Analytic results concerning the loss and queueing behaviour of optical buffer systems are given in, amongst others, [8,9,10]. In [8], an asynchronous FDL buffer with Poisson arrivals of bursts with exponentially distributed lengths is investigated. Synchronous FDL buffers are studied in [9] and [10] by means of a probability generating functions (pgf's) approach. In particular, the latter paper concerns the performance analysis of a synchronous optical buffer. The results obtained therein will constitute a starting point for our current contribution, dealing with the asynchronous case. Compared to a slotted (i.e., synchronous) network, an unslotted network is expected to be technologically more complex due to control issues. However, an unslotted network could turn out more opportune, for reasons of robustness and flexibility. We refer to e.g. [11] for details.

2 Model

We study a single outgoing channel, where contention is resolved by means of an FDL buffer. The FDL buffer cannot delay bursts for an arbitrary period of time, but only for multiples of a basic unit D, called the granularity. Buffers of this type are said to be degenerate, and contain $(N+1)$ FDLs with lengths kD for $k = 0, \ldots, N$. Each incoming burst is routed to the shortest of these FDLs such that the burst will not overlap on departure with bursts from the other FDLs. If such an FDL cannot be found, the burst is dropped. Typically, bursts are delayed for more time than strictly needed. This extra delay results in so-called voids, i.e., periods during which the output channel remains unused, despite of the fact that the system is not empty. Void-filling policies could be used to minimize this loss in throughput, but we do not consider these here.

For ease of analysis, we will assume in Section 3 that the buffer has an infinite amount of FDLs at its disposal, i.e., $N = \infty$. Heuristics are then obtained in Section 4 for finite N, based on the results for $N = \infty$.

Considering the evolution of the buffer contents over time, one can distinguish three important variables. Numbering bursts in the order of their arrival, the first variable, the burst inter-arrival time τ_k, captures the time between the k^{th} arrival instant and the next. The second variable is the burst size B_k, measuring the time needed for transmission of the k^{th} burst. The third important variable is the scheduling horizon H_k as observed by the k^{th} burst upon arrival. This quantity represents the time between the instant of arrival, and the earliest instant by which the previous burst (and all its predecessors) will have left the system.

The relation between these variables can be described by the following equation:

$$H_{k+1} = \left[B_k + D \left\lceil \frac{H_k}{D} \right\rceil - \tau_k \right]^+ \tag{1}$$

The expression $\lceil x \rceil$ is the ceiling of x, i.e., the smallest integer greater than or equal to x. The notation $[x]^+$ is standard shorthand for $\max(x, 0)$. When the k^{th} burst sees a scheduling horizon H_k upon arrival, it will have to be delayed for at least that amount to avoid contention. Since the buffer is degenerate, this delay cannot be realized exactly (in general), the closest match being given by $D \lceil H_k/D \rceil$. Delaying and transmitting this burst pushes the scheduling horizon (just after arrival) to $B_k + D \lceil H_k/D \rceil$. Taking then into account the burst inter-arrival time τ_k, and the possibility that the system becomes empty, one easily obtains equation (1). Note that it is valid for both continuous-time (CT) and discrete-time (DT) systems.

