Wavelength Reuse for Efficient Packet–Switched Transport in an AWG–Based Metro WDM Network

Michael Scheutzow Martin Maier Martin Reisslein Adam Wolisz

Abstract

Metro WDM networks play an important role in the emerging Internet hierarchy; they interconnect the backbone WDM networks and the local access networks. The current circuit–switched SONET/SDH–over–WDM–ring metro networks are expected to become a serious bottleneck — the so–called metro gap — as they are faced with an increasing amount of bursty packet data traffic and quickly increasing bandwidths in the backbone networks and access networks. Innovative metro WDM networks that are highly efficient and able to handle variable–size packets are needed to alleviate the metro gap. In this paper we study an AWG–based single–hop WDM metro network. We analyze the photonic switching of variable–size packets with spatial wavelength reuse. We derive computationally efficient and accurate expressions for the network throughput and delay. Our extensive numerical investigations — based on our analytical results and simulations — reveal that spatial wavelength reuse is crucial for efficient photonic packet switching. In typical scenarios, spatial wavelength reuse increases the throughput by 60% while reducing the delay by 40%.

Keywords

Arrayed–Waveguide Grating; Medium Access Control; Metro WDM Network; Multiple Free Spectral Ranges; Photonic Packet Switching; Spatial Wavelength Reuse.

I. Introduction

THE INTERNET of the future may be viewed as a three–level hierarchy consisting of backbone networks, metro networks, and access networks. Backbone networks will provide almost infinite bandwidth based on Wavelength Division Multiplexing (WDM) links. These WDM links are connected with reconfigurable all–Optical Add–Drop Multiplexers (OADMs) and all–Optical Cross Connects (OXCs) controlled by Multiprotocol Lambda Switching (MPλS) [1], Optical Burst Switching (OBS) [2], and Optical Packet Switching (OPS) [3], [4], [5] mechanisms. Access networks transport data to (and from) individual users. By employing advanced LAN technologies, such as Gigabit Ethernet, broadband access, such as xDSL and cable modems, as well as high–speed next–generation wireless systems, such as UMTS, access networks provide an ever increasing amount of bandwidth. Metro networks interconnect the high–speed WDM backbone networks and the high–speed access networks. Current metro networks are typically SONET/SDH–over–WDM rings which are based on circuit–switching and carry the ever increasing amount of bursty data traffic only inefficiently. In addition, content providers increasingly place proxy caches in metro networks. These proxies further increase the load on metro networks. Metro networks are therefore expected to become a serious bottleneck — the so–called metro gap — in the future Internet. For these reasons there is an urgent need for innovative metro network architectures and protocols [6].

This work was supported in part by the Federal German Ministry of Education and Research within the TransiNet project.

A shorter version of this paper appears in *Proc. of IEEE Infocom*, New York, NY, June 2002.

Please direct correspondence to M. Reisslein.

M. Scheutzow is with the Dept. of Mathematics, Technical University Berlin, (email: ms@math.tu-berlin.de).

M. Maier and A. Wolisz are with the Telecommunication Networks Group, Technical University Berlin, 10587 Berlin, Germany, (email: maier@ee.tu-berlin.de, wolisz@ee.tu-berlin.de).

M. Reisslein is with the Telecommunications Research Center, Dept. of Electrical Engineering, Arizona State University, Goldwater Center, MC7206, Tempe AZ 85287–7206, Phone: (480)965–8593, Fax: (480)965–8325, (email: reisslein@asu.edu, web: http://www.eas.asu.edu/~mre).

Two key requirements for metro networks are (1) flexibility, and (2) efficiency. Flexibility is required since metro networks have to support a wide range of heterogeneous protocols, such as ATM, Frame Relay, SONET/SDH, and IP. This requires, in particular, that the metro networks are able to transport packets of *different sizes*. Efficiency is required because metro networks are highly cost–sensitive. Therefore, the deployed WDM networking components and the WDM networking resources (in particular wavelengths) must be utilized efficiently. As we demonstrate in this paper, a crucial technique for achieving high efficiency is *spatial wavelength reuse*. By spatial wavelength reuse we mean that in our Arrayed–Waveguide Grating (AWG) based metro WDM network (outlined in Section II) all wavelengths are used at all AWG ports simultaneously.

This paper builds on earlier work [7], in which we have proposed a novel AWG–based single–hop WDM network, which provides a dramatically increased degree of concurrency by (i) using multiple Free Spectral Ranges (FSRs), (ii) spatially reusing all wavelengths at each AWG port, and (iii) exploiting spreading techniques to enable simultaneous transmission of control and data. This earlier work focused primarily on the network architecture and the Medium Access Control (MAC) protocol. The elementary analysis conducted in [7] provided very limited insights into the performance of the proposed network. The performance analysis in [7] is limited in that it considered only fixed–size packets and did not consider spatial wavelength reuse. However, the efficient transmission of variable–size packets is of paramount importance for future metro networks. In this paper we study (1) the photonic switching of packets of different sizes, and (2) the spatial wavelength reuse in the AWG–based network proposed in earlier work. The main contribution of this paper is to develop a stochastic model to evaluate the performance of photonic packet switching with spatial wavelength reuse in the AWG–based network. Our analytical model gives computationally efficient and accurate expressions for the throughput and delay in the network. Our numerical results indicate that spatial wavelength reuse is crucial for efficient photonic packet switching. For typical scenarios, spatial wavelength reuse increases the throughput by 60% while reducing the delay by 40%.

This paper is organized as follows. In the following subsection we give a quick overview of related work. In Section II we briefly review the architecture of the studied AWG–based metro network as well as the reasoning for selecting this architecture. In Section III we briefly review the MAC protocol for the studied network. In Section IV we develop a stochastic model for the performance evaluation of the transmission of variable–size packets with spatial wavelength reuse. This model and performance evaluation are our main contributions. In Section V we use our analytical results to conduct numerical investigations. We also conduct simulations to verify the accuracy of our analytical results. We conclude in Section VII.

A. Related Work

Single–hop metro WDM networks based on the Passive Star Coupler (PSC) have been studied extensively, see for instance [8], [9], [10], [11], [12], [13], [14], [15], [16], [17], [18] as well as the surveys [19], [20]. The studied networks are so–called *broadcast–and–select* networks. Each transmission is broadcast over the PSC to all nodes. The indented recipient processes the transmission, whereas the other nodes ignore the transmission. The primary limitation of the PSC–based networks is that each wavelength provides only one communication channel between any pair of network nodes at any point in time. Wavelengths, however, are scarce, especially with the low–cost coarse WDM technology that is used in cost–sensitive metro networks. By partitioning the network into several PSC–based clusters and interconnecting the clusters in some fashion, the same set of wavelengths can be reused at each cluster PSC [21], [22], [23].

The network studied in this paper is fundamentally different from the PSC–based networks in that it exploits the wavelength routing property of the AWG. In the studied AWG–based network, transmissions are not broadcast. Instead, they are selectively forwarded by using the appropriate wavelength, as discussed in detail below. This results in a *switched* network as opposed to the aforementioned broadcast networks. In addition, in our network, wavelengths are spatially reused at the different AWG input ports, which allows for a high degree of concurrency and efficient use of the scarce wavelength resources.

Apart from a few older studies, such as [8], [24], [25], metro WDM networks have just recently begun to attract the interest from the research community [26]. A metro network based on optical add–drop multiplexers (OADMs) has recently been studied in [27]. This network is geared towards optical *circuit* switching. In [28] a ring network employing dynamic wavelength add–drop multiplexers is studied. Slotted packet–switched ring networks are studied in [29], [30]. In [31] a ring WDM network with AWG–based OADMs has been proposed whose frequency–cyclic nature can be used for easily upgrading the network capacity. For interconnecting such WDM rings in an efficient and cost–effective way the AWG was used as a passive wavelength routing hub in [32]. General design principles and architectures for networks based on AWGs are studied in [33], [34], [35], [36], [37], [38]. On–line scheduling algorithms for an AWG–based network with static wavelength assignment and static time–division multiple access, which is fundamentally different from the on–demand reservations in our network, are studied in [39]. A protection routing strategy employing AWGs is developed in [40] and a packet switch based on an AWG is studied in [41].

The HORNET metro network [42], [43] allows for optical *packet* switching. HORNET [42], [43] and the metro networks studied in [27], [28], [29], [30], [31] have a ring topology, i.e., transmissions typically have to traverse multiple network nodes. These networks are therefore fundamentally different from the packet–switched single–hop network studied here. We note that AWG–based single–hop WDM networks are also studied in [44], [45], [46]. However, these networks have either a more complex hub structure with wavelength converters and a centralized resource management [44], [45] or require a large number of transceivers at each node [46]. In contrast, our network is completely *passive* and resources are allocated in a *distributed* fashion (as described in the subsequent sections). Moreover, *multiple* free spectral ranges (FSRs) of the underlying AWG are used to increase the degree of concurrency and thereby to improve the throughput–delay performance significantly. Each node is equipped with one *single* tunable transceiver and an off–the–shelf broadband light source (e.g., light emitting diode (LED)). Finally, the novel node architecture of the proposed metro network allows for simultaneous transmission of data and control *without* requiring a separate control channel or an additional receiver at each node. Therefore, all wavelengths can be used for data transmission.

II. Architecture

In this section we outline the architecture of the studied AWG–based network. First, we briefly review the physical properties of the Arrayed–Waveguide Grating (AWG). The AWG [47], [48], [49] is a passive wavelength routing device with a wide range of applications [50], [51]. In our network we use the AWG as wavelength router [52]. We note that the recent development of AWGs with a crosstalk as low as -40 dB makes it possible to spatially reuse all wavelengths at each AWG input port simultaneously, which in turn increases the capacity of AWG–based networks significantly. Moreover, by deploying *athermal* AWGs which do not require any temperature control the costs and management of AWG–based networks can be reduced significantly [53].

The wavelength routing properties of the AWG are illustrated in Figure 1, where six wavelengths (from laser diodes (LDs)) are sent into each input port of an AWG with degree $D = 2$. The AWG routes every second wavelength to the same output port. This period of the wavelength routing is referred to as *Free Spectral Range (FSR)*. In the depicted example we exploit $R = 3$ FSRs, each consisting of $D = 2$ wavelengths. Note that each FSR provides one wavelength channel for communication between a given AWG input port and a given AWG output port. In addition to the LD wavelengths, a broadband signal (e.g., from an LED) which spans all six wavelength channels is sent into each AWG input port. As illustrated in Figure 1, the AWG slices the broadband signal and routes the slices to the respective output ports. Thus, spectral slicing can be used to broadcast low–speed information, e.g., control information, from a given input port to all output ports.

The considered metro WDM network is based on a $D \times D$ AWG, as shown in Fig. 2. At each AWG input port a wavelength–insensitive $S \times 1$ combiner collects data from S attached nodes. Similarly, at each AWG output port signals are distributed to S nodes by a wavelength–insensitive $1 \times S$ splitter (note that

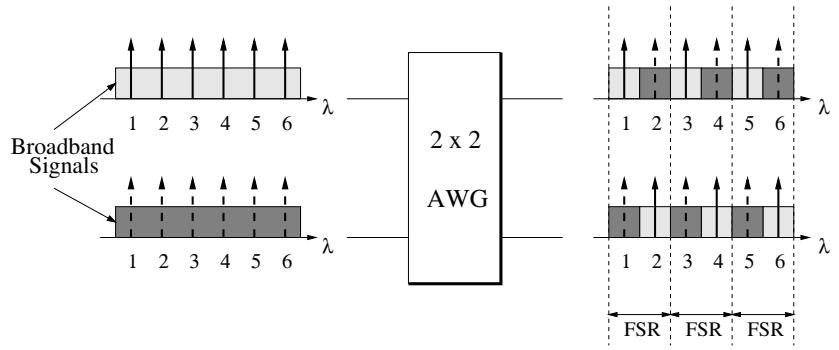

Fig. 1. Wavelength routing in AWG

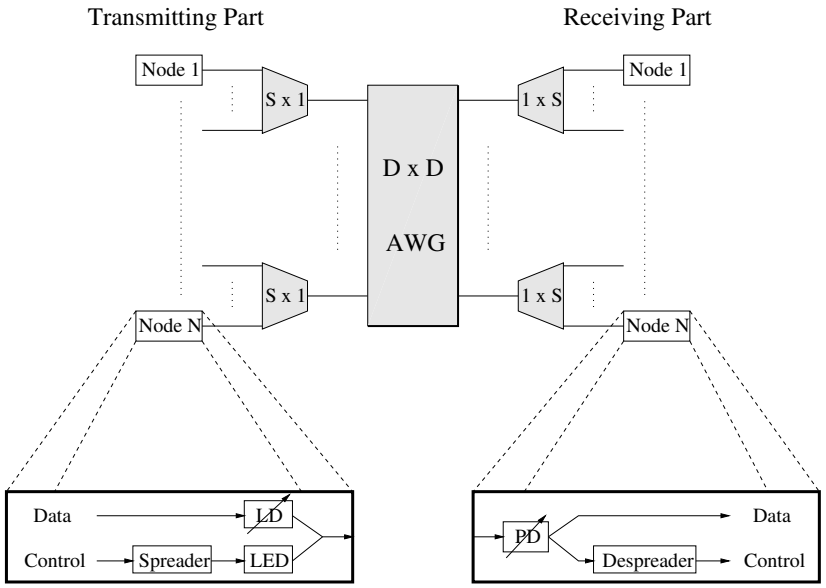

Fig. 2. Network and node architecture

these splitters can also be used for optical multicasting). Each node is composed of a transmitting part and a receiving part. The transmitting part of a node is attached to one of the combiner ports. The receiving part of the same node is located at the opposite splitter port. Thus, the network connects $N = D \cdot S$ nodes. Each node contains a tunable laser diode (LD) and a tunable photodiode (PD) for data transmission and reception, respectively. In addition, each node uses a broadband light source, e.g., LED, for broadcasting control packets by means of spectral slicing. The control information is spread in the electrical domain by means of direct sequence spread spectrum techniques [54], [55]. Both data and control signals are combined and then passed through the AWG based network, as illustrated in Fig. 3. At the output of the network, a PD is tuned to the same wavelength as the LD. The PD detects the LD wavelength and corresponding slice of the broadband signal and converts the combined optical signal into the electrical domain. The modulation speed and launch power of the broadband signal are such that the control signal has (i) a smaller bandwidth, and (ii) a smaller power level than the data signal [56], as illustrated in the lower right corner of Fig. 3. Thus, the control signal does not significantly distort the data signal. The data signal is received without requiring any further processing (except for possibly some highpass filtering). The control signal is retrieved by lowpass filtering the combined data and (spread) control signal. The filtered signal is despread by a decorrelator which multiplies the signal with the corresponding spreading sequence. The spreading of the control signal has the advantage that data and control are sent simultaneously (i.e., there is no time division between signalling and data transmission), resulting in increased bandwidth efficiency.

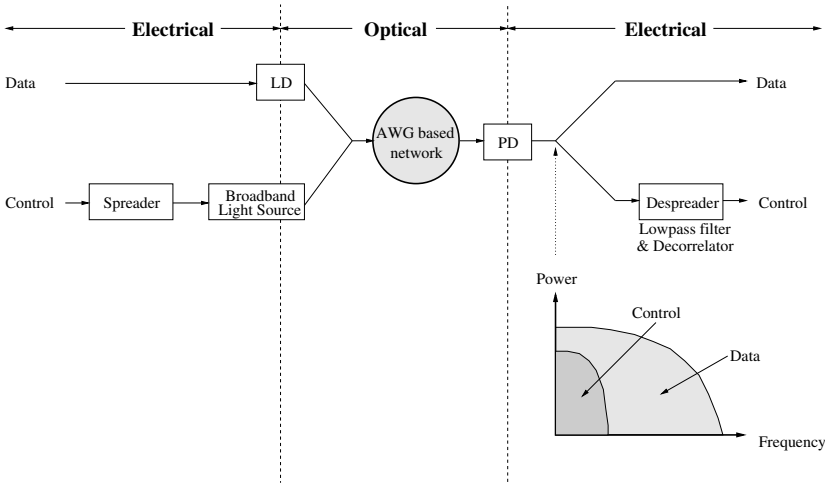

Fig. 3. Illustration of spectral spreading of control information

III. MAC PROTOCOL

In this section we give a brief overview of our MAC protocol; we refer the interested reader to [7] for more details. In our network each node has a tunable transmitter and a tunable receiver (TT–TR). This TT–TR structure allows for high flexibility in the data transmissions and receptions and has the potential to achieve load balancing, improved channel utilization and throughput-delay performance. However, with TT–TR nodes not only channel collisions but also receiver collisions may occur. Typically a MAC protocol is employed to arbitrate the access to the wavelengths and thus to mitigate collisions. Generally, MAC protocols for single-hop WDM networks fall into the three main categories of (i) preallocation protocols, (ii) random access protocols, and (iii) reservation (pretransmission coordination) protocols, comprehensively surveyed in [20]. (Since MAC protocols for AWG based networks have received little attention so far, the survey [20] studies the large body of literature on MAC protocols for PSC based networks. Some key learned lessons from this literature, however, are considered generally valid and guide the design of the MAC protocol for our AWG based network.) Preallocation protocols statically assign a wavelength to a node during a periodically recurring time slot. Preallocation generally gives high utilization only for uniform non-burst traffic and is thus poorly suited for the bursty traffic in future metro networks. Random access protocols do not require any preallocations to nodes or coordination among nodes. For medium to high traffic loads, however, collisions become very frequent resulting in small throughout and large delay. Reservation protocols employ pretransmission coordination (reservation signalling) to assign wavelengths and receivers on demand. With the so-called attempt-and-defer type of reservation protocol, data packets are only transmitted after a successful reservation. Thus, attempt-and-defer protocols completely avoid channel and receiver collisions. This approach is generally preferable in a TT–TR system with bursty traffic and we adopt it for our data packet transmissions. For the pretransmission coordination we transmit small control packets according to a random access protocol, namely a modified slotted Aloha protocol. This approach is adopted since (i) random access control packet transmission, as opposed to fixed assignments, makes the network scalable, and (ii) for the typical large propagation delay to control packet transmission delay ratio, slotted Aloha is superior to carrier sensing based access.

In our MAC protocol time is divided into *cycles*, as illustrated in Fig. 4. Each cycle consists of D *frames*. Each frame contains F *slots*, as illustrated in Fig. 5. The slot length is equal to the transmission time of a control packet. Each frame is partitioned into the first M, $1 \leq M < F$, slots and the remaining $(F - M)$ slots. In the first M slots (the shaded area in Figs. 4 and 5), control packet transmissions take place simultaneously with data packet transmissions which do not exploit spatial wavelength reuse. In order to avoid receiver collisions of control packets, the receivers at all nodes must be tuned (locked) to one of the

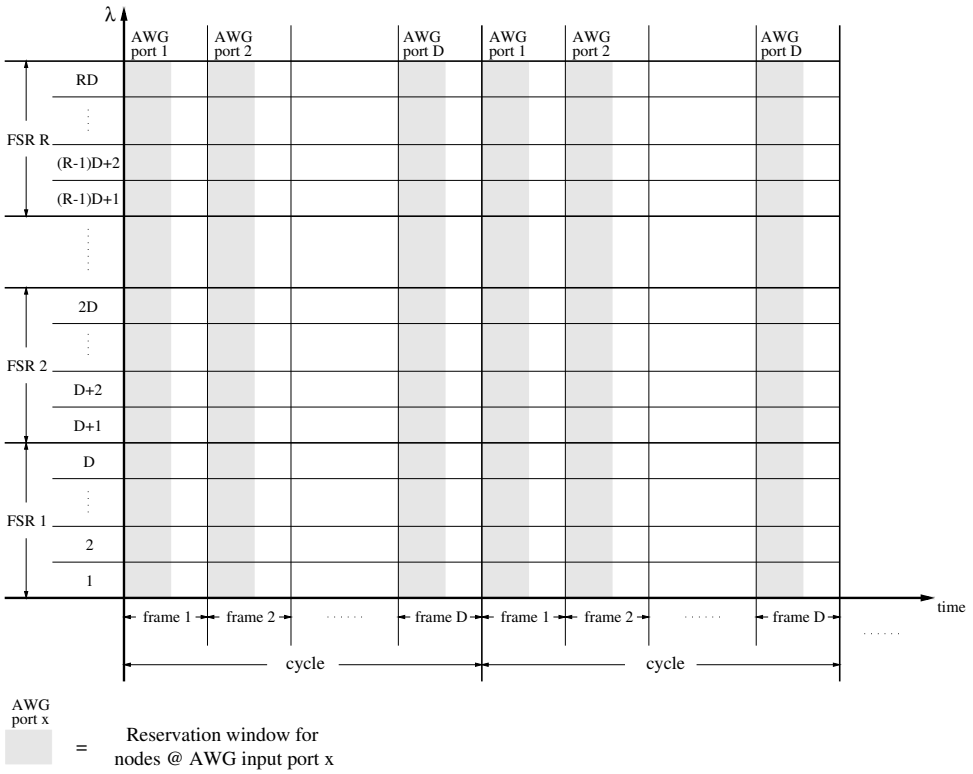

Fig. 4. Wavelength assignment and timing structure of MAC protocol at a given AWG input port

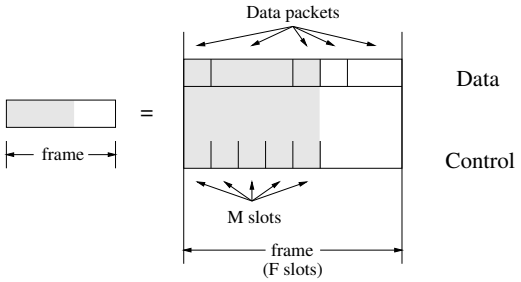

Fig. 5. Frame format

LED slices carrying the control information during these first M slots. Owing to the routing characteristics of the AWG only nodes attached to the same combiner can send control packets in a given frame. Nodes attached to different AWG input ports send their control packets in different frames. Specifically, all nodes attached to AWG input port o, $1 \leq o \leq D$, (via a common combiner) send their control packets in frame o of the cycle. Thus, nodes can transmit control packets only in one frame per cycle. During the first M slots of frame o control and data packets can be transmitted simultaneously by the nodes attached to AWG input port o. Due to the routing characteristics of the AWG transmissions from the other AWG input ports can not be received during this time interval. Consequently, the transmitters at the other ports can not send data during this time interval. Thus, wavelengths *can not* be spatially reused during this time interval. Therefore, data packets that are longer than $(F-M)$ slots can be transmitted only in those frames in which the corresponding source node is allowed to send its control packets. In the last $(F-M)$ slots of each frame no control packets are sent. The receivers are unlocked, allowing transmission between *any* pair of nodes in *each* frame provided that the corresponding data packet is not longer than $(F-M)$ slots. This allows for spatial wavelength reuse — the main focus of this paper. Note that in general long data packets (length $> F-M$ slots) are harder to

schedule than short ones (length $\leq F - M$ slots).

