha d\rfloor w}+\frac{2^{\lfloor\alpha d\rfloor}}{(\lfloor\alpha d\rfloor-2)w}+2+\lfloor\alpha d\rfloor-\Delta-\sum_{i=\Delta}^{\lfloor\alpha d\rfloor}\frac{1}{iw}\\
&\le \frac{2^{\lfloor\alpha d\rfloor}}{w}\left(\frac{1}{\lfloor\alpha d\rfloor-2}-\frac{1}{\lfloor\alpha d\rfloor}\right)+\frac{d}{\lfloor\alpha d\rfloor}+2+\lfloor\alpha d\rfloor-\Delta-\sum_{i=\Delta}^{\lfloor\alpha d\rfloor-1}\frac{1}{iw}\\
&= \frac{2^{\lfloor\alpha d\rfloor}}{w}\left(\frac{2}{\lfloor\alpha d\rfloor(\lfloor\alpha d\rfloor-2)}\right)+\frac{d}{\lfloor\alpha d\rfloor}+\lfloor\alpha d\rfloor-\Delta+2-\sum_{i=\Delta}^{\lfloor\alpha d\rfloor-1}\frac{1}{iw}
\end{aligned}$$

Noticing that $wd+1=2^{\alpha d}\ge 2^{\lfloor\alpha d\rfloor}$

$$\mathtt{tO}(H_d,w)\le\frac{d}{\lfloor\alpha d\rfloor}+\frac{2d}{\lfloor\alpha d\rfloor(\lfloor\alpha d\rfloor-2)}+\lfloor\alpha d\rfloor-\Delta+2. \tag{27}$$

Noticing that the function $f(x)=(2^x-1)/x$ is increasing, and $f(\lfloor\log(w\log w)\rfloor)\le w$, by the definition of Δ we can deduce that $\Delta\ge\lfloor\log(w\log w)\rfloor$; therefore $\lfloor\alpha d\rfloor-\Delta\le\log wd+1-\log(w\log w)$ and

$$\begin{aligned}
\mathtt{tO}(H_d,w) &\le \frac{d}{\lfloor\alpha d\rfloor}\left(1+\frac{2}{(\lfloor\alpha d\rfloor-2)}\right)+\log\left(\frac{wd+1}{w\log w}\right)+2\\
&= \frac{d}{\lfloor\log(wd+1)\rfloor}\left(1+\frac{2}{(\lfloor\alpha d\rfloor-2)}\right)+\log\left(\frac{wd+1}{w\log w}\right)+2\\
&\le \frac{d}{\lfloor\alpha d\rfloor}\left(1+\frac{2}{\lfloor\alpha d\rfloor-2}+\frac{\log(wd+1)}{d}\log\left(\frac{wd+1}{w\log w}\right)\right)+2.
\end{aligned}$$

Putting $c(d,w)=1+\frac{2}{\lfloor\alpha d\rfloor-2}+\frac{\log(wd+1)}{d}\log\left(\frac{wd+1}{w\log w}\right)$ and evaluating the limit as $d\to\infty$ gives the theorem. The upper bound $c(d,w)\le 4$ follows from the definition for $d\ge 8$ and by direct evaluation of upper bound (25) in the remaining cases. □

We conclude this section by mentioning the following result, where the lower bound follows from Lemma 4.2 and the upper bound from a straightforward recursive algorithm, which at each step partitions the network of maximum side k into meshes of maximum side $k/\lfloor\sqrt{4w+1}\rfloor$.

Theorem 4.4 *Let M_{k_1,k_2} and C_{k_1,k_2} be the $k_1\times k_2$ mesh and torus, respectively, on the $n=k_1k_2$ nodes in the set $\{(x_1,x_2)\ :\ 0\le x_i<k_i,\ i=1,2\}$. Let $k=\max\{k_1,k_2\}$, for each w,*

$$\left\lceil\frac{\log(2n-1)}{\log(4w+1)}\right\rceil\le\mathtt{tO}(M_{k_1,k_2},w)\le\left\lceil\frac{\log k}{\log\lfloor\sqrt{4w+1}\rfloor}\right\rceil+1,$$

$$\left\lceil\frac{\log n}{\log(4w+1)}\right\rceil\le\mathtt{tO}(C_{k_1,k_2},w)\le\left\lceil\frac{\log k}{\log\lfloor\sqrt{4w+1}\rfloor}\right\rceil.$$

5 Conclusions and Open Problems

In this paper we have initiated the study of efficient collective communication in switched optical networks. Although we have obtained a number of results, several open problems can be investigated for future lines of research. We list the most important of them here.