To analyze equation (1), we need to impose certain restrictions on the distribution of τ_k and B_k. We assume the τ_k to form a sequence of iid (independent and identically distributed) random variables (rv's), having a common memoryless distribution. The burst sizes B_k also form a sequence of iid rv's, and can have a general distribution. In the below, every time both CT and DT variables occur under the same name, we denote DT variables with a prime, e.g. D', and leave CT variables unchanged, e.g. D. In DT, we will use the probability generating function (pgf) of the probability mass function (pmf), that is, for the burst sizes B_k, $B(z) = \mathrm{E}[z^{B'_k}] = \sum_{n=1}^{\infty} z^n \mathrm{Pr}[B'_k = n]$. In CT we use the Laplace-Stieltjes transform (LST) of the probability density function (pdf), where we have $B^*(s) = E[e^{-sB_k}] = \int_0^\infty e^{-sx} b(x)dx$. As for the inter-arrival times, being memoryless, we have a geometric distribution in DT, with mean $1/p$: $\mathrm{Pr}[\tau'_k = n] = p \cdot \bar{p}^{n-1}$, $n = 1, 2, ...$ (We use the standard notation $\bar{p} = 1 - p$.) In CT, we have an exponential distribution, with arrival intensity λ: $\mathrm{Pr}[\tau_k \leq x] = 1 - e^{-\lambda x}$, $x \geq 0$. The expressions for the corresponding pgf and Laplace transform are $\tau(z) = E[z^{\tau'_k}] = \frac{pz}{1-\bar{p}z}$ and $\tau^*(s) = E[e^{-s\tau_k}] = \frac{\lambda}{\lambda+s}$.

3 Analysis

In this section, we take a look at the infinite system. All derivations assume the system is stable. On this condition, the distributions of H_k converge, for $k \to \infty$, to a unique stochastic equilibrium distribution, independent of the initial system conditions. The pgf and Laplace transforms obtained are associated with this equilibrium. By H we will denote a generic rv following that distribution (and likewise for other rv's involved). Stability requires the offered load ρ to be below some maximum value ρ_{max} that is typically less than unity (unlike in conventional queues, see e.g. [12]). This is also commented upon below.

We present a limit procedure to obtain, among others, the Laplace transform of H in the continuous-time setting. This consists in taking appropriate limits for the slot size becoming infinitely small, mapping results from the discrete-time setting to the continuous-time setting. Before continuing, we note that two separate non-linear effects can be observed in (1): the operation $[x]^+$ and $\lceil x \rceil$. The former effect, related to the non-negativeness of the buffer content, one could call the queueing effect, and requires us to analyze $H = [B + F - \tau]^+$. The latter effect, related to the finite granularity of the FDLs, one could call the FDL effect, and calls for analysis of $F = D \left\lceil \frac{H}{D} \right\rceil$. Note how values of H are mapped to multiples of D. Below, both effects will first be analyzed separately. Results will then be combined to yield the overall solution.

3.1 Results for synchronous systems

In [10], both the queueing effect $[x]^+$ and the FDL effect $\lceil x \rceil$ were studied in a discrete-time setting. The queueing effect yielded the following relation between the pgfs of the variables involved:

$$H(z) = \frac{p}{z - \bar{p}} B(z)F(z) + K' \frac{z-1}{z-\bar{p}} \tag{2}$$

The FDL effect leads to following relation:

$$F(z) = \sum_k \frac{1}{D'} \frac{z^{D'} - 1}{z\epsilon_k - 1} H(z\epsilon_k) \tag{3}$$

where the symbols $\epsilon_k = e^{j\,2\pi k/D'}$ represent the D' different complex D'[th] roots of unity, and the summation index k runs over $-D'/2 < k \le D'/2$ (taking on integer values only). Note that in [10], the summation ran over $0 \le k < D'$. For our present purposes, however, using $-D'/2 < k \le D'/2$, turns out to be more convenient.

Using the property that $F(z\epsilon_k) = F(z)$, which follows directly from the fact that the random variable F is always an integer multiple of D', one can combine (2) and (3). With the identity

$$\frac{x^{D'-1}}{z^{D'} - x^{D'}} = \sum_k \frac{1}{D'} \frac{1}{(z\epsilon_k) - x} \tag{4}$$

(evaluated at $x = \bar{p}$), the expression simplifies to

$$F(z) = K' \left(\frac{\bar{p}^{D'-1}(z^{D'} - 1)}{z^{D'} - \bar{p}^{D'}} \right) \cdot \left(1 - \sum_k \frac{1}{D'} \cdot \frac{z^{D'} - 1}{z\epsilon_k - 1} \frac{pB(z\epsilon_k)}{(z\epsilon_k) - \bar{p}} \right)^{-1} \tag{5}$$

