A Closed Form Formula for Long-lived TCP Connections Throughput

Augustin Chaintreau, Danny de Vleeschauwer

HAL Id: inria-00072145
https://inria.hal.science/inria-00072145

Submitted on 23 May 2006

INRIA

A Closed Form Formula for Long-lived TCP Connections Throughput

Augustin Chaintreau — Danny De Vleeschauwer

N° 4443

Avril 2002

THÈME 1

R apport de recherche

A Closed Form Formula for Long-lived TCP Connections Throughput

* Augustin Chaintreau[†] , Danny De Vleeschauwer[‡]

Thème 1 — Réseaux et systèmes
Projet TREC

Abstract: In this paper, we study the variation of the throughput achieved by TCP resulting from both the individual behavior of a connection and the interaction with all other connections sharing the same link. In particular, we calculate the Tail Distribution Function (TDF) of the instantaneous throughput seen by one TCP connection in the Additive Increase Multiplicative Decrease (AIMD) framework. For the particular case that each TCP connection experiences the same Round Trip Time (RTT) and under the many user approximation we prove that this TDF is given by a closed-form formula that solely depends on the network parameters (number of sources, capacity and buffer size of the bottleneck link). This formula can then be used as a dimensioning tool, where throughput is guaranteed to each user to be "larger than a given value for at least a certain percentage of the time". In the context defined here, this formula plays the same role for the dimensioning of an IP router as the Erlang B formula does for the dimensioning of a PSTN switch.

Key-words: TCP Performance fluctuation, QoS, Dimensioning, Effect of synchronization between flows, Mean field, Laplace calculus

Une formule close pour le débit des connections durables contrôlées par le protocole TCP

Résumé : Le débit variable offert à une connection durable par le protocole TCP est ici étudié comme le résultat de son comportement individuel - additivement incrémental multiplicativement décrémental - et de la coopération de tous les flots qui partagent un même lien de goulot d'étranglement. Dans un cas homogène où toutes les connections, observent le même délai de bout en bout, nous montrons ici que le débit instantané peut être explicitement caractérisé sous une approximation de grands nombres d'usagers. Cette dernière hypothèse ne nécessite pas comme nous le justifions expérimentalement un nombre d'utilisateur considérable. Une formule close décrivant la fonction de répartition du débit obtenu par un utilisateur en état stationnaire comme la somme d'une série en est ensuite déduite par un calcul de Laplace. Ce résultat est ensuite utilisé pour étudier l'effet de la taille du buffer du routeur goulot d'étranglement sur le débit obtenu par les connections TCP sur ce lien, étude que nous comparons ensuite à des résultats obtenus avec le simulateur NS. Nous analysons ensuite à l'aide de la même formule un cas de dimensionnement de réseau contrôlé par TCP pour une garantie de performance probabiliste (un débit instantané garantie pour une portion du temps). Cette formule joue dans le contexte défini ici le même rôle que la formule d'Erlang pour un commutateur téléphonique.

Mots-clés : Controle de congestion, Fluctuation de performance autorisée par TCP, Garantie de Performance, Dimensionnement de réseaux, effet de Synchronisation entre connections en boucle close, Champ moyen, Calcul de Laplace

Contents

1   Introduction

Today, most Internet flows use the TCP protocol on top of the IP packet level to provide reliable transmission of data, and to avoid congestion. The throughput achieved by each TCP connection sharing available resources on the network is not easy to predict as it depends on both the individual behavior of the connection and the packet losses occuring in a router, created by the interaction of all the connections that use it.

In particular, dimensioning a link to provide performance guarantees to users is difficult as resources are shared by all the connections. Fluctuations of the bandwidth used by any connection implies fluctuation on the performance seen by a user, and this needs to be taken into account to offer a probabilistic performance guarantee, i.e. achieving a throughput larger than a certain value for at least a percentage of the time.

Modeling the transmission of packets through the routers of a connection with other flows in the network, and the reaction of the congestion control algorithm as implemented in TCP as a discrete event system has turned out to be very complex, as was already shown in [4]. In particular the expected throughput of a connection does not only depend on the arrival rate of packets from other flows sharing the same router, but on fine statistical properties of this process. In [5] a formula has been proposed to link the expected value of the window with the probability of packet loss.

Here we propose a fluid flow approximation to capture the control loop feedback that governs the behavior of TCP connections sharing a link. In this framework we do not only derive the expectation of the throughput but also its steady state probability distribution. To the best of our knowledge this is the first time where the effect of the TCP behavior and the interaction of all connections leads to a closed form formula giving the fluctuations of the throughput seen by one connection.

As in our analysis we assume that the TCP connections are persistent and use the maximum throughput allowed by TCP, i.e. their number (denoted by N) is fixed and each user remains transmitting data at its maximum rate, we can only apply our theory to FTP downloads and Peer-to-Peer applications. Our analysis cannot capture non-persistency which occurs for example in HTTP traffic, where the traffic demand per user follows an on-off pattern. We analyze the simple case where the connections are all sharing one bottleneck link with capacity C and buffer size B, which governs their throughput. Traffic here is assumed to be homogeneous (i.e. each of the considered connections has the same RTT R). These points are discussed further in the conclusion.

The end-to-end congestion control implemented by the connections is modeled as a fluid flow Additive Increase Multiplicative Decrease (AIMD) evolution of the instantaneous throughput achieved by each connection, already introduced in [1]. In other terms we are not taking into account the occurrence of Time-Outs and their consecutive Slow-Starts.

We analyze the Many Users Approximation of the Fluid Flow AIMD model as introduced in [2]. This assumption will be clarified in section 4 where experiments are presented to justify this approach. In particular we observe that the number of connections considered does not need to be very large ; N greater than 16 is sufficient in many cases.

We show in this context that the law of the throughput achieved by a connection can be characterized explicitly. A closed form formula is then deduced that gives the Tail Distribution Function (TDF) of the instantaneous throughput achieved by one connection as a 'Fourier type' sum, whose convergence and exploitation are then described.

We use our method to study some aspects of a network with TCP connections. First, estimating the impact of the buffer size on synchronization between connections, we observe that increasing the buffer size may in some cases imply a decrease of expected throughput due to higher synchronization, as opposed to the case of an open loop model. We then describe the fluctuation of instantaneous throughput and study a case of statistical performance guarantee.

 In section 2, we develop our model, which captures the feedback loop of TCP controlling all the connections as random matrix products. We then introduce the key concept of *synchronization rate*, playing the role of a parameter in our model. In the following section, the effect of the buffer size is studied and included in our framework as a correction of the value of the link capacity. We then propose a method to estimate both this correction and the synchronization rate, based on a calculation for a Markovian queue. The many users approximation is then introduced in section 4 and justified by simulations ; this assumption allows us to express the steady state distribution of the instantaneous throughput given to one connection as a sum of independent random variables whose laws are known.

   The main new result of the paper is shown and proved in details in section 5. It allows us to express the TDF of the stationary law of the instantaneous throughput seen by one connection as a closed form formula on the network parameters (Number of sources, capacity and buffer size of the bottleneck link). As this expression includes infinite sums and products, a numerical algorithm is described in section 6 to compute the formula giving the TDF with an arbitrary precision. Numerical examples, as well as the exploitation of this formula, and a comparison with NS simulations, are then described in section 7.

 At some stage of the analytical study of section 5, mathematical notions such as the Stieltjes and Contour Integral are necessary to understand the proof. References are given in the text when these notions have to be introduced. This mathematical section can be skipped by readers who are only interested in using the formula.

2   Fluid Flow AIMD Model

2.1   Description

The Fluid Flow AIMD model, as introduced in [1], considers for a connection (denoted by index n), its instantaneous throughput at time $t \in \mathbb{R}$ denoted by $X^{(n)}(t)$, which increases linearly with a slope $\frac{1}{R^2}$ for each connection. As a result of the finite bandwidth available on the link (denoted by C), a congestion epoch occurs when the sum of their throughput reaches the capacity, where some connections may suffer packet losses and halve their throughput. We first consider the buffer-less case in this section. We come back to the case of a bottleneck router with a nonzero buffer in section 3.

   We introduce T_i the time of the $i\_$th congestion epoch, and $X_i^{(n)}$ the instantaneous throughput achieved by connection n taken immediately after the $i\_$th congestion epoch : $X_i^{(n)} = X^{(n)}(T_i+)$. From the AIMD behavior of the simplified TCP protocol, illustrated on Figure 1, we deduce the following recurrence equation :

$$X_i^{(n)} = \gamma_i^{(n)} \left(X_{i-1}^{(n)} + q_i \right) , \quad \text{where} \quad q_i = \frac{1}{R^2}(T_i - T_{i-1}) , \tag{1}$$

- where $\gamma_i^{(n)}$ is a random variable equal to 1, except if the connection n suffers a loss at time T_i, in which case it is $\frac{1}{2}$.

- q_i as defined here is the throughput growth of the connection since the last congestion epoch. This is the same for each connection, as each of them has the same RTT.

 By expressing that at the $(i+1)\_$th congestion epoch the capacity C is just consumed, it can be seen that the increase q_i does not only depend on the throughput of the n-th connection, but also on all the throughputs of all other connections, hence, resulting in an interaction between TCP sources.

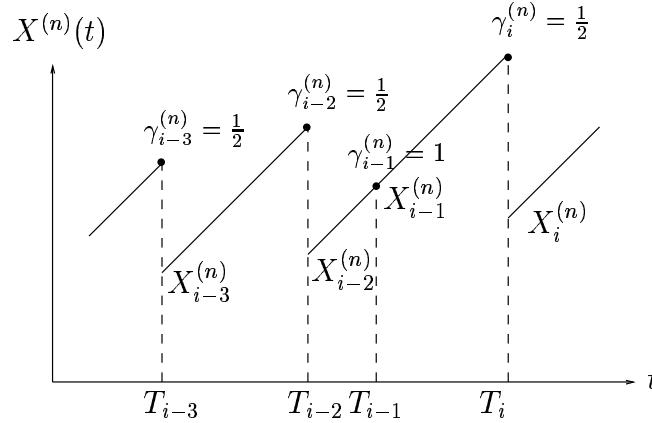$$\sum_{n=1}^{N} \left(X_{i-1}^{(n)} + q_i \right) = C , \tag{2}$$

Figure 1: Evolution of the instantaneous throughput of a connection

$$\text{hence } q_i = \frac{C - S_{i-1}}{N} \text{ where } S_i = \sum_{n=1}^{N} X_i^{(n)} \ . \tag{3}$$

Replacing the expression for q_i in Equation 1, we capture the AIMD competition between TCP connections on the link in a recurrence equation with a product of random matrices, whose elements contain the random variables $(\gamma_i^{(n)})_{N=1...N, i \in \mathbb{Z}}$.

2.2   Synchronization rate

In our model, packet losses are taken into account in the value $\frac{1}{2}$ taken occasionally by variables $(\gamma_i^{(n)})_{n=1...N, i \in \mathbb{Z}}$.

We assume that values of these variables are independent between different connections, and also that they are independent in time. By homogeneity, we assume that they have the same law for all connections, and for all congestion epochs.

In this case, the distribution of all random variables in the system is fixed if we choose the value of the probability $p = P(\gamma_i^{(n)} = \frac{1}{2})$ (this does not depend on n neither i) that is referred to as *synchronization rate*. It represents both the proportion of connections suffering at least one packet loss in a congestion epoch, and the probability for one connection to lose at least one packet when congestion appears.

Another model has been introduced in [2] where, for each connection epoch T_i, the variables $\gamma_i^{(1)}, \ldots, \gamma_i^{(N)}$ are taken independent, but with different law which varies with the throughput of the corresponding connection before the congestion (i.e. $P(\gamma_i^{(n)} = \frac{1}{2})$ is a function of $X_{i-1}^{(n)} + q_i$).

We have considered here the simplest model to be able to provide explicit characterization of the stationary law of the throughput, and to derive an analytical formula.

2.3   Expectation and stationary distribution

The parameter p is now fixed. Iterating the recurrence equation, we find the expansion of an infinite sum.

$$\begin{aligned} X_i^{(n)} &= \gamma_i^{(n)} q_i + \gamma_i^{(n)} X_{i-1}^{(n)} , \\ X_i^{(n)} &= \gamma_i^{(n)} q_i + \gamma_i^{(n)} \gamma_{i-1}^{(n)} q_{i-1} + \gamma_i^{(n)} \gamma_{i-1}^{(n)} X_{i-2}^{(n)} , \\ &\vdots \end{aligned}$$

By a result shown in [1] this infinite sum converges and the limit expresses the steady state distribution of the sequence

$$X_{\infty}^{(n)} = \sum_{j \geq 0} \gamma_i^{(n)} \ldots \gamma_{i-j}^{(n)} q_{i-j} \; . \tag{4}$$

Expectation of the throughput seen by one source can be easily calculated :

From $\;q_i = \dfrac{C}{N} - \dfrac{1}{N} S_{i-1}\;$ we deduce $\;\mathbb{E}[q_i] = \dfrac{C}{N} - \dfrac{\mathbb{E}[X_{i-1}^{(1)}] + \ldots + \mathbb{E}[X_{i-1}^{(N)}]}{N} = \dfrac{C}{N} - \mathbb{E}[X_{i-1}^{(n)}]\; .$

So that $\mathbb{E}[q_i + X_{i-1}^{(n)}] = C/N$ and, as $\gamma_i^{(n)}$ is independent from $X_{i-1}^{(n)}$ and q_i (expressed with the sum of the throughput immediately after $i\!-\!1\_$th congestion epoch), we have by taking expectations on both sides of the recurrence equation :

$$\mathbb{E}[X_i^{(n)}] = \mathbb{E}[X_{i-1}^{(n)} + q_i]\mathbb{E}[\gamma_i^{(n)}] = \frac{C}{N}\left(1 - \frac{p}{2}\right) \; . \tag{5}$$

Note that this expression can be interpreted intuitively : As the total throughput immediately before the congestion is the capacity C, and a proportion p of the connections have halved their throughput during the congestion, this equation only reflects that on average the total throughput immediately after congestion should be $C(1 - \frac{p}{2})$.

3 Buffer and Losses

3.1 Influence of the buffer size

When the sum of the throughputs reached a value larger than the link capacity the buffer occupancy increases. Packet losses occur *as soon as the buffer is full*. Equation 2 hence holds as it is written only in the case where the buffer is of size zero. The general case can be expressed by the same expression where the value of C has to be corrected to $C' = \rho C$ with ρ greater than one due to the filling of the buffer before the losses.

