Efficient Routing and Scheduling Algorithms for Optical Networks*

Alok Aggarwal, Amotz Bar-Noy, Don Coppersmith
Rajiv Ramaswami, Baruch Schieber, Madhu Sudan

IBM – Research Division
T. J. Watson Research Center
Yorktown Heights, NY 10598

Abstract

This paper studies the problems of dedicating routes and scheduling transmissions in optical networks. In optical networks, the vast bandwidth available in an optical fiber is utilized by partitioning it into several channels, each at a different optical wavelength. A connection between two nodes is assigned a specific wavelength, with the constraint that no two connections sharing a link in the network can be assigned the same wavelength. This paper classifies several models related to optical networks and presents optimal or near-optimal algorithms for permutation routing and/or scheduling problems in many of these models. some scheduling problems in one specific model.

*This work was supported by a grant No. MDA 972-92-C-0084 from ARPA.

1. Introduction

Fiber-optic networking technology using wavelength division multiplexing (WDM) offers the potential of building large wide-area networks capable of supporting thousands of nodes and providing capacities of the order of gigabits-per-second to each node in the network [Gre92, Ram93, CNW90]. In WDM optical networks, the vast bandwidth available in optical fiber is utilized by partitioning it into several *channels*, each at a different optical *wavelength*. Each wavelength can carry data modulated at bit rates of several gigabits per second.

In general, such a network consists of *routing nodes* interconnected by point-to-point fiber-optic links (Figure 1). 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, independent of the other wavelengths. However, the same wavelength on two input ports cannot be routed to a common output port. The first class of networks that we consider are *non-reconfigurable*, or *switchless*; i.e., the routing patterns at each of these routing nodes is fixed. Such a routing node is shown in Figure 2. These networks are practically important because the entire network can be constructed out of passive (unpowered) optical components and hence made reliable as well as easy to operate, with all the control being done outside the network.

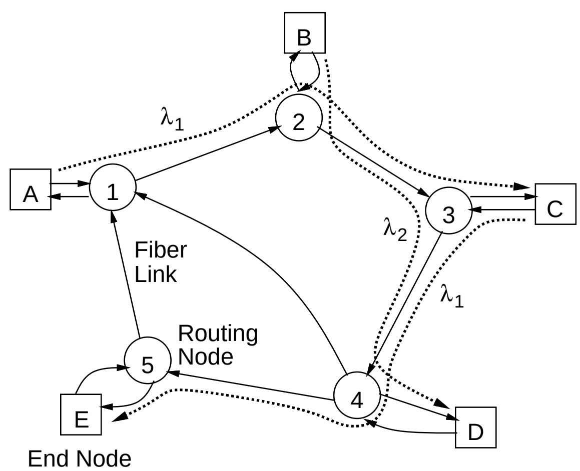

Figure 1: A WDM network consisting of routing nodes interconnected by point-to-point fiber-optic links. Some of the routing nodes have end-nodes attached to them that form the sources and destinations for network traffic.

The second class of networks, which we call *reconfigurable networks*, have *optical switches* at the routing nodes. By reconfiguring the switches, the routing pattern at a routing node can be changed. Optical switches will be required to build large networks because the switchless network requires a large number of wavelengths to support even simple traffic patterns (as will be seen later in this paper). A reconfigurable routing node is shown in Figure 3.

Some of the routing nodes in the network have *end nodes* attached to them via fiber-optic links.

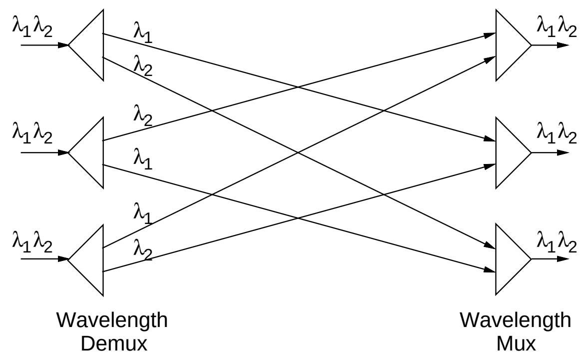

Figure 2: Structure of a non-reconfigurable (switchless) routing node.

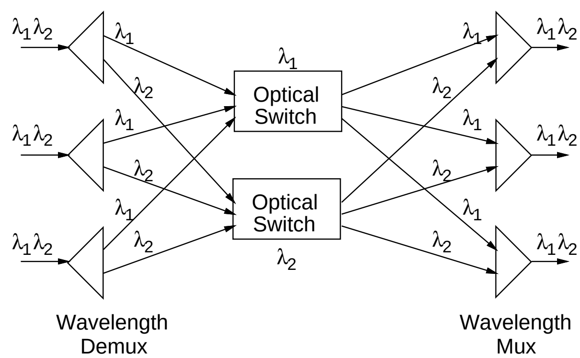

Figure 3: Structure of a reconfigurable routing node. The node can switch each wavelength at its input ports independent of the other wavelengths. The switch can be reconfigured to allow different interconnection patterns.

We refer to end nodes as nodes. Each end node has a tunable optical transmitter and a tunable optical receiver. The transmitter can be tuned to transmit at any one of the available wavelengths and the receiver can be tuned to receive on any one of the available wavelengths.

We are interested in the problem of setting up connections between different node pairs and would like to determine the number of wavelengths and switches required to support different traffic patterns in these networks. Each connection is assigned a wavelength. The constraint imposed by the network is that there can be at most one connection using a given wavelength on any link. In switchless networks, once the routing pattern is set, the only choice remaining is in selecting the wavelength at which each node transmits and the wavelength at which it receives. In networks with switches, additional degrees of freedom are obtained by changing the setting of the switches.

1.1. Preliminaries

A *permutation network* is a network that can successfully route all sets of connections that are a permutation of the network nodes.

A *non-blocking network* (abbreviated NBN) is a network that allows communication routes between pairs of transmitters and receivers to be connected and terminated dynamically. It handles two kinds of requests: *connection requests* and *termination requests*. A connection request specifies a pair of (transmitter,receiver) that has to be connected. It is assumed that both the transmitter and the receiver are idle when the request is initiated. A termination request specifies a pair (transmitter,receiver) that are currently connected and terminates this connection. Following Beneš [Ben62], we distinguish between two types of non-blocking networks.

Rearrangeably NBN: This is the weakest type of an NBN. In such networks whenever a new request arrives, all existing connections can be rerouted.

Wide-sense NBN: In such networks once a route is dedicated for a request, it cannot be rerouted in the future to accommodate later requests.

We consider several variations of the routing problem. First, we distinguish between the *off-line* and *on-line* cases. In the off-line case all the requests are known in advance, whereas in the on-line case, future requests are not known in advance. Note that this distinction is relevant only for a wide-sense NBN. Then, we consider *oblivious* routing schemes in which a connection request (t, r) will always be satisfied using the same wavelength, regardless of the rest of the requests. An oblivious scheme is clearly an on-line scheme. We extend this setting to the case of *partial obliviousness*; i.e., to the case where the number of possible wavelengths that are available to satisfy a connection request is bounded.

The *congestion* of a routing algorithm is the maximum number of paths that go over a single edge in the network. The *dilation* of a routing algorithm is the maximum number of edges in a path used by the routing algorithm.

1.2. Previous Work

The routing problem in these networks has been studied by Barry and Humblet[BH92, BH93], Pankaj[Pan92], and Pieris and Sasaki[PS93a]. Barry and Humblet [BH92] derived an information-theoretic lower bound on the number of wavelengths required to support a given number of traffic states in networks with and without switches. For example, permutation routing in a switchless network requires $\Omega(\sqrt{n})$ wavelengths, where n is the number of nodes in the network. They also showed that oblivious permutation routing could be done using $\lceil n/2 \rceil + 2$ wavelengths.