The constant K' follows from the normalization condition $F(1) = 1$, as

$$K' = \left(\frac{1}{p} - \mathrm{E}[B'] - \frac{D'-1}{2} - \sum_{k \neq 0} \frac{1}{\epsilon_k - 1} \frac{p}{\epsilon_k - \bar{p}} B(z\epsilon_k) \right) \cdot \left(\frac{D\bar{p}^{D'-1}}{1 - \bar{p}^{D'}} \right)^{-1}$$

Having determined $F(z)$, $H(z)$ then follows readily from (2).

3.2 Limit procedure

Our scope is to derive, for the asynchronous system, the LST $H^*(s)$ of the equilibrium distribution of the scheduling horizon H as seen by arrivals. In this subsection, we discuss a limit procedure, to retrieve $H^*(s)$ from $H(z)$.

To correctly convert results from discrete time to continuous time, one should first observe quantities in the discrete domain. A distinction can be made between time-related quantities and counting-related quantities. The former scale with the slot size Δ, the latter do not. For example, the scheduling horizon H' (in slots) actually represents $H'\Delta$ in absolute time. We can also find a simple relation between the pgf and LST, as

$$H^*(s) = \mathrm{E}[e^{-sH}] = \lim_{\Delta \to 0} \mathrm{E}[(e^{-s\Delta})^{H'}] = \lim_{\Delta \to 0} H(e^{-s\Delta})$$

i.e., we need to substitute z by $e^{-s\Delta}$ in the pgf $H(z)$ before taking the limit $\Delta \to 0$. The average inter-arrival time $1/p$ scales as $1/(\lambda\Delta)$, the granularity size as $D'\Delta = D$.

Applying this limit procedure on the DT solution (2) for the queueing effect yields

$$H^*(s) = \frac{\lambda}{\lambda - s} F^*(s) B^*(s) - K \frac{s}{\lambda - s} \tag{6}$$

(In taking the limit, here and in the following, the rules of de l'Hôpital need to be applied frequently, to deal with e.g. indeterminate forms of type $0/0$.)

Concerning the FDL effect, equation (3) results in

$$F^*(s) = \sum_k \frac{1}{D} \frac{1 - e^{-sD}}{s + j\,2\pi k/D} H^*(s + j\,2\pi k/D) \tag{7}$$

where k now runs from $-\infty$ to $+\infty$ (taking on integer values only).

Note that $F^*(s)$ is periodical too, in the sense that

$$F^*(s) = F^*(s + j\,2\pi n/D)$$

for any $n \in \mathbb{Z}$. This property now allows combining (6) and (7) to yield

$$F^*(s) = \left(-K\sum_k \frac{1}{D}\frac{1-e^{-sD}}{\lambda-(s+j\,2\pi k/D)}\right)\cdot\left(1-\sum_k \frac{1}{D}\frac{1-e^{-sD}}{t}\frac{\lambda B^*(t)}{\lambda-t}\bigg|_{t=s+j\,2\pi k/D}\right)^{-1}$$

A further simplification can be made by using

$$\sum_k \frac{1}{D}\frac{1}{\lambda-(s+j\,2\pi k/D)} = -\frac{1}{1-e^{(\lambda-s)d}}$$

which follows from applying the limit procedure on (4). We then find

$$F^*(s) = \left(K\frac{1-e^{-sD}}{1-e^{(\lambda-s)D}}\right)\cdot\left(1-\sum_k \frac{1}{D}\frac{1-e^{-sD}}{t}\frac{\lambda B^*(t)}{\lambda-t}\bigg|_{t=s+j\,2\pi k/D}\right)^{-1} \quad (8)$$

This is exactly the expression we would have found applying the limit procedure directly to (5).

