Serie Research Memoranda

How to Solve Numerically the Equilibrium Equations of a Markov Chain with Infinitely Many States

Henk C. Tijms
Michel C.T. van de Coevering

Research Memorandum 1990-46
September 1990
HOW TO SOLVE NUMERICALLY THE EQUILIBRIUM EQUATIONS OF A MARKOV CHAIN WITH INFINITELY MANY STATES

by
Henk C. Tijms and Michel C.T. Van de Coevering
Department of Econometrics, Vrije University
1007 MC Amsterdam, The Netherlands, BITNET: tijms@econ.vu.nl

Subject classification: Markov chains: numerical solution;
Queues: asymptotics.

Abstract This paper presents a simple and practical method to solve the equilibrium equations of a Markov chain when the number of states is infinite. The basic idea is to reduce the infinite system of linear equations to a finite system using the geometric tail behavior of the equilibrium probabilities. Conditions are given under which such a reduction is possible. The reduction leads to a remarkably small system of linear equations which can be routinely solved by a Gaussian elimination method. An application is given to the D/G/1 queue with scheduled arrivals.

How to solve numerically the equilibrium equations of a Markov chain is an important question that arises in numerous problems in applied probability, computer science and operations research. In many of these applications the state space of the Markov chain is infinite. What one usually does to solve numerically the infinite set of equilibrium equations is approximate the infinite Markov model by a truncated model with a finite number of states so that the equilibrium probability of the set of deleted states is very small. Indeed, for a finite-state truncation with a sufficiently large number of states, the difference between the two models will be negligible. However, such a truncation usually leads to a finite but very large system of linear equations whose numerical solution will be quite time-consuming, though an arsenal of good methods are available to solve the equilibrium equations of a finite Markov chain, see [3], [4], [5], [8] and references therein. Moreover, it is somewhat disconcerting that we need a brute-force approximation to solve numerically the infinite-state model. Usually we introduce infinite-state models to obtain mathematical simplification, and now in its numerical analysis using a brute-force truncation we are proceeding in the reverse direction.
The purpose of this paper is to give an elegant and generally applicable approach for solving numerically an infinite system of equilibrium equations underlying a Markov chain. The basic idea of the approach is to reduce the infinite system of linear equations to a finite one by exploiting the geometric tail of the state probabilities. This results in a finite system of linear equations whose size is typically much and much smaller than the size of the finite system obtained from a brute-force truncation. It is our experience that in practical applications the approach based on the geometric tail behavior leads to a system of linear equations with 1-100 variables. This is also true for applications in which a brute-force truncation would lead to systems with many thousands of variables. The system of linear equations based on the geometric tail behavior can be routinely solved on a PC within seconds using a standard Gaussian elimination procedure.

The paper is organized as follows. The general idea of the reduction using the geometric tail behavior of the state probabilities is given in section 1. In section 2 the powerfulness of the approach is demonstrated for the D/G/1 queueing model with scheduled arrivals. A simple algorithm for the computation of the waiting-time probabilities in this important queueing model will be derived.

1 The General Approach

Let \( \{X_n, n=0,1,\ldots\} \) be a discrete-time Markov chain with a denumerable state space to be denoted by \( I=\{0,1,\ldots\} \). Suppose that the Markov chain \( \{X_n\} \) is irreducible and has a unique stationary probability distribution \( \{\pi_i, i=0,1,\ldots\} \). The state probabilities \( \pi_i \) are the unique solution to the system of linear equations (see e.g. [8]),

\[
\pi_i = \sum_{j=0}^{\infty} \pi_j p_{ji}, \quad i=0,1,\ldots, \tag{1}
\]

\[
\sum_{i=0}^{\infty} \pi_i = 1. \tag{2}
\]

How to solve numerically this infinite system of linear equations without using a finite-state model approximation? Therefore we assume that the following asymptotic relation holds,
for some number \( \tau \) with \( 0 < \tau < 1 \). That is, the state probabilities exhibit a geometric tail behavior. Under which conditions the result (3) holds and how to compute \( \tau \) will be discussed afterwards. For the moment let us assume that (3) holds. Then, for a sufficiently large integer \( N \),

\[
\pi_j \approx \pi_N \tau^{j-N} \quad \text{for all } j \geq N.
\]

