

# A Family of Performance Bounds for QoS Measures in Packet-Based Networks <sup>\*</sup>

József J. Bíró, Zalán Heszberger, and Mátyás Martinecz

Budapest University of Technology and Economics  
Department of Telecommunications and Media Informatics  
H-1117, Magyar tudósok körútja 2., Budapest, Hungary  
biro@tmit.bme.hu

**Abstract.** In characterizing statistical multiplexing models of network nodes important QoS measures are often to be estimated such as saturation probability and workload loss ratio. In the two main streams of multiplexer models (bufferless and buffered statistical multiplexers) the meaning of these measures and their possible estimations are different. In this paper, we set up a family of closed-form performance bounds which can be used for estimating QoS measures in both multiplexing models. The underlying approximation technique in the framework of bufferless fluid flow multiplexing model (*bffm*) is systematically described, and besides the resultant new QoS measure estimates previously known ones are also identified. The bounds in *bffm* have been analyzed and compared based on numerical investigations. The applicability of the performance bounds for buffered multiplexers is also briefly discussed.

## 1 Introduction

Setting and estimating QoS measures play central role in traffic flow control and management functions in QoS guaranteed packet-based networks. As regards the available resources (e.g. transmission link capacities, buffers, processing capacities) at the nodes in the network, an important question is how often such resources are overloaded due to high volume of packets arrived. The class of QoS measures which quantify these overloading phenomena is often referred to as resource-based QoS measures [1]. One representative measure within this family is the link saturation probability which corresponds to the fraction of time when the sum of the instantaneous (or average over a sufficiently small time interval) arrival rate of traffic flows exceeds the transmission link capacity. The use of this measure assumes the bufferless fluid flow multiplexing framework which turned out to be powerful dimensioning tool in either the case of elastic traffic in certain access network scenarios [2] or the case of stream-like traffic [3].

Another characteristic member of the family of resource-based QoS measures is the buffer saturation (or buffer overflow) probability. The analysis of this apparently requires a buffered statistical multiplexing model. Several asymptotic

---

<sup>\*</sup> This work was supported by the Inter University Centre for Telecommunications and Informatics, Hungary.

and approximate results have been formulated for buffer overflow probability under different assumptions. Recent result in [4] incorporates bounds for buffer overflow probability provided the arrival traffic flows are regulated (the regulation is characterized by arrival curves) and the service offered to the traffic by the nodes is described by a so-called service curves. These bounds are based on Hoeffding's result on the tail probability estimation of sum of partial backlogs as bounded random variables [5].

Besides the fraction of resource overload periods, it is also important to identify quantities (measures) which are based on the amount of traffic becoming unconformant due to the resource overload. The class of these measures is referred to as stream-based<sup>1</sup> measures. Although the unconformant packets can be either downgraded into lower level QoS class (e.g. best effort) or simply discarded, for simplicity, the ratio between the unconformant traffic and the whole offered traffic is called as workload loss ratio (*WLR*). In the *bffm* modeling framework *WLR* corresponds to the fraction of traffic which can not be transmitted due to the link saturation. In buffered statistical multiplexers *WLR* means the fraction of packets which can not be placed into the buffer due to buffer overflow.<sup>2</sup>

In this paper, we set up a family of conservative upper bounds for saturation probability and workload loss ratio, which can directly be applied in both (bufferless and buffered) multiplexing framework. For this purpose, the so-called Chernoff-Hoeffding bounding method as an approximation technique has been used which enables to treat the estimates for saturation probability and workload loss ratio in a common way. This underlying approximation method in the framework of bufferless fluid flow multiplexing model (*bffm*) is systematically described, and besides the resultant new QoS measure estimates previously known ones are also identified. The very attractive properties of the bounds that they are expressed in closed-form formulae and use few characteristic information (known a priori and/or measurable) on the traffic flows.

In Section 2 we introduce the general concept of Chernoff-Hoeffding bounding technique in the context of bufferless fluid flow multiplexing model, which originally motivates the design of the bounds on the related QoS measures. In Section 3 the underlying probability generating function approximations have been performed. Previously known and newly developed bounds has been presented under a common framework in Section 4. After that analysis and comparisons based on extensive numerical investigations have been highlighted. The applicability of the performance bounds for buffered multiplexers is also briefly discussed.

---

<sup>1</sup> It does not necessarily mean that these measures can be applied only for streaming traffic.

