Transient Solution of Markov Chains Using the Uniformized Vandermonde Method

Llorenç Cerdà-Alabern

Universitat Politècnica de Catalunya Computer Architecture Dep.
Jordi Girona 1-3 E-08034, Barcelona, Spain
llorenc@ac.upc.edu

January 13, 2011

Abstract

This paper investigates the transient solution of discrete and continuous time Markov chains (DTMC and CTMC) using the class of methods based on the *undetermined coefficients* approach. Two methods that belong to this class are investigated: The well known Eigenvectors method, and what it is referred in the paper as the *Vandermonde method*, because it leads to the solution of a Vandermonde system of equations. Even if the Vandermonde method is possibly the simplest method to obtain the transient solution of a Markov chain, it has received few attention in the literature. This paper fills this gap, showing its advantages. The Vandermonde method is also exploited to derive interesting relations between a CTMC and its uniformized chain. Based on them, it is proposed a simple, yet powerful method to compute the transient solution of Markov chains referred to as the *Uniformized Vandermonde* method.

1 Introduction

Probabilistic modeling using Markov chains has been successfully exploited in almost all fields of modern applied mathematics. The *stationary* solution of a Markov chain is more easy to compute than the transient solution, and it is enough in many cases. However, some applications, as reliability modeling, are primarily interested in the transient solution (see e.g. [6]).

In the literature there have been proposed many methods to compute the transient solution of Markov chains. Some examples are the approaches based on Laplace transform techniques [6], the exponential matrix [7], differential equation solvers, etc. The chapter 8 of the classic book of Stewart [12] is dedicated to this topic. However, most of these algorithms are not general, are difficult to implement, or are applicable only to chains with a small number of states.

This paper focuses on the class of methods based on the *undetermined coefficients* approach. This approach consists on making an intelligent guess of a function with constant coefficients for the solution of an equation, and then use boundary conditions to solve these coefficients. The undetermined coefficients approach has been successfully used to find the solution of many difference and differential equations. The transient solution of a Discrete or Continuous Time Markov Chain, DTMC and CTMC, is the solution of a difference or differential equation respectively. Therefore, DTMC and CTMC are suited to be solved using algorithms based on the undetermined coefficients approach.

The well known Eigenvectors method, explained later in this paper, belongs to the class of undetermined coefficients approaches. Another possibility is applying the undetermined coefficients approach directly to the solution of the Markov chain, expressed in terms of its eigenvalues. By doing so, the undetermined coefficients are obtained by solving a confluent Vandermonde system of equations. In this paper the pros and cons of these methods are analyzed. They will be referred to respectively as the *Eigenvectors* and *Vandermonde* methods. Furthermore, in this paper it is proposed a novel method that combines the Vandermonde method with the well known technique of Uniformization. Thorough the paper, this novel proposal will be referred to as the *Uniformized Vandermonde* method. Numerical results show that the Uniformized Vandermonde method yields a simple yet powerful technique to compute the transient solution of a CTMC.

The rest of the paper is organized as follows. First, the function with constant coefficients for the solution of a DTMC is presented in section 2, and in section 3 it is solved using the Eigenvectors and Vandermonde methods. In sections 4 and 5, the same it is done for a CTMC. Notice that most of the contents of these sections are not a contribution of this paper. For instance, the Eigenvectors method applied to Markov chains can be found in classic books as [4, sec. 4.8]. The expression for the exponential matrix having confluent eigenvalues can be found in early works as [3, 5]. Nevertheless, to my best knowledge, the general solution for a DTMC (equation (3)), has been not reported previously in the literature. The CTMC counterparts have been included for the sake of completeness, and because they are the basis of Uniformized Vandermonde method presented in section 6. An additional reason is that, even if the Vandermonde method is the simplest way to compute the transient

solution of a Markov chain, it is not covered, or it is done superficially, in the books that I have found in the literature. This paper tries to fill this gap, giving examples and highlighting the striking parallelism that exists when this method is used to solve DTMC and CTMC. In section 7 are carried out numerical experiments analyzing the different methods described in this paper. Finally, concluding remarks are given in Section 8.

2 Discrete Time Markov Chains

Let $X(n)$ be an homogeneous finite-state discrete-time Markov chain with k states, and one step transition probability matrix of size k, $\mathbf{P}^{k\times k}$. The transient solution is given by the powers of $\mathbf{P}$ [4]:

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0)\,\mathbf{P}^n \tag{1}$$

where $\boldsymbol{\pi}(0)$ is the initial distribution, and the components of the row vector $\boldsymbol{\pi}(n)$ are the probabilities

$$\pi_j(n) = Prob\left\{X(n) = j \;\middle|\; \boldsymbol{\pi}(0)\right\} \tag{2}$$

If we are interested in computing $\pi_j(n)$ for only a reduced number of states, the following expression may be more efficient that computing the powers of the matrix $\mathbf{P}$.

Theorem 2.1 Let λ_l, $l = 1, \cdots L$ be the eigenvalues of $\mathbf{P}$, each with multiplicity k_l ($\sum_l k_l = k$). Without loss of generality, assume a possible eigenvalue $\lambda_1 = 0$ with multiplicity k_1. Then:

$$\pi_j(n) = \sum_{m=0}^{k_1-1} a_j^{(1,m)}\delta_{n-m} + \sum_{l=2}^{L}\lambda_l^n \sum_{m=0}^{k_l-1} a_j^{(l,m)}\, n^m \tag{3}$$

where δ_k is the Kronecker's delta ($\delta_k = 1$ for $k = 0$ and 0 otherwise).

Proof. Using the Jordan Form of the matrix $\mathbf{P} = \mathbf{A}\,\mathbf{J}\,\mathbf{A}^{-1}$, where $\mathbf{A}$ is some invertible matrix, and $\mathbf{J} = \mathrm{diag}(\mathbf{J}_1, \mathbf{J}_2, \cdots \mathbf{J}_L)$ is a diagonal block matrix, with the Jordan blocks $\mathbf{J}_l$ ($\mathbf{J}_l$ has λ_l in the main diagonal, 1s in the first upper diagonal, and size equal to the multiplicity of λ_l, k_l):

$$\mathbf{J}_l^{k_l\times k_l} = \begin{bmatrix} \lambda_l & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_l & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \lambda_l & 1 & \cdots & 0 & 0 \\ \cdots & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & \lambda_l & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & \lambda_l \end{bmatrix} \tag{4}$$

We have that $\mathbf{J}^n = \mathrm{diag}(\mathbf{J}_1^n, \mathbf{J}_2^n, \cdots \mathbf{J}_L^n)$. To compute the powers of a Jordan block, we write:

$$\mathbf{J}_l = \lambda_l\,\mathbf{I} + \mathbf{U}_1^{k_l\times k_l} \tag{5}$$

where $\mathbf{U}_1^{k_l\times k_l}$ is a matrix with the first upper diagonal equal to 1, and the other elements equal to zero:

$$\mathbf{U}_1^{k_l\times k_l} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ \cdots & & & & & & \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix} \tag{6}$$

It can be easily obtained that:

$$(\mathbf{U}_1^{k_l\times k_l})^n = \mathbf{U}_n^{k_l\times k_l} \tag{7}$$

where $\mathbf{U}_n^{k_l\times k_l}$ is a matrix with the n-th upper diagonal equal to 1, and the other elements equal to zero. $\mathbf{U}_n^{k_l\times k_l} = \mathbf{0}$ if $n \geq k_l$.

