GERAD

GROUPE D'ÉTUDES ET DE RECHERCHE
EN ANALYSE DES DÉCISIONS

Les Cahiers du GERAD

G-2014-92
November 2014

CITATION ORIGINALE / ORIGINAL CITATION

H. Dbira, A. Girard, B. Sansò, Calculation of packet jitter for non-poisson traffic, Annals of Telecommunications, 71(5), 223-237, 2016.

GERAD HEC Montréal
3000, ch. de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Les Cahiers du GERAD ISSN: 0711–2440

Calculation of packet jitter for non-Poisson traffic

H. Dbira, A. Girard,
B. Sansò

G–2014–92

November 2014

GERAD HEC Montréal
3000, chemin de la Côte-Sainte-Catherine
Montréal (Québec) Canada H3T 2A7

Tél. : 514 340-6053
Téléc. : 514 340-5665
info@gerad.ca
www.gerad.ca

Calculation of packet jitter for non-Poisson traffic

Hadhami Dbira[a]
André Girard[b]
Brunilde Sansò[a]

[a] *GERAD & Electrical Engineering Department, Polytechnique Montréal, Montréal (Québec) Canada, H3C 3A7*

[b] *GERAD & Polytechnique Montréal & INRS-Énergie Matériaux Télécommunications, Montréal (Québec) Canada, H5A 1K6*

hadhami.dbira@polymtl.ca
andre.girard@gerad.ca
brunilde.sanso@polymtl.ca

November 2014

Les Cahiers du GERAD

G–2014–92

Abstract: The packet delay variation, commonly called delay jitter, is an important quality of service parameter in IP networks especially for real-time applications. In this paper, we propose exact and approximate models to compute the jitter for some non-Poisson FCFS queues with a single flow that are important for recent IP network. We show that the approximate models are sufficiently accurate for design purposes. We also show that these models can be computed sufficiently fast to be usable within some iterative procedure, e.g., for dimensioning a playback buffer or for flow assignment in a network.

Key Words: Jitter, queue, inter-arrival distribution, service time distribution, analytic model.

Acknowledgments: This work was supported by Grant No. 121949-2012 from the Natural Sciences and Engineering Research Council of Canada.

1 Introduction

The Internet is now the main medium for a large number of applications and multimedia services. These applications require a good level of quality of service (QoS) simply to provide an adequate quality of experience to the customer. This is a challenge for network operators who must provide good transmission while dealing with fast-changing technology and increasing traffic growth so that computing QoS parameters quickly and accurately is a very important issue.

Usually, IP network planning and design use standard metrics based on queuing theory like the average delay and packet loss to optimize the network cost and performance [1, 2]. On the other hand, there is little work on computing the packet delay variation, also known as *jitter*, in IP networks. But this measure of QoS is particularly important for real-time services such as video conferencing, VoIP or video streaming and can have a greater impact on the user experience than latency and packet loss.

Standard bodies give different definitions of jitter. The IETF uses the mean absolute value of the packet delay variation [3] while the ITU-T used the variation of delay from some minimum value in the stream [4]. Irrespective of the particular definition used, a basic difficulty is that delay jitter is based on the distribution of packet delay, also called transit time, something that is often difficult to compute or even simply not known.

This is not a problem for M/M/1 queues where the transit time distribution is known. This has been used in a recent study [5] based on the IETF definition where a general definition of jitter is evaluated for an isolated queue and an approximate model is presented for a network of M/M/1 queues. Because the M/M/1 queue is not a good model for IP traffic, we want to extend these results to other queues.

In this paper, we focus on calculating the jitter, as defined by the IETF, in a single FCFS queue with a single flow but with non-Poisson processes. We make use of the transit time distribution whenever it is known or provide approximations when it is not. We also pay attention to the computation times since this is important for use in iterative procedures, say for optimizing the parameters of a queue or in a network design algorithm, or for controlling jitter for real-time applications.

In order to have some insight on jitter estimation in IP networks, we present in Section 2 a brief literature review. The section also justifies our interest in jitter estimation for non Poisson queues. We recall in Section 3 the general formula to compute jitter in a FCFS queue for a single flow of arbitrary statistics. This formula has two weak points. First, it depends on the distribution of the transit time, which is not known except for a limited number of queues. Also, it is in the form of a triple integral, which often requires long computation times and can suffer from poor convergence in some cases. We then propose in Section 4 an exact analytic model for the G/M/1 queue where the transit time distribution is known. We show how it can be evaluated quickly by solving a simple nonlinear equation. We present in Section 5 two important special cases for low and high load with arbitrary traffic, where the jitter does not depend at all on the transit time distribution so that we can derive exact formulas. Next, we use the high and low traffic values to propose approximations for two important classes of queues where the transit time is not known. We show in Section 6 that a linear approximation is a very accurate model for the M/G/1 queue and we propose in Section 7 a piece-wise linear approximation for the G/D/1 queue. In these two cases, the accuracy of the approximations is evaluated by comparing with simulation results. Finally, we present in Section 8 a sample of the computation requirements for various processes.

2 Related work

There has been much work in the last decades on the estimation of the delay jitter of ATM networks. The jitter distribution for a periodic traffic was derived in [6] and has been used to shape traffic by dimensioning a leaky bucket. A complete characterization of the jitter for a constant bit rate traffic is provided by [7, 8]. A similar analysis [9] with periodic background traffic focuses on per-stream jitter. There is also some work on approximations for jitter in DiffServ networks. This is done for slot-based TDM traffic in [10] and for periodic arrivals with constant packet length in [11]. Authors in [12] computed the jitter generating function for a DiffServ queue where the EF traffic is an ON/OFF stream and the BE stream is Poisson.

This work cannot be used directly since ATM networks are quite different from IP networks. The packets are small and of constant length, the sources are generally modelled as periodic and the only measure of QoS is the cell loss probability, which has to be very low, typically less than 10^{-6} so that results can be derived based on the Chernoff bound. This is very different from IP, where packets have different lengths, the arrival processes are far from periodic and the QoS requirements are based on delay, loss, jitter and bandwidth.

Recently, a simple formula has been proposed [5] for computing the jitter both for a single queue and multiple queues in tandem for Poisson packet arrivals and exponential holding times. Several approximations are presented for a single queue for small, large and intermediate arrival rates of a tagged stream.

These models are then used in [13] to evaluate the effect of jitter on network routing. They find the optimal routing of IP packets subject to jitter constraints and evaluate the impact on network performance of taking jitter into account. The authors claimed that even though the Poisson model might not be very accurate in general, it *is* realistic in some kinds of access networks. Also, the accuracy of the *difference* between the two cases, with and without jitter constraints, might be good enough to draw some conclusions.

Still, there is a large body of work [14, 15] showing that packet traffic in local and wide area Internet networks, and especially traffic from real-time applications, is definitely not Poisson. The study of the Internet traffic in [14, 16] has shown that it can be modelled by a self-similar process and found that the inter-arrival distribution is best represented by heavy-tailed distributions. For WLANs, the study of packet inter-arrival time [17] has shown it to be more consistent with a Pareto, Weibull or Lognormal distribution. An analysis for BitTorrent traffic through IPv4 and IPv6 networks is given by [18] and shows that the distribution of inter-arrival time is Weibull in IPv4 and Gamma in IPv6. There are also many cases where the packet length distribution is not exponential. The authors of [19] show that packets in IP backbones have a trimodal distribution with one packet size corresponding to TCP acknowledgment and two Maximum Transmission Units MTU packet with sizes 572 and 1500 bytes. For all these reasons, we want to extend the previous results to non Poisson traffic processes.

3 General formula

First we define the notation we will be using later and then provide a general formula for computing the jitter.