The remaining unknown constant K can be determined, either by applying the limit procedure once more, or by imposing the normalization condition $F^*(0) = 1$. The final result reads

$$K = \left(\frac{1}{\lambda} - \mathrm{E}[B] - \frac{D}{2} - \sum_{k\neq 0}\frac{\lambda}{t-\lambda}\frac{B^*(t)}{t}\bigg|_{t=s+j\,2\pi k/D}\right)\cdot\left(-\frac{D}{1-e^{-\lambda D}}\right)^{-1} \quad (9)$$

Equations (6), (8) and (9) together fully specify $H^*(s)$. This result was derived through a limit procedure. Note, however, that a direct approach yields identical results. One can do such by directly solving the queueing effect and FDL effect consecutively. Here, this method is not further discussed, and we refer to [13] for this approach.

4 Heuristics for the Burst Loss Probability

Results up to now related to an optical buffer of infinite size. In order to obtain the burst loss probability (BLP) in a finite system, i.e., a system with only $(N+1)$ fiber delay lines (realizing delays in the set $\{0, D, \ldots, ND\}$) one can rely on heuristics, as explained next.

For conventional queues, fed by a Poisson process of bursts of iid size, a relation exists between (the distributions of) the unfinished work in an infinite system and that in a finite system of, say, capacity M, see e.g. [14]. This relation leads to an expression for the loss ratio (LR) in the finite system of the form

$$LR = \frac{(1-\rho)}{\rho}\frac{\Pr[W_\infty > M]}{1-\Pr[W_\infty > M]}$$

where W_∞ denotes the unfinished work in the infinite system (as seen by arrivals). When dealing with degenerate buffers, one can translate this into a heuristic for the BLP

$$BLP \approx \frac{(1-\rho_{eq})}{\rho_{eq}} \frac{\text{Pr}[H_\infty > ND]}{1 - \text{Pr}[H_\infty > ND]} \tag{10}$$

Here, H_∞, the scheduling horizon in an infinite optical buffer, fulfills the role of W_∞, ND is the capacity of the system and ρ_{eq} is the so-called equivalent load, i.e., the load on the system taking into account the overhead created by the voids.

We can again combine results of the synchronous FDL buffer [10] with the limit procedure to find expressions for the unknown quantities ρ_{eq} and Pr$[H_\infty > ND]$ in heuristic (10).

The equivalent load in the synchronous setting is given by

$$\rho'_{eq} = \lambda \text{E}[B'_{eq}] = p\left(\text{E}[B'] + \frac{D'-1}{2} + \sum_{k\neq 0} \frac{1}{\epsilon_k - 1} \frac{pB(\epsilon_k)}{\epsilon_k - \bar{p}}\right)$$

The limit procedure then easily leads to the equivalent load in the asynchronous setting. One finds

$$\rho_{eq} = \lambda \text{E}[B_{eq}] = \lambda\left(E[B] + \frac{D}{2} + \sum_{k\neq 0} \frac{\lambda}{t-\lambda} \left.\frac{B^*(t)}{t}\right|_{t=s+j2\pi k/D}\right)$$

The mean equivalent burst size (E$[B'_{eq}]$ or E$[B_{eq}]$) consists of the mean burst size, about half the delay line granularity, and a term taking into account the finer details of the burst size distribution (through its pgf $B(z)$ or LST $B^*(t)$). One can show that the asynchronous system becomes unstable when λ is such that $\rho_{eq} = 100\%$.

The tail probabilities Pr$[H_\infty > ND]$ that appear in heuristic (10) can be computed by an (approximate) inversion of the LST $H^*(s)$, using a dominant-pole approximation. It was shown in [10] that for synchronous buffers one has