Assume first that $\lambda_l \neq 0, \forall l$. We can write:

$$\mathbf{J}_l^n = (\lambda_l\,\mathbf{I} + \mathbf{U}_1^{k_l\times k_l})^n = \sum_{i=0}^{n}\binom{n}{i}\lambda_l^{n-i}\,(\mathbf{U}_1^{k_l\times k_l})^i = \lambda_l^n \times$$
$$\begin{bmatrix} 1 & \lambda_l^{-1}\binom{n}{1} & \lambda_l^{-2}\binom{n}{2} & \cdots & \lambda_l^{-(k_l-1)}\binom{n}{k_l-1} \\ 0 & 1 & \lambda_l^{-1}\binom{n}{1} & \cdots & \lambda_l^{-(k_l-2)}\binom{n}{k_l-2} \\ 0 & 0 & 1 & \cdots & \lambda_l^{-(k_l-3)}\binom{n}{k_l-3} \\ \cdots & & & & \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{8}$$

where it is assumed $\binom{k}{j} = 0$ for $j > k$, and $(\mathbf{U}_1^{k_l\times k_l})^0 = \mathbf{I}$.

Since $\binom{n}{k_l-1} = \frac{n\,(n-1)\cdots(n-k_l+2)}{(k_l-1)!} = \frac{1}{(k_l-1)!}(n^{k_l-1} + \cdots)$ is a polynomial in n of degree $k_l - 1$, we conclude that the elements of the matrix $\mathbf{P}^n = \mathbf{A}\,\mathbf{J}^n\,\mathbf{A}^{-1}$ will be linear combinations of the eigenvalues of $\mathbf{P}$ to the power of n, times polynomials in n of degree equal, at most, to their multiplicity-1, and so it will be $\boldsymbol{\pi}(0)\,\mathbf{P}^n$, as stated in equation (3).

If there is an eigenvalue $\lambda_1 = 0$ with multiplicity k_1, the Jordan block associated with this eigenvalue will be: $\mathbf{J}_1 = \mathbf{U}_1^{k_1\times k_1}$. Therefore, the result of $\mathbf{A}\,\mathbf{J}^n\,\mathbf{A}^{-1}$ associated with this eigenvalue will be a constant, possibly different, for each $n = 0, \cdots, k_1 - 1$, and 0 for $n \geq k_1$ (recall that $(\mathbf{U}_1^{k_1\times k_1})^n = 0$ for $n \geq k_1$). Note also that if $k_1 = 1$, then it is $(\mathbf{U}_1^{1\times 1})^n = 0^n = \delta_n$. Here it is assumed $0^0 = 1$ for consistency with the fact that $\boldsymbol{\pi}(n)|_{n=0} = \boldsymbol{\pi}(0)\,\mathbf{P}^0 = \boldsymbol{\pi}(0)\,\mathbf{I} = \boldsymbol{\pi}(0)$. □

Remark 2.1 Assume that the Markov chain has the eigenvalue $\lambda_1 = 1$ with multiplicity $k_1 > 1$ (see appendix A). Clearly, the polynomial corresponding to this eigenvalue:

$$p_j^{(1)}(n) = \sum_{m=0}^{k_1-1} a_j^{(1,m)}\, n^m \tag{9}$$

must have the coefficients $a_j^{(1,m)} = 0$, for $m = 1, \cdots, k_1 - 1$. Otherwise, the solution (3) will be unbounded. Therefore, the eigenvalue $\lambda_1 = 1$ has only 1 undetermined coefficient ($p_j^{(1)}(n) = a_j^{(1,0)}$). A consequence of this fact is that the geometric multiplicity of the eigenvalue $\lambda_1 = 1$

must be equal to its algebraic multiplicity, k_1 (the geometric multiplicity is defined as the number of linearly independent eigenvectors associated with the eigenvalue). Thus, the eigenvalue $\lambda_1 = 1$ has k_1 Jordan blocks of size 1. There may be other eigenvalues with geometric multiplicity larger than 1. However, it is not necessary to compute the geometric multiplicity of the eigenvalues: If the geometric multiplicity of an eigenvalue λ_l is g_l, when solving for $a_j^{(l,m)}$ it will be obtained that $a_j^{(l,m)} = 0$ for $m > k_l - g_l$. ■

3 Undetermined Coefficients of a DTMC

In equation (3) there are up to k undetermined coefficients $a_j^{(l,m)}$ of $\pi_j(n)$ to be determined. This paper focuses on the computation of these coefficients. Thorough the paper the sentence *the UC of* $\pi_j(n)$, or simply *the UC*, will be used to refer to the undetermined coefficients $a_j^{(l,m)}$ of $\pi_j(n)$. Additionally, to denote the UC of $\pi_j(n)$ it will be used notation:

$$\boldsymbol{u}_j = \begin{bmatrix} \boldsymbol{u}_j^{(1)} & \cdots & \boldsymbol{u}_j^{(L)} \end{bmatrix}, \tag{10}$$

where $\boldsymbol{u}_j$ is a column vector with $\boldsymbol{u}_j^{(l)} = \begin{bmatrix} a_j^{(l,0)} & \cdots & a_j^{(l,k_l-1)} \end{bmatrix}$, $l = 1, \cdots L$. In the following two methods to compute the UC of $\pi_j(n)$ are described.

3.1 Eigenvectors Method

If the matrix $\mathbf{P}$ is diagonalizable, the Jordan blocs are reduced to scalars: $\mathbf{J}_l = \lambda_l$, and the matrix $\mathbf{P}$ admits the spectral decomposition: $\mathbf{P} = \mathbf{L}^{-1}\,\mathbf{\Lambda}\,\mathbf{L}$, where $\mathbf{\Lambda}$ is the diagonal matrix $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \cdots \lambda_k)$, and $\mathbf{L}$ is a matrix whose rows, $\mathbf{l}_1, \cdots \mathbf{l}_k$, are the left-hand eigenvectors of $\mathbf{P}$. Since $\mathbf{\Lambda}^n = \mathrm{diag}(\lambda_1^n, \cdots \lambda_k^n)$, substituting into equation (1) we have

$$\boldsymbol{\pi}(n) = \boldsymbol{\pi}(0)\,\mathbf{P}^n = \boldsymbol{\pi}(0)\,\mathbf{L}^{-1}\,\mathbf{\Lambda}^n\,\mathbf{L} = \sum_{i=1}^{k} a_i\,\lambda_i^n\,\mathbf{l}_i \tag{11}$$

where we assume $\lambda_i^n = \delta_n$ for any $\lambda_i = 0$. Defining the row vector $\mathbf{a} = \begin{bmatrix} a_1 & \cdots & a_k \end{bmatrix}$, the constants a_i can be obtained solving the system of equations:

$$\mathbf{a}\,\mathbf{L} = \boldsymbol{\pi}(0) \tag{12}$$

and, by (11), the UC of $\pi_j(n)$ (the vector $\mathbf{u}_j$) are given by the element wise product of the vector $\mathbf{a}$ that solves the system (12), and the j column of $\mathbf{L}$.