3.1 Notation

Throughout the paper, we denote a random variable by an upper case symbol like X, its mean by $m = E[X]$, its standard deviation by s_t and its variance by $v = s_t^2$. We write the corresponding pdf as $f_X(x; \mu, \sigma, \ldots)$ with the appropriate list of parameters. The cdf is written as $F_X(x; \mu, \sigma, \ldots)$ and the Laplace transform of f_X as $\mathcal{F}_X(s; \mu, \sigma, \ldots)$. We use the simplified form $f_X(x)$ whenever the value of the parameters is clear from the context. We will be using a number of distributions with pdfs denoted as follows:

$LogN(x; \mu_L, \sigma_L)$	Log-normal distribution with location μ_L and scale σ_L
$Nt(x; \mu, \sigma)$	Truncated normal distribution with support $[0, \infty]$ with location μ_n and scale σ_n
$Gm(x; , k, \theta)$	Gamma distribution with scale θ and shape k
$P(x; x_m, \alpha)$	Pareto type I distribution with scale x_m and shape α

More information on these distributions is contained in Appendix A.

3.2 Definition of jitter

In this paper, we adopt the IETF [3] definition of jitter. It is based on the transit delay of successive packets between two measurement points. For a single queue, these are the entry into the buffer and the exit from the server.

First define for packet i

t_i	arrival time,
R_i	Inter-arrival time, $R_i = t_i - t_{i-1}$
λ	Average arrival rate $\lambda = 1/E[R_i]$
r_i	departure time,
W_i	waiting time,
S_i	service time,
μ	average service rate $\mu = 1/E[S_i]$
T_i	transit time, $T_i = W_i + S_i$,
η	transit rate $= 1/E[T_i]$

The algebraic difference of transit time between two consecutive packets is written as

$$G = T_{i+1} - T_i. \tag{1}$$

The end-to-end jitter is defined as the expected absolute value of G

$$J = E\left[|T_{i+1} - T_i|\right]. \tag{2}$$

Note that this is not the mean difference since the two random variables are not independent. We then have

$$\begin{aligned} T_{i+1} - T_i &= W_{i+1} + S_{i+1} - (W_i + S_i) \\ J &= E\left[|T_{i+1} - T_i|\right] \\ &= E\left[|(W_{i+1} - W_i) + (S_{i+1} - S_i)|\right]. \end{aligned} \tag{3}$$

We now recall the exact formula for jitter from [5]

$$W_{i+1} = \begin{cases} 0 & \text{if } r_i \leq t_{i+1}, \\ W_i + S_i - (t_{i+1} - t_i) & \text{if } r_i > t_{i+1} \end{cases} \tag{4}$$

so that we have

$$T_{i+1} - T_i = \max\left(S_{i+1} - T_i, S_{i+1} - R_{i+1}\right). \tag{5}$$

Note that the random variables T_i, S_{i+1} and R_{i+1} are independent and also that their distribution is the same for all values of i. Define the probability density functions

$f_R(y)$	for the inter-arrival times
$f_S(z)$	for the service times
$f_T(x)$	for the transit times.

Based on (5), we can then express J in terms of these three distributions

$$\begin{aligned} J &= E\left[|\, T_{i+1} - T_i|\right] \\ &= \int_0^\infty f_R(y) \left\{ \int_0^\infty f_S(z) \left[\int_0^y |z - x| f_T(x) dx + |z - y| \int_y^\infty f_T(x) dx \right] dz \right\} dy. \end{aligned} \tag{6}$$

There are two difficulties with using (6). A fundamental problem is that in many cases, we don't even know the form of f_T so that it is simply not possible to use the formula directly. Even in cases where f_T is known, or can be approximated, it is always time consuming to compute a triple integral numerically so that we want something simpler when speed is important. For these reasons, the focus of the paper is to find simple and fast jitter models derived from (6) when $f_T(x)$ is known and techniques to compute good approximations when it is not, at least for some important kinds of queues.

4 The G/M/1 queue: Exact value

The simplest case where we can get an exact value for the jitter is that of the G/M/1 queue with an exponential service time with parameter μ. Here we use the fact that [20] the transit time T has an exponential distribution

$$f_T(x) = \eta e^{-\eta x} \tag{7}$$
$$\eta = \mu(1-\tau) \tag{8}$$

where η is the transit rate. τ is the probability of waiting and is given given by the unique root in the interval $(0,1)$ of

$$\tau = \mathcal{F}_R(\mu - \mu\tau) \tag{9}$$

where $\mathcal{F}_R(s)$ is the Laplace transform of the inter-arrival distribution $R(y)$. Based on this, we can in principle compute the jitter using (6). Still, a direct computation of the triple integral will most likely require large computation times so that we now derive some simplified formulas and show that the running time is quite suitable for use in some iterative algorithm.

4.1 Simplified model

We use the fact that both the service and transit times are exponential to simplify (6) since, the last two integrals can be computed explicitly. We get an expression for the jitter that depends only on the form of R

$$\begin{aligned}
J &= \int_0^\infty f_R(y)dy \left\{ \int_0^\infty (\mu e^{-\mu s})ds \left[\int_0^y |s-x|(\eta e^{-\eta x})dx + |s-y| \int_y^\infty (\eta e^{-\eta x})dx \right] \right\} \\
&= \int_0^\infty f_R(y)dy \left\{ \frac{2 f_T(y) f_S(y)}{\eta\mu(\eta+\mu)} - \frac{f_T(y)}{\eta^2} + \frac{(\eta+\mu)}{\eta\mu} - \frac{2}{(\eta+\mu)} \right\} \\
&= \frac{(\eta^2+\mu^2)}{\eta\mu(\eta+\mu)} + \int_0^\infty f_R(y) \left\{ \frac{2 f_T(y) f_S(y)}{\eta\mu(\eta+\mu)} - \frac{f_T(y)}{\eta^2} \right\} dy \\
&= \frac{(\eta^2+\mu^2)}{\eta\mu(\eta+\mu)} + \frac{2}{(\eta+\mu)} \int_0^\infty f_R(y) e^{-(\eta+\mu)y} dy - \frac{1}{\eta} \int_0^\infty f_R(y) e^{-\eta y} dy \\
&= \frac{(\eta^2+\mu^2)}{\eta\mu(\eta+\mu)} + \frac{2}{(\eta+\mu)} g(\eta+\mu) - \frac{1}{\eta} g(\eta)
\end{aligned} \tag{10}$$

$$g(c) = \int_0^\infty f_R(y) e^{-cy} dy \tag{11}$$

where η is the transit rate given by (8). We can compute (11) exactly when R is a mixture of polynomials and exponentials.

4.2 Jitter vs load for some distributions

In this subsection, we plot the jitter as a function of load for some important inter arrival time distributions. All the results are presented in units of service time. The load $\rho = \lambda/\mu$ is given on the horizontal axis and the jitter, measured in multiples of the service time, is plotted on the vertical axis.

4.2.1 Exponential inter-arrival

We can get an exact formula [5] when the service time is exponential. In that case, Eq. (6) becomes the integral of an exponential and we can easily compute the exact value and get after some algebra

$$J = E[|\ T_{j+1} - T_j\ |] = \frac{1}{\mu}. \tag{12}$$

For the M/M/1 queue, the jitter is the average service time over the whole range of traffic.

4.2.2 Deterministic inter-arrival

We now consider a deterministic arrival process. This is useful to model some codecs that emit packets or frames at fixed intervals. In this case, the inter-arrival time is equal to a constant $m = 1/\lambda$ and its probability density function is

$$f_R(y) = \delta\,(y-m) \tag{13}$$

which gives

$$g(c) = e^{-c/\lambda}.$$

Replacing in (10), we get the jitter

$$J_D = \frac{(\eta^2 + \mu^2)}{\eta\mu(\eta + \mu)} + \frac{2}{(\eta + \mu)} e^{-(\eta+\mu)/\lambda} - \frac{1}{\eta} e^{-\eta/\lambda} \tag{14}$$

The plot of the jitter as a function of the load is shown in Figure 1.

