Sensitivity analysis and fast computation of packet loss probabilities in multiplexer models

A. Cuyt[*1], **R.B. Lenin**[**1], and **K. Van der Borght**[1]

[1] Department of Mathematics and Computer Science, University of Antwerp, Middelheimlaan 1, B2020 Antwerp, Belgium

Received 30 June 2003, accepted 2 December 2003
Published online 15 March 2004

Key words Packet loss probability function, rational interpolation, interval arithmetic, multiplexer model
Subject classification 41A20, 90B12, 65G10

A statistical multiplexer is a basic model used in the design and the dimensioning of communication networks. The multiplexer model consists of a finite buffer, to store incoming packets, served by a single server with constant service time, and a more or less complicated arrival process. The aim is to determine the packet loss probability as a function of the capacity of the buffer. An exact analytic approach is unfeasible in real time, and hence we show how techniques from rational approximation theory can be applied to the computation of the packet loss. Since the parameters used in such networks may not produce precise probabilities of interest without having to introduce drastic assumptions, we also carry out a perturbation analysis with respect to these parameters. Using techniques from interval arithmetic, we can deliver sharp bounds for the true uncertainity effect. The latter is very important because the packet loss probability function can be very sensitive to relatively small changes in the network parameters.

1 Introduction

Fixed length packet switches have been studied extensively in the context of ATM switching models. However, since the Internet is primarily TCP/IP with variable length packets, it is even more important to analyze switching in the new context. Variable bit rate (VBR) communications with real time constraints in general, and video communication services (video phone, video conferencing, television distribution) in particular, are expected to be a major class of services provided by the future Quality of Service (QoS) enabled Internet.

The introduction of statistical multiplexing techniques offers the capability to efficiently support VBR connections by taking advantage of the variability of the bandwidth requirements of individual connections. These techniques will handle a variety of traffic types such as video, voice, still images and data, each with their own QoS. Accurate traffic modelling and analysis of the QoS parameters in the multiplexer environment will enable the admission controller to make decisions that ensure the integrity of the traffic sources and that are efficient to the network. An important QoS measure that we will study in this paper is the packet (or cell) loss probability (PLP).

To compute the PLP $P_L(N)$ as a function of the buffer size N, several approaches have been developed in recent years, based on exact analytical techniques, approximate techniques or simulation.

In exact analysis, the traffic is described by Markovian arrival processes, leading to a Markov model of finite M/G/1-type [11, 23]. Queues of finite M/G/1-type give rise to finite embedded Markov chains whose transition matrices are upper block Hessenberg [12]. The complexity of the algorithms used to find the stationary probabilities of M/G/1-type queues is $O(cN_1^3N_2^2)$, where c is the number of servers, N_1 is the dimension of a block and N_2 is the number of block rows. This order of complexity does not allow one to compute $P_L(N)$ for large values of N in real time.

*Corresponding author: e-mail: Annie.Cuyt@ua.ac.be, Phone: +00 323 265 3898, Fax: +00 323 265 3777
**e-mail: Lenin@saii.ac.in, Phone: +91 40 5582 6155, Fax: +91 40 2326 0013

In approximate analysis, models from fluid queues have been used [7]. However, the computational requirements of the algorithms grow quite rapidly as the system's complexity grows.

Monte Carlo simulation is also an option to compute the packet loss probabilities. However, if the desired probability is in the range of 10^{-6} to 10^{-12} (rare event probability), it is computationally impossible to use the conventional Monte Carlo simulation. A simulation technique called Importance Sampling (IS) can speed up simulations involving rare events. However, because of the complicated nature of multiplexing queueing models, applying the IS technique is not straightforward.

The rational approximation technique proposed in [4] is a kind of "divide and conquer" technique, in the sense that :

– for small values of the buffer capacity N, the exact value $P_L(N)$ is computed;
– the function $\log P_L(N)$ is being approximated by a suitable rational function $r_n(N)$;
– and the approximate model is validated by simulation for one larger value of the buffer length.

The motivation to compute the packet loss probabilities using rational approximation comes from the works of Gong *et al.* [10] and Yang [24]. They compute these probabilities for large buffer sizes, from sampled values of $\log P_L(N)$ for small buffer sizes and its decay rate. The technique was at first applied to multiplexer models with little or no correlation between the cells.

In [4], an automatic procedure to select the sample points (also called support points) is proposed and used for the efficient computation of models $r_n(N)$ in case there is more correlation between the cells. The procedure selects the support points in a region which we determine from the system parameters, until the model $r_n(N)$ is sufficiently accurate, meaning that $|r_n(N) - r_{n+1}(N)|/|r_{n+1}(N)|$ does not exceed a prescribed error threshold. But when encountering positive real poles in the model $r_n(N)$, the procedure has to add more support points and increase n to achieve the desired accuracy.

In [3], the technique of [4] is perfected by making very good use of the knowledge of $P_L(N)$ for extremely small values of N. The approach maximizes the information that can be extracted from the data, while minimizing the number of date samples to be collected. Extremely small is to be interpreted as $N = 1, 2, 3, \ldots, M$, where M is the number of sources. The modified technique allows to construct a model $r_n(N)$ which is free from positive real poles, and is extremely efficient.

Since the parameters used in such networks are measured values based on the application requirements, precise probabilities may not be known. Hence it is interesting and important to carry out a perturbation analysis with respect to these parameters. Instead of taking point values for some real-valued parameters, we will consider intervals and carry out the computation of the PLP for interval data. First, we compute the pair of packet loss probabilities for small buffer sizes using interval arithmetic corresponding to the minimal and maximal loads of the networks. These minimal and maximal loads are computed from the interval parameter values using theorem 5.1. We also compute an interval enclosure for the pair of decay rates corresponding to the minimal and maximal loads. Second, we compute the Newton-Padé type approximant using the technique developed in [3] in combination with the method of Markov *et al.* [14] for the computation of the interval coefficients of the numerator polynomial. This is done by using the interval packet loss probabilities at small buffer sizes as support points and the interval decay rate.

2 Analytical Model

2.1 Notation and definitions

In a multiplexer environment, the arrival of packets to the switch happens in discrete time, with discrete service time, which makes the discrete time Markov chain a natural modelling choice. We assume that the arrival of packets transmitted by M independent and non-identical information sources to the multiplexer, can be modelled as a discrete time batch Markovian Arrival Process (D-BMAP), the discrete-time version of BMAP. The BMAP is a convenient representation of the versatile Markovian point process which generalizes the Markovian arrival process (MAP) [15]. The D-BMAP is a general process used to model a number of arrival processes, for example video [5] and periodic processes [9].

Each information source is controlled by a Markov chain, called the background Markov chain. The basic queueing system which models the multiplexer is a D-BMAP/D/c/N queue with c discrete time servers, where each server can serve at most one cell per time unit. These servers serve a buffer with a capacity of N cells which is fed by M independent information sources. When the server is busy, a maximum of c cells will depart in each slot. Service starts at the beginning of each time slot.

The arrival process associated with a single source is modelled as an Interrupted Bernoulli Process (IBP). This process has two states 0 and 1. Source i generates a packet with probability $d_i(m)$ when it is in state m $(m = 0, 1)$ and its transition probability matrix is given by

$$\boldsymbol{Q_i} = \begin{pmatrix} 1 - p_i & p_i \\ q_i & 1 - q_i \end{pmatrix}. \tag{1}$$

The D-BMAP queueing model is an M/G/1 type queue which is basically a two-dimensional discrete time Markov chain $\{(X_n, Y_n), n \geq 0\}$, where X_n is the number of cells in the buffer and Y_n represents the state of the M sources during the nth time slot. We are interested in the steady state behavior $(X, Y) := \lim_{n\to\infty} (X_n, Y_n)$.

The state space of the process (X, Y) is $S := S_X \times S_Y$, where

$$S_X = \{0, 1, 2, \ldots, N\}, \quad S_Y = \{(m_1, m_2, \ldots, m_M) \mid m_i = 0 \text{ or } 1\}.$$

A significant reduction of the state space S_Y is possible when the sources are homogeneous (identical). We discuss this case in subsection 2.2.

The $2^M \times 2^M$ matrix $\boldsymbol{D}$, the transition probability matrix of the process Y, is given by

$$\boldsymbol{D} = \bigotimes_{i=1}^{M} \boldsymbol{Q_i}. \tag{2}$$

Since we assume that each source can generate at most one cell during a time slot and there are M sources, at most M cells can arrive at the multiplexer during a time slot. Therefore, there are $M + 1$ matrices governing the arrivals, namely $\boldsymbol{D_0}, \boldsymbol{D_1}, \ldots, \boldsymbol{D_M}$. Then by [23] the probability matrix $\boldsymbol{D_M}$ of m arrivals during a time slot is given by

$$\boldsymbol{D_m} = \sum_{\substack{j_1, j_2, \ldots, j_M \\ j_i = 0 \text{ or } 1,\ 1 \leq i \leq M \\ j_1 + j_2 + \cdots + j_M = m}} \bigotimes_{i=1}^{M} \left[(1 - j_i)\boldsymbol{I} + (-1)^{(1-j_i)} \boldsymbol{P_i}\right] \boldsymbol{Q_i}, \tag{3}$$

where

$$\boldsymbol{P_i} = \begin{pmatrix} d_i(0) & 0 \\ 0 & d_i(1) \end{pmatrix}, \quad i = 1, 2, \ldots, M \tag{4}$$

and $\boldsymbol{I}$ is the identity matrix of order 2×2. The dimension of the matrix $\boldsymbol{D_m}$ is $2^M \times 2^M$.

