Performance Analysis of a Photonic Single-Hop ATM Switch Architecture with Tunable Transmitters and Fixed Frequency Receivers

Martin W. McKinnon *George N. Rouskas* *Harry G. Perros*

ARTS
P. O. Box 13164
RTP, NC 27709

Department of Computer Science
North Carolina State University
Raleigh, NC 27695-8206

Abstract

We consider a photonic ATM switch based on the single-hop WDM architecture with tunable transmitters and fixed frequency receivers. The switch operates under a schedule that masks the transceiver tuning latency. We analyze approximately a queueing model of the switch in order to obtain the queue-length distribution and the cell loss probability at the input and output ports. The analysis is carried out assuming two-state MMBP sources that capture the notion of burstiness and correlation, two important characteristics of ATM traffic, and non-uniform destination probabilities. Numerical results are presented to show the performance of the switch as a function of various system parameters, such as the load balancing and scheduling algorithms, the number of available channels, and buffer capacity.

Keywords: Optical networks, Photonic ATM switch architecture, Markov modulated Bernoulli process (MMBP), Wavelength Division Multiplexing (WDM), Discrete-time queueing networks

1 Introduction

One of the issues in evolving today's Asynchronous Transfer Mode (ATM) networks is that of developing switch architectures that can effectively switch cells at very high data rates (currently, data rates on the order of a few tens of Gigabits per second per port are envisioned). Over the last decade, a great deal of research has been devoted to the design of fast cell switches suitable to a broadband integrated services environment; surveys of some of these architectures may be found in [1, 2]. Mainstream research and development activities in the area of broadband switching are focused exclusively on electronics-based technologies which have attained a high level of maturity. On the other hand, the deployment of optics is limited to mere point-to-point transmission where the technology has proven successful in a short time span.

Given the continued rapid progress in lightwave technology (including the demonstration of fast tunable transceivers [3, 4], the development of erbium-doped fiber amplifiers [5], and guided-wave optical switching [6]), and the anticipated total dominance of optical fiber in the wired network, the issue of deeper penetration of optics naturally arises. Given the potential of optical solutions to cell switching, the possibility of employing photonics to implement switching functions hitherto reserved for electronics is currently being explored (see [7] and references thereof). However, there remain at least two major technical challenges to be overcome before one can contemplate the design of *all-optical* switches. First, there is the difficulty of "controlling light by light", and secondly, the technologies for implementing buffering in the optical domain are not yet mature enough. Consequently, the most likely scenarios for near-term photonic cell switching will involve an optical switching fabric with electronic control and buffering.

It has long been recognized that Wavelength Division Multiplexing (WDM) will be instrumental in bridging the gap between the speed of electronics and the virtually unlimited bandwidth available within the optical medium. The wavelength domain adds a significant new degree of freedom to network design, allowing new network concepts to be developed. With a few exceptions (e.g., [8, 9, 10]), however, most broadcast WDM architectures that have appeared in the literature require a large number of wavelengths and/or very fast tunable transceivers [11, 12, 13, 14, 15, 16, 17, 18]. Furthermore, the performance analysis of these architectures has been typically carried out assuming uniform traffic and memoryless arrival processes (see most of the above references, as well as [19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33]). However, it has been shown that, in order to study correctly the performance of a switch, one needs to use traffic models that capture the notion of burstiness and correlation, and which permit non-uniform output port destinations [34, 35].

In this paper we revisit the well known and widely studied single-hop broadcast-and-select WDM architecture [36]. Unlike previous work, however, we develop a queueing-based decomposition algorithm to study the performance of a single-hop ATM switch architecture operating under schedules that mask the transceiver tuning latency [9]. The analysis is carried out using arrival models that capture the important characteristics of ATM type of traffic, and non-uniform destination probabilities. Our work makes it possible to capture the complex interaction among the various system parameters such as the arrival processes, the number of available channels, and the scheduling and load balancing algorithms. To the best of our

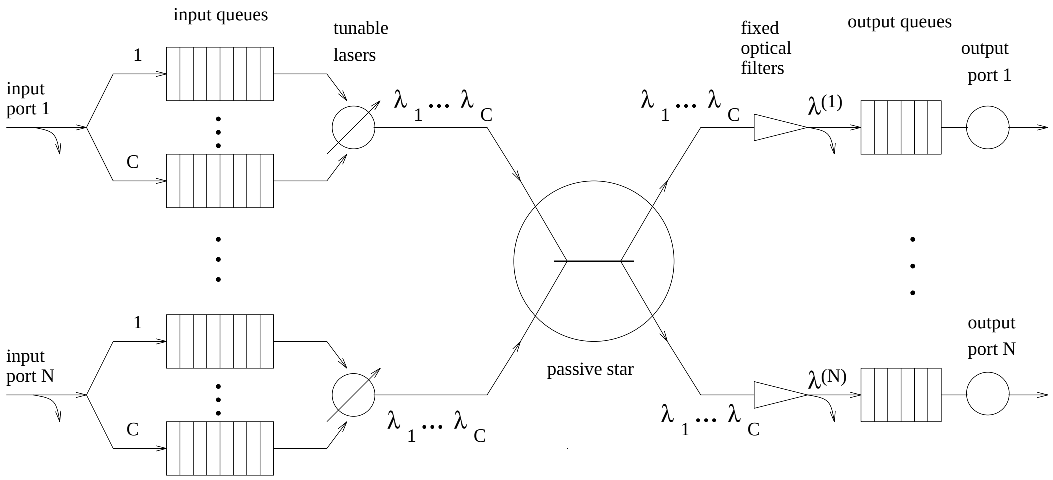

Figure 1: Queueing model of a switch architecture with N ports and C wavelengths

knowledge, such a comprehensive performance analysis of a single-hop WDM architecture has not been done before.

The next section presents our system model and provides some background information. The performance analysis of the switch is presented in Sections 3 and 4, numerical results are given in Section 5, and we conclude the paper in Section 6.

2 The ATM Switch Under Study

2.1 The Switch Architecture

We consider an optical switch architecture with N input ports and N output ports interconnected through a broadcast passive star (the switch fabric) that can support $C \leq N$ wavelengths $\lambda_1, \cdots, \lambda_C$ (see Figure 1). Each input port is equipped with a laser that enables it to inject signals into the optical medium. Similarly, each output port is capable of receiving optical signals through an optical filter. The laser at each input port is assumed to be tunable over all available wavelengths. The optical filters, on the other hand, are fixed to a given wavelength. Let $\lambda(j)$ denote the receive wavelength of output port j. Since $C \leq N$, a set $\mathcal{R}_c$ of output ports may be sharing a single receive wavelength λ_c:

$$\mathcal{R}_c \quad = \quad \{j \; : \; \lambda(j) = \lambda_c\}, \qquad c = 1, \cdots, C \tag{1}$$

Sets $\mathcal{R}_c$ will typically be obtained by running a load balancing algorithm [37].

The switch operates in a slotted mode. Since there are N ports but $C \leq N$ channels, each channel must run at a rate $\frac{N}{C}$ times faster than the rate of the input links ($\frac{N}{C}$ need not be an integer). The rate of an output link is equal to the rate of an input link. Thus, we distinguish between *arrival* slots (which correspond to the ATM cell transmission time at the input/output link rate) and *service* slots (which are equal to the cell transmission time at the channel rate within the switch). Obviously, the duration of a service slot is equal to $\frac{C}{N}$ times that of an arrival slot. Without loss of generality, we assume that all input links are synchronized at arrival slot boundaries; similarly for output links. On the other hand, all C channels internal to the switch are synchronized at service slot boundaries.

The switch employs electronic queueing at both the input and output ports, as Figure 1 illustrates. Cells arrive at an input port i and are buffered at a finite capacity queue, if the queue is not full. Otherwise, they are dropped. As Figure 1 indicates, the buffer space at each input port is assumed to be partitioned into C independent queues. Each queue c at input port i contains cells destined for the output ports which listen to a particular wavelength λ_c, $c = 1, \cdots, C$. This arrangement eliminates the head-of-line problem, and permits an input port to send a number of cells back-to-back when tuned to a certain wavelength. We let $B_{ic}^{(in)}$ denote the capacity of the queue at input port i corresponding to wavelength λ_c.

Cells buffered at an input port are transmitted on a FIFO basis onto the optical medium by the port's laser. This transmission takes place on an appropriate service slot which guarantees that the cell will be correctly received by its destination output port. Upon arriving at the output port, the cell is once again placed in a finite capacity buffer. Let $B_j^{(out)}$ denote the buffer capacity of output port j. Cells arriving at an output port to find a full buffer are lost. Cells in an output buffer are also served on a FIFO basis.

Interest in such a photonic switch architecture arises for several reasons:

- it is highly modular, allowing the switch to grow relatively easily by adding ports and wavelengths;
- it is scalable, since the number of wavelengths need not be equal to the number of ports, and since the data rate within the switch needs only be $\frac{N}{C}$ times the rate of the input/output links;
- it provides end-to-end optical paths;
- its hardware requirements, in terms of the number of transceivers per port, is minimum;
- it can be reconfigured [37] to adapt to changing traffic patterns or to overcome failures of ports or transceivers; and,
- it does not require extremely fast tunable transmitters (as explained below), and thus can be built using *currently available* tunable optical devices.

