

# Retrial Queuing Models of Multi-Wavelength FDL Feedback Optical Buffers

Nail Akar, *Member, IEEE*, and Khosrow Sohraby, *Senior Member, IEEE*

**Abstract**—Optical buffers based on Fiber Delay Lines (FDL) have been proposed for contention resolution in optical packet/burst switching systems. In this article, we propose a retrial queuing model for FDL optical buffers in asynchronous optical switching nodes. In the considered system, the reservation model employed is of post-reservation type and optical packets are allowed to re-circulate over the FDLs in a probabilistic manner. We combine the MMPP-based overflow traffic models of the classical circuit switching literature and fixed-point iterations to devise an algorithmic procedure to accurately estimate blocking probabilities as a function of various buffer parameters in the system when packet arrivals are Poisson and packet lengths are exponentially distributed. The proposed algorithm is both accurate and fast, allowing one to use the procedure to dimension optical buffers in next-generation optical packet switching systems.

**Index Terms**—Fiber Delay Line (FDL) re-circulation buffer, optical packet switching, optical burst switching, Markov modulated Poisson process, retrial queues.

## I. INTRODUCTION

TWO packet switching-based paradigms have emerged for flexible use of bandwidth in optical networks, namely Optical Packet Switching (OPS) [1] and Optical Burst Switching (OBS) [2]. In this article, we analytically study the performance of an OPS/OBS node, hereafter called OPR (Optical Packet Router), that uses FDLs (Fiber Delay Line) for optical buffering irrespective of whether the router supports the OPS or OBS paradigm. OPS/OBS can either be synchronous (time-slotted) in which packets have either fixed sizes or variable sizes that are integer multiples of a time unit (called a slot) or asynchronous (unslotted) with arbitrarily variable packet lengths. Most of current literature on OPS is based on synchronous switching [3] whereas asynchronous OBS is the most dominant mode of OBS operation [2]. However, exceptions exist in the literature; see [4],[5]. The current article focuses on asynchronous OPRs only and the use of FDL buffers in overcoming the performance limitations of asynchronous switching systems.

In optical packet switching-based (OPS or OBS) networks, contention arises when two or more incoming optical packets contend for the same output port/wavelength pair. If the wavelength of an incoming packet is occupied at the destination

port, then its wavelength can be converted to one which is idle [6]. If all wavelength channels at a given port are occupied, then incoming packets can use optical buffering to resolve contention [7]. If the capabilities of the OPR (in terms of wavelength converters and buffers) come short of resolving the contention, deflection routing can be used to send one or more of the contending packets to a different output port [2]. In this article, we study the combined use of wavelength converters and optical buffers for contention resolution. In electronic buffering, packets are stored in RAMs (Random Access Memory) and wait for their turns for transmission until they reach the head of their corresponding queues. Since optical RAMs are not feasible today, the most popular optical buffering technique is the use of FDLs in which a contending packet is sent over a coil of fiber that provides the packet with a fixed delay in order to resolve contention. Such a mechanism is referred to as FDL buffering. The alternative mechanism of SLDLs (Slow Light Delay Lines) [8] that promise variable and controllable delays are outside the scope of this article.

There are different architectures proposed for using FDL buffers; we refer the reader for an extensive discussion of such architectures in [9],[10],[11]. Buffer architectures are generally classified as feed-forward (FF) or feedback (FB) [12]. In feed-forward buffering, optical packets are delayed at the output ports so as to leave the node after traversal through the FDL. On the other hand, in feedback buffering, optical packets are delayed while being fed back to the node after the propagation delay of the FDL. FDL buffers can also be classified based on their lengths; the Fixed-length FDL (F-FDL) architecture uses FDLs of the same length and therefore all FDLs realize the same delay  $D$  making it necessary for the packets to re-circulate if longer delays are required. On the other hand, Increasing-length FDL (I-FDL) architecture comprises FDLs of increasing length which are generally taken to be integer multiples of a basic granularity unit, i.e.,  $D, 2D, \dots, BD$ , where  $B$  denotes the buffer size [3],[13]. In this regard, F-FDLs present a better fit for the FB architecture whereas FF architecture is more likely to use I-FDLs. Another classification is made by [12] based on the type of reservation; if the output port is reserved prior to entering the buffer, this architecture is called Pre-Reservation (PreRes). If the output reservation is attempted at the end of each FDL re-circulation of the packet, the architecture is called Post-Reservation (PostRes). In PostRes, all blocked packets will use buffers if available whereas in PreRes, a blocked packet will only be admitted into the buffer only when an output port can be reserved at the epoch of exit from the buffer. One other classification can also be made based on whether buffers are shared or dedicated for each output port.

Paper approved by W. Shieh, the Editor for Optical Transmission and Switching of the IEEE Communications Society. Manuscript received August 26, 2010; revised January 25, 2011.

N. Akar is with the Electrical and Electronics Engineering Dept., Bilkent University, Bilkent 06800, Ankara, Turkey (e-mail: akar@ee.bilkent.edu.tr).

K. Sohraby is with the School of Computing and Engineering, University of Missouri-Kansas City, 5100 Rockhill Road, Kansas City, MO 64110-2499, USA (e-mail: sohrabyk@umkc.edu).

Digital Object Identifier 10.1109/TCOMM.2011.071111.100521

Our goal is to introduce a new generic queuing model for FDL buffers based on [14] as opposed to modeling one of the specific architectures already proposed in the literature. By doing so, we not only aim to gain insight on the operational principles of optical buffers but also our proposed models can be used as building blocks for analytical modeling of more general and realistic architectures. For this purpose, we focus on a single output port of an OPR with  $K$  available wavelength channels and  $B$  FDLs of the same length  $D$ . We assume exponentially distributed packet lengths with mean set to unity, i.e., the time unit is taken to be mean packet length, and Poisson packet arrivals with rate  $\lambda$  at the designated output port. The system load is denoted by  $\rho = \lambda/K$ . The assumptions of Poisson call arrivals and exponentially distributed call-holding times have been successfully used for circuit-switched networks handling call-oriented traffic. In our generic model, an optical packet which finds all  $K$  wavelength channels occupied is said to be blocked. A blocked packet will always use a buffer (if available) to retry entry into the output port. Such behavior is consistent with the PostRes model which is simpler to implement than the PreRes scheme since the decision of buffer admission only depends on whether any one of the  $B$  FDLs is available at the epoch of arrival and not on the availability information at the epoch of exit from the FDLs. A packet is said to be a retrial packet once it traverses one of the  $B$  FDLs. A retrial packet will use one of the free wavelength channels if available at the epoch of exit from the FDL. Otherwise, it will attempt to re-circulate through one of the FDLs with probability  $\kappa \leq 1$  and otherwise with probability  $1 - \kappa$ , it will be discarded, i.e., lost. Such a probabilistic re-circulation policy not only limits the average number of re-circulations in the system but is also amenable to analytical modeling. Although a deterministic policy of allowing at most a given number of re-circulations is more common in the literature [15], we leave analytical modeling of such systems for future research. In this setting, the problem at hand provides a model for feedback FDL buffers with (i) F-FDLs, (ii) PostRes reservation model (iii) dedicated output buffers. The reason behind the interest in this particular system is its relative simplicity since the switch controller only needs to keep track of binary variables that are indicative of the occupancies of wavelength channels as well as the occupancies of the FDLs at their entrance points. Note that in this model, two or more packets can use the same FDL at the same time using spatial multiplexing.