If we want to use right eigenvectors, perhaps because our numerical tool only computes them, then we can proceed as follows. Let $\mathbf{R}$ be a matrix whose columns, $\mathbf{r}_1, \cdots \mathbf{r}_k$, are the right-hand eigenvectors of $(\mathbf{P})^{\mathrm{T}}$, where $()^{\mathrm{T}}$ is the transpose operator. Then, $(\mathbf{P})^{\mathrm{T}} = \mathbf{R}\,\mathbf{\Lambda}\,\mathbf{R}^{-1}$, and:

$$(\boldsymbol{\pi}(n))^{\mathrm{T}} = \sum_{i=1}^{k} b_i\,\lambda_i^n\,\mathbf{r}_i \tag{13}$$

Defining the column vector $\mathbf{b} = \begin{bmatrix} b_1 & \cdots & b_k \end{bmatrix}$, then, the constants b_i can be obtained solving the system of equations:

$$\mathbf{R}\,\mathbf{b} = (\boldsymbol{\pi}(0))^{\mathrm{T}} \tag{14}$$

and, by (13), the vector $\boldsymbol{u}_j$ (see equation (10)) is given by the element wise product of the vector $\mathbf{b}$ that solves the system (14), and the j row of $\mathbf{R}$.

Note that $\mathbf{P}$ is diagonalizable only if the geometric multiplicity of all its eigenvalues is equal to its multiplicity (the matrix is said to be not *defective*). This is an important restriction of this method.

3.2 Vandermonde Method

The UC of $\pi_j(n)$ can be obtained solving the system of equations that results from imposing the boundary conditions to equation (3):

$$\pi_j(n) = (\boldsymbol{\pi}(0)\,\mathbf{P}^n)_j \tag{15}$$

for $n = 0,\ 1, \cdots$, up to the number of UC to be determined. If we don't exploit the fact that the geometric multiplicity of some eigenvalues may be larger that 1, then there will be k UC, where k is the number of states of the Markov chain, i.e. the size of the square matrix $\mathbf{P}$ (see remark 2.1). The notation $(\boldsymbol{x})_j$ refers to the j component of the vector $\boldsymbol{x}$. In other words, the vector $\boldsymbol{u}_j$ with the UC of $\pi_j(n)$, defined by (10), can be obtained solving the system of equations:

$$\begin{bmatrix} \mathbf{A}_1 & \cdots & \mathbf{A}_L \end{bmatrix} \boldsymbol{u}_j = \mathbf{B} \tag{16}$$

where the sub-matrices $\mathbf{A}_l^{k \times k_l}$, $l = 1, \cdots L$, ($\lambda_1 = 0$, $\lambda_l \neq 0$) are given by:

$$\mathbf{A}_1^{k \times k_1} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \cdots & & & \end{bmatrix} \tag{17}$$

$$\mathbf{A}_l^{k \times k_l} = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ \lambda_l & \lambda_l & \cdots & \lambda_l \\ \lambda_l^2 & 2\,\lambda_l^2 & \cdots & 2^{k_l-1}\,\lambda_l^2 \\ \cdots & & & \\ \lambda_l^n & n\,\lambda_l^n & \cdots & n^{k_l-1}\,\lambda_l^n \\ \cdots & & & \end{bmatrix} \tag{18}$$

and $\mathbf{B}$ is the column vector

$$\mathbf{B} = \begin{bmatrix} \pi_j(0) & (\boldsymbol{\pi}(0)\,\mathbf{P})_j & (\boldsymbol{\pi}(0)\,\mathbf{P}^2)_j \cdots \end{bmatrix} \tag{19}$$

Note that if the chain starts in state i with probability 1 (i.e., $\boldsymbol{\pi}(0)$ is a probability vector with the probability 1 in the component i), then $\mathbf{B} = [\delta_{ij}\ (\mathbf{P})_{ij}\ (\mathbf{P}^2)_{ij} \cdots]$, where δ_{ij} is the Kronecker's delta ($\delta_{ij} = 1$ for $i = j$, and 0 otherwise).

Using this approach, $\boldsymbol{u}_j$ is obtained solving a confluent Vandermonde system, for which there exist abundant literature and fast numerical methods. One example is the method of Björ and Pereyra [1]. However, for large matrices this method fails to give accurate results, due to rounding errors. Therefore, in the numerical experiments given in

section 7 it was found more convenient building the Vandermonde matrix and solving the system using a QR decomposition [11].

This second approach has the advantage over the previous one that the matrix $\mathbf{P}$ can be defective. Therefore, this approach gives a more general solution.

Remark 3.1 In order to compute the vector $\mathbf{B}$ it is not necessary the computation of the powers of the matrix $\mathbf{P}$, as it may seem from equation (19). This would be costly for large matrices. Note that equation (19) can be implemented as the product of a row vector, $\mathbf{r}$, $k-1$ times the matrix $\mathbf{P}$, as shown in the algorithm 3.1. Note also that if we are interested in $\pi_j(n)$ for $j = \{j_1, \cdots j_n\}$, then it is convenient to solve the system (16) for the matrix $\mathbf{B}^{k\times|j|}$ that would be computed using algorithm 3.1 with the indexes $j = \{j_1, \cdots j_n\}$. ■

Algorithm 3.1 Computation of vector $\mathbf{B}$.

1: $\mathbf{B}[1] \leftarrow \boldsymbol{\pi}_0[j]$
2: $\mathbf{r} \leftarrow \boldsymbol{\pi}_0$
3: **for all** i in $2 : k$ **do**
4: $\quad \mathbf{r} \leftarrow \mathbf{r}\,\mathbf{P}$
5: $\quad \mathbf{B}[i] \leftarrow \mathbf{r}[j]$
6: **end for**

In the following there are some examples of the Vandermonde method. The examples illustrate as the Vandermonde method allows obtaining the solution of small chains without the help of a computer. In these simple examples the notation is simplified, using capital letters to refer to the UC.