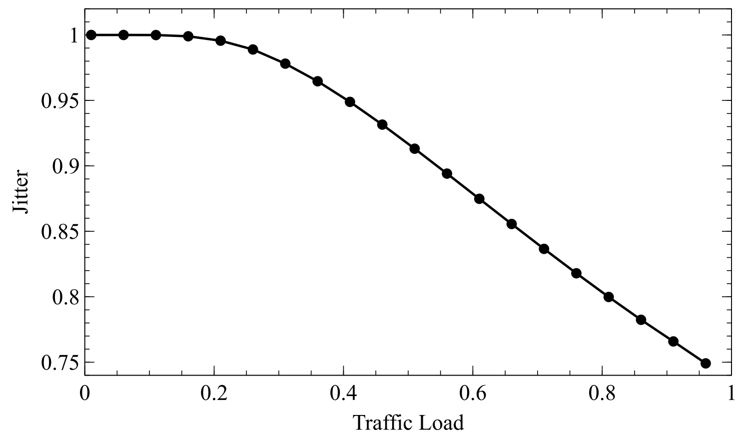

Figure 1: Jitter for a D/M/1 queue

4.2.3 Gamma inter-arrival

We can get an exact value of (11) when the arrival process follows a gamma distribution $Gm(k, \theta)$. (See Appendix A.1 for details.) Replacing (59) in (11), we get

$$g(c) = \frac{1}{(1 + c\theta)^k} \tag{15}$$

from which we can get the value of the jitter from (10). We present in Figure 2 the values of jitter computed for two values of the standard deviation $s_t = 0.5$ and 3. This shows that the variance of the process can have a significant effect of the jitter, as one would expect.

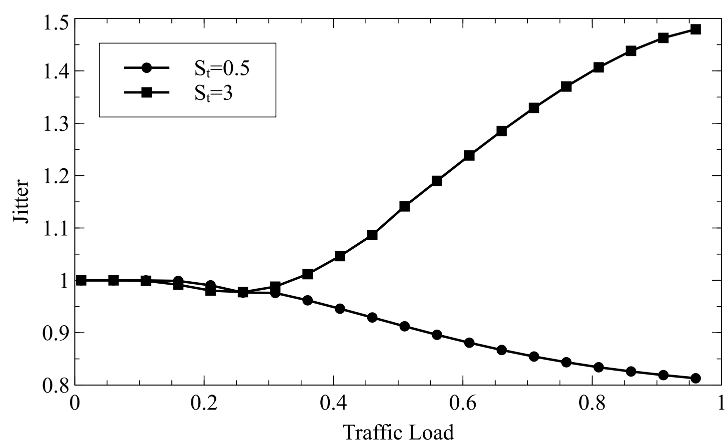

Figure 2: Jitter for a Gm/M/1 Queue

4.2.4 Pareto type I inter-arrival

There is a large body of work [21, 14] showing that the packet inter-arrival time in wide area or local networks is modeled by heavy-tailed distributions. One good model is the Pareto type I $P(x_m, \alpha)$ with scale x_m and shape α (Appendix A.2). Equation (11) can then be written as

$$g(c) = \alpha E_{\alpha+1}(cx_m) \tag{16}$$

which we can replace in (10) to get the jitter. Here, $E_n(x)$ is the exponential integral

$$E_n(x) = \int_1^\infty \frac{e^{-xt}}{t^n} dt. \tag{17}$$

We present in Figure 3 the value of the jitter as a function of load for the standard deviation $s_t = 0.25$.

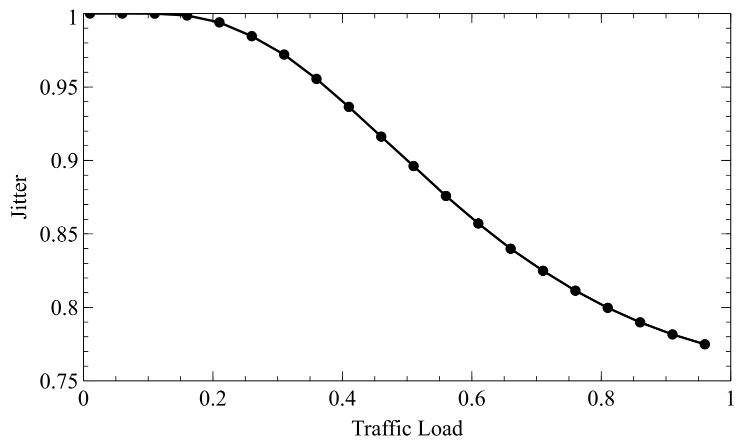

Figure 3: Jitter for a Pareto inter-arrival process

5 The G/G/1 queue: Exact limit values

There is another case where we can get exact results: For high and low traffic limits, the jitter does not depend on T. These results are interesting in their own right since they are valid for G/G/1 queues with a single flow so that we can use them to get insight on some important proprieties of jitter. Because they are relatively simple, it is possible to get either analytic formulas or some fast numerical evaluation. Also, as we discuss in Sections 6 and 7, these limits will be used to build some approximations when f_T is not known.

5.1 Small arrival rate

Suppose that the packet average arrival rate λ is so low that packets arriving to the queue almost always find an empty queue. We then have $W_i \approx 0$ in (3) and in that case,

$$\lim_{\lambda \to 0} J = E\left[|S_{i+1} - S_i|\right]. \tag{18}$$

We then get the general result

Proposition 1 *For a G/G/1 queue, in the low-traffic limit, the jitter depends only on the service time distribution and is given by the expectation of the absolute value of the difference, or mean difference, of the service times of two consecutive packets.*

Note that for some queues, like batch arrivals, the condition $\lambda \to 0$ would apply to the arrival of batches. We cannot conclude that $W_i \approx 0$ and the result does not apply in these cases. Also, we see that at low load, jitter need not go to zero, which is very different from the behavior of delay which does go to zero at low load.

5.2 Large arrival rate

We can also get an approximation for high load when $\lambda \to \mu$. In this case, we assume that when packet i arrives to the queue, packet $i-1$ has not yet started service. Let τ_i be the instant just before packet i arrives to the queue and $W_i(\tau_i)$ the amount of time packet i will have to wait, also called the workload, at time τ_i. We have

$$T_i = W_i(\tau_i) + S_i. \tag{19}$$

Next, consider the queue just before packet $i-1$ arrives. The workload at that time is $W_{i-1}(\tau_{i-1})$. During the interval $R_i = t_i - t_{i-1}$, the workload will increase by S_{i-1} since packet $i-1$ has joined the queue, and will decrease by R_i since the server is busy during the whole interval from τ_{i-1} to τ_i. The workload just before packet i arrives is then given by

$$W_i(\tau_i) = W_{i-1}(\tau_{i-1}) - R_i + S_{i-1}. \tag{20}$$

Replacing (20) in (19), we get

$$\begin{aligned} T_i &= W_{i-1}(\tau_{i-1}) - R_i + S_{i-1} + S_i \\ &= T_{i-1} - R_i + S_i \end{aligned} \tag{21}$$

from which we get

$$\lim_{\lambda \to \mu} J = E\left[|S_i - R_i|\right]. \tag{22}$$

We then have the proposition

Proposition 2 *For a G/G/1 queue, in the high traffic limit, the jitter is given by the expectation of the absolute value of the difference between the service and inter-arrival times of a packet.*

This is another important insight that shows how jitter remains finite at high load. This behavior is also very different from that of delay which becomes infinite at $\rho \approx 1$.

6 The M/G/1 queue: Linear approximation

It is known [19] that the distribution of packet sizes in the Internet is generally not exponential so that we now present some models for these cases. We focus on M/G/1 queues where the service time follows a general distribution. It is well known that network traffic in the core does not generally have Poisson arrivals. Still, we have done measurements in the access portion of the network of a large ISP showing that the traffic arrival process behaves as Poisson after only a few levels of aggregation of connections. This is why we feel that it is important to examine the M/G/1 queue.

In principle, we could use (6) since we know [20] that the Laplace transform of the sojourn time distribution in a M/G/1 queue is given by

$$\mathcal{F}_T(s) = \mathcal{F}_S(s)\frac{s(1-\rho)}{(s-\lambda) + \lambda \mathcal{F}_S(s)}. \tag{23}$$

In order to use this, we need first to compute $\mathcal{F}_S(s)$ either analytically, if possible, or numerically. This would produce a set of values for a suitable range of s. Next, we need to invert the Laplace transform numerically to get $\mathcal{F}_T(t)$, again producing a set of values for a suitable range of t, something that may not be all that easy to do since the numerical calculation of the inverse Laplace transform is notoriously difficult. We can then perform the integration over f_R in (6) which leaves us with a double integral in z and x.