2.2 Transmission Schedules

One of the potentially difficult issues that arises in a WDM environment, is that of coordinating the various transmitters/receivers. Some form of coordination is necessary because (a) a transmitter and a receiver must both be tuned to the same channel for the duration of a cell's transmission, and (b) a simultaneous transmission by one or more input ports on the same channel will result in a *collision*. The issue of coordination is further complicated by the fact that tunable transceivers need a non-negligible amount of time to switch between wavelengths. For the Gigabit per second rates envisioned here, and for 53-byte ATM cells, the

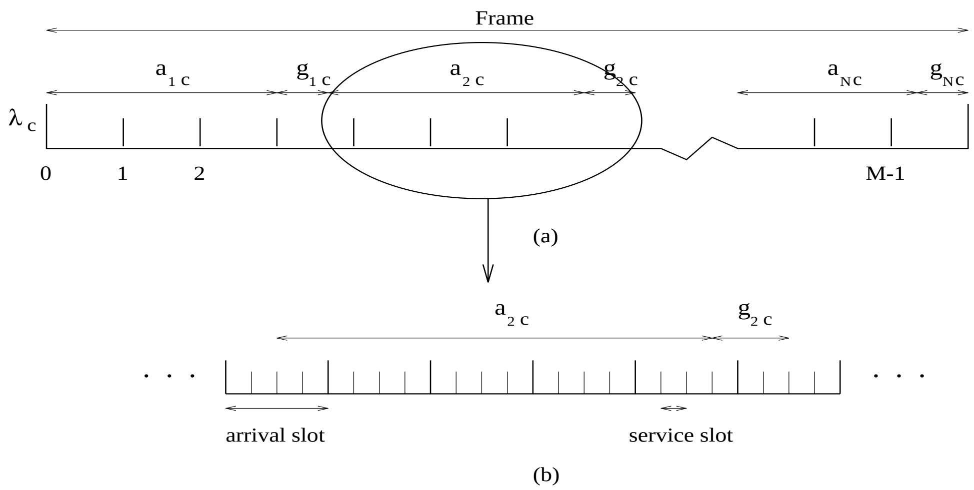

Figure 2: (a) Schedule for channel λ_c, and (b) detail corresponding to input port 2

tuning latency of state-of-the-art tunable lasers or filters can be as long as several times the size of a service slot [3]. Consequently, approaches that require each tunable transmitter to send a single cell and then switch to a new wavelength, will suffer a high tuning overhead and will result in a very low throughput.

In a recent paper [9], it was shown that careful scheduling can mask the effects of arbitrarily long tuning latencies, making it possible to build high-throughput photonic ATM switches using *currently available* lightwave technology. The key idea is to have each tunable transmitter send a *block* of cells on a wavelength before switching to another one. The main result of [9] was a set of new algorithms for constructing near-optimal (and, under certain conditions, optimal) schedules for transmitting a set of traffic demands $\{a_{ic}\}$. Quantity a_{ic} represents the number of cells to be transmitted by input port i onto channel λ_c per frame. The schedules are such that no collisions occur. Furthermore, they are easy to implement in a high speed environment, since the order in which the various input ports transmit is the same for all channels [9].

Quantity $a_{ic},\ i = 1, \cdots, N,\ c = 1, \cdots, C$, can be seen as the number of service slots per frame allocated to input port i, so that the port can satisfy the required quality of service of its incoming traffic intended for wavelength λ_c. By fixing a_{ic}, we indirectly allocate a certain amount of the bandwidth of wavelength λ_c to port i. This bandwidth could be equal to the effective bandwidth of the total traffic carried by input port i on wavelength λ_c. In general, the estimation of the quantities $a_{ic},\ i = 1, \cdots, N,\ c = 1, \cdots, C$, is part of the call admission algorithm, and it is beyond the scope of this paper. We note that as the traffic varies, a_{ic} may vary as well. In this paper, we assume that quantities a_{ic} are fixed, since this variation will more likely take place over larger scales in time.

We assume that transmissions by the input ports onto wavelength λ_c follow a schedule as shown in Figure 2. This schedule repeats over time. Each frame of the schedule consists of M arrival slots. Within each frame, input port i is assigned a_{ic} *contiguous* service slots for transmitting cells on channel λ_c. These a_{ic} slots are followed by a *gap* of $g_{ic} \geq 0$ slots during which no port can transmit on λ_c. This gap may be necessary to ensure that input port $i+1$ has sufficient time to tune from wavelength λ_{c-1} to λ_c before it starts transmission. The algorithms in [9] are such that the number of slots in most of the gaps is equal to either zero or a small integer. Thus, the length of the schedule is very close to the lower bound $\max_i\{\sum_{c=1}^{C} a_{ic}\}$. Note that in Figure 2 we have assumed that an arrival slot is an integer multiple of service slots. This may not be true in general, and it is not a necessary assumption for our model. Observe also that, although the frame begins and ends on *arrival* slot boundaries, the beginning or end of transmissions by a port does not necessarily coincide with the beginning or end of an *arrival* slot (although it is, obviously, synchronized with *service* slots).

2.3 Traffic Model

The arrival process to each input port of the switch is characterized by a two-state Markov Modulated Bernoulli Process (MMBP), hereafter referred to as 2-MMBP. A 2-MMBP is a Bernoulli process whose arrival rate varies according to a two-state Markov chain. It captures the notion of burstiness and the correlation of successive interarrival times, two important characteristics of ATM type of traffic. For details on the properties of the 2-MMBP, the reader is referred to [38]. (We note that the algorithm for analyzing the switch was developed so that it can be readily extended to MMBPs with more than two states.)

We assume that the arrival process to port i, $i = 1, \cdots, N$, is given by a 2-MMBP characterized by the transition probability matrix $\mathbf{Q}_i$, and by $\mathbf{A}_i$ as follows:

$$\mathbf{Q}_i = \begin{bmatrix} q_i^{(00)} & q_i^{(01)} \\ q_i^{(10)} & q_i^{(11)} \end{bmatrix} \qquad \text{and} \qquad \mathbf{A}_i = \begin{bmatrix} \alpha_i^{(0)} & 0 \\ 0 & \alpha_i^{(1)} \end{bmatrix} \tag{2}$$

In (2), $q_i^{(kl)}, k, l = 0, 1$, is the probability that the 2-MMBP will make a transition to state l, given that it is currently at state k. Obviously, $q_i^{(k0)} + q_i^{(k1)} = 1,\ k = 0, 1$. Also, $\alpha_i^{(0)}$ $(\alpha_i^{(1)})$ is the probability that an arrival will occur in a slot at state 0 (1). Transitions between states of the 2-MMBP occur only at the boundaries of *arrival* slots. We assume that the arrival process to each input port is given by a different 2-MMBP.

Let r_{ij} denote the probability that a cell arriving to input port i will have j as its destination output port. We will refer to $\{r_{ij}\}$ as the *routing* probabilities; this description implies that the routing probabilities can be input port dependent and non-uniformly distributed. The destination probabilities of successive cells are not correlated. That is, in an input port, the destination of one cell does not affect the destination of the cell behind it. This assumption is reasonable when the switch is used as part of a backbone network. Given these assumptions, the probability that a cell arriving to port i will have to be transmitted on wavelength λ_c is:

$$r_{ic} = \sum_{j \in \mathcal{R}_c} r_{ij}, \qquad i = 1, \cdots, N \tag{3}$$

3 Queueing Analysis

In this section we analyze the queueing network shown in Figure 1, which represents the tunable-transmitter, fixed-receiver switch under study. The arrival process to each input port is assumed to be a 2-MMBP, and the access of the input ports to the wavelengths is governed by the schedule described in Section 2.2. We analyze this queueing network in order to obtain the queue-length distribution in an input or output port, from which performance measures such as the cell-loss probability can be obtained.

3.1 Input Side Analysis

In this section, we obtain the queue length distribution of an input queue. We first sketch an exact decomposition of the corresponding queueing network which, however, is not scalable to large systems. Then, we present in detail an approximation method which, as we will show later, gives accurate results.

3.1.1 Exact Queueing Analysis

We first observe that we can analyze the input side of th switch by decomposing it into N sub-systems, each corresponding to an input port, and analyzing each sub-system in isolation. Because of the fact that (a) the arrival processes to the various input queues are independent, (b) the way the schedule is constructed (i.e., that different inputs transmit to the same wavelength at different times), and (c) the operation of the input ports is independent of the operation of output ports, this decomposition is exact. Furthermore, we can analyze the sub-system corresponding to input queue i by defining a $(C+2)$-dimensional stochastic process $(x, y_1, \cdots, y_C, z)$, where

- x represents the arrival slot number within a frame ($x = 0, 1, \cdots, M-1$),
- y_c indicates the number of cells in the input queue servicing λ_c ($y_c = 0, 1, \cdots, B_{ic}^{(in)}$; $c = 1, \cdots, C$), and
- z indicates the state of the 2-MMBP describing the arrival process to this queue, that is, $z = 0, 1$.

It is easy to verify that this process defines a Markov chain and, thus, the steady state joint occupancy distribution of the C queues of input port i can be obtained. Unfortunately, the state space of the Markov chain grows in size as $O(2M \prod_{c=1}^{C} B_{ic}^{(in)})$. As a result, this analysis can only be applied to trivial systems. In the next subsection, we describe an approximate decomposition that can be applied to large systems.

3.1.2 Approximate Queueing Analysis

Our main approximation is to assume that arrivals to each queue of a given input port are independent and are generated by the original 2-MMBP (which characterizes the arrival process to the input port) appropriately thinned using the routing probabilities r_{ic}.