In a reconfigurable network with w available wavelengths, Barry and Humblet [BH93] showed that the number of 2×2 switches required to support permutation routing is $\Omega(n \log(n/w^2))$. When the transmitters are fixed-tuned and the receivers are tunable, Pieris and Sasaki [PS93a] showed that the number of 2×2 switches required for permutation routing is $\Omega(n \log(n/w))$, and constructed such a network using $O(n \log(n/w))$ switches.

Pankaj [Pan92] obtained bounds on the number of wavelengths required for permutation routing in certain network topologies. His network model assumed a multi-wavelength switch at each

Category	Lower Bound	Previous Upper Bound	Our Upper Bound
Non-Reconfigurable (Switchless) Networks			
Oblivious	$\lceil \frac{n}{2} \rceil + 1$ [BH93]		$\lceil \frac{n}{2} \rceil + 2$ [BH92]
Non-oblivious existential	$\Omega(\sqrt{n})$ [BH92]		$O(\sqrt{n \log n})$
Non-oblivious constructive	$\Omega(\sqrt{n})$ [BH92]		$O\left(\sqrt{n} 2^{(\log n)^{0.8+o(1)}}\right)$
Partial oblivious ($k \geq 3$)	$\Omega(\sqrt{n})$ [BH92]		$O\left(n^{\frac{k+1}{2k-1}} k^{\frac{k}{2k-1}}\right)$
Star network, n wavelengths	$\Omega(n\sqrt{D})$	$n(\sqrt{D}+1)$ [PS93b]	$n(\sqrt{D}+0.5)$
Reconfigurable Networks			
Elementary switches existential	$n \log \frac{n}{w^2}$ [BH92]		$n \log \frac{n \log w}{w^2}$
Elementary switches constructive	$O\left(n \log \frac{n}{w^2}\right)$ [BH92]		$O\left(n \log \frac{n 2^{(\log w)^{0.8+o(1)}}}{w^2}\right)$
$O(n)$ generalized switches	$\Omega(\log n)$ [Pan92]	$O(\log^2 n)$, $O(\log^3 n)$ [Pan92][1]	$O(\log n)$
$O(n)$ generalized switches	$\Omega[c \min\{d, \sqrt{m}\}]$		$O[c \min\{d, \sqrt{m}\}]$

Table 1: Main results on permutation routing and scheduling. The results hold for both rearrangeable and wide-sense non-blocking cases unless specified otherwise. The number of nodes is n, c denotes congestion, d denotes dilation, m is the number of edges in the network, and k denotes the number of wavelengths available for a connection between a pair of nodes. Except for the star network where the schedule length is given and D denotes the time taken to tune between wavelengths, all other results are for the number of wavelengths.

routing node; such a switch is a *generalized* switch in that it can permute each wavelength independent of the other wavelengths. For this model, Pankaj showed that $\Omega(\log n)$ wavelengths are required for permutation routing. He also showed that rearrangeably non-blocking permutation routing can be done with $O(\log^2 n)$ wavelengths and wide-sense non-blocking permutation routing can be done with $O(\log^3 n)$ wavelengths in popular interconnection networks such as the shuffle exchange network, the DeBruijn network, and the hypercube network.

1.3. Contributions of This Work

We present almost tight bounds for most of the problems considered in the earlier papers. Our results are summarized in Table 1.

We show that oblivious permutation routing in a switchless network requires $\lfloor n/2 \rfloor + 1$ wave-

[1] $O(\log^2 n)$ for the rearrangeable case and $O(\log^3 n)$ for the wide sense non-blocking case.

lengths and we demonstrate the existence of a switchless permutation network using $O(\sqrt{n \log n})$ wavelengths. Both these results have been obtained independently by Barry and Humblet [BH93], the latter result however was proved for networks where *wavelength converter* devices are used. (A wavelength converter is a device that performs a transformation of a data stream coming in at a specific wavelength onto an outgoing data stream at a different specific wavelength.) The *non-constructive* aspect of the above result may be viewed as a drawback. We complement this result with a constructive version which is only slightly weaker; In particular, we show how to construct a switchless permutation network using $O\left(\sqrt{n}\, 2^{(\log n)^{0.8+o(1)}}\right)$ wavelengths. We also provide non-trivial upper bounds for partially oblivious networks.

For reconfigurable networks, we show the existence of a wide-sense non-blocking network using $O(n \log \frac{n \log w}{w^2})$ switches and construct a wide-sense non-blocking network using $O\left(n \log \frac{n 2^{\log w^{0.8+o(1)}}}{w^2}\right)$ switches. Clearly all of these results apply also for rearrangeably non-blocking networks.

For the class of networks considered by Pankaj, we show a tight bound of $\Theta(\log n)$ wavelengths for both rearrangeable and wide-sense non-blocking permutation routings. As in his thesis, we only use n generalized switches that are configured as a shuffle exchange network, or a DeBruijn network, or, in fact, as any $O(\log n)$ depth permutation network.

We also derive an upper bound on the number of wavelengths required for any routing scheme in terms of congestion and dilation for the given routing and the given network. We show that these exist a class of networks for which this bound cannot be improved.

Finally, we deal with an important special case: that of a star broadcast network with n nodes and n wavelengths available. The star network is the first network to have been prototyped and undergone field tests [JRS93]. In this case, we consider optical transmitters and receivers that take D units of time (slots) to tune between two wavelengths. We show that any schedule where each node pair is assigned a single slot for communication, must have at least $\Omega(n\sqrt{D})$ slots, thus indicating that an upper bound of $n(\sqrt{D}+1)$ derived in [PS93b] and improved by us to $n(\sqrt{D}+1/2)$ is almost tight.

Section 2 deals with non-reconfigurable (switchless) networks and Section 3 with reconfigurable networks (with switches). Section 4 presents some results on a scheduling problem in a simple broadcast network. There are several open problems remaining to be solved; Section 5 gives a few such problems.

2. Non-Reconfigurable Optical Networks

In this section we consider non-reconfigurable (or switchless) optical networks. The network is modeled as a bipartite multigraph [2] $G(T, R, E)$, where T is the set of *transmitters* and R is the set of *receivers*. An edge from a transmitter t to a receiver r does not represent an actual link in the

[2] A multigraph is a graph with multiple edges allowed between nodes.

network, but represents a wavelength using which t can establish a connection to r. Since t may talk to r using many possible wavelengths, there can be multiple edges between t and r. Thus, each edge $e \in E$ is labelled with a wavelength denoted $\ell(e)$, and two or more edges between a transmitter t and a receiver r must have different labels. If e connects transmitter $t \in T$ to receiver $r \in R$, then whenever t transmits using wavelength $\ell(e)$, receiver r may receive this information only if it tunes to this wavelength. Moreover, since the network has no switches, *all* the receivers connected to t with edges labelled $\ell(e)$ receive t's message if they tune to this wavelength. Consequently, if a receiver r is tuned to a wavelength λ, then only *one* transmitter that is connected to r by an edge labelled λ may use this wavelength. Note that once the graph G and the labelling are determined, the only choice remaining is in the tuning of the transmitters and receivers.

The assumption that G is bipartite is made only to make the presentation clearer. We can achieve the same results for networks where some of the nodes are both transmitters and receivers. In this paper we consider the special case of $|T| = |R| = n$. However, most of the results can be extended to the case where $|T| \neq |R|$.