We remark that the buffer size B also has an influence on the RTT R. Indeed, the buffer occupancy can in principle vary between 0 and B. Although we neglect the packet behavior in this paper, as we consider a fluid flow model to describe TCP, in reality the IP packets carry the data. Some of these IP packets hit an empty buffer in the bottleneck router under consideration, while others hit a buffer that is practically full. The latter ones need to queue for an additional time $\frac{B}{C}$ with respect to the former ones. As we have assumed a constant RTT R in our analysis, this additional queuing time needs to be much smaller than R, otherwise our analysis does no longer apply. Hence, the approach described in this paper is only valid if $B \ll RC$. This boils down to assuming that the RTTs are solely due to constant delays encountered by each IP packet and are not heavily influenced by queuing.

To compute the value taken by ρ we study the **buffer evolution between two successive congestion epochs T_{i-1} and T_i :**

At time $t = T_{i-1}$ a congestion epoch occurs, as the total throughput have reached the value $C' = \rho C$, and the buffer is full containing B packets. Immediately after losses occuring at this time, the total throughput is on average $C'(1 - \frac{p}{2})$ as some of the connection (in the proportion p) have halved their throughput. All connection are then increasing their throughput and the total throughput in the link is getting higher with rate N/R^2, such that we have :

$$\text{for } t \text{ in }]T_{i-1}; T_i], \; X^{(\mathtt{total})}(t) = C'(1 - \frac{p}{2}) + \frac{N}{R^2}(t - T_{i-1}) \; .$$

We cut the interval $]T_{i-1}; T_i]$ into two part, $]T_{i-1}; \tilde{T}_i]$ and $]\tilde{T}_i; T_i]$, introducing $\tilde{T}_i$ the time of this interval where the total throughput is the actual value of the link capacity C. We have :

$$X^{(\mathtt{total})}(\tilde{T}_i) = C \;\text{ so that }\; \frac{N}{R^2}(\tilde{T}_i - T_{i-1}) = C - C'(1 - \frac{p}{2}) \; .$$

During the first part of the interval, the number of packet in the queue is decreasing as $X^{(\mathtt{total})}$ is then smaller than C. The decrease of the queue, that is naturally bounded by B, is then given by :

$$\mathtt{QDec}_i = B(T_{i-1}) - B(\tilde{T}_i) = \min\left(B, \int_{T_{i-1}}^{\tilde{T}_i} C - X^{(\mathtt{total})}(t)\,dt\right) \ ,$$

$$\mathtt{QDec}_i = \min\left(B, \left(C - C'(1-\frac{p}{2}) - \frac{1}{2}\frac{N}{R^2}(\tilde{T}_i - T_{i-1})\right)(\tilde{T}_i - T_{i-1})\right) \ ,$$

$$\mathtt{QDec}_i = \min\left(B, \frac{R^2}{N}\frac{C^2}{2}\left(1 - \rho(1 - \frac{p}{2})\right)^2\right) \ . \tag{6}$$

During the second part of the interval, From time $\tilde{T}_i$ to T_i the total throughput is increasing from C to the value C' attained at time T_i where the buffer is full :

$$X^{(\mathtt{total})}(T_i) = C' \ \text{ so that } \ \frac{N}{R^2}(T_i - \tilde{T}_i) = C' - C \ .$$

The increase of the buffer in this period is then given by :

$$\mathtt{QInc}_i = B(T_i) - B(\tilde{T}_i) = \int_{\tilde{T}_i}^{T_i} X^{(\mathtt{total})}(t) - C\,dt \ ,$$

$$\mathtt{QInc}_i = \frac{1}{2}\frac{N}{R^2}(T_i - \tilde{T}_i)^2 = \frac{R^2}{N}\frac{C^2}{2}(\rho - 1)^2 \ . \tag{7}$$

As buffer is full at times T_{i-1} and at time T_i, decrease and increase of the queue exactly compensate. So that the value of ρ verifies :

$$\rho - 1 = \min\left(\chi\,,\ 1 - \rho(1 - \frac{p}{2})\right) \ \text{ where } \ \chi = \frac{\sqrt{2BN}}{RC} \ . \tag{8}$$

It can then be shown (see section A.1 in appendix A), that the value taken by ρ is given by the following

$$\rho = \min\left(1 + \chi\,,\ \frac{1}{1 - \frac{p}{4}}\right) \ , \tag{9}$$

where the term chosen by the minimum exactly corresponds to the term chosen in Equation 8. This means that we have to distinguish between two functioning cases :

- `Empty buffer case` : This corresponds to the case where the min in the expression of `QDec` (Equation 6) takes the value B. In this case the buffer empties before time $\tilde{T}_i$. Since at this time $\tilde{T}_i$ the buffer is certainly empty, we refer to this case as the empty buffer case. It follows easily that

$$C' = C + \frac{\sqrt{2BN}}{R} \ ,$$

  corresponding to $\rho = 1 + \chi$.

- `Never Empty buffer case` : This is the functioning case where the buffer never empties between two congestion epochs and the minimum in Equation 6 is given by the right term. Hence

$$C - C'(1 - \frac{p}{2}) = C' - C \ , \ \text{ such that } \ \rho = \frac{1}{1 - \frac{p}{4}} \ .$$

  Note that in this case the corrected capacity is depending on the synchronization rate.

Which one of these two cases is valid is determined by the minimum chosen in the expression of ρ. (Note that, as $0 \leq p \leq 1$, the right term is bound by $\frac{4}{3}$. Consequently, for $\chi \geq \frac{1}{3}$, the buffer does never empty, and the value is given by the right term.)

3.2   Estimating the synchronization rate

Here we propose a method to choose the value of the parameter p, which is described in full detail in [3]. The formula found for p is then applied to the two different functioning cases that were previously considered (empty buffer case, and never empty buffer case) ; p is in general given by the minimum.

From a study of the Markovian queue M/M/1/B with service rate C and arrival rate $C' = \rho C$ with $\rho > 1$, a calculation shows that the probability for a connection to lose at least one packet in a congestion epoch is

$$p = \frac{1 - \exp\left(-\frac{RC}{N}\frac{\rho-1}{1-(\frac{1}{\rho})^{B+1}}\right)}{1 - \exp\left(-RC\frac{\rho-1}{1-(\frac{1}{\rho})^{B+1}}\right)} = f(\rho) \ . \tag{10}$$

- In the **Empty buffer case**, the value of ρ is known equal to $1 + \chi$, which gives directly a value for p.

- In the **Never empty buffer case** this is slightly more complex as the value of ρ depends on p itself. p is then the solution of a fixed point equation $p = g(p)$ where $g(p) = f(\frac{1}{1-\frac{p}{4}})$. Very generally, this function admits a unique fixed point in $[0; 1]$, which can be evaluated as the limit of the following iterative scheme.

$$\begin{cases} p_0 = \frac{1-\exp\left(-\frac{RC}{N}\frac{1}{B+1}\right)}{1-\exp\left(-RC\frac{1}{B+1}\right)}, & \text{which is the limit of } g(p) \text{ in } p \to 0 \ , \\ p_{m+1} = g(p_m) \ . \end{cases} \tag{11}$$

Again we have observed that p is more generally given by the minimum of these two possible values (see section A.2 in appendix A). Which one is the smallest determines whether the empty buffer case or never empty buffer case applies.

The value of parameter p is then entirely fixed by characteristics of the network C, R, B, N. We illustrate this method on examples in Section 7. Although in Section 7 we use this scheme to estimate p and ρ, variants on this scheme to estimate these parameters exist. Since the theory developed further to calculate the TDF of the throughput is independent of the scheme with which the parameters p and ρ are estimated, any of these alternative schemes could have been used.

4   Many Users Approximation

4.1   Description

Here we study the behavior of our model for a large number of connections, and see how this simplifies the expression of the steady state law. Experimental validations of this approximation have been conducted and are presented at the end of this section.

The recurrence equation gives relations between values of throughput for different connections and different epochs. In particular one flow interacts with another one only in the expression of the inter-congestion time, through the average $\frac{1}{N}S_{i-1}$, found in the expression of q_i, made by averaging N variables with same expectation equal to $\frac{C'}{N}(1 - \frac{p}{2})$.

When N is increased while keeping the same capacity ratio C'/N, the variable S_{i-1} is obtained as an average of more and more random variables with same expectation. From the study of their inter-correlation we can show that the variance of the average vanishes and that this average can be taken in the model to be equal to its expected value (proof can be found in [2]) :

$$\frac{S_{i-1}}{N} \sim \frac{C'}{N}(1 - \frac{p}{2}) \text{ if } N \text{ is large enough, hence } q_i = \frac{C' - S_{i-1}}{N} \sim \frac{C'}{N}\frac{p}{2} = q \ .$$

Under this approximation, the variables q_i converge for all i to their expected values which are equal to the same constant q . The recurrence simplifies to a one-dimensional equation :

$$X_i^{(n)} = \gamma_i^{(n)} \left(X_{i-1}^{(n)} + q \right) \quad \text{and we have} \quad X_\infty^{(n)} = q \sum_{j \geq 0} \gamma_i^{(n)} \ldots \gamma_{i-j}^{(n)} . \tag{12}$$

4.2   Steady state law as the result of an infinite geometric walk

In the infinite sum defining the steady state law in Equation 12, we can deduce the $j+1$ terms from the term number j by an independent coin tossing ; with probability $1-p$ they are equal and with probability p the latter term is half the first one. Consequently, the steady state law is the result of the "infinite geometric walk" represented on the Figure 2, where after each step we decide independently with probability p to halve the step length or else keep the same value.
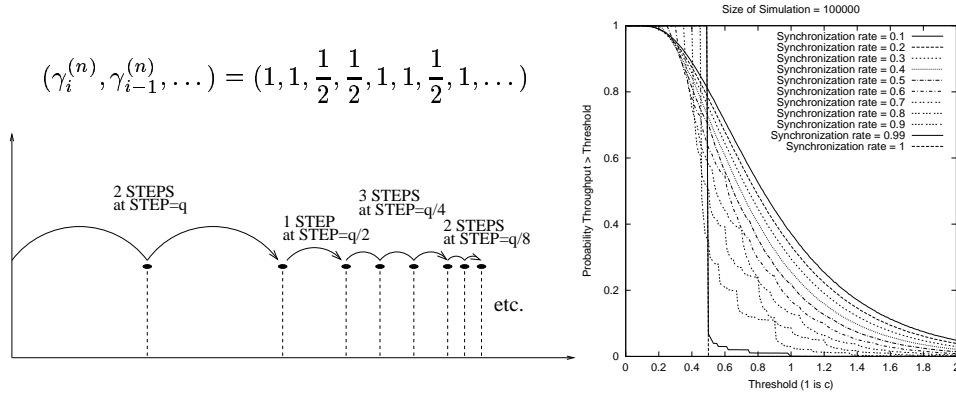

Figure 2: Description of the infinite geometric walk (left), and its TDF for different values of p (right)

We can observe that the number of steps made in this walk with a certain fixed length $\frac{q}{2^k}$ is thus geometrically distributed with a ratio $1-p$. The variable may be zero if we consider the step length q (a division can occur immediately in the first term), but for any smaller length $\frac{q}{2^k}$ with $k \geq 1$, there is at least one step with this length. Also these variables are independent. The steady state distribution of the sequence $(X_i^{(n)})_{i \in \mathbb{Z}}$ is then the same as the distribution of :

$$X_\infty = qA_0 + \sum_{k \geq 1} \frac{q}{2^k}(1 + A_k) = q + q \sum_{k \geq 0} \frac{A_k}{2^k} , \tag{13}$$

- where $A_0, A_1, \ldots$ are independent geometric random variables starting in 0 with ratio $(1-p)$:

$$\text{Prob}[A_k = m] = p(1-p)^m, \forall m \geq 0.$$

We have represented this distribution, for different values of p and capacity such that $c = \frac{C'}{N} = 1$, on the right of Figure 2. One can observe for p close to 1 a self similar nature of the TDF.

4.3   Throughput in continuous time

From the continuous time process of the instantaneous throughput $(X^{(n)}(t))_{t \in \mathbb{R}}$, only the embedded chain made of samples of this process taken immediately after congestion epoch $(X^{(n)}(T_i+))_{i \in \mathbb{Z}}$ have been considered. The entire continuous time process can easily be retrieved as the throughput of each connection increases linearly with a constant slope between congestion times.

In the Many Users Approximation, the value q_i of the throughput increase between T_{i-1} and T_i is equal for all i to the same constant q. Inter-congestion time are then constant. Adding the

variable qU -where U is independent, uniformly distributed on $[0;1]$ - to the steady state law of the embedded chain then leads to the stationary distribution of the continuous time process.

Characterization of the stationary distribution : Under the many users approximation, the stationary distribution of the throughput obtained by a connection has the same distribution as the variable :

$$X(\infty) = q + qU + q \sum_{k \geq 0} \frac{A_k}{2^k} \; , \tag{14}$$

- where q is a constant equal to $\frac{C'}{N} \frac{p}{2}$,

- U is uniformly distributed in $[0;1]$,

- A_k has a geometric distribution starting in 0, with ratio $(1-p)$,

- all the variables here are independent.

This expression will be used later on to write a closed form formula giving the TDF of the law of $X(\infty)$.

We can already deduce the expectation for throughput obtained in the steady state by one source for any time t :

$$E[X(\infty)] = E[X_\infty] + qE[U] \;\; \text{so that} \;\; E[X(\infty)] = \frac{C'}{N}(1 - \frac{p}{4}) \; . \tag{15}$$

4.4   Experimental validation

We have conducted experiments to justify the use of the many users approximation, where we have compared for different values of N, as well as for the asymptotic many users approximation, the TDF function observed by simulation of the model, while keeping the capacity per user in the link $c = \frac{C'}{N}$ constant equal to 1.

For finite values of N, we have simulated a large number of congestion epochs step by step, and the consecutive value of the throughput in discrete time (given immediately after congestion). Counting the proportion where a source gets higher throughput than a threshold gives the stationary law's TDF value for this threshold, if the simulation is large enough. For the many users, only one source need to be simulated, as the increase resulting from the interaction of all sources is always equal to a constant q. Results are shown in Figure 3, where again a self similar nature of the TDF for the stationary law of discrete time throughput are observed for high value of p.