<sup>2</sup> In buffered multiplexing models the buffer size is usually assumed to be infinite, hence, *WLR* is often identified as the fraction of workload in the buffer being above a certain threshold.

## 2 The Chernoff-Hoeffding Bounding Method

Bufferless fluid flow multiplexing is often used in the literature to analyze QoS measures, e.g., packet loss probability in a multiplexer [4, 1, 3]. Because this approach assumes no buffer at burst time scales, it is able to provide conservative estimates for the QoS measures under question. For modeling purposes under *bffm*, let us assume that we have  $n$  fluid flows to be multiplexed on a communication link with transmission capacity  $C$ . Let the instantaneous stationary (that is time dependence can be eliminated) arrival rate of flow  $i$  be noted by  $X_i$ , as a random variable. Because every flow has a peak rate  $p_i$  we also have  $0 \leq X_i \leq p_i$ . Further, let the aggregate flow arrival rate be  $X = \sum_{i=1}^n X_i$ .

The link saturation probability can now be defined as

$$P_{\text{sat}} \stackrel{\text{def}}{=} \mathbb{P}(X > C).$$

This probability reflects the fraction of time when the link is overloaded (provided the system is ergodic), i.e. the frequency that the combined arrival rate exceeds the link capacity. This resource-based congestion measure could be important from network operation point of view. The workload loss ratio can be identified as

$$WLR \stackrel{\text{def}}{=} \frac{\mathbb{E}[(X - C)^+]}{\mathbb{E}[X]}, \quad (1)$$

where  $\mathbb{E}[\cdot]$  stands for the expectation value operator and  $(X - C)^+ = \max(X - C, 0)$ . The estimation of this quantity can provide more accurate loss performance analysis. This measure better characterizes the expected loss rate and could also contribute to determining the users' satisfaction. From traffic management (e.g. connection admission control) point of view an important question can arise: Whether the ongoing session (possibly together with a newcomer) satisfies a predefined QoS constraint related to some quality of service measure. In a more formal way, the inequalities

$$\mathbb{P}(X > C) \leq e^{-\gamma}, \quad \frac{\mathbb{E}[(X - C)^+]}{\mathbb{E}[X]} \leq e^{-\gamma} \quad (2)$$

represent the fulfillment of the constraint on saturation probability and workload loss ratio, respectively.

The Chernoff bound of  $P_{\text{sat}}$  and  $WLR$  are as follows [6]:

$$\mathbb{P}(X > C) \leq \inf_{s>0} \frac{G_X(s)}{e^{sC}} = \inf_{s>0} \exp(\Lambda_X(s) - sC), \quad (3)$$

$$WLR \leq \exp(\Lambda_X(s^*) - s^*C - \log(s^*M)), \quad (4)$$

where

$$s^* = \operatorname{arginf}_s (\Lambda_X(s) - sC), \quad (5)$$

and  $G_X(s) \stackrel{\text{def}}{=} \mathbb{E}[\exp(sX)]$  and  $\Lambda_X(s) \stackrel{\text{def}}{=} \log G_X(s)$  are the probability generating function (PGF) and the cumulant generating function (CGF) of  $X$ , respectively.

The computation of these bounds is usually not possible, because the underlying generating functions would require all the moments of  $X$  to be known. Instead, the CGF's are to be further bounded based on the available information (moments) on  $X$  and embedded into the Chernoff bound. This is called the Chernoff-Hoeffding bounding method.

### 3 Approximations of Probability Generating Functions (PGF)

In the previous section we have seen that the cumulant generating function plays important role in QoS measure approximation through the Chernoff bounding method. In this section, we provide three conservative bounds of the PGF of aggregate traffic rate distribution, provided only the following pieces of information are available on  $X$ : the number of traffic flows multiplexed ( $n$ ), the peak rates of the traffic flows ( $p_i$ ) and the aggregate mean arrival rate ( $M \stackrel{\text{def}}{=} \mathbb{E}[X]$ ).

#### 3.1 Approximations Based on Hoeffding's result

The following lemmas due to Hoeffding (1963), is on the PGF approximation of bounded random variables.