• The computation complexity of the quantities $\mathtt{wO}(G,t)$, $\mathtt{wA}(G,t)$, $\mathtt{tO}(G,w)$, $\mathtt{tA}(G,w)$ deserves to be investigated. It is likely that for some of them it is NP–hard.

• Our algorithms require a centralised control. This seems not to be a severe limitation in that the major applications for optical networks require connections that last for long periods once set up; therefore, the initial overhead is acceptable as long as sustained throughput at high data rates is subsequently available [43]. Still distributed algorithms are worth investigating.

• We did not consider fault tolerant issues here. See the recent survey [40] for an account of the vast literature on fault–tolerant communication in traditional networks.

Acknowledgements

We wish to thank the referees for their comments which have helped us to improve the presentation of the results of this paper.

References

[1] A. Aggarwal, A. Bar-Noy, D. Coppersmith, R. Ramaswami, B. Schieber, M. Sudan, "Efficient Routing and Scheduling Algorithms for Optical Networks", in: *Proceedings of the 5th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'94)*, (1994), 412–423.

[2] M. Ajmone Marsan, A. Bianco, E. Leonardi, and F. Neri, "Topologies for Wavelength–Routing for All-Optical Networks", *IEEE/ACM Transactions on Networking*, vol. 1, No. 5, (1993), 534–546.

[3] Y. Aumann and Y. Rabani, "Improved Bounds for All Optical Routing", in: *Proceedings of the 6th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA'95)*, (1995), 567–576.

[4] R. A. Barry and P. A. Humblet, "Bounds on the Number of Wavelengths Needed in WDM Networks", in: *LEOS '92 Summer Topical Mtg. Digest*, (1992), 21–22.

[5] R. A. Barry and P. A. Humblet, "On the Number of Wavelengths and Switches in All–Optical Networks", *IEEE Trans. on Communications*, Part I, 1994, pp. 523–591.

[6] K. Bala, T.E. Stern, D. Simchi–Levi, and K. Bala, "Routing in Linear Lightwave Networks", *IEEE/ACM Trans. on Networking*, vol. 3, (1995), 459–469.

[7] B. Beauquier, J.-C. Bermond, L. Gargano, P. Hell, S. Perennes, and U. Vaccaro, "Graph Problems Arising from Wavelength–Routing in All–Optical Networks", *2nd Workshop on Optics and Computer Science (WOCS)*, Geneve, Switzerland, April 1997,

[8] C. Berge, *Graphs*, North–Holland.

[9] J.-C. Bermond, N. Homobono, and C. Peyrat, "Large Fault–Tolerant Interconnection Networks", *Graphs and Combinatorics*, vol. 5, (1989), 107–123.

[10] D. P. Bertsekas, and J. N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Prentice–Hall, Englewood Cliffs, NJ, 1989.

[11] K. W. Cheng, "Acousto-optic Tunable Filters in Narrowband WDM Networks", *IEEE J. Selected Areas in Comm.*, vol. 8, (1990), 1015–1025.

[12] N.K. Cheung *et al.*, *IEEE JSAC*: Special Issue on Dense WDM Networks, vol. 8 (1990).

[13] I. Chlamtac, A. Ganz, and G. Karmi, "Purely Optical Network for Terabit Communication", *Proc. IEEE Infocom '89*, (1989), 887–896.

[14] E. Desurvire, *Erbium–Doped Fiber Amplifiers — Principles and Applications*, NY Wiley 1994.

[15] J. J. Dongarra and D. W. Walker, "Software Libraries for Linear Algebra Computation on High Performances Computers", *SIAM Review*, vol. 37, (1995), 151–180.