How large \( N \) should be chosen has to be determined experimentally and depends of course on the required accuracy in the state probabilities. However, empirical investigations show that it is generally true that for all practical purposes the ratio \( \pi_j / \pi_{j-1} \) can be replaced by \( \tau \) already for remarkably small values of \( j \). Replacing \( \pi_j \) by \( \pi_N \tau^{j-N} \) for \( j \geq N \) in (1)-(2) leads to the following finite set of linear equations:

\[
\pi_i = \sum_{j=0}^{N} a_{ij} \pi_j, \quad i=0,1,...,N-1 \tag{4}
\]

\[
\sum_{i=0}^{N-1} \pi_i + \frac{1}{1-\tau} \pi_N = 1, \tag{5}
\]

where for any \( i=0,...,N-1 \) the coefficients \( a_{ij} \) are given by

\[
a_{ij} = \begin{cases} 
  p_{j1} & \text{for } j=0,...,N-1 \\
  \sum_{k=N}^{\infty} \tau^{k-N} p_{k1} & \text{for } j=N
\end{cases}
\tag{6}
\]

In the many numerical examples we studied it appeared that this system of linear equations is nonsingular even for very small values of \( N \) yielding the state probabilities only in a few decimals exact. It is our experience that the value of \( N \) is usually below 100 when the calculated values of the state probabilities are required to be exact in at least in seven decimals. Taking into account the typical size of the system (4)-(5), we found that a Gaussian elimination method such as the Gauss-Jordan method is a reliable and fast method to solve this system of linear equations. The source book [6]
gives a code that can be directly used by the nonspecialist.

It remains to outline under which conditions the asymptotic expansion (3) is true and how to compute the decay factor $\tau$. Therefore we define the generating function $\Pi(z)$ by

$$
\Pi(z) = \sum_{j=0}^{\infty} \pi_j z^j \quad \text{for } |z| \leq 1.
$$

We now need that the generating function $\Pi(z)$ is also defined in some region outside the unit circle $|z| \leq 1$. To be more precise, let us assume that $\Pi(z)$ can be represented as

$$
\Pi(z) = \frac{A(z)}{B(z)},
$$

where $A(z)$ and $B(z)$ are analytic functions whose domains of definitions can be extended to a region $|z| < R$ in the complex plane for some $R > 1$. In many applications the representation (7) is possible. It is no restriction to assume that $A(z)$ and $B(z)$ have no common roots. Let us further assume that the following conditions can be verified:

C1 The equation

$$
B(x) = 0
$$

has a real root $x_0$ on the interval $(1, R)$.

C2 The function $B(z)$ has no zeros in the domain $1 < |z| < x_0$ of the complex plane.

C3 The zero $z = x_0$ of $B(z)$ is of order one and is the only zero of $B(z)$ on the circle $|z| = x_0$.

Then, by the Laurent expansion from complex function theory, it follows that for some number $R_1$ with $x_0 < R_1 < R$,

$$
\Pi(z) = \frac{a_{-1}}{z-x_0} + U(z) \quad \text{for } |z| \leq R_1, \ z \neq x_0,
$$

where $U(z)$ is some analytical function in the domain $|z| \leq R_1$ and the residue $a_{-1} \neq 0$ is given by $\lim_{z \to x_0} (z-x_0) \Pi(z)$. Since $U(z)$ is analytic for $|z| \leq R_1$ a Taylor series expansion $U(z) = \sum_{j=0}^{\infty} u_j z^j$ is true for $|z| \leq R_1$. Note that
As follows from the observation that the power series \( \sum_{j=0}^{\infty} u_j z^j \) is convergent for \( z=R \). Observing that the power series representation \( \Pi(z)=\sum_{j=0}^{\infty} \pi_j z^j \) extends to \( |z|<x_0 \) by analytic continuation, it now follows from (9) that
\[
\sum_{j=0}^{\infty} \pi_j z^j = -a^{-1} \sum_{j=0}^{\infty} x_0^{-j} z^j + \sum_{j=0}^{\infty} u_j z^j \quad \text{for } |z|<x_0.
\]
Equating coefficients yields
\[
\pi_j = -a^{-1} x_0^{-j-1} + u_j \quad \text{for all } j \geq 0.
\]
Since \( u_j = O(R^{-j}) \) and \( R>x_0 \), the coefficient \( u_j \) tends faster to zero than \( x_0^{-j} \) as \( j \to \infty \). Hence we find, for some constant \( c>0 \),
\[
\pi_j \approx c x_0^{-j} \quad \text{for } j \text{ large enough. (10)}
\]
The asymptotic expansion gives the desired result (3) with \( \tau=1/x_0 \).