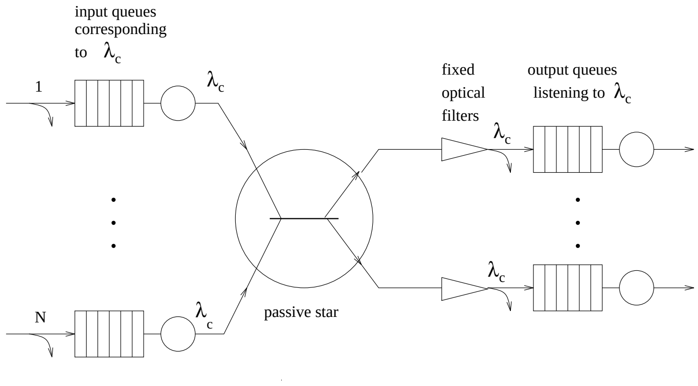

Figure 3: Queueing sub-network for wavelength λ_c

Assuming independence of arrivals among the queues of each input port, the original queueing network can now be decomposed into C sub-networks, one per wavelength, as in Figure 3. For each wavelength λ_c, the corresponding sub-network consists of N input queues, and all the output queues that listen to wavelength λ_c. Each input queue of the sub-network is the queue associated with wavelength λ_c in each input port of the switch. That is, the i-th input queue of this sub-network is the c-th queue of input port i. Since throughout this section we only consider the sub-network corresponding to λ_c, we will simply refer to this queue as "input queue i". These input queues will transmit to the output queues of the sub-network over wavelength λ_c. In view of this decomposition, it suffices to analyze a single sub-network, since the same analysis can be applied to all other sub-networks.

Consider now the sub-network for wavelength λ_c. We will analyze this sub-network by decomposing it into individual input and output ports. As discussed in the previous section, each input queue i of the sub-network is only served for a_{ic} consecutive service slots per frame. During that time, no other input port is served. Input queue i is not served in the remaining slots of the frame. In view of this, there is no dependence among the input queues of the sub-network, and consequently each one can be analyzed in isolation in order to obtain its queue-length distribution.

From the queueing point of view, the queueing network shown in Figure 3 can be seen as a polling system in discrete time. Despite the fact that polling systems have been extensively analyzed, we note that very little work has been done within the context of discrete time (see, for example, [39]). In addition, this particular problem differs from the typical polling system since we consider output queues, which are not typically analyzed in polling systems.

3.1.3 The Queue-Length Distribution of an Input Queue

Consider input queue i of the sub-network for λ_c in isolation. This input queue receives exactly a_{ic} service slots on wavelength λ_c, as shown in Figure 4(a). The block of a_{ic} service slots may not be aligned with the boundaries of the arrival slots. For instance, in the example shown in Figure 4(a), the block of a_{ic} service slots begins at the second service slot of arrival slot $x-1$, and it ends at the end of the second service slot

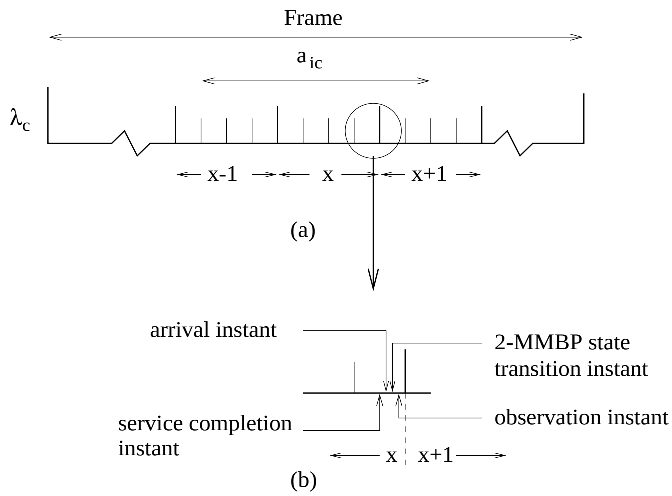

Figure 4: (a) Service period of input port i on channel λ_c, and (b) detail showing the relationship among service completion, arrival, 2-MMBP state transition, and observation instants within a service and an arrival slot

in arrival slot $x+1$. Here, $x-1$, x, and $x+1$ represent the arrival slot number within a frame.

For each arrival slot, define $v_{ic}(x)$ as the number of service slots allocated to input queue i, that lie within arrival slot x [1]. Then, in the example in Figure 4(a), we have: $v_{ic}(x-1) = 3, v_{ic}(x) = 4, v_{ic}(x+1) = 2$, and $v_{ic}(x') = 0$ for all other x'. Obviously we have,

$$\sum_{x=0}^{M-1} v_{ic}(x) = a_{ic} \tag{4}$$

We analyze input queue i by constructing its underlying Markov chain embedded at arrival slot boundaries. The order of events is as follows. The service (i.e., transmission) completion of a cell occurs at an instant just before the end of a service slot. An arrival may occur at an instant just before the end of an arrival slot, but after the service completion instant of a service slot whose end is aligned with the end of an arrival slot. The 2-MMBP describing the arrival process to the queue makes a state transition immediately after the arrival instant. Finally, the Markov chain is observed at the boundary of each arrival slot, *after* the state transition by the 2-MMBP. The order of these events is shown in Figure 4(b).

The state of the input queue is described by the tuple (x, y, z), where:

- x represents the arrival slot number within a frame ($x = 0, 1, \cdots, M-1$),
- y indicates the number of cells in the input queue ($y = 0, 1, \cdots, B_{ic}^{(in)}$), and
- z indicates the state of the 2-MMBP describing the arrival process to this queue, that is, $z = 0, 1$.

[1] In Figure 4, we assume that each arrival slot contains an integral number of service slots. If this is not the case, $v_{ic}(x)$ is defined as the number of service slots that are concluded within arrival slot x (i.e., if there is a service slot that lies partially within arrival slots x and $x+1$, it will be counted in $v_{ic}(x+1)$).

Table 1: Transition probabilities out of state (x, y, z) of the Markov chain

Current State	Next State	Transition Probability
(x, y, z)	$(x \oplus 1, \max\{0, y - v_{ic}(x \oplus 1)\}, z')$	$q_i^{(zz')}(1 - \alpha_i^{(z)} r_{ic})$
(x, y, z)	$(x \oplus 1, \min\{B_{ic}^{(in)}, \max\{0, y - v_{ic}(x \oplus 1)\} + 1\}, z')$	$q_i^{(zz')} \alpha_i^{(z)} r_{ic}$

It is straightforward to verify that, as the state of the queue evolves in time, it defines a Markov chain. Let $\oplus$ denote modulo-M addition, where M is the number of arrival slots per frame. Then, the transition probabilities out of state (x, y, z) are given in Table 1. Note that, the next state after (x, y, z) always has an arrival slot number equal to $x \oplus 1$. In the first row of Table 1 we assume that the 2-MMBP makes a transition from state z to state z' (from (2), this event has a probability $q_i^{(zz')}$ of occurring), and that no cell arrives to this queue during the current slot (from (2) and (3), this occurs with probability $1 - \alpha_i^{(z)} r_{ic}$). Since at most $v_{ic}(x \oplus 1)$ cells are serviced during arrival slot $x \oplus 1$, and since no cell arrives, the queue length at the end of the slot is equal to $\max\{0, y - v_{ic}(x \oplus 1)\}$. In the second row of Table 1 we assume that the 2-MMBP makes a transition from state z to state z' and a cell arrives to the queue. This arriving cell cannot be serviced during this slot, and has to be added to the queue. Finally, the expression for the new queue length ensures that it will not exceed the capacity $B_{ic}^{(in)}$ of the input queue.

We observe that the probability transition matrix of this Markov chain has the following block form:

$$
\mathbf{S}_{ic} \quad = \quad \begin{array}{c} \left[\begin{array}{cccccc} 0 & \mathbf{R}_{ic}(0) & 0 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{R}_{ic}(1) & 0 & \cdots & 0 \\ 0 & 0 & 0 & \mathbf{R}_{ic}(2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \mathbf{R}_{ic}(M-2) \\ \mathbf{R}_{ic}(M-1) & 0 & 0 & 0 & \cdots & 0 \end{array} \right] \end{array} \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ M-2 \\ M-1 \end{array} \tag{5}
$$

This block form is due to the fact that at each transition instant (i.e., at each arrival slot boundary), the random variable x changes to $x \oplus 1$. Changes in the other two random variables, y and z, of the state of the queue are governed by the matrices $\mathbf{R}_{ic}(x)$. There are M different $\mathbf{R}_{ic}$ matrices, one for each arrival slot x in the frame. Let us define matrices $\mathbf{X}_{ic}$ and $\mathbf{Y}_{ic}$ as follows:

$$
\mathbf{X}_{ic} \;=\; r_{ic}\,\mathbf{A}_i\,\mathbf{Q}_i \qquad \text{and} \qquad \mathbf{Y}_{ic} \;=\; (\mathbf{I} - r_{ic}\,\mathbf{A}_i)\,\mathbf{Q}_i, \tag{6}
$$

where $\mathbf{I}$ is the identity matrix. Then, the transition matrix $\mathbf{R}_{ic}(x)$ associated with arrival slot x can be written as:

$$
\mathbf{R}_{ic}(x) = \left[\begin{array}{ccccccccc} \mathbf{Y}_{ic} & \mathbf{X}_{ic} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{Y}_{ic} & \mathbf{X}_{ic} & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \mathbf{Y}_{ic} & \mathbf{X}_{ic} & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{Y}_{ic} & \mathbf{X}_{ic} & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \mathbf{Y}_{ic} & \mathbf{X}_{ic} & 0 & \cdots & 0 \end{array} \right] \begin{array}{c} 0 \\ \vdots \\ v_{ic}(x \oplus 1) \\ v_{ic}(x \oplus 1) + 1 \\ v_{ic}(x \oplus 1) + 2 \\ \vdots \\ B_{ic}^{(in)} \end{array} \tag{7}
$$

The structure of matrix $\mathbf{R}_{ic}(x)$ given in (7) can be explained as follows. Suppose that the number of cells y in the queue at the end of slot x is at most $v_{ic}(x \oplus 1)$. Since up to $v_{ic}(x \oplus 1)$ cells can be served within slot $x \oplus 1$, the number in the queue at the end of that slot will be 1 or 0, depending on whether an arrival occurred or not. This is indicated by the transitions in rows 0 through $v_{ic}(x \oplus 1)$ of matrix $\mathbf{R}_{ic}(x)$. However, if at the end of slot x we have $y > v_{ic}(x \oplus 1)$, then the number in the queue at the next transition will be $y - v_{ic}(x \oplus 1)$ (plus one if an arrival occurred). This is indicated by the transitions in rows $v_{ic}(x \oplus 1) + 1$ through B_{ic} of $\mathbf{R}_{ic}(x)$. Of course, y cannot exceed the queue capacity $B_{ic}^{(in)}$. Since the number of service slots $v_{ic}(x \oplus 1)$ depends on the particular slot $x \oplus 1$ within the frame, $\mathbf{R}_{ic}(x)$ is a function of x.