We consider the problem of constructing a non-blocking non-reconfigurable optical network. Our goal is to minimize the number of different wavelengths used in the network.

2.1. Rearrangeably Non-Blocking Networks

In this subsection we consider the minimum number of wavelengths required in a non-reconfigurable rearrangeably NBN. Barry and Humblet [BH92] proved that any such network requires at least $(1+\epsilon(n))\sqrt{n/e}$ wavelengths, where e is the base of the natural logarithm, and $\epsilon(n)$ goes to zero faster than $(\ln\sqrt{n})/\sqrt{n}$. We show that there exists a rearrangeably NBN that uses $O(\sqrt{n\log n})$ wavelengths. Define $\mathrm{p}(n) = 2^{(\log n)^{0.8+o(1)}}$. We also show how to *construct* a network that uses $\sqrt{n}\mathrm{p}(n)$ wavelengths.

In our upper bounds, the network has the following structure. The transmitter set T is partitioned into b blocks $T_0, \ldots, T_{b-1}$ each of size at most n/b, where b is a parameter to be fixed later. (For clarity of exposition we omit the $\lceil . \rceil$ $\lfloor . \rfloor$ operators.) The receiver set R is partitioned k times, where k is to be fixed later. Each partition $1 \leq i \leq k$ partitions R into b blocks $R_0^i, \ldots, R_{b-1}^i$. (The size of each such block may vary.) Our construction will use $w = b \cdot k$ wavelengths, denoted $\lambda_{i,j}$, for $1 \leq i \leq k$, and $0 \leq j \leq b-1$. The edges of the network are labelled as follows: for $1 \leq i \leq k$, $0 \leq j \leq b-1$, and $0 \leq a \leq b-1$, all the transmitters in T_a are connected to all the receivers in $R^i_{(a+j) \bmod b}$ by edges labelled $\lambda_{i,j}$.

The construction above has the following two properties:

G1 Transmitters in each block are identically connected to all of the receivers.

G2 For any wavelength λ, if transmitters t_1 and t_2 belong to different blocks, then the set of receivers connected to t_1 by edges labelled λ is disjoint from the set of receivers connected to t_2 by edges labelled λ.

It is not difficult to see that to get a rearrangeably NBN, it is necessary and sufficient to construct a network such that for any permutation $\Pi = \pi(1), \ldots, \pi(n)$, there is a way to tune the transmitters and receivers such that the connection requests $(t, \pi(t))$, for $1 \le t \le n$, are satisfied. A given tuning satisfies these connection requests if the following two properties are satisfied for all $1 \le t \le n$. (i) Both t and $\pi(t)$ are tuned to the same wavelength λ, and there exists an edge e connecting t to $\pi(t)$ with $\ell(e) = \lambda$. (ii) For all transmitters $t' \neq t$ such that there exists an edge e' labelled λ connecting t' to $\pi(t)$, t' is not tuned to λ.

Consider a permutation $\Pi = \pi(1), \ldots, \pi(n)$ that is to be routed. Property [G2] of our network implies that we can tune the transmitters of each block independently from the transmitters of other blocks. This is because transmitters from different blocks do not interfere. Property [G1] implies that in order to route Π, for any block of transmitters T_i, we have to use n/b different wavelengths. For this, the n/b destination receivers of the transmitters in T_i have to belong to n/b different blocks. Note that these blocks may belong to different partitions.

Given a network G, define the bipartite graph $H(S, B, F)$, where S, the *input* set, corresponds to the set of receivers, and B, the *output* set, corresponds to the set of blocks of receivers. A node $r \in S$ is connected by an edge in F to $v_j^i \in B$ if and only if the corresponding receiver r belongs to the corresponding block R_j^i.

It is not hard to verify that the graph H has the following two properties and that each graph having these properties defines a network.

H1 The degree of each node in S is at most k.

H2 Each node in S is connected to exactly one node in $v_0^i, \ldots, v_{b-1}^i$, for any fixed i.

Theorem 1: The network $G(T, R, E)$ is non-blocking if the corresponding graph $H(S, B, F)$ has the following *matching* property:

H3 The subgraph induced by any subset of n/b receivers and their neighbors in B contains a matching of size n/b.

Proof: Consider a permutation Π. Recall that Property [G2] of the construction implies that transmitters of each block can be tuned independently from the transmitters of other blocks. Fix a block of transmitters T_i. The destination receivers of the transmitters in T_i, may be any subset of n/b receivers. Thus for any subset of n/b receivers, the receivers have to belong to different blocks. By the definition of H, this translates to the matching property [H3]. $\square$

In the rest of this section we prove the existence of a graph H with properties [H1], [H2], and [H3] for $b = \sqrt{n/\log n}$ and $k = O(\log n)$. Then, we show how to construct such a graph with $b = \sqrt{n}$ and $k = \mathrm{p}(n)$.

The results of [FFP88] imply the existence of a graph H having all three properties in which $b = \sqrt{n}$ and $k = O(\log n)$. The following theorem improves this result by a factor of $\sqrt{\log n}$ using a different proof.

Theorem 2: There exists a graph $H(S, B, F)$ with properties [H1], [H2], and [H3] in which $b = \sqrt{n/\log n}$ and $k = O(\log n)$.

Proof: We construct $H(S, B, F)$ probabilistically as follows. Let $|S| = n$, and let B be partitioned into k blocks $B_1, \ldots, B_k$ of size b each. We let each vertex in S pick k neighbors – one in each B_i independently and at random. We now analyze the probability that this graph has the matching property: i.e., any subset of up to n/b vertices from S is contained is some matching.

By Hall's Theorem [Hal35], such a matching does not exist if and only if there exists a set A of a vertices from S ($a \leq n/b$), such that $|N(A)| < |A|$, where $N(A)$ denotes the set of neighbors of A. For a fixed set $A \subset S$ and for sets $A_i \subseteq B_i$ such that $|\cup_i A_i| < a$, we estimate the probability that $N(A) \subseteq \cup_i A_i$. Let $a_i = |A_i|$. Then this probability is at most $\prod_{i=1}^{k}(\frac{a_i}{b})^a$. Thus, the probability that there exist A and A_i's of size a and a_i respectively such that $N(A)| < |A|$ is at most

$$\binom{n}{a}\prod_{i=1}^{k}\binom{b}{a_i}\left(\frac{a_i}{b}\right)^a \leq \left(\frac{na^{k-2}}{k^{k-1}b^{k-1}}\right)^a .$$

Thus *if* $kb \geq cn/b \geq ca$ for some constant $c > 2$ and $k \geq \Omega(\log n)$, then this probability goes to zero as $n^{-O(a)}$. The probability that there exist a and a_i's such that this happens can now be bounded by $o(1)$.

Thus under the conditions $k = \Theta(\log n)$ and $cn/b \leq kb$ (or $b = \Theta(\sqrt{\frac{n}{\log n}})$), we get that with a positive probability $H(S, B, F)$ has the required three properties. □

We now show how to construct such a graph with $b = \sqrt{n}$ and $k = \mathrm{p}(n)$. First, we define a concentrator.

Definition: An (n, m, ℓ)-concentrator with expansion α is a network with n inputs and m outputs such that every set of $t \leq \ell$ inputs expands to at least ct outputs.

We use the following result from Wigderson and Zuckerman [WZ93].

Theorem 3: For all n, there are explicitly constructible $(n, 2\alpha\sqrt{n}, \sqrt{n})$-concentrators with expansion α, depth one and size $\alpha n \cdot \mathrm{p}(n)$.