Our goal in this paper is to introduce a simple queuing model for FDL optical buffers that accurately capture the behavior of this system for which related work already exists in the literature. Let us first start with the particular case of  $\kappa = 0$  which is studied more in the context of feed-forward I-FDL buffers using the PreRes reservation model which amounts to the balking property of FDLs, i.e., a packet will not be admitted into the FDL buffer if the maximum delay provided by all available FDLs is not sufficient to reserve an output port. Most of the research in this area focuses on the single-wavelength case, i.e.,  $K = 1$ ; the reference [13] presents an approximate method using an iterative procedure which is simple to implement for Poisson arrivals and exponential packet lengths. Again, for the same scenario, the work in

[16] provides closed-form expressions for loss probabilities and expected delays. Similarly, the authors derive closed-form expressions in [17] that allow one to optimize the fiber lengths of an optical buffer for variable length packets. For the same I-FDL type feed-forward buffer, [18] relaxes the traffic models for interarrivals and service times while still being able to provide an accurate approximative method which turns out to be exact in some certain sub-cases. Recently, [19] proposed an exact analytical procedure for the single-wavelength case for I-FDL feed-forward buffers using the paradigm of feedback fluid queues even with more general arrival processes such as Markovian Arrival Process (MAP) and phase-type distributed packet lengths. The reference [15] introduces a queuing model for a single-wavelength feedback optical buffer when there is a limit on the number of FDL circulations. Results associated with  $\kappa = 0$  but for the multi-wavelength case, i.e.,  $K > 1$ , are rather rare. One of the earlier works in this area is [14] that uses an M/M/K/K+B queuing model with  $K$  servers and  $B$  additional waiting places whose main assumption is that a stored burst can immediately start to use a channel whenever it becomes available. However, FDL delays are deterministic and a packet cannot leave the FDL buffer arbitrarily. Therefore, the proposed method only provides a lower bound for the loss probability. With this observation in place, [20] provides upper and lower bounds for the loss probability for a related system of interest using I-FDLs via the Erlang loss formula (no buffers) and the method of [14], respectively. On the other hand, [21] proposes an approximate Markovian model to capture the balking property but in a system with variable delays which is more of a property of SLDLs. The closest work to ours in the literature is [12] which provides a simulation-based study of I-FDL feed-forward PreRes buffers in comparison with a system employing F-FDL feedback PostRes buffers. Our goal in this article is to introduce a queuing model for the latter which is computationally efficient even for systems with large number of wavelengths and FDLs.

Retrial queuing systems have extensively been studied in the literature; see [22],[23],[24] for an exhaustive survey of the literature on retrial queues. An M/M/K retrial queue comprises a  $K$ -server queuing system in which the primary customers arrive according to a Poisson flow and service times are exponentially distributed. A primary customer finding a server free will occupy the server and leave the system after the service time. On the other hand, a primary customer finding all servers busy upon arrival will join a so-called orbit so as to repeat his demand after an exponential time with parameter  $\mu$ . Typically, orbit capacities are infinite and such a system is well described by a bi-variate process  $\{(C(t), N(t)), t \geq 0\}$ , where  $C(t)$  is the number of busy servers at time  $t$  and  $N(t)$  is the number of customers in the orbit at time  $t$  [22]. Classical Markov chain solution techniques do not apply since the orbit capacity is infinite and most existing literature on M/M/K retrial queues addresses accurate and approximate calculation of the steady-state probabilities of this bi-variate process. Variations of this basic model are available such as a retrial finding the system full will re-enter orbit with probability  $\kappa$  and leave the system forever with probability  $1 - \kappa$  independent of the number of trials the customer has

already made [24]. Although most well-known methods rely on the solution of large Markov chains, there are also relatively simple and easy-to-implement methods for approximating system behavior such as RTA (Retrials see Time Averages) or FRA (Fredericks and Reisner Approximation) [24],[25]. There are several differences between the classical retrial queue and the FDL buffer:

- The orbit time is exponential in retrial queues whereas it is deterministic in PostRes FDL buffers,
- The orbit capacity is typically infinite in retrial queues whereas we have a finite number  $B$  of FDLs in FDL buffers,
- Even in case the orbit capacity can be taken to be finite, say  $J$ , in a retrial queue, the value  $J$  limits the number of customers in the orbit. However, the value  $B$  limits the number of FDLs and not the number of packets traveling through the FDLs.

Although there are differences between the two models, we introduce a queuing model inspired by retrial queues in the current article. The main idea is to characterize the retrying traffic with an MMPP (Markov Modulated Poisson Process) and feeding back this process to the system using an iterative procedure. While doing so, we benefit from MMPP-based overflow models of the classical circuit-switching literature [26],[27]. With the model at hand, we address the provisioning problem of selecting the FDL delay  $D$  as well as the buffer size  $B$  and re-circulation parameter  $\kappa$  that meets delay, loss, and re-circulation requirements.

The remainder of this article is organized as follows. Section 2 presents the queuing model we propose for multi-wavelength feedback FDL buffers. In Section 3, we provide numerical examples for validating the accuracy of this approach as well as the use of these models for engineering and dimensioning purposes. Finally, we conclude.

## II. ANALYTICAL MODEL

Recall that we focus on a single output port of an OPR with  $K$  available wavelength channels and  $B$  FDLs of the same length  $D$  and we use the PostRes reservation model. Fig. 1 illustrates the operation for an FDL buffer with  $K = 2$  channels and  $B = 2$  FDLs with each FDL being of length  $D = 4$  time units and no re-circulations, i.e.,  $\kappa = 0$ . For illustrative purposes, all packets have the same size which is three time units. Consider the situation depicted in Fig. 1(a) at  $t = 0^-$ . On channel 1, packet 1 is under transmission upon a repeated attempt whereas on channel 2, packet 2 is under transmission upon its first attempt, therefore this packet is not using any FDL resources. The FDL 1 has the tail of packet 1 at its head but another packet 4 has just started to traverse through FDL 1 since it arrived one time unit ago but found both channels occupied. On FDL 2, packet 3 has just completed its journey on FDL 2 and is about to look for an opportunity to join one of the two wavelength channels. Since  $\kappa = 0$  and re-circulations are not allowed for this example and both channels are busy at  $t = 0^+$ , this packet 3 will be dropped by the system leading to packet loss. Note that such a packet drop is specific to the PostRes reservation model since the output channel would not be reserved immediately when the packet