The next simulations presented here consider the throughput in continuous time. After each congestion, the value of the inter-congestion time is computed as in Equation 2, so that we deduce the proportion of the time over all the simulation where the continuous time process of the throughput of one sources was above a threshold, giving the values taken by the TDF of the continuous time process stationary law. Simulation is similar for the many-users asymptotic except that the inter-congestion time is equal to a constant that can be computed directly, and only one connection needs to be evaluated. The TDF as observed by simulations are shown on Figure 4 and 5 where extremal values of p are presented.

These experiments show that the many user approximation is very accurate with a very small number of connections. For a number of connections greater than 16, whatever be the value of p, it is very hard to distinguish between the TDF and the value given by its many users approximation.
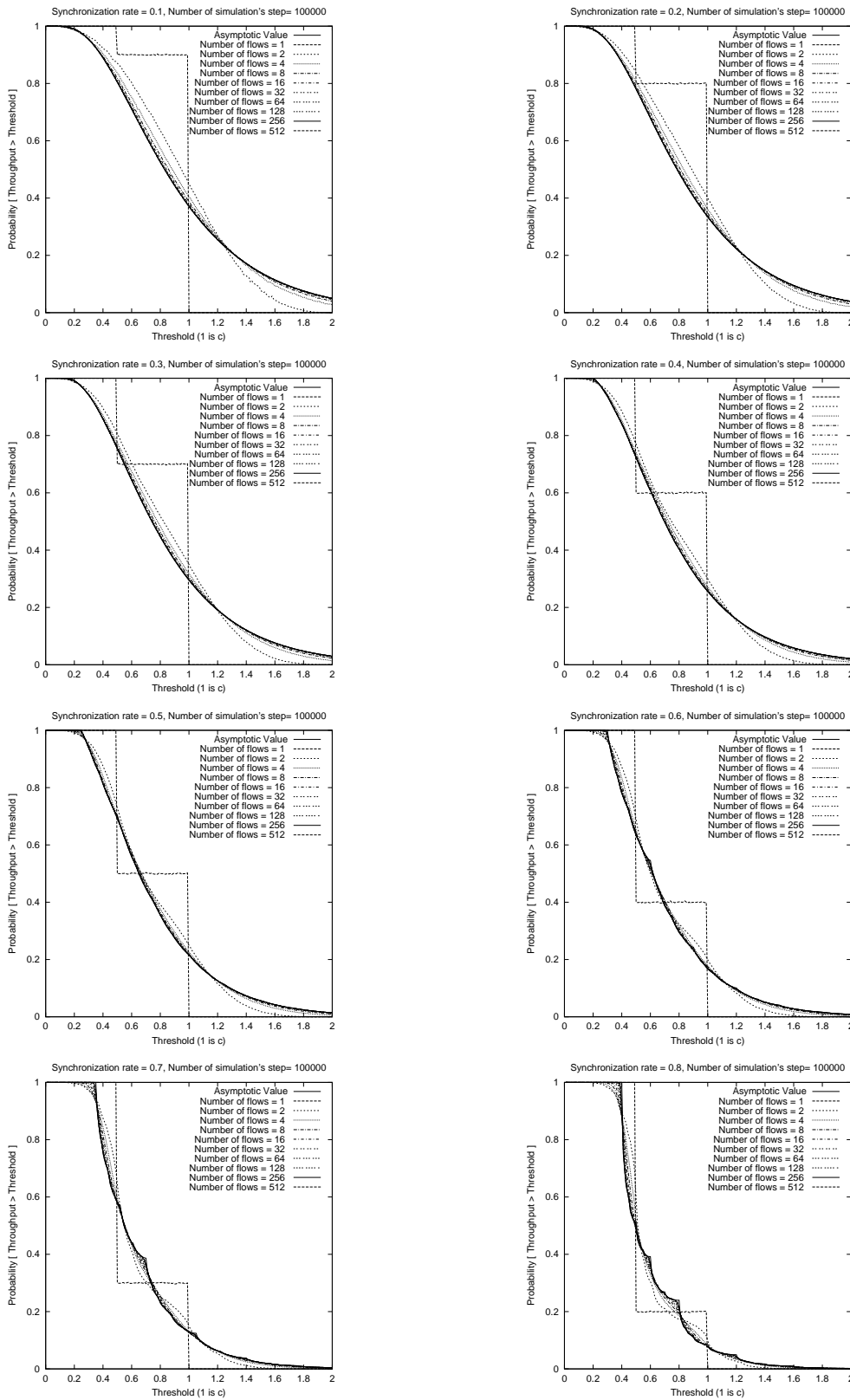
Figure 3: TDF in finite case and many users asymptotic for discrete time throughput
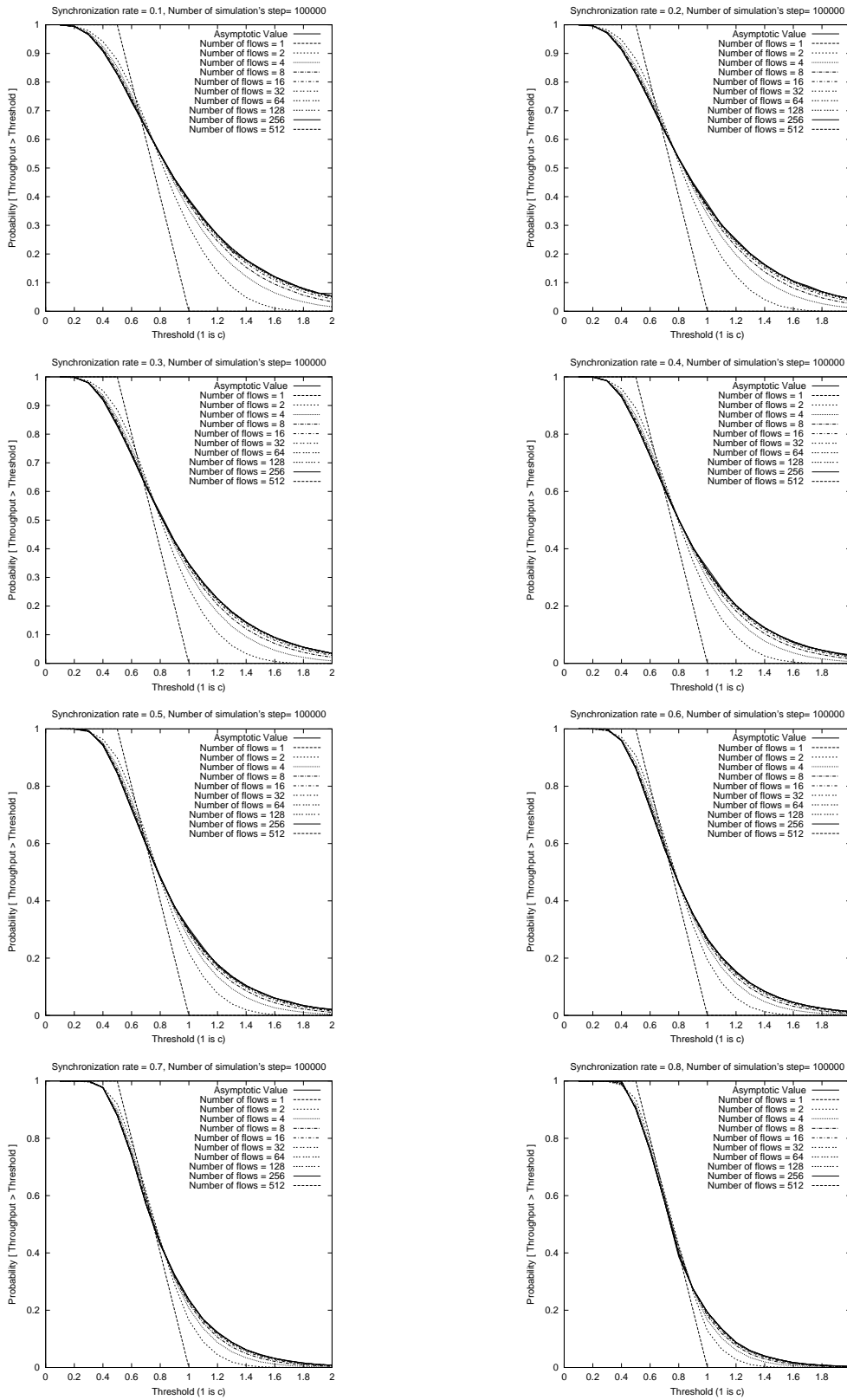
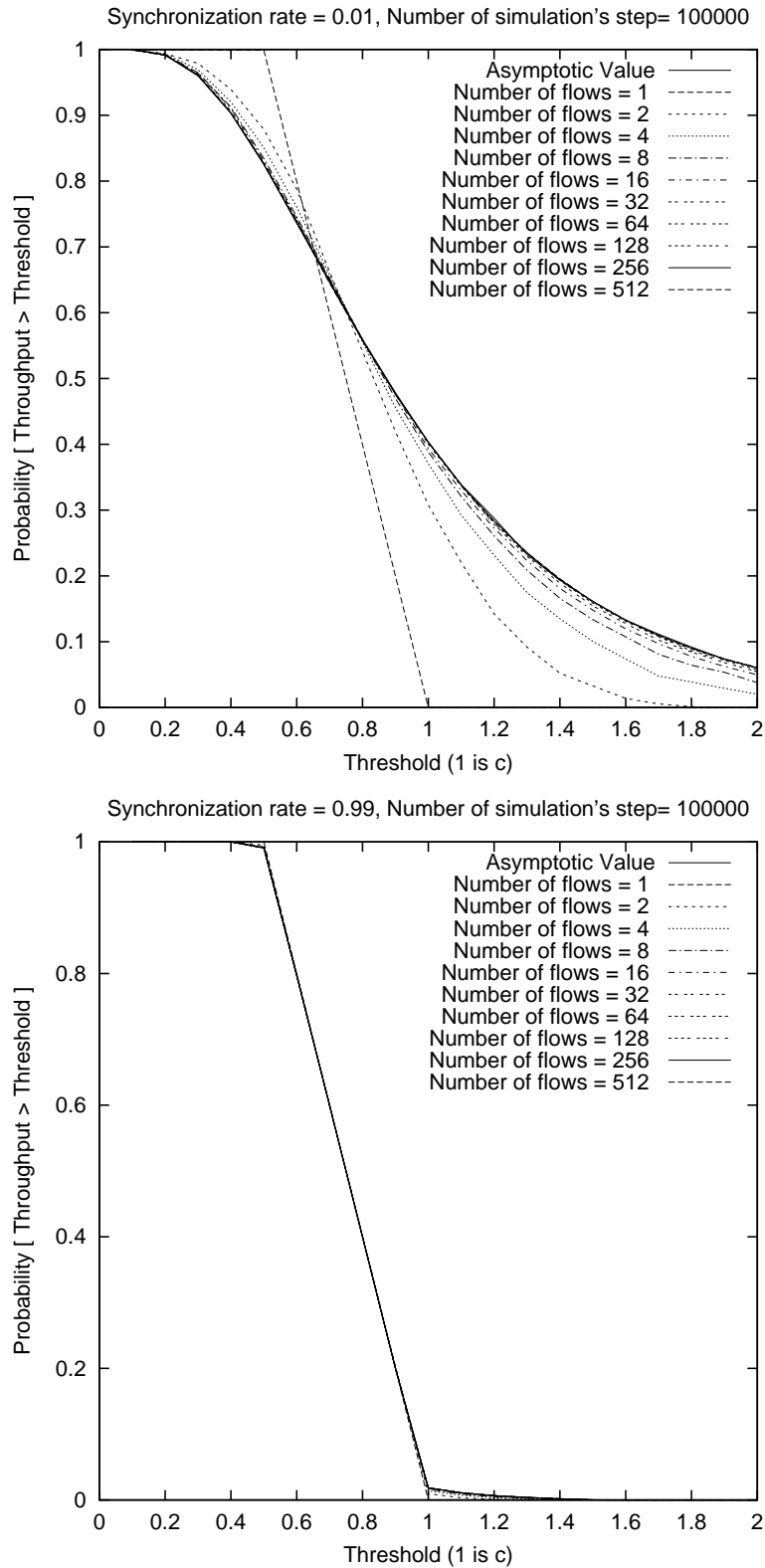Figure 4: Throughput TDF in finite case and many users asymptotic (p=0.1 to p=0.8)

Figure 5: Throughput TDF in finite case and many users asymptotic (Extremal values of p)

5 A Closed Form Formula for the Throughput

We derive based on Equation 14 the following formula that gives the TDF of the stationary throughput obtained by one TCP connection as an infinite sum containing infinite products :

$$P(X(\infty) \geq q + qx) = \sum_{k \geq 0} \alpha_k(x)\phi^{(+\infty)}(2^k x) \,, \tag{16}$$

where $\phi^{(+\infty)}$ is a Fourier series with coefficients $(\beta_l^{(+\infty)})_{l \in \mathbb{Z}}$: $\phi^{(+\infty)}(x) = \sum_{l \in \mathbb{Z}} \beta_l^{(+\infty)} e^{2i\pi l x}$,

$$^1 \text{ and } \begin{cases} \alpha_k(x) = \dfrac{p(1-p)^{2^k x}((\frac{1}{1-p})^{2^k} - 1)}{2^k} \displaystyle\prod_{m=1}^{k} a_m \,, & \text{with } a_m = \frac{p}{1-(1-p)^{1-2^m}} \,, \\[4mm] \beta_l^{(+\infty)} = \dfrac{1}{(\ln(1-p) + 2i\pi l)^2} \displaystyle\prod_{m \geq 1} b_{m,l} \,, & \text{with } b_{m,l} = \frac{p}{1 - e^{-i\frac{2\pi}{2^m}l}(1-p)^{1-\frac{1}{2^m}}} \,. \end{cases} \tag{17}$$

Sketch of the proof : We consider in this section the Cumulative Distribution Function (CDF) of a random variables instead of its TDF, to be consistent with classical mathematical notation. One can be easily deduced from the other as they always sum to one.

Section 5.1 : A general inversion formula expresses the CDF of a random variable as a complex integral of its Laplace transform over a straight line ($s = c + i\mathbb{R}$) that lies in the area where the Laplace transform exists. This Laplace transform is easily computed as $X(\infty)$ is written as a sum of independent random variables.

Section 5.2 : To compute this integral we use the classical Residue method, as the Laplace transform can be extended to an analytic function defined on the whole complex plane excluding its singularities. The path of integration can then be included in a closed contour of the complex plane which may contain some singularities. It is then sufficient to show that the integral over other paths in the contour can be asymptotically neglected. This is not easy to prove directly for the Laplace transform of the variable $X(\infty)$ itself, we then first give the argument for a uniformly distributed random variable, and for a geometric random variable. The proof is then adapted to the case of a truncated version of $X(\infty)$.