Matrix $\mathbf{R}_{ic}(x)$ is slightly different when $v_{ic}(x \oplus 1) = 0$. This is because, in this case, if the state of the input queue is such that $y = B_{ic}^{(in)}$, a new arrival will be discarded. So when $y = B_{ic}^{(in)}$, the 2-MMBP is allowed to make a transition, but regardless of whether or not an arrival is generated, the number of cells in the queue will remain equal to $B_{ic}^{(in)}$. Thus, the the last row of $\mathbf{R}_{ic}(x)$ will be: $[0 \ 0 \ \cdots \ 0 \ \mathbf{Q}_i]$.

It is now straightforward to verify that the Markov chain with transition matrix $\mathbf{S}_{ic}$ is irreducible, and therefore a steady-state distribution exists. Transition matrix $\mathbf{S}_{ic}$ defines a *p-cyclic* Markov chain [40], and therefore it can be solved using any of the techniques for p-cyclic Markov chains in [40, ch. 7]. We have used the LU decomposition method in [40] to obtain the steady state probability $\pi_{ic}(x, y, z)$ that at the end of arrival slot x, the 2-MMBP is in state z and the input queue has y cells. The steady-state probability that the queue has y cells at the end of slot x, independent of the state of the 2-MMBP is:

$$
\pi_{ic}(x, y) = \sum_{z=0,1} \pi_{ic}(x, y, z) \tag{8}
$$

Finally, we note that all of the results obtained in this subsection can be readily extended to MMBP-type arrival processes with more than two states. For this, it would suffice to appropriately modify matrices $\mathbf{X}_{ic}$ and $\mathbf{Y}_{ic}$.

3.2 Output Side Analysis

We now obtain the queue-length distribution of an output queue. Our analysis follows steps similar to the input side case.

3.2.1 Exact Queueing Analysis

Let us suppose that the (exact or approximate) queue-length distribution of the input queues is known. Given that transmissions on different channels are independent, and that output queue receivers operate on one frequence, the output side of the switch may be decomposed into C independent sub-systems, one per wavelength. Let us consider the sub-network corresponding to channel λ_c, and let k_c be the number of output ports sharing this channel. We can then define a $(k_c + 1)$-dimensional Markov chain, where

- x indicates the arrival slot number within the frame ($x = 0, 1, \cdots, M - 1$), and
- $w_n \ \forall \ n = 1, \cdots, k_c$ indicates the number of cells at the n^{th} output queue which shares λ_c ($w_n = 0, 1, \cdots, B_j^{(out)}$).

The transitions out of state $(x, w_1, \cdots, w_{k_c})$ can be computed given the schedule and the queue length distribution of the input ports. However, for realistic switch dimensions, this method will lead to a state space explosion since the total number of states is of the order of $M \times \prod_{j \in \mathcal{R}_c} B_j^{(out)}$. We now proceed to describe an approximation method that can be used for large systems.

3.2.2 Approximate Queueing Analysis

Consider the sub-network for wavelength λ_c, and observe that the arrival process to the output queues sharing λ_c is the combination of the departure processes from the input queues corresponding to λ_c. An interesting aspect of the departure process from the input queues is that for each frame, during the sub-period a_{ic} we only have departures from the i-th input queue. This period is then followed by a gap g_{ic} during which no departure occurs. This cycle repeats for the next input queue. Thus, in order to characterize the overall departure process offered as the arrival process to these output queues, it suffices to characterize the departure process from each input queue, and then combine them. (We note that this overall departure process is quite different from the typical superposition of a number of departure processess into a single stream, where, at each slot, more than one cell may be departing.)

However, the arrival processes to the output queues listening on λ_c are not independent. Specifically, if j and j' are two output ports on λ_c, and there is a transmission from input port i to output port j in a given service slot, then there can be no arrival to output port j' in the same service slot. As in the input side case, we will nevertheless make the assumption that these arrival processes are indeed independent, and that each is an appropriately thinned (based on the routing probabilities) version of the departure process from the input queues. Note that this is an approximation only when there are multiple output ports with receivers fixed on channel λ_c.

3.2.3 The Queue-Length Distribution of an Output Queue

As in the previous section, we obtain the queue-length distribution of output port j at arrival slot boundaries. Recall that an arrival slot to an input queue is equal to a departure slot from an output queue. Also, arrival

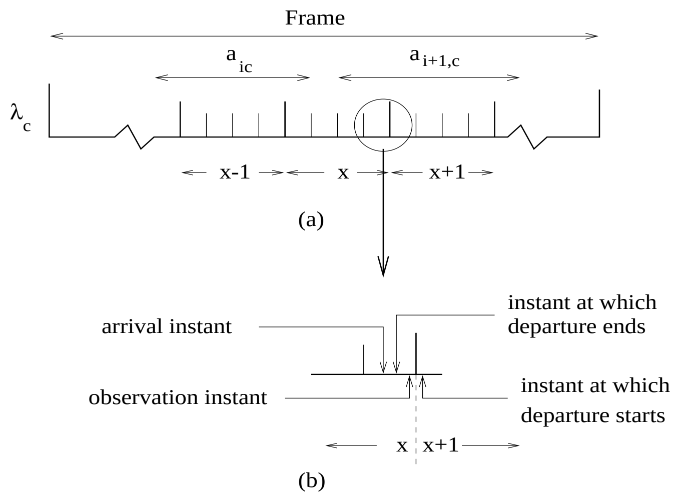

Figure 5: (a) Arrivals to output port j from input ports i and $i+1$, and (b) detail showing the relationship of departure, arrival, and observation instants

and departure slots are synchronized. Therefore, during an arrival slot x a cell may be transmitted to the outgoing link from the output queue. However, during slot x, there may be several arrivals to the output queue from the input queues.

Let (x, w) be the state associated with output port j, where

- x indicates the arrival slot number within the frame ($x = 0, 1, \cdots, M-1$), and
- w indicates the number of cells at the output queue ($w = 0, 1, \cdots, B_j^{(out)}$).

We assume the following order of events. A cell will begin to depart from the output queue at an instant immediately after the beginning of an arrival slot and the departure will be completed just before the end of the slot. A cell from an input port arrives at an instant just before the end of a service slot, but before the end-of-departure instant of an arrival slot whose end is aligned with the end of the service slot. Finally, the state of the queue is observed just before the end of an arrival slot and after the arrival associated with the last service slot has occurred (see Figure 5(b)).

Let $u_j(x)$ be the number of service slots of any input queue on wavelength λ_c within arrival slot x. We have that

$$u_j(x) \;=\; \sum_{i=1}^{N} v_{ic}(x) \tag{9}$$

where $v_{ic}(x)$ is as defined in (4). Quantity $u_j(x)$ represents the maximum number of cells that may arrive to output port j within slot x. In the example of Figure 5(a) where we show the arrival slots during which cells from input ports i and $i+1$ may arrive to output port j, we have: $u_j(x-1) = v_{ic}(x-1) = 4$, $u_j(x) = v_{ic}(x) + v_{i+1,c}(x) = 1 + 2 = 3$, and $u_j(x+1) = v_{i+1,c}(x+1) = 4$.

Observe now that (a) at each state transition x advances by one (modulo-M), (b) exactly one cell departs

Table 2: Transition probabilities out of state (x, w) of the Markov chain

Current State	Next State	Transition Probability
$(x, 0)$	$(x \oplus 1, \min\{B_j^{(out)}, s\}), 0 \leq s \leq u_j(x \oplus 1)$	$\sum_{s_1+\cdots+s_N=s} \prod_{i=1}^{N} L_i(s_i \mid x)$
$(x, w), w > 0$	$(x \oplus 1, \min\{B_j^{(out)}, w+s\} - 1), 0 \leq s \leq u_j(x \oplus 1)$	$\sum_{s_1+\cdots+s_N=s} \prod_{i=1}^{N} L_i(s_i \mid x)$

from the queue as long as the queue is not empty, (c) a number $s \leq u_j(x \oplus 1)$ of cells may be transmitted from the input ports to output port j within arrival slot $x \oplus 1$, and that (d) the queue capacity is $B_j^{(out)}$. Then, the transition probabilities out of state (x, w) for this Markov chain are given in Table 2.

In Table 2, $L_i(s_i \mid x)$ is the probability that input port i transmits s_i cells to output port j given that the system is at the end of arrival slot x (in other words, it is the probability that s_i cells are transmitted within slot $x \oplus 1$) [2]. To obtain $L_i(s_i \mid x)$, define r'_{ij} as the conditional probability that a cell is destined for output port j, given that the cell is destined to be transmitted on λ_c, the receive wavelength of output port j:

$$r'_{ij} = \frac{r_{ij}}{\sum_{k \in \mathcal{R}_c} r_{ik}} = \frac{r_{ij}}{r_{ic}} \tag{10}$$

This "thinning" of the arrival processes using the r'_{ij} routing probabilities discounts the correlation between arrival streams and is the crux of the approximation for the output side of the switch. The error introduced by this approximation will be discussed later in this work.