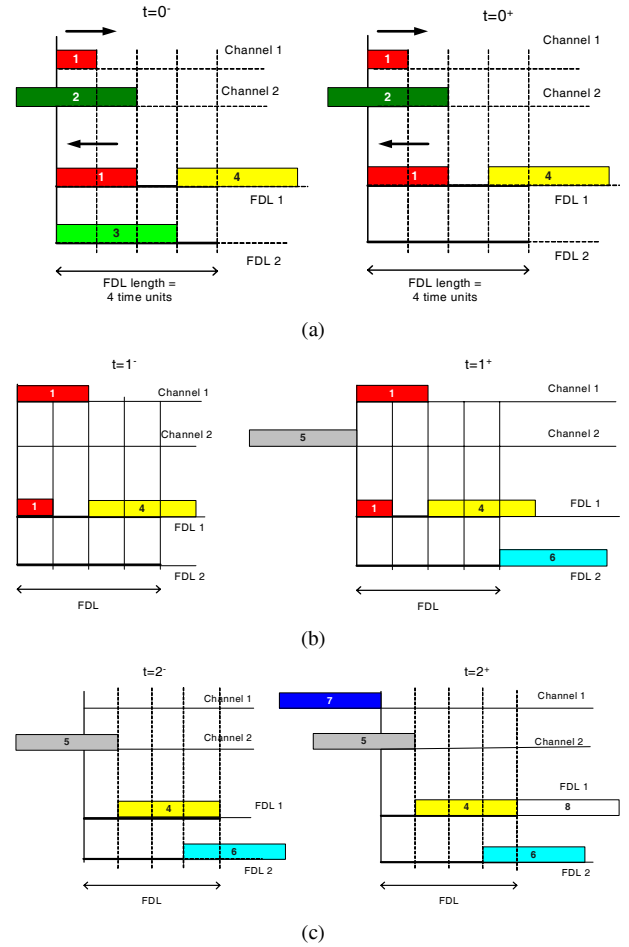


Fig. 1. The evolution of a feedback PostRes FDL buffer with  $K = 2$  wavelength channels and  $B = 2$  FDLs each with length  $D = 4$  time units where all packets have a length of three time units at three different time epochs (a)  $t = 0$ , (b)  $t = 1$ , (c)  $t = 2$ . For a given time epoch  $\tau$ , packet arrivals are said to occur at  $\tau^-$  and at time epoch  $\tau^+$ , decisions for these packets, i.e., whether they are admitted to the wavelength channels or admitted to the FDL buffer or they are simply discarded, are made.

is accepted into the buffer in the PostRes model. Due to this packet drop, the system state is illustrated at  $t = 0^+$  in the right figure of Fig. 1(a). We let the system evolve until  $t = 1^-$  when two packets (numbered 5 and 6) arrive. The system state just before the arrivals is given in Fig. 1(b). Compared to Fig. 1(a), packet 2 has just left channel 2, the transmission of packet 1 over channel 1 and that of packet 4 over FDL 1 are in progress at  $t = 1^-$ . Since at  $t = 1^-$ , channel 2 is idle, one of the arriving packets, say packet 5, would be accepted into channel 2, and the other arrival 6 would be directed to the idle FDL 2. The system state just after the arrivals is illustrated in the right figure of Fig. 1(b). We now let the system evolve until  $t = 2^-$  when three packets (numbered 7, 8, and 9) arrive. At  $t = 2^-$ , packet 7 is assigned to channel 1, packet 8 is directed to FDL 1, but packet 9 finds both servers and FDLs busy and is dropped. The system state before and after these three arrivals is illustrated in Fig. 1(c). For this particular example, we used fixed packet sizes and instantaneous packet arrivals as in slotted systems only for illustrative purposes but the analytical model will be developed for asynchronous systems.

Based on the operational principles outlined in Fig. 1, it

is not possible to exactly represent the system as a Markov process as in retrial queues. Therefore, it is clear that approximations are required for analyzing this system. Our approach is based on the stochastic characterization of the traffic coming out of the FDL buffer based on MMPP (Markov Modulated Poisson Process)-based models that have successfully been used in modeling overflow traffic in circuit-switched networks. An MMPP is characterized with a matrix pair  $(Q, R)$  where  $Q$  is the infinitesimal generator matrix of a modulating finite-state continuous-time Markov chain and  $R = \text{diag}\{\lambda_i\}$  is a diagonal matrix of rates where  $\lambda_i$  denotes the Poisson rate of arrivals when the state  $i$  of the modulating chain is visited [28]. For more details on the properties of MMPPs, we refer the reader to [29]. Of particular interest to this paper is a two-state MMPP which has two states 1 and 2 with arrival rates  $\lambda_1$  and  $\lambda_2$  in these two states, respectively. The transition rate from state 1 to state 2 (state 2 to state 1) is denoted by  $\sigma_1$  ( $\sigma_2$ ). The two-state MMPP is then said to be characterized by the ordered quadruple  $(\lambda_1, \lambda_2, \sigma_1, \sigma_2)$ . In this paper, we assume that the traffic coming out of the FDL buffer is modeled with a two-state MMPP called  $X$  comprising two states, namely the HIGH and LOW states, which is characterized by the quadruple  $(r_H, r_L, \eta, \gamma)$ . In the HIGH state, packet arrivals are Poisson with rate  $r_H$  and in the LOW state with rate  $r_L$  such that  $r_H \geq r_L$ . The two-state MMPP  $X$  is helpful in representing the burstiness of this particular traffic stream while maintaining simplicity. We do not know the parameters of MMPP  $X$  yet but we will obtain these parameters as an outcome of an iterative process. However, this MMPP process is not independent from the channel occupancy process that keeps track of the number of busy channels at the designated port. This stems from the observation that a retrying optical packet will see a system quite correlated with the one the same packet had attempted to join but failed  $D$  time units back. For the purpose of characterizing this inter-dependence, let  $p_{i,H}$  ( $p_{i,L}$ ) for  $0 \leq i \leq K$  denote the joint probability that the two-state MMPP is in state HIGH (LOW) and there are  $0 \leq i \leq K$  occupied channels at the designated port at an arbitrary epoch. Similarly, let  $\pi_{i,H}$  ( $\pi_{i,L}$ ) denote the probability that the two-state MMPP is in state HIGH (LOW) and there are  $0 \leq i \leq K$  occupied channels at the designated port at the epoch of an arbitrary retrial. Again, we do not know the probabilities  $p_{i,H}$ ,  $p_{i,L}$ ,  $\pi_{i,H}$ , and  $\pi_{i,L}$  at this point but we will obtain these probabilities through the same iterative procedure. Moreover, we define

$$p_H = \sum_{i=0}^K p_{i,H}, p_L = \sum_{i=0}^K p_{i,L}, \pi_H = \sum_{i=0}^K \pi_{i,H}, \pi_L = \sum_{i=0}^K \pi_{i,L}.$$

Let us now define the bi-variate process  $\mathcal{X}(t) = \{(C_P(t), S_P(t)); t \geq 0\}$ , where  $0 \leq C_P(t) \leq K$  denotes the number of busy channels at the designated port and  $S_P(t) \in \{H, L\}$  denotes the state of the two-state MMPP  $X$  (whether it is in the HIGH or LOW state) at time  $t$ . The rate of repeated (i.e., retrying) arrivals when  $C_P(t) = i$  and  $S_P(t) = H$  is denoted by  $r_H^{(i)}$ ,  $0 \leq i \leq K$ . Note that  $r_H^{(i)}$  should be proportional with  $\pi_{i,H}/\pi_H$  and moreover it should hold that  $p_H r_H = \sum_{i=0}^K p_{i,H} r_H^{(i)}$ . From these two

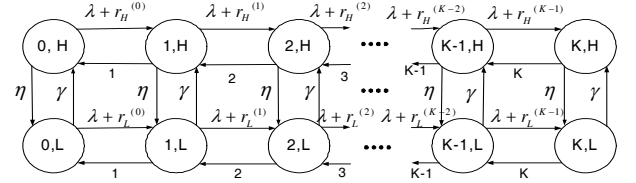


Fig. 2. The transition diagram of the Markov chain governing the behavior of the bi-variate process  $\mathcal{X}(t)$ .

observations, we conclude that

$$r_H^{(i)} = r_H \frac{p_H \pi_{i,H}}{\pi_H p_{i,H}}. \quad (1)$$

Similarly, if we define  $r_L^{(i)}$ ,  $0 \leq i \leq K$ , to be the rate of repeated arrivals when  $C_P(t) = i$  and  $S_P(t) = L$ , we obtain

$$r_L^{(i)} = r_L \frac{p_L \pi_{i,L}}{\pi_L p_{i,L}}. \quad (2)$$

Assuming that retrials at a given state of  $\mathcal{X}(t)$  are of Poisson nature, we have a continuous-time Markov chain of size  $2(K+1)$  whose transition diagram is given in Fig. 2 which approximates the behavior of the occupancy process of the port, called hereafter the port process.