Example 3.1 Assume a DTMC with

$$\mathbf{P} = \begin{bmatrix} 4/5 & 1/5 & 0 \\ 0 & 4/5 & 1/5 \\ 1/5 & 1/5 & 3/5 \end{bmatrix}$$

We want the probability of being in state 3 in n steps starting from state 1 (i.e. $\pi_3(n)$ with $\boldsymbol{\pi}(0) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$). It can be easily found that the eigenvalues of $\mathbf{P}$ are $\lambda_1 = 1$ with multiplicity 1 (as expected, since $\mathbf{P}$ is an irreducible stochastic matrix), and $\lambda_2 = 3/5$ with multiplicity 2. Following equation (3) we guess the solution:

$$\begin{aligned} \pi_3(n) &= \lambda_1^n A + \lambda_2^n (B + C\,n) \\ &= A + (3/5)^n\,(B + C\,n) \end{aligned}$$

Imposing the boundary conditions:

$$\begin{aligned} &\pi_3(0) = A + B = (\mathbf{P}^0)_{13} = 0 \\ &\pi_3(1) = A + (3/5)\,(B + C) = (\mathbf{P}^1)_{13} = 0 \\ &\pi_3(2) = A + (3/5)^2\,(B + 2\,C) = (\mathbf{P}^2)_{13} = 1/25 \end{aligned}$$

we have that $A = -B = 1/4$, $C = -1/6$, thus:

$$\pi_3(n) = 1/4 - (3/5)^n\,(1/4 + n/6), \quad n \geq 0$$

It can be checked that the geometric multiplicity of the eigenvalue $\lambda_2 = 3/5$ is 1. Thus, this problem would be ill-conditioned if solved with the Eigenvectors method. ■

Example 3.2 Assume a DTMC with two transient states:

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2/3 & 1/3 \\ 0 & 0 & 1/3 & 2/3 \end{bmatrix}$$

We want the probability of being in state 3 in n steps starting from state 1 (i.e. $\pi_3(n)$ with $\boldsymbol{\pi}(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$). It can be easily found that the eigenvalues of $\mathbf{P}$ are $\lambda_1 = 0$ with multiplicity 2 (due to the two transient states), and $\lambda_2 = 1$, $\lambda_3 = 1/3$. We guess the solution:

$$\pi_3(n) = A\,\delta_n + B\,\delta_{n-1} + C + D\,(1/3)^n$$

Imposing the boundary conditions:

$$\begin{aligned} \pi_3(0) &= A + C + D &&= (\mathbf{P}^0)_{13} = 0 \\ \pi_3(1) &= B + C + D/3 &&= (\mathbf{P}^1)_{13} = 0 \\ \pi_3(2) &= C + D/3^2 &&= (\mathbf{P}^2)_{13} = 1 \\ \pi_3(3) &= C + D/3^3 &&= (\mathbf{P}^3)_{13} = 2/3 \end{aligned}$$

we have that $A = -5$, $B = -2$, $C = 1/2$, $D = 9/2$ thus:

$$\pi_3(n) = -5\,\delta_n - 2\,\delta_{n-1} + 1/2 + 9/2\,(1/3)^n, \quad n \geq 0$$

It can be checked that the geometric multiplicity of the eigenvalue $\lambda_1 = 0$ is 1. Thus, this problem would be ill-conditioned if solved with the Eigenvectors method. ■

Example 3.3 Assume a DTMC with two absorbing states:

$$\mathbf{P} = \begin{bmatrix} 3/5 & 1/5 & 1/5 & 0 \\ 1/5 & 3/5 & 0 & 1/5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

We want the probability of being in state 3 in n steps starting from state 1 (i.e. $\pi_3(n)$ with $\boldsymbol{\pi}(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$). It can be easily found that the eigenvalues of $\mathbf{P}$ are $\lambda_1 = 1$ with multiplicity 2 (due to the two absorbing states), and $\lambda_2 = 4/5$, $\lambda_3 = 2/5$. Following the remark 2.1 we guess the solution:

$$\pi_3(n) = A + B\,(4/5)^n + C\,(2/5)^n$$

Imposing the boundary conditions:

$$\begin{aligned} &\pi_3(0) = A + B + C = (\mathbf{P}^0)_{13} = 0 \\ &\pi_3(1) = A + (4/5)\,B + (2/5)\,C = (\mathbf{P}^1)_{13} = 1/5 \\ &\pi_3(2) = A + (4/5)^2\,B + (2/5)^2\,C = (\mathbf{P}^2)_{13} = 8/25 \end{aligned}$$

we have that $A = 2/3$, $B = -1/2$, $C = -1/6$, thus:

$$\pi_3(n) = 2/3 - 1/2\,(4/5)^n - 1/6\,(2/5)^n, \quad n \geq 0$$

Of course, 2/3 is the probability of being absorbed by state 3, starting from state 1. ■

4 Continuous Time Markov Chains

Let $X(t)$ be an homogeneous finite-state continuous-time Markov Chain with k states, and infinitesimal generator of size k, $\mathbf{Q}^{k\times k}$. The transient solution is given by the exponential matrix $\mathrm{e}^{\mathbf{Q}\,t}$ [4]:

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)\,\mathrm{e}^{\mathbf{Q}\,t} \tag{20}$$

where $\boldsymbol{\pi}(0)$ is the initial distribution and the components of the row vector $\boldsymbol{\pi}(t)$ are the probabilities

$$\pi_j(t) = Prob\left\{X(t) = j \;\middle|\; \boldsymbol{\pi}(0)\right\} \tag{21}$$

An expression for $\pi_j(t)$ can be obtained analogously to the discrete time case.

Theorem 4.1 Let λ_l, $l = 1, \cdots L$ be the eigenvalues of $\mathbf{Q}$, each with multiplicity k_l ($\sum_l k_l = k$), then:

$$\pi_j(t) = \sum_{l=1}^{L} \mathrm{e}^{\lambda_l\,t} \sum_{m=0}^{k_l-1} a_j^{(l,m)}\,t^m \tag{22}$$

Proof. Similarly to section 2, using the Jordan Form of the matrix $\mathbf{Q} = \mathbf{A}\,\mathbf{J}\,\mathbf{A}^{-1}$, where $\mathbf{A}$ is some invertible matrix, and $\mathbf{J} = \mathrm{diag}(\mathbf{J}_1, \mathbf{J}_2, \cdots \mathbf{J}_L)$ is a diagonal block matrix, with the Jordan blocks $\mathbf{J}_l$. We have:

$$\mathrm{e}^{\mathbf{Q}\,t} = \sum_{n=0}^{\infty} \frac{(\mathbf{Q}\,t)^n}{n!} = \mathbf{A}\,\mathrm{e}^{\mathbf{J}\,t}\,\mathbf{A}^{-1} \tag{23}$$

where $\mathrm{e}^{\mathbf{J}\,t} = \mathrm{diag}(\mathrm{e}^{\mathbf{J}_1\,t}, \mathrm{e}^{\mathbf{J}_2\,t}, \cdots \mathrm{e}^{\mathbf{J}_L\,t})$. To compute the exponential matrices of the Jordan blocks, we write: $\mathbf{J}_l = \lambda_l\,\mathbf{I} + \mathbf{U}_1^{k_l\times k_l}$, where the matrix $\mathbf{U}_1^{k_l\times k_l}$ is as defined in section 2. Thus:

$$\begin{aligned}
\mathrm{e}^{\mathbf{J}_l\,t} = \sum_{n=0}^{\infty} \frac{(\mathbf{J}_l\,t)^n}{n!} = \sum_{n=0}^{\infty} \frac{t^n}{n!}(\lambda_l\,\mathbf{I} + \mathbf{U}_1^{k_l\times k_l})^n = \\
\sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{i=0}^{n} \binom{n}{i} \lambda_l^{n-i}\,(\mathbf{U}_1^{k_l\times k_l})^i = \\
\sum_{i=0}^{\infty}\sum_{n=i}^{\infty} \frac{(\lambda_l\,t)^{n-i}}{n!} \binom{n}{i}\,(\mathbf{U}_1^{k_l\times k_l}\,t)^i = \\
\sum_{i=0}^{\infty} \frac{(\mathbf{U}_1^{k_l\times k_l}\,t)^i}{i!} \sum_{j=0}^{\infty} \frac{(\lambda_l\,t)^j}{j!} = \\
\mathrm{e}^{\lambda_l\,t} \sum_{i=0}^{k_l-1} \frac{(\mathbf{U}_1^{k_l\times k_l})^i\,t^i}{i!}
\end{aligned} \tag{24}$$

where it was used the change $j = n - i$, and the fact that $(\mathbf{U}_1^{k_l\times k_l})^i = \mathbf{0}$ if $i \geq k_l$. Using (24), we conclude that the elements of the matrix $\mathrm{e}^{\mathbf{Q}\,t} = \mathbf{A}\,\mathrm{e}^{\mathbf{J}\,t}\,\mathbf{A}^{-1}$ will be linear combinations of $\mathrm{e}^{\lambda_l\,t}$ times polynomials in t of degree equal, at most, to $k_l - 1$, as stated in equation (22). □

Remark 4.1 Note that, in contrast to case of a DTMC, now it is not necessary a special attention to the eigenvalue $\lambda_l = 0$. In fact, since $\mathrm{e}^{\lambda_l\,t} = 1$ for $\lambda_l = 0$, the UC associated with $\lambda_l = 0$ in a CTMC is the only one that do not vanish when $t \to \infty$. Thus, $\lambda_l = 0$ is for a CTMC equivalent to what the eigenvalue $\lambda_l = 1$ is for a DTMC. Additionally, the remark 2.1, regarding to the eigenvalue $\lambda_1 = 1$ in a DTMC, has its counterpart for the eigenvalue $\lambda_1 = 0$ in a CTMC: Assume that the Markov chain has the eigenvalue $\lambda_1 = 0$ with multiplicity $k_1 > 1$ (see appendix A). Clearly, the polynomial corresponding to this eigenvalue:

$$p_j^{(1)}(n) = \sum_{m=0}^{k_1-1} a_j^{(1,m)}\,t^m \tag{25}$$

must have the coefficients $a_j^{(1,m)} = 0$, for $m = 1, \cdots, k_1 - 1$. Otherwise, the solution (22) will be unbounded. Therefore, for the eigenvalue $\lambda_1 = 0$ there will be only 1 undetermined coefficient ($p_j^{(1)}(n) = a_j^{(1,0)}$). ■

5 Undetermined Coefficients of a CTMC

In equation (22) there are up to k UCs $a_j^{(l,m)}$ of $\pi_j(t)$ to be determined. The two methods described in section 3 can be applied to obtain the UC of $\pi_j^{(l)}(t)$. The Eigenvectors method described in section 3.1 is almost identical for a CTMC, thus, it will not be repeated here. The only difference is that λ_i^n in equations (11) and (13), are now $\mathrm{e}^{\lambda_i\,t}$ (and, in contrast to the DTMC, it is not necessary a special attention to the eigenvalue $\lambda_l = 0$).

5.1 Vandermonde Method

The UC of $\pi_j(t)$ can be obtained solving the system of equations that results from imposing the boundary conditions to equation (22):

$$\frac{\partial^n \pi_j(0)}{\partial t^n} = (\boldsymbol{\pi}(0)\,\mathbf{Q}^n)_j \tag{26}$$

for $n = 0, 1, \cdots$, up to the number of UC to be determined. As in the discrete case, the vector $\boldsymbol{u}_j$ with the UC of $\pi_j(t)$, defined by (10), can be obtained solving the system of equations: $\begin{bmatrix}\mathbf{A}_1 & \cdots & \mathbf{A}_L\end{bmatrix}\boldsymbol{u}_j = \mathbf{B}$, where the sub-matrices $\mathbf{A}_l^{k\times k_l}$, $l = 1, \cdots L$ are given by (see appendix B):

$$\mathbf{A}_l^{k\times k_l} = \begin{bmatrix}
1 & 0 & \cdots & 0 \\
\lambda_l & 1 & \cdots & 0 \\
\lambda_l^2 & 2\,\lambda_l & \cdots & 0 \\
\cdots & & & \\
\lambda_l^n & n\,\lambda_l^{n-1} & \cdots & n^{\underline{k_l-1}}\,\lambda_l^{n-(k_l-1)} \\
\cdots & & &
\end{bmatrix} \tag{27}$$

where $n^{\underline{m}} = n\,(n-1)\cdots(n-m+1)$, $n^{\underline{m}} = 0$ for $n < m$, and $\mathbf{B}$ is the column vector:

$$\mathbf{B} = \begin{bmatrix}\pi_j(0) & (\boldsymbol{\pi}(0)\,\mathbf{Q})_j & (\boldsymbol{\pi}(0)\,\mathbf{Q}^2)_j \cdots\end{bmatrix} \tag{28}$$

Note that if the chain starts in state i with probability 1 (i.e., $\boldsymbol{\pi}(0)$ is a probability vector with the probability 1 in the component i), then $\mathbf{B} = [\delta_{ij}\ (\mathbf{Q})_{ij}\ (\mathbf{Q}^2)_{ij} \cdots]$. Note also

that the vector $\mathbf{B}$ given by (28) as exactly the same form than in the DTMC (equation (19)). Thus, the remark 3.1 regarding the computation of the vector $\mathbf{B}$ for the DTMC, is applicable now changing the matrix $\mathbf{P}$ by $\mathbf{Q}$.

Example 5.1 Assume a CTMC with

$$\mathbf{Q} = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 1 & -2 \end{bmatrix}$$

We want the probability of being in state 3 at time t starting from state 1 (i.e. $\pi_3(t)$ with $\boldsymbol{\pi}(0) = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$).