Application of Theorem 3

For our case we set α to be one, and get that there exists a bipartite graph $H'(S', B', F')$, where $|S'| = 2n$, $B' = 4\sqrt{n}$, and $|F| = n \cdot \mathrm{p}(n)$ with the desired matching property. However, graph H' does not satisfy Properties [H1] and [H2]. We modify H' so that it satisfies these two properties. First, we consider a subgraph of H' which excludes all input nodes whose degree is more has than two times the average degree in H'. Specifically, the degree of each input node in this subgraph is at most $\mathrm{p}(n)$. Clearly, this subgraph still has the desired matching property. Because the size of the graph is $n \cdot \mathrm{p}(n)$, there are at least n input nodes in this subgraph. Next, we duplicate each output node $\mathrm{p}(n)$ times and split the neighbors of each output node among the copies as follows.

We number the edges outgoing from each input node with the numbers 1 to $\mathrm{p}(n)$. Now, the first copy will have as edges the subset of the edges of the original node numbered one, the second will have the subset numbered two, and so forth. It is easy to see that the resulting graph has all the three properties.

2.2. Wide-Sense Non-Blocking Networks

First, define a-limited wide-sense one-sided NBNs.

Definition: A connection request is *one-sided* if it specifies only an input. It is satisfied by connecting the input to *any* of its neighboring outputs. A network H is wide-sense one-sided NBN if it can satisfy any sequence of *one-sided* connection and termination requests without rerouting.

Definition: A network H is a-limited wide-sense one-sided NBN if it can satisfy any sequence of requests in which at most a transmitters are connected simultaneously.

We show that to get a wide-sense NBN G, it is sufficient to make the corresponding graph H n/b-limited wide-sense one-sided NBN.

Theorem 4: The network $G(T, R, E)$ is wide-sense NBN if the corresponding graph $H(S, B, F)$ is n/b-limited wide-sense one-sided NBN.

Proof: Recall that Property [G2] of our construction implies that transmitters of each block can be tuned independently from the transmitters of other blocks. Fix a block of transmitters T_i. The destination receivers of the transmitters in T_i at any given time, may be any subset of at most n/b receivers. Thus, in order to satisfy any sequence of requests for this block in G, at any given time, all the receivers connected to transmitters in T_i have to belong to different blocks. By the definition of H, this translates to the property that H is n/b-limited wide-sense one-sided NBN. □

The following theorem is from [FFP88].

Theorem 5: A network $H(S, B, F)$ is a-limited wide-sense one-sided NBN if it has the following property:

H4 Every set X of inputs of size at most $2a$ has at least $2|X|$ neighbors.

We prove the existence of a graph H with properties [H1], [H2], [H3], and [H4], for $b = \sqrt{n/\log n}$ and $k = O(\log n)$. Then, we show how to construct such a graph with $b = \sqrt{n}$ and $k = \mathrm{p}(n)$. Details omitted.

There is one problem with our construction. Any algorithm which decides how to tune the transmitters and receivers seems not to be polynomial. Borrowing terminology from [FFP88] we have to maintain the maximum *critical* set of inputs. For this, it seems that after each request, we have to check all subsets of idle inputs. There are two ways to alleviate this problem, one that works for the off-line case and the other for the on-line case.

If the sequence of requests is given in advance then the tuning decisions can be done in polynomial time. Roughly speaking, given a connection request, we decide which of the edges in H to use as follows. For each possible edge we check if after using it, the rest of the requests can be satisfied. If so, we use the edge. Because of our construction, we are guaranteed to have at least one such edge.

Suppose now that future requests are not known in advance. In this case we show how to get a polynomial decision algorithm by strengthening the properties H has to satisfy. We note that previous research on non-blocking networks did not address the problem of designing a network that can be operated by a polynomial time algorithm.

Theorem 6: A network $H(S, B, F)$ is a-limited wide-sense one-sided NBN and has a polynomial time decision algorithm if it has the following property:

H5 For every subset X of inputs of size at most $2a$, even after we arbitrarily erase half of the edges adjacent to each input in X, subset X has at least $2|X|$ neighbors.

Proof: The algorithm for making the assignment decision maintains the following invariant:

Let A be the set of inputs for which connections requests are currently active; and let the set $M(A)$ denote the set of outputs to which these inputs are connected. Then, for any subset S of the remaining inputs of cardinality at most a, the size of the neighborhood of S outside of $M(A)$ is at least $|S|$. (*In other words, Hall's condition is satisfied by all potential sets of inputs.*)

It is clear that the invariant above is maintained upon termination of requests. We now show how to satisfy a new request while maintaining the above invariant.

Let v be a new input for which a connection is requested. We tentatively match v to some neighbour outside $M(A)$ - say $t(v)$. We maintain a set C of "critical" inputs - inputs which pose bottlenecks to the invariant. Initially C consists of only v. For every vertex in C we maintain a tentative match $t(C)$. While there exist inputs w outside $A \cup C$ such that more than half the neighbours of w are in $M(A) \cup t(C)$, we add w to the set C; and find an output w' such that $C \cup \{w\}$ can be matched into $t(C) \cup \{w'\}$. *Note that such an output w' must exist because of the invariant and the property of matchings.* We repeat this step till no such vertex w is found. By property **H5** we must terminate before the size of C becomes larger than the size of A. Details omitted. □

It can be shown that such a network that uses $O(\sqrt{n \log n})$ wavelengths exists. We remark that a similar construction is used in [BRSU93] for a different purpose.

2.3. Oblivious and Partial Oblivious Routing

In this subsection we assume that whenever a transmitter has to communicate with a receiver it would use one out of a fixed number k of wavelengths. The case $k = 1$ is called the **oblivious** routing problem since there is no freedom in choosing wavelengths. Notice that this implies that G

is a graph rather than multigraph. The case $k \geq 1$ is called the **partial oblivious** routing problem. In this case, G is a multigraph with bounded multiplicity.

An oblivious routing network can be described by an $n \times n$ matrix M. The entry $M(i,j)$ in the matrix is an integer in the range $1, \ldots, w$ where w is the total number of wavelengths in the solution. The entry $M(i,j)$ indicates that i transmits to j using wavelength $M(i,j)$ in any permutation Π for which $\pi(i) = j$.

Let the matrix M be a solution to the oblivious routing.

Lemma 7: If $\lambda = M(i,j) = M(i',j')$ for $i \neq i'$ and $j \neq j'$, then $M(i,j') \neq \lambda$ and $M(i',j) \neq \lambda$.

Proof: If either $M(i,j')$ or $M(i',j)$ is λ then any permutation Π such that $\pi(i) = j$ and $\pi(i') = j'$ can not be satisfied. $\square$

Define $L(\lambda)$ to be the number of entries in the matrix that are equal to λ.

Lemma 8: $\mathrm{L}(\lambda) \leq 2n - 2$.

Proof: If a row has two or more λ-entries, then these entries belong to columns that have a single λ-entry. Therefore, the number of λ-entries coming from rows which have 2 or more λ-entries is bounded by n. Denote this number by α. If $\alpha = n$ then there is no row with a single λ-entry and we get that $L(\lambda) \leq n$. Otherwise $\alpha \leq n-1$. The number of λ-entries coming from rows which have exactly a single λ-entry is also bounded by n. Denote this number by β. If $\beta = n$ then obviously $L(\lambda) \leq n$. Otherwise $\beta \leq n - 1$. Together we get that $L(\lambda) \leq 2n - 2$. $\square$

The following lower bound follows directly from the above lemma.

Theorem 9: A solution matrix M for the oblivious routing problem must contains at least $\left\lceil \frac{n^2}{2n-2} \right\rceil \geq \lceil \frac{n+1}{2} \rceil = \lfloor n/2 \rfloor + 1$ wavelengths.