By definition, the probabilities we defined  $p_{i,H}$  and  $p_{i,L}$  for  $0 \leq i \leq K$  are the steady-state probabilities of the port process described by its state-transition diagram in Fig. 2. Let  $P$  denote the infinitesimal generator of this Markov chain with the states enumerated as

$$\{(0, H), (1, H), \dots, (K, H), (0, L), \dots, (K, L)\}.$$

Packets are blocked and directed to the FDL buffer only in the two states  $(K, H)$  and  $(K, L)$ . In particular, the rate of traffic directed to the FDL buffer is  $\lambda + r_H^{(K)} \kappa$  at state  $(K, H)$  and  $\lambda + r_L^{(K)} \kappa$  at state  $(K, L)$ . Assuming an MMPP model, the traffic comprising packets directed to the FDL buffer can be modeled with an MMPP with  $2(K+1)$  states with infinitesimal generator  $P$  and rate matrix

$$R = \text{diag}\{[0, \dots, 0, \lambda + r_H^{(K)} \kappa, 0, \dots, 0, \lambda + r_L^{(K)} \kappa]\}. \quad (3)$$

Although a large-scale MMPP model with  $2(K+1)$  states is now at hand, we propose to further use the model reduction algorithm of [26] to reduce this MMPP characterized with the pair  $(P, R)$  to a two-state MMPP denoted by  $Y$ . The algorithm proposed in [26] for model reduction matches the first three marginal moments and the area under the auto-covariance function of the original process to those of the two-state approximating MMPP  $Y$ . This particular algorithm is given in Appendix A for the sake of completeness. The MMPP  $Y$  is assumed to be characterized with the quadruple  $(r_1, r_2, \eta_1, \gamma_1)$ .

We have now obtained a stochastic characterization of the traffic appearing at the input of the FDL buffers. However, some of the packets directed to the FDL buffer will not be admitted due to the lack of an idle FDL. In order to find a characterization for the actual traffic carried through the FDL buffer, we define another bi-variate process  $\mathcal{Y}(t) = \{(C_F(t), S_F(t)); t \geq 0\}$ , where  $0 \leq C_F(t) \leq B$  denotes the number of busy FDLs in the FDL buffer and  $S_F(t) \in \{1, 2\}$

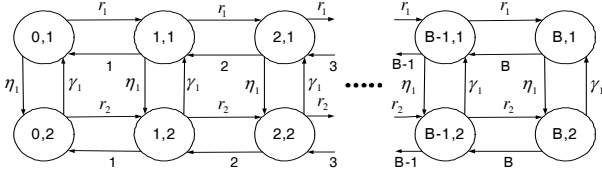


Fig. 3. The transition diagram of the Markov chain governing the behavior of the bi-variate process  $\mathcal{Y}(t)$ .

denotes the state of the two-state MMPP  $Y$  at time  $t$ . We then have a continuous-time Markov chain of size  $2(B+1)$  whose transition diagram is given in Fig. 3 which hereafter is referred to as the buffer process. Let  $Q$  denote the infinitesimal generator of this buffer process with the states enumerated as  $\{(0,0), (1,1), \dots, (B,1), (0,2), \dots, (B,2)\}$ . Packets are not admitted into the buffer only in the two states  $(B,1)$  and  $(B,2)$ . In particular, the rate of traffic using the FDL buffer is  $r_1$  at state  $(i,1)$  and  $r_2$  at state  $(i,2)$  for all  $i$  such that  $0 \leq i < B$ . Assuming an MMPP model, the traffic comprising packets using the FDL buffer can be modeled with an MMPP with  $2(B+1)$  states with infinitesimal generator  $Q$  and rate matrix

$$\Lambda = \text{diag}\{[r_1, \dots, r_1, 0, r_2, \dots, r_2, 0]\}. \quad (4)$$

We again use the procedure of [26] for reducing this MMPP to a two-state MMPP say  $Z$  which is characterized by the quadruple  $(\hat{r}_H, \hat{r}_L, \eta, \gamma)$  such that  $\hat{r}_H \geq \hat{r}_L$ . Moreover, let  $q$  denote the steady-state vector of  $Q$ :

$$q = [q_{0,1}, \dots, q_{B,1}, q_{0,2}, \dots, q_{B,2}]. \quad (5)$$

Let us now revisit the port process whose transition diagram was given in Fig. 2. Blocking arises when an arriving packet finds the system in one of the two states  $(K,H)$  and  $(K,L)$ . A blocked packet should have visited the state  $(K,H)$  with probability  $\alpha_H$  and state  $(K,L)$  with probability  $\alpha_L$ :

$$\alpha_H = \frac{(\lambda + r_H^{(K)} \kappa) p_{K,H}}{(\lambda + r_H^{(K)} \kappa) p_{K,H} + (\lambda + r_L^{(K)} \kappa) p_{K,L}}, \alpha_L = 1 - \alpha_H. \quad (6)$$

Assume that the initial distribution of the Markov chain in Fig. 2 is given by the row vector

$$\pi_0 = [0, \dots, 0, \alpha_H, 0, \dots, 0, \alpha_L]. \quad (7)$$

Then, it is clear that the vector defined by

$$\pi = \pi_0 e^{QD} \quad (8)$$

gives the conditional probabilities  $\pi_{i,H}$  and  $\pi_{i,L}$  such that

$$\pi = [\pi_{0,H}, \dots, \pi_{K,H}, \pi_{0,L}, \dots, \pi_{K,L}]. \quad (9)$$

We now have to link the two-state MMPP  $Z$  to the MMPP  $X$ . Let us consider the two extreme cases as the FDL delay  $D \rightarrow \infty$  and  $D \rightarrow 0$ . As  $D \rightarrow \infty$ , the traffic coming out of the FDL buffer should be independent of the port process. Note that  $\pi \rightarrow p$  as  $D \rightarrow \infty$  where

$$p = [p_{0,H}, \dots, p_{K,H}, p_{0,L}, \dots, p_{K,L}] \quad (10)$$

and therefore  $r_i^{(H)} \rightarrow r_H$  and  $r_i^{(L)} \rightarrow r_L$ . Consequently, as  $D \rightarrow \infty$ ,  $r_H$  and  $r_L$  should be chosen so that  $r_H = \hat{r}_H$  and

TABLE I  
ITERATIVE ALGORITHM FOR FINDING VARIOUS PERFORMANCE MEASURES OF THE FDL BUFFERING SYSTEM

Input: $K, B, D, \kappa$
1. First start with arbitrary probability vectors $p$ and $\pi$ as well as the two-state MMPP $X$ with zero rates in each state with arbitrary $\eta$ and $\gamma$ .
2. Given the vectors $p$ and $\pi$ and the parameters $r_H, r_L, \eta$ , and $\gamma$ , construct the Markov chain depicted in Fig. 2 with generator $P$ and solve for its steady-state probability vector $p$ which in turn gives the probabilities $p_{i,H}$ and $p_{i,L}$ for $0 \leq i \leq K$ .
3. Find the vector $\pi$ from (8) and define the rate matrix $R$ as in (3).
4. Use the procedure in Appendix A to obtain the two-state MMPP model $Y$ characterized by the quadruple $(r_1, r_2, \eta_1, \gamma_1)$ from the $2(K+1)$ -dimensional MMPP characterized with the pair $(P, R)$ .
5. Given the parameters $r_1, r_2, \eta_1$ , and $\gamma_1$ , construct the Markov chain depicted in Fig. 3 with generator $Q$ and solve for its steady-state probability vector $q$ .
6. First define the rate matrix $\Lambda$ as in (4) and use the procedure in Appendix A to obtain the two-state MMPP model $Z$ given from the $2(B+1)$ -dimensional MMPP characterized with the pair $(Q, \Lambda)$ .
7. Employ (11) to update $r_H$ and $r_L$ which completes the characterization of the MMPP $X$ since the parameters $\eta$ and $\gamma$ are already obtained in the previous step.
8. If the normalized difference between two successive values of the vector $p$ is less than a tolerance parameter $\epsilon$ , then exit the loop.
9. Go to step 2.
Output: loss probability $P_L$ , expected delay $E[W]$ , etc.