It can be easily found that the eigenvalues of $\mathbf{Q}$ are $\lambda_1 = 0$ with multiplicity 1 (as expected, since $\mathbf{Q}$ is irreducible), and $\lambda_2 = -2$ with multiplicity 2. Following equation (22) we guess the solution:

$$\begin{aligned} \pi_3(t) &= \mathrm{e}^{\lambda_1 t} A + \mathrm{e}^{\lambda_2 t}(B + C\,t) \\ &= A + \mathrm{e}^{-2\,t}\,(B + C\,t) \end{aligned}$$

Imposing the boundary conditions:

$$\begin{aligned} \pi_3(0) &= A + B = (\mathbf{Q}^0)_{13} = 0 \\ \pi_3'(0) &= -2\,B + C = (\mathbf{Q}^1)_{13} = 0 \\ \pi_3''(0) &= (-2)^2\,B + 2\,(-2)\,C = (\mathbf{Q}^2)_{13} = 1 \end{aligned}$$

we have that $A = -B = 1/4$, $C = -1/2$, thus:

$$\pi_3(t) = 1/4 - \mathrm{e}^{-2\,t}\,(1/4 + t/2), \quad t \geq 0$$

It can be checked that the geometric multiplicity of the eigenvalue $\lambda_2 = -2$ is 1. Thus, this problem would be ill-conditioned if solved with the Eigenvectors method. ■

Example 5.2 Assume a CTMC with two absorbing states:

$$\mathbf{Q} = \begin{bmatrix} -5 & 4 & 1 & 0 \\ 4 & -5 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

We want the probability of being in state 3 at time t starting from state 1 (i.e. $\pi_3(t)$ with $\boldsymbol{\pi}(0) = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$). It can be easily found that the eigenvalues of $\mathbf{Q}$ are $\lambda_1 = 0$ with multiplicity 2 (due to the two absorbing states), and $\lambda_2 = -1$, $\lambda_3 = -9$. Following the remark 4.1 we guess the solution:

$$\pi_3(t) = A + B\,\mathrm{e}^{-t} + C\,\mathrm{e}^{-9\,t}$$

Imposing the boundary conditions:

$$\begin{aligned} \pi_3(0) &= A + B + C = (\mathbf{Q}^0)_{13} = 0 \\ \pi_3'(0) &= -B - 9\,C = (\mathbf{Q}^1)_{13} = 1 \\ \pi_3''(0) &= (-1)^2\,B + (-9)^2\,C = (\mathbf{Q}^2)_{13} = -5 \end{aligned}$$

we have that $A = 5/9$, $B = -1/2$, $C = -1/18$, thus:

$$\pi_3(t) = 5/9 - 1/2\,\mathrm{e}^{-\,t} - 1/18\,\mathrm{e}^{-9\,t}, \quad t \geq 0$$ ■

6 Uniformized Vandermonde Method

Computational roundoff errors may lead the CTMC solution methods described in section 5 to give inaccurate results, specially for chains with a large number of states. The Eigenvectors method fails, particularly for asymmetric matrices, because the numerical tool may not be able to obtain linearly independent eigenvectors (see [7]). In case of the Vandermonde method, because the eigenvalues of the matrix $\mathbf{Q}$ may be out of the unit circle, and thus, the matrix (27) used to compute the undetermined coefficients may have a large norm and be ill-conditioned. This problem can be alleviated using the well known uniformization method (see e.g. [9, 12]). This method consists of considering the *uniformized matrix*:

$$\mathbf{P} = \mathbf{I} + \frac{1}{q}\mathbf{Q} \tag{29}$$

where q is a constant such that $\max_i |(\mathbf{Q})_{ii}| \leq q < \infty$, and using the equation (conveniently truncated):

$$\boldsymbol{\pi}(t) = \boldsymbol{\pi}(0)\,\mathrm{e}^{\mathbf{Q}\,t} = \sum_{n=0}^{\infty} \mathrm{e}^{-q\,t}\frac{(q\,t)^n}{n!}\,\boldsymbol{\pi}(0)\,\mathbf{P}^n \tag{30}$$

Note that the matrix $\mathbf{P}$ defined by (29) is stochastic, and the equation (30) can be easily proved by direct substitution of the definition of the matrix $\mathbf{P}$ (29). The problem of directly applying the uniformization formula (30) is that computing the powers $\mathbf{P}^n$ can be costly for large number of states. Additionally, storing these powers will be infeasible for large matrices, thus, the powers $\mathbf{P}^n$ will have to be computed for each time t that we wish to evaluate.

To cope with the above mentioned problems we can use the DTMC defined by $\mathbf{P}$ to compute the *uniformized chain probabilities* $\pi_j^{(P)}(n)$ for each of the states j we are interested, as described in section 2. Thorough this section the indexes $^{(P)}$ and $^{(Q)}$ will be used to distinguish the uniformized chain and the original CTMC chain with infinitesimal generator $\mathbf{Q}$, respectively. Note that the initial conditions, $\boldsymbol{\pi}(0)$, are the same for both chains. So, $^{(P)}$ $^{(Q)}$ will not be used with $\boldsymbol{\pi}(0)$. From equation (30) we have:

$$\pi_j^{(Q)}(t) = \sum_{n=0}^{\infty} \mathrm{e}^{-q\,t}\frac{(q\,t)^n}{n!}\,\pi_j^{(P)}(n) \tag{31}$$

Let $\lambda_l^{(P)}$, $l = 1, \cdots L$ be the eigenvalues of the uniformized matrix $\mathbf{P}$ given by equation (29), each with multiplicity k_l. Without loss of generality, let $\lambda_1^{(P)}$ be the eigenvalue equal to 1, and $\lambda_2^{(P)}$ a possible eigenvalue equal to 0. Let $b_j^{(l,m)}$ be the UC of $\pi_j^{(P)}(n)$. Using equation (3) and

remark 2.1, and substituting into (31) we have:

$$\pi_j^{(Q)}(t) = \sum_{n=0}^{\infty} \mathrm{e}^{-q\,t}\frac{(q\,t)^n}{n!}\left[b_j^{(1,0)} + \sum_{m=0}^{k_2-1} b_j^{(2,m)}\delta_{n-m} + \sum_{l=3}^{L}(\lambda_l^{(P)})^n \sum_{m=0}^{k_l-1} b_j^{(l,m)}\, n^m\right] =$$

$$b_j^{(1,0)} + \mathrm{e}^{-q\,t}\left[b_j^{(2,0)} + \sum_{m=1}^{k_2-1}\frac{(q\,t)^m}{m!}\, b_j^{(2,m)}\right] + \sum_{l=3}^{L}\mathrm{e}^{-q\,t\,(1-\lambda_l^{(P)})}\left[b_j^{(l,0)} + \sum_{m=1}^{k_l-1} b_j^{(l,m)}\, q^{(l,m)}(t)\right] \quad (32)$$

where $q^{(l,m)}(t)$ are polynomials in t of degree m (see appendix C):

$$q^{(l,m)}(t) = \sum_{i=1}^{m} q_i^{(m)}\,(q\,\lambda_l^{(P)})^i\,t^i \quad (33)$$

with the coefficients $q_i^{(m)}$ given by the recurrence relation:

$$q_i^{(m)} = \begin{cases} 1, & i = 1 \\ \sum_{k=i-1}^{m-1}\binom{m-1}{k}\, q_{i-1}^{(k)}, & i = 2,\cdots m \end{cases} \quad (34)$$