[16] T. Erlebach and K. Jansen. "Scheduling of Virtual Connections in Fast Networks", *Proc. of 4th Workshop on Parallel Systems and Algorithms PASA '96*, (1996), 13–32.

[17] A.M. Farley, "Minimum–Time Line Broadcast Networks", *NETWORKS*, vol. 10 (1980), 59-70.

[18] R. Feldmann, J. Hromkovic, S. Madhavapeddy, B. Monien, P. Mysliwietz, "Optimal Algorithms for Dissemination of Information in Generalised Communication Modes", in: *Proc. of Parallel Architectures and Languages Europe (PARLE '92)*, Springer LNCS 605, (1992) 115–130.

[19] G. Fox, M. Johnsson, G. Lyzenga, S. Otto, J. Salmon, and D. Walker, *Solving Problems on Concurrent Processors, Volume I*, Prentice Hall, Englewood Cliffs, NJ, 1988.

[20] P. Fraignaud, E. Lazard, "Methods and Problems of Communication in Usual Networks", *Discrete Applied Math.*, 53 (1994), 79–134.

[21] P. E. Green, *Fiber-Optic Communication Networks*, Prentice–Hall, 1992.

[22] P. E. Green, "Optical Networking Update", *IEEE J. Selected Areas in Comm.*, vol. 14, (1996), 764–779.

[23] S. M. Hedetniemi, S. T. Hedetniemi, and A. Liestman, "A Survey of Gossiping and Broadcasting in Communication Networks", *NETWORKS*, 18 (1988), 129–134.

[24] M.C. Heydemann, J. C. Meyer, and D. Sotteau, "On Forwarding Indices of Networks", *Discrete Applied Mathematics*, vol. 23, (1989), 103–123.

[25] C.-T. Ho and M.-Y. Kao, "Optimal Broadcast in All-Port Wormhole-Routed Hypercubes", *IEEE Transactions on Parallel and Distributed Systems*, vol. 6, No. 2, (1995), 200–204.

[26] J. Hromkovič, R. Klasing, W. Unger, H. Wagener, "Optimal Algorithms for Broadcast and Gossip in the Edge-Disjoint Path Modes", in: *Proc. of the 4th Scandinavian Workshop on Algorithm Theory (SWAT'94)*, Springer LNCS 824,(1994), pp. 219-230.

[27] J. Hromkovič, R. Klasing, B. Monien, and R. Peine, "Dissemination of Information in Interconnection Networks (Broadcasting and Gossiping)", in: Ding-Zhu Du and D. Frank Hsu (Eds.) *Combinatorial Network Theory*, Kluwer Academic Publishers, 1995, pp. 125-212.

[28] K. Kaklamanis and G. Persiano, "Efficient Wavelength Routing in Directed Fiber Trees", in: Diaz and Serna (Eds.) *Proc. of 4-th Annual European Symposium on Algorithms*, LNCS, Springer Verlag, 1996, pp. 460–470.

[29] R. Klasing, "The Relationship Between Gossiping in Vertex-Disjoint Paths Mode and Bisection Width", *Proc. of the 19th International Symposium on Mathematical Foundations of Computer Science (MFCS'94)*, Springer LNCS 841, pp. 473-483, 1994.

[30] D. W. Krumme, K.N. Venkataraman, and G. Cybenko, "Gossiping in Minimal Time", *SIAM J. on Computing*, 21 (1992), 111–139.

[31] F. T. Leighton and S. Rao, "An Approximate Max-Flow Min-Cut Theorem for Uniform Multicommodity Flow Problems with Applications to Approximation Algorithms", in: *Proceedings of the 29th Annual IEEE Symposium on Foundations of Computer Science (FOCS'88)*, (1988), 422–431.

[32] K. Liu, R. Ramaswami, "Analysis of Optical Bus Networks using Doped–Fiber Amplifiers", Proc. *IEEE/LEOS Topical Meeting on Optical Multi–Access Networks*, July 1990.

[33] L. Lovász, *Combinatorial Problem and Exercises*, 2nd edition, Elsevier, 1993.

[34] W. Mader, "Minimale n-fach Kantenzusammenhangende Graphen", *Math. Ann.*, 191 (1971), 21-28.