$r_L = \hat{r}_L$ . Consequently, the Markov chain in Fig. 2 reduces to the modeling of a  $K$ -server system fed with Poisson traffic with rate  $\lambda$  and independent MMPP  $Z$  traffic coming out of the FDL buffer. The situation is different as  $D \rightarrow 0$  since the traffic retrying to join the port is completely dependent on the port process and the burstiness of this stream should not play any additional role. Therefore, as  $D \rightarrow 0$ , we propose to replace the MMPP  $Z$  with a pure Poisson process with the same average rate, i.e.,  $r_H = r_L = \frac{\hat{r}_H \gamma + \hat{r}_L \eta}{\eta + \gamma}$ . We now have a way of characterizing the traffic coming out of the FDL buffer in the two extreme scenarios. For arbitrary values of  $D$ , we propose the following adjustment:

$$\begin{aligned} r_H &= \hat{r}_L + (\hat{r}_H - \hat{r}_L)(1 - \beta) + (\hat{r}_H - \hat{r}_L)\beta \frac{\gamma}{\eta + \gamma}, \\ r_L &= \hat{r}_L + (\hat{r}_H - \hat{r}_L)\beta \frac{\gamma}{\eta + \gamma}, \end{aligned} \quad (11)$$

where  $\beta = \|\pi - p\|_2 / \|\pi_0 - p\|_2$ , where  $\|x\|_2$  is the 2-norm of the vector  $x$ , i.e., square root of the sum of the squares of the entries of the vector  $x$ . It is also well-known that  $\beta$  approaches zero exponentially fast as  $D \rightarrow \infty$ . This method ensures that the average rate of the MMPP  $X$  matches the average rate of MMPP  $Z$  for all choices of  $\beta$  whereas the two MMPPs are exactly the same as  $D \rightarrow \infty$  as desired but the approximate MMPP is replaced with a Poisson process with the same average rate as  $D \rightarrow 0$ . The method described above leads us to a fixed-point iterative procedure which is provided in Table I.

The most important performance measure of interest is the loss probability  $P_L$  which is the probability that a packet is dropped by the system. Upon convergence of the algorithm given in Table I, we write  $P_L$  as follows:

$$P_L = \frac{(r_H^{(K)} p_{K,H} + r_L^{(K)} p_{K,L})(1 - \kappa) + r_1 q_{B,1} + r_2 q_{B,2}}{\lambda}. \quad (12)$$

On the other hand, the mean waiting time in the system is given by  $E[W]$  where  $W$  denotes the waiting time of a successfully transmitted packet. It is not difficult to show that  $E[W] = \frac{Dr_f s_r}{(1-P_L)(1-r_r)}$ , where  $r_f$  is the probability that a fresh arrival uses the FDL buffer, i.e., packet retries at least once,  $s_r$  is the probability that a retrial attempt is successful, and  $r_r$  is the probability that a retrial attempt leads to another retrial, i.e., retrying packet finds all channels occupied, therefore gets directed to the FDL buffer, and finds one idle FDL. We write these probabilities below:

$$r_f = \frac{(p_{K,H} + p_{K,L}) \sum_{i=0}^{B-1} (r_1 q_{i,1} + r_2 q_{i,2})}{\frac{r_1 \gamma_1 + r_2 \eta_1}{\eta_1 + \gamma_1}},$$

$$s_r = \frac{\sum_{i=0}^{K-1} (r_H^{(i)} p_{i,H} + r_L^{(i)} p_{i,L})}{\frac{\hat{r}_H \gamma + \hat{r}_L \eta}{\eta + \gamma}},$$

$$r_r = \kappa \frac{(r_H^{(K)} p_{K,H} + r_L^{(K)} p_{L,H}) \sum_{i=0}^{B-1} (r_1 q_{i,1} + r_2 q_{i,2})}{\frac{\hat{r}_H \gamma + \hat{r}_L \eta}{\eta + \gamma} \frac{r_1 \gamma_1 + r_2 \eta_1}{\eta_1 + \gamma_1}}.$$

Another performance measure we are interested in is the probability of strictly more than  $N$  circulations for a successful packet which is denoted by  $P_R(N) = P[W > ND]$ . We can write this probability as  $P_R(N) = \frac{r_f s_r r_r^N}{(1-P_L)(1-r_r)}$ . Finally, the computational complexity of the proposed method per iteration is  $O(M^3)$  where  $M = \max(K, B)$  which is quite reasonable when compared to traditional multi-dimensional retrial queues.

### III. NUMERICAL RESULTS

We validate the accuracy of the proposed approach by comparing them against simulations. The iteration tolerance parameter is set to  $\varepsilon = 0.000001$  throughout all the numerical examples. As a first example, we fix  $K = 16$ ,  $\rho = 0.6$  and  $\kappa = 0$  (i.e., only one FDL circulation is allowed). We then plot the loss probability  $P_L$  as a function of the FDL delay  $D$  in Fig. 4 obtained through simulations and the approach proposed in this article. For comparison purposes, we also test three other methods:

- *Method A*: This method is the same as the proposed algorithm except that the adjustment proposed in (11) is not made and we therefore set  $r_H = \hat{r}_H$  and  $r_L = \hat{r}_L$  irrespective of the FDL delay  $D$ .
- *Method B*: Instead of modeling the traffic directed to the FDL buffer as a two-state MMPP, we use a Poisson process with the same average rate and therefore ignore the burstiness of this traffic stream.
- *Method C*: This method is the M/M/K/K+B traditional queuing model also described in [14] in the context of FDL buffers.

The results are depicted in Fig. 4. We have several observations:

- For this example, the loss probability  $P_L$  appears to saturate beyond  $D = 2$  and there appears to be minimal need to further increase FDL lengths since the loss probability would not reduce much. This observation is consistent with the observation of [12] on required FDL delays based on simulation only.

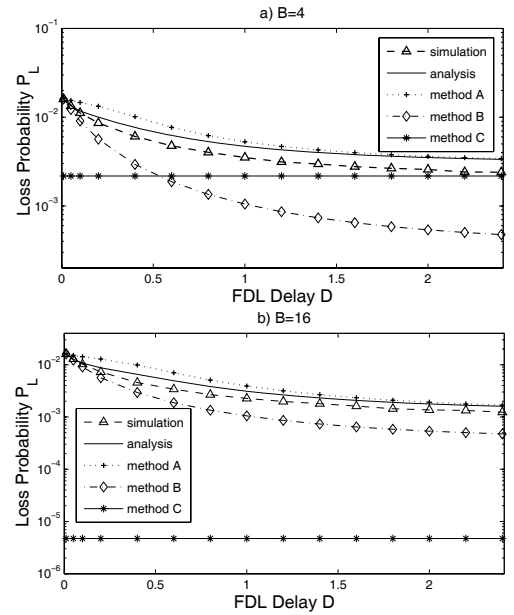


Fig. 4. Loss probability  $P_L$  as a function of the FDL delay  $D$  for the particular case  $K = 16$ ,  $\rho = 0.6$ ,  $\kappa = 0$ .