When a data packet arrives to a node attached to AWG input port o, the node's LED broadcasts the corresponding control packet in one of the first M slots of the frame assigned to AWG input port o. The control packet has four fields (see [7]): Destination address (unicast/multicast), length, type (packet/circuit switched), and Forward Error Correction (FEC). The control packet is transmitted on a contention basis using a modified version of slotted Aloha. Note that using a derivative of slotted Aloha keeps the control packet contention simple, which is of paramount importance in very high–speed networks. Furthermore, by not assigning each node a dedicated reservation slot, new nodes can join the network without service disruptions. Every node (including the sending node) collects all control packets by locking its receiver to one of the LED slices carrying the control information during the first M slots of every frame. Thus, each node maintains global knowledge of all the other nodes' activities (and also learns whether its own control packet collided in the control packet contention or not). Note that this approach introduces only a one–way end–to–end propagation delay (i.e., half the round trip time). Whereas the conventional approach with control packets and explicit acknowledgements introduces the two–way end–to–end propagation delay (i.e., the full round trip time). Also, note that the control packet broadcast completely avoids receiver collisions of control packets and allows each node to learn immediately whether the control packet suffered a channel collision. If a control packet collides, it is retransmitted in the next cycle with probability p; with probability $(1 - p)$ the retransmission is deferred by one cycle. The successfully received control packets are processed by all nodes in a distributed fashion. Each node applies the same scheduling algorithm and thus comes to the same conclusion. The scheduling algorithm tries to schedule the data packets within the scheduling window of D frames (i.e., one cycle). If the scheduling fails, the source node retransmits the control packet. Given the very high–speed nature of optical networks and that each node has to process all other nodes' control packets, we employ a simple first–come–first–served and first–fit scheduling policy to avoid a computational bottleneck.

IV. Analysis

A. System and Traffic Model

In our analysis we consider a system with a large S. Our analysis is approximate for finite S and exact in the asymptotic limit $S \to \infty$. Throughout our analysis we assume that the propagation delay is no larger than one cycle (this is reasonable for a metropolitan area network). Thus, if a control packet is sent in a given frame, the corresponding data packet could be scheduled for transmission one cycle later. We assume that all nodes are equidistant from the AWG, i.e., the propagation delay is the same for all nodes (which could easily be achieved with low–loss fiber delay lines in real networks).

We assume that each node has a buffer that can hold a single data packet and a single control packet. We make the following assumptions about the traffic generation process. Suppose that a node's control packet has just been (1) successfully transmitted, and (2) the corresponding data packet has been successfully scheduled (within the scheduling window of one cycle; see Section IV-C). With probability σ this node then generates the control packet for the next data packet right before the beginning of the next frame in which the node can send the next control packet (i.e., one cycle after the previous control packet was sent). If no control packet is generated, then the node waits for one cycle and then generates a new control packet with probability σ, and so on. The node's buffer may hold the scheduled (but not yet transmitted) data packet and the next control packet at the same time. This next control packet is sent with probability one in the next frame assigned to the node's AWG input port (possibly simultaneously with the scheduled data packet). A data packet is purged from the node's buffer at the end of the frame during which it is transmitted. After a data packet is purged from the buffer, the next data packet is placed in the buffer, provided the corresponding control packet is already in the buffer.

If a control packet fails in the slotted Aloha contention or the data packet scheduling, then the node re–transmits the control packet in the next frame assigned to the node's AWG input port with probability p, with probability $(1-p)$ it defers the re–transmission by one cycle. In this next cycle the node transmits the control

packet with probability p and defers the transmission with probability $(1-p)$, and so on. We define

$$\tilde{\alpha} := \frac{S\sigma}{M} \text{ and } \alpha := \frac{Sp}{M}. \tag{1}$$

We conduct an approximate analysis for large S. Our analysis becomes asymptotically exact when $S \to \infty$ and $\tilde{\alpha}$ as well as α (and also M) are fixed (with σ and p chosen so as to satisfy (1)).

We assume uniform unicast traffic. A data packet is destined to any one of the N nodes (including the sending node, for simplicity) with equal probability $1/N$. Let L denote the length of a data packet in slots. A data packet is long (has size $L = F$) with probability q, i.e., $P(L = F) = q$. A data packet is short (has size $L = K,\ 1 \leq K \leq F - M$) with probability $(1-q)$, i.e., $P(L = K) = 1 - q$. Our model can be extended to more complex packet size distributions at the expense of introducing more notation and a more complex arbitration policy in Section IV-C.

If a control packet fails (either in the slotted Aloha or the scheduling) the size of the corresponding data packet is not changed. However, we do assume nonpersistency for the destination in our analysis; i.e., a new random destination is drawn for each attempt of transmitting a control packet. We note that our comparisons with simulations with persistent destinations (see Section V) clearly show that the analysis provides very accurate results despite the destination node nonpersistency assumption. We also note that our analysis can be extended to persistent destinations in a straightforward manner by maintaining the random variables introduced below for each AWG output port.

Now consider the nodes attached to a given (fixed) AWG input port $o,\ 1 \leq o \leq D$. These nodes send their control packets in frame o of a given cycle. We refer to the nodes that at the beginning of frame o hold an old packet, that is, a control packet that has failed in slotted Aloha or scheduling, as *"old"*. We refer to all the other nodes as *"new"*. Note that the set of "new" nodes comprises both the nodes that have generated a new (never before transmitted) control packet as well as the nodes that have deferred the generation of a new control packet. Let η be a random variable denoting the number of "new" nodes at AWG input port o, and let

$$\nu := \frac{E[\eta]}{S}. \tag{2}$$

Let λ_l be a random variable denoting the number of nodes at port o that are to send a control packet corresponding to a long data packet next (irrespective of whether a given node is "old" or "new", and keeping in mind that the set of "new" nodes also comprises those nodes that have deferred the generation of the next control packet; those nodes are accounted for in λ_l if the next generated control packet corresponds to a long data packet). Let λ_k be a random variable denoting the number of nodes at port o that are to send a control packet corresponding to a short data packet next. By definition, $\lambda_k = S - \lambda_l$. Let

$$\tilde{q} := \frac{E[\lambda_l]}{S} \tag{3}$$

denote the expected fraction of long packets to be sent. We expect that $\tilde{q}$ is typically larger than q since long packets are harder to schedule and thus typically require more re–transmissions (of control packets).

B. Analysis of Control Packet Contention

First, we calculate the number of control packets from nodes attached to AWG input port $o,\ 1 \leq o \leq D$, that are successful in the slotted Aloha contention in frame o. Let $Y_i^n,\ i = 1, \ldots, M$, be a random variable denoting the number of control packets that were randomly transmitted in slot $i,\ i = 1, \ldots, M$, by "new" nodes. Recall that each of the η "new" nodes sends a control packet with probability σ in the frame. Thus,

$$P(Y_i^n = k) = \binom{\eta}{k} \left(\frac{\sigma}{M}\right)^k \left(1 - \frac{\sigma}{M}\right)^{\eta - k}, \quad k = 0, 1, \ldots, \eta. \tag{4}$$

Throughout our analysis we assume that S is large and that $\tilde{\alpha}$ and α are fixed. We may therefore reasonably approximate the BIN$(\eta,\ \sigma/M)$ distribution with a Poisson $(\eta\sigma/M)$ distribution, that is,

$$P(Y_i^n = k) \approx e^{-\eta\sigma/M}\,\frac{(\eta\sigma/M)^k}{k!}, \quad k = 0, 1, \ldots, \tag{5}$$

which is exact for $\eta \to \infty$ with $\eta\sigma/M$ fixed. (A refined analysis that does not approximate the binomial distribution by the Poisson distribution is given in Appendix A.) We now recall the definition $\tilde{\alpha} := \sigma S/M$. We also approximate η/S by its expectation ν; this is reasonable since η/S has only small fluctuations in steady state for large S. Thus,

$$P(Y_i^n = k) \approx e^{-\tilde{\alpha}\nu}\,\frac{(\tilde{\alpha}\nu)^k}{k!}, \quad k = 0, 1, \ldots. \tag{6}$$

We note that for $S \to \infty$ the random variables $Y_1^n, Y_2^n, \ldots, Y_M^n$ are mutually independent. This is because a given node places with the miniscule probability σ/M a control packet in a given slot, say slot 1. (Note in particular that the expected value of Y_1^n is small compared to the number of "new" nodes, that is, $\tilde{\alpha}\nu \ll \eta$; this is because in the considered asymptotic limit $S \to \infty$ with α fixed, we have $1 \gg \sigma/M = \tilde{\alpha}\nu/\eta$.) Thus, Y_1^n has almost no impact on $Y_2^n, \ldots, Y_M^n$ (see Appendix B for a formal proof).

Let $Y_i^o,\ i = 1, \ldots, M$, be a random variable denoting the number of control packets in slot $i,\ i = 1, \ldots, M$, that originate from "old" nodes. Each of the $(S - \eta)$ "old" nodes sends a control packet with probability p in the frame. Thus,

$$P(Y_i^o = k) = \binom{S-\eta}{k}\left(\frac{p}{M}\right)^k\left(1 - \frac{p}{M}\right)^{S-\eta-k},\ k = 0, 1, \ldots, S - \eta. \tag{7}$$

Approximating this BIN$(S - \eta,\ p/M)$ distribution by the Poisson $((S - \eta)p/M)$ distribution we have

$$P(Y_i^o = k) \approx e^{-(S-\eta)p/M}\,\frac{[(S-\eta)p/M]^k}{k!},\ k = 0, 1, \ldots. \tag{8}$$

With $\alpha := pS/M$ and approximating $(S - \eta)/S$ by its expectation $(1 - \nu)$ we get

$$P(Y_i^o = k) \approx e^{-\alpha(1-\nu)}\,\frac{[\alpha(1-\nu)]^k}{k!},\ k = 0, 1, \ldots. \tag{9}$$

We note again that for $S \to \infty$ the random variables $Y_1^o, Y_2^o, \ldots, Y_M^o$ are mutually independent. They are also independent of $Y_1^n, Y_2^n, \ldots, Y_M^n$. Hence, we obtain for $Y_i = Y_i^n + Y_i^o$

$$P(Y_i = k) \approx e^{-[\tilde{\alpha}\nu+\alpha(1-\nu)]}\,\frac{[\tilde{\alpha}\nu + \alpha(1-\nu)]^k}{k!},\ k = 0, 1, \ldots. \tag{10}$$

Henceforth, we let for notational convenience

$$\beta := \tilde{\alpha}\nu + \alpha(1-\nu), \tag{11}$$

i.e.,

$$P(Y_i = k) \approx e^{-\beta}\,\frac{\beta^k}{k!},\ k = 0, 1, \ldots. \tag{12}$$

Let $X_i,\ i = 1, \ldots, M$, be a random variable indicating whether or not slot i contains a successful control packet. Specifically, let

$$X_i = \begin{cases} 1 & \text{if } Y_i = 1 \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

From (12) clearly, $P(X_i = 1) = \beta e^{-\beta}$ and $P(X_i = 0) = 1 - \beta e^{-\beta}$ for $i = 1, \ldots, M$. The total number of successful control packets in the considered frame is $\sum_{i=1}^{M} X_i$, which has a BIN$(M,\ \beta e^{-\beta})$ distribution, that is,

$$P\left(\sum_{i=1}^{M} X_i = l\right) = \binom{M}{l} \left(\beta e^{-\beta}\right)^l \left(1 - \beta e^{-\beta}\right)^{M-l},\ l = 0, 1, \ldots, M. \tag{14}$$

Recall from Section IV-A that each packet is destined to any one of the D AWG output ports with equal probability $1/D$. Let Z denote the number of successful control packets — in the considered frame — that are destined to a given (fixed) AWG output port $d,\ d = 1, \ldots, D$. Clearly, from (14)

$$P(Z = k) = \binom{M}{k} \left(\frac{\beta e^{-\beta}}{D}\right)^k \left(1 - \frac{\beta e^{-\beta}}{D}\right)^{M-k},\ k = 0, 1, \ldots, M. \tag{15}$$

(A refined approximation of $P(Z = k)$ which does not approximate the binomial distributions of the Y_i^n's and Y_i^o's by Poisson distributions is given in Appendix A.) Let Z_l denote the number of successful control packets that correspond to long data packets destined to a given AWG output port d. Recall that $\tilde{q}$ is the expected fraction of long packets to be sent. Hence, $Z_l \sim BIN(M,\ \beta e^{-\beta} \frac{1}{D} \tilde{q})$. Similarly, let Z_k denote the number of control packets that are successful in the slotted Aloha contention and correspond to short data packets destined to a given AWG output port d. Clearly, $Z_k \sim BIN(M,\ \beta e^{-\beta} \frac{1}{D}(1 - \tilde{q}))$.

C. Analysis of Packet Scheduling

In this section we calculate the expected number of packets that are successfully scheduled. Recall from the previous section that the total number of long packets that (1) originate from a given AWG input port $o,\ 1 \leq o \leq D$, (2) are successful in the slotted Aloha contention of frame o (of a given cycle), and (3) are destined to a given AWG output port $d,\ 1 \leq d \leq D$, is $Z_l \sim BIN(M,\ \beta e^{-\beta} \frac{1}{D} \tilde{q})$. For short packets we have $Z_k \sim BIN(M,\ \beta e^{-\beta} \frac{1}{D}(1 - \tilde{q}))$. Note that these two random variables are not independent. Let $\mathcal{L}$ $(\mathcal{S})$ be a random variable denoting the number of long (short) packets that (1) originate from a given AWG input port $o,\ 1 \leq o \leq D$, (2) are successful in the slotted Aloha contention of frame o (of a given cycle), (3) are destined to a given AWG output port $d,\ 1 \leq d \leq D$, and (4) are successfully scheduled within the scheduling window of D frames (i.e., one cycle).

Consider the scheduling of packets from a given (fixed) AWG input port o to a given (fixed) AWG output port d over the scheduling window (i.e., D frames). Clearly, we can schedule at most R long packets (i.e., $\mathcal{L} \leq R$) because the receivers at output port d must tune to the appropriate spectral slices during the first M slots of every frame. Thus they can tune to a node at AWG input port o for F consecutive slots, only in the frame, during which the nodes at AWG input port o send their control packets.

Now, suppose that $\mathcal{L}$ $(\leq R)$ long packets are scheduled (how $\mathcal{L}$ is determined is discussed shortly). With $\mathcal{L}$ long packets already scheduled, we can schedule at most

$$\mathcal{S} \leq (D - 1) \cdot R \cdot \left\lfloor \frac{F - M}{K} \right\rfloor + (R - \mathcal{L}) \left\lfloor \frac{F}{K} \right\rfloor \tag{16}$$

short packets. To see this, note that in the frame during which the nodes at AWG input port o send their control packets, there are $(R - \mathcal{L})$ FSRs — channels between AWG input port o and AWG output port d — free for a duration of F consecutive slots. Furthermore, there are $(D - 1)$ frames in the scheduling window during which the nodes at AWG output port d must tune (are locked) to the nodes sending control packets from the other AWG input ports for the first M slots of the frame. During each of these frames, the receivers are unlocked for $(F - M)$ slots. The R utilized FSRs provide R parallel channels between AWG input port o and AWG output port d. Note that the $(D - 1)R\lfloor (F - M)/K \rfloor$ component in (16) is due to the spatial reuse of wavelengths at the considered AWG input port. Without spatial wavelength reuse this component would

be zero and we could schedule at most $(R-\mathcal{L})\lfloor F/K \rfloor$ short packets. Continuing our analysis for a network with spatial wavelength reuse, we have

$$\mathcal{S} = \min\left\{ Z_k,\ (D-1)\cdot R\cdot \left\lfloor \frac{F-M}{K} \right\rfloor + (R-\mathcal{L})\left\lfloor \frac{F}{K} \right\rfloor \right\}. \tag{17}$$

In (17) we neglect receiver collisions, that is, we do not account for situations where a packet can not be scheduled because its receiver is already scheduled to receive a different packet. This assumption is reasonable as receiver collisions are rather unlikely for moderately large S. This is verified by our simulations which take receiver collisions into consideration, see Section V.

In this paper we consider a first–come–first–served–first–fit scheduling policy. Data packets are scheduled for the first possible slot(s) at the lowest available wavelength. To arbitrate the access to the frame which allows transmission for F contiguous slots and the $(D-1)$ frames which allow transmission for $(F-M)$ contiguous slots we adopt the following *arbitration policy*. Our arbitration policy proceeds in one round if there are R or less successful control packets in the slotted Aloha contention. In case there are more than R successful control packets in the slotted Aloha contention, our arbitration policy proceeds in two rounds. First, consider the case where R or less control packets are successful in the slotted Aloha contention and we have one round of arbitration. In this case all the successful packets are scheduled in the frame with F available transmission slots. Next, consider the case where more than R control packets are successful in the slotted Aloha contention and we have two rounds of arbitration. In this case we scan the M slotted Aloha slots from index 1 through M. In the first round we schedule the first R successful packets out of the slotted Aloha contention in the R long (F slots) transmission slots. In this round we schedule only one packet for each of the long transmission slots, irrespective of whether the packet is long or short. At this point (having filled each of the long transmission slots with one data packet) all the remaining successful control packets that correspond to long data packets fail in the scheduling and the transmitting node has to re–transmit the control packet. We then proceed with the second round. In the second round we schedule the remaining successful control packets that correspond to short data packets. Provided $F/K > 2$, we schedule these short data packets for the long transmission slots that hold only one short data packet from the first round. We also schedule these short data packets for the short ($F-M$ slots) transmission slots. After all the long and short transmission slots have been filled, the remaining short data packets fail in the scheduling and the transmitting node has to re–transmit the control packet. We note that our adopted arbitration policy is just one out of many possible arbitration policies, all of which can be analyzed in a similar fashion. The first–come–first–served and first–fit scheduling algorithm was chosen to meet the stringent timing requirements of reservation–based very high–speed networks.

With the adopted arbitration policy the expected number of scheduled long packets is

$$\begin{aligned}
E[\mathcal{L}] &= \sum_{k=0}^{M} E[\mathcal{L}|Z=k]\cdot P(Z=k) && (18)\\
&= \sum_{k=0}^{M} \min(k,\ R)\cdot \tilde{q}\cdot \binom{M}{k}\left(\frac{\beta e^{-\beta}}{D}\right)^{k}\left(1-\frac{\beta e^{-\beta}}{D}\right)^{M-k} && (19)\\
&= \tilde{q}\left\{ R - \sum_{k=0}^{\min(R,\ M)} \binom{M}{k}\left(\frac{\beta e^{-\beta}}{D}\right)^{k}\left(1-\frac{\beta e^{-\beta}}{D}\right)^{M-k}(R-k) \right\}. && (20)
\end{aligned}$$

To see this, note that in case there are $k \le R$ successful control packets in the slotted Aloha contention, then on average $\tilde{q}k$ of these correspond to long data packets. In case $k \ge R$, then there are on average $\tilde{q}R$ long packets among the first R control packets. (If the arbitration policy does not schedule the long (F slots) transmission slots first, but, say after l frames that allow transmission for $(F-M)$ slots have been scheduled, then an expected number of $\min\{R,\ \tilde{q}(lR\lfloor\frac{F-M}{K}\rfloor + R)\}$ long data packets are scheduled

given $(l\lfloor\frac{F-M}{K}\rfloor+1)R$ or more successful control packets in the slotted Aloha contention.) For notational convenience let

$$\varphi(\beta) := R - \sum_{k=0}^{\min(R,\ M)} \binom{M}{k} \left(\frac{\beta e^{-\beta}}{D}\right)^k \left(1-\frac{\beta e^{-\beta}}{D}\right)^{M-k} (R-k). \tag{21}$$

Thus,

$$E[\mathcal{L}] = \tilde{q}\cdot\varphi(\beta). \tag{22}$$

We now calculate the expected number of scheduled short packets. Generally,

$$E[\mathcal{S}] \quad = \quad \sum_{k=0}^{M} E[\mathcal{S}|Z=k]\cdot P(Z=k). \tag{23}$$

First, consider the case that there are no more than R successful control packets in the slotted Aloha contention, i.e., $k \le R$. In this case we have with (17)

$$\begin{aligned} E[\mathcal{S}|Z=k] &= E[Z_k|Z=k] &(24)\\ &= (1-\tilde{q})\cdot k, &(25)\end{aligned}$$

since all the successful control packets are scheduled in the long transmission slots.