Define $\pi_{ic}(y \mid x)$ as the conditional probability of having y cells at the i-th input queue given that the system is observed at the end of slot x:

$$\pi_{ic}(y \mid x) = \frac{\pi_{ic}(x, y)}{\pi_{ic}(x)} = M\ \pi_{ic}(x, y) \tag{11}$$

Then, for $r'_{ij} < 1$, the probability $L_i(s_i \mid x)$ is given by

$$L_i(s_i \mid x) = \begin{cases} 0, & s_i > v_{ic}(x \oplus 1) \\ \sum_{y=s_i}^{B_{ic}^{(in)}} \pi_{ic}(y \mid x) \begin{pmatrix} \min\{y, v_{ic}(x \oplus 1)\} \\ s_i \end{pmatrix} (r'_{ij})^{s_i} (1 - r'_{ij})^{\min\{y, v_{ic}(x \oplus 1)\} - s_i}, & s_i \leq v_{ic}(x \oplus 1) \end{cases} \tag{12}$$

Expression (12) can be explained by noting that input port i will transmit s_i cells to output port j during arrival slot $x \oplus 1$ if (a) $v_{ic}(x \oplus 1) \geq s_i$, (b) input port i has $y \geq s_i$ cells in its queue for λ_c at the beginning of the slot (equivalently, at the end of slot x), and (c) exactly s_i of $\min\{y, v_{ic}(x \oplus 1)\}$ cells that will be transmitted by this queue in this arrival slot are for output j.

[2] Since in most cases only one or two input ports will transmit to the same channel within an arrival slot (refer also to Figure 2), the summation and product in the expression in the last column of Table 2 does not necessarily run over all N values of i, only over one or two values of i. Thus, this expression can be computed very fast, not in exponential time as implied by the general form presented in the table.

If $r'_{ij} = 1$, in which case j is the only port listening on wavelength λ_c, the expression for $L_i(s_i \mid x)$ must be modified as follows:

$$L_i(s_i \mid x) \quad = \quad \begin{cases} 0, & s_i > v_{ic}(x \oplus 1) \\ \pi_{ic}(s_i \mid x), & s_i < v_{ic}(x \oplus 1) \\ \sum_{y=s_i}^{B_{ic}^{(in)}} \pi_{ic}(y \mid x), & s_i = v_{ic}(x \oplus 1) \end{cases} \tag{13}$$

Expressions (12) and (13) are based on the assumption that $v_{ic}(x \oplus 1) < B_{ic}^{(in)}$ which we believe is a reasonable one. In the general case, quantity $v_{ic}(x \oplus 1)$ in both expressions must be replaced by $\min\{v_{ic}(x \oplus 1), B_{ic}^{(in)}\}$.

The transition matrix $\mathbf{T}_j$ of the Markov chain defined by the evolution of the state (x, w) of output queue j has the following form, which is similar to that of matrix $\mathbf{S}_{ic}$ given by (5):

$$\mathbf{T}_j \quad = \quad \begin{array}{c c} \left[\begin{array}{cccccc} 0 & \mathbf{U}_j(0) & 0 & 0 & \cdots & 0 \\ 0 & 0 & \mathbf{U}_j(1) & 0 & \cdots & 0 \\ 0 & 0 & 0 & \mathbf{U}_j(2) & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \mathbf{U}_j(M-2) \\ \mathbf{U}_j(M-1) & 0 & 0 & 0 & \cdots & 0 \end{array} \right] & \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \\ M-2 \\ M-1 \end{array} \end{array} \tag{14}$$

$\mathbf{U}_j(x)$ is a $(B_j^{(out)} + 1) \times (B_j^{(out)} + 1)$ matrix that governs changes in random variable w of the state of the output queue. The elements of this matrix can be determined using Table 2 and expression (12) or (13). Since $L_i(s_i \mid x)$ depends on $v_{ic}(x)$ and $u_j(x)$, $\mathbf{U}_j(x)$ also depends on x, the slot number within the frame.

We observe that $\mathbf{T}_j$ also defines a *p-cyclic* Markov chain. We have used the LU decomposition method as prescribed in [40] to obtain $\pi_j(x, w)$, the steady-state probability that output queue j has w cells at the end of slot x.

3.3 Summary of the Decomposition Algorithm

Below we summarize our approach to analyzing the sub-network of Figure 3 corresponding to wavelength λ_c. We assume that quantities $\{a_{ic}\}$ and the corresponding schedule (see [9]) are given.

1. For each arrival slot x, use the schedule and expressions (4) and (9) to compute the quantities $v_{ic}(x)$ and $u_j(x)$, $i = 1, \cdots, N,\ j : \lambda(j) = \lambda_c$.

2. For each input queue i, construct the transition probability matrix $\mathbf{S}_{ic}$ from (2), (3), (5), (6), and (7). Solve this matrix and use (8) to obtain the steady-state probability $\pi_{ic}(x, y)$ that input queue i has y cells at the end of the x-th slot of the frame.

3. For each output port $j \in \mathcal{R}_c$, use $\pi_{ic}(x, y)$ derived in Step 2, and (12)-(13) to construct the transition matrix $\mathbf{T}_j$ given by (14). Solve the matrix as in Step 2 to obtain $\pi_j(x, w)$, the steady-state probability that port j has w cells in its queue at the end of slot x.

Note that the complexity of this approach is dominated by Step 2. For each of the N input queues we have to solve a matrix of dimensions $\left[2M(B_{ic}^{(in)}+1)\right] \times \left[2M(B_{ic}^{(in)}+1)\right]$, where M is the length of the schedule (in arrival slots) and $B_{ic}^{(in)}$ is the capacity of the respective queue. (Inverting a $K \times K$ matrix takes time $O(K^3)$, although we can take advantage of the fact that the matrix is sparse to solve for the queue-length distributions at a significantly faster rate.) Thus, in the worst case, the overall complexity of our algorithm is $O(NM^3B^3)$, where $B = \max_i\{B_{ic}^{(in)}\}$.

4 Cell-Loss Probability

We now use the queue-length distributions for the input and output ports, $\pi_{ic}(x, y)$ and $\pi_j(x, w)$, respectively, derived in the previous section, to obtain the cell-loss probability at the input and output ports.

4.1 The Cell-Loss Probability at an Input Port

Let Ω_{ic} be the cell-loss probability at the c-th queue of input port i, i.e., the probability that a cell arriving to that queue will be lost. Ω_{ic} can be expressed as:

$$\Omega_{ic} \quad = \quad \frac{E[\text{number of cells lost per frame at queue } c \text{ of port } i]}{E[\text{number of arrivals per frame at queue } c \text{ of port } i]} \tag{15}$$

Obtaining the expectation in the denominator is easy. From (2) and [38], the steady-state arrival probability for the arrival process to this queue is

$$\gamma_i \quad = \quad \frac{q_i^{(10)}\alpha_i^{(0)} + q_i^{(01)}\alpha_i^{(1)}}{q_i^{(01)} + q_i^{(10)}} \tag{16}$$

Then,

$$E[\text{number of arrivals per frame at queue } c \text{ of port } i] \quad = \quad M\ \gamma_i\ r_{ic} \tag{17}$$

To obtain the expectation in the numerator, let us refer to Figure 4(b) which shows the service completion, arrival, and observation instants within slot x. We observe that, due to the fact that at most one cell may arrive in slot x, if the number $v_{ic}(x)$ of slots during which this queue is serviced within arrival slot x is not zero (i.e., $v_{ic}(x) > 0$), no arriving cell will be lost. Even if the c-th queue at input port i is full at the beginning of slot x, $v_{ic}(x) \geq 1$ cells will be serviced during this slot, and the order of service completion and arrival instants in Figure 4(b) guarantees that an arriving cell will be accepted. On the other hand, if $v_{ic}(x) = 0$ for slot x, then an arriving cell will be discarded if and only if the queue is full at the beginning of x (equivalently, at the end of the slot before x). Since the 2-MMBP can be in one of two states, we have that

$$E[\text{number of cells lost per frame at queue } c \text{ of port } i] \quad = \quad \sum_{x:v_{ic}(x)=0} \sum_{z=0}^{1} \alpha_i^{(z)} r_{ic} \pi_{ic}(B_{ic}, z \mid x \ominus 1) \tag{18}$$

In (18), $\ominus$ denotes regular subtraction with the exception that, if $x = 0$, then $x \ominus 1 = M - 1$, and the summation runs over all x for which $v_{ic}(x) = 0$. Using (15), (17), (18), and the fact that $\pi_{ic}(x) = \frac{1}{M}$ for all x, we obtain an expression for Ω_{ic} as follows:

$$\Omega_{ic} = \frac{\sum_{x:v_{ic}(x)=0} \sum_{z=0}^{1} \alpha_i^{(z)} \pi_{ic}(x \ominus 1, B_{ic}^{(in)}, z)}{\gamma_i} \tag{19}$$

4.2 The Cell-Loss Probability at an Output Port

The cell-loss probability at an output port is more complicated to calculate, since we may have multiple cell arrivals to the given output port within a single arrival slot (refer to Figure 5(a)). Let us define $\Omega_j(n \mid x)$ as the conditional probability that n cells will be lost at output queue j given that the current arrival slot is x. An output port will lose n cells in slot x if (a) the port had $w, 0 \le w \le B_j^{(out)}$ cells at the beginning of slot x, and (b) exactly $B_j^{(out)} - w + n$ cells arrived during slot x. We can then write:

$$\Omega_j(n \mid x) = \sum_{w=0}^{B_j^{(out)}} \pi_j(w \mid x \ominus 1)\ Pr[B_j^{(out)} - w + n \text{ cells arrive to } j \mid x] \tag{20}$$

where $\pi_j(w \mid x \ominus 1) = M\pi(x \ominus 1, w)$ similar to (11). The last probability in (20) can be obtained using (12) or (13), as in Table 2:

$$Pr[s \text{ cells arrive to output port } j \mid x] = \sum_{s_1+\cdots+s_N=s} \prod_{i=1}^{N} L_i(s_i \mid x \ominus 1) \tag{21}$$

Note that at most $u_j(x)$ cells may arrive (and get lost) in arrival slot x. Using (20), we can then compute the expected number of cells lost in slot x as:

$$E[\text{number of cells lost at } j \mid x] = \sum_{n=1}^{u_j(x)} n\Omega_j(n \mid x) \tag{22}$$

The expected number of arrivals to port j in slot x can be computed using (21):

$$E[\text{number of arrivals to } j \mid x] = \sum_{s=1}^{u_j(x)} s\ Pr[s \text{ cells arrive to } j \mid x] \tag{23}$$

Finally, the probability Ω_j that an arriving cell to output port j will be lost regardless of the arrival slot x can be found as follows:

$$\Omega_j = \frac{E[\text{number of lost cells in a frame}]}{E[\text{number of arrivals in a frame}]} = \frac{\sum_{x=0}^{M-1} E[\text{number of lost cells at } j \mid x]}{\sum_{x=0}^{M-1} E[\text{number of arrivals to } j \mid x]} \tag{24}$$

5 Numerical Results

We now demonstrate the accuracy of out analysis by applying the decomposition algorithm to a 16×16 switch and comparing it to simulation results.