A limit argument is presented in Section 5.3 to conclude.

5.1 Laplace transform, Laplace inversion

The Stieltjes Integral is the natural tool to be used here as it allows one to handle within the same framework both a continuous law with density (as for variable U), and a discrete law (as for variables $(A_k)_{k \geq 0}$), as found in the expression of $X(\infty)$. Proofs of the following results, with an introduction to the Stieltjes Integration may be found in [8] (Chap. I and II).

Laplace Transform in Stieltjes Integral : Consider a random variable Y whose CDF μ is of bounded variation, its Laplace transform can be defined by the following Stieltjes integral on the function μ :

$$f_Y(s) = \int_0^{+\infty} e^{-sx} d\mu(x) \text{ for } s = \sigma + i\tau \in \mathbb{C} \,, \tag{18}$$

μ being bounded, this integral is well defined for $\sigma > 0$. And we have, for any $c > 0$, the following convergence known as the **Laplace Inversion formula** :

$$\lim_{T \to +\infty} \frac{1}{2i\pi} \int_{c-iT}^{c+iT} \frac{e^{sx}}{s} f_Y(s)\, ds = \frac{\mu(x+) + \mu(x-)}{2} \quad (= \mu(x) \text{ if } \mu \text{ is continuous at this point}) \,. \tag{19}$$

Laplace Transform in the model : Since the density of the random variable $Y = U + \sum_{k \geq 0} \frac{A_k}{2^k}$ exists, its CDF is continuous. The Laplace transform f_Y is expressed as an infinite product and we have by inversion for all $c > 0$ and $x \geq 0$,

$$\text{as} \quad f_U(s) = \frac{1 - e^{-s}}{s}, \quad \text{and} \quad f_{A_k}(s) = \frac{p}{1 - (1-p)e^{-s}}, \quad \text{the following expression :}$$

$$P(Y \leq x) = \mu(x) = \lim_{T \to +\infty} \frac{1}{2i\pi} \int_{c-iT}^{c+iT} \frac{e^{sx}}{s} \frac{(1 - e^{-s})}{s} \prod_{k \geq 0} \frac{p}{1 - (1-p)e^{\frac{s}{2^k}}} \, ds \, . \qquad (20)$$

Singularities : f_U has no singularity as $\lim_{s \to 0} f_U(s) = 1$, as opposed to f_{A_k} which possesses an infinite number of them, all located on the axis $\sigma = \ln(1-p)$ in points $s = \ln(1-p) + 2il\pi$, for $l \in \mathbb{Z}$, as shown in Figure 6. Consequently, the Laplace transform for variable Y has singularities in points $s = 2^k \ln(1-p) + 2i(l2^k)\pi$, for $k \geq 0, l \in \mathbb{Z}$.

Figure 6: Geometric random variables : singularities of the laplace transform

5.2   Computing the integral as an evaluation of residues

It does not seem easy to justify directly to compute the indefinite complex integral that is found in Equation 20 as an evaluation of the residues of the considered analytic function. One difficulty is that the function to be integrated is given as an infinite product, which is difficult to bound.

First we illustrate the Jordan method on two classical examples made of uniform and geometric distribution. Combination of arguments developed in this two cases leads to a justification for a truncated version of the variable $X(\infty)$.

5.2.1 Preliminary study

A Uniform random variable

Consider U a uniformly distributed random variable on $[0;1]$, its Laplace transform, that converges for any s is :

$$f_U : s \mapsto \frac{1 - e^{-s}}{s}, \text{ extended to 1 in } s = 0.$$
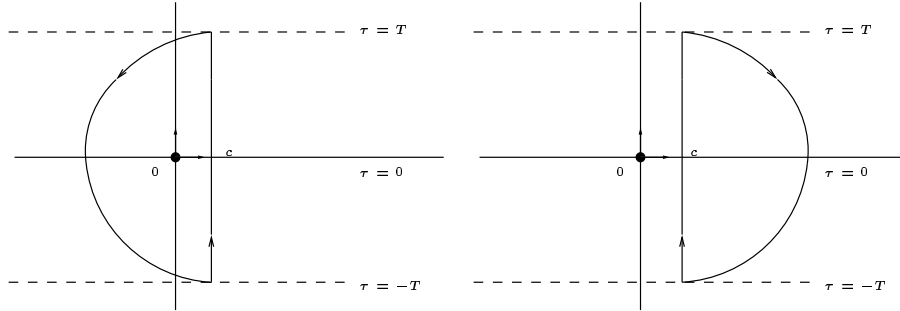
Its CDF μ_U is then given by the following integral (for $c > 0$),

$$\mu_U(x) = \lim_{T \to +\infty} \frac{1}{2i\pi} \int_{c-iT}^{c+iT} \frac{1 - e^{-s}}{s^2} e^{sx} \, ds = J(x) - J(x-1)$$

$$\text{where } \quad J(x) = \lim_{T \to +\infty} \frac{1}{2i\pi} \int_{c-iT}^{c+iT} \frac{e^{sx}}{s^2} \, ds$$

To calculate the integral defining $J(x)$ for any $x \in \mathbb{R}$: one need only to observe that $s \mapsto \frac{1}{s^2}$ is going very fast to zero as the modulus of s becomes large, and that the function $s \mapsto e^{sx}$ is bounded on the left side of the complex plane if x is positive (or similarly on the right side of the complex plane if x is negative).

$\bullet$: Singularities of the function

Figure 7: Contours used to compute the integral defining $J(x)$ as a residue calculus for x positive (left), and for x negative (right)

For $x \geq 0$: The integral $J(x)$ is equal to x.

We complete the vertical line $\{s = c + i\tau, \ -T \leq \tau \leq T\}$ by the semi circle centered in c leading from $c + iT$ to $c - iT$ on the left side on the complex plane, as shown on Figure 7 (left).

The contour lies in $\sigma \leq c$, the modulus of e^{sx} is then bounded by e^{cx}. Besides the modulus of $\frac{1}{s^2}$ is bound by $\frac{1}{(T-c)^2}$ on the semi circle, whose length is πT. Integral over this part of the contour is then bounded by $e^{cx} \frac{\pi T}{(T-c)^2}$ which vanishes for any x when T goes to infinity.

Since the integrandum is analytic except in its singularities, the integral over the contour is equal to the sum all residues that are associated with a singularity inside the contour. For T tending to infinity, this sum of residues tends to the limit of the integral over the straight line that defines $J(x)$.

One singularity of the function lies inside the closed contour considered. We develop the function $s \mapsto \frac{e^{sx}}{s^2}$ as a Laurent Series :

$$\frac{e^{sx}}{s^2} = \frac{1 + (sx) + \frac{(sx)^2}{2} + \ldots + \frac{(sx)^n}{n!} + \ldots}{s^2} = \frac{1}{s^2} + \frac{x}{s} + \frac{x^2}{2} + \ldots$$

So that the Residue is the coefficient for power -1 which is x. Consequently $J(x) = x$ in this case.

For $x \leq 0$ **:** The integral defining $J(x)$ is 0.

A similar closed contour will be considered, where the semi circle lies on the right side of the complex plane, as shown on Figure 7 (right).

On this contour which lies in $\sigma \geq 0$, the modulus $|e^{sx}|$ is bounded by 1, and in particular on the semi circle the modulus of $\frac{1}{s^2}$ is bounded by $\frac{1}{T^2}$. The integral of the function over this part of the contour hence vanishes when T is going to infinity. Because the closed contour does not encircle any singularity and because the integrandum is analytic, the integral over this contour is 0. So, letting T tend to infinity, we find that $J(x)$ is 0.

Conclusion : $J(x) = x$ for $x \geq 0$, else it is null, hence $\mu_U(x) = J(x) - J(x-1) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } 0 < x < 1 \\ 1 & \text{if } x \geq 1 \end{cases}$

Remark : Note that in this case the expression of the CDF μ_U is very easy to deduce from the definition of the considered distribution.

B Geometric random variable

We would like to compute in a similar way the integral giving the CDF of a geometric random variable :

$$\mu_A(x) = \lim_{T \to +\infty} \frac{1}{2i\pi} \int_{c-iT}^{c+iT} \frac{p}{1 - (1-p)e^{-s}} \frac{e^{sx}}{s} ds$$

The contour to consider will be different in this case as singularities of the analytic integrandum do not lie in a bounded domain. We consider the rectangular contour whose right edge is the vertical line of integration of this integral for $T = \pi + 2\pi L$ with $L \in \mathbb{N}$, as represented on Figure 8.
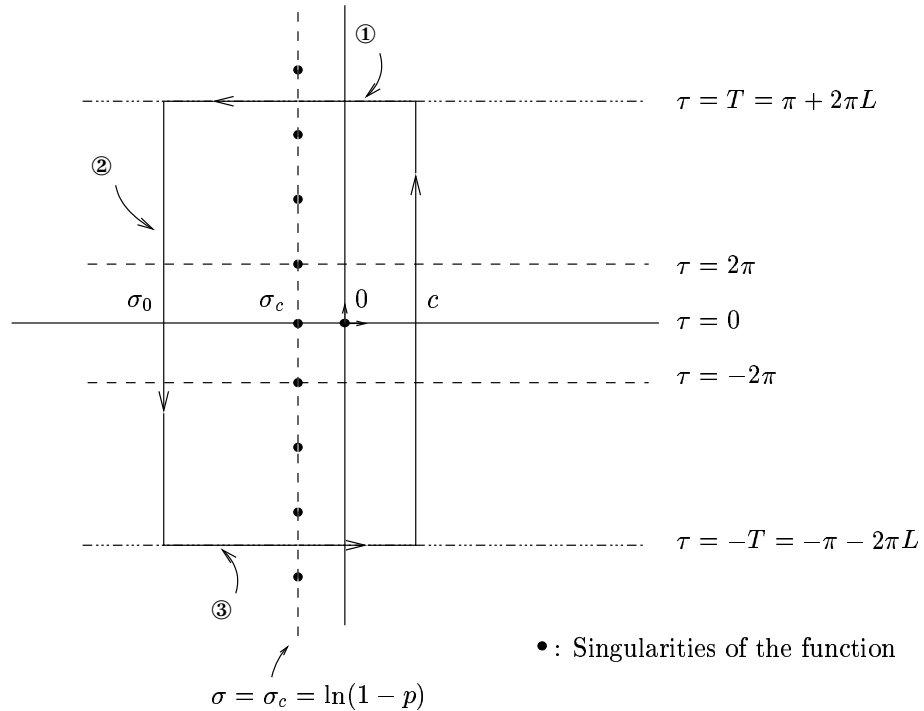

Figure 8: Closed contour used in the computation of the integral giving CDF of a geometric random variable

Integral over ① and ③ : for the top edge of the rectangle we obtain

$$\int_{①} = \frac{1}{2i\pi} \int_{c+iT}^{\sigma_0+iT} \frac{p}{1-(1-p)e^{-s}} \frac{e^{sx}}{s} \, ds = -\frac{1}{2i\pi} \int_{\sigma_0}^{c} \frac{p}{1-(1-p)e^{-\sigma}e^{-iT}} \frac{e^{\sigma x}e^{iTx}}{\sigma+iT} \, d\sigma \; .$$

As T has been chosen to be equal to $2i\pi L + i\pi$, e^{iT} is equal to -1, hence the denominator of the left factor is equal to 1 plus a positive number, and all the left term is bounded in modulus by p. This integral is then bounded in modulus by :

$$\left| \int_{①} \right| \le \frac{p}{2\pi} \int_{-c}^{-\sigma_0} \frac{e^{-\sigma x}}{\sqrt{\sigma^2+T^2}} \, d\sigma \le \frac{p}{2\pi} \int_{-c}^{+\infty} \frac{e^{-\sigma x}}{\sqrt{\sigma^2+T^2}} \, d\sigma \; ,$$

where the second inegality can be written only for $x > 0$. The same bound can be shown for the integral ③. Note that for any x the function to be integrated in this last inegality is monotically decreasing to zero when T increase. As a consequence, this bound vanishes uniformly with the value of σ_0 when T goes to infinity.

Integral over ② : On this edge of the rectangle, the integrandum can be written as

$$\int_{②} = \frac{1}{2i\pi} \int_{\sigma_0+iT}^{\sigma_0-iT} \frac{p}{1-(1-p)e^{-s}} \frac{e^{sx}}{s} ds \quad \text{so that} \quad \left| \int_{②} \right| \le \frac{e^{\sigma_0 x}}{2\pi} \frac{p}{(1-p)e^{-\sigma_0}-1} \int_{-T}^{+T} \frac{1}{\sqrt{\sigma_0^2+\tau^2}} d\tau \; ,$$

as when τ varies, $(1-p)e^{-\sigma_0}e^{-i\tau}$ describes the zero centered circle with radius $(1-p)e^{-\sigma_0} > 1$, so that the distance of this point to 1 is at least $(1-p)e^{-\sigma_0} - 1$.

When T is fixed, this vanishes with σ_0 going to negative infinity, which will be enough to conclude.

Once T is chosen large enough to have the integrals over ① and ③ arbitrary small, and the integral over the right edge to be arbitrary close to the limit (which exists as a consequence of the Laplace inversion formula) the value of σ_0 can be taken arbitrary large to have the integral over ② arbitrary small. This proves both the convergence of the sum of all residues of the function, and the fact that the infinite sum is equal to the limit of integral over the right edge.

When $s \mapsto s_l = \ln(1-p) + 2i\pi l$ we have $\dfrac{s-s_l}{1-(1-p)e^{-s}} = \dfrac{s-s_l}{1-e^{-(s-s_l)}}$ which converge to 1. The residue of the function in this point is then equal to $p\dfrac{e^{s_l x}}{s_l}$. In zero the residue is equal to $\lim_{s\to 0} \dfrac{p}{1-(1-p)e^{-s}} e^{sx} = 1$ We can then write for any $x > 0$:

$$\mu_A(x) = 1 + p(1-p)^x \sum_{l\in\mathbb{Z}} \frac{e^{2i\pi l x}}{\ln(1-p) + 2i\pi l}$$

5.2.2 Justification of the residues calculus for $K\_$order truncation

$K\_$**order truncation :** We consider here the variable Y where the estimation of the infinite sum has been stopped at the order K,

$$\bar{Y}_K = U + \sum_{k=0\ldots K} \frac{A_k}{2^k}$$

Its Laplace transform is given by the expression defining f_Y where the infinite product has been reduced to its $K+1$ first factors.

$$f_{\bar{Y}_K}(s) = \frac{1-e^{-s}}{s} \frac{p}{1-(1-p)e^{-s}} \cdots \frac{p}{1-(1-p)e^{-s/2^K}} \; .$$

In particular $f_{\bar{Y}_K}$ possesses no singularities in $s = \sigma + i\tau$ for σ strictly smaller than $2^K \ln(1-p)$.