Next, consider the case $R \le k \le M$. Let Θ denote the number of control packets that correspond to short data packets to be scheduled in the second round of arbitration. Note that $Z_k = R - \mathcal{L} + \Theta$, because $(R-\mathcal{L})$ short data packets have been scheduled in the first round of arbitration. With (17) we obtain

$$\begin{aligned} E[\mathcal{S}|Z=k] &= E\left[\min\left(R-\mathcal{L}+\Theta,\ (D-1)R\left\lfloor\frac{F-M}{K}\right\rfloor + (R-\mathcal{L})\left\lfloor\frac{F}{K}\right\rfloor\right) \mid Z=k\right] &(26)\\ &= E[R-\mathcal{L}|Z=k] + \\ &\quad E\left[\min\left(\Theta,\ (D-1)R\left\lfloor\frac{F-M}{K}\right\rfloor + (R-\mathcal{L})\left(\left\lfloor\frac{F}{K}\right\rfloor-1\right)\right) \mid Z=k\right]. &(27)\end{aligned}$$

Note that for a network without spatial wavelength reuse the $(D-1)R\lfloor(F-M)/K\rfloor$ term has to be replaced by zero in (26) and (27), as well as all the following expressions in this section.

Clearly, $E[R-\mathcal{L}|Z=k] = (1-\tilde{q})R$, since $k \ge R$. Moreover, note that conditional on $Z=k,\ k\ge R$, the random variables $\mathcal{L}$ and Θ are independent. This is because the first R successful control packets in the slotted Aloha slots determine $\mathcal{L}$; Θ is determined by the subsequent $(k-R)$ successful control packets. Hence,

$$\begin{aligned} &E\left[\min\left(\Theta,\ (D-1)R\left\lfloor\frac{F-M}{K}\right\rfloor + (R-\mathcal{L})\left(\left\lfloor\frac{F}{K}\right\rfloor-1\right)\right) \mid Z=k\right] &(28)\\ &= \sum_{j=1}^{k-R} P\left(\min\left(\Theta,\ (D-1)R\left\lfloor\frac{F-M}{K}\right\rfloor + (R-\mathcal{L})\left(\left\lfloor\frac{F}{K}\right\rfloor-1\right)\right) \ge j \mid Z=k\right) &(29)\\ &= \sum_{j=1}^{k-R} P(\Theta\ge j|Z=k)\cdot P\left((D-1)\left\lfloor\frac{F-M}{K}\right\rfloor + (R-\mathcal{L})\left(\left\lfloor\frac{F}{K}\right\rfloor-1\right) \ge j|Z=k\right). &(30)\end{aligned}$$

Now,

$$\begin{aligned} P(\Theta\ge j|Z=k) &= \sum_{m=j}^{k-R} P(\Theta=m|Z=k) &(31)\\ &= \sum_{m=j}^{k-R} \binom{k-R}{m} (1-\tilde{q})^m\,\tilde{q}^{k-R-m}. &(32)\end{aligned}$$

For notational convenience let

$$\gamma_j := P\left((D-1)\,R\left\lfloor\frac{F-M}{K}\right\rfloor + (R-\mathcal{L})\left(\left\lfloor\frac{F}{K}\right\rfloor - 1\right) \geq j | Z=k\right), \tag{33}$$

and

$$v_j := \min\left(R,\ \frac{(D-1)R\left\lfloor\frac{F-M}{K}\right\rfloor - j}{\left\lfloor\frac{F}{K}\right\rfloor - 1} + R\right). \tag{34}$$

If $\lfloor F/K \rfloor - 1 > 0$ then

$$\begin{aligned}
\gamma_j &= P(\mathcal{L} \leq v_j | Z=k) && (35)\\
&= \sum_{\{m:\ m \leq v_j\}} P(\mathcal{L}=m|Z=k) && (36)\\
&= \sum_{\{m:\ m \leq v_j\}} \binom{R}{m} \tilde{q}^m\,(1-\tilde{q})^{R-m}. && (37)
\end{aligned}$$

In case $\lfloor F/K \rfloor = 1$ we have

$$\gamma_j = \begin{cases} 1 & \text{if } j \leq (D-1)R\left\lfloor\frac{F-M}{K}\right\rfloor \\ 0 & \text{otherwise.} \end{cases} \tag{38}$$

Combining the cases $k \leq R$ and $R \leq k \leq M$, we obtain for $M \geq R$, which is typical for practical networks,

$$\begin{aligned}
E[\mathcal{S}] &= \sum_{k=0}^{R}(1-\tilde{q})k \cdot P(Z=k) + (1-\tilde{q})R \cdot \sum_{k=R+1}^{M} P(Z=k) + && (39)\\
&\quad \sum_{k=R+1}^{M}\left(\sum_{j=1}^{k-R}\gamma_j \sum_{m=j}^{k-R}\binom{k-R}{m}(1-\tilde{q})^m\,\tilde{q}^{k-R-m}\right)\cdot P(Z=k)\\
&= (1-\tilde{q})\left[R - \sum_{k=0}^{R}(R-k)\cdot P(Z=k)\right] + && (40)\\
&\quad \sum_{j=1}^{M-R}\gamma_j \sum_{m=j}^{M-R}\sum_{k=m+R}^{M}\binom{k-R}{m}(1-\tilde{q})^m\,\tilde{q}^{k-R-m}\cdot P(Z=k)\\
&=: h(\tilde{q},\ \beta). && (41)
\end{aligned}$$

D. Network/System Analysis

In this section we put the analyses for the individual components of the considered network, namely traffic model, slotted Aloha contention, and packet scheduling, together. We establish two equilibrium conditions and solve for the two unknowns $\tilde{q}$ and β. (Alternatively, we may consider the two unknowns $\tilde{q}$ and ν, noting that $\nu = (\beta-\alpha)/(\tilde{\alpha}-\alpha)$ for $\tilde{\alpha} \neq \alpha$; the case $\tilde{\alpha} = \alpha$ is discussed at the end of this section.)

In steady state the system satisfies the equilibrium condition

$$E[\mathcal{L}] = q(E[\mathcal{L}] + E[\mathcal{S}]). \tag{42}$$

To see this, note that in equilibrium the mean number of scheduled long packets from a given (fixed) AWG input port destined to a given (fixed) AWG output port (LHS) is equal to the mean number of newly generated long packets (RHS). Inserting (22) and (41) in (42) gives

$$\begin{aligned}
\tilde{q}\cdot\varphi(\beta) &= q[\tilde{q}\cdot\varphi(\beta) + h(\tilde{q},\ \beta)] && (43)\\
\Leftrightarrow (1-q)\cdot\tilde{q}\cdot\varphi(\beta) &= q\cdot h(\tilde{q},\ \beta). && (44)
\end{aligned}$$

The second equilibrium condition is

$$\frac{\sigma}{D} \cdot E[\eta] = E[\mathcal{L} + \mathcal{S}]. \tag{45}$$

This is because $\sigma \cdot \eta$ new packets are generated in each frame at the nodes attached to a given AWG input port. With probability $1/D$ each of the generated packets is destined to a given (fixed) AWG output port. On the other hand, $E[\mathcal{L} + \mathcal{S}]$ packets are scheduled (and transmitted) on average from a given AWG input port to a given AWG output port in one cycle; in equilibrium as many new packets must be generated. Inserting (1) and (2) in the LHS of (45) and (22) and (41) in the RHS of (45) we obtain

$$\frac{\tilde{\alpha} \cdot M}{D} \cdot \frac{\beta - \alpha}{\tilde{\alpha} - \alpha} = \tilde{q} \cdot \varphi(\beta) + h(\tilde{q},\ \beta). \tag{46}$$

Inserting (46) in (44) we obtain

$$\tilde{q} = \frac{q \cdot \tilde{\alpha} \cdot M}{D \cdot \varphi(\beta)} \cdot \frac{\beta - \alpha}{\tilde{\alpha} - \alpha}. \tag{47}$$

Inserting (47) in (46) we obtain

$$(1 - q) \cdot \frac{\tilde{\alpha} \cdot M}{D} \cdot \frac{\beta - \alpha}{\tilde{\alpha} - \alpha} = h\left(\frac{q \cdot \tilde{\alpha} \cdot M}{D \cdot \varphi(\beta)} \cdot \frac{\beta - \alpha}{\tilde{\alpha} - \alpha},\ \beta\right). \tag{48}$$

We solve Equation (48) numerically to obtain β (noting that by (11), $\min(\tilde{\alpha},\ \alpha) \leq \beta \leq \max(\tilde{\alpha},\ \alpha)$). We then insert β in (47) to obtain $\tilde{q}$. With β and $\tilde{q}$ we calculate $E[\mathcal{L}]$ (22) and $E[\mathcal{S}]$ (41). We define the mean throughput as the average number of transmitting nodes in steady state (which may also be interpreted as the average number of successfully transmitted data packets per frame). The mean throughput from a given (fixed) AWG input port to a given (fixed) AWG output port, i.e., the average number of nodes transmitting from a given AWG input port to a given AWG output port in steady state, is then given by

$$TH_{\text{port}} = \frac{F \cdot E[\mathcal{L}] + K \cdot E[\mathcal{S}]}{F \cdot D}. \tag{49}$$

The mean aggregate throughput of the network is

$$TH_{\text{net}} = D^2 \cdot TH_{\text{port}}. \tag{50}$$

Note that $\mathcal{L} \leq R$ and that $\mathcal{S}$ is bounded by the expression in Eqn. (16). Thus, if $(F - M)/K$ and F/K are integers, the aggregate throughput is bounded by $D^2 \cdot R \cdot [1 - M/F \cdot (1 - 1/D)]$.

We note that in case $\tilde{\alpha} = \alpha$, i.e., $\sigma = p$, we have from (11) $\beta = \alpha$. Inserting this in (44) gives an equation for $\tilde{q}$, which we solve numerically.

We now espouse the mean packet delay in the network. We define the mean delay as the average time period in cycles from the generation of the control packet corresponding to a data packet until the transmission of the data packet. Recall from Section IV-C that $E[\mathcal{L}] + E[\mathcal{S}]$ is the expected number of data packets that the nodes at a given AWG input port transmit to the nodes at a given AWG output port per cycle. Now, consider a given (fixed) node $m,\ 1 \leq m \leq N$. In the assumed uniform packet traffic scenario, this node m transmits on average $(E[\mathcal{L}] + E[\mathcal{S}])/S$ data packets to the nodes at a given AWG output port per cycle. Thus, node m transmits on average $(E[\mathcal{L}] + E[\mathcal{S}])D/S$ data packets to the N nodes attached to the D AWG output ports per cycle. The average time period in cycles from the generation of a control packet at node m until the generation of the next control packet is therefore $S/[D \cdot (E[\mathcal{L}] + E[\mathcal{S}])]$. Note that the time period from the successful scheduling of a data packet until the generation of the control packet for the next data packet is geometrically distributed with mean $(1 - \sigma)/\sigma$ cycles. Hence, the average delay in the network in cycles is

$$\text{Delay} = \frac{S}{D \cdot (E[\mathcal{L}] + E[\mathcal{S}])} - \frac{1 - \sigma}{\sigma}. \tag{51}$$

where $E[\mathcal{L}]$ and $E[\mathcal{S}]$ are known from the evaluation of the throughput (49).

V. Numerical Results

In this section, we show the benefit of spatial wavelength reuse and the impact of the system parameters on the throughput–delay performance of the network. Data packets can have one of two lengths. A data packet is F slots long with probability q and $K = F - M$ slots long with probability $(1-q)$. We consider the system parameters number of used FSRs R, fraction of long data packets q, number of reservation slots per frame M, physical degree of the AWG D, number of nodes N, and retransmission probability p. By default the parameters take on the following values: $R = 2$, $q = 0.25$, $M = 30$, $D = 4$, $N = 200$, $p = 0.8$, $F = 200$, and $K = 170$. (For these default parameters the aggregate throughput is bounded by 28.4.) Each cycle is assumed to have a constant length of $D \cdot F = 800$ slots. All numerical results in this section are obtained using the expression (15), which approximates the number of successful control packets in the slotted Aloha contention by a Poisson distribution. (For a numerical evaluation of the refined approximation, that does not use the Poisson distribution, but uses directly the binomial distribution, we refer the interested reader to Appendix A. In summary, we find that using the approximate expression (15) gives very accurate results for a wide range of parameter values, as is also demonstrated by the numerical results in this section.) We also provide extensive simulation results of a more realistic network in order to verify the accuracy of the analysis. As opposed to the analysis, in the simulation a given node cannot transmit data packets to itself and both length and destination of a given data packet are not renewed when retransmitting the corresponding control packet. In addition, the simulation takes receiver collisions into account, i.e., a given data packet is not scheduled if the receiver of the intended destination node is busy. Each simulation was run for 10^7 slots including a warm–up phase of 10^6 slots. Using the method of batch means we also calculated the 98% confidence intervals for the mean aggregate throughput and the mean delay whose variations from the sample mean were less than 1% for all simulation results.

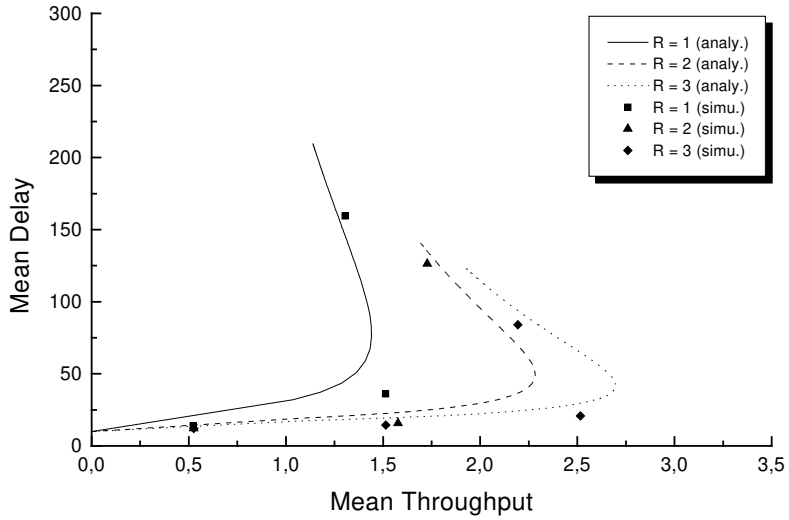

Fig. 6. Mean delay (cycles) vs. mean aggregate throughput (packets/frame) for different number of used FSRs $R \in \{1,\ 2\}$

Fig. 6 shows the mean delay vs. the mean aggregate throughput as the mean arrival rate σ is varied from 0 to 1. As one would expect, using two FSRs instead of one (leaving all other parameters unchanged) dramatically increases the mean aggregate throughput while decreasing the mean delay. This is due to the fact that an additional FSR increases the degree of concurrency and thereby mitigates the scheduling bottleneck resulting in more successfully transmitted data packets and fewer retransmissions. Note that the number of

used FSRs is limited and is determined by the transceiver tuning range, the degree of the underlying AWG, and the channel spacing. To avoid tuning penalties we deploy fast tunable transceivers whose tuning range is typically $10 - 15$ nanometers (nm). All results presented in this section assume a channel spacing of 200 GHz, i.e., 1.6 nm at 1.55 μm. Thus, we can use $7 - 10$ wavelengths at each AWG input port depending on the transceiver tuning range. For all subsequent results the number of wavelengths is assumed to be eight. Consequently, with a 4×4 AWG we deploy two FSRs for concurrent transmission/reception of data packets.

Figs. 7 and 8 illustrate that spatial wavelength reuse dramatically improves the throughput–delay performance of the network for variable–size data packets. Fig. 7 shows the mean aggregate throughput vs. the mean arrival rate σ with and without spatial wavelength reuse for different fraction of long data packets $q \in \{0,\ 0.5,\ 1.0\}$. Simulation and analysis results match very well. For $q = 1.0$, i.e., all data packets have a length of F slots, the mean aggregate throughput is the same no matter whether wavelengths are spatially reused or not. This is because the data packets are too long for being scheduled in the $(D-1)$ frames in which the corresponding nodes do not send control packets and spatial wavelength reuse would be possible in the last $(F - M)$ slots of the frame. Thus, these frames remain unused for $q = 1.0$. For $q = 0.5$, 50% of the data packets are long (F slots) and the other 50% are short (K slots). Allowing for spatial wavelength reuse the latter ones can now be scheduled in all frames including the aforementioned $(D - 1)$ frames. Consequently, with wavelength reuse more data packets are successfully transmitted, resulting in a higher throughput. In contrast, without wavelength reuse data packets can be scheduled only in one frame per cycle in which the corresponding nodes also transmit their control packets. Furthermore, since for $q = 0.5$ some successfully transmitted data packets are short (K slots), wavelengths are not fully utilized resulting in a lower throughput compared to $q = 1.0$. For $q = 0$ the benefit of spatial wavelength reuse becomes even more dramatic. In this case there are only short data packets (K slots) which fill up a large number of frames leading to a further increased mean aggregate throughput. Note that for $q = 0$ spatial wavelength reuse significantly increases the maximum aggregate throughput by more than 60%. All curves in Fig. 7 run into saturation since for increasing σ no additional data packets can be scheduled due to busy channels and receivers and an increasing number of colliding control packets.

Fig. 8 depicts the mean delay vs. σ with and without wavelength reuse for different fraction of long data packets $q \in \{0,\ 0.5,\ 1.0\}$. We observe that the simulation gives slightly larger delays than the analysis. This is because the simulation takes also the transmission time of data packets into account as opposed to the analysis. In the analysis the mean delay is equal to the time interval between the generation of a given data packet and the end of the cycle in which the given data packet is successfully scheduled but not yet transmitted. All curves have in common that at very light traffic the mean delay is equal to one cycle owing to the propagation delay of the control packet. With increasing σ the mean delay increases due to more unsuccessful control packets. These control packets have to be retransmitted, resulting in an increased mean delay. Note that we obtain the largest delay if the aforementioned $(D - 1)$ frames per cycle cannot be used for data transmission. This holds not only for the cases where wavelength reuse is not deployed but also for $q = 1.0$ with spatial wavelength reuse. This is due to the fact that for $q = 1.0$ the data packets are too long and do not fit in the last $(F - M)$ slots of the aforementioned $(D - 1)$ frames. As a consequence, for these cases fewer data packets can be successfully scheduled and the corresponding control packets have to be retransmitted more often, leading to a higher mean delay. With decreasing q there are more short data packets which can easily be scheduled in the aforementioned $(D - 1)$ frames. Due to the resulting wavelength reuse more data packets can be successfully scheduled. Therefore, fewer control packets have to be retransmitted leading to a decreased mean delay. In particular, for $q = 0$, wavelengths are used very efficiently resulting in the lowest mean delay.