Summarizing, under the assumption that the representation (7) holds and the conditions C1-C3 are satisfied, the asymptotic expansion (3) holds. The decay factor \( \tau \) is found by computing the zero \( x_0 \) of the equation (8) and next putting \( \tau=1/x_0 \).

2 Application to the D/G/1 Queue

An important queueing model is the single server D/G/1 queue with scheduled arrivals and generally distributed service times. This model is amongst others useful in the analysis of appointment systems and clocked schedules for computer systems. In many applications it is of great importance to compute the waiting-time probabilities, particularly the small tail probabilities. The waiting-time probabilities can be quickly computed for larger values of the squared coefficient of variation of the service time assuming that the service time has a hyperexponential distribution, see [8]. However, in many cases of interest, the service time has a low coefficient of variation. Then the exact computation of the waiting-time probabilities becomes computationally very expensive when using standard methods. To avoid the expensive exact calculations some approximations have been proposed in
the paper [2]. Though this paper gives excellent approximations for the mean waiting time and the delay probability, the approximations for the waiting-time probabilities are only of limited value. In this section it will be shown that the approach discussed in the previous section enables us to compute quickly and routinely the waiting-time probabilities. The computation times are a matter of seconds on a PC.

2.1 The Waiting-Time Probabilities

In what follows we restrict ourselves to the case that the service time $S$ of a customer has a squared coefficient of variation $c_s^2$ satisfying

$$0 < c_s^2 \leq 1,$$

where $c_s^2$ denotes the ratio of the variance and the squared mean of the service time. As pointed out before, for the case of $c_s^2 \leq 1$ the computation of the waiting-time probabilities offers no difficulties at all when the service time has a hyperexponential distribution, see [8]. In our analysis it will be assumed that the service time density is a mixture of Erlang densities. From a practical point of view it is no restriction to assume a service density of this form since extensive numerical experiments indicate that the waiting-time probabilities are rather insensitive to more than the first two moments of the service time density provided that $c_s^2 \leq 1$. To any service time $S$ with given mean $E(S)$ and squared coefficient of variation $c_s^2(\leq 1)$, we can fit a density of the form

$$p \mu^r \frac{t^{r-2}}{(r-2)!} e^{-\mu t} + (1-p) \mu^r \frac{t^{r-1}}{(r-1)!} e^{-\mu t}, \quad t \geq 0. \quad (11)$$

By choosing the integer $r$ and the numbers $p$ and $\mu$ according to

$$\frac{1}{r} \leq c_s^2 < \frac{1}{r-1} \quad (12)$$

and

$$p = \frac{1}{1+c_s^2} \left\{ r c_s^2 - \sqrt{r(1+c_s^2)-r^2 c_s^2} \right\}, \quad \mu = \frac{r-p}{E(S)} \quad (13)$$
this mixture of $E_{r-1}$ and $E_r$ densities has the same first two moments as the service time $S$. The reason for using a mixture of Erlangians for the service time is that it enables us to compute the waiting-time probabilities through Markov chain analysis. A service time $S$ with density (11) has the following probabilistic interpretation. With respective probabilities $p$ and $1-p$ the service time consists of $r-1$ respectively $r$ consecutive phases, where the completion times of the phases are independent random variables each having an exponential distribution with the same mean $\mu$.

Assuming the service time density (11) and using the memoryless property of the exponential distribution, we can now define the Markov chain

$$X_n = \text{the number of uncompleted service phases in the system just prior to the } n\text{th arrival epoch.}$$

This Markov chain has the infinite state space $\{0,1,\ldots\}$. Denoting by

$$D = \text{the fixed time between two consecutive arrivals},$$

and assuming that the offered load $\rho$, defined by

$$\rho = \lambda E(S) \quad \text{with} \quad \lambda = 1/D,$$

is smaller than 1, the Markov chain $\{X_n\}$ has a unique stationary probability distribution $\{\pi_i, i=0,1,\ldots\}$. Once the state probabilities $\pi_i$ have been computed, the stationary waiting-time probabilities can be calculated from

$$P(W > x) = \sum_{i=1}^{\infty} \pi_i \sum_{k=0}^{i-1} e^{-\mu x} \frac{(\mu x)^k}{k!}, \quad x \geq 0,$$

where the generic variable $W$ denotes the waiting time of a customer arriving in steady-state. To explain this expression, note that the conditional waiting time of a customer finding upon arrival $i$ uncompleted service phases is the sum of $i$ independent exponentials with the same mean and thus has an Erlang-$i$ density. For computational purposes the expression (14) should be rewritten as

$$P(W > x) = \sum_{k=0}^{\infty} e^{-\mu x} \frac{(\mu x)^k}{k!} \left(1 - \sum_{j=0}^{k} \pi_j\right), \quad x \geq 0.$$  