**Lemma 1 ([5]).**

Let  $X_i$ ,  $i = 1 \dots n$  be independent random variables with  $X = \sum_{i=1}^n X_i$ ,  $M = \mathbb{E}[X]$  and  $0 \leq X_i \leq p$ . Then, for  $s > 0$

$$G_X(s) \leq \left(1 - \frac{M}{np} + \frac{M}{np} \exp(sp)\right)^n \quad (6)$$

**Lemma 2 ([5]).** Let  $Y$  be a random variable with  $\mathbb{E}[Y] = 0$ ,  $a \leq Y \leq b$ . Then for  $s > 0$ ,  $\mathbb{E}[\exp(sY)] \leq \exp\left(\frac{s^2(b-a)^2}{8}\right)$

Based on this lemma it can be easily constructed an upper bound for the PGF of sums of independent and bounded random variables.

**Corollary 1.** Let  $X_i$ ,  $i = 1 \dots n$  be independent random variables with  $X = \sum_{i=1}^n X_i$ ,  $M = \mathbb{E}[X]$  and  $0 \leq X_i \leq p_i$ . Then for  $s > 0$ ,

$$\mathbb{E}[\exp(sX)] \leq \exp(sM) \exp\left(\frac{s^2 \sum_{i=1}^n p_i^2}{8}\right) \quad (7)$$

The PGF bound in (6) is applicable for random variable bounded uniformly (but with not necessarily identical distribution), while that in (7) covers a more general case with non-uniformly bounded random variables. Nevertheless, the latter one does not coincide the former one in the special case of  $p_1 = p_2 = \dots = p_n$ . This fact motivated the construction of the following upper bound.

### 3.2 An Improved Hoeffding-type Approximation

Applying the Chernoff-Hoeffding bounding method on  $X$  in a different way, we have obtained the following conservative bound for  $\mathbb{E}[\exp(sX)]$ .

**Theorem 1** ([1]). *Let  $X_i$  be independent bounded random variables with  $0 \leq X_i \leq p_i$ ,  $X = \sum_{i=1}^n X_i$  and  $M = \mathbb{E}[X]$ . Then for  $s > 0$ ,*

$$\mathbb{E}[e^{sX}] \leq \prod_{i=1}^n \left( \frac{e^{sp_i} - 1}{p_i} \right) \left( \frac{M + \sum_{k=1}^n \frac{p_k}{e^{sp_k} - 1}}{n} \right)^n. \quad (8)$$

The proof of this theorem is not detailed here, it can be found in [1]. Nevertheless, an important step in the chain of bounding formulae is worth repeating here:

$$G_X(s) = \mathbb{E}[e^{sX}] \leq \prod_{i=1}^n \left( 1 + m_i \frac{e^{sp_i} - 1}{p_i} \right), \quad (9)$$

where  $m_i \stackrel{\text{def}}{=} \mathbb{E}[X_i]$ .

**Corollary 2.** *The PGF bound on the right-hand side in (9) is the exact generating function of the sum of heterogeneous on-off random variables with the distribution*

$$\mathbb{P}(X_i^{\text{onoff}} = p_i) = \frac{m_i}{p_i}, \quad \mathbb{P}(X_i^{\text{onoff}} = 0) = 1 - \frac{m_i}{p_i}.$$

This is because

$$\mathbb{E}[e^{sX_i^{\text{onoff}}}] = \left( 1 - \frac{m_i}{p_i} + \frac{m_i}{p_i} e^{sp_i} \right).$$

It can also be seen that the formula on the right hand side in (8) gives back the bound in (6) in the case of uniformly bounded random variables. In this way, this improved Hoeffding-type approximation is a consistent extension of the result presented in Lemma 1.

### 3.3 A PGF Approximation Based on Stochastic Ordering

In this subsection let us recall the essential definitions and properties of a certain type of stochastic ordering of random variables to be applied for PGF approximation.

**Definition 1** ([4]). *Given two random variables  $X$  and  $Y$  with distribution function  $F_X$  and  $F_Y$ , respectively. Then,  $X$  is said to be smaller than  $Y$  with respect to increasing convex ordering, written as*

$$X <_{icx} Y,$$

if the condition

$$\int_{-\infty}^{\infty} \phi(x) dF_X(x) \leq \int_{-\infty}^{\infty} \phi(x) dF_Y(x)$$

holds for all increasing convex function  $\phi$ , for which the integral exists.

An important consequence of this definition for probability generating functions of random variables is the following:

**Lemma 3.** *Let  $X$  and  $Y$  be two random variables with the relation  $X <_{icx} Y$ . Then for  $s > 0$ ,  $G_X(s) \leq G_Y(s)$ .*

This can be justified by the substitution  $\phi(x) = \exp(sx)$ .