The impact of the number of reservation slots M per frame on the network throughput–delay performance is shown in Figs. 9 and 10. The mean aggregate throughput and the mean delay are depicted as a function of σ for $M \in \{15,\ 20,\ 30,\ 40\}$. Recall that by default the frame length F is set to 200 slots. Each frame is composed of M reservation slots and $K = F - M$ slots which can be used for transmitting short

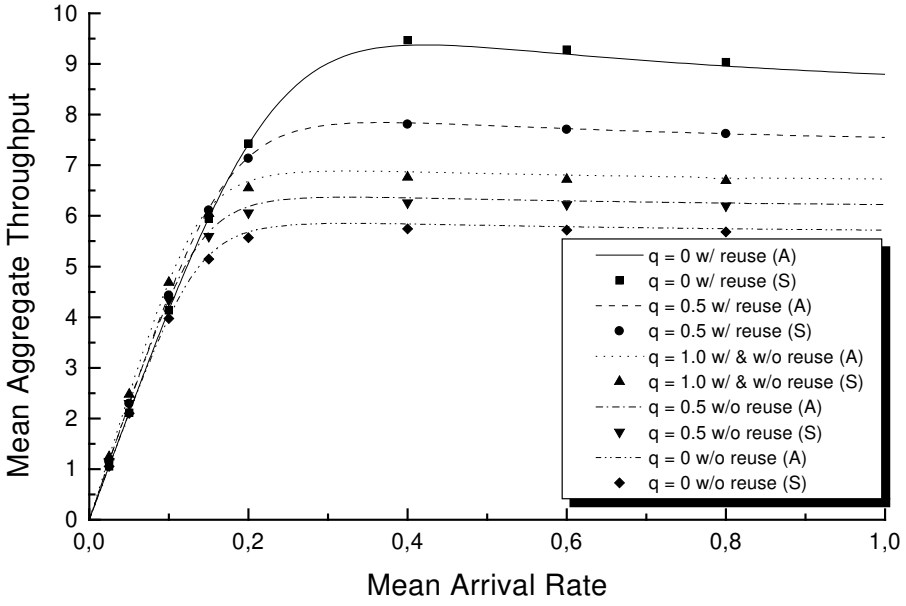

Fig. 7. Mean aggregate throughput (packets/frame) vs. mean arrival rate σ with and without wavelength reuse for different fraction of long data packets $q \in \{0,\ 0.5,\ 1.0\}$

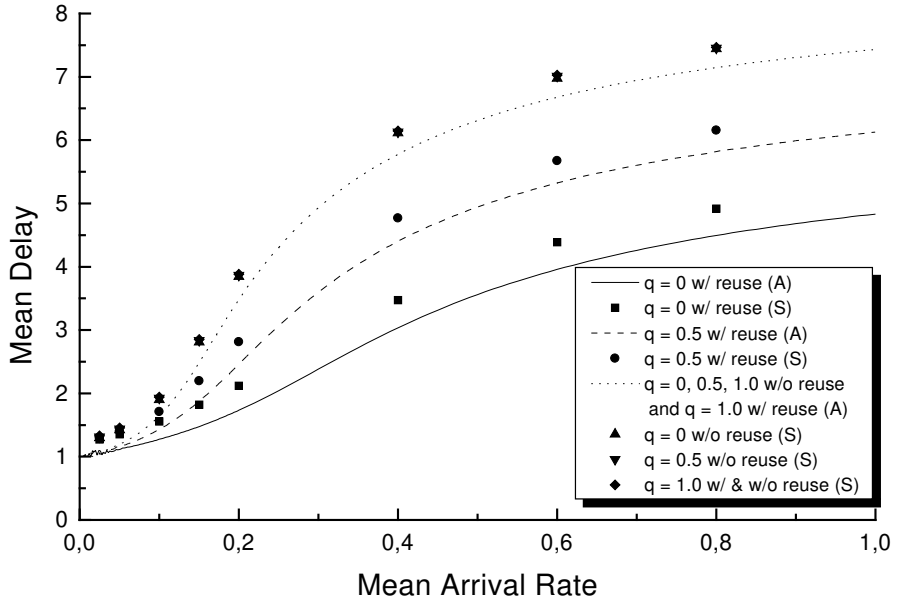

Fig. 8. Mean delay (cycles) vs. mean arrival rate σ with and without wavelength reuse for different fraction of long data packets $q \in \{0,\ 0.5,\ 1.0\}$

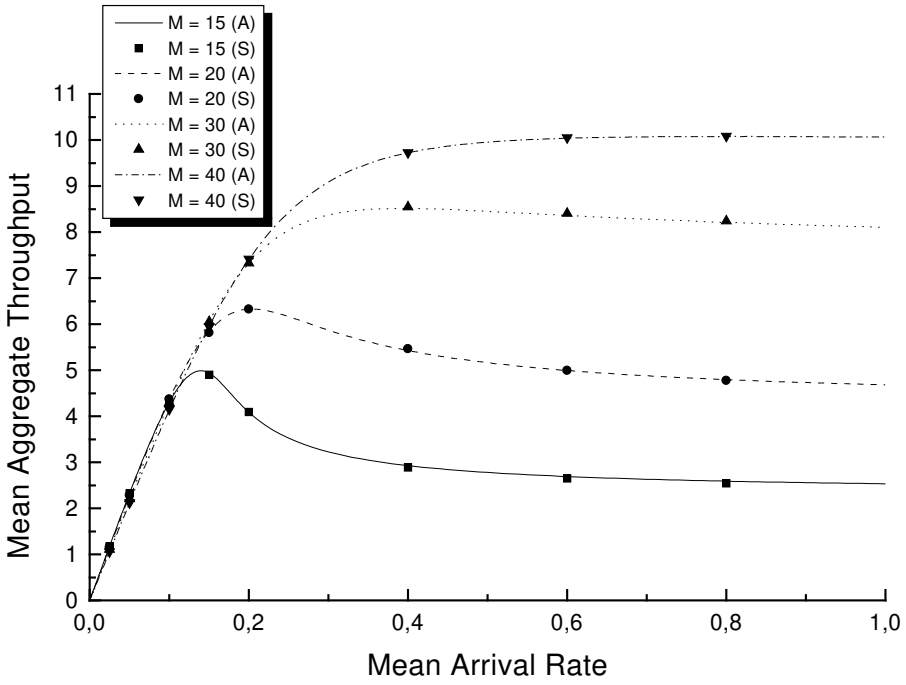

Fig. 9. Mean aggregate throughput (packets/frame) vs. mean arrival rate σ for different number of reservation slots $M \in \{15,\ 20,\ 30,\ 40\}$

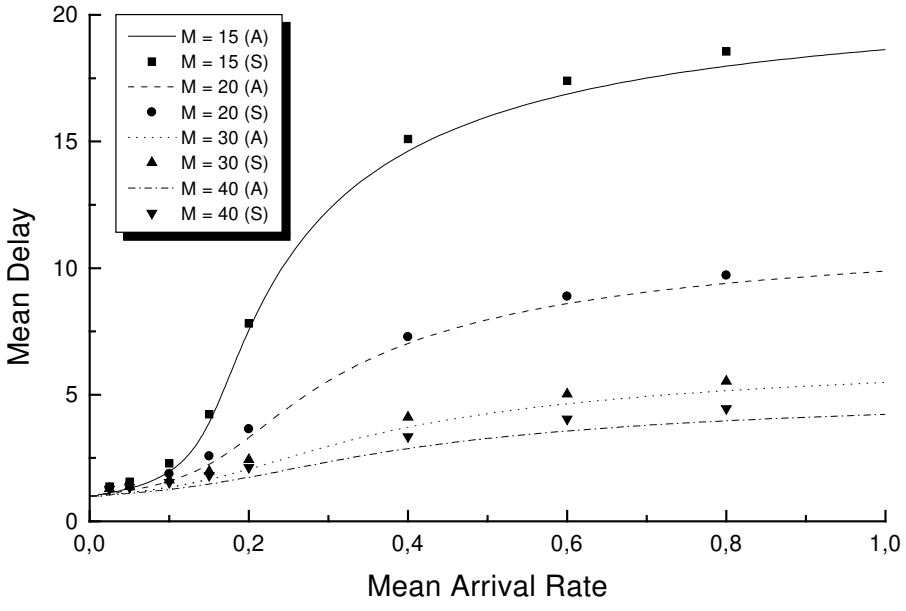

Fig. 10. Mean delay (cycles) vs. mean arrival rate σ for different number of reservation slots $M \in \{15,\ 20,\ 30,\ 40\}$

packets by means of spatial wavelength reuse. Clearly, for a fixed F, increasing M decreases the length of short packets K, and reduces the contribution of the short packets to the throughput. At the same time, increasing M increases the probability of successful control packet contention. We observe from Figs. 9 and 10 that the effect of increasing the probability of successful control packet contention dominates in the considered scenario, that is, the mean throughput increases and the mean delay decreases as M increases. This indicates that the random access reservation scheme can be a severe bottleneck. Indeed, for a very small number of $M = 15$ control slots, we observe the typical bi–stable behavior of slotted Aloha — the underlying mechanism of our control packet contention — in the throughput results. Note that the network throughput–delay performance could be easily improved by replacing the random access of the reservation slots with a dedicated assignment of the reservation slots. However, such a dedicated slot assignment requires reconfiguration when adding new nodes.

A given number of nodes can be connected by AWGs with different physical degree D. Figs. 11 and 12 depict for $D \in \{2,\ 4,\ 8\}$ the mean aggregate throughput and the mean delay as a function of σ, respectively. Recall that we have chosen the transceiver tuning range and the channel spacing such that we make use of eight wavelengths. The number of used FSRs R is then determined only by the physical degree D of the underlying AWG and is given by $R = 8/D$. Consequently, for a smaller D more FSRs can be exploited, and vice versa for a larger D. Furthermore, for a smaller D each cycle contains fewer but longer frames, and vice versa for a larger D.

As shown in Fig. 11, $D = 2$ provides the largest maximum mean aggregate throughput at light traffic. However, with increasing σ the mean aggregate throughput decreases. This is due to the fact that for $D = 2$ short data packets are rather long ($K = 800/D - M = 370$ slots) resulting in a higher channel utilization and thereby a higher throughput at small traffic loads. But a small D also implies that for a given population N more nodes are attached to the same combiner since $S = N/D$. All these S nodes make their reservations in the same frame. For an increasing σ this leads to more collisions of control packets resulting in a lower mean aggregate throughput and a higher mean delay due to more retransmissions of the corresponding control packets (Fig. 12).

This problem is alleviated by deploying a 4×4 or 8×8 AWG. For a larger D fewer nodes send control packets in the same frame causing fewer collisions at high traffic loads. However, for $D = 4$ and $D = 8$ only 2 FSRs and 1 FSR can be deployed, respectively. Moreover, a larger D reduces the length of short data packets as well. Fig. 11 shows that for $D = 4$ the mean aggregate throughput is rather high for a wide range of σ. Whereas for $D = 8$ the throughput is rather low due to the small number of control packets per frame and the low channel utilization owing to the reduced length of short data packets. Note that for $D = 4$ the mean aggregate throughput gradually decreases for increasing σ. This is because at high traffic loads control packets suffer from collisions and have to be retransmitted, resulting in a slightly higher mean delay compared to $D = 8$. Concluding, in terms of throughput–delay performance choosing $D = 4$ seems to provide the best solution for a wide range of traffic loads.

Figs. 13 and 14 depict the throughput–delay performance of the network for different population $N \in \{40,\ 100,\ 200,\ 300\}$. As shown in Fig. 13, increasing N improves the mean aggregate throughput due to more reservation requests and successfully scheduled data packets. However, for $N = 200$ and especially $N = 300$ the throughput decreases for increasing σ. This is because for large populations more control packets suffer from channel collisions resulting in a lower mean aggregate throughput. Accordingly, this leads to higher mean delays as shown in Fig. 14. Note that simulation and analysis results match very well even for small populations despite the fact that (i) we have conducted an asymptotic analysis for large S, and (ii) the analysis does not take receiver collisions into account (while the simulation does). Also, the analysis assumes non–persistent destinations, whereas the destinations are persistent in the simulation.

The impact of different AWG degree $D \in \{2,\ 4,\ 8\}$ on the system throughput–delay performance is shown in Figs. 15 and 16, respectively. Note that throughput and delay are not given as a function of σ but as a function of the fraction of long data packets $q \in [0,\ 1]$. Recall that we have assumed a constant cycle

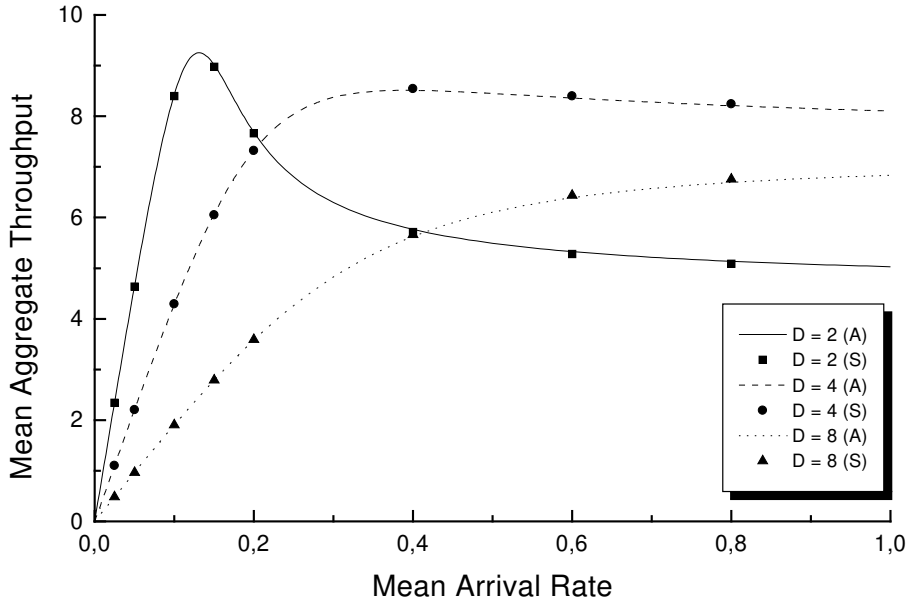

Fig. 11. Mean aggregate throughput (packets/frame) vs. mean arrival rate σ for different AWG degree $D \in \{2,\ 4,\ 8\}$

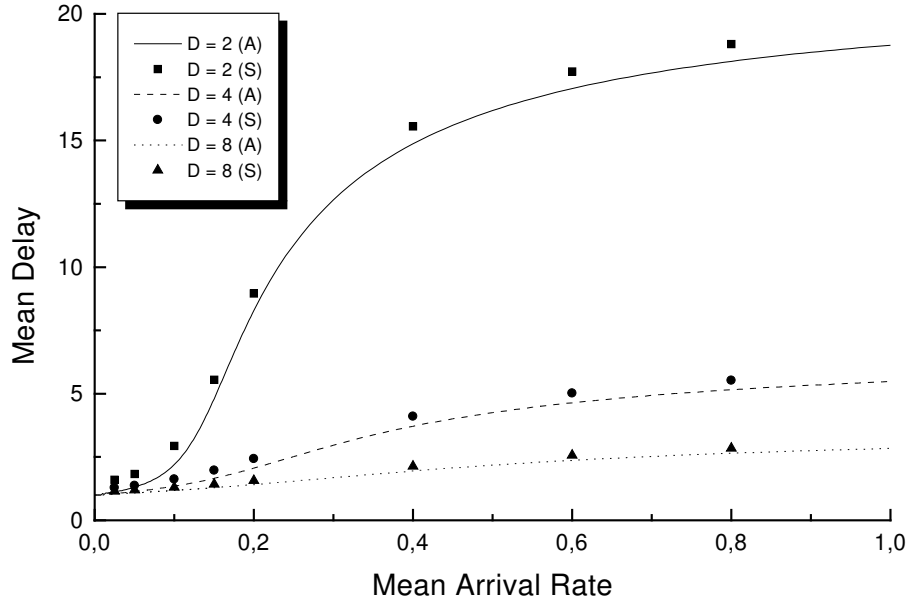

Fig. 12. Mean delay (cycles) vs. mean arrival rate σ for different AWG degree $D \in \{2,\ 4,\ 8\}$

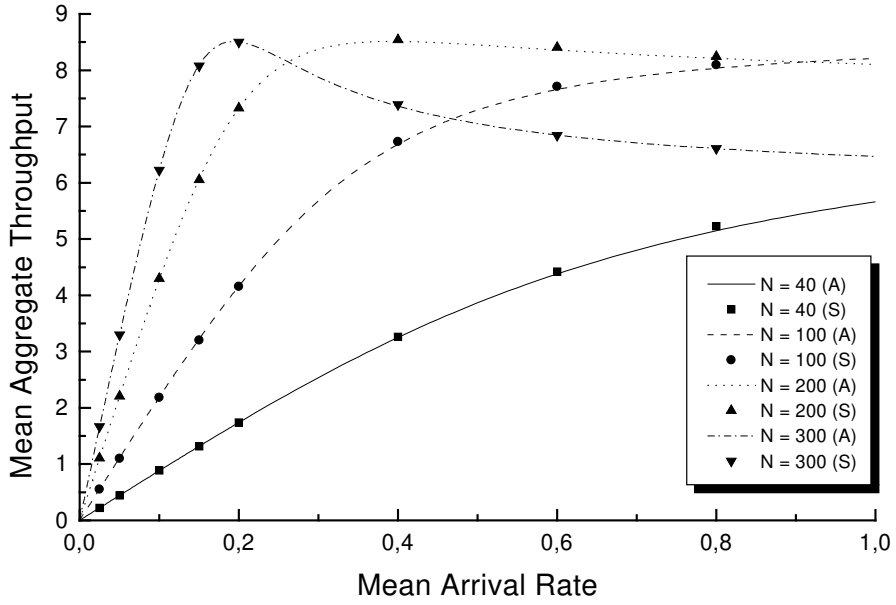

Fig. 13. Mean aggregate throughput (packets/frame) vs. mean arrival rate σ for different population $N \in \{40,\ 100,\ 200,\ 300\}$

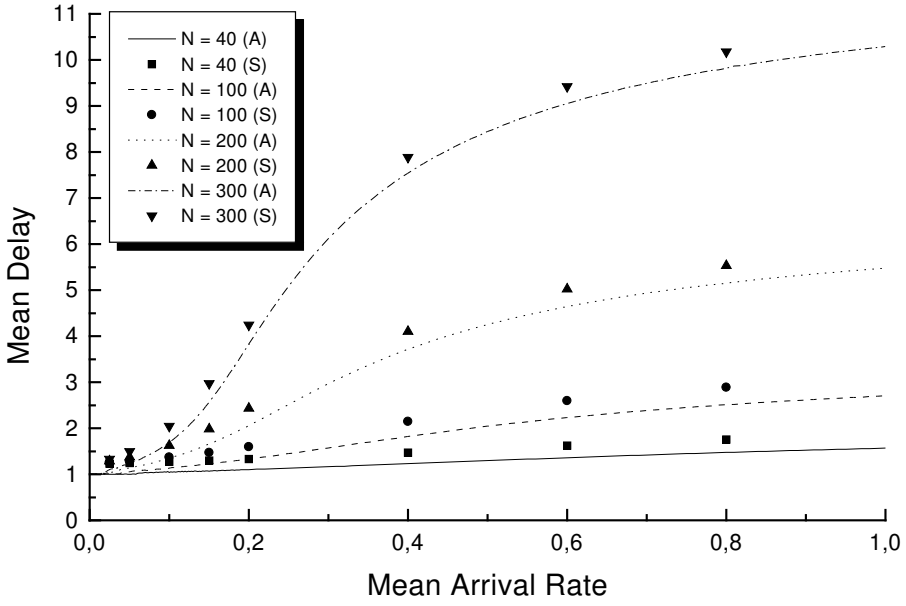

Fig. 14. Mean delay (cycles) vs. mean arrival rate σ for different population $N \in \{40,\ 100,\ 200,\ 300\}$

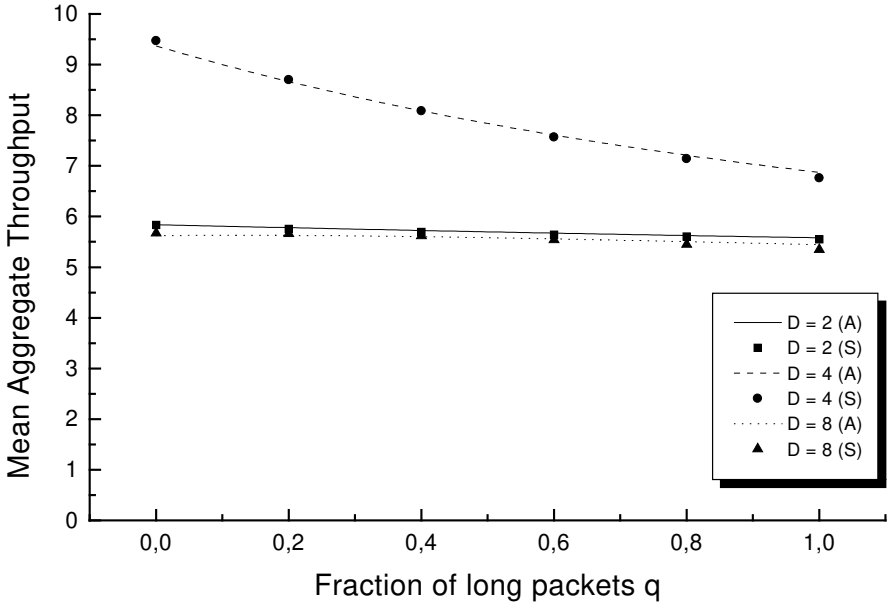

Fig. 15. Mean aggregate throughput (packets/frame) vs. fraction of long data packets q for different AWG degree $D \in \{2,\ 4,\ 8\}$.

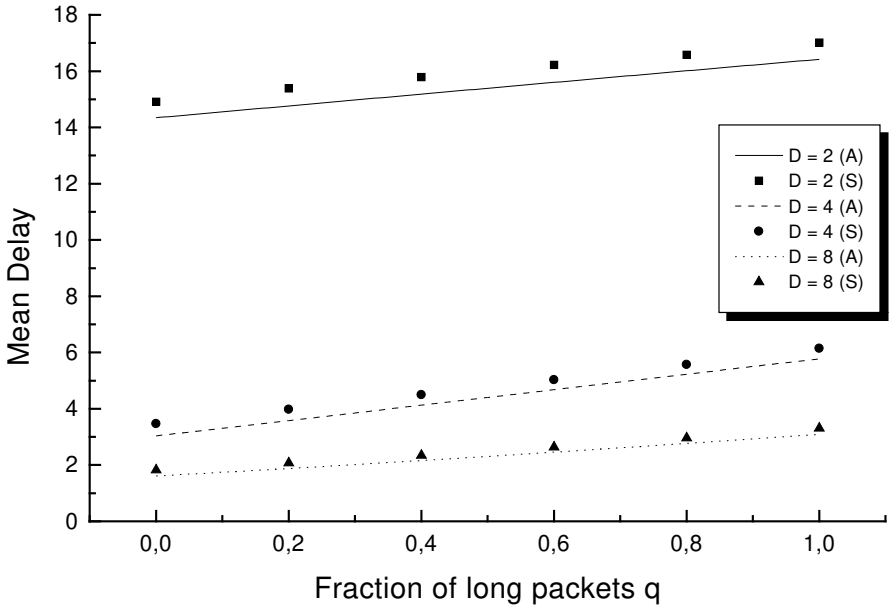

Fig. 16. Mean delay (cycles) vs. fraction of long data packets q for different AWG degree $D \in \{2,\ 4,\ 8\}$.

length of 800 slots and a fixed number of reservation slots $M = 30$. As a consequence, the frame length is given by $F = 800/D$ slots and the length of short data packets is equal to $K = F - M = 800/D - 30$ slots. Moreover, the number of used FSRs of the underlying AWG is given by $R = 8/D$.