The average arrival rate of the packets at the multiplexer is given by

$$\lambda = \sum_{i=1}^{M} \frac{p_i d_i(1) + q_i d_i(0)}{c(p_i + q_i)} \tag{5}$$

The sources are said to be of ON-OFF type with state 0 corresponding to the OFF state and state 1 corresponding to the ON state if $d_i(0) = 0,\ \forall i$. That is, when a source is in OFF state, it does not transmit any cell. In this case we denote $d_i(1)$ by d_i for $i = 1, 2, \ldots, M$.

The load (traffic intensity) of the network $\rho := \lambda / c$. Under the condition of ergodicity ($\rho < 1$) of the chain (X, Y), the stationary distribution vector $\boldsymbol{\Pi} := \{\boldsymbol{\pi_0}, \boldsymbol{\pi_1}, \ldots, \boldsymbol{\pi_N}\}$ where $\boldsymbol{\pi_i} \in {}^{2^M}$, satisfies

$$\boldsymbol{\Pi P} = \boldsymbol{\Pi} \tag{6}$$

$$\boldsymbol{\Pi e} = 1 \tag{7}$$

where the transition probability matrix $\boldsymbol{P}$ [11] of the process (X, Y) is given by

$$\boldsymbol{P} = \begin{pmatrix} \boldsymbol{D_0} & \boldsymbol{D_1} & \dots & \boldsymbol{D_{N-c}} & \dots & \boldsymbol{D_{N-1}} & \boldsymbol{B_N} \\ \boldsymbol{D_0} & \boldsymbol{D_1} & \dots & \boldsymbol{D_{N-c}} & \dots & \boldsymbol{D_{N-1}} & \boldsymbol{B_N} \\ \vdots & & & & & & \vdots \\ \boldsymbol{D_0} & \boldsymbol{D_1} & \dots & \boldsymbol{D_{N-c}} & \dots & \boldsymbol{D_{N-1}} & \boldsymbol{B_N} \\ \boldsymbol{0} & \boldsymbol{D_0} & \dots & \boldsymbol{D_{N-c-1}} & \dots & \boldsymbol{D_{N-2}} & \boldsymbol{B_{N-1}} \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{D_{N-c-2}} & \dots & \boldsymbol{D_{N-3}} & \boldsymbol{B_{N-2}} \\ \vdots & & & & & & \vdots \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{D_0} & \dots & \boldsymbol{D_{c-1}} & \boldsymbol{B_c} \end{pmatrix}_{(N+1)2^M \times (N+1)2^M} \tag{8}$$

with

$$\boldsymbol{B_n} := \sum_{j=n}^{M} \boldsymbol{D_j} \tag{9}$$

and $\boldsymbol{e}$ is a column vector of ones. The complexity of matrix-analytic methods to find the stationary probabilities $\boldsymbol{\Pi}$ is $O(c2^{3M}N^2)$ (see for example [2, 11]).
The packet loss probability function, as a function of the buffer size N, is then given by

$$P_L(N) = 1 - \frac{\text{average number of served packets}}{\text{average number of arrived packets}} = 1 - \frac{1}{\lambda}\sum_{i=0}^{N} \min(i, c)\boldsymbol{\pi_i e}. \tag{10}$$

2.2 Particular case

Suppose all the sources are homogeneous (identical) and of ON-OFF type. Then

$$p_i = p,\ q_i = q,\ d_i(0) = 0,\ d_i(1) = d,$$

where p is the probability of an ON source becoming OFF, q is the probability of an OFF source becoming ON, and d is the probability of an ON source transmitting a cell.

The average arrival rate for this case can easily be verified as

$$\lambda = \frac{Mpd}{c(p+q)}.$$

For this case the state space

$$S_Y = \{0, 1, 2, \dots, M\},$$

where $i \in S_Y$ denotes the number of ON sources. The state space S_X remains the same.

For example, consider a population of voice sessions serviced by a single T1 channel (1.536 Mbps). We discretize time into 16ms slots and model each voice source as an ON-OFF source controlled by a 2-state Markov chain. The service capacity of the channel is $c = 48$, corresponding to each voice source generating cells at a peak rate of 32Kbps ($= 1.536$Mbps$/48$). Suppose the peak ON and OFF periods are 352ms and 650ms, respectively and a voice packet will be transmitted if the source is in ON state. Then the probabilities p and q are given by

$$p = \frac{16}{650} \approx 0.02462 \text{ and } q = \frac{16}{352} \approx 0.04546$$

The packet generation probability $d = 1$. With the presence of $M = 100$ sources, the average arrival rate $\lambda = 35.1297$.

The (i, j)-th element d_{ij} of the transition probability matrix $\boldsymbol{D}$ of Y is given by

$$d_{ij} = \sum_{k=0}^{i} \binom{i}{k} \binom{M-i}{k+j-i} q^k (1-q)^{i-k} p^{j+k-i} (1-p)^{M-j-k}. \tag{11}$$

When the parameters p and q are very small (more correlation between the arriving cells), then the d_{ij} in (11) can be approximated by the following formula:

$$d_{ij} = \begin{cases} 1-(M-i)p-iq, & \text{if } j=i \\ (M-i)p, & \text{if } j=i+1 \\ iq, & \text{if } j=i-1 \end{cases} \tag{12}$$

That is, the d_{ij} are one-step transition probabilities and the matrix $\boldsymbol{D}$ corresponds to the transition probability matrix of a birth-death process with birth rate $(M-i)p$ and death rate iq when the process is in state i.

The matrices $\boldsymbol{D_m}$ are given by

$$\boldsymbol{D_m} = \text{diag}(c_m(0), c_m(1), \ldots, c_m(M))\boldsymbol{D}, \quad m = 0, 1, \ldots, M, \tag{13}$$

where $c_m(k)$, the probability of m arrivals during a time slot when the process Y is in state k, is given by

$$c_m(k) = \begin{cases} \binom{k}{m} d^m(1-d)^{k-m}, & \text{if } d \neq 1 \\ \delta_{mk}, & \text{if } d = 1 \end{cases} \tag{14}$$

The (i,j)-th element of $\boldsymbol{D_m}$ equals the probability of m arrivals at the buffer during a time slot when the background Markov chain changes from state i to j.

The formulae to compute $\boldsymbol{P}$ and $P_L(N)$ remain the same, namely (8) and (10), respectively. The matrix $\boldsymbol{P}$ is now a square matrix of order $(N+1)(M+1)$.

2.3 Asymptotic behaviour of $\log P_L(N)$

It has been proved that for infinite M/G/1-type queues, the buffer overflow probability decays exponentially [8]. In [13], the authors have shown that for Markov modulated queueing models with multi-server and infinite buffer, the queue length distribution has exponential bounds. In [1], the exponential decay of the loss probability of the finite MAP/G/1/K queue is studied. The asymptotic decay rate in a finite buffer queueing system is the same as in the equivalent infinite buffer system [18].

We now briefly discuss the approach to compute the decay rate from the knowledge of the parameters for a given model. We first show how we arrange the blocks in the matrix $\boldsymbol{P}$ for the multi-server case so that the structure of $\boldsymbol{P}$ is similar to that of a finite M/G/1-type Markov chain.
Define for $i = 0, 1, \ldots, K$,

$$\boldsymbol{A_i} := \begin{pmatrix} \boldsymbol{D_{i\times c}} & \boldsymbol{D_{i\times c+1}} & \cdots & \boldsymbol{D_{i\times c+c-1}} \\ \boldsymbol{D_{i\times c-1}} & \boldsymbol{D_{i\times c}} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \boldsymbol{D_{i\times c+1}} \\ \boldsymbol{D_{i\times c-c+1}} & \cdots & \boldsymbol{D_{i\times c-1}} & \boldsymbol{D_{i\times c}} \end{pmatrix}$$

with $\boldsymbol{D_k} = \boldsymbol{0}$ for $k < 0$ and $K = \lceil M/c \rceil$ ($\lceil 2^M/c \rceil$) for homogeneous (heterogeneous) sources. The matrix $\boldsymbol{A_i}$ is a square matrix of size $c(M+1)$ if the sources are homogeneous and size $2^M c$ if the sources are heterogeneous. If $c = 1$, then $\boldsymbol{A_i} = \boldsymbol{D_i}$ for $i = 0, 1, \ldots, K$.
Define

$$\boldsymbol{A}(z) := \sum_{n=0}^{K} \boldsymbol{A_n} z^n, \quad 0 < z < R_A, \tag{15}$$