Let's consider the rectangular contour represented on the Figure 9, defined with parameters $T = \pi(2L+1)$ and $\sigma_0 < 2^K \ln(1-p)$.
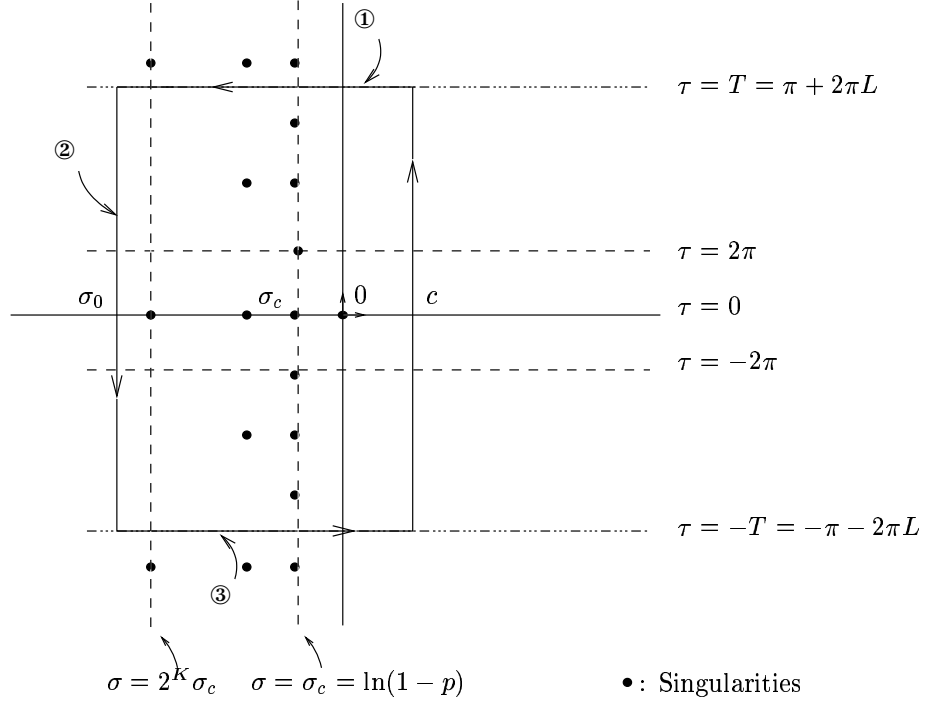
Figure 9: Contour user to compute the integral giving the CDF of finite $K\_$order truncation (here $K = 2$)

- Along the right edge of the rectangle, the integral for T large enough tends to $P(\bar{Y}_K \leq x)$.

- Along edge ②, $s = \sigma_0 + i\tau$ with $-T \leq \tau \leq T$, and the integral is

$$\int_{②} = \frac{1}{2i\pi} \int_{\sigma_0+iT}^{\sigma_0-iT} \frac{e^{sx}(1-e^{-s})}{s^2} \frac{p}{1-(1-p)e^{-s}} \cdots \frac{p}{1-(1-p)e^{-s/2^K}} ds \; ;$$

for any $k = 0\ldots K$, when τ varies $(1-p)e^{-\frac{s}{2^k}}$ is in the circle with center zero and radius $(1-p)e^{-\frac{\sigma_0}{2^k}}$, which is larger than 1. Consequently, its distance to the point 1 in the complex plane is bigger than its radius minus 1, and we have :

$$\left| \int_{②} \right| \leq \frac{e^{\sigma_0 x} + e^{\sigma_0(x-1)}}{2\pi} \left(\frac{p}{(1-p)e^{-\sigma_0}-1} \cdots \frac{p}{(1-p)e^{-\sigma_0/2^K}-1} \right) \int_{-T}^{T} \frac{1}{|\sigma_0+i\tau|^2} d\tau \; . \quad (21)$$

In this last bound, for $x \geq 0$ the coefficient before the integral remains bounded for $\sigma_0 \to -\infty$, as the product of $K + 1$ terms in brackets is equivalent to $\left(\frac{p}{1-p}\right)^{K+1} e^{\sigma_0(1+\frac{1}{2}+\ldots+\frac{1}{2^K})}$. The integral on the right can be bounded for any value of T by the integral of the same function on the infinite domain $]-\infty, +\infty[$, which vanishes when $\sigma_0 \to -\infty$ by monotone convergence. To conclude, integral over ② vanishes for very large negative values of σ_0, uniformly with respect to T.

- On edge ① (integral over ③ is treated similarly) : $s = \sigma + iT$ with $\sigma_0 \leq \sigma \leq c$, and the integral is

$$\int_{①} = -\frac{1}{2i\pi} \int_{\sigma_0+iT}^{c+iT} \left(\frac{p}{1-(1-p)e^{-s}} \cdots \frac{p}{1-(1-p)e^{-s/2^K}} \right) \frac{e^{sx}-e^{s(x-1)}}{(\sigma+iT)^2} ds \; .$$

The product of $K + 1$ terms in brackets can be bounded, as for any $k = 0 \ldots K$ the argument or $z = (1 - p)e^{-\frac{s}{2^k}}$ cannot be closer than $\frac{\pi}{2^K}$ from zero (as $T = \pi$ modulo 2π). In other terms, the complex z cannot be in the domain $\{re^{i\theta}, r \geq 0, |\theta| \leq \frac{\pi}{2^K}\}$ and this implies in particular that its distance to 1 is at least $\sqrt{2}\sqrt{1 - \cos\frac{\pi}{2^K}} = \mathtt{dist}_K > 0$,

$$\text{and thus } |\int_{①}| \leq \frac{1}{2\pi} \left(\frac{p}{\mathtt{dist}_K}\right)^{K+1} \int_{\sigma_0}^{c} \frac{e^{-\sigma x} + e^{-\sigma(x-1)}}{\sigma^2 + T^2} d\sigma . \tag{22}$$

For σ_0 fixed, this integral vanishes with $T = \pi(2L + 1)$ becoming large for $L \to +\infty$.

Asymptotically, for any arbitrary precision $\varepsilon > 0$, the σ_0 can be taken negative and large enough to have integral over ② smaller than ε for any value of T. If T is $\pi(2m + 1)$ with m large enough, integrals over ① and ③ are both smaller than ε. Residue calculus can then be used as the integral for the right edge of the contour extended to infinity is equal to the sums of Residues on the left side of this line.

5.2.3   Computing the residue, finite order formula

We want to compute all the residues for function : $s \mapsto \frac{e^{sx}(1 - e^{-s})}{s^2}\left(\frac{p}{1 - (1-p)e^{-s}} \cdots \frac{p}{1 - (1-p)e^{-s/2^K}}\right)$.

- For $s \to 0$: every term in the product of $K + 1$ terms in brackets is converging to one, the residue is given by a development in Laurent's Series of the coefficient on the left, and is equal to 1.

- For $s \to s_{k,l} = 2^k \ln(1 - p) + 2i\pi(2^k l)$, where $k = 0 \ldots K$, $l \in \mathbb{Z}$, the coefficient on the left is equal to :
$$\frac{(1 - p)^{2^k x} - (1 - p)^{2^k(x-1)}}{(2^k)^2} \frac{1}{(\ln(1 - p) + 2i\pi l)^2} e^{2^k 2i\pi l x} .$$

Terms appearing in the product are $\frac{p}{1 - (1-p)e^{-s/2^j}}$ with $j = 0 \ldots K$:

 - for j equal to k, $\lim_{s \to s_{k,l}} (s - s_{k,l})\frac{p}{1 - (1 - p)e^{-s/2^k}} = p(2^k)$,

 - else for $j = 0 \ldots (k - 1)$, the value does not depend on l and is a_{k-j}, where $a_m = \frac{p}{1 - (1 - p)^{1 - 2^m}}$,

 - for $j = (k + 1) \ldots K$, the factor is equal to $b_{j-k,l}$, where $b_{m,l} = \frac{p}{1 - e^{i\frac{2\pi}{2^m}l}(1 - p)^{1 - \frac{1}{2^m}}}$.

Then the residue in this point is equal to

$$R_{k,l}^{(K)} = - \underbrace{\frac{p\left((1 - p)^{2^k(x-1)} - (1 - p)^{2^k x}\right)}{2^k} \prod_{m=1}^{k} a_m}_{=\alpha_K(x) \text{ and does not depend on } l} \underbrace{\frac{1}{(\ln(1 - p) + 2i\pi l)^2} \prod_{m=1}^{K-k} b_{m,l} \, e^{2i\pi l 2^k x}}_{=\beta_l^{(K-k)}} .$$

So, we find a formula, similar to Formula 16, that is valid for $K\_$ order truncation :

$$P(\bar{Y}_K \leq x) = 1 + \sum_{k=0\ldots K} \sum_{l \in \mathbb{Z}} R_{k,l}^{(K)} = 1 - \sum_{k=0\ldots K} \alpha_K(x)\phi^{(K-k)}(2^k x) , \tag{23}$$

where $\phi^{(K-k)}$ is given by a Fourier series, $\phi^{(K-k)}(x) = \sum_{l \in \mathbb{Z}} \beta_l^{(K-k)} e^{2i\pi l x}$.

5.3   From the Finite order case to the exact formula

Random variables $\bar{Y}_K$ are monotonically converging for $K \to +\infty$ to Y. Consequently, if CDF $\bar{\mu}_K$ for a truncated variable pointwise converges to a function μ, this implies that the limit function is the limit variable's CDF (i.e. $\mu(x) = P(Y \le x)$). The rest of the section will prove that pointwise convergence ; Equation 16 follows.

5.3.1   Study of the coefficient :

- The coefficient $\alpha_k(x)$ converges quickly to zero when k increases. Each of them is derived from the sequence $(a_m)_{m \ge 1}$:

$$a_m = \frac{p}{1 - (1-p)^{1-2^n}} = \frac{1}{1 - \frac{1-p}{p}((\frac{1}{1-p})^{2^n} - 1)} \ .$$

  Coefficients $(a_m)_{m \ge 1}$ are negative except for a finite number of terms in the beginning. For $m \to +\infty$, they are equivalent to $-\frac{p}{1-p}(1-p)^{2^m}$ and approach zero very quickly. In particular we have $|a_m| \le 1$ for m large enough, thus the product of them stay stay bounded in modulus.

  We have the following asymptotic comparison,

$$\alpha_k(x) = \frac{p(1-p)^{2^k x}((\frac{1}{1-p})^{2^k} - 1)}{2^k} a_k \prod_{m=1}^{k-1} a_m = O(\frac{1}{2^k}) \ \text{ when } \ k \to +\infty \ ,$$

  by the previous equivalency. This implies that the sum made of the coefficients $(\alpha_k(x))_{k \ge 0}$ is **absolutely convergent** for any x.

- For $k \ge 0$, $l \in \mathbb{Z}$, The coefficient $\beta_l^{(K-k)}$ **is converging for K going to $+\infty$ to the coefficient** $\beta_l^{(+\infty)}$ given by product extended to an infinite number of terms : This is a consequence from the convergence of the infinite product $\prod_{m \ge 0} b_{m,l}$,

$$\text{As } \ b_{m,l} \ \text{ can be rewritten as } \frac{1}{1 - d_{m,l}} \ , \text{ where } \ d_{m,l} = \frac{p}{1-p}(e^{\frac{\ln(1-p)+2i\pi l}{2^m}} - 1) \ ,$$

$$\text{and we have } \ d_{m,l} \sim \frac{1}{2^m} \cdot \frac{p}{1-p}(\ln(1-p) + 2i\pi l) \ \text{ for } \ m \to +\infty \ .$$

  From $\sum_{m \ge 0} |d_m| < +\infty$, we deduce that the product $\prod_{m \ge 0}(1 - d_m)$ is absolutely convergent, and thus converges to a finite non null limit (see [7] Chap.I for more details on the matter).

- For all $m \ge 1$ and for all values of l, $b_{m,l} = \dfrac{p}{1 - (1-p)^{1 - \frac{1}{2^m}}\theta}$ where θ is a complex number with modulus 1, hence we have the comparison : $|b_{m,l}| \le \dfrac{p}{1 - (1-p)^{1 - \frac{1}{2^m}}} = b_{m,0}$, As $b_{m,0} \ge 1$, this implies the following **general product comparison**

$$\text{for any } M < M' \ \left| \prod_{m=M}^{M'} b_{m,l} \right| \le \prod_{m=M}^{M'} b_{m,0} \le \prod_{m \ge 1} b_{m,0} = P \tag{24}$$

  As a consequence, we have the following **uniform bound** :

$$|\beta_l^{(K-k)}| \le P\frac{1}{|l|^2} \ \text{ for } l \in \mathbb{Z} \text{ and any value of } K - k, \text{ this implies } \ |\beta_l^{(+\infty)}| \le P\frac{1}{|l|^2}. \tag{25}$$

5.3.2 Convergence of the CDF

From the uniform bound found in the previous section (Equation 5.3.1), we can deduce that the function $\phi^{(+\infty)}$ given as the Fourier series of $(\beta_l^{(+\infty)})_{l\in\mathbb{Z}}$ is well defined. For $k \geq 0$, this is also the pointwise limit of the function $\phi^{(K-k)}$ when K goes to infinity, as it now proved :

In fact, for any $\varepsilon > 0$ one can find L large enough to have the following comparisons :

$$\sum_{|l|\geq L} |\beta_l^{(K-k)}| < \varepsilon \text{ for any } K-k \text{ and } \sum_{|l|\geq L} |\beta_l^{(+\infty)}| < \varepsilon$$

then only terms with $|l| < L$ need to be considered in the definition of $\phi^{(K-k)}$. They are in finite number, and all of them pointwise converge to the equivalent term for $\phi^{(+\infty)}$.

We can also deduce from the same comparison Equation 5.3.1 that $\phi^{(K-k)}(x)$ can be bounded by a constant for any value of x and any value of $K-k$, possibly infinite. Thus, as the series of the coefficients $(\alpha_k(x))_{k\geq 0}$ is absolutely convergent, the rest of the series

$$\sum_{k=K'}^{K} \alpha_k(x)\phi^{(K-k)}(2^k x)$$

can be arbitrary small for any value of K, possibly infinite, if K' is chosen large enough.