Fig. 15 depicts the mean aggregate throughput vs. q. For $D = 4$ we observe that the throughput monotonously decreases for increasing q. For $q = 0$ all data packets are short and can be scheduled in any frame resulting in a high mean aggregate throughput. For increasing q more and more data packets are long ($q = 1$ implies that there are only long data packets). However, long data packets can be scheduled only in one frame per cycle. In addition, at most two of them can be scheduled since $R = 2$. Consequently, for increasing q fewer data packets can be scheduled resulting in a decreasing mean aggregate throughput (Fig. 15) and a higher mean delay as shown in Fig. 16. For $D \in \{2,\ 8\}$ the mean aggregate throughput is smaller than for $D = 4$ and, more interestingly, almost independent from q. For $D = 2$, twice as many nodes are attached to each AWG input port compared to $D = 4$. As a consequence, more control packets suffer from collisions and fewer data packets are available for scheduling, resulting in a smaller mean aggregate throughput. Moreover, there are not enough control packets to fully capitalize on spatial wavelength reuse. Thus, for $q = 0$ only slightly more data packets are successfully scheduled than for $q = 1$. However, since for $q = 1$ all successfully scheduled data packets are long as opposed to $q = 0$ the mean aggregate throughput is about the same in both cases. Similarly, since for $D = 8$ fewer nodes are attached to each AWG input port there are fewer reservation requests per frame than for $D = 4$ resulting in a smaller throughput. However, these reservation requests experience fewer collisions significantly decreasing the mean delay as illustrated in Fig. 16. In contrast, $D = 2$ provides the highest mean delay due to the large number of collided control packets and their retransmissions. Note that for all $D \in \{2,\ 4,\ 8\}$ the mean delay grows with increasing q since for larger q fewer data packets can be scheduled owing to the lack of spatial wavelength reuse. This leads to more retransmissions of control packets and thereby to an increased mean delay. Concluding, while $D \in \{2,\ 8\}$ suffers from a relatively small throughput and $D = 2$ exhibits a too large mean delay, choosing $D = 4$ seems to provide the best compromise in terms of throughput–delay performance.

Figs. 17 and 18 depict the throughput–delay performance of the network as a function of σ for different retransmission probability $p \in \{0.3,\ 0.6,\ 0.9\}$. As shown in Fig. 17, for $p = 0.3$ the mean aggregate throughput grows monotonously for increasing σ. We observe that at high traffic loads the slope decreases due to increasingly busy channels and transceivers. For $p = 0.6$ the mean aggregate throughput is larger than for $p = 0.3$. This is because with a larger p nodes retransmit collided control packets with a higher probability resulting in more successful control packets and an increased mean aggregate throughput. However, further increasing p has a detrimental impact on the throughput. For $p = 0.9$ nodes retransmit collided control packets after a small backoff period. As a consequence, at medium to high traffic loads an increasing number of control packets collide leading to a decreased mean aggregate throughput. Fig. 18 shows that for all $\sigma \in (0,\ 1]$, $p = 0.3$ yield larger mean delays than $p = 0.6$. With a smaller p nodes defer retransmissions of collided control packets for a larger time interval which in turn increases the mean delay. Note that for $p = 0.9$ nodes experience the smallest mean delay at light to medium traffic loads. At high loads the mean delay becomes the largest one due to the increasing number of retransmissions of control packets.

VI. Supplementary simulation results

In this section, we investigate by means of extensive simulations additional aspects of our AWG based network which have not been analyzed yet in the previous three sections. Prior to the investigations we have validated the correctness of our simulator. We have run simulations making the same assumptions as in Section V. In extensive tests we have compared the obtained results with the analytical ones presented in Section V. Throughout our tests we have achieved very good matches between simulation and analysis. In the following simulations we not only examine additional performance metrics but, more importantly, also relax some of the simplifying assumptions made in the previous analyses in order to make our investigations more realistic. Specifically, in Section VI-A we consider the self–stability of our MAC protocol. We thereby

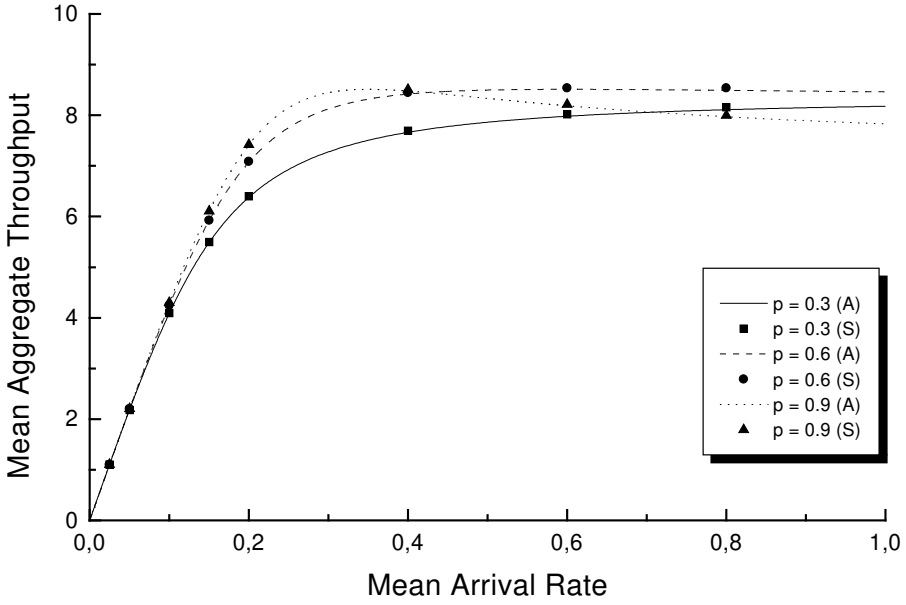

Fig. 17. Mean aggregate throughput (packets/frame) vs. mean arrival rate σ for different retransmission probability $p \in \{0.3,\ 0.6,\ 0.9\}$

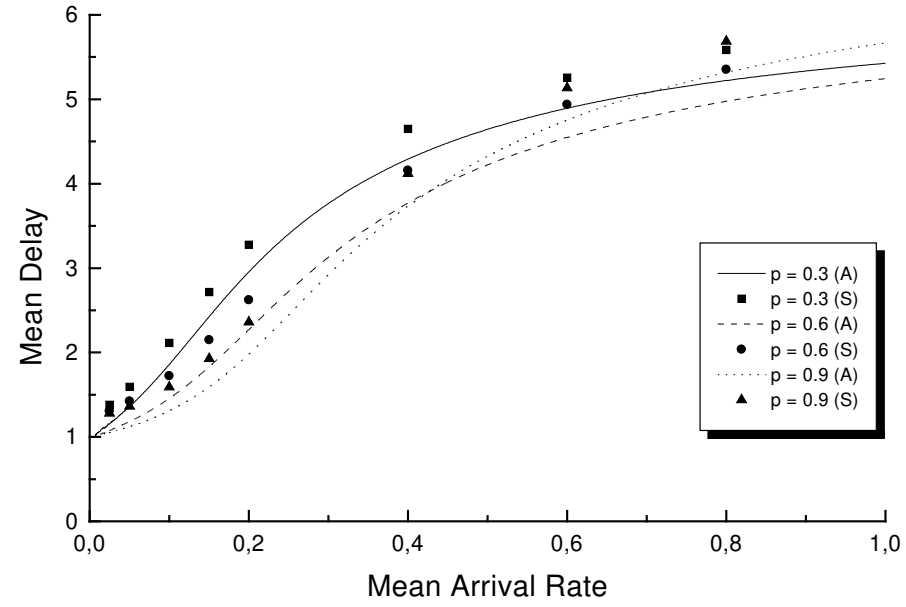

Fig. 18. Mean delay (cycles) vs. mean arrival rate σ for different retransmission probability $p \in \{0.3,\ 0.6,\ 0.9\}$

drop the assumption of a constant retransmission probability p of unsuccessful control packets. Instead, p is reduced each time a given reservation request fails. Section VI-B investigates several approaches to reduce packet loss. Among others, packet loss can be decreased by extending the scheduling window size and/or equipping each node with a finite buffer. In doing so, we relax the assumptions of a finite window size and a single–packet buffer made in our analytical models. Finally, in Section VI-C we also take circuit switching into consideration beside packet switching. Section VI-D discusses the efficiency of our network and also conducts a benchmark comparison with a previously reported reservation protocol designed for a PSC based single–hop metro WDM network. In the following simulative studies we concentrate on unicast traffic.

In the simulation the network parameters take on the following values by default: Number of nodes $N = 200$, physical degree of AWG $D = 2$, number of used FSRs $R = 4$, number of slots per frame $F = 200$, number of reservation slots per frame $M = 30$, and retransmission probability $p = 1.0$. The fraction of long data packets q equals 0.25, i.e., a generated data packet is long ($L = F = 200$ slots) with probability $q = 0.25$ and short ($L = K = (F - M) = 170$ slots) with probability $(1 - q) = 0.75$. By default the size of the scheduling window is one cycle and each node's single–packet buffer is able to store either a long or a short data packet. Each cycle is assumed to have a constant length of $D \cdot F = 400$ slots. The propagation delay is assumed to be less than one cycle. A node that has made a successful reservation can send the corresponding data packet in the next cycle. In the simulation the mean arrival rate denotes the probability that a given node generates a data packet at the beginning of that frame in which the node is allowed to send control packets. In the simulation nodes can not transmit data packets to themselves. A given data packet is destined to any of the other $(N - 1)$ nodes with equal probability $1/(N - 1)$. Furthermore, in the simulation the length and destination of a given data packet are not changed if the corresponding control packet has to be retransmitted. As in Section IV, the mean aggregate throughput is given in packets per frame. The mean delay is equal to the time interval between the generation and the end of transmission of a given data packet and is given in cycles. Each simulation was run for $6 \cdot 10^5$ slots including a warm–up phase of $6 \cdot 10^4$ slots. The remaining $5.4 \cdot 10^5$ slots were divided into 90 batches, each comprising $6 \cdot 10^3$ slots. Using the method of batch means we calculated the 95% confidence intervals for the mean aggregate throughput, mean delay, and relative packet loss.

A. *Self–stability*

So far, we have not addressed *backoff* in our MAC protocol. Without backoff the retransmission probability p remains constant irrespective of how many times a given control packet has already been retransmitted. As the mean arrival rate increases more nodes are backlogged and try to make a reservation in their assigned frame. The network becomes increasingly congested and more nodes have to retransmit their unsuccessful control packets (with constant probability $p = 1$ in our case). This leads to an increased number of control packet collisions on slotted ALOHA and retransmissions. As a result, the mean aggregate throughput decreases while the mean delay increases dramatically with an increasing mean arrival rate, as depicted in Figs. 19 and 20, respectively.

By deploying backoff this instability can be alleviated. With backoff the retransmission probability p is reduced each time the reservation fails. More precisely, a given control packet which for the first time fails in making a successful reservation is retransmitted with probability $p = 1$ in the next cycle. If the reservation again fails p is reduced by 50%. Thus, the corresponding control packet is retransmitted in the next cycle with probability $p = 0.5$. With probability $(1-p) = 0.5$ the reservation is deferred by one cycle. The control packet is then retransmitted with probability $p = 0.5$ and deferred with probability $(1 - p) = 0.5$, and so on. Each time the reservation fails p is further halved. In general, p is reduced by 50% at most b times until p is equal to a given minimum retransmission probability p_{min}, where $b \geq 1$ denotes the backoff limit. Once p_{min} is reached the corresponding control packet is retransmitted in a cycle with probability p_{min} until the reservation is successful, i.e., there is no attempt limit.

Figs. 19 and 20 show the positive impact of backoff on the throughput–delay performance of the network

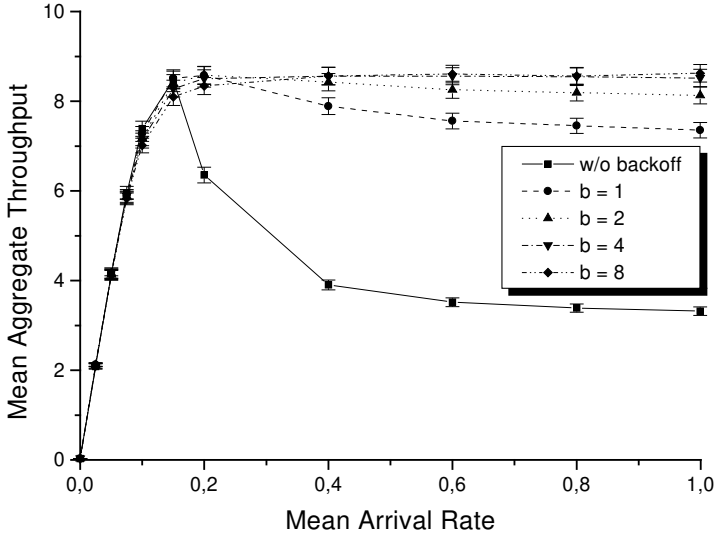

Fig. 19. Mean aggregate throughput (packets/frame) vs. mean arrival rate for different backoff limits $b \in \{1, 2, 4, 8\}$.

for $b \in \{1, 2, 4, 8\}$. We observe that already with $b = 1$, i.e., p is halved not more than once, the mean aggregate throughput is significantly increased for medium to high traffic loads while decreasing the mean delay considerably. It can be seen in Fig. 19 that with $b = 4$ the mean aggregate throughput does not decrease for an increasing mean arrival rate. This property of MAC protocols is known as *self–stability*. Further increasing b does not yield a better throughput–delay performance. Therefore, we set $b = 4$ in the subsequent simulations.

B. Packet loss

Since so far we assume that each node is equipped with a single–packet buffer new generated data packets are dropped when they find the buffer full resulting in packet loss. Fig. 21 depicts the relative packet loss (ratio of number of dropped data packets and number of generated data packets) as a function of the mean arrival rate for $b = 4$ and different number of reservation slots per frame $M \in \{30, 40, 60, 80\}$. Obviously, the relative packet loss increases monotonously for an increasing mean arrival rate. For increasing M the relative packet loss is decreased. Hence, choosing M as large as possible appears to be the best solution to reduce the packet loss. However, from the throughput perspective it is more advantageous to choose M neither too small nor too large. This can be seen in Fig. 22 where $M = 60$ achieves the largest mean aggregate throughput for a wide range of mean arrival rate. While small values of M result in more control packet collisions on slotted ALOHA and thereby a decreased mean aggregate throughput, $M = 80$ implies that more control packets are sent collisionfree but the length of the corresponding short data packets is reduced since the length of short data packets is given by $K = (F - M) = (200 - M)$ slots. Fig. 23 illustrates the mean delay vs. mean arrival rate. Since the mean delay for $M = 60$ and $M = 80$ do not differ significantly and $M = 60$ achieves the maximum mean aggregate throughput we set $M = 60$ in the subsequent simulations.

After considering the slotted ALOHA bottleneck we now turn our attention to two other bottlenecks of the proposed network: (i) Number of wavelength channels, and (ii) scheduling window size. Fig. 24 shows the relative packet loss vs. mean arrival rate for $b = 4$, $M = 60$, and different number of used FSRs $R \in \{2, 4, 6, 8, 16\}$. We observe that the packet loss is decreased significantly for increasing R. Figs. 25 and

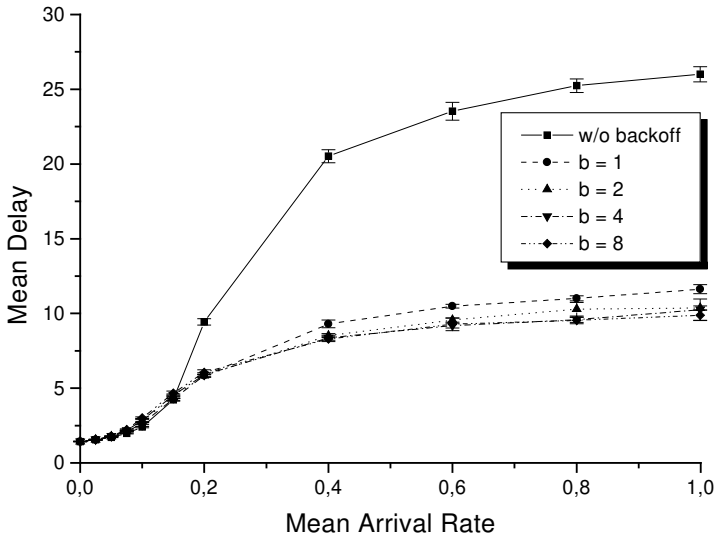

Fig. 20. Mean delay (cycles) vs. mean arrival rate for different backoff limits $b \in \{1, 2, 4, 8\}$.

26 show that a large R also improves the throughput–delay performance of the network. This is due to the fact that by increasing R, i.e., by using more FSRs of the underlying AWG, the degree of concurrency is increased. As a consequence, data packets are simultaneously transmitted over a larger number of wavelengths leading to a decreased packet loss, increased throughput, and decreased delay. Note that for a given transceiver tuning range $\Lambda = D \cdot R$ and a given AWG degree D increasing R implies a smaller channel spacing. Generally, AWGs with a smaller channel spacing exhibit a larger crosstalk. In order to achieve acceptable crosstalk values we set $R = 8$ in the subsequent simulations. For $D = 2$ and a typical fast transceiver tuning range of 12 nm setting $R = 8$ translates into a channel spacing of 100 GHz.

Fig. 27 depicts the packet loss as a function of the mean arrival rate for different scheduling window size $W \in \{2, 4, 6, 8\}$, where W is given in frames. We observe that the packet loss is decreased when the scheduling window is enlarged from two frames (i.e., one cycle for $D = 2$) to four frames (i.e., two cycles for $D = 2$). Further increasing W has no impact on the packet loss. The same observation can be made for the mean aggregate throughput and the mean delay, as shown in Figs. 28 and 29, respectively. Using a scheduling window of two cycles increases the throughput and decreases the delay. However, further increasing W does not affect the throughput–delay performance of the network. This is due to the fact that for $W \geq 4$ the network resources are almost fully utilized. As a consequence, no additional data packets can be transmitted which in turn leads to a stagnating packet loss, throughput, and delay for increasing W.

Note that all curves are rather close to each other. The packet loss, throughput, and delay do not change very much for increasing W. Therefore, we can state that the assumption of a scheduling window size equal to one cycle as made in the analyses so far is reasonable. For the subsequent simulations we set $W = 8$. This is because we are going to consider also the case $D = 4$ below; in order to benefit from a scheduling window of two cycles the parameter W must be equal to eight for $D = 4$.

Another approach to reduce the packet loss is increasing the buffer at each node. The single–packet buffer is replaced with a buffer that is able to store up to B long data packets, where $B \geq 1$. (The queueing discipline is assumed to be first–in–first–out (FIFO)). Fig. 30 illustrates the positive impact of larger buffer sizes on the packet loss. With increasing B more arriving data packets can be stored and do not have to be dropped resulting in a decreased packet loss. In addition, each node is less likely to be idle which leads to an increased mean aggregate throughput, as depicted in Fig. 31. However, Fig. 32 shows that the mean delay

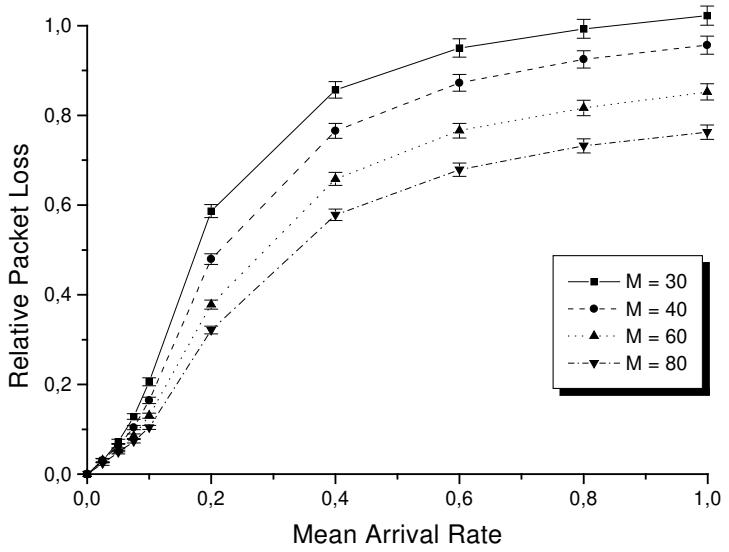

Fig. 21. Relative packet loss vs. mean arrival rate for $b = 4$ and different $M \in \{30, 40, 60, 80\}$.