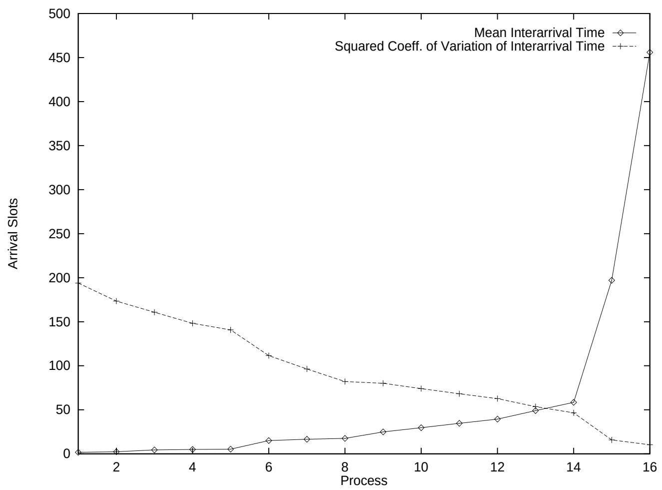

Figure 6: Mean arrival rate and squared coefficient of variation of the interarrival time for the arrival processes to the 16 input ports of the switch

We have selected the following set of parameters for our study case. The arrival process to each of the ports of the switch is described by a different 2-MMBP. In Figure 6, we plot two important parameters of each of the 16 2-MMBPs we have used: the mean interarrival time in slots (γ_i^{-1} in (16)), and the squared coefficient of variation of the interarrival time given in [38]. As can be seen, the arrival processes exhibit a wide range of behavior in terms of these two parameters. The routing probabilities used are:

$$r_{ij} \quad = \quad \begin{cases} 0.10, & i = 1, \cdots, 16, \; j = 1 \\ 0.06, & i = 1, \cdots, 16, \; j = 2, \cdots, 16 \end{cases} \tag{25}$$

That is, output port 1 is a hot spot, receiving 10% of the total traffic, while the remaining traffic is evenly distributed to the other 15 ports. The total arrival rate to the network is $\gamma = \sum_i \gamma_i = 1.98$ cells per arrival slot. Most of the traffic enters the switch at input port 1, as the arrival rate to that port is $\gamma_1 = 0.59$ cells per arrival slot.

For load balancing purposes, we have allocated one of the C channels exclusively to output port 1, since this port receives a considerable fraction of the total traffic. The remaining $C - 1$ channels are shared by the other 15 output ports. The allocation of the output ports to the remaining frequencies was performed in a round-robin fashion, and is given in Table 3 for $C = 4, 6, 8$. The quantities a_{ic} of the schedule, i.e., the number of cells to be transmitted by input port i onto channel λ_c per frame (refer to Section 2.2 and Figure 2) were fixed to be as close to 0.5 arrival slots as possible. Recall that, while the length of an arrival slot is independent of C and is taken as our unit of time, the length of a service slot depends on the number of channels. In cases in which this value was not an integral number of service slots, the value a_{ic} was rounded up to ensure that every queue was granted at least 0.5 arrival slots of service during each frame [3] (i.e., $a_{ic} = \lceil \frac{N}{2C} \rceil \;\; \forall \; i, c$). In constructing the schedules, we have assumed that the time it takes a laser to

[3] Other schemes for allocating a_{ic} have been implemented, including setting a_{ic} proportional to r_{ic}, setting a_{ic} proportional to $\max_z \alpha_i^{(z)}$, and setting a_{ic} to the effective bandwidth [41] of port i's total traffic carried on channel

Table 3: Channel sharing for $C = 4, 6, 8$

	$C = 4$	$C = 6$	$C = 8$
$\mathcal{R}_1$	$\{1\}$	$\{1\}$	$\{1\}$
$\mathcal{R}_2$	$\{2, 5, 8, 11, 14\}$	$\{2, 7, 12\}$	$\{2, 9, 16\}$
$\mathcal{R}_3$	$\{3, 6, 9, 12, 15\}$	$\{3, 8, 13\}$	$\{3, 10\}$
$\mathcal{R}_4$	$\{4, 7, 10, 13, 16\}$	$\{4, 9, 14\}$	$\{4, 11\}$
$\mathcal{R}_5$		$\{5, 10, 15\}$	$\{5, 12\}$
$\mathcal{R}_6$		$\{6, 11, 16\}$	$\{6, 13\}$
$\mathcal{R}_7$			$\{7, 14\}$
$\mathcal{R}_8$			$\{8, 15\}$

tune from one channel to another is equal to one arrival slot [4]. Finally, for all the results we present in this section we have let all input and output queues have the same buffer capacity B (i.e., $B_{ic}^{(in)} = B_j^{(out)} = B$) so that we can vary a single parameter.

Figures 8, 9, 10, and 11 show the cell loss probability (CLP), obtained through the analysis, at the input queues as a function of the buffer size B for $C = 4, 6, 8$. The simulation results and confidence intervals are not given as they do not provide a meaningful comparison for the scale given. The difference between analysis and simulation later in this section. We only present two cases, namely, the highest intensity port (input port 1) in Figures 8 and 10 and a representative intermediate port (input port 8) in Figures 9 and 11. We also consider only queues 1 and 2 (out of C) at each port. Queue 1 at each input port is for traffic to be carried on wavelength λ_1, which is dedicated to output port 1 (the hot spot). Thus, the amount of traffic received by this queue does not change as we vary the number of channels. Queue 2 at each input port is for traffic to be carried on wavelength λ_2. The amount of traffic received by this queue will decrease as the number of channels increases, since channel λ_2 will be shared by fewer output ports. The behavior of queue 2 is representative of the behavior of the other $C - 2$ queues, 3 through C. The portions of the schedules corresponding to λ_1 and λ_2 for $C = 4, 6$, and 8 are shown in Figure 7.

In Figure 8, the $\Omega_{1,1}$ decreases predictably as the buffer size, B, increases. The cell loss across frequencies, however, changes due to the changing ratio of arrival slots per frame to service slots for this queue per frame. In the case in which $C = 4$, $M = 8$ and $a_{1,1} = 2$, therefore, as many as 8 cells may arrive during a frame while 2 will be serviced. For $C = 6$, $M = 12$ and $a_{1,1} = 2$, indicating an increase in the offered traffic (which is fixed per slot) between service periods. Similarly, for $C = 8$, $M = 12$ and $a_{1,1} = 1$, and the decrease in available service per frame for this queue increases the CLP. Similar behavior is observed in Figure 9, where

λ_c. Although the cell loss probability results do depend on how a_{ic} is determined, the overall conclusions drawn regarding our analysis are very similar. Thus, we have decided to include only the simplest case.

[4] Again, due to the synchronous nature of this switch, if one arrival slot is not an integral number of service slots, the number of service slots for which a transmitter cannot transmit is rounded up, thereby setting the required time for tuning to some value slightly greater than one arrival slot. In fact, the tuning time will be $\lceil \frac{N}{C} \rceil$ service slots.

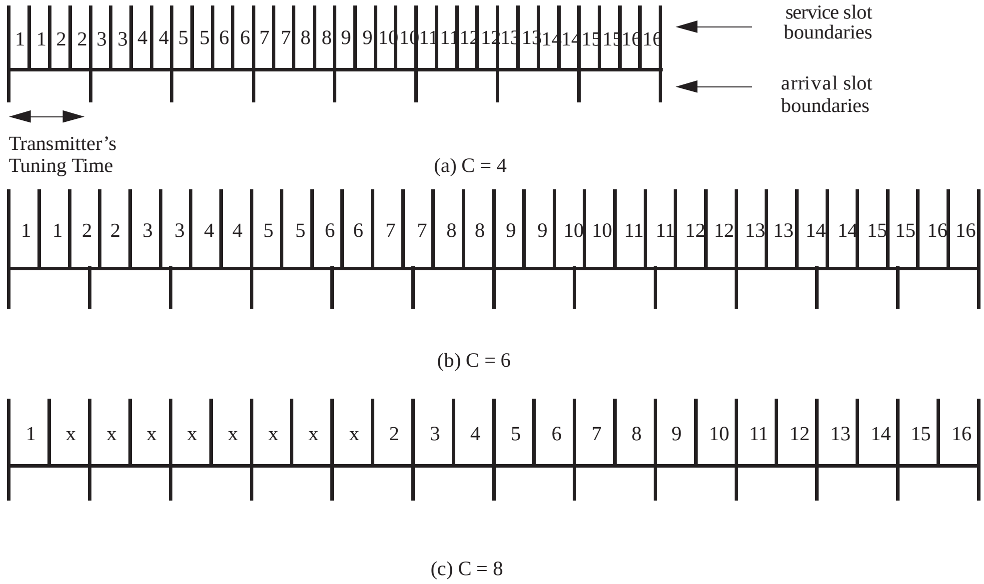

Figure 7: Transmission schedules for λ_1 and λ_2 for $C = 4, 6, 8$

we plot $\Omega_{8,1}$. We were not able to establish a reliable comparison between the analytic and simulated results for queue 1 of input port 8: the cell loss calculated analytically for queue 1 of input port 8 was 6×10^{-6}, whereas the simulation, after 30 replications of $100,000$ service slots, showed no loss.