Is it now possible to conclude : for any $\varepsilon > 0$, one can choose K' large enough, for the rest of the sums defining $\mu_K(x)$ and $\mu(x)$ - sums of the terms for $k \geq K'$ - to be smaller than ε.

for any $k = 0..K'$, the term $\alpha_k(x)\phi^{(K-k)}(2^k x)$ is converging to $\alpha_k(x)\phi^{(+\infty)}(2^k x)$ when K is going to infinity.There is a finite number of these terms, their sums are thus converging to the first K' terms of the definition of $\mu(x)$, which proves the result.

6 Numerical Aspects

As the formula given in the last section describes the quantity to be estimated by means of two infinite sums and an infinite product, we are describing in this section a numerical algorithm to compute numerically the formula with an arbitrary precision.

6.1 Proposed truncations for the algorithm

Knowing the value of p and q, for a choice of the parameter x, we want to compute the value of $P(X(\infty) \geq q + qx)$ with an error smaller than ϵ.

The exact value is given by Equation 16, but in practice one needs to estimate how many terms should be evaluated. We propose in this section an algorithm to compute values of K, L and $M^{(k,l)}$ such that for a chosen $\varepsilon > 0$ we have :

$$P(X(\infty) \geq q + qx) = \sum_{k=0}^{K} \alpha_k(x) \left(\sum_{|l|\leq L} \frac{1}{(\ln(1-p) + 2i\pi l)^2} \left(\prod_{m=1}^{M^{(k,l)}} b_{m,l} \right) e^{2i\pi l 2^k x} \right) \pm \varepsilon \qquad (26)$$

Brief description of the algorithm

- In STEP 1, we compute numerically a value P', which only depends on p and plays the same role than P as found in Equation 24 : it bounds the modulus of products of the shape $|\prod_{m=M}^{M'} b_{m,l}|$ (where M' might be $+\infty$).

- In STEP 2, values of the sequence $(\alpha_k(x))_{k=0,1,...}$ are being computed. At some stage, the process is stopped with a criteria that we give, as the rest of the series in Equation 16 will be smaller than $\varepsilon/3$. Only terms $k = 0..K$ will then have to be considered.

- STEP 3 allows us to stop the infinite sum for the index l in all of these $(K+1)$ terms, such that each of them has a precision $\varepsilon/(3(K+1))$.

- We are now left with $K(2L+1)$ infinite product $\prod_{m\geq 1} b_{m,l}$ to be computed. In STEP 4 We precise, depending on the index l of the term, the convergence of this infinite product, and conclude by giving the appropriate truncation, to have each term with a precision $\varepsilon/(3(K+1)(2L+1))$.

- The value can now be computed in a finite number of steps and the sums of all possible errors is less than ε.

Details and Proof are given in the next paragraph, with all the mathematical arguments. All notation and criteria are given again for each step in a conclusion. These conclusions contain all the necessary results to compute the entire algorithm.

6.2   Details of the algorithm

6.2.1   STEP 1 : Bounding the value of the products

From Equation 24 we deduce for any $M < M'$ the following bound :

$$\left| \prod_{m=M}^{M'} b_{m,l} \right| \leq \prod_{m=M}^{M'} b_{m,0} \leq \prod_{m\geq 1} b_{m,0} = \prod_{m\geq 1} \frac{1}{1-d_{m,0}} \quad \text{with} \quad d_{m,0} = \frac{1-p}{p}\left(e^{\frac{\ln(\frac{1}{1-p})}{2^m}} - 1\right) > 0.$$

Let M_0 be chosen large enough to have $\frac{\ln(\frac{1}{1-p})}{2^{M_0}}$ smaller than 1, then we have for any $m \geq M_0$:

$$d_{m,0} \leq \frac{1-p}{p}e\ln(\frac{1}{1-p})\frac{1}{2^m} \quad \text{as} \quad e^x \leq 1 + ex \quad \text{for } 0 \leq x \leq 1,$$

$$\text{such that} \quad \sum_{m>M} d_{m,0} \leq \frac{1-p}{p}e\ln(\frac{1}{1-p})\frac{1}{2^{M_{P'}}} \quad \text{for any } M+1 \geq M_0.$$

If we choose $M_{P'}$ is chosen large enough to have this last right terms to be smaller than $1/2$ we then have,

$$\text{for any } M \geq M_{P'} + 1 :^2 \quad (1 - d_{M_{P'}+1,0})\ldots(1 - d_{M,0}) \geq 1 - \sum_{m=M_{P'}+1}^{M} d_{m,0} \geq \frac{1}{2},$$

$$\text{hence by letting } M \text{ tend to infinity we deduce} \quad \prod_{m>M_{P'}} b_{m,0} \leq 2.$$

Conclusion of the STEP 1 : Let M_0 be chosen to have $2^{M_0} \geq \ln(\frac{1}{1-p})$.

For $M_{P'} + 1 \geq M_0$ taken large enough to have $2^{M_{P'}} \geq 2\frac{1-p}{p}e\left(\ln(\frac{1}{1-p})\right)$,

$$\text{we have :} \quad \left| \prod_{m=M}^{M'} b_{m,l} \right| \leq 2 \prod_{m=1}^{M_{P'}} b_{m,0} = P' \text{ for all } l \text{ and } M \leq M' \text{ (}M' \text{ might be } +\infty\text{)}.$$

---

[2]The first comparison is true by recurrence as all $d_{m,0}$ are positive real numbers.

6.2.2 STEP 2 : Choosing the right truncation for index k

Let's first deduce the following bound : for all $x \geq 0$,

$$\left| \phi^{(+\infty)}(x) \right| = \left| \sum_{l \in \mathbb{Z}} \left(\prod_{m \geq 1} b_{m,l} \right) \frac{1}{(\ln(1-p) + 2i\pi l)^2} e^{2i\pi l x} \right| \leq \sum_{l \in \mathbb{Z}} \left| \prod_{m \geq 1} b_{m,l} \right| \frac{1}{\ln^2(\frac{1}{1-p}) + 4\pi^2 |l|^2},$$

hence $\quad \left| \phi^{(+\infty)}(x) \right| \leq P' \left(\frac{1}{\ln^2(\frac{1}{1-p})} + \sum_{l \in \mathbb{Z}, l > 0} \frac{1}{|l|^2} + \sum_{l \in \mathbb{Z}, l < 0} \frac{1}{|l|^2} \right) = P' \left(\frac{1}{\ln^2(\frac{1}{1-p})} + \frac{\pi^2}{6} + \frac{\pi^2}{6} \right) = P''.$

Consequently all we require here is to choose a proper K to have the rest of the series $\sum_{k > K} |\alpha_k(x)|$ smaller than $(\varepsilon/3)/P''$.

$$\text{We start from} \quad |\alpha_k(x)| = \left| \frac{p(1-p)^{2^k x}}{2^k} \right| \left| \left((\frac{1}{1-p})^{2^k} - 1 \right) a_k \right| |a_{k-1} \ldots a_1| \,,$$

$$a_j \text{ being given by } \frac{1}{1 - \frac{1-p}{p}((\frac{1}{1-p})^{2^j} - 1)} = \frac{1}{1 - \frac{1-p}{p} U_j} \text{ with } U_j = (\frac{1}{1-p})^{2^j} - 1.$$

Note that $(U_j)_{j=1\ldots}$ are positive, they increase with j up to $+\infty$.

Let's consider K_0 such that $\frac{1-p}{p} U_{K_0} \geq 1$. For any $j \geq K_0$, the modulus of a_j is less than 1, consequently the last terms appearing in $|\alpha_k(x)|$ is decreasing with k for $k \geq K_0 - 2$.

Also for $k \geq K_0$, we have that $\left| \left(\frac{1}{1-p} \right)^{2^k} - 1 \right) a_k \right| = \frac{U_k}{\frac{1-p}{p} U_k - 1}$ is decreasing in k.

From the above results, we deduce for $k' \geq k \geq K_0$: $|\alpha_{k'}(x)| \leq \frac{1}{2^{k'-k}} |\alpha_k(x)| \,,$

$$\text{hence} \quad \sum_{k > K} |\alpha_k(x)| \leq 2|\alpha_{K+1}(x)| \text{ for any } K + 1 \geq K_0.$$

Conclusion of the STEP 2 : Let's K_0 be chosen such that $(\frac{1}{1-p})^{2^{K_0}} \geq \frac{p}{1-p} + 1$. In Equation 16, the infinite series with index k can be replaced by its terms $k = 0..K$ - as the rest is smaller than $\varepsilon/3$ - where $K + 1 \geq K_0$ has been chosen large enough to have :

$$|\alpha_{K+1}(x)| \leq \frac{1}{2} \frac{\varepsilon/3}{P''} \text{ where } P'' = P' \left(\frac{1}{\ln^2(1-p)} + \frac{\pi^2}{3} \right).$$

6.2.3 STEP 3 : Choosing the right truncation for index l

Let's denote by $K + 1$ the number of terms that is computed in the series of index k (this is known after computing the STEP 2). We want to be able to estimate all of this terms with an error smaller than $(\varepsilon/3)/(K + 1)$.

From the bound given in STEP 1 on the infinite product we can estimate.

$$\sum_{|l| > L} \left| \beta_l^{+\infty} e^{i2\pi l 2^k x} \right| \leq P' \sum_{|l| > L} \frac{1}{\ln^2(\frac{1}{1-p}) + 4\pi^2 l^2} \leq \frac{P'}{2\pi^2} \sum_{l > L} \frac{1}{l^2} = \frac{P'}{2\pi^2} \left(\frac{\pi^2}{6} - \sum_{l=1}^{L} \frac{1}{l^2} \right).$$

Conclusion of the STEP 3 : When computing the series $\sum_l \beta_l^{+\infty} e^{i2\pi l 2^k x}$, one can consider only terms for $|l| \leq L$ where L is chosen such that

$$\sum_{l=1}^{L} \frac{1}{l^2} \geq \frac{\pi^2}{6} - \frac{2\pi^2}{P'} \frac{\varepsilon/3}{K+1}$$

6.2.4   STEP 4 : **Computing the infinite product**

After STEP 1,2 and 3, we need to estimate $(K+1) \times (2L+1)$ terms containing an infinite product $\prod_{m \geq 1} b_{m,l}$. We describe here how to compute the term with index k and index l in the double sum :

Each of this product are, as it was shown in the proof, absolutely convergent :

$$b_{m,l} = \frac{1}{1 - d_{m,l}} \quad \text{with} \quad d_{m,l} = \frac{1-p}{p}(e^{\frac{\ln(\frac{1}{1-p})+2i\pi l}{2^m}} - 1) \quad \text{such that} \quad \sum_{m \geq 1} |d_{m,l}| < +\infty.$$

It suffices to find M large enough to have the following smaller than $(\varepsilon/3)/(K+1)(2L+1)$:

$$\left| \alpha_k(x) \frac{1}{(\ln(1-p)+2i\pi l)^2} e^{2i\pi l 2^k x} \right| \left| \prod_{m=1}^{M} b_{m,l} - \prod_{m \geq 1} b_{m,l} \right| = \frac{|\alpha_k(x)|}{(\ln(\frac{1}{1-p}))^2 + 4\pi^2 l^2} \left| \prod_{m=1}^{M} b_{m,l} - \prod_{m \geq 1} b_{m,l} \right|$$

We use the Abelian method (see [7] Chap.I) to have for any $M' \geq M$:

$$\left| \prod_{m=1}^{M} b_{m,l} - \prod_{m=1}^{M'} b_{m,l} \right| = \left| \prod_{m=1}^{M} b_{m,l} - \prod_{m=1}^{M+1} b_{m,l} + \prod_{m=1}^{M+1} b_{m,l} - \prod_{m=1}^{M+2} b_{m,l} + \ldots + \prod_{m=1}^{M'-1} b_{m,l} - \prod_{m=1}^{M'} b_{m,l} \right|,$$

$$\left| \prod_{m=1}^{M} b_{m,l} - \prod_{m=1}^{M'} b_{m,l} \right| \leq \sum_{j=M+1}^{M'} \left| \prod_{m=1}^{j-1} b_{m,l} - \prod_{m=1}^{j} b_{m,l} \right| = \sum_{j=M+1}^{M'} \left| \prod_{m=1}^{j} b_{m,l} \right| |(1 - d_{j,l}) - 1|$$

hence $\left| \prod_{m=1}^{M} b_{m,l} - \prod_{m=1}^{M'} b_{m,l} \right| \leq P' \sum_{j=M+1}^{M'} |d_{j,l}|$ and $\left| \prod_{m=1}^{M} b_{m,l} - \prod_{m \geq 1} b_{m,l} \right| \leq P' \sum_{j>M} |d_{j,l}|$.

The convergence of the product can then be directly derived from the convergence of the series made of modulus of coefficients $(d_{m,l})_{m \geq 1}$.

For any $z \in \mathbb{C}$ such that $|z| \leq 1$ the following comparison holds :

$$|e^z - 1| \leq |z| + \frac{|z|^2}{2!} + \frac{|z|^3}{3!} + \ldots \leq |z| + \frac{|z|}{2!} + \frac{|z|}{3!} + \ldots \leq |z|(1 + \frac{1}{2!} + \frac{1}{3!} + \ldots) \leq e|z|.$$

Such that for $M_0^{(l)}$ large enough to satisfy $\dfrac{|\ln(\frac{1}{1-p}) + 2i\pi l|}{2^{M_0^{(l)}}} \leq 1$,

we have, for $m \geq M_0^{(l)}$, the comparison : $|d_{m,l}| \leq \dfrac{1-p}{p} e |\ln(\frac{1}{1-p}) + 2i\pi l| \dfrac{1}{2^m}$

and finally, for $M + 1 \geq M_0^{(l)}$: $\sum_{m>M} |d_{m,l}| \leq \dfrac{1-p}{p} e |\ln(\frac{1}{1-p}) + 2i\pi l| \dfrac{1}{2^M}$.