$$\text{Pr}[H_\infty > ND'] \approx \frac{cst'}{z_0^{ND'}}$$

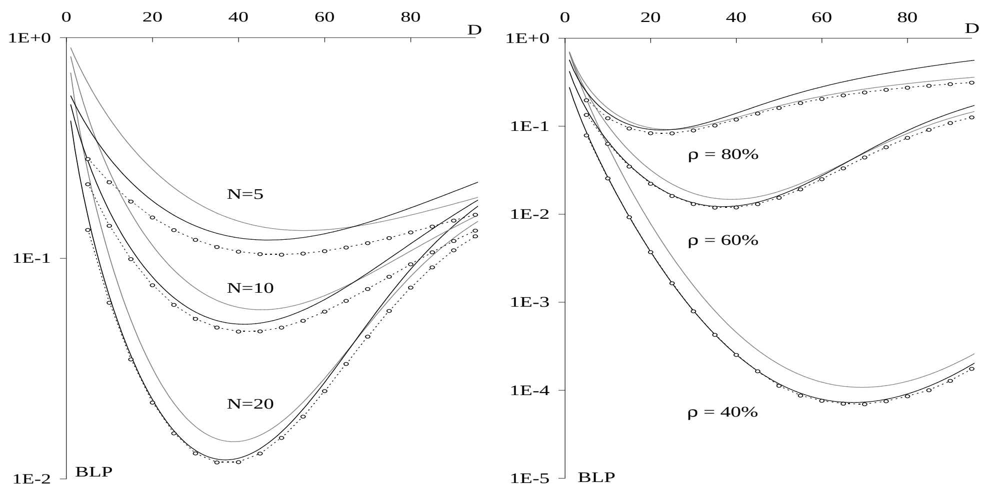
Here, z_0 is the dominant pole of $H(z)$ (and of $F(z)$). It is real, positive and larger than 1. This approximately geometric behaviour occurs under rather mild conditions on the burst size distribution, a sufficient (yet not necessary) condition for the burst size distribution is to have a rational pgf. The constant cst' follows from residue theory and is given by

$$cst' = -\frac{1}{z_0} \frac{D'}{z_0^{D'} - 1} \left(\lim_{z\to z_0} F(z)(z - z_0)\right)$$

Applying the limit procedure once more, we find that for asynchronous buffers

$$\text{Pr}[H_\infty > ND] \approx \frac{cst}{\gamma^N}$$

where we introduced $\gamma = e^{-s_0 D} = \lim_{\Delta\to 0} z_0^{D'}$ for convenience. Here, s_0 denotes the dominant pole of $H^*(s)$ and $F^*(s)$ along the negative real line. In general, a simple bisection algorithm



(a) $\rho = 60\%$ fixed (b) $N = 20$ fixed

Fig. 1. BLP vs. granularity for exponentially distributed burst sizes

(with possibly an initial search for the appropriate starting interval) suffices to determine γ numerically. In some cases, an explicit expression can also be found.

For small BLP, a modified heuristic

$$BLP \approx (1 - \rho_{eq}) \frac{\Pr[H_\infty > ND]}{1 - \Pr[H_\infty > ND]} \tag{11}$$

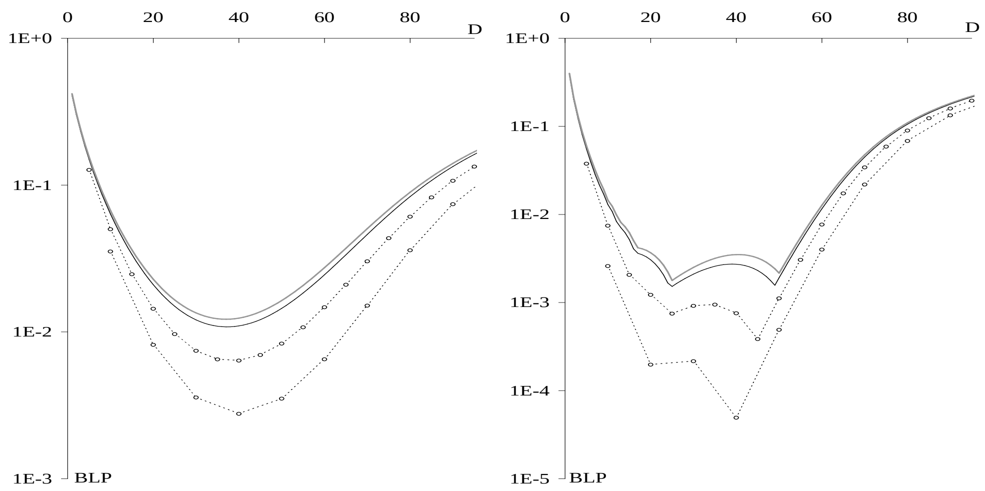
(i.e., dropping the factor ρ_{eq} in the denominator in (10)) turns out to be more accurate. In the following, we will refer to (10) as "heuristic A" and to (11) as "heuristic B" respectively.