[35] A. D. McAulay, *Optical Computer Architectures*, John Wiley, 1991.

[36] M. Mihail, K. Kaklamanis, S. Rao, "Efficient Access to Optical Bandwidth", in: *Proceedings of 36th Annual IEEE Symposium on Foundations of Computer Science (FOCS'95)*, (1995), 548–557.

[37] B. Mukherjee, "WDM–Based Local Lightwave Networks, Part I: Single–Hop Systems", *IEEE Networks*, vol. 6 (1992), 12–27.

[38] B. Mukherjee, "WDM–Based Local Lightwave Networks, Part II: Multihop Systems", *IEEE Networks*, vol. 6 (1992), 20–32.

[39] R.K. Pankaj, *Architectures for Linear Lightwave Networks*, PhD Thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 1992.

[40] A. Pelc, "Fault Tolerant Broadcasting and Gossiping in Communication Networks", *NETWORKS*, vol. 28, 1996, 143–156.

[41] S. Personick, "Review of Fundamentals of Optical Fiber Systems", *IEEE J. Selected Areas in Comm.*, vol. 3, (1983), 373–380.

[42] G. R. Pieris and G. H. Sasaki, "A Linear Lightwave Beneš Network", *IEEE/ACM Trans. on Networking*, vol.1 , (1993).

[43] P. Raghavan and E. Upfal, "Efficient Routing in All–Optical Networks", in: *Proceedings of the 26th Annual ACM Symposium on Theory of Computing (STOC'94))*, (1994), 133-143.

[44] R. Ramaswami, "Multi-Wavelength Lightwave Networks for Computer Communication", *IEEE Communication Magazine*, vol. 31, (1993), 78–88.

[45] R. Ramaswami and K. Sivarajan, "Routing and Wavelength Assignement in All–Optical Networks", *IEEE/ACM Trans. on Networking*, vol. 3, (1995), 489–500.

[46] R. Ramaswami and K. Sivarajan, "Design of Logical Topologies for Wavelength–Routed Optical Networks", *IEEE JSAC/JLT Special Issue on Optical Networks*, June 1996.

[47] P. Solé, "Expanding and Forwarding", *Discrete Applied Mathematics*, vol. 58, (1995), 67–78.

[48] R.J. Vetter and D.H.C. Du, "Distributed Computing with High-Speed Optical Networks", *IEEE Computer*, vol. 26, (1993), 8–18.

[49] O. Wolfson and A. Segall, "The Communication Complexity of Atomic Commitment and Gossiping", *SIAM J. on Computing*, 20 (1991), 423–450.

A Appendix

Proof of Lemma 4.4 Fix s and consider any spanning tree T of G. It is obvious that we can limit ourselves to construct an s–tree cover of T. We will need the following simple and known fact, which can be easily proved by induction: There exist a node in T such that each subtree T_i formed by removing from T this node and all incident edges, satisfies $|T_i| \leq n/2$. In the sequel we denote by r such a node and by $T_1, \ldots, T_{t-1}, T_t = \{r\}$ the subtrees obtained by removing all edges incident on r; such subtrees are indexed in order of non increasing number of nodes, that is,

$$n/2 \geq |T_1| \geq \ldots \geq |T_t| = 1. \tag{28}$$

Moreover, we indicate by $m \geq 0$ the largest index such that

$$|T_1| + |T_2| + \ldots + |T_m| < s \tag{29}$$

If $n \leq s$ then a 1-tree cover of T consists of T itself. Let

$$s \leq n < 3s/2.$$

In this case we will consider the s–tree cover $\mathcal{F} = \{F_1, F_2\}$, where:

F_1 is the induced subtree of T consisting of all nodes in the trees $T_1, \ldots, T_m, T_t$ and

F_2 is the induced subtree of T consisting of all nodes in the trees $T_{m+1}, \ldots, T_t$.