It is obvious that queue 2 at input port 1 will be subject to not only the interactions discussed above, but also the effects of the changing traffic patterns due to the reallocation of output ports with the addition of new channels (see Table 3). In Figure 10, the traffic allocated to this queue decreases from 30% of that offered to input port 1 to 18% as C increases from 4 to 6. This change in the traffic arrival rate (from a mean normalized offered load of 0.175 to 0.105 per arrival slot) is enough to offset the change in the service rate. The CLP values for $C = 6$ are less than that for $C = 4$ in Figure 11 due to the fact that the change in the mean normalized offered load (from 0.121 to 0.0724 per arrival slot) is not substantial enough to offset the change in the service rate. As C increases to 8, however, there is no change in the offered traffic for either queue; as expected, the CLP rises with the decrease in the normalized service rate. The maximum absolute error observed between the anlaytic results and the mean simulated results for queue 2 of input port 1 was 17.6×10^{-3}; the loss for input queue 2 of port 8 could not be estimated for the same reason as that for queue 1 of input port 8 [5].

The CLP at output port 1 is similarly affected by the schedule parameters (see Figure 12). In the case where $C = 4$, we have $M = 8$ and $\sum_x u_1(x) = 32$, meaning that as many as 32 arrivals may be presented to this output port during a frame while only 8 may be serviced. $M = 12$ for $C = 6$ and $C = 8$ with $\sum_x u_1(x) = 32$ and $\sum_x u_1(x) = 16$, respectively. The dramatic change in $\sum_x u_1(x)$ is due to two major

[5] The maximum loss calculated for queue 2 of input port 8 was 7×10^{-6}.

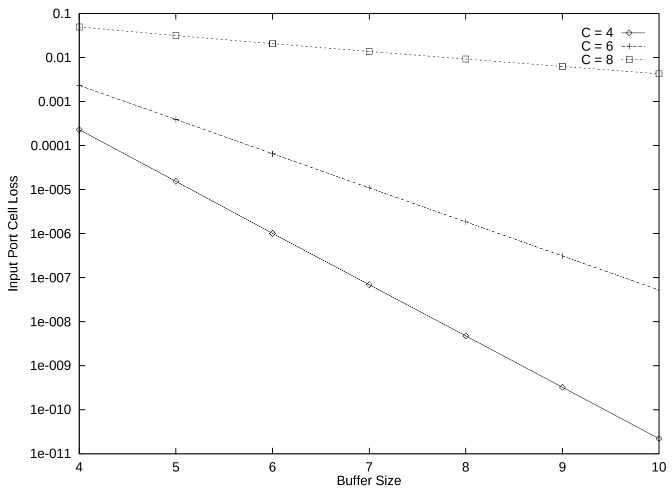

Figure 8: Input port cell loss probability $\Omega_{1,1}$ for $C = 4, 6, 8$ as a function of buffer size B

changes in the schedule's structure. For $C = 8$, the tuning time for a transmitter changes from 4 service slots to 2 service slots and the values of a_{ic} change from 2 service slots to 1 service slot. The effect of these parameter changes on the schedule is that a significant amount of the bandwidth in the 8 channel case is unused due to transmitters having to tune to the many frequencies. The reader is directed to [42] for a complete discussion of this phenomena. The CLP for output port 8 behaves similarly in Figure 13. The maximum absolute error between the simulated and analytic loss values for output ports 1 and 8 were 1.45×10^{-4} and 1.6×10^{-3}, respectively.

Finally, in order to validate the accuracy of the approximation algorithm, we ran several experiments involving different switch configurations. The switches varied from four ports and two frequencies to 16 ports with 10 frequencies. We observed that the smallest relative error[6] for cell loss occurred at an output port which was allocated to a dedicated frequency. The relative error observed was approximately 1×10^{-3}. The relative error increased as the number of ports sharing a single frequency increased. The worst relative error observed was 5×10^{-2}. We observed a similar behavior for the cell loss at the input ports.

6 Concluding Remarks

We have developed a decomposition algorithm to obtain the queue-length distributions at the input and output ports of a photonic ATM switch architecture based on the single-hop WDM architecture. We have also obtained analytic expressions for the cell-loss probability at the input and output ports. Finally, we have presented numerical results which illustrate the complex interactions between crucial system parameters and explained non-intuitive results.

[6] Relative error is taken here to mean the ratio of the absolute difference between the mean simulated value and the value calculated by our algorithm divided by the algorithm's value.

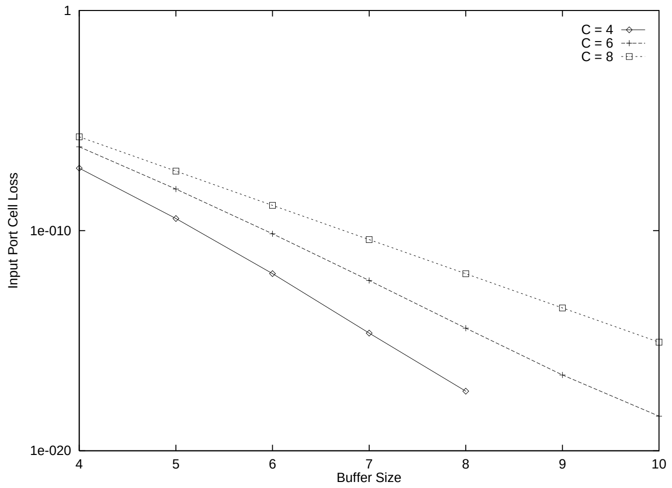

Figure 9: Input port cell loss probability $\Omega_{8,1}$ for $C = 4, 6, 8$ as a function of buffer size B

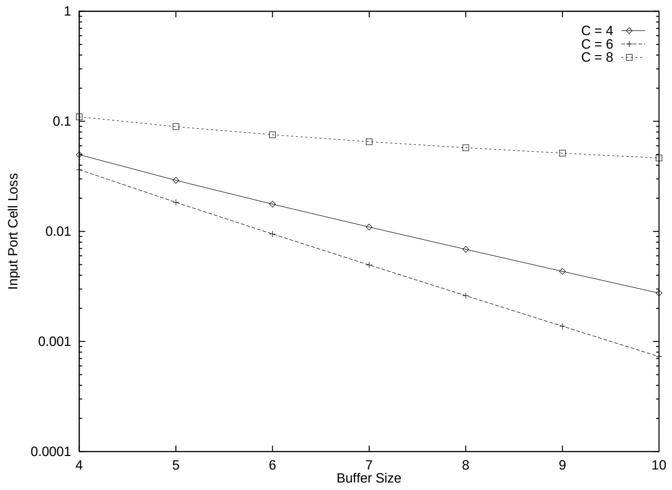

Figure 10: Input port cell loss probability $\Omega_{1,2}$ for $C = 4, 6, 8$ as a function of buffer size B

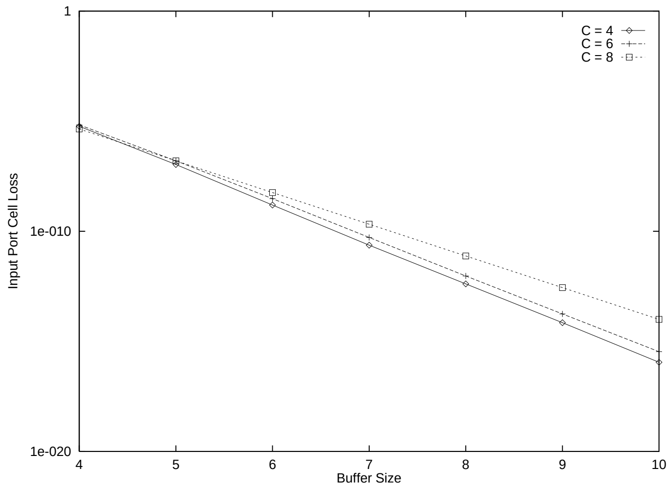

Figure 11: Input port cell loss probability $\Omega_{8,2}$ for $C = 4, 6, 8$ as a function of buffer size B

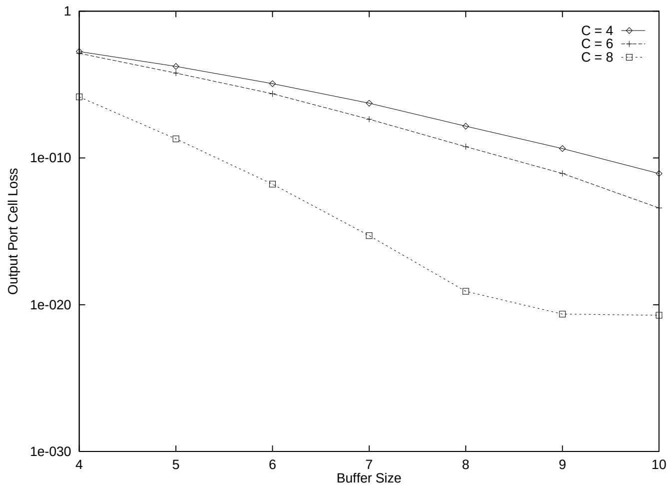

Figure 12: Output port cell loss probability Ω_1 for $C = 4, 6, 8$ as a function of buffer size B

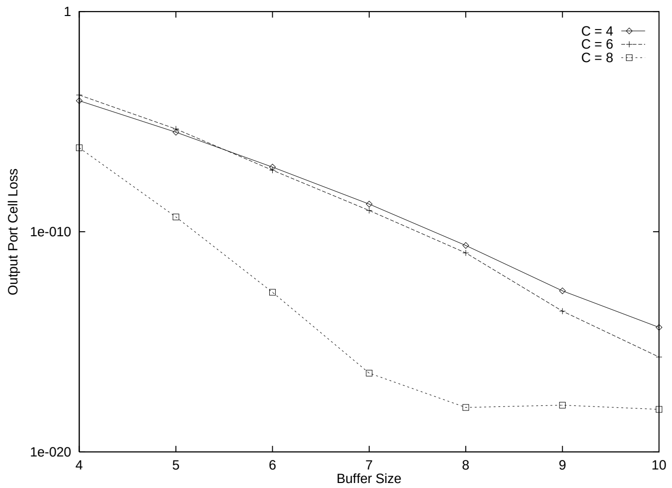

Figure 13: Output port cell loss probability Ω_8 for $C = 4, 6, 8$ as a function of buffer size B

References

[1] F. A. Tobagi. Fast packet switch architectures for broadband integrated services digital networks. *Proceedings of IEEE*, pages 133–167, January 1990.