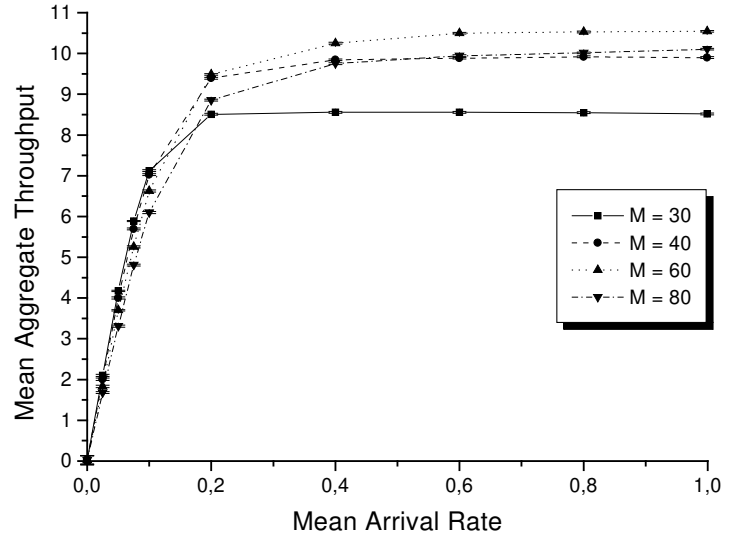

Fig. 22. Mean aggregate throughput (packets/frame) vs. mean arrival rate for $b = 4$ and different $M \in \{30, 40, 60, 80\}$.

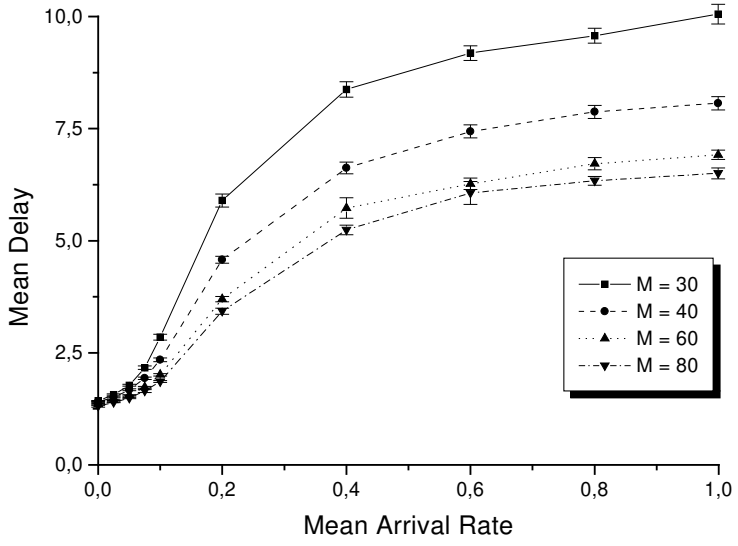

Fig. 23. Mean delay (cycles) vs. mean arrival rate for $b = 4$ and different $M \in \{30, 40, 60, 80\}$.

significantly increases for larger B. This is because in larger buffers there are packets which have to wait for a longer time interval until they can be transmitted. Clearly, there is a trade–off between packet loss and delay. To avoid large delays and provide a reasonable throughput–loss performance we set $B = 10$ in the following simulations. Fig. 33 shows that the packet loss can be further reduced if the $N = 200$ nodes are connected by a 4×4 AWG instead of a 2×2 one. Note that changing the AWG degree D implies also a different R, F, and B. For a given transceiver tuning range Λ, a larger D translates into a smaller R since $R = \Lambda/D$. Moreover, we have assumed a fixed cycle length of $D \cdot F = 400$ slots. For $D = 4$ we get $F = 100$ slots compared to $F = 200$ slots for $D = 2$. As a consequence, the length of data packets is decreased for increasing D. For $D = 4$ short data packets are $K = (F - M) = 40$ slots and long data packets are $F = 100$ slots long compared to $K = 140$ and $F = 200$ slots for $D = 2$, respectively. This results in a decreased mean aggregate throughput, as depicted in Fig. 34. Intuitively, we expect $D = 4$ to provide a smaller mean delay than $D = 2$ since in the former case each cycle contains as twice as many reservation slots as in the case $D = 2$ and the spatial reuse factor is doubled. We observe from Fig. 35 that this is true only for light to medium traffic loads. For a mean arrival rate larger than about 0.45 we can see that $D = 2$ provides smaller delays. This is due to the fact that each buffer can store up to 10 long data packets for $D = 2$ whereas for $D = 4$ each buffer is able to store up to 20 long data packets owing to the halved length of long data packets. As a consequence, for $D = 4$ at medium to high loads there are packets which have to wait for a longer time interval until they are transmitted resulting in an increased mean delay (and a decreased packet loss, as shown in Fig. 33). Again, we witness the trade–off between packet loss and delay. In the following simulations we set $D = 2$ for providing high throughput and low delay for a wide range of traffic loads. Note that the larger packet loss for $D = 2$ can be reduced by higher protocol layers. Protocols above the MAC layer have to make sure that packets are put in the MAC buffer such that overflow is completely avoided or kept under an acceptable level.

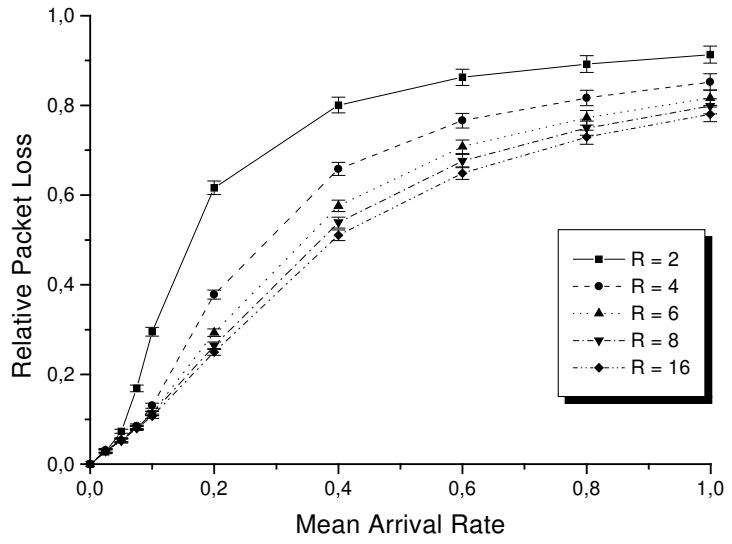

Fig. 24. Relative packet loss vs. mean arrival rate for $b = 4$, $M = 60$, and different $R \in \{2, 4, 6, 8, 16\}$.

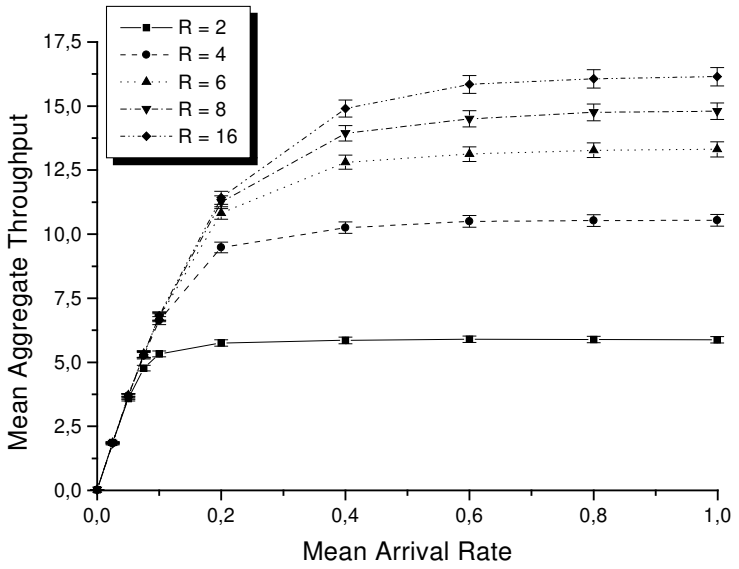

Fig. 25. Mean aggregate throughput (packets/frame) vs. mean arrival rate for $b = 4$, $M = 60$, and different $R \in \{2, 4, 6, 8, 16\}$.

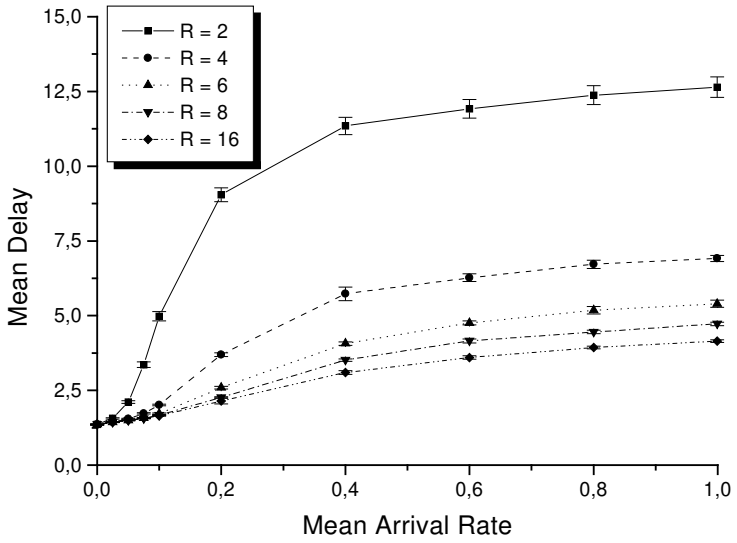

Fig. 26. Mean delay (cycles) vs. mean arrival rate for $b = 4$, $M = 60$, and different $R \in \{2, 4, 6, 8, 16\}$.

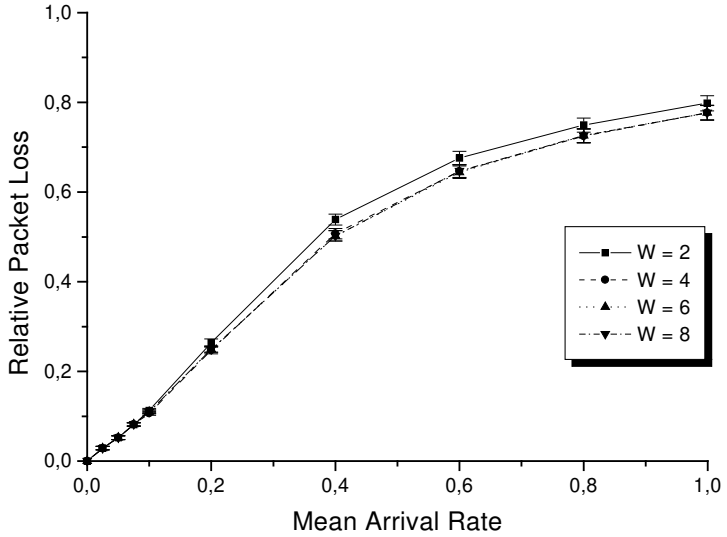

Fig. 27. Relative packet loss vs. mean arrival rate for $b = 4$, $M = 60$, $R = 8$, and different $W \in \{2, 4, 6, 8\}$.

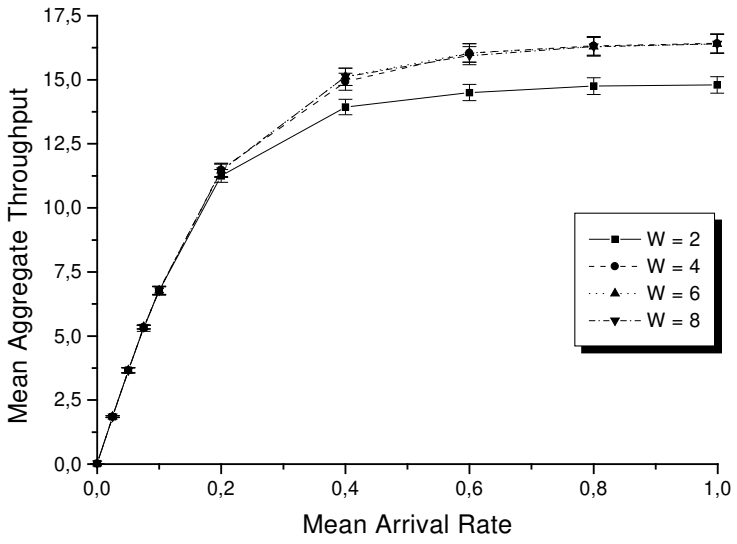

Fig. 28. Mean aggregate throughput (packets/frame) vs. mean arrival rate for $b = 4$, $M = 60$, $R = 8$, and different $W \in \{2, 4, 6, 8\}$.

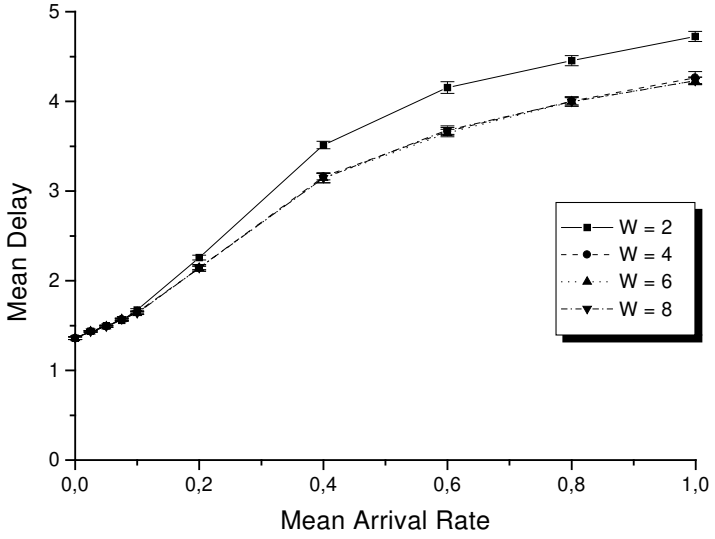

Fig. 29. Mean delay (cycles) vs. mean arrival rate for $b = 4$, $M = 60$, $R = 8$, and different $W \in \{2, 4, 6, 8\}$.

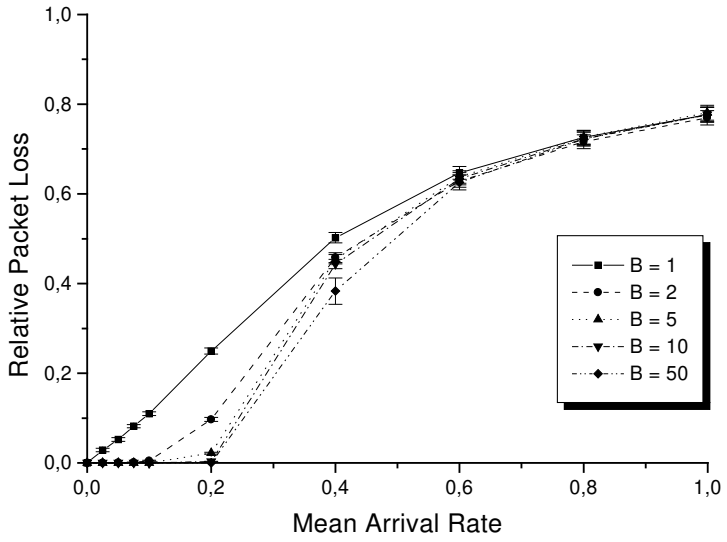

Fig. 30. Relative packet loss vs. mean arrival rate for $b = 4$, $M = 60$, $R = 8$, $W = 8$, and different $B \in \{1, 2, 5, 10, 50\}$.

C. Circuit switching

So far, we have considered only packet switching. We now take beside packet switching also circuit switching into account and show the positive impact of reservation ALOHA (R–ALOHA, see Section III) on the throughput. For this purpose, we introduce two additional parameters r and ρ:

- A generated data packet belongs to a circuit with probability r, where $0 \leq r \leq 1$. Note that the case considered so far where no packet sets up a circuit corresponds to $r = 0$.
- The circuit holding time is geometrically distributed with parameter ρ, where $0 \leq \rho < 1$. The parameter ρ denotes the probability that a given circuit continues in the next cycle. Thus, the average circuit holding time is equal to $\frac{1}{1-\rho}$ cycles. Note that $\rho = 0$ corresponds to packet switching.

In the following, the mean aggregate throughput counts for both packet and circuit–switched data packets. Whereas the mean delay takes only nodes into account which are involved in the reservation process. Nodes which have set up a circuit do not contribute to the mean delay. This is because these nodes use fixed assigned slots and do not encounter any delay caused by pretransmission coordination.

Fig. 36 shows the mean aggregate throughput vs. the mean arrival rate for $b = 4$, $M = 60$, $R = 8$, $W = 8$, $B = 10$, $r = 0.3$, and different $\rho \in \{0, 0.5, 0.8, 0.9, 0.95\}$. It can be seen that with increasing ρ the mean aggregate throughput increases. This is due to the fact that a larger ρ translates into an increased mean holding time. After making a successful setup each circuit is hold for a longer time interval resulting in a higher channel utilization and mean aggregate throughput. However, since each node keeping an active circuit indicates this by using one dedicated reservation slot (R–ALOHA, see Section III), for an increasing ρ reservation slots are busy for a longer time interval. Consequently, for increasing ρ nodes which try to make a reservation can access fewer reservation slots resulting in more collisions of control packets. This in turn leads to more retransmissions and an increased delay encountered by these nodes, as depicted in Fig. 37.

Fig. 38 illustrates the mean aggregate throughput vs. the mean arrival rate for $b = 4$, $M = 60$, $R = 8$, $W = 8$, $B = 10$, $\rho = 0.9$, and different $r \in \{0, 0.3, 0.6, 0.9\}$. Apparently, with circuit switching the wavelengths are used more efficiently than without leading to an increased mean aggregate throughput. This is because in R–ALOHA nodes which have successfully set up a circuit can send data packets in each cycle

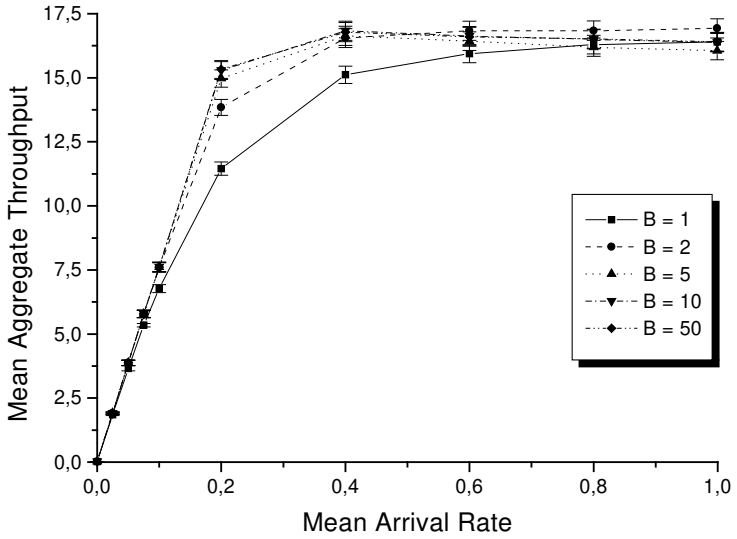

Fig. 31. Mean aggregate throughput (packets/frame) vs. mean arrival rate for $b = 4$, $M = 60$, $R = 8$, $W = 8$, and different $B \in \{1, 2, 5, 10, 50\}$.

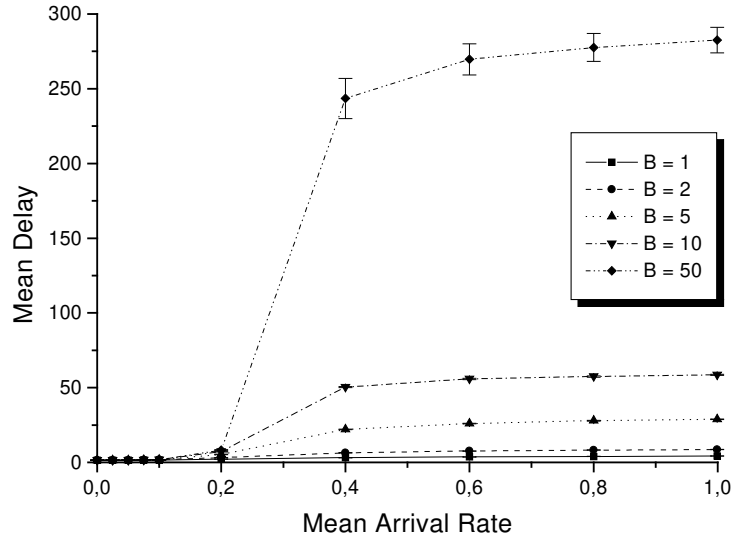

Fig. 32. Mean delay (cycles) vs. mean arrival rate for $b = 4$, $M = 60$, $R = 8$, $W = 8$, and different $B \in \{1, 2, 5, 10, 50\}$.