The following results presented in [4] leads us to construct a new PGF bound.

**Lemma 4.** *Let the random variables  $X_1^{\text{onoff}}, \dots, X_n^{\text{onoff}}$  represent  $n$  independent heterogeneous on-off sources with peak rates  $p_1, \dots, p_n$  and mean rates  $m_1, \dots, m_n$ . Let  $Y_1^{\text{onoff}}, \dots, Y_{n_Y}^{\text{onoff}}$  be  $n_Y$  independent homogeneous on-off sources with the identical peak rate  $p = \max(p_i, i = 1, \dots, n)$ ,  $n_Y = \lceil \sum_{i=1}^n p_i/p \rceil$ , and identical mean rate  $m = \sum_{i=1}^n m_i/n_Y$ . Then  $X_{\text{onoff}} <_{icx} Y_{\text{onoff}}$ , where*

$$X_{\text{onoff}} \stackrel{\text{def}}{=} \sum_{i=1}^n X_i^{\text{onoff}} \quad \text{and} \quad Y_{\text{onoff}} \stackrel{\text{def}}{=} \sum_{i=1}^{n_Y} Y_i^{\text{onoff}}. \quad (10)$$

For a proof of Lemma 4 see [4].

Now, the PGF bound based on increasing convex stochastic ordering can be formulated in the following theorem [7]:

**Theorem 2.** *Let  $X_1, \dots, X_n$  indicate  $n$  independent random variables with  $0 \leq X_i \leq p_i$ ,  $X = \sum_{i=1}^n X_i$  and  $M = \mathbb{E}[X]$ . Then for  $s > 0$ ,*

$$G_X(s) \leq \left( 1 - \frac{M}{n_Y p} + \frac{M}{n_Y p} e^{sp} \right)^{n_Y} \quad (11)$$

*Proof of Theorem 2:* By Corollary 2 we have  $G_X(s) \leq G_{X_{\text{onoff}}}(s)$ ,  $\forall s > 0$ . Further, by combining Lemma 4 and Lemma 3 the following relation also holds:  $G_{X_{\text{onoff}}}(s) \leq G_{Y_{\text{onoff}}}(s)$ ,  $\forall s > 0$ . The two inequalities above give the statement of the theorem, because  $G_{Y_{\text{onoff}}}(s) = \left( 1 - \frac{M}{n_Y p} + \frac{M}{n_Y p} e^{sp} \right)^{n_Y}$ . Q.E.D.

Let the PGF approximations presented in (7), (8) and (11), be designated by  $\tilde{G}_{X,\text{hoe}}(s)$ ,  $\tilde{G}_{X,\text{ih}}(s)$  and  $\tilde{G}_{X,\text{so}}(s)$ , respectively. The corresponding cumulant generation functions (CGF's) are  $\tilde{\Lambda}_{X,\text{hoe}}(s)$ ,  $\tilde{\Lambda}_{X,\text{ih}}(s)$  and  $\tilde{\Lambda}_{X,\text{so}}(s)$ .

## 4 Conservative Upper Bounds Based on the PGF Approximations

### 4.1 Bounds for the saturation probability $P_{\text{sat}}$

Applying the Chernoff bound for the saturation probability of uniformly bounded random variables with the CGF bound in (6) the following bound can be obtained [5]:

**Theorem 3.** Let  $X_i$ ,  $i = 1 \dots n$  be independent random variables with  $X = \sum_{i=1}^n X_i$ ,  $M = \mathbb{E}[X]$  and  $0 \leq X_i \leq p$ . Then, for  $C > M$

$$\mathbb{P}(X > C) \leq \left(\frac{M}{C}\right)^{\frac{C}{p}} \left(\frac{np - M}{np - C}\right)^{n - \frac{C}{p}} \quad (12)$$

In this case the optimal  $s$  parameter can also be expressed as

$$s^* = \frac{1}{p} \log \frac{C}{M} \frac{np - M}{np - C}. \quad (13)$$

In the more interesting case of non-uniformly bounded random variables the following bounds can be obtained by the substitution of  $\tilde{\Lambda}_{\text{hoe}}(s)$ ,  $\tilde{\Lambda}_{\text{ih}}(s)$  and  $\tilde{\Lambda}_{\text{so}}(s)$  into the Chernoff bound of  $P_{\text{sat}}$  (3).