- MMPP-based models accurately estimate the loss probability in the asymptotic regime of  $D \rightarrow \infty$  as opposed to Poisson process-based models, i.e., Method B, that come short of capturing the burstiness of the traffic stream associated with retrials.
- Method C provides very loose lower bounds on the loss probability especially for larger  $B$  and it is evident that the traditional M/M/K/K+B based queuing models can not effectively be used for dimensioning FDL buffers.
- The adjustment of MMPP parameters proposed in (11) is especially useful for relatively small FDL delays and improves accuracy with respect to Method A that does not employ this adjustment procedure.

In the next example, we fix  $K = 16$  and  $D = 2$ , and we plot the loss probability  $P_L$  and expected delay  $E[W]$  for two different values of  $B$  in Fig. 5a and Fig. 5b, respectively, for  $\rho = 0.7$  and  $\rho = 0.9$ , as a function of the re-circulation parameter  $\kappa$ . It is clear that the proposed analysis method accurately captures the behavior of the system as a function of the re-circulation parameter  $\kappa$  as well as the other parameters such as  $\rho$  and  $B$  although we note the existence of a slight discrepancy between the results of simulations and analysis. The major factors for this discrepancy are the reduced order MMPP modeling of the actual traffic directed to the FDL buffer for computational feasibility and the inter-dependence between the port and buffer processes that is ignored.

In the current example, we compare two policies of re-circulation control; Policy *Fixed* allowing a fixed number of circulations ( $N$  circulations at most) as in [15] and Policy *Prob* using probabilistic re-circulation as proposed and studied in the current article. Policy *Fixed* is commonly proposed in the literature since a limit is placed on the number of FDL circulations in order not to require signal regeneration. In order to compare against Policy *Fixed* with  $N$  allowable circulations, we choose the re-circulation parameter  $\kappa(N)$  for Policy *Prob*

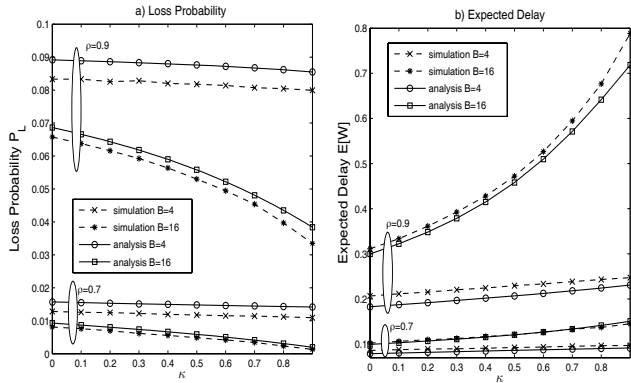


Fig. 5. Loss probability  $P_L$  and expected delay  $E[W]$  as a function of the re-circulation parameter  $\kappa$  for  $K = 16$ ,  $D = 2$ , and for two different values of  $B$  and  $\rho$ .

by analysis (using binary search) such that  $P_R(N)$  is small enough. In this example, we set  $P_R(N) = 0.001$ . In case this identity can not be satisfied, we set  $\kappa(N) = 1$ . In this way, we place a statistical limit on the maximum number of FDL circulations with Policy *Prob*. We report our findings in Table II for the case  $K = 16$ ,  $D = 2$ ,  $\rho = 0.8$ , and for various values of the buffer size  $B$ . Note that only simulation results are reported for Policy *Fixed* whereas we have both analysis and simulation results for Policy *Prob*. We have the following observations:

- Table II clearly demonstrates that the analysis method is able to capture the performance measure  $P_R(N)$ ; note the last column obtained via simulations with  $P_R(N)$  values very close to 0.001 as desired when  $\kappa < 1$ .
- For  $\kappa(N) < 0.9$ , the analysis method gives quite accurate results when compared to simulations for all three performance measures we studied. However, accuracy appears to be reduced when  $\kappa(N)$  is closer to unity.
- Although policies *Fixed* and *Prob* present comparable  $P_R(N)$ , the way we select  $\kappa(N)$  leads to slightly larger loss probability  $P_L$  but slightly lower delay  $E[W]$  for Policy *Prob* when compared to Policy *Fixed* for a given  $N$  for  $\kappa(N) < 1$ . This situation appears to be slightly reversed when  $\kappa(N) = 1$ . When  $\kappa(N) < 1$ , in order to reduce the loss probability for Policy *Prob*, one needs to further increase  $\kappa(N)$  which however also increases  $P_R(N)$ . The results show that Policy *Fixed* outperforms Policy *Prob* if a delay constraint is given in terms of  $P_R(N)$  only but there are scenarios for which Policy *Prob* may outperform Policy *Fixed* when there are further constraints on delay for example in terms of  $E[W]$ . Another advantage of probabilistic circulation is the existence of infinitely many choices for  $\kappa$  as opposed to a few values for  $N$  to choose from, to optimize system behavior in terms of  $P_L$ ,  $E[W]$ , and  $P_R(N)$ .

In the next example, for a given parameter set  $(K, D, B, \kappa)$ , we find the maximum load  $\rho_m(K, D, B, \kappa)$ , so-called throughput, on the associated system that meets the specific QoS requirements, i.e.,  $P_L \leq k_L, P_R(N) \leq k_R(N), E[W] \leq k_W$ . To obtain the quantity  $\rho_m(K, D, B, \kappa)$ , we iteratively vary the load on the system and use the proposed analysis method

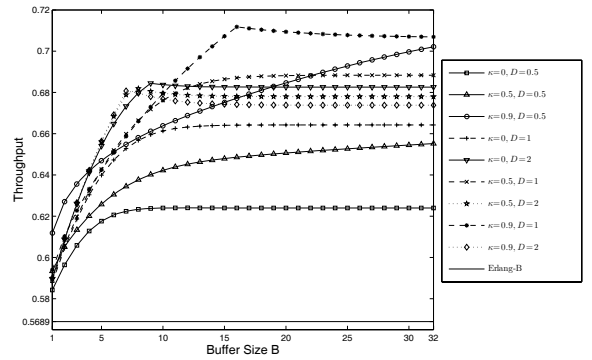


Fig. 6. The throughput of the system as a function of  $B$  for the case  $K = 32$  and for three different values of  $\kappa$  and  $D$

until all QoS requirements are just met. The low computational complexity of the analysis method allows us to obtain the quantity  $\rho_m(K, D, B, \kappa)$  for a wide range of problem parameters in acceptable time. For this example, we fix  $K = 32$ ,  $k_L = 0.001$ ,  $N = 3$ ,  $k_R(3) = 0.001$ , and  $k_W = 0.02$ . Note that the expected delay requirement is relatively strict for this example. We then plot the throughput of the system as a function of  $B$  for  $\kappa = 0, 0.5, 0.9$  and for  $D = 0.5, 1, 2$  in Fig. 6. If no buffers are used, we obtain the throughput of the system as 0.5689 using the Erlang-B formula which is also plotted in Fig. 6 as a lower bound. Our goal in this example is to quantify the gain in throughput by using FDL buffers but still meeting the QoS requirements of the system. If there are no delay or re-circulation constraints, one can increase  $B$  and  $\kappa$  together to arbitrarily increase the throughput but at the expense of very high expected delays. In this example, delay and re-circulation constraints are also taken into account. For small  $D$ , the throughput increases monotonically with respect to  $B$  and  $\kappa$  since the delay and re-circulation constraints are not violated for all parameters we tried. However, for larger values of  $D$ , the throughput increases until a certain value of  $B$  and then starts to decrease for fixed  $\kappa$ . A similar argument can be made for larger values of  $\kappa$  for which throughput is reduced beyond a certain value of  $\kappa$  for large  $D$ . This observation leads us to the conclusion that provisioning optical buffers requires analysis and optimization on provisioning parameters such as  $B, \kappa$ , and  $D$ .