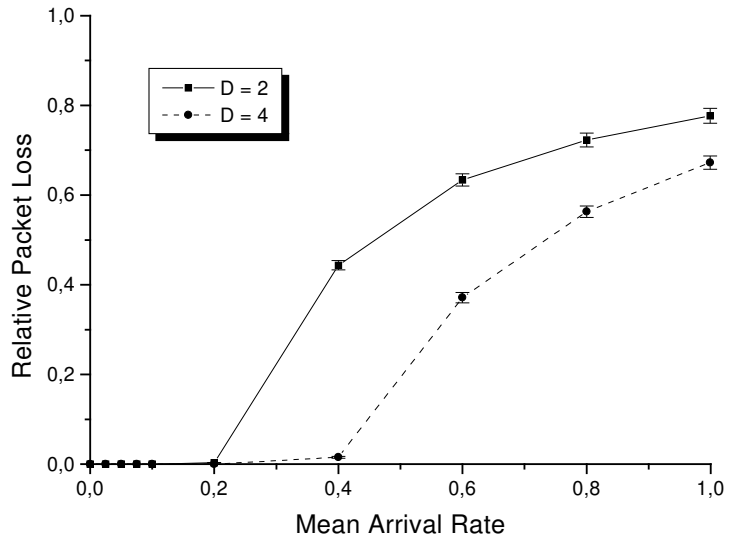

Fig. 33. Relative packet loss vs. mean arrival rate for $b = 4$, $M = 60$, $W = 8$, and different $D \in \{2, 4\}$.

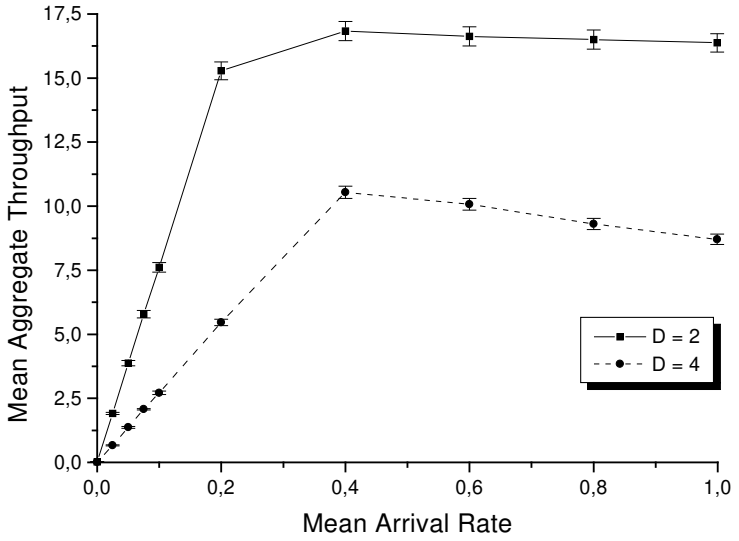

Fig. 34. Mean aggregate throughput (packets/frame) vs. mean arrival rate for $b = 4$, $M = 60$, $W = 8$, and different $D \in \{2, 4\}$.

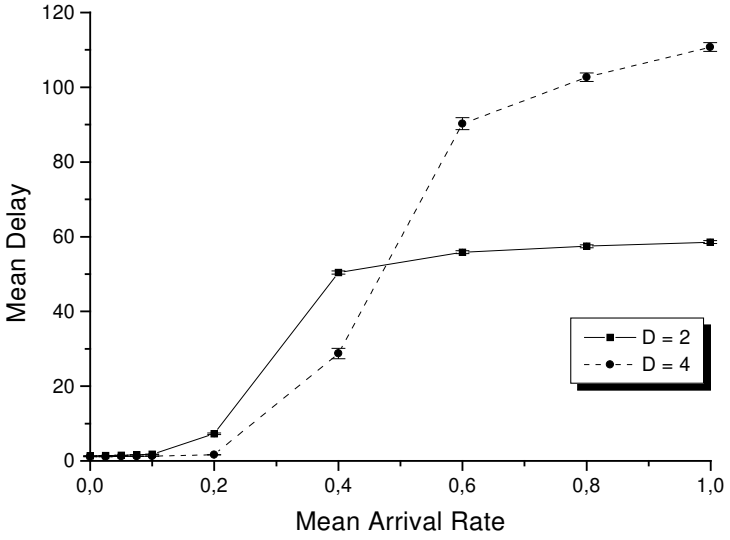

Fig. 35. Mean delay (cycles) vs. mean arrival rate for $b = 4$, $M = 60$, $W = 8$, and different $D \in \{2, 4\}$.

since the corresponding control packets do not suffer from channel collisions. In Fig. 39 we observe that with increasing r, i.e., an increasing fraction of circuit–switched data packets, the mean delay increases significantly. Again, this is because other nodes which try to make a reservation can access fewer reservation slots resulting in more control packet retransmissions and an increased mean delay. However, this is achieved at the expense of high delays for nodes trying to make a reservation, as shown in Figs. 37 and 39.

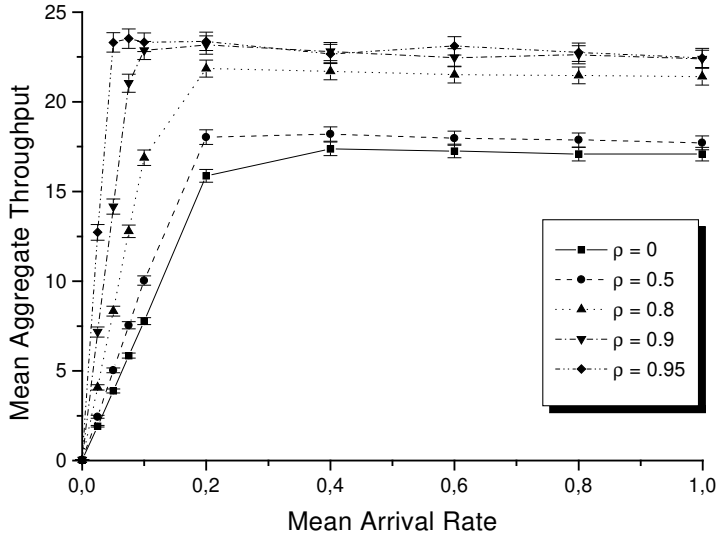

Fig. 36. Mean aggregate throughput (packets/frame) vs. mean arrival rate for $b = 4$, $M = 60$, $R = 8$, $W = 8$, $B = 10$, $r = 0.3$, and different $\rho \in \{0, 0.5, 0.8, 0.9, 0.95\}$.

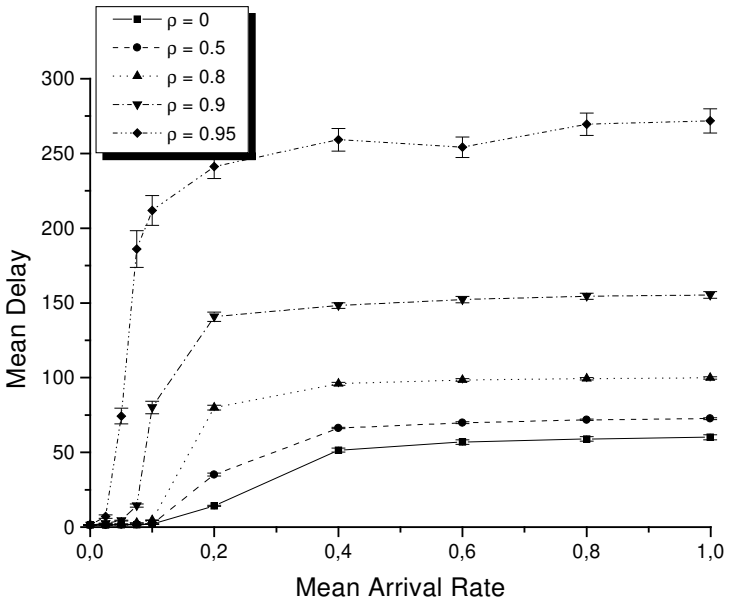

Fig. 37. Mean delay (cycles) vs. mean arrival rate for $b = 4$, $M = 60$, $R = 8$, $W = 8$, $B = 10$, $r = 0.3$, and different $\rho \in \{0, 0.5, 0.8, 0.9, 0.95\}$.

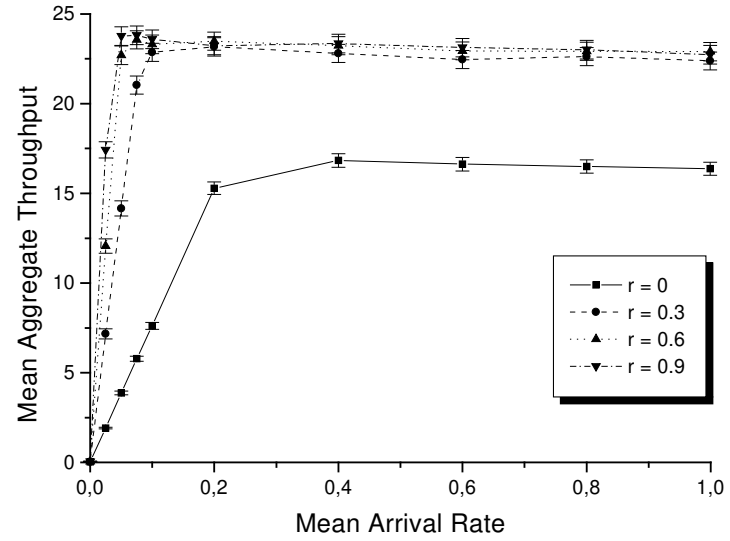

Fig. 38. Mean aggregate throughput (packets/frame) vs. mean arrival rate for $b = 4$, $M = 60$, $R = 8$, $W = 8$, $B = 10$, $\rho = 0.9$, and different $r \in \{0, 0.3, 0.6, 0.9\}$.

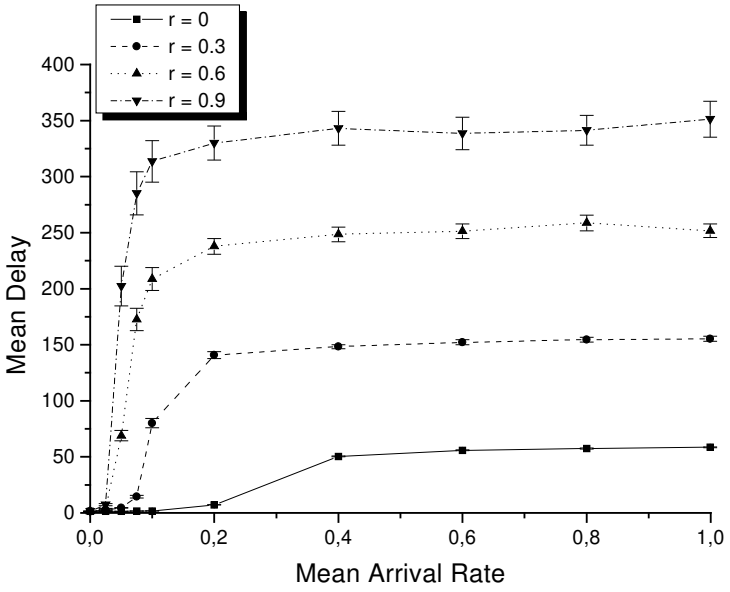

Fig. 39. Mean delay (cycles) vs. mean arrival rate for $b = 4$, $M = 60$, $R = 8$, $W = 8$, $B = 10$, $\rho = 0.9$, and different $r \in \{0, 0.3, 0.6, 0.9\}$.

D. Benchmark comparison

In this section, we discuss the efficiency of our AWG based network and MAC protocol and compare it with a previously reported single–hop metro WDM network that is based on a PSC. In the latter one nodes deploy the so–called *dynamic time–wavelength division multiaccess* (DT–WDMA) protocol for resolving packet collisions [8]. For convenience we briefly review the underlying architecture and the main features of DT–WDMA below. We have chosen DT–WDMA since among the MAC protocols designed specifically for multiwavelength single–hop WDM networks based on a PSC, DT–WDMA has perhaps been the *most influential* one and has spurred a number of research papers on modified and improved versions [57]. Moreover, DT–WDMA has the following properties in common with our protocol, which allow for a reasonably fair benchmark comparison:

- For data transmission/reception each node is equipped with *one single* transceiver.
- Each node's receiver is *tunable*. Consequently, receiver collisions can potentially occur and have to be resolved by the access protocol.
- DT–WDMA belongs to the category of *reservation* protocols.
- Resources are dynamically on–demand assigned on a *per–packet* basis.
- Nodes are able to acquire and maintain *global knowledge*.
- Explicit acknowledgements are *not* required.

Next, we briefly describe the network architecture and DT–WDMA (for detailed informations the interested reader is referred to [8]). DT–WDMA is proposed for a metropolitan–sized single–hop WDM network employing fixed–tuned and tunable transceivers attached to a PSC. More precisely, each node is equipped with one transceiver fixed tuned to a common control channel. For data each node deploys one transmitter fixed–tuned to a separate wavelength, i.e., each node has its own home channel for transmission, and one tunable receiver. Control information is sent over a dedicated signalling channel. Time is divided into slots (which correspond to frames in our protocol) on each channel and slots on the control channel are further split into mini–slots (which correspond to slots in our protocol). Fixed time–division multiaccess is used within each slot on the control channel where one mini–slot is dedicated to each node. Transmitters indicate

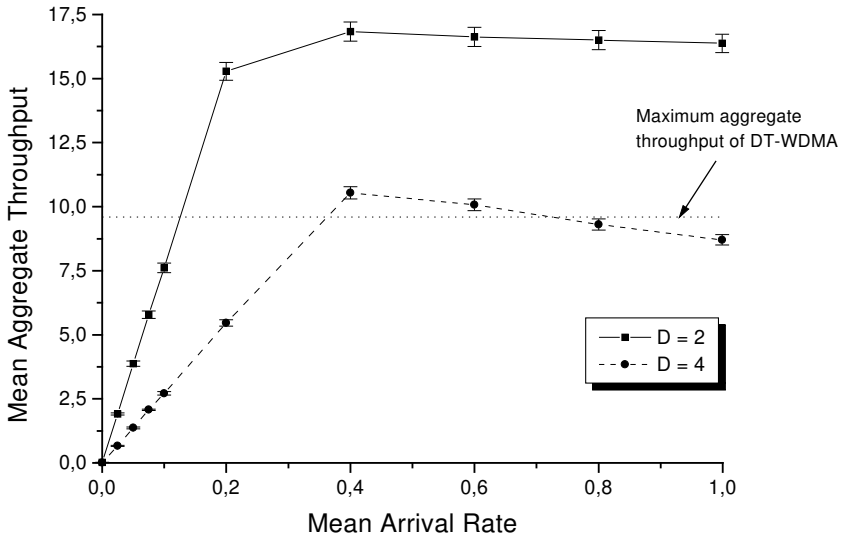

Fig. 40. Mean aggregate throughput (packets/frame) for $b = 4$, $M = 60$, $W = 8$, and $D \in \{2, 4\}$, compared to maximum aggregate throughput (packets/frame) of DT–WDMA vs. mean arrival rate.

their intention to transmit a packet by transmitting the destination address during their appropriate mini–slot in the control channel and then transmit their packet in the next slot on their data channel, i.e., DT–WDMA belongs to the family of tell–and–go reservation protocols. Receivers listen to the control channel and tune to the appropriate channel to receive packets addressed to them. A common distributed arbitration algorithm is used to resolve conflicts when packets from multiple transmitters contend for the same receiver. Each receiver executes the same deterministic algorithm to choose one of the contending packets. Each transmitter uses the same algorithm to determine the success or failure of its packet. This eliminates the need for explicit acknowledgements. In case of failure the corresponding node re–starts the reservation procedure.

Note that in DT–WDMA each node has its own home channel (wavelength) for transmission, i.e., the number of nodes equals the number of wavelengths, while receivers are assumed to be tunable over all these wavelengths. Thus, for large populations receivers with a large tuning range are required whose large tuning time significantly decreases the channel utilization. In our comparison we use 16 wavelengths which allows for deploying fast tunable receivers in DT–WDMA (and fast tunable transceivers in our network) whose tuning time is negligible. Furthermore, to compare our above simulation studies with DT–WDMA we also set the number of nodes to $N = 200$ in both the AWG and PSC based networks. To accommodate 200 nodes in DT–WDMA we let all data wavelengths be equally shared among the nodes while we assume that each of the 200 nodes has its own mini–slot on the control channel. In our comparison we consider only packet switching and focus on the throughput performance. It was shown in [8] that under the same assumptions on packet arrival process (Bernoulli) and traffic pattern (uniform traffic) as made in our above simulation studies the maximum utilization of each of the 16 wavelength channels equals 0.6 in DT–WDMA. This translates into a *maximum* aggregate throughput of $0.6 \cdot 16 = 9.6$ packets/slot, which is identical to 9.6 packets/frame for one slot in DT–WDMA corresponds to one frame in our protocol.

Fig. 40 depicts the mean aggregate throughput of our network vs. mean arrival rate for $D \in \{2, 4\}$ (as already shown in Fig. 34) and compares it with the maximum aggregate throughput of DT–WDMA. We observe that for $D = 4$ the mean aggregate throughput of our network is approximately equal to the maximum aggregate throughput of DT–WDMA at medium to high traffic loads. However, for $D = 2$ our network clearly outperforms DT–WDMA in terms of throughput. For a wide range of arrival rates

our proposed network provides a mean aggregate throughput that is about 70% larger than the maximum aggregate throughput of DT–WDMA.

Next, let us consider the efficiency of DT–WDMA and our protocol. As mentioned above, the maximum *wavelength utilization* of DT–WDMA is 60%. Fig. 40 shows that for $D = 2$ our protocol yields a mean aggregate throughput of up to 16.8 packets/frame. Note that this results in a mean wavelength utilization of more than 100%. To see this, recall that 16 wavelengths are deployed in both PSC and AWG based networks. However, due to spatial wavelength reuse the 2×2 AWG provides twice as many communication channels than the PSC leading to a larger aggregate throughput. With this in mind, we take a look at the *channel utilization* of both networks. Clearly, in DT–WDMA the channel utilization and the wavelength utilization are the same, since each wavelength creates one channel. Whereas in our 2×2 AWG based network each wavelength creates two channels resulting in a mean channel utilization of 53%, approximately. Recall that in our protocol wavelengths can not be spatially reused during the first M slots of each frame when nodes are involved in the reservation process. Correspondingly, our protocol can not fully capitalize on spatial wavelength reuse (in [58] we present a modified architecture which allows for *full* spatial wavelength reuse at any given time). However, we have seen in [59] that these M reservation slots can be used very efficiently for multicasting. Furthermore, with circuit switching the channel utilization of our network can be increased significantly, as illustrated in Figs. 36 and 38. Finally, we note that further channel utilization improvements could be achieved by replacing our simple first–come–first–served–first–fit algorithm with more efficient scheduling algorithms.

VII. Conclusion

We have analyzed the photonic switching of variable–size packets with spatial wavelength reuse in an AWG–based metro WDM network. We have obtained computationally efficient and accurate expressions for the throughput and delay in the network. Based on our analytical results we have conducted extensive numerical investigations of the performance characteristics of the network. We have also conducted extensive simulations to verify the accuracy of the analytical results. Our numerical results indicate that the AWG–based single–hop network, originally proposed in [7], can efficiently transport packets of different sizes. We found that spatial wavelength reuse is crucial for efficient photonic packet switching. Spatial wavelength reuse significantly increases the throughput while dramatically reducing the delay.

In our ongoing work we are studying the optimal trade–offs of the network parameters, e.g., the AWG degree that maximizes the throughput (and minimizes the delay) for a given number of nodes (and traffic load). We are also studying multicasting in the AWG–based metro single–hop network.

Appendix

A: Refined Approximation of $P(Z = k)$

In this appendix we derive a refined approximation for the distribution of Z, the number of successful control packets destined to a given AWG output port in a given frame. This refined approximation does not approximate the binomial distributions of the random variables Y_i^n (see Eqn. (4)) and Y_i^o (see Eqn. (7)) with Poisson distributions. From (13) we note that $P(X_i = 1) = P(Y_i = 1)$. Recalling that $Y_i = Y_i^n + Y_i^o$, we have

$$P(Y_i = 1) = P(Y_i^n = 1,\ Y_i^o = 0) + P(Y_i^n = 0,\ Y_i^o = 1). \tag{52}$$

By the independence of the Y_i^n's and Y_i^o's we have

$$P(Y_i = 1) = P(Y_i^n = 1)P(Y_i^o = 0) + P(Y_i^n = 0)P(Y_i^o = 1). \tag{53}$$

Hence, with (4) and (7) we obtain

$$P(Y_i = 1) \quad = \quad \eta\frac{\sigma}{M}\left(1 - \frac{\sigma}{M}\right)^{\eta-1}\left(1 - \frac{p}{M}\right)^{S-\eta} + \left(1 - \frac{\sigma}{M}\right)^{\eta}(S-\eta)\frac{p}{M}\left(1 - \frac{p}{M}\right)^{S-\eta-1} \tag{54}$$

$$= \frac{1}{M}\left(1-\frac{\sigma}{M}\right)^{\eta-1}\left(1-\frac{p}{M}\right)^{S-\eta-1}\left[\eta\sigma\left(1-\frac{p}{M}\right)+p(S-\eta)\left(1-\frac{\sigma}{M}\right)\right] \quad (55)$$

$$\approx \frac{S}{M}\left(1-\frac{\sigma}{M}\right)^{\nu S-1}\left(1-\frac{p}{M}\right)^{S(1-\nu)-1}\left[\nu\sigma\left(1-\frac{p}{M}\right)+p(1-\nu)\left(1-\frac{\sigma}{M}\right)\right] \quad (56)$$

$$=: \kappa, \quad (57)$$

where (56) follows by approximating η/S by its expectation ν. Thus, the refined approximation of the distribution of Z is given by

$$P(Z=k) = \binom{M}{k}\left(\frac{\kappa}{D}\right)^{k}\left(1-\frac{\kappa}{D}\right)^{M-k},\ k=0,1,\ldots,M. \quad (58)$$

A.1 Numerical Evaluation of Refined Approximation for $P(Z=k)$

In this section we evaluate the refined approximation for $P(Z=k)$ and compare it with the approximation (15), which we henceforth refer to as "unrefined". First, note that in the unrefined approximation, the binomial distribution $BIN(\eta,\ \sigma/M)$ is approximated by the Poisson distribution with parameter $\eta\sigma/M$ (and the $BIN((S-\eta),\ p/M)$ distribution is approximated by the Poisson distribution with parameter $(S-\eta)p/M$). Clearly, this approximation is accurate when (i) S is large, (ii) σ is small (and also p), and (iii) M is large. We also note that the evaluation of the refined approximation is computationally slightly more demanding, as it involves the evaluation of the expression in (56) as compared to the single exponential term in (15).