This computation will most likely be time-consuming and instead, we propose a method based on linear interpolation using the fact that it is relatively easy to compute the asymptotic values (18) and (22). This is equivalent to the following proposition

Assumption 1 *In an M/G/1 queue, the jitter is approximately linear as a function of the traffic load ρ.*

Under this assumption, all we need is the values for J at $\rho = 0$ and $\rho = 1$ for which we can use the results of Section 5. We present results for the M/D/1, M/Gm/1 and M/Nt/1 queues in Figures 4, 5 and 6 where we check the accuracy of the linear assumption by comparing the calculated values, labeled `Line`, with the simulated values, labeled `sim`. The 95% confidence interval for the simulated values is smaller than the plot marker and does not show up on the graph. In all three cases, as well as for other distributions not shown here, we find that the agreement is excellent over the whole range of traffic.

To summarize, we see that for an M/G/1 queue, the linear approximation is very accurate for many different forms of the service time distribution. More importantly, this model is based on simple formulas for the two end points from Section 5 that can sometimes be calculated analytically or, when this is not possible, can be evaluated numerically. In both cases, the approximation will be very helpful for any iterative procedure that needs to compute the jitter as a function of the load.

7 The G/D/1 queue: Piece-wise linear approximation

We now examine queues with a constant service time since, as we mentioned in Section 2, a significant amount of traffic is made up of packets of constant length [19] so that

$$f_S(s) = \delta(s - \overline{S}).$$

Our first attempts were based on extending the techniques of the previous sections. First, we confirmed that using an exponential sojourn time $f_T(x)$ in (6) is not very accurate. The linear model that was such a good fit for the M/G/1 queue also turned out to be very inaccurate.

We then ran a number of simulations whose results are presented in Figures 7 and 8 and observed that the jitter is zero not only at very low load but for a significant range of low traffic values, sometimes as high as 0.5. After this point, we also saw that the jitter increases more or less linearly with load.

7.1 Two-segment approximation

This suggests an approximation based on two linear segments, one with zero slope and the other increasing linearly. We need to be able to compute three points P_0, P_1 and P_2, either analytically or numerically. We can use propositions (1) and (2) to compute the jitter for the two end points

$$P_0 = (0, J_0)$$
$$P_2 = (1, J_2)$$

where J_0 and J_2 are given by equations (18) and (22). Note that for the G/D/1 queue, $J_0 = 0$.

The intermediate point ρ_1 can be estimated based on the following argument. In the low-traffic limit, packets don't wait so that their transit time is simply the service time. This will happen as long as the interval between arrivals is significantly larger than the service time $\overline{S}$. More precisely, we consider that $J \approx 0$ as long as

$$Pr\left[R_i < \overline{S}\right] < \epsilon \tag{24}$$

for some small ϵ, typically 10^{-2} and where R_i is the inter-arrival time of packet i. The point ρ_1 is defined by solving (24) as an equality. Let F_R be the cdf of R. The condition (24) can be written as

$$F_R(\overline{S}; m, s_t, \ldots) = \epsilon \tag{25}$$

which we can solve for m and we get $\rho_1 = 1/(m\mu)$. From this, we can make a two-segment interpolation

$$J(\rho) = \begin{cases} 0 & \text{if } \rho \in [0, \rho_1] \\ \dfrac{J_2}{1-\rho_1}\left[\rho - \rho_1\right] & \text{if } \rho \in [\rho_1, 1]. \end{cases} \tag{26}$$

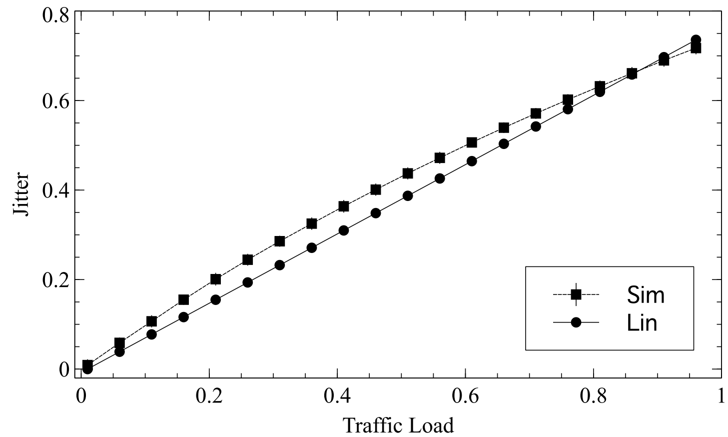

Figure 4: Jitter for a M/D/1 Queue

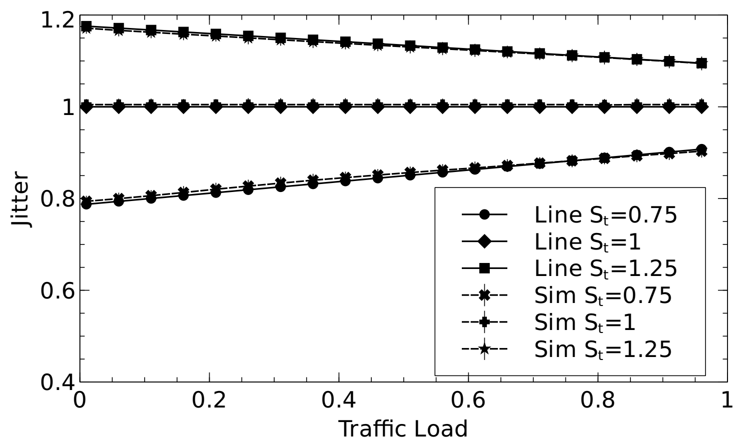

Figure 5: Jitter for a M/Gm/1 Queue

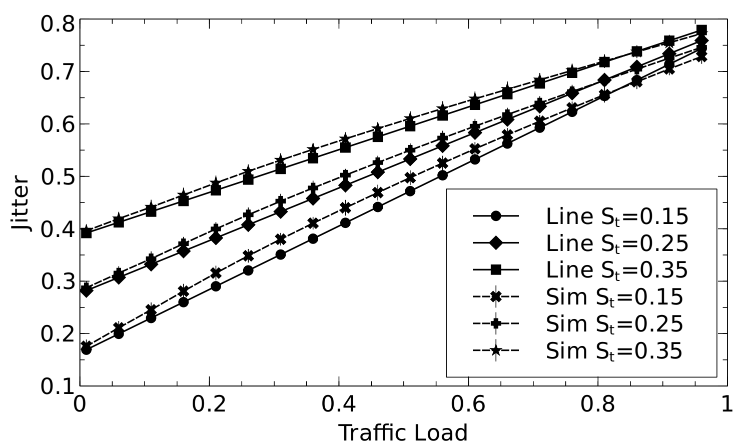

Figure 6: Jitter for M/Nt/1 Queue

7.2 Three-segment approximation

Although the interpolation (26) is continuous in $[0,1]$, its derivative is not continuous at $\rho = \rho_1$. This will be a problem if we want to use the formula within any algorithm that requires continuous derivatives wrt λ.

We now modify the technique of Section 7.1 to get a continuous and differentiable function. First we choose a small $\delta > 0$, typically 0.1, and define two points close to ρ_1

$$\rho^+ = \rho_1(1+\delta)$$
$$\rho^- = \rho_1(1-\delta)$$

with the corresponding jitter values

$$J(\rho^-) = J_0$$
$$J(\rho^+) = \frac{J_2}{1-\rho_1}\left[\rho^+ - \rho_1\right]$$

where $J(\rho^+)$ is the value of J at $\rho = \rho^+$ on the linear segment from (26).