(15)
This representation has the computational advantage of a faster converging series. In addition, using basic properties of the Poisson distribution, the infinite sum $\sum_{k=0}^{\infty}$ in (15) can be easily replaced by a finite sum $\sum_{k=0}^{M}$. For example, using the truncation integer

$$M = 20 + [\mu x + 6\sqrt{\mu x}]$$,

the absolute error introduced in the computation of $P(W > x)$ is smaller than $10^{-8}$ for any value of $x \geq 0$. In fact, for sufficiently large values of $x$, one can avoid the series representation (15) and use the asymptotic expansion

$$P(W > x) \approx \gamma e^{-\delta x} \quad \text{for } x \text{ large}, \quad (16)$$

where the constants $\gamma$ and $\delta$ are easily computed from the geometric tail behavior of the state probabilities $\pi_i$, as will be shown in the next subsection.

The above discussion clearly shows that an effective computation of the waiting-time probabilities $P(W > x)$ stands or falls with the computation of the state probabilities $\pi_i$. In the next subsection we give a simple and fast algorithm for the computation of the state probabilities.

2.2 The Computation of the State Probabilities

The state probabilities $\pi_i$ are the equilibrium distribution of the embedded Markov chain $\{X_n\}$ describing the number of uncompleted service phases in the system just prior to the arrival epochs. For any $i \geq 0$, the one-step transition probabilities $p_{ij}$ of this embedded Markov chain are given by

$$p_{ij} = p a(i+r-1-j) + (1-p) a(i+r-j) \quad \text{for } j \geq 1, \quad (17)$$

where $a(k)$ is a shorthand notation for

$$a(k) = e^{-\mu D} \frac{(\mu D)^k}{k!} \quad \text{for } k=0,1,... ,$$

with the convention $a(k) = 0$ for $k<0$. Note that $a(k)$ is the probability of exactly $k$ phase completions during an interarrival time assuming that ample
service phases are present. In formula (17) it should be required that \( j \neq 0 \).

The one-step probability \( p_{10} \) is computed from

\[
p_{10} = 1 - \sum_{j=1}^{1+r} p_{1j} \quad \text{for any } i=0,1,\ldots.
\]

Once we have specified the one-step transition probabilities \( p_{1j} \), we can derive the generating function \( \Pi(z) = \sum_{i=0}^{\infty} \pi_i z^i \) of the equilibrium probabilities \( \pi_i \). In this particular model it can be shown after tedious algebra that (see also [7])

\[
\Pi(z) = \frac{\sum_{n=0}^{\infty} \alpha_n (1-z^{-n})}{1 - (pz^{-1} + (1-p)z^r) e^{-(|D|-1/z)}}.
\]

where \( \{\alpha_n\} \) is some sequence of nonnegative numbers which need not be specified.

The generating function \( \Pi(z) \) is indeed the ratio of two analytic functions \( A(z) \) and \( B(z) \) whose domains of definition can be extended outside the unit circle. The functions \( A(z) \) and \( B(z) \) are analytic for all \( |z|<R \) with \( R=\omega \). It can be verified that the conditions C1-C3 in section 1 are satisfied. In particular, using the transformation \( y=1/x \) in the equation (7), it follows that \( \tau \) can be computed as the unique root of the equation

\[
(py + 1 - p)e^{-(\mu D(1-y))} \cdot y^r = 0 \quad (18)
\]

on the interval (0,1). The root \( \tau \) describes the geometric tail behavior of the tail probabilities \( \pi_i \) accordingly to \( \pi_i / \pi_{i-1} \approx \tau \) for \( i \) sufficiently large.