Let $\lambda_l^{(Q)}, l = 1,\cdots L$ be the eigenvalues of the infinitesimal generator $\mathbf{Q}$, and $a_j^{(l,m)}$ the UC of $\pi_j^{(Q)}(t)$. Comparing equation (22) and (32), it turns out that it must be:

$$\lambda_l^{(Q)} = -q\,(1-\lambda_l^{(P)}) \quad (35)$$

$$\lambda_l^{(P)} = 1 + \frac{\lambda_l^{(Q)}}{q} \quad (36)$$

and $a_j^{(l,0)} = b_j^{(l,0)}$, $l = 1,\cdots L$. Thus, if there are not confluent eigenvalues, the UC of $\pi_j^{(P)}(n)$ and $\pi_j^{(Q)}(t)$ are the same (note that thorough the paper it is used *confluent* and *multiple* eigenvalues interchangeably). In the confluent case, from equations (32) and (33) we have that:

$$a_j^{(2,m)} = \frac{q^m}{m!}\, b_j^{(2,m)},\ m = 1,\cdots k_2 - 1 \quad (37)$$

$$a_j^{(l,m)} = (q\,\lambda_l^{(P)})^m \sum_{k=m}^{k_l-1} q_m^{(k)}\, b_j^{(l,k)},\ m = 1,\cdots k_l - 1 \quad (38)$$

Using the definition of eigenvector it is easy to prove that equations (35) and (36) imply that the eigenvectors of $\mathbf{Q}$ and its uniformized matrix $\mathbf{P}$ given by (29), must be the same. Therefore, if the system of equations (14) is ill-conditioned for the eigenvectors of $\mathbf{Q}$, so it will be for the eigenvectors of $\mathbf{P}$. Thus, there is no much advantage in using the Eigenvectors method with the uniformized matrix $\mathbf{P}$ (except that the numerical tool may compute the eigenvalues and eigenvectors more accurately using $\mathbf{P}$ than $\mathbf{Q}$).

On the other hand, using the Vandermonde method with the uniformized matrix $\mathbf{P}$ implies building the Vandermonde matrix using $\lambda_l^{(P)}$. This makes a big difference with respect to the Vandermonde matrix that would be obtained using $\lambda_l^{(Q)}$. In the following the Vandermonde matrices obtained using $\lambda_l^{(P)}$ and $\lambda_l^{(Q)}$ will be denoted as $\mathbf{V}^{(P)}$ and $\mathbf{V}^{(Q)}$, respectively. Notice that $\lambda_l^{(P)}$ must be inside the unit disk, thus, $\mathbf{V}^{(P)}$ will be possibly much better conditioned than $\mathbf{V}^{(Q)}$. Additionally, by choosing an appropriate value for q, we can tune the condition number of $\mathbf{V}^{(P)}$.

Summing up, the *Uniformized Vandermonde Method* proposed in this paper consists of the following steps:

1. Compute the eigenvalues of $\mathbf{Q}$, $\lambda_l^{(Q)}$.
2. Process the eigenvalues $\lambda_l^{(Q)} < 0$, merging those that are *near confluent*, if any, and limiting the maximum multiplicity (see remark 6.2). The multiplicity of the eigenvalue $\lambda_1^{(Q)} = 0$, k_1, should be determined by the number of irreducible closed sets (see appendix A and remark 4.1).
3. Choose an appropriate value for the uniformization parameter q (see section 6.1).
4. Compute the uniformized matrix $\mathbf{P} = \mathbf{I} + \frac{1}{q}\mathbf{Q}$. To save memory space, the matrix $\mathbf{P}$ can overwrite matrix $\mathbf{Q}$ in this step, since $\mathbf{Q}$ is not needed anymore.
5. Using $\mathbf{P}$ and $\boldsymbol{\pi}(0)$, compute the matrix $\mathbf{B}$ using the algorithm 3.1 (page 4). Then, the matrix $\mathbf{P}$ can be removed, since it is not needed anymore.
6. Compute the eigenvalues of $\mathbf{P}$, $\lambda_l^{(P)}$, using equation (36).
7. Using $\lambda_l^{(P)}$, construct the Vandermonde matrix, as explained in section 3.2.
8. Solve the resulting Vandermonde system to obtain the UC of $\pi_j^{(P)}(n)$. If there aren't confluent eigenvalues, these are the UC of $\pi_j^{(Q)}(t)$, and we are done. In case of confluent eigenvalues, use equations (37) and (38) to compute the remaining UC of $\pi_j^{(Q)}(t)$.
9. Use UC of $\pi_j^{(Q)}(t)$ and the eigenvalues $\lambda_l^{(Q)}$ in equation (22) to evaluate the desired transient solution.

Remark 6.1 It is interesting to note that if the uniformized chain probabilities $\pi_j^{(P)}(n)$ given by the matrix $\mathbf{P}$ (29) have a limiting distribution, they converge to $\pi_j^{(Q)}(t)$ at the points $n = [q\,t]$, where $[x]$ stands for the integral value of x. This comes from the fact:

$$\lim_{n\to\infty}\mathbf{P}^n = \lim_{n\to\infty}\left(\mathbf{I} + \frac{1}{q}\mathbf{Q}\right)^n = \lim_{n\to\infty}\left(\mathbf{I} + \frac{t\,\mathbf{Q}}{n}\right)^n = \mathrm{e}^{\mathbf{Q}\,t} \quad (39)$$

Figure 1 gives a pictorial example. The figure shows $\pi_j^{(Q)}(q\,t)$, $j = 0,\cdots N$ for the M/M/1/N queue analyzed

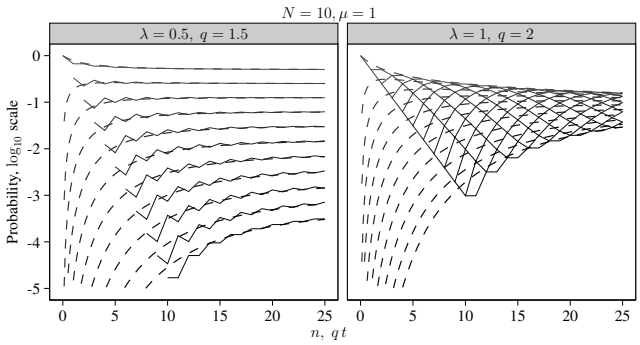

Figure 1: $\pi_j^{(Q)}(t)$ of an M/M/1/10 queue (dashed lines) and their uniformized chain probabilities $\pi_j^{(P)}(n)$ (solid lines).