where R_A is the radius of convergence of $\boldsymbol{A}(z)$. Then for $z \in]1, R_A[$, the exponential decay rate ξ is the Perron-Frobenius eigenvalue of $\boldsymbol{A}(z)$ satisfying the condition $\xi = z$ [8]. Since $P_L(N)$ decays exponentially with decay rate ξ, we have as $N \to \infty$,

$$\log P_L(N) \sim \xi N. \tag{16}$$

3 Rational approximation without pole placement

Because of the fact that the function $\log P_L(N)$ asymptotically behaves as ξN for large N, polynomial approximation techniques for $\log P_L(N)$ are not suitable. Every polynomial model of degree larger than one, would blow up for large N. However, a rational function $r_n(N)$ of numerator degree $n+1$ and denominator degree n,

$$r_n(N) = \frac{p_n(N)}{q_n(N)} = \frac{\sum_{i=0}^{n+1} a_i N^i}{\sum_{i=0}^{n} b_i N^i}, \tag{17}$$

has a similar asymptotic behavior as that of $\log P_L(N)$. Remains to compute the coefficients in numerator and denominator of the rational function from sampled function values $\log P_L(N_j)$ for chosen N_j and to fit its asymptotic behaviour to ξ. A rational approximant $r_n(N)$ can be obtained as the $2n^{th}$ convergent of a so-called Thiele type continued fraction [19]:

$$\begin{aligned} r_n(N) &= \varphi[N_0] + \sum_{j=0}^{2n} \left. \frac{N-N_j}{\varphi[N_0,\ldots,N_{j+1}]} \right| \\ &= \varphi[N_0] + \cfrac{N-N_0}{\varphi[N_0,N_1] + \cfrac{N-N_1}{\varphi[N_0,N_1,N_2] + \cfrac{N-N_2}{\cdots}}}, \end{aligned}$$

where the inverse differences $\varphi[N_0,\ldots,N_{j+1}]$ are computed recursively from

$$\begin{aligned} \varphi[N_j] &= \log P_L(N_j) \\ \varphi[N_0,\ldots,N_{j+1}] &= \frac{N_{j+1}-N_j}{\varphi[N_0,\ldots,N_{j-1},N_{j+1}] - \varphi[N_0,\ldots,N_{j-1},N_j]}. \end{aligned} \tag{18}$$

The rational function $r_n(N)$ is sometimes also known as a Newton-Padé or multipoint Padé approximant.

In order to fit the asymptotic behavior of $r_n(N)$ to that of $\log P_L(N)$, we only compute $\varphi[N_0,\ldots,N_{j+1}]$ with $j = 0,\ldots,2n-1$ from (18). The last inverse difference $\varphi[N_0,\ldots,N_{2n+1}]$ is computed from the following property. The coefficient of highest degree in the numerator of $r_n(N)$, namely a_{n+1} equals

$$a_{n+1} = \frac{1}{\sum_{j=0}^{n} \varphi[N_0,\ldots,N_{2j+1}]}, \quad b_n = 1.$$

For $r_n(N)$ to behave asymptotically like ξN, we need to require $a_{n+1}/b_n = \xi$ or in other words

$$\varphi[N_0,\ldots,N_{2n+1}] = \frac{1}{\xi} - \sum_{j=0}^{n-1} \varphi[x_0,\ldots,x_{2j+1}].$$

The accuracy of the model $r_n(N)$ is assessed by looking at

$$||r_n(N) - r_{n+1}(N)|| = \sup_{N\in} |r_n(N) - r_{n+1}(N)|$$

which tends to zero if $r_n(N)$ converges to $\log P_L(N)$ [21].

From the knowledge of the transition probabilities of the background Markov chain, the load, the decay rate and the number of servers of a given system, we determine a range $[K, L]$ for N, in which the function $\log P_L(N)$ switches from a fast decreasing to a slowly decreasing function. The initial support points N_j are chosen to satisfy $K \le N_j \le L$ as detailed in the flowgraphs (Figures 1, 2) given below. We use `supp` as an abbreviation for the set of support points.

Successive support points are added in the following way. A discrete approximation

$$\max_{N=K,\ldots,L} |r_n(N) - r_{n+1}(N)|, \qquad r_0(N) = (\log P_L(N_0) - \xi N_0) + \xi N$$

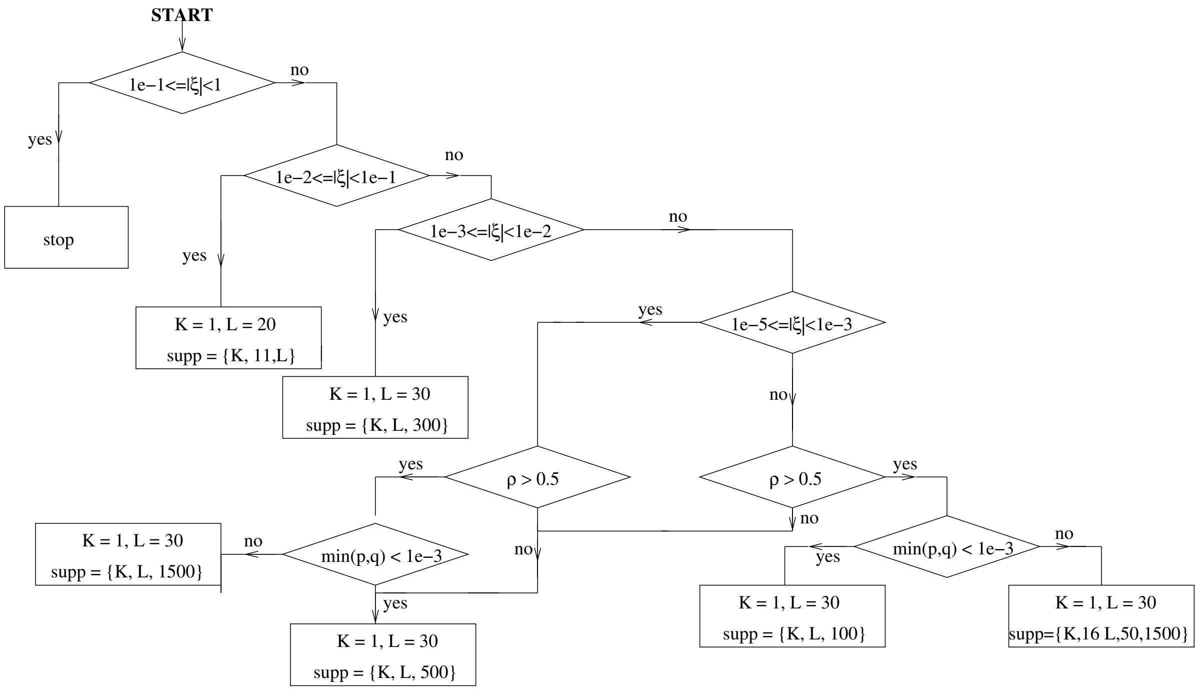

Fig. 1 Strategy for networks with heterogeneous sources.

of $||r_n(N) - r_{n+1}(N)||$ is computed. The values of N in $[K, L]$ for which the maximum and the second largest value are attained are chosen to be the next two support points.

To compare the model $r_n(N)$ to the true $\log P_L(N)$ in the Figures 3(a)-3(d), the latter is computed using the algorithm from [11].

In all figures, the values obtained at support points are circled, the computed function $\log P_L(N)$ is graphed using a full line and the approximation $r_n(N)$ is graphed using a dotted line. An additional simulation point, used merely for validation, is denoted by a $\star$. When only the full line is visible, this means that on the displayed figure the approximation and the function $\log P_L(N)$ are graphically indistinguishable. In Table 1 the parameters of the networks and the corresponding figures are tabulated, where for Figure 3(d)

$$\boldsymbol{p} = \begin{pmatrix} 6.984e-5 \\ 2.1e-7 \\ 8.366e-5 \\ 8.8894e-5 \\ 1.98e-6 \end{pmatrix}, \qquad \boldsymbol{q} = \begin{pmatrix} 9.84e-6 \\ 3.747e-5 \\ 9.675e-5 \\ 6.196e-5 \\ 6.7e-5 \end{pmatrix} \text{ and } \boldsymbol{d} = \begin{pmatrix} 0.4562 & 0.2953 \\ 0.8380 & 0.6022 \\ 0.8231 & 0.1828 \\ 0.5421 & 0.7332 \\ 0.0924 & 0.5489 \end{pmatrix}. \tag{19}$$

The displayed rational approximants satisfy

$$||r_n(N) - r_{n+1}(N)|| \leq \epsilon ||r_{n+1}(N)||$$

with $\epsilon = 0.01$.