The three segments are defined as follows. In the interval $[0, \rho^-]$, we take $J = J_0 = 0$. In the interval $[\rho^+, 1]$, J is given by the linear approximation in the second part of (26). In the segment $[\rho^-, \rho^+]$, J is modelled by a third degree polynomial

$$P(\rho) = a\rho^3 + b\rho^2 + c\rho^1 + d. \tag{27}$$

We choose the coefficients so that both the function and its derivative are continuous on the whole interval

$$P(\rho^-) = 0 \tag{28}$$
$$P'(\rho^-) = 0 \tag{29}$$
$$P(\rho^+) = \frac{J_2}{1-\rho_1}\left[\rho^+ - \rho_1\right] \tag{30}$$
$$P'(\rho^+) = \frac{J_2}{1-\rho_1}. \tag{31}$$

Conditions (28) and (30) guarantee the continuity at ρ^- and ρ^+ and the two others ensure the differentiability. To summarize, the4 following 3-segments model is

$$J(\rho) = \begin{cases} 0 & \text{if } \rho \in [0, \rho^-] \\ P(\rho) & \text{if } \rho \in [\rho^-, \rho^+] \\ \frac{J_2}{1-\rho_1}\left[\rho^+ - \rho_1\right] & \text{if } \rho \in [\rho^+, 1]\ . \end{cases} \tag{32}$$

We check the accuracy of the model by comparing the values from (32) with simulation results for some distributions of G.

7.2.1 Gamma inter-arrival

We start with the case where the inter-arrival time has a gamma distribution $Gm(k,\theta)$. Assume that the variance s_t^2 is given. In that case, the two parameters k and θ depend only on the mean m through (62–63). We can then write (25) as

$$F_R\left[\overline{S}, k(m), \theta(m)\right] = \epsilon$$

which we can solve numerically for m. From this value, we get $\rho_1 = \overline{S}/m$. Figure 7 shows the results of the approximation with $\delta = 0.1$. Curves produced by simulations for different values of the variance are labelled `sim` and they are compared to the 3-segments curves labelled `Fit`. As in other cases, the agreement is very good for most cases, except when $s_t = 3$, which has a very large variance.

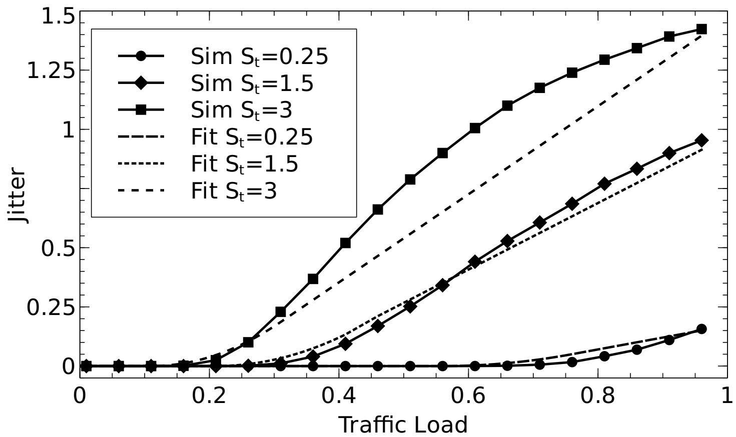

Figure 7: Jitter for Gm/D/1 Queue

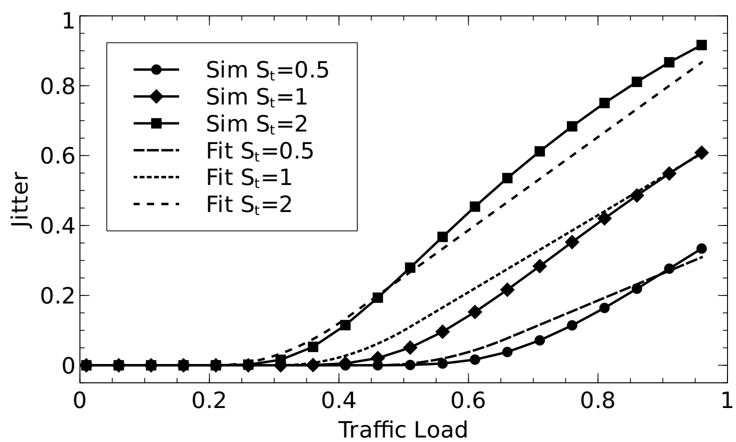

Figure 8: Jitter for LogN/D/1 Queue

7.2.2 Log-normal inter-arrival

We get similar results when the inter-arrival time follows a log-normal distribution. We compare the approximation to some simulation results in Figure 8 for the same value of δ as for the Gm/D/1. Here again the agreement is found to be quite good.

8 Computation times

The results of the previous sections show that it is possible to compute the jitter for some interesting classes of queues. Here we discuss the computation times required for different traffic processes. We used the Python `SciPy` module for scientific calculations with default parameter settings. The only Python code was for calling the evaluation of the functions by the `numpy` or `SciPy.stats` modules. The actual integration used the `quad` or `dblquad` function for single or double integration. from the `QuadPack` module of `SciPy`. We computed the solution of nonlinear equations with the `fsolve` function from the `Scipy.optimize` module. All these functions are the Netlib functions and are run as compiled code so that the cpu times reported here should not be too different from a full implementation in a compiled language. All calculations are done on an Intel core i5 with 2 cores at 2.53 GHz.

8.1 G/M/1

The computation of the jitter is made up of two separate parts: Evaluating the transit rate η from (9) and then computing the jitter itself from (11). We give a short discussion of the computation time required for these two parts

8.1.1 Computing the transit rate

The transit rate is given implicitly by the solution of the nonlinear equation (9). For this, we need the Laplace transform $\mathcal{F}_R$ of the arrival process R. The computation time will thus depend strongly on whether there exists an analytic expression for $\mathcal{F}_R$. Note that in general, it is not possible to solve (9) analytically even when such an analytic expression exists. For instance, for the Gamma distribution, we have

$$\mathcal{F}_R(s) = (1 + \theta s)^{-k} \tag{33}$$

but replacing in (9) yields

$$\tau = [1 + \theta\mu\,(1 - \tau)]^{-k}$$

which cannot be solved analytically. In other cases, like the log-normal [22] or Pareto [23], there exist good approximations so that solving (9) with a numerical algorithm is not too time-consuming.

If there is no exact or approximate value, we need to compute the Laplace transform numerically every time the solution procedure needs a value of $\mathcal{F}_R$ for some value of the argument τ. This requires in each case an integration of F_R which can lead to large computation times. For instance, in the case of the Gamma distribution, a solution using (33) will take less than one millisecond while the computation time when evaluating the Laplace transform by numerical integration would take nearly 300 ms.

8.1.2 Computing the jitter

Once the value of η has been computed, the calculation of the jitter reduces to the evaluation of the integral (11). The computational effort is similar to what is discussed in Section 5. In some cases, the integral can be computed analytically, in which case the computation time is negligible. In other, more difficult cases, one has to do a single numerical integration and the computation times are of the same magnitude as the ones reported in Table 2 in Section 8.2.2.

8.2 Limit cases

We now present some detailed results for the computation of the bounds. They give some insight on the computational load required by single and double integration. We also give some analytic expressions for the jitter when they are available. From (18) and (22), we can compute the jitter for very general queues in the two limit cases. Still, there remains the question of how fast this can be done. First we recall some useful relations. Let X and Y be two independent random variables with probability density functions $f_X(x)$ and $f_Y(y)$ and define Z as

$$Z = X - Y. \tag{34}$$

The pdf of Z is the cross-correlation of X and Y

$$f_Z(z) = \int_{\Re} f_X(z+y) f_Y(y) dy = \int_{\Re} f_X(x) f_Y(x-z) dx \tag{35}$$

where we assume that a distribution $f_X(x) = 0$ for x outside its support. We are interested in the pdf $f_{|Z|}$ of $|Z|$ which is given by

$$\begin{aligned} f_{|Z|}(z) &= f_Z(z) + f_Z(-z) \\ &= f_Z^+(z) + f_Z^-(z) \end{aligned} \tag{36}$$

where f^+ and f^- are called the positive and negative parts of f_Z in the following. Replacing in (35), we get

$$f_{|Z|}(z) = \int_{\Re} f_X(z+y) f_Y(y) dy + \int_{\Re} f_X(y-z) f_Y(y) dy. \tag{37}$$

If f_X and f_Y are defined for $x, y \geq 0$, as will be the case here, we get

$$\begin{aligned} f_{|Z|}(z) &= f_Z^+(z) + f_Z^-(z) \\ &= \int_0^\infty f_X(z+y) f_Y(y) dy + \int_z^\infty f_X(y-z) f_Y(y) dy \\ &= \int_0^\infty f_X(z+y) f_Y(y) dy + \int_0^\infty f_X(y) f_Y(z+y) dy \end{aligned} \tag{38}$$

for $0 \leq z < \infty$. If $X = Y$, this reduces to

$$f_{|Z|}(z) = 2 \int_0^\infty f_X(z+y) f_Y(y) dy. \tag{39}$$

Finally, we get the average as

$$E[|Z|] = \int_0^\infty z f_{|Z|}(z) dz. \tag{40}$$

We can get a simplified form for J in the important case of G/D/1 queues where f_S is deterministic of the form

$$f_S(y) = \delta(y - \overline{S}). \tag{41}$$

Replacing this in (38), we get

$$\begin{aligned} E\left[|Z|\right] &= \int_0^\infty z f_R(z + \overline{S}) + \int_0^\infty z f_R(\overline{S} - z) \\ &= \int_0^\infty z f_R(z + \overline{S}) + \int_0^{\overline{S}} z f_R(\overline{S} - z) \end{aligned} \tag{42}$$

since f_R has positive support only.