**Theorem 4 ([5]).** Let  $X_i$  be independent bounded random variables with  $0 \leq X_i \leq p_i$ ,  $X = \sum_{i=1}^n X_i$  and  $M = \mathbb{E}[X]$ , then

$$\mathbb{P}(X > C) \leq \exp\left(\frac{-2(C - M)^2}{\sum_{i=1}^n p_i^2}\right) \quad (14)$$

The optimizing parameter  $s$  can be formulated here as

$$s^* = \frac{4(C - M)}{\sum_{i=1}^n p_i^2}. \quad (15)$$

**Theorem 5 ([1]).** If  $X_1, X_2, \dots, X_n$  are independent (and not necessarily identically distributed) random variables, for which  $0 \leq X_i \leq p_i$  holds, then

$$\mathbb{P}(X \geq C) \leq e^{-s^*C} \left(\frac{M + \sum_{j=1}^n \frac{p_j}{e^{s^* p_j} - 1}}{n}\right)^n \prod_{k=1}^n \frac{e^{s^* p_k} - 1}{p_k}, \quad (16)$$

where  $s^*$  is the solution of the following equation.

$$\sum_{k=1}^n \frac{e^{s^* p_k} p_k}{e^{s^* p_k} - 1} - \frac{n \sum_{j=1}^n \frac{e^{s^* p_j} p_j^2}{(e^{s^* p_j} - 1)^2}}{M + \sum_{j=1}^n \frac{p_j}{e^{s^* p_j} - 1}} - C = 0. \quad (17)$$

Unfortunately, in this case neither the optimizing parameter  $s^*$  nor the bound of  $P_{\text{sat}}$  can not be expressed in closed form. In [1] closed-form solutions have been developed for  $P_{\text{sat}}$  through finding closed-form suboptimal solutions of the equation above with respect to  $s$ . One of them is repeated here:

$$\mathbb{P}(X \geq C) \leq e^{-\tilde{s}^*C} \left(\frac{M + \sum_{j=1}^n \frac{p_j}{e^{\tilde{s}^* p_j} - 1}}{n}\right)^n \prod_{k=1}^n \frac{e^{\tilde{s}^* p_k} - 1}{p_k}, \quad (18)$$

where

$$\tilde{s}^* = \frac{C - M}{\frac{1}{4} \sum_{i=1}^n p_i^2 - \frac{1}{n} \left(M - \frac{1}{2} \sum_{i=1}^n p_i\right)^2} \quad (19)$$

As a further new contribution of this paper, applying the CGF approximation  $\tilde{\Lambda}_{\text{so}}(s)$  based on stochastic ordering, inherently a closed form upper bound can be obtained for  $P_{\text{sat}}$ :

**Theorem 6.** Let  $X_i$  be independent bounded random variables with  $0 \leq X_i \leq p_i$ ,  $X = \sum_{i=1}^n X_i$  and  $M = \mathbb{E}[X]$ . Further, let  $p = \max(p_i, i = 1, \dots, n)$ ,  $n_Y = \lceil \sum_{i=1}^n p_i/p \rceil$ , and  $m = \sum_{i=1}^n m_i/n_y$ , then

$$\mathbb{P}(X > C) \leq \left(\frac{M}{C}\right)^{\frac{C}{p}} \left(\frac{n_Y p - M}{n_Y p - C}\right)^{n_Y - \frac{C}{p}} \quad (20)$$

*Proof sketch of Theorem 6:* Combining the result of Theorem 2 and Theorem 3 gives the required statement.

In this case the optimal  $s$  parameter can also be expressed as

$$s^* = \frac{1}{p} \log \frac{C}{M} \frac{n_Y p - M}{n_Y p - C}. \quad (21)$$

## 4.2 Bounds for the workload loss ratio *WLR*

Turning to the *WLR* approximation, here it is worth using formula (4)<sup>3</sup>, because in this case the optimizing parameter  $s$  and hence the resulted conservative upper bounds can be expressed directly in closed-form (like in the previously presented  $P_{\text{sat}}$  bounds) when the CGF approximations  $\tilde{\Lambda}_{X,\text{hoe}}(s)$  and  $\tilde{\Lambda}_{X,\text{so}}(s)$  are embedded in (4). When the CGF bound  $\tilde{\Lambda}_{X,\text{ih}}(s)$  is used similar sub-optimal solutions can be obtained as in the corresponding  $P_{\text{sat}}$  bound in (18).