4 Rational approximation with optimal pole placement

4.1 Problem of undesirable poles

Functions with poles or at most a countable number of isolated essential singularities, in particular, allow nice convergence properties when approximated by rational functions. In that case, the singularities of the function under consideration, attract the poles of the rational approximant to their position. Although the behaviour of the

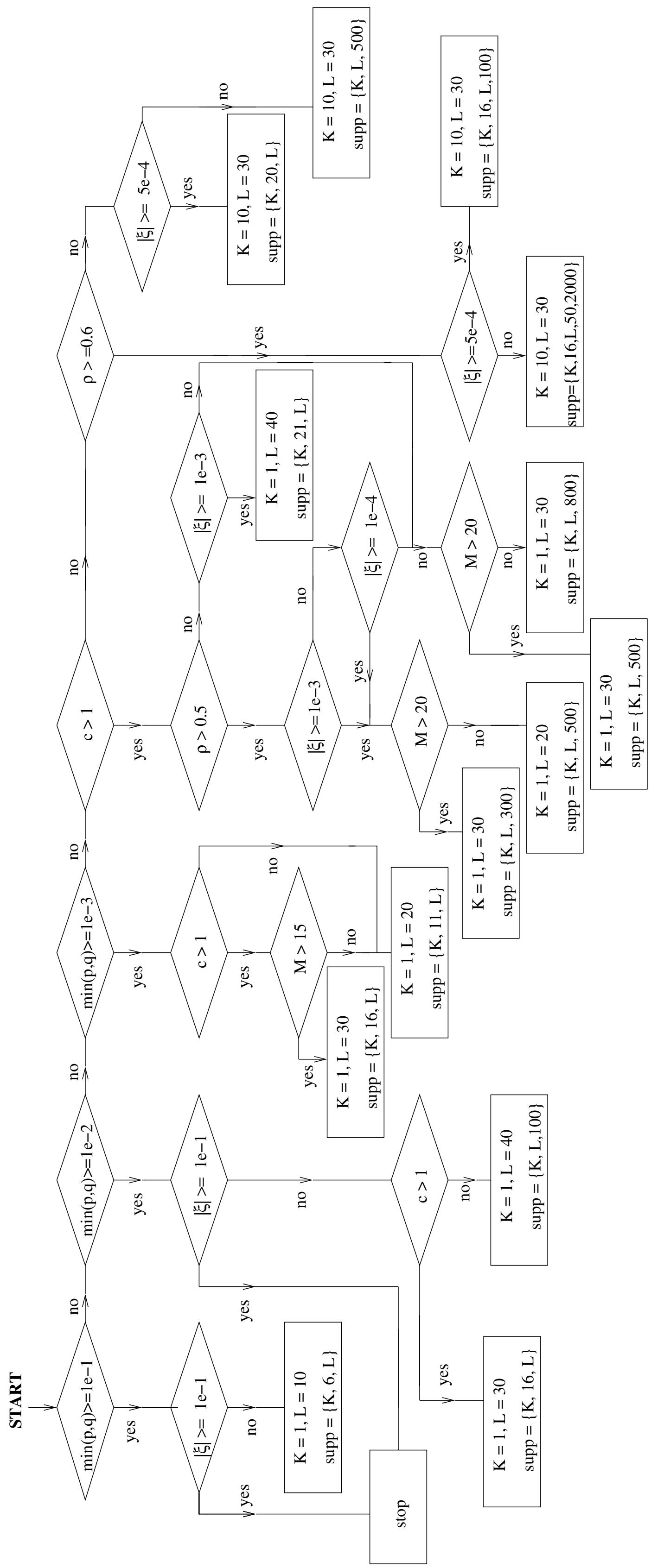

Fig. 2 Strategy for networks with homogeneous sources

Table 1 Examples from section 3

Figure	M	c	p	q	d	ρ	ξ	approximant
3(a)	30	1	$2.19e-5$	$4.3e-5$	$7.6923e-2$	0.7787	$-1.43e-4$	r_9
3(b)	25	15	$2.5e-3$	$1.15e-3$	$6.5e-1$	0.742	$-7.6923e-4$	r_{12}
3(c)	15	9	$2.19e-4$	$5.5e-6$	$6.0e-1$	0.9755	$-2.628e-3$	r_7
3(d)	5	3	(19)	(19)	(19)	0.8128	$-2.271e-4$	r_6

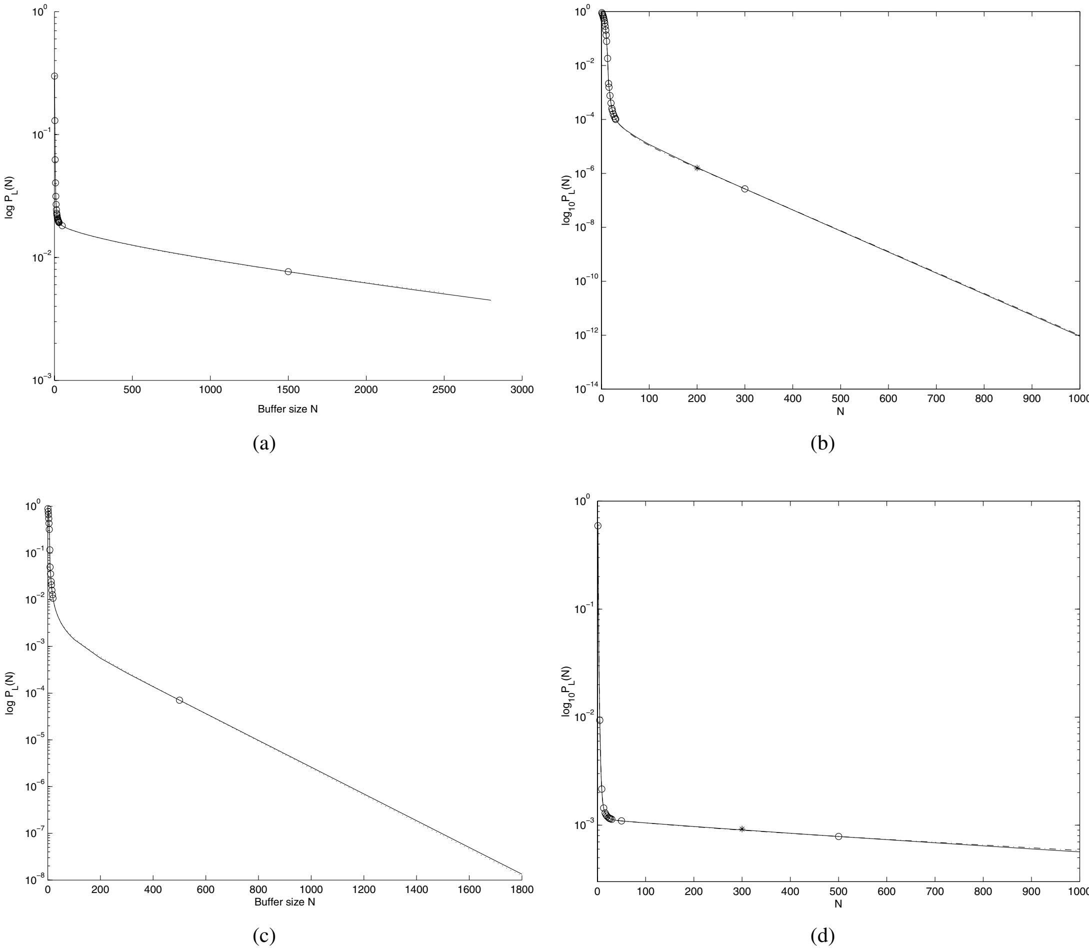

Fig. 3 Packet loss probabilities without optimal pole placement of r_n

function $\log P_L(N)$ is not such that a rational approximant $r_n(N)$ "naturally" attracts real positive poles, once in a while it may happen that $r_n(N)$ has one or more poles in the region of interest for N. For instance, with the system parameters given by $M = 30,\ p = 2.19e-5, q = 4.3e-5, d = 7.6923e-2,\ c = 1$ the approximant r_8, which can be seen in Figure 4, exhibits a pole around $N \approx 100$. This is of course undesirable. Since a suitable value for the denominator degree n is determined from comparing $||r_{n+1} - r_n||/||r_{n+1}||$ on the positive real axis to a threshold value ϵ,

$$\sup_{0<N<\infty} |r_{n+1}(N) - r_n(N)| \leq \epsilon \sup_{0<N<\infty} |r_{n+1}(N)|, \tag{20}$$

the occurrence of an undesirable pole slows the method down. Fortunately, it does not make the method unsuitable. Indeed, when $\log P_L(N)$ itself does not have any singularities on the positive real axis, while $r_n(N)$ has a

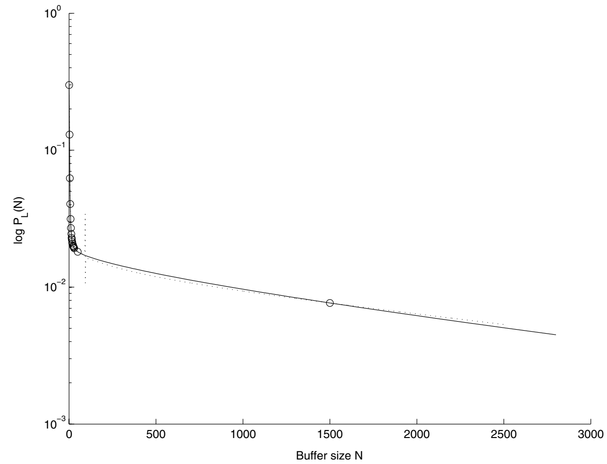

Fig. 4 $\log P_L(N)$

pole at $N = N^\star > 0$, then there is no mathematical argument for the next approximant $r_{n+1}(N)$ to have a pole in the neighbourhood of $N = N^\star$ as well. While (20) is not satisfied, it may be satisfied for a larger value of n. Let us now explain how the occurrence of these undesirable poles can be avoided.