In our final example, we fix  $K = 32$  and  $\kappa = 0.9$  and searched for the optimum value of  $D$ , say  $D_{opt}$ , in the interval  $0.2 \leq D \leq 4$  that yields the maximum throughput denoted by  $\rho_{opt}$  as a function of the delay requirement  $k_W$ . For this example, we set  $k_L = 0.001$ ,  $N = 3$ , and  $k_R(3) = 0.001$ . The results are given for three different choices of  $B$  in Table III. The results show that increasing the buffer size  $B$  beyond  $K$  is not very helpful. The maximum throughput strictly depends on the delay requirement whereas the optimum delay  $D_{opt}$  increases with increasing delay requirement  $k_W$  as expected but decreases with buffer size  $B$ .

#### IV. CONCLUSION

We propose a retrieval-queuing model along with fixed-point iterations to evaluate the performance of multi-wavelength feedback FDL optical buffers. Simulation results show that

TABLE II  
COMPARISON OF POLICIES *Fixed* and *Prob* FOR THE CASE  $K = 16$ ,  $D = 2$ ,  $\rho = 0.8$ , AND FOR VARIOUS VALUES OF THE BUFFER SIZE  $B = 8, 16, 64$ . BOTH ANALYSIS AND SIMULATION RESULTS ARE REPORTED FOR POLICY *Prob*.

$B$	Policy <i>Fixed</i>			Policy <i>Prob</i>					
	$N$	$P_L$	$E[W]$	$\kappa(N)$	$P_L$ (analy.)	$E[W]$ (analy.)	$P_L$ (sim.)	$E[W]$ (sim.)	$P_R(N)$ (sim.)
8	2	0.0189	0.2768	0.438	0.0278	0.2238	0.0236	0.2343	0.0008
	3	0.0168	0.3019	0.938	0.0226	0.2751	0.0175	0.3037	0.0010
	4	0.0159	0.3070	1	0.0220	0.2806	0.0159	0.3103	0.0003
	5	0.0159	0.3100	1	0.0220	0.2806	0.0159	0.3103	0.0001
16	2	0.0128	0.3221	0.361	0.0247	0.2401	0.0223	0.2451	0.0009
	3	0.0068	0.3862	0.675	0.0177	0.3071	0.0155	0.3114	0.0009
	4	0.0044	0.4194	0.909	0.0109	0.3805	0.0073	0.3957	0.0012
	5	0.0037	0.4387	1	0.0085	0.4027	0.0029	0.4470	0.0009
64	2	0.0126	0.3237	0.360	0.0246	0.2403	0.0221	0.2440	0.0009
	3	0.0061	0.3942	0.664	0.0177	0.3082	0.0154	0.3092	0.0009
	4	0.0031	0.4369	0.844	0.0112	0.3848	0.0093	0.3771	0.0010
	5	0.0017	0.4620	0.938	0.0058	0.4620	0.0045	0.4396	0.0010

TABLE III  
THE VALUES  $D_{opt}$  AND  $\rho_{opt}$  AS A FUNCTION OF THE DELAY REQUIREMENT  $k_W$  FOR  $K = 32$  AND  $\kappa = 0.9$  AND FOR THREE VALUES OF  $B$ .

$B$	Delay Requirement $k_W$	$D_{opt}$	$\rho_{opt}$
16	0.005	0.28	0.6774
	0.01	0.73	0.6881
	0.02	1.00	0.7118
	0.05	1.68	0.7397
	0.1	2.93	0.7509
	0.2	4.00	0.7527
32	0.005	0.30	0.6933
	0.01	0.53	0.7039
	0.02	0.72	0.7251
	0.05	1.18	0.7562
	0.1	2.05	0.7702
	0.2	3.92	0.7766
64	0.005	0.20	0.7095
	0.01	0.40	0.7188
	0.02	0.62	0.7324
	0.05	1.15	0.7574
	0.1	2.03	0.7706
	0.2	3.89	0.7750

the proposed model allows one to accurately estimate the loss probabilities, expected delays, and re-circulation probabilities in this system. Provisioning procedures based on the analytical model we develop are also presented to validate the effectiveness of the proposed approach. We leave the queuing modeling of shared FDL feedback buffers and the alternative scenario of employing a fixed limit on the number of FDL circulations for future research.

#### APPENDIX A TWO-STATE MMPP APPROXIMATION

One of the crucial components of the proposed method in this article is a two-state MMPP approximation characterized with the pair  $(A, \Lambda)$  to an original MMPP characterized with the pair  $(Q, R)$  with  $n > 2$  states. Let

$$A = \begin{bmatrix} -\sigma_1 & \sigma_1 \\ \sigma_2 & -\sigma_2 \end{bmatrix}, \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}.$$

Also let  $x$  denote the steady-state vector of the modulating Markov chain such that  $xQ = 0, xe = 1$  where  $e$  denotes a column vector of ones of appropriate size. The  $r$ th central moment of the arrival rate of the original MMPP is denoted

by  $m_r$  and is given by  $m_r = xR^r e, r \geq 1$ . Let  $v$  denote the variance of the original MMPP that is given by  $v = m_2 - m_1^2$ . The time constant  $\tau_c$  of the original MMPP is expressed as  $\tau_c = \frac{1}{v} \int_0^\infty r(t) dt$ , where  $r(t)$  is the covariance function of the arrival rate. The reference [27] gives an expression for  $\tau_c$ :

$$\tau_c = \frac{1}{v} [xR(ex - Q)^{-1}Re - m_1^2].$$

The following choices for the parameters of the approximating MMPP are known to match the first three non-central moments as well as the time constant [26]:

$$\sigma_1 = \frac{1}{\tau_c(1 + \eta)}, \sigma_2 = \frac{\eta}{\tau_c(1 + \eta)},$$

$$\lambda_1 = m_1 + \sqrt{v/\eta}, \lambda_2 = m_1 - \sqrt{v\eta},$$

where

$$\eta = 1 + \frac{\delta}{2} (\delta - \sqrt{4 + \delta^2}), \delta = \frac{m_3 - 3m_1v - m_1^3}{v^{3/2}}.$$

#### ACKNOWLEDGMENT

This work is done when Nail Akar was visiting UMKC as a Fulbright research scholar. The work of Nail Akar was supported in part by the support of the BONE-project ("Building the Future Optical Network in Europe"), a Network of Excellence funded by the European Commission through the 7th ICT-Framework Programme.