Altogether, the complexity of the computation depends on whether we can get an analytic expression for (38) and (40), which in turn depends on the particular form of the density functions f_R and f_S. We present computational results for three cases depending on which expression can be evaluated analytically: Both (38) and (40), only (38) or neither of them.

8.2.1 Analytic expressions for J

In some cases, we can find an explicit expression both for the probability density function (38) and for the average (40). This yields an analytic formula for the jitter and the computation time is negligible.

G/M/1 at $\rho = 0$ A simple case where this happens is for the G/M/1 queue. Here, the limit (18) at $\rho \approx 0$ is given by the difference of two identical independent exponential variables with parameter μ. This has a symmetric Laplace distribution centered at the origin and the distribution of the absolute value will be an exponential with parameter μ. We have the same situation for the limit $\rho \approx 1$ where the two exponential variables R and S are exponentials with parameters λ and μ where $\lambda \to \mu$. In both the low and high limits

$$J = 1/\mu. \tag{43}$$

M/D/1 at $\rho = 1$ We can get an exact result for the M/D/1 queue at $\rho \approx 1$ by replacing $f_R(x) = \lambda \exp(-\lambda x)$ into (42). We get

$$J = \lambda \int_0^\infty z e^{-(z+\overline{S})} + \int_0^{\overline{S}} z e^{-(\overline{S}-z)}$$
$$= \frac{\lambda \overline{S} - 1}{\lambda} + \frac{2e^{-\lambda \overline{S}}}{\lambda}, \tag{44}$$
$$\lim_{\lambda \overline{S} \to 1} J = \frac{2}{e}. \tag{45}$$

G/Gm/1 at $\rho = 0$ Another case where we can get a complete analytic formula at $\rho \approx 0$ is the G/Gm/1 queue. We can compute the integral (38) and get

$$f_{|Z|}(z) = \frac{1}{\Gamma(k)^2 \theta^{2k}} \left[e^{-z/\theta} \int_0^\infty (z+x)^{k-1} x^{k-1} e^{-2x/\theta} dx + e^{z/\theta} \int_z^\infty (x-z)^{k-1} x^{k-1} e^{-2x/\theta} dx \right]$$

which evaluates to [24]

$$f_{|Z|}(z) = \frac{z^{k-1/2}}{2^{k-1/2}\Gamma(k)^2 \theta^{k+1/2}} \times \left[\frac{\sqrt{\pi} e^{z/\theta}}{\sin(k\pi)\Gamma(1-k)} + \frac{\Gamma(k) e^{-z/\theta}}{\sqrt{\pi}} \right] K_{k-1/2}(z/\theta) \tag{46}$$

where $K_\nu(x)$ is the modified Bessel function of the second kind. Replacing (46) in (40), we get

$$J = \frac{2\theta \Gamma(1/2 + k)}{\sqrt{\pi}\Gamma(k)} \quad \forall k > 0. \tag{47}$$

Gm/D/1 at $\rho = 1$ We have a similar result for the Gm/D/1 queue at $\rho \approx 1$. When the arrival distribution is $Gm(k, \theta)$, we can evaluate J using the two terms of (42). The positive part is given by

$$\int_0^\infty z f_R(z + \overline{S}) = \frac{1}{\Gamma(k)\theta} \int_0^\infty z \left(\frac{\overline{S} + z}{\theta} \right)^{k-1} e^{-(\overline{S}+z)/\theta} \, dz \tag{48}$$

Setting $y = (\overline{S} + z)/\theta$, this is equivalent to

$$\int_0^\infty z f_R(z + \overline{S}) = \frac{1}{\Gamma(k)\theta} \int_0^\infty \left(\theta y - \overline{S}\right) y^{k-1} e^{-y} \, \theta dy$$
$$= \frac{1}{\Gamma(k)} \left[\theta \Gamma\left(k+1, \frac{\overline{S}}{\theta}\right) - \overline{S}\Gamma\left(k, \frac{\overline{S}}{\theta}\right) \right]. \tag{49}$$

For the negative part, we get

$$\int_0^{\overline{S}} z f_R(\overline{S} - z) = \int_0^{\overline{S}} z \left(\frac{\overline{S} - z}{\theta} \right)^{k-1} e^{-(\overline{S}-z)/\theta} \, dz \tag{50}$$

and using the transformation $y = (\overline{S} - z)/\theta$, we get

$$\int_0^{\overline{S}} z f_R(\overline{S} - z) = \frac{1}{\Gamma(k)\theta} \int_0^{\overline{S}/\theta} \left(\overline{S} - \theta y\right) y^{k-1} e^{-y} \, dy$$
$$= \frac{1}{\Gamma(k)} \left[\overline{S}\gamma\left(k, \frac{\overline{S}}{\theta}\right) - \theta\gamma\left(k+1, \frac{\overline{S}}{\theta}\right) \right] \tag{51}$$

where $\gamma(k, x)$ and $\Gamma(k, x)$ are the lower and upper incomplete gamma functions. These cases are summarized in Table 1.

8.2.2 Analytic expression for $f_{|Z|}$ only

More difficult cases happen when we can get an analytic expression for the distribution of the absolute value from (38) but where it is not possible to compute the average (40) analytically and we need to compute the integral (40) numerically. This is also the case for the G/D/1 at $\rho = 1$ where we cannot compute (42) analytically.

Table 1: Analytic expression

Queue	ρ	Calculation
G/M/1	0	(43)
M/D/1	1	(44)
G/Gm/1	0	(47)
Gm/D/1	1	(49–51)

M/Nt/1 at $\rho = 1$ The first example is the M/Nt/1 queue where the service time S follows a normal distribution $Nt(x; \mu_n, \sigma_n)$ truncated at 0 with location μ_n and scale σ_n. We get for the two terms of (38)

$$f_Z^+(z) = \frac{\lambda}{\Phi\left(\frac{\mu_n}{\sigma_n}\right)} \exp\left(\frac{\lambda^2\sigma_n^2}{2} - \lambda(\mu_n + z)\right) \Phi\left(\frac{\mu_n - \lambda\sigma_n^2}{\sigma_n}\right) \tag{52}$$

$$f_Z^-(z) = \frac{\lambda}{\Phi\left(\frac{\mu_n}{\sigma_n}\right)} \exp\left(\frac{\lambda^2\sigma_n^2}{2} - \lambda(\mu_n - z)\right) \times \Phi\left(\frac{\mu_n - z - \lambda\sigma_n^2}{\sigma_n}\right) \tag{53}$$

but it is not possible to get an analytic expression for (40) from this and we need to compute the integral numerically.

LogN/D/1 at $\rho = 1$ We get the same case for the LogN/D/1 queue for $\rho \to 1$ where R follows a log-normal distribution $LogN(x; , \mu_L, \sigma_L)$ and S is deterministic. We cannot evaluate (42) analytically so we need to compute the integral (42) numerically.

$$f_Z^+ = e^{\left(\frac{1}{2}\sigma_L^2 + \mu_L\right)} \Phi\left(\frac{\sigma_L^2 + \mu_L - \log(\overline{S})}{\sigma_L}\right) \tag{54}$$

$$f_Z^- = e^{\left(\frac{1}{2}\sigma_L^2 + \mu_L\right)} \Phi\left(-\frac{\sigma_L^2 + \mu_L - \log(\overline{S})}{\sigma_L}\right) \tag{55}$$

We can get the distribution of J by (36) from the sum of the two terms (54–55) but here again it is not possible to compute the mean value analytically.