4.2 Curvature maximum of $\log P_L(N)$

Besides the typical asymptotic behaviour of $\log P_L(N)$, we also want to "grab" the neighbourhood of N for which the graph turns from a steep descent towards its linear asymptotic look. In the neighbourhood of the transition from burst region to cell region, the curvature of $\log P_L(N)$ attains its maximal value. The curvature of a function $f(x)$ is given by

$$\kappa(x) = \frac{f''(x)}{\left(\sqrt{1 + (f'(x))^2}\right)^3}.$$

Here $'$ and $''$ denote the first and second derivatives, respectively. A discretised version of $\kappa(x)$, which can also be used for $f(N) = \log P_L(N)$, is given by

$$\kappa_j = \frac{(f_{j+1} - 2f_j + f_{j-1})}{\left(\sqrt{1 + (f_{j+1} - f_j)^2}\right)^3}, \qquad f_j = \log P_L(j+1), \qquad j \geq 1. \tag{21}$$

We are interested in the point of maximal curvature of the function $\log P_L(N)$, which can be estimated by monitoring κ_j for successive values of j. Let us denote by $\tilde{N}$ the value of $j+1$ for which κ_j attains its maximum. That the computation of $\tilde{N}$ can be put to good use, becomes clear from the following observations.

A rational function of the form $1/(N^2 + R^2)$ with $R > 0$ "resembles" $\log P_L(N)$ on the positive real axis. It has its maximal curvature for positive N in the immediate neighbourhood of the point $N = R$, namely in the interval $[R, 1.002R]$ if $R \geq 2$. A typical evolution of the curvature of $\log P_L(N)$ and that of $1/(N^2 + R^2)$ can be found in Figure 5. A rational function of the form

$$r_2(N) = \frac{\xi N^3 + a_2 N^2 + a_1 N + a_0}{N^2 + R^2}$$

also exhibits a limiting behaviour of the required type, namely $\lim_{N\to\infty} r_2(N) \approx \xi N$. And it achieves maximal curvature in the neighbourhood of $N = R$.

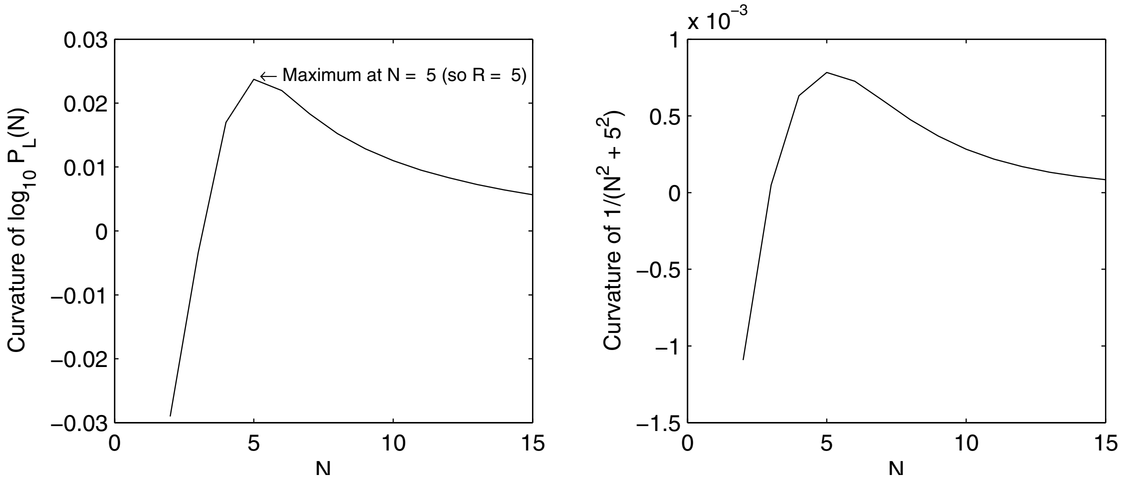

Fig. 5 Curvature of $\log P_L(N)$ and of $\frac{1}{N^2+R^2}$

In all numerical experiments we observed that the maximum of the curvature of $\log P_L(N)$ is attained in the interval $[1, M]$, where M is the number of sources. That is, either the curvature increases from $N = 1$ on until it reaches its global maximum and then decreases, or, for some networks, the curvature function is a decreasing function in the interval $[1, M]$ and then we return $N = 2$ as the argument of the maximum. Since M is relatively speaking rather small, the computation of $R = \tilde{N}$ is almost negligible. The detailed algorithm for the estimation of R goes as follows:

1. compute $\log P_L(j + 1)$ for $j = 0, 1, 2, 3$
2. compute κ_j for $j = 1, 2$
3. set $i = 2$
4. while $((\kappa_i - \kappa_{i-1} > 0)$ or $(i < M)$)
   compute $\log P_L(i + 3)$
   set $i = i + 1$
   compute κ_i
   end
5. if $(i < M)$ then $R = i$ else $R = 2$

4.3 Optimally placed poles

So, when detecting an approximation N_j of the point of maximal curvature of $\log P_L(N)$, through the computation of (21), we can choose $R = N_j$ in $r_2(N)$ and hence introduce two complex conjugate poles, thereby preventing the occurrence of real poles in $r_2(N)$. Remains to point out which strategy can be followed for larger denominator degrees n.

A rational function $r_n(N)$ with denominator polynomial of the form

$$q_n(N) = \begin{cases} (N^2 + R^2) \displaystyle\prod_{j=1}^{k-1} \left(N - Re^{\imath\theta_j}\right)\left(N - Re^{-\imath\theta_j}\right), & n = 2k \\ (N + R)(N^2 + R^2) \displaystyle\prod_{j=1}^{k-1} \left(N - Re^{\imath\theta_j}\right)\left(N - Re^{-\imath\theta_j}\right), & n = 2k + 1 \end{cases} \tag{22}$$

has its point of maximal curvature in the neighbourhood of $N = R$ only if $\theta_j \approx \pi/2$. So, when increasing the denominator degree of $r_n(N)$, more complex conjugate poles of modulus R can be prescribed, by choosing different $\theta_j \approx \pi/2$. Complex conjugate poles further away from the imaginary axis "pull" the point of maximal curvature away from $N = R$.

So far we have explained how to use the curvature of $\log P_L(N)$ to fix the coefficients $b_0, \ldots, b_n$ in the rational model. The coefficient a_{n+1} of the numerator polynomial is determined from the asymptotic behaviour

$$\lim_{N\to\infty} \log P_L(N) \approx \xi N = \frac{a_{n+1}}{b_n} N, \tag{23}$$

where the usual normalization $b_0 = 1$ is replaced by $b_n = 1$, an equally simple choice. The remaining coefficients, being the numerator coefficients $a_0, \ldots, a_n$, can be computed from the polynomial data fitting conditions

$$(q_n \log P_L)(N_j) = p_n(N_j) \qquad j = 0, \ldots, n \tag{24}$$

which, under the condition that $q_n(N_j) \neq 0$, are equivalent to

$$r_n(N_j) = \log P_L(N_j) \qquad j = 0, \ldots, n.$$

The values to be chosen for N_j are detailed in section 3. This technique is called Newton-Padé type or multipoint Padé-type approximation [20]. It differs from the standard multipoint Padé approximation because the denominator $q_n(N)$ is not determined by the interpolation conditions but is prechosen.

The convergence results obtained in [20] underline that:

1. the rational approximants $r_n(N)$ need to be uniformly bounded on bounded subsets of the region of interest (here the natural numbers);

2. the interpolation points N_j cannot be scattered around but must be centered in one location (as described in section 3.

We reconsider the examples discussed in section 3 which are illustrated in the Figures 3. The modified rational approximants with preassigned poles are shown in the Figures 6. Besides being pole free, they also require less support points. The numerical results are compared in Table 2.

Table 2 Examples from section 4

Figures	# support points		Optimal $q_n(N)$
	Figure 3	Figure 6	
6(a)	19	5	$(N^2+25)\left(N+5e^{(\pi/2+\pi/12)\imath}\right)\left(N+5e^{-(\pi/2+\pi/12)\imath}\right)$
6(b)	25	3	N^2+225
6(c)	15	4	$(N^2+81)(N+9)$
6(d)	13	4	$(N^2+9)(N+3)$

5 Uncertainty analysis and interval arithmetic

5.1 Uncertainty in the parameters

For the type of networks under investigation, the following set of parameters completely determines the behaviour of a network:

$(M,\ c,\ p,\ q,\ d)$ if the sources are homogeneous,

$(M,\ c,\ \boldsymbol{p},\ \boldsymbol{q},\ \boldsymbol{d})$ if the sources are heterogeneous.