Once the root \( \tau \) has been computed, the following system of linear equations is solved:

\[
\pi_i = \sum_{j=0}^{N} a_{ij} \pi_j \quad \text{for } i=0,\ldots,N-1 \quad (19)
\]

\[
\sum_{i=0}^{N-1} \pi_i + \frac{1}{\tau} \pi_N = 1, \quad (20)
\]

where, for any \( i=0,\ldots,N-1 \),
\[
    a_{ij} = \begin{cases} 
        p_{ji} & \text{for } j=0,\ldots,N-1 \\
        \sum_{k=N}^{\infty} p_{kl} \tau^{k-N} & \text{for } j=N.
    \end{cases}
\]

Here \( N \) is a sufficiently large integer so that
\[
    \pi_j \approx \pi_N \tau^{j-N} \quad \text{for all } j \leq N.
\]  

(21)

In practical situations one will be usually satisfied with a choice of \( N \) for which the calculated values of the state probabilities \( \pi_j \) are exact in at least seven decimals. Conservative choices for \( N \) for which this accuracy is achieved are displayed in Table 1 for a coarse grid of values for \( \rho \) and \( c_s^2 \).

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Choices for ( N ) to achieve a seven-decimal accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \rho \leq 0.5 )</td>
<td>( 0.5 \leq \rho \leq 0.7 )</td>
</tr>
<tr>
<td>( 0.2 \leq c_s^2 \leq 1 )</td>
<td>10</td>
</tr>
<tr>
<td>( 0.1 \leq c_s^2 \leq 0.2 )</td>
<td>10</td>
</tr>
<tr>
<td>( 0.05 \leq c_s^2 \leq 0.1 )</td>
<td>10</td>
</tr>
<tr>
<td>( 0.02 \leq c_s^2 \leq 0.05 )</td>
<td>10</td>
</tr>
<tr>
<td>( 0.01 \leq c_s^2 \leq 0.02 )</td>
<td>10</td>
</tr>
</tbody>
</table>

For very light traffic it is not necessary to compute \( \tau \) and to solve the linear equations. Then, for all practical purposes, we can directly put \( \pi_0 = 1 \) and \( P(W > t) = 0 \) for all \( t \geq 0 \). In fact, for very light traffic the state probabilities \( \pi_j \) are practically zero before the asymptotic expansion \( \pi_j / \pi_{j-1} \approx \tau \) applies. It turns out that the difference between \( \pi_0 \) and 1 is less than \( 10^{-2} \) for \( \rho \) in the range \( c_s^2 \) when \( \rho(c_s^2) \) has the respective values 0.05, 0.10, 0.20, 0.30, and 0.40 for the respective ranges of \( c_s^2 \) in Table 1.

It is remarkable to see from Table 1 that already for relatively small values of \( N \) the state probabilities \( \pi_j \) can be computed at a very high accuracy. The value of \( N \) increases only slightly when \( \rho \) becomes very close to 1. This is in sharp contrast with the brute-force truncation for which the integer \( M \) satisfying \( \sum_{j=M}^{\infty} \pi_j \) sharply increases when \( \rho \) gets close to 1. For example, suppose \( c = 10^{-7} \). Then for \( \rho = 0.9 \) the integer \( M \) has the respective values 76, 75, and 71 for \( c_s^2 = 1, 0.5, \) and 0.01, while for \( \rho = 0.99 \) the inte-
Remark 1. In implementing an algorithm, it is always useful to have an accuracy check. An accuracy check for the calculated values of the state probabilities $\pi_\ell$ is provided by the relation

$$\rho = \frac{1}{D} \sum_{j=0}^{\infty} \pi_j \left[ \sum_{i=1}^{r} \sum_{k=0}^{1+j-1} e^{-\mu D} \frac{(\mu D)^k}{k!} \right]$$

where $p_\ell r = p_\ell$, $p = 1 - p_\ell$, and $p_\ell = 0$ otherwise. This relation can be explained as follows. The long-run fraction of time the server is busy equals $\rho$. On the other hand, this long-run fraction is also equal to the average expected amount of time the server is busy during an interarrival time divided by the average length of the interarrival time. Using this observation, the relation easily follows.

Remark 2. In addition to the asymptotic expansion for the state probabilities, an asymptotic expansion for the waiting-time probabilities applies as well. The relations (15) and (21) imply that, for $x$ large,

$$P(W > x) \approx e^{-\mu x} \sum_{k=0}^{\infty} \pi_k \sum_{j=k+1}^{1+j-1} \frac{e^{-\mu D} (\mu D)^k}{k!}$$

yielding the asymptotic expansion

$$P(W > x) \approx \pi_\ell^N \frac{\tau^{-N+1}}{1-\tau} e^{-\mu(1-\tau)x} \text{ for } x \text{ large enough.} \quad (23)$$