G/Nt/1 at $\rho = 0$ The pdf of the service time is that of a truncated normal random variable S. We can evaluate the two parts of (38) analytically and we get

$$\int_0^\infty f_S(z+x) f_S(x) dx = \frac{1}{2\sigma_n\sqrt{\pi}\left[\Phi\left(\frac{\mu_n}{\sigma_n}\right)\right]^2} \times \exp\left(-\frac{z^2}{4\sigma_n^2}\right) \Phi\left(\frac{2\mu_n - z}{\sigma_n\sqrt{2}}\right) \tag{56}$$

$$\int_{-z}^\infty f_S(z+x) f_S(x) dx = \frac{1}{2\sigma_n\sqrt{\pi}\left[\Phi\left(\frac{\mu_n}{\sigma_n}\right)\right]^2} \times \exp\left(-\frac{z^2}{4\sigma_n^2}\right) \Phi\left(\frac{2\mu_n + z}{\sigma_n\sqrt{2}}\right) \tag{57}$$

We can compute the pdf of $|Z|$ by replacing (56) and (57) in (36) which gives

$$f_{|Z|}(z) = \frac{1}{\sigma_n\sqrt{\pi}\left[\Phi\left(\frac{\mu_n}{\sigma_n}\right)\right]^2} \exp\left(-\frac{z^2}{4\sigma_n^2}\right) \Phi\left(\frac{2\mu_n + z}{\sigma_n\sqrt{2}}\right). \tag{58}$$

In all these cases, Eq. (40) has to be evaluated numerically with a significant computation time. Some values for the computation times are presented in Table 2. We fix the mean of R to 0 or 1 as the case may be and measure the cpu time for different values of the standard deviation s_t. The worst case is for the M/Nt/1 queue at $\rho = 1$ which needs about 200 ms to compute. We see that in most cases, the integration is quite fast so that it would be possible to use this within some iterative procedure.

8.2.3 No analytic formulas

An even more difficult situation arises when it is not possible to evaluate (38) analytically. In this case, we evaluate (40) by computing numerically a double integral: one for the distribution function and one for the

Table 2: Cpu times for single integration

Queue	ρ	s_t	cpu sec
G/Nt/1	0	0.15	0.037
G/Nt/1	0	0.25	0.036
G/Nt/1	0	0.35	0.036
M/Nt/1	1	0.15	0.20
M/Nt/1	1	0.25	0.25
M/Nt/1	1	0.35	0.21
Gm/D/1	1	0.25	0.002
Gm/D/1	1	1.5	0.006
Gm/D/1	1	3.5	0.008
Logn/D/1	1	0.5	0.004
Logn/D/1	1	1.	0.005
Logn/D/1	1	2.	0.008

average. We present in Table 3 some numerical results for the M/Nt/1 queue in both limit cases. We use the same computation tool as in Section 8.2.2 but use the `dblquad` function for double integration.

We repeat the calculation for the G/Nt/1 queue at $\rho = 0$ that is shown in Table 2 to estimate the difference in cpu time when using double as opposed to single integration. Typical computation times by double integration are about 3 orders of magnitude larger than for single integration, around 10 seconds. It would be difficult to use double integration iteratively or in real time. Still, we could compute these values off-line and use them as input to the approximation algorithms for the estimation of jitter as a function of traffic when f_T is not known.

Table 3: Cpu times for double integration

Queue	ρ	s_t	cpu sec
G/Nt/1	0	0.15	11.8
G/Nt/1	0	0.25	13.0
G/Nt/1	0	0.35	12.4
M/Nt/1	1	0.15	15.5
M/Nt/1	1	0.25	10.2
M/Nt/1	1	0.35	10.3

8.3 G/D/1

For the three-segment model, we need the three points P_0, P_1 and P_2. The first point P_0 is determined immediately, since for G/D/1 queues,

$$\lim_{\rho \to 0} J(\rho) = 0.$$

The time required to compute P_2 when $\rho \approx 1$ is equivalent to time evaluated in Section 8.2. Finally, computing the intermediate point P_1 requires the solution of equation (25). It takes about 3 ms when the inter-arrival time R follows a Gamma distribution and approximately the same for a Lognormal distribution. Altogether, the computation is dominated by the computation of P_2 and despite the fact that the approximation may not be very accurate in some cases , the simplicity and the computation speed could still make it very useful.

9 Conclusion

We proposed some techniques for the calculation of jitter for different types of FCFS queues with a single stream. First we recalled a general formula proposed previously and pointed out two problems: One is that it depends on the distribution of the sojourn time, which is often not known, and two, that it takes the form of a triple integral, which can be difficult to solve numerically.

We first proposed an exact model for G/M/1 queue since the sojourn time follows an exponential distribution whose parameter is the solution of a nonlinear equation involving the Laplace transform of the inter-arrival distribution. In the worst case, when the transform has to be computed numerically, solving this equation required in the order of a few hundred milliseconds. We then reduced the calculation of the triple integral to a single integral which can be computed numerically quickly. This approach is feasible because the sojourn time parameter can be evaluated offline outside the iteration procedure.

We then gave two expressions for a G/G/1 queue at low and high traffic that do not depend on the sojourn time. We also gave examples where it is possible to compute these values analytically or with a single numerical integration. More complex cases require a double integral which require about 10 to 20 seconds. These limit points are then used to construct approximate models for queues where the sojourn time distribution is not known.

Next, we showed that for the M/G/1 queue, a linear interpolation between the low and high limits is an excellent approximation of the actual jitter so that the computation time is essentially that of the limit points.

Finally, we presented a good approximation for the G/D/1 queue using a three-segment interpolation. The computational requirement is the solution of a nonlinear equation for the cdf of the arrival process. The results were compared with simulation results indicating a reasonably good accuracy with poorer results with processes with a high variance.

Our conclusion is that we can use these expressions to evaluate quickly the jitter in a queue with reasonable accuracy and for a number of important processes. The computation speed is small enough that we can use them to investigate the effect of jitter in network design and optimization models.

We also get a better understanding of the properties of jitter as compared with delay. We find that jitter does not go always to zero at low load and does not go to infinity at high load where it can sometimes decrease with increasing load. In fact, the values seldom exceed the holding time and when they do, only by a small value. This seems to happen when processes have a large variance, something that is not totally unexpected. An even more counter-intuitive result is that in some cases, jitter can actually *decrease* when the load increases. These properties seem to hold for a range of arrival and service processes, which lead us to believe that they may in fact be quite general.

A Notation

For the sake of completeness, we recall the definitions of the distributions used in the paper.

A.1 Gamma

The Gamma distribution $Gm(x; k, \theta)$ with shape $k > 0$ and scale $\theta > 0$ has a pdf given by

$$f_R(x) = x^{k-1} \frac{e^{-x/\theta}}{\theta^k \Gamma(k)}. \tag{59}$$

The cdf is the regularized gamma function

$$F_R(x, k, \theta) = \frac{1}{\Gamma(k)} \int_0^{x/\theta} t^{k-1} e^{-t} dt.$$

The parameters are related to the mean m and variance $v = s_t^2$ of the distribution by

$$m = k\theta \tag{60}$$

$$v = k\theta^2 \tag{61}$$

which we can solve to get

$$\theta = \frac{v}{m} \tag{62}$$

$$k = \frac{m^2}{v}. \tag{63}$$

A.2 Pareto

The Pareto Type I $P(x; x_m, \alpha)$ with scale x_m and shape α has a density function given by

$$f_R(x) = \begin{cases} \alpha \dfrac{x_m^\alpha}{x^{\alpha+1}} & \text{if } x \geq x_m \\ 0 & \text{otherwise.} \end{cases} \tag{64}$$