#### REFERENCES

- [1] P. Gambini, M. Renaud, C. Guillemot, F. Callegati, I. Andonovic, B. Bostica, D. Chiaroni, G. Corazza, S. L. Danielsen, P. Gravey, P. B. Hansen, M. Henry, C. Janz, A. Kloch, R. Krahenbuhl, C. Raffaelli, M. Schilling, A. Talneau, and L. Zucchelli, "Transparent optical packet switching: network architecture and demonstrators in the KEOPS project," *IEEE J. Sel. Areas Commun.*, vol. 16, pp. 1245–1259, 1998.
- [2] C. Qiao and M. Yoo, "Optical burst switching (OBS)—a new paradigm for an optical Internet," *J. High Speed Netw.*, vol. 8, no. 1, pp. 69–84, 1999.
- [3] F. Callegati, W. Cerroni, G. Corazza, C. Develder, M. Pickavet, and P. Demeester, "Scheduling algorithms for a slotted packet switch with either fixed or variable length packets," *Photonic Netw. Commun.*, vol. 8, no. 2, pp. 163–176, 2004.
- [4] W. A. Vanderbauwhede and D. A. Harle, "Architecture, design, and modeling of the OPSnet asynchronous optical packet switching node," *J. Lightwave Technol.*, vol. 23, pp. 2215–2228, 2005.
- [5] O. Ozturk, E. Karasan, and N. Akar, "Performance evaluation of slotted optical burst switching systems with quality of service differentiation," *J. Lightwave Technol.*, vol. 27, no. 14, pp. 2621–2633, 2009.



- [6] R. Barry and P. Humblet, "Models of blocking probability in all-optical networks with and without wavelength changers," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 5, pp. 858–867, June 1996.
- [7] I. Chlamtac, A. Fumagalli, L. Kazovsky, P. Melman, W. Nelson, P. Poggiolini, M. Cerisola, A. Choudhury, T. Fong, R. Hofmeister, C.-L. Lu, A. Mekikittikul, I. Sabido, D. J. M., C.-J. Suh, and E. Wong, "CORD: contention resolution by delay lines," *IEEE J. Sel. Areas Commun.*, vol. 14, no. 5, pp. 1014–1029, June 1996.
- [8] R. S. Tucker, P.-C. Ku, and C. J. Chang-Hasnain, "Slow-light optical buffers: capabilities and fundamental limitations," *Lightwave Technol., J.*, vol. 23, no. 12, pp. 4046–4066, Dec. 2005.
- [9] D. Hunter, M. Chia, and I. Andonovic, "Buffering in optical packet switches," *J. Lightwave Technol.*, vol. 16, no. 12, pp. 2081–2094, Dec 1998.
- [10] W. D. Zhong and R. S. Tucker, "Wavelength routing-based photonic packet buffers and their applications in photonic packet switching systems," *J. Lightwave Technol.*, vol. 16, no. 10, pp. 1737–1745, Oct. 1998.
- [11] L. Xu, H. Perros, and G. Rouskas, "Techniques for optical packet switching and optical burst switching," *IEEE Commun. Mag.*, vol. 39, no. 1, pp. 136–142, Jan. 2001.
- [12] C. M. Gauger, "Dimensioning of FDL buffers for optical burst switching nodes," in *Proc. 6th IFIP Working Conference on Optical Network Design and Modelling*, Feb. 2002.
- [13] F. Callegati, "Optical buffers for variable length packets," *IEEE Commun. Lett.*, vol. 4, no. 9, pp. 292–294, Sep. 2000.
- [14] J. S. Turner, "Terabit burst switching," *J. High Speed Netw. (JHSN)*, vol. 8, no. 1, pp. 3–16, 1999.
- [15] A. Rostami and S. Chakraborty, "On performance of optical buffers with specific number of circulations," *IEEE Photonics Technol. Lett.*, vol. 17, no. 7, pp. 1570–1572, 2005.
- [16] W. Rogiest, J. Lambert, D. Fiems, B. van Houdt, H. Bruneel, and C. Blondia, "A unified model for synchronous and asynchronous FDL buffers allowing closed-form solution," *Performance Evaluation*, vol. 66, no. 7, pp. 343–355, 2009.
- [17] W. Rogiest and H. Bruneel, "Exact optimization method for an FDL buffer with variable packet length," *IEEE Photon. Technol. Lett.*, vol. 22, no. 4, pp. 242–244, 2011.
- [18] J. Almeida, J. Pelegrini, and H. Waldman, "A generic-traffic optical buffer modeling for asynchronous optical switching networks," *IEEE Commun. Lett.*, vol. 9, no. 2, pp. 175–177, Feb. 2005.
- [19] H. Kankaya and N. Akar, "Exact analysis of single-wavelength optical buffers with feedback Markov fluid queues," *IEEE/OSA J. Optical Commun. Netw.*, vol. 1, no. 6, pp. 530–542, Nov. 2009.
- [20] M. Yoo, C. Qiao, and S. Dixit, "QoS performance of optical burst switching in IP-over-WDM networks," *IEEE J. Sel. Areas Commun.*, vol. 18, pp. 2062–2071, Oct. 2000.
- [21] X. Lu and B. Mark, "Performance modeling of optical-burst switching with fiber delay lines," *IEEE Trans. Commun.*, vol. 52, no. 12, pp. 2175–2183, Dec. 2004.
- [22] J. R. Artalejo and A. Gomez-Corral, *Retrial Queueing Systems*. Springer, 2008.
- [23] T. Yang and J. G. C. Templeton, "A survey on retrial queues," *Queueing Systems*, vol. 2, no. 3, pp. 201–233, 1987.
- [24] R. W. Wolff, *Stochastic Modeling and the Theory of Queues*. Prentice-Hall, 1988.
- [25] A. A. Fredericks and G. A. Reisner, "Approximations to stochastic service systems with application to a retrial model," *Bell Syst. Tech. J.*, vol. 58, pp. 557–576, 1979.
- [26] H. Heffes, "A class of data traffic processes-covariance function characterization and related queuing results," *Bell Syst. Tech. J.*, vol. 59, no. 6, pp. 897–929, 1980.
- [27] K. S. Meier-Hellstern, "The analysis of a queue arising in overflow models," *IEEE Trans. Commun.*, vol. 37, no. 4, pp. 367–372, 1989.
- [28] M. F. Neuts, *Structured Stochastic Matrices of M/G/1 Type and Their Applications*. Marcel Dekker, 1989.
- [29] W. Fischer and K. Meier-Hellstern, "The Markov-modulated Poisson process (MMPP) cookbook," *Performance Evaluation*, vol. 18, no. 2, pp. 149–171, 1993.



**Nail Akar** received his B.S. degree from Middle East Technical University, Turkey, in 1987 and M.S. and Ph.D. degrees from Bilkent University, Turkey, in 1989 and 1994, respectively, all in electrical and electronics engineering. From 1994 to 1996, he was a visiting scholar and a visiting assistant professor in the Computer Science Telecommunications program at the University of Missouri-Kansas City (UMKC). In 1996, he joined the Technology Planning and Integration group at the Long Distance Division, Sprint, where he held a senior member of technical

staff position from 1999 to 2000. Since 2000, he has been a faculty member at Bilkent University, currently as an associate professor. He spent six months in 2010 at the School of Computing and Engineering, UMKC, as a Fulbright scholar. His current research interests include performance evaluation of computer and communication networks, optical packet/burst switching, and network engineering. Dr. Akar is an active participant of the European Commission FP7 Network of Excellence project BONE.



**Khosrow Sohraby** is currently the Curators' Professor of Computer Science and Electrical Engineering at the University of Missouri-Kansas City. He received B.Eng and M.Eng degrees from McGill University, Montreal, Canada in 1979 and 1981, respectively, and Ph.D. degree in 1985 from the University of Toronto, Canada, all in Electrical Engineering. His current research interests include design and analysis of high speed computer and communications networks, traffic management and analysis, modern queuing theory, large-scale computations in performance analysis, and networking aspects of wireless and mobile communications.