In the following theorem we summarize these new closed form conservative bounds:

**Theorem 7.** Let  $X_i$  be independent bounded (and not necessarily identically distributed) random variables with  $0 \leq X_i \leq p_i$ ,  $X = \sum_{i=1}^n X_i$  and  $M = \mathbb{E}[X]$ . Further, let  $K = \frac{1}{4} \sum_{i=1}^n p_i^2 - \frac{1}{n} (M - \frac{1}{2} \sum_{i=1}^n p_i)^2$ ,  $p = \max(p_i, i = 1, \dots, n)$ ,  $n_Y = \lceil \sum_{i=1}^n p_i/p \rceil$ , and  $m = \sum_{i=1}^n m_i/n_y$ , then the following three inequalities hold for *WLR*:

$$WLR \leq \frac{\sum_{i=1}^n p_i^2}{4(C-M)M} \exp\left(\frac{-2(C-M)^2}{\sum_{i=1}^n p_i^2}\right), \quad (22)$$

$$WLR \leq \left(\frac{M + \sum_{j=1}^n \frac{p_j}{e^{\frac{C-M}{K} p_j} - 1}}{n}\right)^n \frac{K e^{\frac{(M-C)C}{K}}}{(C-M)M} \prod_{k=1}^n \frac{e^{\frac{C-M}{K} p_k} - 1}{p_k}, \quad (23)$$

$$WLR \leq \left(\frac{M}{C}\right)^{C/p} \left(\frac{n_Y p - M}{n_Y p - C}\right)^{n_Y - C/p} \frac{p}{M \log \frac{C}{M} \frac{n_Y p - M}{n_Y p - C}}. \quad (24)$$

<sup>3</sup> A slight improvement of formula (4) would be  $\inf_{s>0} \exp(\Lambda_X(s) - sC - \log(sM))$ , but the use of this would require further approximations to arrive closed-form expression.



*Proof sketch of Theorem 7* : Substituting the three CGF approximations  $\tilde{\Lambda}_{X,\text{hoe}}(s)$ ,  $\tilde{\Lambda}_{X,\text{ih}}(s)$  and  $\tilde{\Lambda}_{X,\text{so}}(s)$  and the corresponding optimization parameters  $s^*$  performed in equations (15), (19), (21) into the Chernoff bound of  $WLR$  (4), the three bounds above are obtained.

Finally, it is worth again emphasizing that in the set of bounds presented above the ones in (14) and (18) are already known from [5] and [1], but, to the authors best knowledge, the bounds in (20), (22), (23) and (24) are neither presented nor analyzed previously.

## 5 Performance Analysis

In this section the performance of the bounds are analyzed and illustrated through numerical examples. For this purpose a simple two-class on-off traffic mix has been defined. The number of sources within the classes are represented by  $n_1$  and  $n_2$ , respectively. The mean arrival rate and the peak rate of a source within a class are assumed to be identical and indicated by  $m_i$ ,  $p_i$ ,  $i = \{1, 2\}$ . The representative traffic scenarios considered in the paper for illustrating the numerical investigations are summarized in Table 1. The first traffic mix (Mix 1) resembles the aggregation of uncompressed voice and compressed video flows. The second (Mix 2) and third one (Mix 3) represent the multiplexing of uncompressed and compressed voice traffic with low and high peak to mean ratio, respectively.

**Table 1.** Traffic Scenarios

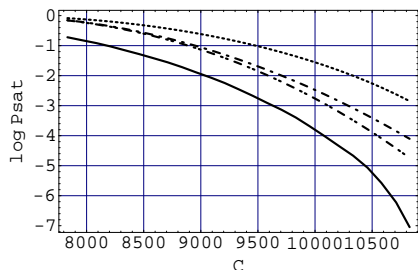
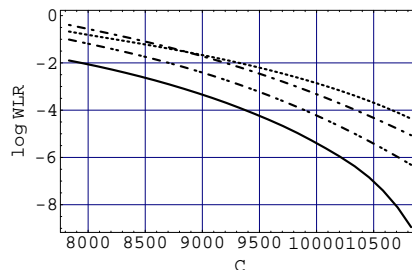
	$n_1$	$m_1$ [kbit/s]	$p_1$ [kbit/s]	$n_2$	$m_2$ [kbit/s]	$p_2$ [kbit/s]	$\frac{P}{M}$
Mix 1	100	51	64	10	$2 \times 10^3$	$5 \times 10^3$	2.24
Mix 2	100	51	64	1000	4.8	5.8	1.34
Mix 3	100	25	64	1000	2.3	5.8	2.54