Now, we construct a solution for an even n using $(n/2) + 2$ wavelengths. The idea of the construction is well demonstrated by the routing matrix presented in Figure 4. The formal details and the case of an odd n will appear in the full version. In the example $n = 12$, and the wavelengths are denoted by a,b,c,d,e,f,1,2. Note that the northwest and the southeast quadrants are identical and so are the northeast and the southwest quadrants. Moreover the northeast quadrant is a reflection of the northwest quadrant along the northwest-southeast diagonal.

Partial Oblivious Routing

We now sketch the case of the partial oblivious routing. Let k be a bound on the number of wavelengths permitted to be used between any pair of transmitter receiver. In case $k = O(\log n)$ then $w = \Omega(\sqrt{n})$ wavelengths are required, $O(\sqrt{n \log n})$ are sufficient (existentially), and $\sqrt{n}\mathrm{p}(n)$ are sufficient (constructively). (See Section 2.1.)

The existential upper bound for $k = o(\log n)$ can be achieved as follows. Assume as in the case of the $k = O(\log n)$ that we are looking at partitions of the receivers. If the degree is k this implies

a	a	a	a	a	2	b	c	d	e	f	1
f	f	f	f	2	a	b	c	d	e	1	f
e	e	e	2	f	a	b	c	d	1	e	e
d	d	2	e	f	a	b	c	1	d	d	d
c	2	d	e	f	a	b	1	c	c	c	c
2	c	d	e	f	a	1	b	b	b	b	b
–	–	–	–	–	–	–	–	–	–	–	–
b	c	d	e	f	1	a	a	a	a	a	2
b	c	d	e	1	f	f	f	f	f	2	a
b	c	d	1	e	e	e	e	e	2	f	a
b	c	1	d	d	d	d	d	2	e	f	a
b	1	c	c	c	c	c	2	d	e	f	a
1	b	b	b	b	b	2	c	d	e	f	a

Figure 4: The solution for the oblivious routing for $n = 12$ with 8 wavelengths

that we are looking at k different partitions. Assume that each such partition is to b blocks each of size n/b. In this case the number of wavelength is $w = \max\{n/b, kb\}$. To get the bound we have to find the maximum w such that the failure probability is less than 1.

Our results are based on the claim that the failure probability is dominated by the expression:

$$\binom{n}{n/b}\binom{kb}{\frac{n}{b}-1}\left(\frac{\frac{n}{b}-1}{kb}\right)^{\frac{kn}{b}}$$

Denote $\alpha = n/b$ we get that the expression is

$$\binom{n}{\alpha}\binom{kn/\alpha}{\alpha-1}\left(\frac{(\alpha-1)\alpha}{kn}\right)^{k\alpha} \leq \frac{\alpha^{(2k-1)r+1}}{k^{(k-1)\alpha+1}n^{(k-2)\alpha+1}} .$$

This expression is less than one if

$$\alpha^{(2k-1)\alpha+1} \leq n^{(k-2)\alpha+1}k^{(k-1)\alpha+1} .$$

Recall that w, the number of wavelength is given by $\max\{\alpha, kn/\alpha\}$. For such values of α it is always the case that $kn/\alpha \geq \alpha$. For example, for $k = 2$ it gives $w = O(n \log\log n/\log n)$. Things look better for $k > 2$. Then, the number of wavelengths is

$$O\left(n^{\frac{k+1}{2k-1}}k^{\frac{k}{2k-1}}\right) .$$

For $k = 3$ it is $O(n^{4/5})$, for $k = 4$ it is $O(n^{5/7})$, and so on. As k increases the exponent of n tends from above to $1/2$.

3. Reconfigurable Optical Networks

In this section we consider reconfigurable optical networks, i.e., networks with optical switches. The network is modeled by a layered multigraph $G(T, R, M)$, where T is the set of *transmitters*, R is the set of *receivers*, and M is an undirected graph (where each edge is considered to be bidirectional) that connects T and R. We assume that M has nodes of degree four, corresponding to two by two switches. (This assumption is made for simplicity, as well as due to current technological limitations that only allow for construction of switches with constant degree.) An edge of G may be used several times each with a different wavelength. However, a routing with congestion c would require at least c wavelengths. Again, we assume that a connection is carried on the same wavelength on all links of the path; i.e., there is no wavelength conversion. The case with wavelength converters will be considered in the final version of this paper.

In this paper we consider the special case of $|T| = |R| = n$. However, most of the results can be extended to the case where $|T| \neq |R|$. We consider the problem of constructing a reconfigurable optical NBN. Our goal is to study the tradeoffs between the number of switches and the number of different wavelengths used in the network. As in Section 2 we differentiate between rearrangeably NBNs and wide-sense NBNs, and consider several variations of this problem. These variants arise because of different capabilities that can be attributed to the transmitters, or receivers, or the switches. Finally, we distinguish between the *off-line* and *on-line* cases.

We consider two kinds of optical switches: *generalized* switches and *elementary* switches. Generalized switches, considered by Pankaj [Pan92] are fairly powerful in that they can change their state differently for different wavelengths. Elementary switches are considered in [BH92, PS93a]; these switches may not set differently for different wavelengths.

3.1. Non-Blocking Networks with Generalized Switches

In his thesis, Pankaj [Pan92] considered networks of generalized switches of constant degree in which each receiver-cum-transmitter (i.e., each end-node) can be tuned to any wavelength. Pankaj showed that in order to permute n messages any network that uses n switches must use $\Omega(\log n)$ wavelengths. He also described permutation routing algorithms for popular networks such as the shuffle-exchange network, the Debruijn network, and the hypercube. These algorithms use $O(\log^2 n)$ wavelengths to route messages in rearrangeable NBNs and $O(\log^3 n)$ wavelengths to route messages in wide sense NBNs. Theorem 11 creates an optical network for routing in rearrangeable NBNs and gives a routing algorithm which routes messages using $\Theta(\log n)$ wavelengths, using any rearrangeable NBN which uses $n \log n$ switches to route n messages. Theorem 12 modifies this algorithm to obtain a similar result for the wide-sense non-blocking network. Indeed, this algorithm is fairly general in that it will work on any $O(\log n)$ depth permutation network.

Our constructions are based on the following proposition given in [Lei92].

Proposition 10: Given any permutation ρ from kl elements to kl elements $\{x_{ij}\}_{i=1,j=1}^{k,l}$, ρ can be expressed as the product of three permutations ρ_1, ρ_2 and ρ_3, where ρ_1 and ρ_3 preserves the row index of the elements and ρ_2 preserves the column index. (That is, if $\rho_1(x_{ij}) = x_{i'j'}$ then $i = i'$ and similarly for ρ_2 and ρ_3.)

Theorem 11: We can construct an optical reconfigurable rearrangeably NBN with n generalized switches and $O(\log n)$ wavelengths.

Proof: We construct a network G with n inputs and n outputs labelled with a pair (i, j) for $i \in \{1, \ldots, \frac{n}{\log n}\}$ and $j \in \{1, \ldots, \log n\}$. Our network uses a traditional rearrangeably non-blocking network H for $\frac{n}{\log n}$ inputs and $\frac{n}{\log n}$ outputs, as a black box. It is well-known that such networks using $O(n)$ switches exist. (See, e.g. [Lei92].) The switches of G are just the switches of H replaced by generalized ones. In addition to the edges of H, for a fixed i, network G carries edges from every input (i, j) in G to the ith input in H, and similarly from the ith output of H to the outputs (i, j) of G. Each edge in G can carry all wavelengths $\lambda_1, \ldots, \lambda_{\log n}$.