In Table I we compare the mean aggregate throughput TH_{net} obtained with unrefined approximation, refined approximation, and simulation for the typical network parameters chosen as default parameters in Section V (i.e., R = 2, q = 0.25, M = 30, D = 4, N = 200 (and thus S = 50), p = 0.8, F = 200, and K = 170). We observe that (i) the throughput obtained by both analytical approximations almost coincides, and (ii)

TABLE I

MEAN AGGREGATE THROUGHPUT TH_{NET} OBTAINED WITH UNREFINED APPROXIMATION, REFINED APPROXIMATION, AND SIMULATION FOR DEFAULT NETWORK PARAMETERS.

σ	0.02	0.04	0.1	0.2	0.5	1.0
TH_{net}, unrefined	0.886	1.77	4.29	7.32	8.45	8.10
TH_{net}, refined	0.888	1.77	4.29	7.37	8.52	8.16
TH_{net}, simul.	0.883	1.77	4.29	7.32	8.48	8.14

the analytical results match very well with the simulation results. (Similar observations hold for the average delay.) In fact, we found in our extensive numerical investigations that these two observations hold for all parameter values considered in Section V. In particular, both analytical approximations essentially coincide and match very well with the simulations for S values as small as 10. We also found that both analytical approximations essentially coincide and match very well with the simulations for M values as small as 15. Table II compares the throughput obtained with unrefined approximation, refined approximation, and simulation for M = 8 and K = 192 and all other network parameters at their default values. For this small M value, we observe that unrefined approximation and refined approximation give identical throughput for small σ values (σ = 0.02, 0.04), but differ for larger σ values. We also observe that the refined approximation matches the simulation results very well. In summary, we find that the unrefined approximation gives accurate results for a wide range of network parameters. The refined approximation is more accurate when M is small and σ is large, at the expense of a slightly more demanding computation. However, small M values typically give poor network performance, as is illustrated in Fig. 9, and may therefore not be desirable in practice.

TABLE II

MEAN AGGREGATE THROUGHPUT TH_{NET} OBTAINED WITH UNREFINED APPROXIMATION, REFINED APPROXIMATION, AND SIMULATION FOR M = 8.

σ	0.02	0.04	0.1	0.2	0.5	1.0
TH_{net}, unrefined	0.966	1.90	0.331	0.285	0.266	0.260
TH_{net}, refined	0.966	1.90	0.274	0.241	0.226	0.221
TH_{net}, simul.	0.966	1.89	0.272	0.242	0.224	0.223

B: PROOF OF INDEPENDENCE OF Y_i^n'S AND Y_i^o'S

Let $\eta = \eta(S)$ be a sequence of positive integers satisfying $\eta(S) \leq S$ and $\lim_{S\to\infty} \eta(S) = \infty$. Let $Y_i,\ i = 1, \ldots, M$, be the number of nodes (among all η nodes) that send a control packet in slot i.

Claim: For all nonnegative integers $k_1, \ldots, k_M$, we have

$$\lim_{S\to\infty} \frac{P(Y_1 = k_1, \ldots, Y_M = k_M)}{\prod_{i=1}^{M} \frac{e^{-\eta\frac{\tilde{\alpha}}{S}}\left(\frac{\eta\tilde{\alpha}}{S}\right)^{k_i}}{k_i!}} = 1, \tag{59}$$

i.e., the joint distribution of the random variables $Y_1, \ldots, Y_M$, approaches a product of M Poisson distributions with identical parameters $\eta\tilde{\alpha}/S$. In particular, $Y_1, \ldots, Y_M$, are asymptotically independent as $S \to \infty$.

Proof: Fix nonnegative integers $k_1, \ldots, k_M$, and assume that $\eta(S) \geq k_1 + \cdots + k_M$. The distribution of the random vector $(Y_1,\ Y_2,\ \ldots, Y_M,\ \eta - Y_1 - \cdots - Y_M)$ is multinomial with parameters $(\eta,\ \varrho, \ldots, \varrho, 1 - \varrho M)$, where $\varrho = \tilde{\alpha}/S\ (= \sigma/M)$ (we assume that $S > \tilde{\alpha}M$, so that $1 - \varrho M > 0$). Therefore,

$$P(Y_1 = k_1, \ldots, Y_M = k_M) = \frac{\eta!}{k_1! \cdots k_M! (\eta - k_1 - \cdots - k_M)!} \varrho^{k_1+\cdots+k_M} (1 - \varrho M)^{\eta - k_1 - \cdots - k_M} \tag{60}$$

$$= \frac{\varrho^{k_1+\cdots+k_M}}{k_1! \cdots k_M!} \frac{1}{(1 - \varrho M)^{k_1+\cdots+k_M}} e^{\eta \ln(1-\varrho M)} \prod_{j=0}^{k_1+\cdots+k_M-1} (\eta - j). \tag{61}$$

Hence,

$$\frac{P(Y_1 = k_1, \ldots, Y_M = k_M)}{\prod_{i=1}^{M} \frac{e^{-\eta\varrho}(\eta\varrho)^{k_i}}{k_i!}} = \frac{1}{(1 - \varrho M)^{k_1+\cdots+k_M}} e^{\eta(\ln(1-\varrho M)+\varrho M)} \prod_{j=0}^{k_1+\cdots+k_M-1} \frac{\eta - j}{\eta}. \tag{62}$$

Now note that

A)

$$\lim_{S\to\infty} \frac{1}{(1 - \varrho M)^{k_1+\cdots+k_M}} = 1 \tag{63}$$

since $\lim_{S\to\infty} \varrho(S) = \lim_{S\to\infty} \frac{\tilde{\alpha}}{S} = 0$.

B)

$$e^{\eta(\ln(1-\varrho M)+\varrho M)} = e^{\eta(-\varrho M + o(1/S^2) + \varrho M)} = e^{\eta o(1/S^2)} \overset{S\to\infty}{\longrightarrow} 1 \tag{64}$$

because $\eta \leq S$.

C) Since $\eta(S) \to \infty$ we have $\frac{\eta - j}{\eta} \to 1$ and therefore

$$\lim_{S\to\infty} \prod_{j=0}^{k_1+\cdots+k_M-1} \frac{\eta(S) - j}{\eta(S)} = 1. \tag{65}$$

This proves the claim.

ACKNOWLEDGMENT

We are grateful to Prof. Hans Mittelmann of Arizona State University for providing us with valuable background on numerical analysis.

REFERENCES

[1] C. Awduche and Y. Rekhter, "Multiprotocol lambda switching: Combining MPLS traffic engineering control with optical crossconnects," *IEEE Communications Magazine*, vol. 39, no. 3, pp. 111–116, Mar. 2001.

[2] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of Optical Burst Switching in IP–over–WDM networks," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 2062–2071, Oct. 2000.

[3] D. K. Hunter, M. H. M. Nizam, M. C. Chia, and I. Andonovic *et al.*, "WASPNET: a wavelength switched packet network," *IEEE Communications Magazine*, vol. 37, no. 3, pp. 120–129, March 1999.

[4] A. Jourdan, D. Chiaroni, E. Dotaro, G. J. Eilenberger, and *et al.*, "The perspective of optical packet switching in IP–dominant backbone and metropolitan networks," *IEEE Communications Magazine*, vol. 39, no. 3, pp. 136–141, March 2001.

[5] M. J. O'Mahony, D. Simeonidou, D. K. Hunter, and A. Tzanakaki, "The application of optical packet switching in future communication networks," *IEEE Communications Magazine*, vol. 39, no. 3, pp. 128–135, March 2001.

[6] B. Mukherjee, "WDM optical communication networks: Progress and challenges," *IEEE Journal on Selected Areas in Communications*, vol. 18, no. 10, pp. 1810–1824, Oct. 2000.

[7] M. Maier, M. Reisslein, and A. Wolisz, "High performance switchless WDM network using multiple free spectral ranges of an arrayed–waveguide grating," in *Proc. of SPIE Terabit Optical Networking: Architecture, Control, and Management Issues*, Boston, MA, Nov. 2000, pp. 101–112, Paper won the *Best Paper Award* of the conference.

[8] M-S. Chen, N.R. Dono, and R. Ramaswami, "A media–access protocol for packet–switched wavelength division multiaccess metroplitan area networks," *IEEE Journal on Selected Areas in Communication*, vol. 8, no. 6, pp. 1048–1057, Aug. 1990.

[9] 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*, vol. 11, no. 4, pp. 579–589, May 1993.

[10] A. Ganz and Y. Gao, "Time–wavelength assignment algorithms for high performance WDM star based systems," *IEEE Trans. on Commun.*, vol. 42, no. 2/3/4, pp. 1827–1836, Feb./March/April 1994.

[11] 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," *IEEE/OSA J. of Lightwave Technol.*, vol. LT–5, no. 12, pp. 1782–1794, Apr. 1987.

[12] M. J. Karol and B. Glance, "A collision–avoidance WDM optical star network," *Computer Networks and ISDN Systems*, vol. 26, pp. 931–943, Mar. 1994.

[13] H. C. Lin and C. H. Wan, "A hybrid multicast scheduling algorithm for single–hop WDM networks," *IEEE/OSA Journal of Lightwave Technology*, vol. 19, no. 11, pp. 1654–1664, Nov. 2001.

[14] J. Lu and L. Kleinrock, "A wavelength division multiple access protocol for high–speed local area networks with a passive star topology," *Performance Evaluation*, vol. 16, no. 1–3, pp. 223–239, Nov. 1992.

[15] N. Mehravari, "Performance and protocol improvements for very high speed optical fiber local area networks using a passive star topology," *IEEE/OSA J. of Lightwave Technol.*, vol. 8, no. 4, pp. 520–530, Apr. 1990.

[16] Y. Ofek and M. Sidi, "Design and analysis of hybrid access control to an optical star using WDM," in *Proc. of IEEE Infocom '91*, Bal Harbour, FL, May 1991, pp. 20–31.

[17] K. M. Sivalingam, "Design and analysis of a media access protocol for star coupled WDM networks with TT–TR architecture," in *Optical WDM Networks — Principles and Practice, Chapter 9*, K. M. Sivalingam and S. Subramaniam, Eds. 2000, Kluwer Academic Publishers.

[18] G. N. M. Sudhakar, N. D. Georganas, and M. Kavehrad, "Slotted Aloha and Reservation Aloha protocols for very high–speed optical fiber local area networks using passive star topology," *IEEE/OSA J. of Lightwave Technol.*, vol. 9, no. 10, pp. 1411–1422, Oct. 1991.

[19] B. Mukherjee, "WDM–based local lightwave networks part I: Single–hop systems," *IEEE Network Magazine*, vol. 6, no. 3, pp. 12–27, May 1992.

[20] M. Maier, M. Reisslein, and A. Wolisz, "Towards efficient packet switching metro WDM networks," *Optical Networks Magazine*, vol. 3, no. 6, pp. 44–62, November/December 2002.

[21] C. J. Chae, H. Park, and Y. Park, "Hybrid optical star coupler suitable for wavelength reuse," *IEEE Photonic Technology Letters*, vol. 10, no. 2, pp. 279–281, Feb. 1998.

[22] B. Kannan, S. Fotedar, and M. Gerla, "A two level optical star WDM metropolitan area network," in *Proc. of IEEE Globecom '94 Communications*, San Francisco, CA, Nov. 1994.

[23] M.W. Janoska and T.D. Todd, "Coupled reservation protocols for hierarchical single–hop photonic networks," *IEE Proceedings Communications*, vol. 144, no. 4, pp. 247–255, Aug. 1997.

[24] P.E. Green, L.A. Coldren, K.M. Johnson, J.G. Lewis, C.M. Miller, J.F. Morrison, R. Olshansky, R. Ramaswami, and E.H. Smithand, "All–optical packet–switched metropolitan-area network proposal," *IEEE/OSA Journal of Lightwave Technology*, vol. 11, no. 5, pp. 754–769, May/June 1993.

[25] E. Hall, J. Kravitz, R. Ramaswami, M. Halvorson, S. Tenbrink, and R. Thomsen, "The Rainbow–II gigabit optical network," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 5, pp. 814–823, June 1996.

[26] S. Yao, S. J. B. Yoo, and B. Mukherjee, "All–optical packet switching for metropolitan area networks: Opportunities and challenges," *IEEE Communications Magazine*, vol. 39, no. 3, pp. 142–148, March 2001.

[27] D. Stoll, P. Leisching, H. Bock, and A. Richter, "Metropolitan DWDM: A dynamically configurable ring for the KomNet field trial in Berlin," *IEEE Communications Magazine*, vol. 39, no. 2, pp. 106–113, Feb. 2001.

[28] W. Cho and B. Mukherjee, "Design of MAC protocols for DWADM–based metropolitan–area optical ring networks," in *Proceedings of IEEE Globecom 2001*, San Antonio, TX, Nov. 2001.

[29] C. S. Jelger and M. H. Elmirghani, "A simple MAC protocol for WDM metropolitan access ring networks," in *Proceedings of IEEE Globecom 2001*, San Antonio, TX, Nov. 2001.

[30] K. Bengi and H. R. van As, "Efficient QoS support in a slotted multihop WDM metro ring," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 1, pp. 216–227, Jan. 2002.

[31] M. D. Feuer, S. L. Woodward, C. F. Lam, and M. L. Boroditsky, "Upgradeable metro networks using frequency–cyclic optical add/drop," in *OFC 2001 Technical Digest, paper WBB5*, Anaheim, CA, Mar. 2001.

[32] F. Ruehl and T. Anderson, "Cost–effective metro WDM network architectures," in *OFC 2001 Technical Digest, paper WL1*, Anaheim, CA, Mar. 2001.

[33] D. Banerjee, J. Frank, and B. Mukherjee, "Passive optical network architecture based on waveguide grating routers," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 7, pp. 1040–1050, Sept. 1998.

[34] B. Glance, I. P. Kaminow, and R. W. Wilson, "Applications of the integrated waveguide grating router," *IEEE/OSA J. Lightwave Technol.*, vol. 12, no. 6, pp. 957–962, June 1994.

[35] B. Glance and M. Karol, "Large capacity multiaccess optical packet network," *IEEE Photonic Technology Letters*, vol. 6, no. 7, pp. 872–875, July 1994.

[36] A.M. Hill, S. Carter, J. Armitage, and M. Shabeer *et al.*, "A scalable and switchless optical network structure, employing a single 32×32 free–space grating multiplexer," *IEEE Photonics Technology Letters*, vol. 8, no. 4, pp. 569–571, Apr. 1996.

[37] D.K. Jung, S.K. Shin, C.H. Lee, and Y.C. Chung, "Wavelength–division–multiplexed passive optical network based on spectrum–slicing techniques," *IEEE Photonics Technology Letters*, vol. 10, no. 9, pp. 1334–1336, Sept. 1998.

[38] M. J. Spencer and M.A. Summerfield, "WRAP: A medium access control protocol for wavelength–routed passive optical networks," *IEEE Journal of Lightwave Technology*, vol. 18, no. 12, pp. 1657–1676, Dec. 2000.

[39] K. Bengi, "Performance of single–hop WDM LANs supporting real–time services," *Photonic Network Communications*, vol. 1, no. 4, pp. 287–301, Dec. 1999.

[40] A.M. Hill, M. Brierley, R.M. Percival, R. Wyatt, D. Pitcher, K.M. Ibrahim Pati, I. Hall, and J.-P. Laude, "Multi–star wavelength–router network and its protection strategy," *IEEE J. Select. Areas Commun.*, vol. 16, no. 7, pp. 1134–1145, Sept. 1998.

[41] M. C. Chia, D. K. Hunter, I. Andonovic, P. Ball, I. Wright, S. P. Ferguson, K. M. Guild, and M. J. O Mahony, "Packet loss and delay performance of feedback and feed–forward arrayed–waveguide gratings–based optical packet switches with WDM inputs outputs," *IEEE Journal of Lightwave Technology*, vol. 19, no. 9, pp. 1241–1254, Sept. 2001.

[42] K. V. Shrikhande, I. M. White, D. Wonglumsom, S. M. Gemelos, and *et al.*, "HORNET: A packet–over–WDM multiple access metropolitan area ring network," *IEEE J. on Sel. Areas in Commun.*, vol. 18, no. 10, pp. 2004–2016, Oct. 2000.

[43] I. M. White, K. Shrikhande, M. S. Rogge, M. Gemelos, and *et al.*, "Architecture and protocols for HORNET: A novel packet–over–WDM multiple–access MAN," in *Proceedings of IEEE Globecom*, San Francisco, CA, Nov./Dec. 2000.

[44] N. P. Caponio, A. M. Hill, F. Neri, and R. Sabella, "Single–layer optical platform based on WDM/TDM multiple access for large–scale 'switchless' networks," *European Trans. on Telecomm.*, vol. 11, no. 1, pp. 73–82, Jan./Feb. 2000.

[45] A. Okada, T. Sakamoto, Y. Sakai, and K. Noguchi *et al.*, "All–optical packet routing by an out–of–band optical label and wavelength conversion in a full–mesh network based on a cyclic–frequency AWG," in *Proc. of OFC 2001 Technical Digest, paper ThG5*, Anaheim, CA, Mar. 2001.

[46] K. Kato, A. Okada, Y. Sakai, and K. Noguchi *et al.*, "10–Tbps full–mesh WDM network based on a cyclic–frequency arrayed–waveguide grating router," in *Proc. of ECOC '00*, Munich, Germany, Sept. 2000, vol. 1, pp. 105–107.

[47] C. Dragone, "Optimum design of a planar array of tapered waveguides," *J. Opt. Soc. Amer.*, vol. 7, pp. 2081–2093, Nov. 1990.

[48] C. Dragone, C.A. Edwards, and R.C. Kistler, "Integrated optics N×N multiplexer on silicon," *IEEE Photon. Techno. Lett.*, vol. 3, no. 10, pp. 896–899, Oct. 1991.

[49] P. Bernasconi, C. Doerr, C. Dragone, and M. Cappuzzo *et al.*, "Large $n \times n$ waveguide grating routers," *IEEE/OSA Journal of Lightwave Technology*, vol. 18, no. 7, pp. 985–991, July 2000.

[50] Y. Hibino, "An array of photonic filtering advantages," *Circuits and Devices*, pp. 23–27, Nov. 2000.

[51] Y. Tachikawa, Y. Inoue, M. Ishii, and T. Nozawa, "Arrayed waveguide grating multiplexer with loop–back optical paths and its applications," *Journal of Lightwave Technology*, vol. 14, no. 6, pp. 977–984, June 1996.

[52] K.A. McGreer, "Arrayed waveguide gratings for wavelength routing," *IEEE Communications Magazine*, vol. 36, no. 12, pp. 62–68, Dec. 1998.

[53] N. Keil, H. H. Yao, C. Zawadzki, J. Bauer, and *et al.*, "Athermal polarization–independent all–polymer arrayed waveguide grating (AWG) multi/demultiplexer," in *OFC 2001 Technical Digest, paper PD7*, Anaheim, CA, Mar. 2001.

[54] L. Giehmann, A. Gladisch, N. Hanik, and J. Rudolph, "The application of code division multiple access for transport overhead information in transparent optical networks," in *OFC 1998 Technical Digest, paper WM42*, San Jose, CA, Feb. 1998, pp. 228–229.

[55] L. Giehmann, A. Gladisch, and J. Rudolph, "Field trial of OAM–signal transport capabilities with a 10 Mchip/s LED–direct sequence spread spectrum systems suited for OAM–signal–transport in transparent optical WDM–networks," in *Proc. of OFC '99, paper TuR3*, 1999.

[56] V. K. Jain and G. DeMarchis, "Performance evaluation of optical code division multiple access networks," *Journal of Optical Communications*, vol. 21, no. 3, pp. 110–115, 2000.

[57] R. Ramaswami and K. N. Sivarajan, *Optical Networks — A Practical Perspective*, Morgan Kaufmann, 1998.

[58] C. Fan, M. Maier, and M. Reisslein, "The AWG||PSC network: A performance enhanced single–hop WDM network with heterogeneous protection," Tech. Rep., Arizona State University, Telecommunications Research Center, July 2002, Available at `http://www.eas.asu.edu/~mre`.

[59] M. Maier, M. Scheutzow, and M. Reisslein, "The arrayed-waveguide grating based single-hop wdm network: An architecture for efficient multicasting," Tech. Rep., Telecommunications Research Center, Arizona State University, Dec. 2002.