In the figures the 10-based logarithm of the exact values of saturation probability and workload loss ratio and their bounds are drawn in the function of the transmission capacity  $C$ . Since the bounds presented give reasonable values when  $M < C < P$  ( $P \stackrel{\text{def}}{=} \sum_{i=1}^n p_i$ ), parts of the interval  $(M, C)$  is considered in the drawing in such a way that the exact values of  $P_{\text{sat}}$  and  $WLR$  should be no smaller than  $10^{-8}$ . The exact values are drawn with continuous lines, while the bounds based on the Hoeffding, improved Hoeffding and stochastic ordering-based CGF bounds are represented by dotted, dash-dotted and dash-dot-dotted lines, respectively.

Common observations and remarks based on our extensive numerical analysis are given, which are partly illustrated by the numerical examples.

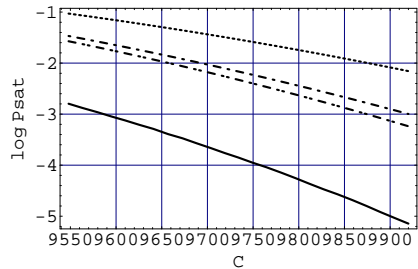
*Observations:*

- The bounds (14), (22) based on the CGF approximation (7) has usually the poorest performance, due to the underlying coarse bound on the cumulant generating function.
- The differences between the improved Hoeffding and stochastic ordering-based  $P_{\text{sat}}$  bounds are usually small, furthermore, it turned out to be negligible when the number of sources are higher than 100 in each traffic class and the peak rates of the traffic classes are in similar order of magnitude (e.g. in Mix2 and Mix 3).
- The superiority of the stochastic ordering-based  $WLR$  bound can be observed in several cases, especially when the aggregate peak to mean ratio ( $P/M$ ) is small (e.g. in Mix 2).
- In case of high peak to mean ratio and high differences between the peak rates of the traffic classes, the improved Hoeffding-based  $P_{\text{sat}}$  bound can outperform (such figures not presented) the stochastic ordering-based  $P_{\text{sat}}$  bound.
- The horizontal and vertical distances between the curves are usually increases with increasing  $\gamma$  (with tightening the QoS constraint).

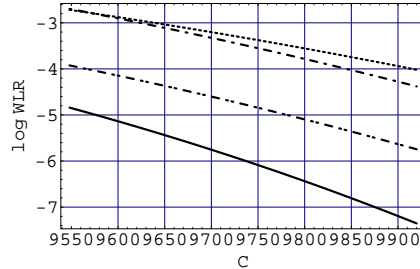
**Fig. 1.** Bounds on  $P_{\text{sat}}$ , Mix 1**Fig. 2.** Bounds on  $WLR$ , Mix 1

*Remarks:* Although all the bounds presented require the same amount of information on the the traffic flows, the complexity of their closed-form formulae are different. The bounds based on the Hoeffding-based CGF approximation appear in the simplest way, however, these have the poorest accuracy. Nevertheless, these bounds can result in saving considerable amount of capacity compared to the plain peak rate reservation scheme (see the horizontal distances between the corresponding curves), and the application of them could be recommended when the simplicity is an exclusive criterion.

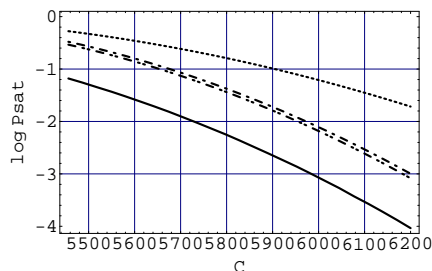
The bounds (18), (23) based on the CGF approximation (8) have the most complicated appearance in the formulae, the implementation of their computation might encounter serious problems due to the presence of the several exponential-like terms. The improved Hoeffding-based saturation probability bound can have