The parameters M and c take positive integer values and the rest take real values. So, we investigate the effect of uncertainties in the parameters $\boldsymbol{p}$ (or p), $\boldsymbol{q}$ (or q) and $\boldsymbol{d}$ (or d) on the PLP on one hand and/or uncertainties in

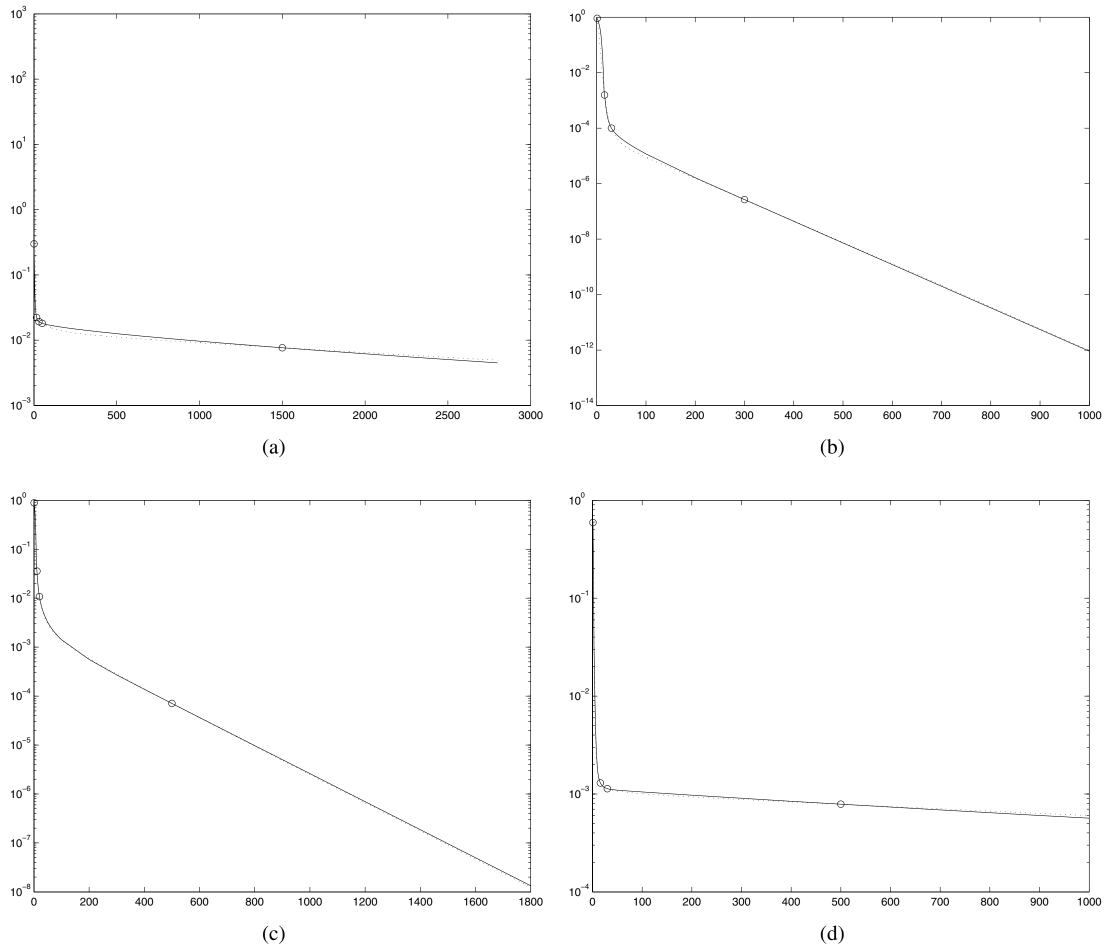

Fig. 6 Packet loss probabilities with optimal pole placement of r_n

the values of $\log P_L(N_j)$ when coming from measurements or verified computations. Assume interval values for $p,\ q$ and d (or $\boldsymbol{p}$, $\boldsymbol{q}$ and $\boldsymbol{d}$).

To automatically incorporate this perturbation analysis in the computation of the rational fitting technique, we use tools from interval arithmetic. The verified computation of $r_n(N)$ requires enclosures for or the verified computation of:

- $\boldsymbol{p}$ (or p), $\boldsymbol{q}$ (or q) and $\boldsymbol{d}$ (or d) which are involved in the matrices $\boldsymbol{D_m}$,
- the samples $\log P_L(N_j)$,
- and the asymptotic slope ξ [16].

One of the important finds is that $P_L(N)$ itself is rather sensitive with respect to small changes in the parameters. This of course has its effect on the data fitting problem. When using a rational model with prescribed poles to fit the interval data, as in the multipoint Padé-type approach, the verified computation of $r_n(N)$ is carried out as in [14].

We denote by $[r]$, an interval enclosure for a real value r and we denote by $\inf{[r]}, \sup{[r]}$ and mid$[r]$ its lower bounds, upper bounds and midpoint, respectively.

When considering interval ranges $[\boldsymbol{p}]$, $[\boldsymbol{q}]$ and $[\boldsymbol{d}]$ for the parameters $\boldsymbol{p}$, $\boldsymbol{q}$ and $\boldsymbol{d}$, the load of the network under consideration varies. The following result expresses $[\rho]$ in terms of $[\boldsymbol{p}]$, $[\boldsymbol{q}]$ and $[\boldsymbol{d}]$.

Theorem 5.1 *For a network with parameters $M,\ c$, $[\boldsymbol{p}]$, $[\boldsymbol{q}]$ $[\boldsymbol{d}]$, the load varies between $\check{\rho}$ and $\hat{\rho}$ given by*

$$\check{\rho} = \begin{cases} \frac{1}{c}\sum_{i=1}^{M}\dfrac{\check{d}_i(1)\hat{p}_i + \check{d}_i(0)\check{q}_i}{\hat{p}_i + \check{q}_i}, & \text{if } \check{d}_i(0) > \check{d}_i(1) \\ \frac{1}{c}\sum_{i=1}^{M}\dfrac{\check{d}_i(1)\check{p}_i + \check{d}_i(0)\hat{q}_i}{\check{p}_i + \hat{q}_i}, & \text{if } \check{d}_i(0) \le \check{d}_i(1) \end{cases} \tag{25}$$

$$\hat{\rho} = \begin{cases} \frac{1}{c}\sum_{i=1}^{M}\dfrac{\hat{d}_i(1)\check{p}_i + \hat{d}_i(0)\hat{q}_i}{\check{p}_i + \hat{q}_i}, & \text{if } \hat{d}_i(0) > \hat{d}_i(1) \\ \frac{1}{c}\sum_{i=1}^{M}\dfrac{\hat{d}_i(1)\hat{p}_i + \hat{d}_i(0)\check{q}_i}{\hat{p}_i + \check{q}_i}, & \text{if } \hat{d}_i(0) \le \hat{d}_i(1) \end{cases} \tag{26}$$

Proof. We derive only the expression for the minimal load as the derivation is completely similar for the maximal load.

Let us consider the source i. The probability ranges corresponding to this source are $[p_i]$, $[q_i]$, $[d_i(0)]$ and $[d_i(1)]$. Since $[d_i(0)]$ and $[d_i(1)]$ are the transmitting probabilities, clearly the minimum load is achieved for $\check{d}_i(0)$ and $\check{d}_i(1)$.

If $\check{d}_i(0) > \check{d}_i(1)$, then the probability for source i to transmit a cell is smaller in state 1 than in state 0, resulting in less load. So by letting source i stay in state 1 for a longer period, which is achieved with $\hat{p}_i$ and $\check{q}_i$, the load diminishes to

$$\check{\rho} = \frac{1}{c}\sum_{i=1}^{M}\frac{\check{d}_i(1)\hat{p}_i + \check{d}_i(0)\check{q}_i}{\hat{p}_i + \check{q}_i}, \qquad \check{d}_i(0) > \check{d}_i(1).$$

If $\check{d}_i(0) < \check{d}_i(1)$, then the probability for source i to transmit a cell is smaller in state 0 than in state 1, resulting in less load in the network. So by letting source i stay in state 0 for a longer period, which is achieved with $\check{p}_i$ and $\hat{q}_i$, the load grows to

$$\hat{\rho} = \frac{1}{c}\sum_{i=1}^{M}\frac{\check{d}_i(1)\check{p}_i + \check{d}_i(0)\hat{q}_i}{\check{p}_i + \hat{q}_i}, \qquad \check{d}_i(0) \le \check{d}_i(1).$$

If $\check{d}_i(0) = \check{d}_i(1)$, (5) becomes independent of the probabilities $\boldsymbol{p}$ and $\boldsymbol{q}$. □

Corollary 5.2 *If the sources are homogeneous and of ON-OFF type, the minimal and maximal loads are given by*

$$\check{\rho} = \frac{M\check{p}\check{d}}{c(\check{p} + \hat{q})} \qquad \textit{and} \qquad \hat{\rho} = \frac{M\hat{p}\hat{d}}{c(\hat{p} + \check{q})}. \tag{27}$$

In the forthcoming analysis we refer to the parameters determining $\check{\rho}$ and $\hat{\rho}$ as minimal and maximal parameters, respectively. We refer to $P_L(N)$ which corresponds to the minimal and maximal parameters, as the lower bound and upper bound loss probabilities $\check{P}_L(N)$ and $\hat{P}_L(N)$, respectively.