The Pareto parameters α and x_m are related to the mean m and variance $v = s_t^2$ by

$$m = \begin{cases} x_m \dfrac{\alpha}{\alpha - 1} & \text{if } \alpha > 1 \\ \infty & \text{if } \alpha \leq 1 \end{cases} \tag{65}$$

$$v = \begin{cases} x_m^2 \dfrac{\alpha}{(\alpha-1)^2(\alpha-2)} & \text{if } \alpha > 2 \\ \infty & \text{if } \alpha \leq 2. \end{cases} \tag{66}$$

For $\alpha > 2$, we can solve (65–66) to get

$$\alpha = 1 + \sqrt{1 + (m^2/v)} \tag{67}$$

$$x_m = m \frac{\alpha - 1}{\alpha}. \tag{68}$$

A.3 Normal

We denote the pdf of the standard normal distribution

$$\phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2) \tag{69}$$

and the corresponding cdf

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} \exp(-t^2/2) dt. \tag{70}$$

The error function is given by

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \tag{71}$$

and we have the relation

$$\Phi(x) = \frac{1}{2} + \frac{1}{2} \text{erf}\left(\frac{x}{\sqrt{2}}\right). \tag{72}$$

The pdf of a normally distributed random variables X with location μ_n and scale σ_n is given by

$$f_X(x; \mu_n, \sigma_n) = \frac{1}{\sigma_n} \phi\left(\frac{x - \mu_n}{\sigma_n}\right) \tag{73}$$

and the corresponding cdf by

$$F_X(x; \mu_n, \sigma_n) = \Phi\left(\frac{x - \mu_n}{\sigma_n}\right). \tag{74}$$

A.4 Truncated normal

The pdf of the truncated normal is that of a normal random variable X with location and scale parameters μ_n and σ_n and truncated to the interval $[0, \infty]$, is given by

$$f_X(x;\mu_n,\sigma_n) = Nt(x;m,s_t) = \frac{1}{\sigma_n \Phi\left(\mu_n/\sigma_n\right)} \phi\left(\frac{x-\mu_n}{\sigma_n}\right) \tag{75}$$

which is simply a gaussian distribution with support $[0, \infty]$ with the appropriate normalization. The mean m and variance $v = s_t^2$ are given by

$$m = \mu_n + \sigma_n A \tag{76}$$

$$v = \sigma_n^2 (1 - B) \tag{77}$$

where we have defined

$$A = \frac{\phi\left(\frac{\mu_n}{\sigma_n}\right)}{\Phi\left(\frac{\mu_n}{\sigma_n}\right)}$$

$$B = A\left(A - \frac{\mu_n}{\sigma_n}\right).$$

Note that the parameters μ_n and σ_n need not exist for an arbitrary choice of m and s_t.

A.5 Log-normal

This is the distribution of a random variable X such that $\log X$ is normally distributed. The pdf for location μ_L and scale σ_L is given by

$$f_X(x;\mu_L,\sigma_L) = \frac{1}{x\sqrt{2\pi}\sigma_L} \exp\left(-\frac{(\ln x - \mu_L)^2}{2\sigma_L^2}\right) \tag{78}$$

The mean m and variance $v = s_t^2$ are given by

$$m = e^{\mu_L + \sigma_L^2/2}$$

$$v = (e^{\sigma_L^2} - 1) e^{2\mu_L + \sigma_L^2}$$

which can be inverted to give

$$\mu_L = \ln\left(\frac{m^2}{\sqrt{v+m^2}}\right)$$

$$\sigma_L = \sqrt{\ln\left(1 + \frac{v}{m^2}\right)}$$

References

[1] T. Hoang, Planning and optimization of multi-service computer networks, in Proc. 10th Communications and Networking Simulation Symposium, Mar. 2007.

[2] P. Barreto and P. Carvalho, Network planning optimization for multimedia networks, in Proc. 8th IEEE International Symposium on Network Computing and Applications, Jul. 2008, 60–67.

[3] C. Demichelis and P. Chimento, RFC 3393: IP packet delay variation metric for IP performance metrics (IPPM), IETF, Nov. 2002.

[4] Internet protocol data communication service–IP packet transfer and availability performance parameters, ITU-T Recommendation Y.1540, Mar. 2011. [Online]. Available: https://www.itu.int/rec/T-REC-Y.1540-201103-I/en

[5] H. Dahmouni, A. Girard, and B. Sansò, An analytical model for jitter in IP networks, Annals of Telecommunications, 67, 81–90, Jan. 2012. [Online]. Available: http://www.springerlink.com/content/5046104335872460/

[6] J. Roberts and F. Guillemin, Jitter in ATM networks and its impact on peak rate enforcement, IEEE/ACM Transactions on Networking, 16, 35–48, 1992.

[7] W. Matragi, C. Bisdikian, and K. Sohraby, Jitter calculus in ATM networks: single node case, in Proc. IEEE INFOCOM, 1, Toronto, 1994, 232–241.

[8] W. Matragi, K. Sohraby, and C. Bisdikian, Jitter calculus in ATM networks: Multiple node case, IEEE/ACM Transactions on Networking, 5, 122–133, 1997.

[9] A. Privalov and K. Sohraby, Per-stream jitter analysis in CBR ATM multiplexors, IEEE/ACM Transactions on Networking, 6, 141–149, 1998.

[10] J.-M. Chung and H.-M. Soo, Jitter analysis of homogeneous traffic in differentiated services networks, IEEE Communications Letters, 7(3), 130–132, 2003.

[11] O. Brun, C. Bockstal, and J. Garcia, Analytical approximation of the jitter incurred by CBR traffics in IP networks, Telecommunication Systems, 33, 23–45, Dec. 2006.

[12] H. Alshaer and J. Elmirghani, Expedited forwarding end-to-end delay and jitter in DiffServ, International Journal of Communication Systems, 21, 815–841, 2008.

[13] H. Dahmouni, A. Girard, M. Ouzineb, and B. Sansò, The impact of jitter on traffic flow optimization in communication networks, IEEE Transactions on Network and Service Management, 9(3), 279–292, Sep. 2012.

[14] V. Paxson and S. Floyd, Wide area traffic: The failure of Poisson modeling, IEEE/ACM Transactions on Networking, 3(3), 226–244, Jun. 1995.

[15] R. Garcia, X. Paneda, V. Garcia, D. Melendi, and M. Vilas, Statiscal characterization of a real video on demand service: User behaviour and streaming-media workload analysis, Simulation Modelling Practice and Theory, 15(6), 672–689, 2007.

[16] M.E. Crovella and A. Bestavros, Self-similarity in World Wide Web traffic: Evidence and possible causes, IEEE/ACM transaction on Networking, 5(6), 835–846, Dec. 1997.

[17] L. Liang, Y. Chen, and Z. Sun, Characterisation of internet traffic in wireless networks, Network Performance Engineering, 5233, 191–202, 2011. [Online]. Available: http://link.springer.com/chapter/10.1007/978-3-642-02742-0_9#

[18] C. Cifliklia, A. Gezera, A. Özahina, and O. Özkasapb, Bittorrent packet traffic features over IPv6 and IPv4, Simulation Modelling Practice and Theory, 18, 1214–1224, Oct. 2010.

[19] C. Fraleigh, S. Moon, B. Lyles, C. Cotton, M. Khan, D. Moll, R. Rockell, T. Seely, and S. Diot, Packet-level traffic measurements from the Sprint IP backbone, Network, 17(6), 6–16, 2003.

[20] L. Kleinrock, Queuing Systems. Wiley, 1975.

[21] W. Leland, M. Taqqu, W. Willinger, and D. Wilson, On the self-similar nature of Ethernet traffic, IEEE/ACM Transactions on Networking, 1–15, Feb. 1994.

[22] C. Tellambura and D. Senaratne, Accurate computation of the MGF of the lognormal distribution and its application to sum of lognormals, IEEE Transactions on Communications, 58, 1568–1577, 2010.

[23] S. Nadarajah and S. Kotz, On the laplace transform of the pareto distribution, IEEE Communications Letters, 10(9), 682, Sep. 2006.

[24] I. Gradshteyn and I. Ryzhik, Table of Integrals, Series, and Products, 7th ed. Elsevier Inc., 2007.