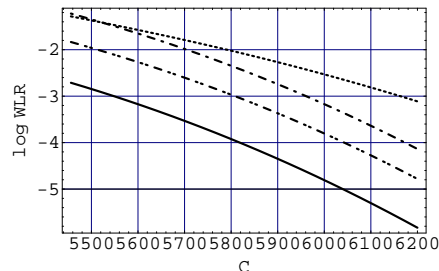
**Fig. 3.** Bounds on  $P_{\text{sat}}$ , Mix 2



**Fig. 4.** Bounds on  $WLR$ , Mix 2



**Fig. 5.** Bounds on  $P_{\text{sat}}$ , Mix 3



**Fig. 6.** Bounds on  $WLR$ , Mix 3

better performance in some cases (not seen in the figures) than the stochastic ordering-based one (see *Observations*), but the gain in capacity savings is not in proportion to the higher complexity of implementation.

The formulae of the stochastic ordering-based bounds are relatively simple, they seem to be implementable (especially the logarithm of the bounds) in a straightforward manner. Consequently, the application of these bounds (especially the  $WLR$  bound) is strongly encouraged, also because of the good performance in accuracy.

## 6 Conclusion

In this paper a family of bounds on saturation probability and workload loss ratio has been set up under the bufferless fluid flow multiplexing framework. This family comprises previously known as well as newly developed bounds. According to the analysis the stochastic ordering-based bounds have the best performance, especially in the case of workload loss ratio approximation. Nevertheless, the simple Hoeffding-based bounds could still form viable alternative from implementation point of view.

## Appendix: Improved Bounds for Independent Regulated Flows Multiplexed into a Service Curve Network Element

In [8] and [9] performance bounds (bounds on buffer overflow probability) have been derived when independent and regulated traffic flows have been multiplexed into a buffer with fixed service rate. These bounds have been extended in [10] to the case of a more general multiplexer model, i.e. the service is characterized by a general service curve<sup>4</sup>.

All the buffer overflow bounds presented in [8], [9] and [10] relies on the use of Hoeffding's inequalities [5], also presented in this paper in equation (12) and (14) for the homogeneous and heterogeneous case, respectively. Our closed form bounds presented in (18) and (20) are also based on and improve (see [1] and Section 4) one of the Hoeffding's inequalities (14), and use the same amount of information. Therefore, they can apparently be used for further improving the results of Vojnovic and Le Boudec in [10] in the case of heterogeneously regulated traffic multiplexed into a service curve network element. For the same reason, improved bounds on buffer workload loss ratio can also be set up by the use of (23) and (24). The performance evaluation of the improved bounds in the buffered multiplexer context is a matter of future work.

## References

1. Heszberger, Z., Zátónyi, J., Bíró, J.: Efficient chernoff-based resource assessment techniques in multi-service networks. *Telecommunication Systems* **20** (2002) 59–80
2. Charzinski, J.: Fun factor dimensioning for elastic traffic. In: *ITC Specialist Seminar on Internet Traffic Measurement, Modeling and Management*, Monterey, CA, USA (2000)
3. Grossglauser, M., Tse, D.: A framework for robust measurement-based admission control. In: *ACM SIGCOMM Symposium on Communications Architectures and Protocols*, Cannes, France (1997) 237–248
4. Mao, G., Habibi, D.: Loss performance analysis for heterogeneous on-off sources with application to connection admission control. *IEEE/ACM Transactions on Networking (TON)* **10** (2002) 125–138
5. Hoeffding, W.: Probability inequalities for sums of bounded random variables. *Journal of the American Statistical Association* **58** (1963) 13–30
6. Hsu, I., Walrand, J.: Admission control for ATM networks. *Stochastic Networks* (1995) 413–429 Springer Verlag, New York.
7. Bíró, J.J., Heszberger, Z., Martinecz, M.: Equivalent capacity estimators for bufferless fluid flow multiplexing. In: *IEEE Globecom*, San Fransisco, CA (2003)
8. Kesidis, G., Tassiulas, L.: Traffic shaping for a loss system. *IEEE Communication Letters* (Dec. 2000) 417–419
9. Chang, C.S.: On the performance of multiplexing independent regulated inputs. In: *Sigmetrics*, Cambridge, Massachusetts, USA (2001)
10. Vojnovic, M., Le Boudec, J.Y.: Bounds for independent regulated inputs multiplexed in a service curve network element. *IEEE Trans. on Communications to appear* (2003)

---

<sup>4</sup> In this model the multiplexer is referred to as service curve network element.