To route a permutation ρ in this network, we decompose ρ into $\rho_1 \cdot \rho_2 \cdot \rho_3$ using Proposition 10 above (with $k = \frac{n}{\log n}$ and $l = \log n$). The wavelength allotted to the message at the (i, j)th input is λ iff $\rho_1(x_{ij}) = x_{i\lambda}$. Similarly, the wavelength allotted to the message at the (i, j)th output is λ iff $\rho_3(x_{i\lambda}) = x_{ij}$. Observe that exactly one message in each wavelength arrives at any input of H, and similarly for the outputs. Notice further that the task of routing the messages within H, which is basically performing the permutation ρ_2, decomposes into $\log n$ different permutation routing tasks – one for each wavelength. We are done now since H can route any such permutation. □

Theorem 12: We can construct an optical reconfigurable wide-sense NBN with n generalized switches and $O(\log n)$ wavelengths.

Proof: The idea here is similar to that of Theorem 11. The network here is the same as the one above, except for two differences: First, we now use a wide-sense NBN H as our black box (rather than the rearrangeable network). Again, it is well-known that such networks using $O(n)$ switches exist [ALM90]. Second, each edge can now carry $2 \log n - 1$ wavelengths (rather than exactly $\log n$). To route a request from input (i, j) to output (i', j') we look for a wavelength λ such that both the ith input of H and the i'th output of H are not currently using the wavelength λ. Such a wavelength must exist since the ith input of H is connected to at most $\log n - 1$ inputs of G (other than (i, j)), and hence must have at least $\log n$ unused wavelengths. Similarly for the i'th output of H. Thus they must have one common unused wavelength. The message can now be routed on the wavelength λ since H is a wide-sense NBN. □

3.2. Tunable Non-Blocking Networks with Elementary Switches

A different class of questions is posed by Barry and Humblet in [BH92], who considered elementary switches. This scenario is also important since currently the estimated cost of making a generalized switch is much more than that of the elementary switch. For this case, Barry and Humblet showed a lower bound trade-off for the number of switches and the number of wavelengths used. More precisely, they showed that in a elementary switch network with w wavelengths, the number of switches must be $\Omega(n \log \frac{n}{w^2})$ for both rearrangeable and wide-sense NBNs. Pieris and Sasaki [PS93a] construct such networks that use $\Omega(n \log \frac{n}{w})$ switches. Here, we show tighter upper bounds on the number of switches required in such networks by combining the arguments from Sections 2 and the previous subsection. We prove the following two theorems.

Theorem 13: Given w wavelengths, there exists an optical rearrangeable NBN of size n that uses $O(n \log \frac{n \log w}{w^2})$ switches. Furthermore, we can *construct* an optical rearrangeable NBN that uses $O(n \log \frac{n \mathrm{p}(n)}{w^2})$ switches.

Proof: Again, the idea is to use Proposition 10 about decompositions of permutations. The network G is constructed in three layers. The first and third layer do the "row" permutations and the second layer does the "column" permutation. For each of the first and third layers we use a rearrangeable non-blocking network of the traditional type with traditional switches replaced by elementary ones. The middle layer consists of the switchless network(s) described in Section 2. For w wavelengths, the middle layer can hence route up to $m = \frac{w^2}{\log w}$ messages existentially and $m = \frac{w^2}{\mathrm{p}(w)}$ messages constructively. This gives a bound on the number of rows. Hence, the number of columns must be at least n/m. To route along the columns we use any traditional rearrangeable network and this will require $O(n/m \log(n/m))$ switches per row. Thus the total number of switches is $O(n \log(n/m))$.

□

Theorem 14: Given w wavelengths, there exists an optical wide-sense NBN of size n that uses $O(n \log \frac{n \log w}{w^2})$ switches. Furthermore, we can *construct* an optical wide-sense NBN that uses $O(n \log \frac{n \mathrm{p}(n)}{w^2})$ switches.

The proof of this theorem follows from the proofs of Theorem 13, 12 and the results in Section 2. Details are omitted.

Finally, we consider the special case where only the transmitters can be tuned whereas the wavelengths of the receivers are fixed (or vice-versa). For this case we have the following theorem.

Theorem 15: Given w wavelengths, we can construct an optical wide-sense NBN of size n in which only the transmitters (receivers) are tunable that uses $\Theta(n \log \frac{n}{w})$ switches.

3.3. Bounding Number of Wavelengths via Congestion and Dilation

In this section, we give bounds on the number of wavelengths required to route a set of messages on optical networks with n nodes, each having a generalized switch. Our bounds here relate the number of wavelengths to two classical parameters associated with routing i.e., the congestion (denoted c) and the dilation (denoted d).

Given a graph G with n nodes and m edges, suppose some messages can be routed in it such that the maximum congestion (i.e., number of messages using any edge) is c and the maximum dilation (i.e., the maximum path length from any source to any sink) is d. Then, observe that by a greedy coloring of the interference graph of paths cd wavelengths are sufficient to achieve this routing. Our first observation shows that for if the dilation is sufficiently large then it is possible to beat this bound. In particular, if $d > \sqrt{m}$ then $2c\sqrt{m}$ wavelengths suffice to route the messages. Next we show that this bound is optimal up to a constant factor, in that there exist graphs and message requests for which $\Omega(c \min\{d, \sqrt{m}\})$ wavelengths are required to route the given messages.

Lemma 16: For any graph G, and any set of routing requests with congestion of c, $2c\sqrt{m}$ wavelengths are sufficient to achieve the given routing.

Proof: The number of paths of length greater than $\sqrt{m}$ is at most $c\sqrt{m}$. Give each such path its own dedicated wavelength. Each of the remaining paths (of length less than $\sqrt{m}$), conflicts with fewer than $c\sqrt{m}$ paths, so again can be given a wavelength without conflicting by a greedy coloring of the interference graph of the paths of length less than $\sqrt{m}$. □

Theorem 17: There exist graphs and message routing requests which can be routed with congestion c and dilation d, but require $w = \Omega(c \min\{d, \sqrt{m}\})$ wavelengths under any routing.

Proof: Define the graph G with n transmitters and n receivers as depicted in Figure 5. G has n columns with $O(n)$ nodes each and the number of edges is $O(n^2)$.

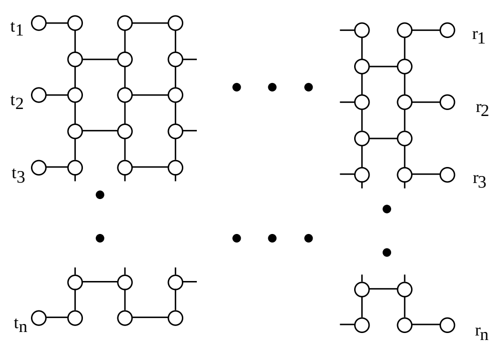

Figure 5: The existential lower bound

There are n requests. The transmitter and the receiver of the ith connection are nodes t_i and r_{n-i} respectively. It is not hard to verify that for any choice of n paths satisfying these n connection, any two paths intersect in at least one edge. Consequently each path needs a unique wavelength, which implies that the number of wavelengths is n.

It is possible to construct n paths such that the congestion is constant and the dilation is $O(n)$. Details are omitted. Since in this graph $n = O(\sqrt{m})$, it follows that $w = \Omega(\min\{d, \sqrt{m}\})$.