Since $|T_1| \leq n/2 < s$ we have that $m \geq 1$. Moreover, by (29) we have

$$|F_1| \leq s.$$

We show now that $|F_2| \leq s$. Consider first the case $m = 1$. If we supposed that

$$|F_2| = n - |T_1| = |T_2| + \ldots + |T_t| > s$$

we get $|T_1| < n - s < s/2$ which implies that $|T_1| + |T_2| \leq 2|T_1| < s$, contradicting the assumption that $m = 1$ is the largest integer such that (29) holds.

Suppose now that $m \geq 2$. We have $|T_{m+2}| + \ldots + |T_t| \leq n - s$ and $|T_{m+1}| \leq |T_3| \leq n/3$. Therefore, $|F_2| = |T_{m+1}| + \ldots + |T_t| \leq n/3 + n - s < s$. Since properties 1., 2., and 3. of Definition 4.1 hold for $\mathcal{F}$, the lemma holds in this case.

Consider now

$$3s/2 \leq n < 2s.$$

In this case we can consider the s–tree cover $\mathcal{F} = \{F_1, F_2, F_3\}$, where:

F_1 is the induced subtree of T consisting of all nodes in the trees $T_1, \ldots, T_m, T_t$,

$F_2 = T_{m+1}$, and

F_3 is the induced subtree of T consisting of all nodes in the trees $T_{m+2}, \ldots, T_t$.

Indeed, by (29) we have $|F_1| = |T_1| + \ldots + |T_m| + 1 \leq s$, and $|F_3| = |T_{m+2}| + \ldots + |T_t| \leq n - s \leq s$; moreover, $|F_2| = |T_{m+1}| \leq n/(m+1) \leq n/2 < s$. Since properties 1., 2., and 3. of Definition 4.1 hold for $\mathcal{F}$, the lemma holds in this case.

The rest of the proof is by induction. Assume that the lemma is true for any n' such that $n' < (i-1)s$, for some $i \geq 3$. We will prove that the lemma is true also for all values of n such that

$$(i-1)s \leq n < i\ s, \quad i \geq 3.$$

We distinguish two cases on the value of $|T_1|$.
If $|T_1| < s$, we can consider the s-tree cover $\mathcal{F} = \{F_1, F_2\} \cup \mathcal{F}'$, where:
F_1 is the induced subtree of T consisting of all nodes in $T_1, \ldots, T_m, T_t$,
$F_2 = T_{m+1}$, and
$\mathcal{F}'$ is the s–tree cover of the induced subtree of T consisting of all nodes in $T_{m+2}, \ldots, T_t$.
By (29) we have $|F_1| \leq s$; moreover $|F_2| = |T_{m+1}| \leq |T_1| < s$. Finally, $|T_{m+2}| + \ldots + |T_t| \leq n - s < (i-1)s$. Therefore, by inductive hypothesis

$$|\mathcal{F}'| \leq \frac{2(|T_{m+2}| + \ldots + |T_t|)}{s} \leq \frac{2n}{s} - 2$$

in case $|T_{m+2}| + \ldots + |T_t| > s$, otherwise $|\mathcal{F}'| = 1$. Therefore, $|\mathcal{F}| = 2 + |\mathcal{F}'| \leq 2n/s$. Moreover, properties 1. and 2. of Definition 4.1 hold for $\mathcal{F}$, and the lemma holds in this case.

If $|T_1| \geq s$, we can consider the s-tree cover $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$, where:
$\mathcal{F}_1$ is the s-tree cover of the tree T_1, and
$\mathcal{F}_2$ is the s-tree cover of the induced subtree of T consisting of all nodes in $T_2, \ldots, T_t$.
We have $s \leq |T_1| \leq n/2 < (i-1)s$. Moreover, $|T_2| + \ldots |T_t| = n - |T_1| \geq n/2 \geq (i-1)s/2 \geq s$ and $|T_2| + \ldots |T_t| = n - |T_1| \leq n - s < (i-1)s$. Therefore, the inductive hypothesis implies

$$|\mathcal{F}| = |\mathcal{F}_1| + |\mathcal{F}_2| \leq \frac{2|T_1|}{s} + \frac{2(n - |T_1|)}{s} = \frac{2n}{s}$$

Since Properties 1., 2., and 3. of Definition 4.1 hold for $\mathcal{F}$, the lemma holds. $\square$