It is worth to point out that the same heuristics can also be used to evaluate the BLP in overloaded systems, i.e., when the equivalent load exceeds 100%. Strictly speaking, no equilibrium distribution then exists for e.g. H_∞. The transform $H^*(s)$ that is used to approximate $\Pr[H_\infty > ND]$, however, remains a proper function. Formally then, one can still compute the quantities $\Pr[H_\infty > ND]$, the only caveat being that γ is then to be found in the interval $[0, 1)$, i.e. $s_0 > 0$. The expression for the constant cst remains the same. (When the equivalent load is exactly 100%, $s_0 = 0$ and $\gamma = 1$. In principle, this requires somewhat modified expressions. Here, we do not pursue this issue further.)

5 Numerical Examples

In this section, we take a look at the BLP of two special cases for the burst size distribution: the exponential and deterministic. For both of them, the infinite sum appearing in e.g. equation (8) or (9), can be removed. One obtains closed-form formulas for the LSTs and performance measures derived therefrom. As for the derivation, we confine ourself to saying that, for both special cases, an efficient method is to derive results in discrete time (see [15]), and obtain continuous-time results appling the limit procedure, as explained in Section 3.2.

In Figure 1, numerical results for the case of exponentially distributed burst sizes are shown. It compares results from simulation (points connected by dotted lines) with those obtained via



(a) memoryless burst sizes (b) determistic burst sizes

Fig. 3. BLP vs. granularity for synchronous and asynchronous systems

curves (lowest BLP curves) show results for a DT setting, with a finite slot length Δ, obtained with heuristic B. The first two sets (points connected by dotted lines) are valid for $\Delta = 10\mu s$, $\Delta = 5\mu s$ respectively. The third set has a slot length $\Delta = 1\mu s$, and is depicted with solid black curves. (The DT step, although not visible, is $\Delta = 1\mu s.$). The last set (highest BLP curves) represents the BLP for a CT setting, and is also obtained with heuristic B. In DT, the memoryless burst size distribution corresponds to the geometric distribution, in CT to the exponential distribution.

Both panes illustrate how higher values for the time slot lengths result in lower losses. For gradually reducing time slot lengths, curves evolve to the limit distribution of zero slot length, i.e. the asynchronous case. For the memoryless burst size distribution, the curves mostly preserve their smooth shape. For the deterministic case, the “local optima” keep occuring.

The above comparison suggests a better performance for synchronized systems in terms of loss probabilities. Note, however, that all curves assumed that at most one arrival can occur per slot. Results from a so-called batch arrival model, where more than one arrival per slot can occur, confirm this trend, but show that the gain in performance is smaller.

6 Conclusions

Expressions for various performance measures for asynchronous optical buffers were derived by taking the limit of corresponding results obtained elsewhere for synchronous ones. Based on (approximate inversion of) the LST of the scheduling horizon in an infinite system, heuristics were developed to determine the BLP in finite systems. Two special cases of burst size distributions were used to establish the accuracy of this approximation, by comparing heuristics to simulation results. The examples given further revealed that the BLP is rather sensitive to the choice of the granularity D, as was the case for synchronous optical buffers. The optimal value depends not only on the burst size distribution, but also on the offered load.