To compute the rational approximant for the packet loss probability function, in the context of uncertainty, we need to determine reliable enclosures for $\log P_L(N)$ at some support points N_j. This can be done by computing interval enclosures $[\check{P}_L(N_j)]$ and $[\hat{P}_L(N_j)]$ for the lower bound and upper bound loss probabilities respectively, using the matrix-analytic method given in [2, 11] and the interval technique described in [17]. In the Figures 7, $[\log \check{P}_L(N)]$ and $[\log \hat{P}_L(N)]$ are plotted using a full line, for as large N as computationally feasible, which is more than needed to collect the samples and is only done for comparison with the rational interval model of which the computation is detailed below. The enclosing intervals are so sharp that the infimum and the supremum of both $[\log \check{P}_L(N)]$ and $[\log \hat{P}_L(N)]$ are visually indistinguishable. The top full line corresponds to $[\log \hat{P}_L(N)]$

while the bottom full line depicts $[\log \check{P}_L(N)]$. The sharpness of the interval enclosures is a result of the iterative validation technique for linear systems of equations as detailed in [17].

For the interval rational interpolant that encloses the lower and upper bound loss probabilities, we further compute:

1. a floating-point approximation for the decay rates $\check{\xi}$ and $\hat{\xi}$ corresponding, respectively, to the minimal and maximal parameters;

2. reliable enclosures $[\check{\xi}]$ and $[\hat{\xi}]$ from the knowledge of $\check{\xi}$ and $\hat{\xi}$ and their corresponding eigenvectors using the technique detailed in [16] and define

$$[\xi] := [\inf\,[\check{\xi}], \sup\,[\hat{\xi}]]. \tag{28}$$

The interval rational interpolant is then given by

$$[r_n(N)] = \frac{[p_n(N)]}{[q_n(N)]}, \qquad N = 1, 2, \ldots, \tag{29}$$

where $[q_n(N)]$ is computed using (22). Using the result of Markov *et al.*[14], the numerator polynomial is given by

$$[p_n(N)] = \sum_{j=1}^{n+1} \left[\left(\prod_{k=1,\ldots,n,\ k\neq j} \frac{N - N_k}{N_j - N_k} \right) \left([\log P_L(N_j)][q_n(N_j)] - [\xi] N_j^{n+1} \right) \right] + [\xi] N^{n+1}, \tag{30}$$

where $N_j,\ j = 1, 2, \ldots, n+1$ are the support points.

Although interval techniques make it computationally feasible to perform a sensitivity analysis and compute $[\check{P}_L(N)]$ and $[\hat{P}_L(N)]$, it is uncomparably faster to compute them only at the support points and construct the rational interval model. However, when only computing the latter the next situation may cause a problem. An interval technique is only reliable if it takes all errors into account, rounding errors (for which interval arithmetic is particularly useful) as well as truncation errors. When the largest buffer size N^* at which both $\log \check{P}_L(N)$ and $\log \hat{P}_L(N)$ are sampled, is already in the range of the asymptotic behaviour of $\log P_L(N)$, then the truncation error is negligible and the interval rational interpolant is trustworthy. If N^* is still too small, then the rational model may not be 100% reliable. Whether or not N^* is in the "correct" range can be checked as follows, if desired. One computes $[\log \check{P}_L(N^*+1)]$ and $[\log \hat{P}_L(N^*+1)]$ and compares the divided differences $[\log \check{P}_L(N^*+1)] - [\log \check{P}_L(N^*)]$ and $[\log \hat{P}_L(N^*+1)] - [\log \hat{P}_L(N^*)]$, which estimate the slope of the lower bound and upper bound loss probabilities, to $[\check{\xi}]$ and $[\hat{\xi}]$. According to our experience, the interval rational model is usually a good estimate of $[\log \check{P}_L(N), \log \hat{P}_L(N)]$.

The interval rational model from section 5 fulfills another role than the floating-point rational model from section 4. While the floating-point multipoint Padé type approximant is built with the aim to accurately model the loss probabilities, the interval rational function is constructed with the aim to return a guaranteed enclosure for the effect of the uncertainty in the parameters. If the latter is not desired, then Theorem 5.1 tells us that a floating-point estimate of the uncertainty effect can be obtained by computing the floating-point multipoint Padé type approximants separately for the maximal and minimal parameters, using the technique detailed in section 4. The one computed with the maximal parameters estimates $\hat{P}_L(N)$ and the one calculated with the minimal parameters approximates $\check{P}_L(N)$. The difference in meaning of the models is also apparent from the simulation output, added as validation in the Figures 3 and 7 (marked by $\star$). In the floating-point model, graphed in the Figures 3, the simulation output is supposed to end up "on" the graph of the rational model. Thus the rational model is validated by the simulation result. In the interval rational model, the simulation results only have to lie inside $[r_n(N)]$. Moreover, when simulating rare event probabilities, such as in the Figures 7(b) and 7(c), the statistical reliability is also less. For instance, for Figure 7(c), the simulated results for $\check{P}_L(1500)$ and $\hat{P}_L(1500)$, which respectively take slightly less and slightly more than 36 hours of simulation time each (on a dual 1 GHz Pentium III machine), come with a relative error estimate of only $\pm 48.5\%$ and $\pm 22.9\%$, repectively. For the simulation output given in Figure 7(b), the technique of importance sampling was used, taking up twice around 6 hours of simulation time for both $\check{P}_L(500)$ and $\hat{P}_L(500)$. Although the computation time was much improved by the importance sampling, the reliability was not, because the simulated results are clearly "off" the true curve (full line) computed with interval arithmetic. As a conclusion we can say that the additional overestimation present in the interval rational model (dotted line), when compared to the true packet loss probability functions $\check{P}_L(N)$ and $\hat{P}_L(N)$, is only a small drawback of the method, taking into account that it is an extremely fast technique. Moreover, the interval rational model is more trustworthy than the simulation.

5.2 Numerical illustration

We use the automatic procedure proposed in section 3 to select the support points. The successive approximants $[r_n(N)]$ with optimal denominator are computed until the following condition is satisfied:

$$\max\left\{\frac{|\sup[r_n]-\sup[r_{n-1}]|}{|\text{mid}[r_n]|},\ \frac{|\inf[r_n]-\inf[r_{n-1}]|}{|\text{mid}[r_n]|}\right\}\le\epsilon, \tag{31}$$

where ϵ is usually chosen to be a small multiple of 0.01. If the functions $[r_{n-1}(N)]$ and $[r_n(N)]$ are visually indistinguishable, then $[r_{n-1}(N)]$ is plotted in the graphical illustrations.