Conclusion of the STEP 4 : Let $M_0^{(l)}$ be chosen to have $2^{M_0^{(l)}} \geq |\ln(\frac{1}{1-p}) + 2i\pi l|$. To compute the infinite product for index (k, l) of the double sum with the required precision, one can consider the $M^{(k,l)}$ first terms, with $M^{(k,l)} + 1 \geq M_0^{(l)}$ and $M^{(k,l)}$ being chosen large enough to have :

$$2^{M^{(k,l)}} \geq \frac{|\alpha_k(x)|}{|\ln(\frac{1}{1-p}) + 2i\pi l|} P' \frac{1-p}{p} e \frac{(K+1)(2L+1)}{\varepsilon/3}.$$

7 Exploitation of the Results

In Section 5 we have developed a closed form formula, i.e. Equation 16, to calculate the TDF of the throughput and in Section 2 we have introduced a method the estimate the two central parameters in that formula : i.e. the synchronization rate p and the load ρ at congestion epochs. The synchronization rate p was approximated based on the M/M/1/B-queue and the load ρ at congestion epochs was determined by considering the time evolution of the buffer occupancy. In this section we describe how this methodology can be used to study the performance of long-lived TCP-controlled connections and dimension networks with such traffic.

In all previous sections the link capacity and the buffer size were expressed in packets per second and packets respectively, while in this section we use bits per second and bytes as units respectively. To relate the latter to the former we need to choose a typical packet size for the IP packets transporting the TCP data. We assume a packet size of 536 bytes.

7.1 Buffer size and synchronization

In this section we consider a bottleneck router of capacity C shared by N users. Each (persistent TCP) user sees the same RTT R of 100 ms. The number of users is always such that the available capacity per user is kept to $\frac{C}{N} = 1$ Mb/s. Three cases are considered : a bottleneck router with a link capacity C of 34 Mb/s (i.e. an E3 link), 100 Mb/s (i.e. a 100BaseT Ethernet link) and 600 Mb/s (i.e. an STM-4 link) which is shared by 34, 100 and 600 (persistent TCP) users respectively. We study first the effect of the buffer size B of this router on the synchronization rate p and its consequence on the instantaneous throughput seen by one user in average as well as its distribution.

As explained in Section 2 our analysis is only valid for $B \ll RC$. We then take the largest considered value of the buffer size to be $B = 0.1RC$, which corresponds to 0.0425 Mbyte, 0.125 Mbyte and 0.75 Mbyte for $C = 34$ Mb/s, 100 Mb/s and 600 Mb/s respectively. The value of the buffer size B_{thr} that separates the empty buffer case from the never empty buffer case is 0.55 Mbyte, 1.62 Mbyte and 9.71 Mbyte for the 34 Mb/s, 100 Mb/s and 600 Mb/s case respectively. Consequently, the empty buffer case always applies to values of B considered in this study.

We define the relative average throughput as the ratio of the average throughput seen by a user at a random time instant in steady state (see Equation 15) over the user's maximum fair share $\frac{C}{N}$. It is given by

$$\frac{\mathbb{E}[X(\infty)]}{\frac{C}{N}} = \rho \left(1 - \frac{p}{4} \right) \, , \tag{27}$$

and as one can see by examining Equation 9 which estimate ρ as a minimum, it can never exceed 1. Moreover, this value of 1 can only occur, if in Equation 9 the right term under the minimum operator is the smallest, corresponding to the case where the buffer never empties between congestion epochs. This conforms with the intuitive argument that any system that apportions resources amongst users, cannot offer each user its maximum fair share if it "wastes" transmission opportunities, i.e. if it allows its buffer to empty. Notice also that the relative average throughput never drops below 0.75 as $\rho \geq 1$ and $0 \leq p \leq 1$. As the empty buffer case always applies here, the load is increasing with the square-root of the buffer size B as :

$$\rho = 1 + \sqrt{2\frac{N}{RC}}\sqrt{\frac{B}{RC}} \, . \tag{28}$$

Figure 10 illustrates how the synchronization rate p and the relative average throughput seen by one user changes as the buffer B increases. The buffer size is expressed relative to RC, where $\frac{B}{RC}$ is always kept smaller than 0.1. In particular, it shows that the average throughput does not always increase as the buffer size B gets bigger, i.e. there may be some (counterintuitive) situations where increasing the buffer size B actually leads to a decrease in average throughput. This can be understood in our model by studying the variation of the synchronization rate with B.
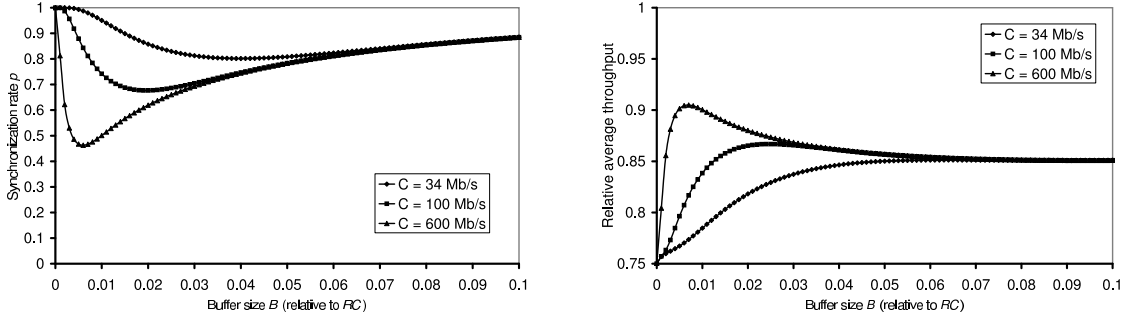
Figure 10: Synchronization rate p (left) and relative average throughput (right) seen by one TCP user as a function of the buffer size B relative to RC for several values of the capacity C, as estimated by the method developed in this paper.

For a very small buffer size the load ρ (at congestion epochs) is barely larger than 1. Under this condition the synchronization rate is given by

$$ p = \frac{1 - \exp\left(-\frac{RC}{N(B+1)}\right)}{1 - \exp\left(-RC\frac{1}{B+1}\right)} . \tag{29} $$

Hence, for B small enough to have $\frac{B}{RC} \ll \frac{1}{N}$, p is very close to 1, and decreases initially with B. Equation 10 shows that this decrease for a certain value of B gets compensated by an increase in ρ, which also increases as B increases. A minimal value of p is then attained for some buffer size - notice that the value for $\frac{B}{RC}$ where the local minimum occurs and the minimal value of p both decrease as the link capacity C increases - and the value of p is then increasing under the influence of the second effect. If this increase of p (just after its local minimum) is sufficient to over-compensate the increase of ρ in the product defining the relative average throughput, and hence, makes the latter one decrease, a local maximum appears in the relative average throughput curve by a value of the buffer. This local maximum is sharper for a larger capacity. For larger buffer sizes - such values are not shown in Figure 10 as the necessary condition $B \ll RC$ is not valid anymore for the cases studied here - the increase in load ρ dominates and results in an overall increase for the average throughput. Ultimately if the buffer size is bigger than than B_{thr}, the never empty buffer case applies, implying a relative average throughput equal to 1.

These results were checked via a NS simulation (a packet-based simulator of TCP networks, available at http://www.isi.edu/nsnam/ns/). The set-up was chosen to be as close as possible to the one above. N users share a router of capacity C - where $\frac{C}{N}$ is kept equal to 1 Mb/s - which introduces a propagation delay of 10 ms. The N users each access this router through a separate link of capacity 1.5 Mb/s where the propagation delay is 30 ms ; after the common router each user is fanned out on a separate link to its destination on a link of capacity 1.5 Mb/s where the propagation delay is 10 ms. Finally, the propagation delay for the acknowledgment packets is 50 ms. Notice that (apart from queueing in the bottleneck router) the RTT is 100 ms. It is expected that the impact of all links except the shared link, which plays the role of the bottleneck, is negligible.

Results from these simulations are shown in Figure 11. There is no curve available for $N = 600$ users as it turned out to require too much computer memory. The overall trends of the average throughput is similar to what our model predicts ; the average relative throughput is roughly increasing with B, but it admits some local maximum and its evolution stays pretty flat after this local maximum except for some small variations. We can also observe some differences with our model. First, a relative average throughput smaller than 75% can be observed when the buffer is very small, as B can then play the role of a $W_{\max}$ (maximal number of packets processing
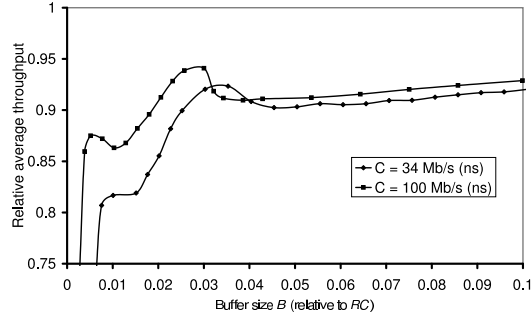
Figure 11: Relative average throughput seen by one user as a function of the buffer size B relative to RC as obtained by NS simulations.
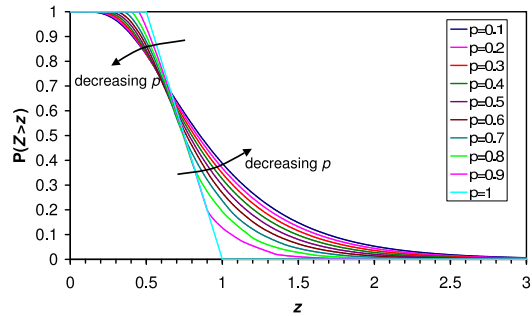

Figure 12: TDF of the variable Z with the synchronization rate p as parameter.

through the network for a connection) if the traffic is bursty ; this packet level effect cannot be captured by our fluid model. Second, we observe that for large values of the buffer size our model underestimates the average throughput (or equivalently overestimate the synchronization rate p). This might be a consequence of the variation of the RTT due to the queuing delay, as for these cases the buffer is sufficiently large to have such an impact, which leads to a smaller value of synchronization between flows than the one predicted by our calculation on Markovian queues.

Next, we consider the TDF of the instantaneous throughput. To study the impact of ρ and p separately we define the relative instantaneous throughput by rewriting Equation 14 as

$$\frac{X(\infty)}{\frac{C}{N}} = \rho Z \quad \text{with} \quad Z = \frac{p}{2} \left(1 + U + \sum_{k \geq 0} \frac{A_k}{2^k} \right) , \tag{30}$$

Notice that the synchronization rate p solely determines the TDF of the random variable Z. For $p = 1$, Z is uniformly distributed between 0.5 and 1. Figure 12 shows the TDF of the variable Z for different values of p. As p decreases from 1, the TDF becomes less steep and a broader range of values can be taken from $\frac{p}{2}$ to $+\infty$. The 0.95-quantile (i.e. that value z for which $P(Z > z) = 0.95$) thus decreases as p increases and similarly, the 0.05-quantile increases with p.

Figure 13 shows the TDF of the throughput seen by one user. In each case four buffer sizes were chosen such that $\frac{B}{RC} = 0.0008$, 0.0024, 0.008 and 0.08. We observe in two cases, corresponding to the lowest buffer size for capacity 34 Mb/s and 100 Mb/s, a law very close to the uniform distribution on the interval $[0.5\text{Mb/s}; 1\text{Mb/s}]$, as the synchronization rate p is then close to 1. The average throughput is in these cases only a little bit larger than 0.75 Mb/s (i.e. only 75% of the
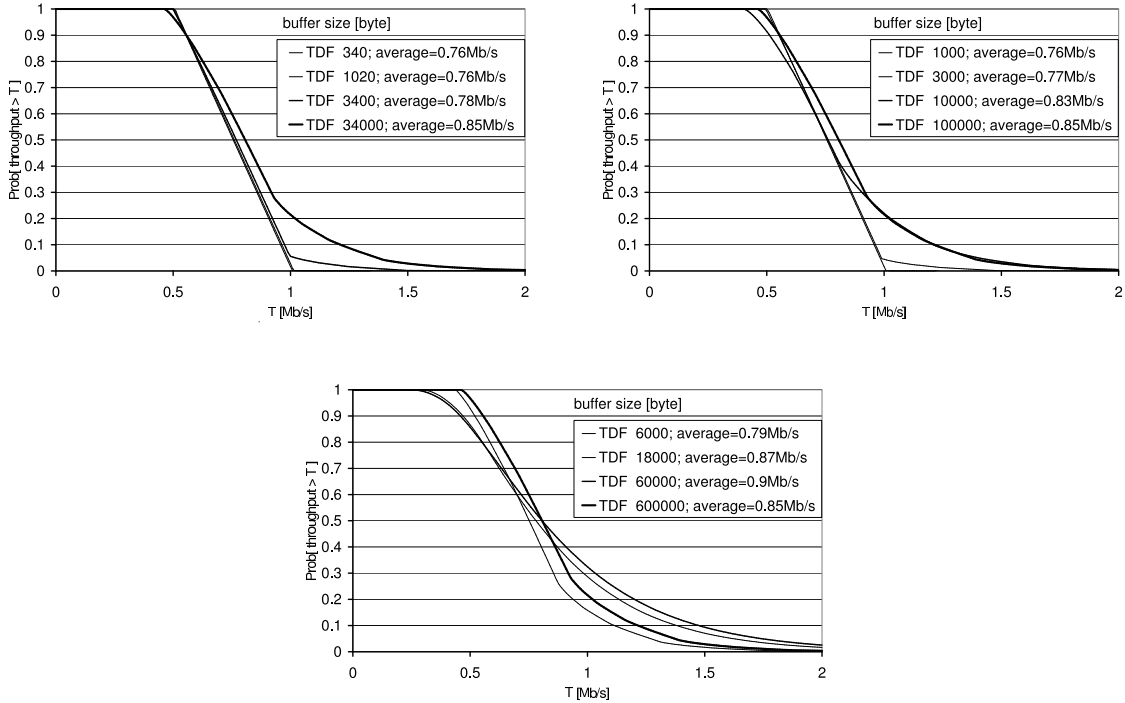
Figure 13: Average and TDF of the throughput seen by one TCP user for several buffer sizes B and for $C = 34$ Mb/s (top left), 100 Mb/s (top right) and 600 Mb/s (bottom).

naively expected value of 1 Mb/s). Remark that this is not the case for the smallest buffer size with capacity 600 Mb/s, as the synchronization rate p has already decreased from 1.

In the case of 600 Mb/s we observe the decrease in average throughput as B increases for the two largest values of the buffer size. This was already explained when discussing Figure 10.

As a conclusion, long-lived TCP connections with same RTT sharing a bottleneck router do not get on average the available resources per user. Its expected throughput depends on the buffer size through the synchronization rate and the load.