It turns out that the higher the traffic load on the system, the earlier this asymptotic expansion applies. An appropriate measure for the traffic load on the system is the delay probability $P_q W = P(W > 0)$. To indicate how large $x$ should be so that the asymptotic expansion (23) can be applied, it is convenient to use the waiting-time percentiles. For any $\alpha$ with $1 - P_q W \leq \alpha < 1$, the $\alpha$th percentile $\xi(\alpha)$ is defined by $P(W \leq \xi(\alpha)) = \alpha$. In practice one will cer-
tainly be satisfied in having the percentiles in four figures exact. Then the asymptotic expansion (23) can be used to compute the percentiles \( \xi(\alpha) \) for \( \alpha = \alpha(P_w) \), where conservative estimates for \( \alpha(P_w) \) are given in Table 2 for a grid of \( P_w \) values. These numbers \( \alpha(P_w) \) are typical for many other queueing models.

Table 2

<table>
<thead>
<tr>
<th>( P_w )</th>
<th>0.5</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8 ≤ ( P_w ) ≤ 0.9</td>
<td>0.8</td>
</tr>
<tr>
<td>0.7 ≤ ( P_w ) ≤ 0.8</td>
<td>0.9</td>
</tr>
<tr>
<td>0.6 ≤ ( P_w ) ≤ 0.7</td>
<td>0.99</td>
</tr>
<tr>
<td>0.5 ≤ ( P_w ) ≤ 0.6</td>
<td>0.999</td>
</tr>
</tbody>
</table>

2.3 Implementation Aspects

The algorithm for the waiting-time probabilities boils down to solving one nonlinear equation in a single variable and a small system of linear equations. In implementing an algorithm one often faces unexpected difficulties. This happened also in the present algorithm, particularly when solving the nonlinear equation (18). Usually, once an interval is known for the root of an equation in a single variable, a safe approach to solve this equation is to use bisection or Brent's method, see [6]. However, the equation (18) presents numerical difficulties for certain parameter combinations. In particular for \( r \) large the left-hand side of (18) is extremely flat in the neighborhood of the actual zero \( \tau \) and assumes values that cannot be numerically distinguished from zero within the machine precision. The trick to circumvent this difficulty is to use logarithms. Divide both sides of (18) by \( y^r \) and bring the constant term 1 to the right. Then the function \( f(y) \) in the resulting equation \( f(y) = 1 \) is positive for all \( y \) and thus solving (18) is equivalent to solving

\[
\ln(py+1-p) - \mu D(1-y) - r \ln(y) = 0. \tag{24}
\]

Another important implementation aspect is the following. To minimize the
computational effort to generate the coefficients \(a_{ij}\) of the linear equations (19)-(20), some thought should be given to the computation of these coefficients. The coefficients \(a_{ij}\) with \(i \neq 0\) are trivial to compute. They depend only on the difference \(i-j\) and are expressed in Poisson probabilities which can be recursively computed. However, the computation of the coefficients \(a_{0j}\) is time-consuming when it is not realized that the probabilities \(p_{10}\) can be recursively computed. It is easy to verify the recursion

\[
p_{i+1,0} = p_{i0} - p a(i+r-1) - (1-p) a(i+r), \quad i=0,1,...
\]

with \(p_{00} = 1 - \sum_{j=1}^{r} (p a(r-l-j) + (1-p) a(r-j))\). Finally, it is worthy to point out that the summation in the infinite series \(a_{0N} = \sum_{i=0}^{\infty} p_{10}^i i^N\) can be stopped as soon as \(\tau^{-N} p_{10} \leq (1-\tau)c\) for some small \(c\) (say, \(c=10^{-12}\)). This stopping criterion is based on the fact that \(p_{10}\) is decreasing in \(i\).

3 Conclusion

Our experiments demonstrate that the approach using the geometric tail behavior of the state probabilities is very effective to reduce an infinite system of equilibrium equations for a Markov chain to a finite system of linear equations with relatively few variables. The proposed method is generally applicable for the computation of the equilibrium distribution of a Markov chain with infinitely many states. The method was applied to the single server D/G/1 queue with scheduled arrivals. We found successful applications to many other congestion problems including the multi-server M/D/c queue. It was pointed out to us by Don Gaver that the early paper [1] uses a very similar approach to solve the equilibrium equations of the M/D/c queue. This interesting paper seems to have remained unnoticed in the literature.

Acknowledgement The first author acknowledges the hospitality of the Research Institute for Applied Knowledge Processing (FAW) at the University of Ulm, BRD, where part of this research was performed.
REFERENCES