By letting each transmitter sending more than 1 message to the same receiver, we can force any value for the congestion c. The theorem follows since the number of wavelengths in this case is $O(cn)$. □

4. The Passive Optical Star Network

Model We consider one of the simplest networks described in the optical network setting - namely the *passive star network* (see Figure 6). The model consists of n transmitters and n receivers transmitting messages over a passive medium. The transmitters can transmit on any wavelength from the set Λ of wavelengths available for tuning. The message is then broadcast over the medium and can be picked up by any of the receivers which is tuned to this wavelength. The only constraints are that only one transmitter may transmit on any one wavelength at any time slot. The important parameter here is the tuning time of D units that it takes for a transmitter or a receiver to change its transmitting or receiving frequency. In the bipartite graph model, for every transmitter t_i and every receiver r_j and every wavelength λ, there exists an edge between t_i and r_j with label λ.

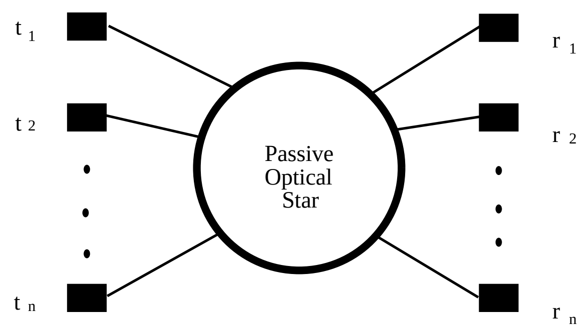

Figure 6: The passive star network

All-to-All transmission problem In this section we consider the complexity of performing all-to-all transmission on the passive optical star network; i.e., each transmitter has to transmit one message to each receiver and it takes one unit of time for any such message to be sent. The objective is to find a schedule of transmission/tuning for these messages which minimizes the total broadcast time.

This problem was considered by [PS93b] who showed a schedule that achieves the all-to-all transmission in $n(\sqrt{D}+1)$ steps. This paper also considered the all-to-all transmission problem in which at any time either all transmitters and receivers are tuning or all are transmitting/receiving. With this constraint [PS93b] showed another schedule which takes $2n\sqrt{D}$ steps, and proved that this is an optimal schedule. This constraint does pose an unnatural restriction on the schedule. [PS93b] also showed a lower bound of $\Omega(n\sqrt{D})$ for the case where only one end is tunable. However, they left open the question of what is the optimal complexity for unrestricted schedules.

Results We show a lower bound for the general scheduling problem which is tight to within constant factors. The bound shows that the schedule described by [PS93b] is nearly optimal. We can also slightly improve the upper bound to be $n(\sqrt{D}+0.5)$. Details for the upper bound are omitted.

The Lower Bound

Definition: A **schedule** S is an assignment of a wavelength and a time slot to each message. For a schedule to be valid, two messages may not be assigned the same (wavelength,time) pair. Furthermore, if a transmitter (receiver) transmits (receives) successive messages on different frequencies, then these messages must be scheduled at least D units of time apart.

Definition: The **waiting time** of a **transmitter** (**receiver**) on a message m is defined to be the time passed since the previous message was sent (received) by the transmitter (receiver). The **waiting time** of a **message** is the sum of the waiting time of the transmitter and the receiver on that message.

It follows that the time taken by the all-to-all broadcast is at least $(\sum_{\text{messages}}$ waiting time of message $)/2n$. Since by an averaging argument there must exist some party (transmitter/receiver) messages from whom take the average time.

To show the lower bound we define a graph G_λ, for every wavelength λ. We associate a vertex in the graph G_λ for each time that a transmitter or a receiver tunes into the wavelength λ. For instance, if the transmitter t tunes k times into the wavelength λ, then there are k vertices $t_1, t_2, \ldots, t_k$ corresponding to these k tunings. The edges of this graph are as follows: $t_i \leftrightarrow r_j$ if and only if the transmitter t transmits a message during its ith tuning into the wavelength λ to the receiver r during its jth tuning into λ.

Notice that a valid schedule S assigns a distinct positive integer as a label to every edge in G_λ. Based on this labels assignments we define the notion of the **stretch** of the graph.

Definition: The stretch $\text{stretch}(u, S)$ of a vertex u in G_λ is the difference between the largest label and the smallest label on the edges incident to u. The stretch $\text{stretch}(G_\lambda, S)$ of the graph G_λ is the sum $\sum_u \text{stretch}(u, S)$.

The graph G_λ naturally captures the two possible causes for a message to wait. Each vertex gives a fixed delay of D steps corresponding to the tuning associated with it. In addition by definition each vertex corresponding to a transmitter (receiver) spends at least $\text{stretch}(u, S)$ units of time to finish transmitting (receiving) its messages before tuning out. Thus for every vertex $u \in G_\lambda$ we have a wait of $D + \text{stretch}(u)$ associated with it.

Lemma 18: The total waiting time for all the messages using the wavelength λ is $\text{stretch}(G_\lambda, S) + v_\lambda \cdot D$ (where v_λ denotes the number of vertices in G_λ).

In the rest of this section we lower bound the stretch of any graph under any schedule.

Lemma 19: For every graph $G = (V, E)$ and every schedule S for the edges of G,

$$\text{stretch}(G, S) \geq \frac{1}{12} \sum_{(u,w) \in E} \min\{d_u, d_w\}.$$

(where d_u denotes the degree of the vertex u in the graph G.)

Proof: Consider the vertex $u_{\max}$ which has the largest stretch.

Claim: $\operatorname{stretch}(u_{\max}, S) \geq \frac{1}{6} \sum_{(u_{\max}, w) \in E} d_w$.

Proof: Consider the set of all edges adjacent to the neighbors of $u_{\max}$. There are at least $\frac{1}{2} \sum_{(u_{\max}, w) \in E} d_u$ many such edges. Thus the difference in time between the earliest scheduled edge, adjacent to u, and the latest scheduled edge, adjacent to u', is at least $\frac{1}{2} \sum_{(u_{\max}, w) \in E} d_w$. This implies that

$$\operatorname{stretch}(u, S) + \operatorname{stretch}(u', S) + \operatorname{stretch}(u_{\max}, S) \geq \frac{1}{2} \sum_{(u_{\max}, w) \in E} d_w .$$

But since $\operatorname{stretch}(u_{\max}, S)$ is no smaller than the other two quantities, it must be at least $\frac{1}{6} \sum_{(u_{\max}, w) \in E} d_w$. □

Let G' be the induced graph on $G - \{u_{\max}\}$, let E' be its edges, and let d' denote degrees of vertices in G'. By induction on the number of vertices we have:

$$\begin{aligned} \operatorname{stretch}(G', S) &\geq \frac{1}{12} \sum_{(u,w) \in E'} \min\{d'_u, d'_w\} \\ &\geq \frac{1}{12} \left(\sum_{(u,w) \in E'} \min\{d_u, d_w\} - \sum_{(u,w) \in E' \colon (u, u_{\max}) \in E \text{ or } (w, u_{\max}) \in E} 1 \right) \\ &\geq \frac{1}{12} \left(\sum_{(u,w) \in E'} \min\{d_u, d_w\} - \sum_{u \leftrightarrow u_{\max}} d_u \right) \end{aligned}$$

By applying the claim about the stretch of $u_{\max}$ we now get:

$$\operatorname{stretch}(G, S) \geq \frac{1}{12} \left(\sum_{(u,w) \in E'} \min\{d_u, d_w\} - \sum_{u \leftrightarrow u_{\max}} d_u \right) + \frac{1}{6} \sum_{u \leftrightarrow u_{\max}} d_u \geq \frac{1}{12} \sum_{(u,w) \in E} \min\{d_u, d_w\} .$$

□

Lemma 20: For any (multi)graph G with m edges and n vertices, the following inequality holds: $\sum_{\{u,v\} \in E} \min\{d_u, d_v\} \geq m^2/n$.