7.2 Dimensioning for statistical performance guarantee

In the previous section we calculated the average (and the TDF of the) throughput seen by a user in a given scenario (i.e. for a given C, N, B and R). Here the object of our study is quite different as we compute for some values of the capacity what is the maximum number of connections that can share the link, if for each of them one needs to satisfy a statistical performance requirement. The link has still a capacity of 34 Mb/s, 100 Mb/s and 600 Mb/s, and the buffer is again a parameter that we vary while keeping $\frac{B}{RC} \leq 0.1$. The object of our study is the maximal value of N to have a throughput always greater than x=100 kb/s, except for at most 5% of the time (i.e. $P(X(\infty) \geq x) \geq 95\%$). Therefore, we need to numerically invert the closed form formula of Equation 16.

As the threshold for the throughput considered here (100 kb/s) is one order of magnitude smaller than than throughput considered in the previous section (around 1 Mb/s), we can intuitively expect that the number of users we consider now is higher. This is indeed the case, and there is for every capacity level between three and five times more connections in this second study. Consequently, the threshold buffer B_{thr} (which decrease when N is higher) may fit in the range of values that we consider for B (to keep $\frac{B}{RC}$ smaller than 0.1).
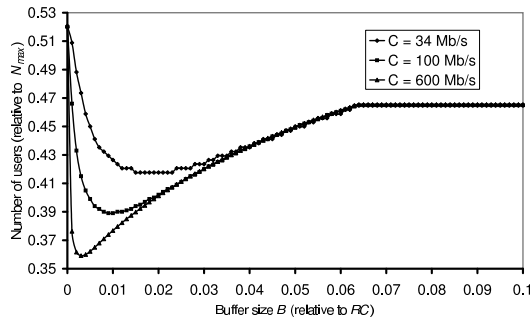
Figure 14: The maximum number of users for which a throughput of 100 kb/s can be guaranteed for 95 % of the time as a function of the buffer size B in the bottleneck router, obtained with the closed formula.

Figure 14 shows the results obtained with the closed formula. For each value of B, we have increased N, while we evaluate p, and $P(X(\infty) \geq x)$ using the Formula 16 ; the process was stop when this probability falls under 0.95. The maximal value of N is given relative to $N_{\max}$ which is the value we can expect if the capacity was equally shared perfectly at any time. This is equal to $\frac{C}{x}$ and for a shared capacity of 34 Mb/s, 100 Mb/s and 600Mb/s, $N_{\max}$ is 340, 1000 and 6000 respectively.

If the buffer size B is very small the synchronization rate p is very close to 1. Hence, the TDF of the throughput is practically uniform between $\frac{1}{2}\frac{C}{N}$ and $\frac{C}{N}$. Hence, the 0.95-quantile hardly exceeds $\frac{1}{2}\frac{C}{N}$, and the maximal number of users is slightly higher than half of $N_{\max}$. Remark however that the NS simulations of the previous section indicated that in this case the packet phenomena are so important that our fluid flow model cannot capture this behavior very well.

For a slightly larger buffer size B the load ρ increases with B while the synchronization rate p decreases. This had a beneficial effect on the (relative) average throughput, as was observed in the previous section. Strangely enough this beneficial effect does not apply to the 0.95-quantile. This was already observed for the relative instantaneous throughput Z on Figure 12. So, in this interval for B where the average throughput increases, the value of the 0.95-quantile actually decreases (and as a consequence less users can be supported if such a guarantee of 100 kb/s needs to be given 95% of the time).

For still larger values of the buffer size B, i.e. for a B larger than the value where the minimal synchronization rate p occurs, but still smaller than the value where the never empty buffer case starts, the synchronization rate increases again and more importantly the load ρ still increases, resulting in a steep increase of the 0.95-quantile of the throughput (and hence, also the number of users that are guaranteed 100 kb/s 95% of the time).

Finally, the buffer is so large that the never empty buffer case applies. Remember that as number of users considered in the link is larger now than it was in the previous section, this happens for a value where our analysis still applies. The average throughput seen by each user is exactly $\frac{C}{N}$. The load ρ is given by $\frac{4}{4-p}$ and the synchronization rate p remains more or less constant. This value of the synchronization rate determines the 0.95-quantile of the relative instantaneous throughput Z and in turn the number of users N that can be supported. As can be seen on Figure 14 for these values of B the number of users that can be supported remains more or less constant.

8 Conclusion

A new framework has been proposed to estimate the throughput obtained by a long-lived TCP connections in a simple networking case : homogeneous traffic sharing one bottleneck link. The model is able to capture interaction with other connections via the synchronization rate, a unique

parameter that together with the load at the congestion epoch completely describes the distribution of the instantaneous throughput. We started by considering the system in discrete time, i.e. where we only consider the throughput given after a congestion epoch, and noticed the tail distribution function of the throughput shows some self similar nature. We then developed a closed form formula for the tail distribution function of the throughput in continuous time, i.e. taken at any point in time in steady state.

A method was proposed to estimate both the synchronization rate and the load during congestion epochs. Two cases were distinguished : the case where the buffer empties between two successive congestion epochs and the case where it does not.

This framework is then used to study the effect of the buffer on synchronization between flows, in particular studying the counterintuitive behavior of TCP-controlled connections, that can also be observed in NS simulations. It is then applied to dimension a network with long-lived TCP controlled connections, for a probabilistic performance requirement, where a throughput larger than a given value needs to be guaranteed a certain percentage of the time. This closed form formula for the TDF, allows us to dimension the routers in an IP network just like the Erlang B formula allows to dimension the switches in a telephony network.

The framework proposed here can be extended in several ways :

- Because in our case the RTT seen by each connection is the same, the synchronization rate for each connection also is. In the heterogeneous case where all TCP connections do not see the same RTT, the throughput seen by a connection will depend on the RTT. The Fluid Flow AIMD model for this case can be found described in [1]. In future research we will extend our methodology to also include this scenario.

- The Fluid Flow AIMD model for a general network topology has been recently studied and is presented in [3].

- As our analysis only applies for a constant RTT, we can only study buffers that are not so large that they do not result in a variable RTT. We are studying ways to extend our analytical framework to also take this effect into account.

- Finally also developing an analytical framework to describe the impact of non persistent connections - which may be responsible for some long range dependence now observed in TCP controlled network (see [6]) - on the throughput TCP offers to a connection is left as future work.

9   Acknowledgment

A   Proof of Some Propositions of Section 3

A.1   From Equation 8 to 9

Proposition 1 *For $\rho_0 \in \mathbb{R}$, we consider the solution of the equation :*

$$\rho = g\left(\max\left(\rho_0, \rho\right)\right) \tag{31}$$

where $g : [1; +\infty[\mapsto \mathbb{R}$ is a decreasing function admitting a fixed point $\tilde{\rho}$ (necessarily unique).

- *Then the unique solution of the equation is given by the minimum :*

$$\rho = \min\left(g(\rho_0), \tilde{\rho}\right) .$$

Proof :

- if $\rho_0 < \tilde{\rho}$, we have $g(\rho_0) > \tilde{\rho} > \rho_0$, hence $g(\rho_0)$ is not a solution whereas $\tilde{\rho}$ is.

- if $\rho_0 > \tilde{\rho}$, we have $g(\rho_0) < \tilde{\rho} < \rho_0$, hence $\tilde{\rho}$ cannot be solution, while $g(\rho_0)$ is a solution.

Remark Note that this expression of the solution of Equation 31 holds with the the more general condition on g (see Figure 15) :

$$
Condition\ 1 \quad \left\{ \begin{array}{l} g \text{ has a unique fixed point } \tilde{\rho}, \\ \text{for } \rho \leq \tilde{\rho}, \quad \text{we have } g(\rho) \geq \tilde{\rho} \\ \text{and for } \rho \geq \tilde{\rho}, \quad \text{we have } g(\rho) \leq \tilde{\rho} \ . \end{array} \right.
$$

A.1.1 Application to the estimation of ρ

Equation 8, i.e.

$$
\rho - 1 = \min \left(\chi \ , \ 1 - \rho(1 - \frac{p}{2}) \right) \ ,
$$

can be rewritten under the form of proposition 1 as

$$
\rho = g \left(\max \left(\frac{1 - \chi}{1 - \frac{p}{2}} \ , \ \rho \right) \right) \ ,
$$

with $g : \rho \mapsto 2 - \rho(1 - \frac{p}{2})$, $\rho_0 = \frac{1 - \chi}{1 - \frac{p}{2}}$. All conditions necessary for Proposition 1 to apply are verified and ρ can then be expressed by the minimum :

$$
\rho = \min \left(g(\rho_0), \tilde{\rho} \right) = \min \left(1 + \chi, \frac{1}{1 - \frac{p}{4}} \right) \ .
$$

A.2 General estimation of the synchronization rate

Proposition 2 *For $p_0 \geq 0$, we consider the solution of the equation :*

$$
p = g \left(\min \left(p_0, p \right) \right) \tag{32}
$$

where $g : [0; 1] \mapsto \mathbb{R}$ verifies the conditions (see Figure 15) :

$$
Condition\ 2 \quad \left\{ \begin{array}{l} g \text{ has a unique fixed point } \tilde{p}, \\ \text{for } p \leq \tilde{p}, \quad \text{we have } g(p) \geq \tilde{p} \geq p, \\ \text{and for } p \geq \tilde{p}, \quad \text{we have } g(p) \leq \tilde{p} \leq p \ . \end{array} \right.
$$

- *Then the unique solution of the equation is given by the minimum :*

$$
p = \min \left(g(p_0), \tilde{p} \right) \ .
$$

Proof :

- if $p_0 < \tilde{p}$, we have $\tilde{p} > g(p_0) > p_0$, hence $g(p_0)$ is a solution whereas $\tilde{p}$ is not.

- if $p_0 > \tilde{p}$, we have $\tilde{p} < g(p_0) < p_0$, hence $\tilde{p}$ is a solution, while $g(p_0)$ is not.

Remark : Note that for g admitting a fixed point, the condition $0 \leq g' < 1$ is sufficient to have Condition 2 .
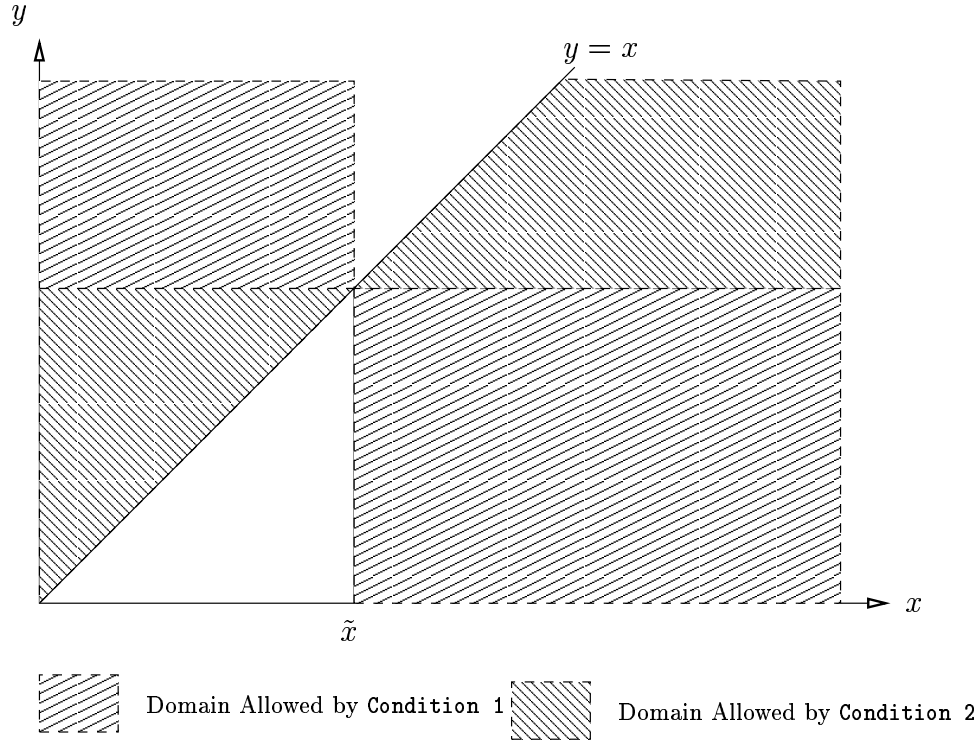
Figure 15: Necessary conditions for proposition 1 and 2

A.2.1 Application to the general estimation of p :

From $p = f(\rho)$ where f is defined by Equation 10, we deduce $p = f\left(\min\left(1 + \chi, \frac{1}{1 - \frac{p}{4}}\right)\right)$

Hence p is the solution of equation given in Proposition 2 for the function g as defined in section 3.2, i.e. $g(p) = f\left(\frac{1}{1 - \frac{p}{4}}\right)$, and $p_0 = \frac{4\chi}{1 + \chi}$.

In all the cases described in section 7 and more generally in all the cases that has been observed, Condition 2 was verified by the function g which depends on R,C,N and B. Note also one consequence from this condition is that the iterative sequence considered in section 3.2 always converges to the fixed point (as for $p < \tilde{p}$ we have $p < g(p) < \tilde{P}$ the sequence is increasing bounded by $\tilde{p}$ which is the only limit possible).

As a consequence, The value of p satisfying Equation 10 is always given by the minimum $p = \min\left(g(p_0), \tilde{p}\right)$.

References

[1] F. Baccelli and D. Hong, *A.I.M.D., Fairness and Fractal Scaling of TCP Traffic*, Proceedings of IEEE Infocom, July 2002, New York.

[2] D. Hong and D. Lebedev, *Many TCP User Asymptotic Analysis of the AIMD Model*, INRIA Research Report number 3971, June 2001.

[3] F. Baccelli and D. Hong, *Interaction of TCP flows as Billiards*, IPAM Workshop on large scale network, March 2002, UCLA (INRIA research report to appear).

[4] F. Baccelli and T. Bonald, *Window flow control in FIFO networks with cross traffic* INRIA Research Report number 3434, May 1998.

[5] J. Padhye, V. Firoiu, D. Towsley and J. Kurose, *Modeling TCP Throughput: A Simple Model and its Empirical Validation* Proceedings of SIGCOMM'98.

[6] M. S. Taqqu, W. Willinger, and R. Sherman, *Proof of a Fundamental Result in Self-Similar Traffic Modeling*, ACM Computer Communication Review 27 (1997).

[7] E.C. Titchmarsh, *The Theory of Functions*, Oxford University Press (1932).

[8] D. Widder, *The Laplace Transform*, Princeton University Press (1946).