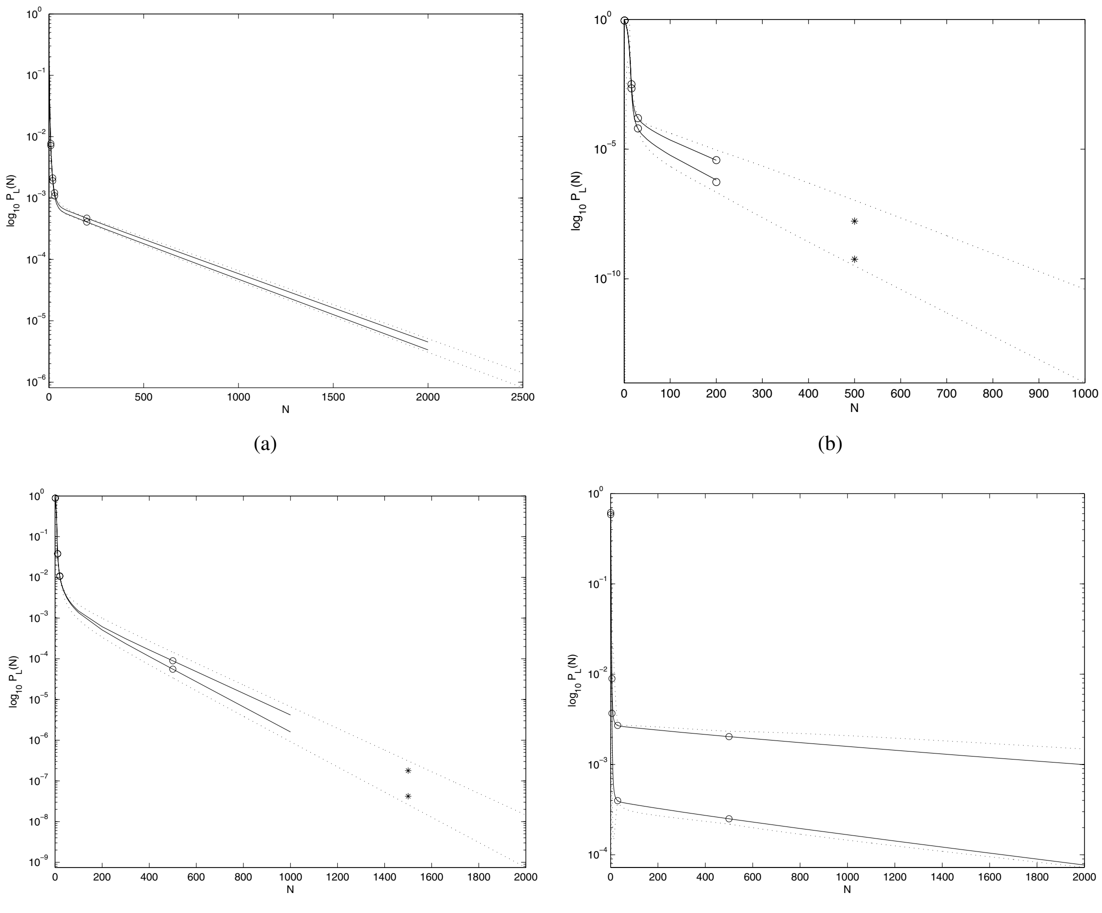

Fig. 7 Packet loss probabilities in intervals

In all figures, the values obtained at support points are circled, the exact value of $\log P_L(N)$ corresponding to minimal and maximal parameters is graphed using full lines, and the approximation $[r_n(N)]$ is graphed using dotted lines. Besides the examples discussed in section 3 and corresponding to the Figures 3(b), 3(c), 3(d), we

Table 3 Examples

Parameter	Figure			
	7(a)	7(b)	7(c)	7(d)
M	15	25	15	5
c	1	15	9	3
$[p]$	$[2.18e-5, 2.2e-5]$	$[2.45e-3, 2.55e-3]$	$[2.18e-4, 2.2e-4]$	(32)
$[q]$	$[6.9e-6, 7.1e-6]$	$[1.145e-3, 1.155e-3]$	$[5.4e-6, 5.6e-6]$	(32)
$[d]$	$[6.8965e-2, 6.8967e-2]$	$[6.45e-1, 6.55e-1]$	$[5.999e-1, 6.001e-1]$	(33)
$[\rho]$	$[0.7803, 0.7875]$	$[0.7306, 0.7534]$	$[0.9748, 0.9762]$	(34)
$[\xi]$	$[-1.1518e-3,\ -1.1143e-3]$	$[-8.9220e-3, -6.9786e-3]$	$[-3.0811e-3, -2.6553e-3]$	(34)
R	2	15	9	3
approximant	r_3	r_3	r_3	r_3

also consider a single-server homogeneous sources network. For the parameters of the examples of section 3, we now consider interval values $[p]$, $[q]$ and $[d]$ for $p,\ q$ and d. For the heterogeneous sources network, we consider intervals $[\boldsymbol{p}]$, $[\boldsymbol{q}]$ and $[\boldsymbol{d}]$ for $\boldsymbol{p}$, $\boldsymbol{q}$ and $\boldsymbol{d}$, given by

$$[\boldsymbol{p}] = \begin{pmatrix} [5.984e-5,\ 7.984e-5] \\ [1.1e-7,\ 3.1e-7] \\ [7.366e-5,\ 9.666e-5] \\ [7.8894e-5,\ 9.8894e-5] \\ [0.98e-6,\ 2.98e-6] \end{pmatrix},\ [\boldsymbol{q}] = \begin{pmatrix} [8.84e-6,\ 1.094e-5] \\ [2.742e-5,\ 4.742e-5] \\ [8.675e-5,\ 1.675e-4] \\ [5.196e-5,\ 7.196e-5] \\ [5.7e-5,\ 7.7e-5] \end{pmatrix} \tag{32}$$

$$[\boldsymbol{d}] = \begin{pmatrix} [0.4552,\ 0.4572], & [0.2943,\ 0.2963] \\ [0.8370,\ 0.8380], & [0.6012,\ 0.6032] \\ [0.8221,\ 0.8241], & [0.1818,\ 0.1838] \\ [0.5411,\ 0.5431], & [0.7322,\ 0.7342] \\ [0.0914,\ 0.0934], & [0.5479,\ 0.5499] \end{pmatrix}. \tag{33}$$

The functions $[\log \hat{P}_L(N)]$ and $[\log \check{P}_L(N)]$ and the interval rational models are shown in the Figures 7. For Figure 7(d)

$$[\rho] = [0.7892, 0.8572], \qquad [\xi] = [-0.2826e-3, -0.1662e-3]. \tag{34}$$

References

[1] A. Baiocchi, Analysis of the loss probability of the MAP/G/1/K queue, Part I: Asymptotic theory, Communications in Statistics-Stochastic Models, **10**, 867(1994).

[2] C. Blondia, A discrete-time batch Markovian arrival process as B-ISDN traffic model, Belgian Journal of Operations Research, Statistics and Computer Science, **32** 3(1992).

[3] A. Cuyt, R.B. Lenin, Computing packet loss probabilities in multiplexer models using adaptive rational interpolation with optimal pole placement, submitted.

[4] A. Cuyt, R.B. Lenin, G. Willems, C. Blondia, and P. Rousseeuw, Rational approximation technique to compute cell loss probabilities in multiplexer models, IEEE Transactions on Computers, **52**, 1(2003).

[5] C. Blondia and O. Casals, Performance analysis of statistical multiplexing of VBR sources: A matrix-analytical approach, Performance Evaluation, **16**, 5(1992).

[6] C.S. Chang, P. Heidelberger, S. Juneja and P. Shahabuddin, Effective bandwidth and fast simulation of ATM intree networks, in: Proceedings of Performance '93, Rome, Italy, October 1993.

[7] A.I. Elwalid and D. Mitra, Analysis, approximations and admission control of a multi-service multiplexing system with priorities, in: Proceedings of INFOCOM '95, 463(1995).

[8] E. Falkenberg, On the asymptotic behavior of the stationary distribution of Markov chains of M/G/1-type, Communications in Statistics-Stochastic Models, **10**, 75(1994).

[9] J. Garcia, O. Casals, A discrete time queueing model to study the cell delay variation in an atm network, Performance Evaluation, **21**, 3(1994).

[10] W. Gong et al., On the convergence of global rational approximants for stochastic discrete event systems, Discrete Event Dynamic Systems - Theory and Applications, **7**, 93(1997).
[11] A.E. Kamal, Efficient solution of multiple server queues with application to the modeling of ATM concentrators, in: Proceedings of IEEE INFOCOM '96, San Francisco, CA, March 1996, 248(1996).
[12] G. Latouche and G. Stewart, Numerical methods for M/G/1 type queues, in: Proceedings of the 1995 International Meeting on Computations with Markov Chains, edited by W. Stewart Kluwer Academic, 571(1995).
[13] Z. Liu et al., Exponential bounds with an application to call admission, Tech. Rep., Univ. Mass, Comp. Sci., Dept., Amherst, MA, TR94-63 (1994).
[14] S. Markov, E. Popova E, U. Schneider and J. Schulze, On linear interpolation under interval data, Mathematics and Computeters in Simulation, **42**, 35(1996).
[15] M.F. Neuts, Structured Stochastic Matrices of M/G/1 Type and their Applications, Marcel Dekker, 1989.
[16] S.M. Rump, Computational error bounds for multiple or nearly multiple eigenvalues, Linear Algebra and Applications, **324**, 209(2001).
[17] S.M. Rump, Verification methods for dense and sparse systems of equations, edited by J. Herzberger, Topics in Validated Computations - Studies in Computational Mathematics, Elsevier, Amsterdam, 63(1994).
[18] N.B. Shroff, IEEE/ACM Trans. Networking, **6**, 441(1998).
[19] T. Thiele, Interpolationsrechnung, Teubner, Leipzig (1909).
[20] J. Van Iseghem, Multipoint Padé-type approximants: convergence, continuity, Computational and applied mathematics. I. Algorithms and theory (Dublin, 1991), C. Brezinski and U. Kulisch (eds), North-Holland, Amsterdam, 237(1992).
[21] H. Wallin, Potential theory and approximation of analytic functions by rational interpolation, in: Proceedings Colloquium Compl. Anal. at Joensuu, Lecture Notes in Mathematics, Vol. 747, Springer Verlag, Berlin, 434(1979).
[22] D. Warner, Hermite interpolation with rational functions, Ph.D. thesis, University of California, 1974.
[23] K. Wuyts and R. Boel, Efficient matrix geometric methods for B-ISDN by using the spectral decomposition of the rate matrices, in: Proceedings of the 2nd conference on Matrix-Analytic Methods in stochastic models, edited by S. Chakravarthy and A. Alfa, Notable Publications, 341(1998).
[24] H. Yang, Global rational approximation for computer systems and communication networks, Ph.D. thesis, University of Massachusetts, Amherst, MA 01003, 1996.