[2] H. Ahmadi and W. E. Denzel. A survey of modern high-performance switching techniques. *IEEE JSAC*, pages 1227–1237, September 1989.

[3] P. E. Green. *Fiber Optic Networks*. Prentice-Hall, Englewood Cliffs, New Jersey, 1993.

[4] A. Sneh and K. M. Johnson. High-speed tunable liquid crystal optical filter for WDM systems. In *Proceedings of IEEE/LEOS Summer Topical Meetings on Optical Networks and their Enabling Technologies*, pages 59–60, Lake Tahoe, CA, July 1994.

[5] K. Nakagawa, S. Nishi, K. Aida, and E. Yoneda. Trunk and distribution network application of erbium-doped fiber amplifier. *Journal of Lightwave Technology*, LT-9, February 1991.

[6] L. Thylen, G. Karlsson, and O. Nilsson. Switching technologies for future guided wave optical networks: Potentials and limitations of photonics and electronics. *IEEE Communications Magazine*, 34(2):106–113, February 1996.

[7] A. Jaszczyk and H. T. Mouftah. Photonic fast packet switching. *IEEE Communications Magazine*, pages 58–65, February 1993.

[8] G. N. Rouskas and M. H. Ammar. Analysis and optimization of transmission schedules for single-hop WDM networks. *IEEE/ACM Transactions on Networking*, 3(2):211–221, April 1995.

[9] G. N. Rouskas and V. Sivaraman. On the design of optimal TDM schedules for broadcast WDM networks with arbitrary transceiver tuning latencies. In *Proceedings of INFOCOM '96*, pages 1217–1224. IEEE, March 1996.

[10] D. A. Levine and I. F. Akyildiz. PROTON: A media access control protocol for optical networks with star topology. *IEEE/ACM Transactions on Networking*, 3(2):158–168, April 1995.

[11] T. T. Lee, M. S. Goodman, and E. Arthurs. A broadband optical multicast switch. In *Proceedings of ISS '90*, 1990.

[12] K. Y. Eng. A photonic knockout switch for high-speed packet networks. *IEEE Journal Selected Areas in Communications*, pages 1107–1116, August 1988.

[13] E. Arthurs, M. S. Goodman, H. Kobrinski, and M. P. Vecchi. HYPASS: An optoelectronic hybrid packet switching system. *IEEE Journal on Selected Areas in Communications*, 6:1500–1510, December 1988.

[14] E. Arthurs, J. M. Cooper, M. S. Goodman, H. Kobrinski, M. Tur, and M. P. Vecchi. Multiwavelength optical crossconnect for parallel-processing computers. *Electronic Letters*, 24:119–120, 1988.

[15] N. R. Dono, P.E. Green Jr., K. Liu, R. Ramaswami, and F. F. Tong. A wavelength division multiple access network for computer communication. *IEEE Journal on Selected Areas in Communications*, 8(6):983–994, August 1990.

[16] M. S. Goodman, H. Kobrinski, M. P. Vecchi, R. M. Bulley, and J. L. Gimlett. The LAMBDANET multiwavelength network: Architecture, applications, and demonstrations. *IEEE Journal on Selected Areas in Communications*, 8(6):995–1004, August 1990.

[17] K. Bogineni, K. M. Sivalingam, and P. W. Dowd. Low-complexity multiple access protocols for wavelength-division multiplexed photonic networks. *IEEE Journal on Selected Areas in Communications*, 11(4):590–604, May 1993.

[18] P. A. Humblet, R. Ramaswami, and K. N. Sivarajan. An efficient communication protocol for high-speed packet-switched multichannel networks. *IEEE Journal on Selected Areas in Communications*, 11(4):568–578, May 1993.

[19] I. M. I. Habbab, M. Kavehrad, and C.-E. W. Sundberg. Protocols for very high-speed optical fiber local area networks using a passive star topology. *Journal of Lightwave Technology*, LT-5(12):1782–1793, December 1987.

[20] N. Mehravari. Performance and protocol improvements for very high-speed optical fiber local area networks using a passive star topology. *Journal of Lightwave Technology*, 8(4):520–530, April 1990.

[21] I. Chlamtac and A. Ganz. Channel allocation protocols in frequency-time controlled high speed networks. *IEEE Transactions on Communications*, 36(4):430–440, April 1988.

[22] P. W. Dowd. Random access protocols for high speed interprocessor communication based on an optical passive star topology. *Journal of Lightwave Technology*, LT-9:799–808, June 1991.

[23] A. Ganz. End-to-end protocols for WDM star networks. In *IFIP/WG6.1-WG6.4 Workshop on Protocols for High-Speed Networks*, pages 219–235, May 1989.

[24] A. Ganz and Z. Koren. WDM passive star - protocols and performance analysis. In *Proceedings of INFOCOM '91*, pages 991–1000. IEEE, April 1991.

[25] S. Tridandapani, J. S. Meditch, and A. K. Somani. The MaTPi protocol: Masking tuning times through pipelining in WDM optical networks. In *Proceedings of INFOCOM '94*, pages 1528–1535. IEEE, June 1994.

[26] Mon-Song Chen, N. R. Dono, and R. Ramaswami. A media-access protocol for packet-switched wavelength division multiaccess metropolitan area networks. *IEEE Journal on Selected Areas in Communications*, 8(6):1048–1057, August 1990.

[27] R. Chipalkatti, Z. Zhang, and A. S. Acampora. Protocols for optical star-coupler network using WDM: Performance and complexity study. *IEEE Journal on Selected Areas in Communications*, 11(4):579–589, May 1993.

[28] E. M. Foo and T. G. Robertazzi. A distributed global queue transmission strategy for a WDM optical fiber network. In *Proceedings of INFOCOM '95*, pages 154–161. IEEE, April 1995.

[29] A. Muir and J. J. Garcia-Luna-Aceves. Distributed queue packet scheduling algorithms for WDM-based networks. In *Proceedings of INFOCOM '96*, pages 938–945. IEEE, March 1996.

[30] G. R. Pieris and G. H. Sasaki. Scheduling transmissions in WDM broadcast-and-select networks. *IEEE/ACM Transactions on Networking*, 2(2):105–110, April 1994.

[31] M. Azizoglu, R. A. Barry, and A. Mokhtar. The effects of tuning time in banwidth-limited optical broadcast networks. In *Proceedings of INFOCOM '95*, pages 138–145. IEEE, April 1995.

[32] R. L. Cruz and J-T. Tsai. COD: Alternative architectures for high speed packet switching. *IEEE/ACM Transactions on Networking*, 4(1):11–21, February 1996.

[33] J. Jue, M. Borella, and B. Mukherjee. Performance analysis of the Rainbow WDM optical network prototype. *IEEE Journal Selected Areas in Communications*, 14(5):945–951, June 1996.

[34] V. Paxson and S. Floyd. Wide area traffic: The failure of poisson modeling. *IEEE/ACM Transactions on Networking*, 3(3):226–244, June 1995.

[35] G. Pujolle and H. G. Perros. Queueing systems for modelling ATM networks. In *Int'l Conf. on the Performance of Distributed Systems and Integrated Comm. Networks*, pages 10–12, Kyoto, Japan, September 1991.

[36] B. Mukherjee. WDM-Based local lightwave networks Part I: Single-hop systems. *IEEE Network Magazine*, pages 12–27, May 1992.

[37] I. Baldine and G. N. Rouskas. Reconfiguration in rapidly tunable transmitter, slowly tunable receiver single-hop wdm networks. Technical Report TR-96-10, North Carolina State University, Raleigh, NC, 1996. (Submitted to Infocom '97.).

[38] D. Park, H. G. Perros, and H. Yamashita. Approximate analysis of discrete-time tandem queueing networks with bursty and correleated input traffic and customer loss. *Operations Research Letters*, 15:95–104, 1994.

[39] A. O. Zaghloul and H. G. Perros. Approximate analysis of a discrete-time polling system with bursty arrivals. *Modelling and Performance Evaluation of ATM Technology (Perros, Pujolle, Takahashi Eds.)*, 1993.

[40] W. Stewart. *Numerical Solutions of Markov Chains.* Princeton University Press, Princeton, New Jersey, 1994.

[41] H. G. Perros and K. M. Elsayed. Call admission control schemes: A review. *IEEE Communications Magazine*, 34:82–91, 1996.

[42] G. N. Rouskas and V. Sivaraman. Packet scheduling in broadcast WDM networks with arbitrary transceiver tuning latencies. *IEEE/ACM Transactions on Networking*, ??(??):??–??, ?? ??