Proof of Lemma 20: Consider the multigraph G which minimizes the summation above, and number the vertices $1, \ldots, n$ in non-decreasing order of degrees. Consider a pair of edges (j_1, j_2) and (k_1, k_2) with $j_1 \leq j_2 \leq k_1 \leq k_2$. We can switch these edges to go from j_1 to k_1 and j_2 to k_2 without increasing the summation (see Figure 7). Consequently, there exists an index i such that if (j, k) is an edge in this graph, with $j < k$, then $j \leq i$ and $k \geq i$. Now in this graph the summation above is the same as $\sum_{k=1}^{i} d_k^2$ where $\sum_{k=1}^{i} d_k \geq m$ (since we count each edge at least once this way). Thus $\sum_{k=1}^{i} d_k^2 \geq \frac{m^2}{i} \geq \frac{m^2}{n}$. □

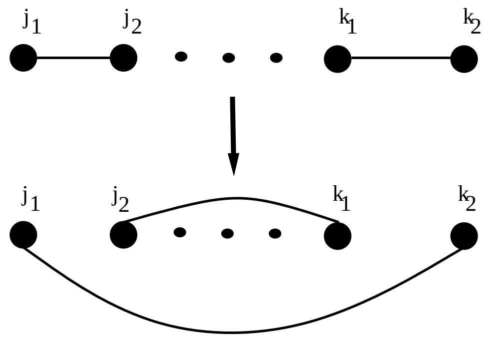

Figure 7: Edge switching

Theorem 21: For every scheduling of all-to-all broadcasts on the passive optical star network, the total broadcast time is at least $\Omega(n\sqrt{D})$.

Proof: Consider the messages scheduled to use the wavelength λ. By Lemmas 18, 19 and 20 we have that the average waiting time per message is at least $\frac{1}{e_\lambda}(v_\lambda \cdot D + \frac{{e_\lambda}^2}{12 v_\lambda})$ which is at least $\sqrt{\frac{D}{3}}$. Thus by an averaging argument there must exist a transmitter or a receiver the waiting time of which is at least $\frac{n^2\sqrt{\frac{D}{3}}}{2n} = \frac{n\sqrt{D}}{2\sqrt{3}}$. This implies that the broadcast time is at least $\Omega(n\sqrt{D})$. $\square$

5. Open Problems

There are several unresolved problems related to the models in this paper; some of them are listed below:

1. We do not have tight bounds for switchless non-oblivious networks. The same is true for the switchless, partially oblivious networks also that allow at most k wavelengths on any edge.
2. Our algorithms for the wide-sense, non-oblivious networks take exponential time in many cases; it would be useful to obtain polynomial time algorithms for these cases.
3. We provided an algorithm to convert any network with m edges that routes messages with congestion c, dilation d, into a network that uses $O(c(\min(d, \sqrt{m})))$ wavelengths, and we also gave a network and a message pattern for which this bound is optimal up to a constant factor. However, a much more interesting and practically useful question is getting a good bound on the number of wavelengths required for a given network and a given message pattern. Here, we have no results and getting even an approximate bound on the number of wavelengths would be very interesting.
4. One research topic not studied in this paper is that of strict-sense non-blocking networks. A strict-sense non-blocking network is one that allows a new connection to be always routed through irrespective of how the previous connections were routed.
5. Another research topic not studied in this paper is that of wavelength converters. It appears that technology in the future will be able to sustain switches that will be able to statically and/or

dynamically convert the wavelength on which an incoming signal is traveling. This implies that the path from a source to a sink need to be of one wavelength (or one color). Clearly all the upper bounds in the paper hold for networks with fixed wavelength converters. We can also show that all the lower bounds hold too. However, with dynamic wavelength converters, we are adding to the number of possible states in the network and the bounds may no longer hold. For instance, if we are allowed to use dynamic wavelength converters arbitrarily, then the number of wavelengths sufficient for permutation routing in a reconfigurable network with generalized switches is c, where c is the congestion of the routing algorithm (the network becomes identical to a classical circuit-switched network). It would be useful to study this topic in detail.

6. The area of fault tolerance in optical networks is an open area for research. Also, there is the set of problems of dynamically maintaining topology of optical networks (especially when the links are created and/or destroyed), and maintaining information regarding link utilization, congestion, etc.

7. An important model not studied here is that of networks that are not all-optical. In these networks, a connection need not be carried on a single wavelength all the way to its destination; it could be carried on one wavelength to an intermediate node, where it is received and switched electronically onto another wavelength enroute to its destination. In this case, it is also possible to multiplex several connections on to a single wavelength, allowing packet switching.

References

[ALM90] S. Arora, T. Leighton, and B. Maggs On-line algorithms for path selection in a non-blocking network. In *Proc. of the ACM Symp. on Theory of Computing*, pages 149–158, 1990.

[Ben62] V.E. Beneš. Heuristic remarks and mathematical problems regarding the theory of switching systems. *Bell System Tech. Journal*, 41:1201–1247, 1962.

[BH92] R. A. Barry and P. A. Humblet. Bounds on the number of wavelengths needed in wdm networks. In *LEOS'92 Summer Topical Mtg. Digest*, pages 21–22, 1992.

[BH93] R. A. Barry and P. A. Humblet. On the number of wavelengths and switches in all-optical networks. *To appear in IEEE Trans. on Commun.*, 1993.

[BRSU93] A. Borodin, P. Raghavan, B. Schieber, and E. Upfal. How much can hardware help routing? In *Proc. of the 25th ACM Symp. on Theory of Computing*, San Diego, CA, May 1993.

[CNW90] N. K. Cheung, K. Nosu and G. Winzer (eds.) *IEEE JSAC: Special Issue on Dense WDM Networks.* volume 8, Aug. 1990.

[FFP88] P. Feldman, J. Friedman, and N. Pippenger. Wide-sense nonblocking networks. *SIAM Journal on Discrete Math.*, 1(2):158–173, May 1988.

[Gre92] P. E. Green. *Fiber-Optic Communication Networks.* Prentice Hall, 1992.

[Hal35] P. Hall. On the representatives of subsets. *Journal of London Mathematical Society*, 10(1):26–30, 1935.

[JRS93] F. J. Janniello, R. Ramaswami, and D. G. Steinberg. A prototype circuit-switched multi-wavelength optical metropolitan-area network. *J. Lightwave Tech.*, To appear, May 1993.

[Lei92] F.T. Leighton. *Introduction to parallel algorithms and architectures: arrays, trees, hypercubes.* Morgan Kaufmann Publishers, San Mateo, CA, 1992.

[Pan92] R.K. Pankaj. *Architectures for linear lightwave networks.* PhD thesis, Dept. of Electrical Engineering and Computer Science, MIT, Cambridge, MA, September 1992.

[PS93a] G. R. Pieris and G. H. Sasaki. A linear lightwave Beneš network. *Submitted to IEEE/ACM Trans. on Networking*, 1993.

[PS93b] G. R. Pieris and G. H. Sasaki. Scheduling transmissions in broadcast and select networks. *Unpublished manuscript*, 1993.

[Ram93] R. Ramaswami. Multi-wavelength lightwave networks for computer communication. *IEEE Communications Magazine*, 31(2):78–88, Feb. 1993.

[WZ93] A. Wigderson and D. Zuckerman. Expanders that beat the eigenvalue bound: explicit construction and applications. In *Proc. of the 25th ACM Symp. on Theory of Computing*, San Diego, CA, May 1993.