in section 7, and $\pi_j^{(P)}(n)$. Here, j is the number in the system, and the initial state is $j = 0$. The points of $\pi_j^{(P)}(n)$ are connected with solid lines, while $\pi_j^{(Q)}(q\,t)$ is depicted with dashed lines. Note that $\pi_j^{(P)}(n) = 0$ for $n < j$, since it takes at least j steps to move from state 0 to j. The parameters are $N = 10$, $\mu = 1$, and two values of λ: 0.5 and 1. The uniformization parameter has been taken as $q = \max_i |(\mathbf{Q})_{ii}| = \lambda + \mu$. Recall that the limiting distribution is given by:

$$\pi_j^{(Q)}(\infty) = \begin{cases} \frac{1}{N+1}, & \lambda = \mu \\ \frac{(1-\rho)\,\rho^j}{1-\rho^{N+1}}, & \lambda \neq \mu \end{cases} \tag{40}$$

where $\rho = \lambda/\mu$. Clearly, in this case the limiting distribution is the same for both $\pi_j^{(P)}(n)$ and $\pi_j^{(Q)}(t)$. If the uniformized chain probabilities $\pi_j^{(P)}(n)$ do not have a limiting distribution, e.g. because it is periodic, then (39) does not apply. However, equation (39) was not used to derive the former equations in this paper, and the Uniformized Vandermonde method explained before. Thus, the previous results are not affected by this fact. The following example illustrates the Uniformized Vandermonde method, and shows a uniformized chain having no limiting distribution. ■

Example 6.1 Assume a CTMC with

$$\mathbf{Q} = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix}$$

We want the probability of being in state $\pi_j^{(Q)}(t)$, $j = 1, 2$ starting from state 1 ($\boldsymbol{\pi}(0) = \begin{bmatrix} 1 & 0 \end{bmatrix}$), using the Uniformized Vandermonde method. The eigenvalues of $\mathbf{Q}$ are $\lambda_1^{(Q)} = 0$, $\lambda_2^{(Q)} = -2$. Choosing the uniformization parameter $q = 1$ we have from (36) $\lambda_1^{(P)} = 1$ and $\lambda_2^{(P)} = -1$, and an uniformized matrix:

$$\mathbf{P} = \mathbf{I} + \frac{1}{q}\mathbf{Q} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

which is clearly periodic. The Vandermonde matrix of the uniformized chain is:

$$\mathbf{V}^{(P)} = \begin{bmatrix} 1 & 1 \\ \lambda_1^{(P)} & \lambda_2^{(P)} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

and the matrix $\mathbf{B}$ (see section 3.2):

$$\mathbf{B} = \begin{bmatrix} \pi_1(0) & \pi_2(0) \\ (\boldsymbol{\pi}(0)\,\mathbf{P})_1 & (\boldsymbol{\pi}(0)\,\mathbf{P})_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus, the UC of the uniformized chain are given by:

$$\begin{bmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 \end{bmatrix} = (\mathbf{V}^{(P)})^{-1}\,\mathbf{B} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

Since the eigenvalues are not confluent, the UC of $\pi_j^{(P)}(n)$ and $\pi_j^{(Q)}(t)$ are the same, so we have:

$$\begin{cases} \pi_1^{(P)}(n) = 1/2 + 1/2\,(-1)^n \\ \pi_2^{(P)}(n) = 1/2 - 1/2\,(-1)^n \end{cases} \tag{41}$$

$$\begin{cases} \pi_1^{(Q)}(t) = 1/2 + 1/2\,\mathrm{e}^{-2\,t} \\ \pi_2^{(Q)}(t) = 1/2 - 1/2\,\mathrm{e}^{-2\,t} \end{cases} \tag{42}$$

In this case the uniformized chain does not have a limiting distribution, and the equation (39) does not apply: Clearly, equations (41) do not converge to equations (42) at the points $n = [q\,t]$. ■

Remark 6.2 Having eigenvalues of $\mathbf{P}$ *near confluent* will make the matrix $\mathbf{V}^{(P)}$ to have repeated rows, and thus, be ill-conditioned. This may occur due to rounding errors, even if the exact eigenvalues are properly spaced. By near confluent we mean eigenvalues $\lambda_a^{(P)}$ and $\lambda_b^{(P)}$ such that $|\Re\lambda_a^{(P)} - \Re\lambda_b^{(P)}| < \epsilon$ and $|\Im\lambda_a^{(P)} - \Im\lambda_b^{(P)}| < \epsilon$, where $\Re$ and $\Im$ are the real and imaginary part respectively, and ϵ is a small positive number. To solve this problem it has been proceed as follows: First, all eigenvalues satisfying the above inequality, have been merged into confluent eigenvalues. Additionally, all complex conjugate eigenvalues pairs having $|\Im\lambda_b^{(P)}| < \epsilon$ have been converted into its real part. In the numerical experiments of section 7 it was found convenient to use $\epsilon = ((1+\epsilon_m)^k - 1)^{0.5}$, where ϵ_m is the smallest positive floating-point number such that $1 + \epsilon_m > 1$ (the variable `.Machine$double.eps` in the numerical tool R [8], used in the experiments), and k is the number of rows of $\mathbf{V}^{(P)}$.

Another problem may arise in case of having eigenvalues with large multiplicity. Due to rounding errors, this may happen even if the original matrix has all its eigenvalues single. This could produce huge elements in $\mathbf{V}^{(P)}$ (see (18)), and the system would become ill-conditioned. To cope with this problem, the multiplicity of the eigenvalues can been limited to a maximum value M. In other words, if the multiplicity k_l of an eigenvalue $\lambda_l^{(P)}$ is $k_l > M$, then it is assumed that the UC $a_j^{(l,m)} = 0$ for $m = M, \cdots k_l - 1$. Note that this is equivalent to assume that the geometric multiplicity of the eigenvalue is $> k_l - M$ (see remark 2.1). An additional advantage of doing so, is that the size of the Vandermonde system to solve is reduced. In the numerical experiments of section 7 it has been set $M = 5$. ■

6.1 Choosing the uniformization parameter

It is common to take the uniformization parameter $q = \max_i |(\mathbf{Q})_{ii}|$. This is the minimum value that q can have

for the matrix $\mathbf{P}$ to be stochastic. However, in order improve the condition number of the Uniformized Vandermonde matrix, $\mathbf{V}^{(P)}$, it may be better choosing a higher value for q (note from (36) that taking $q \to \infty$, all eigenvalues $\lambda_l^{(P)} \to 1$). First, it has been observed in the numerical experiments that $\mathbf{V}^{(P)}$ is better conditioned having all the elements in the non negative x-plane. From equation (36) we have that this goal can be achieved if $q \geq |\min_l \Re\lambda_l^{(Q)}|$, where $\min_l \Re\lambda_l^{(Q)}$ is the most negative real part of the eigenvalues of $\mathbf{Q}$. Additionally, if all the eigenvalues $|\lambda_l^{(P)}| < 1$ are close to the origin, then the terms $(\lambda_l^{(P)})^n$ of $\mathbf{V}^{(P)}$ will vanish when increasing n, and $\mathbf{V}^{(P)}$ will be ill-conditioned. To solve this problem we can proceed as follows. Let k be the number of rows of $\mathbf{V}^{(P)}$, and $\lambda_m^{(P)}$ the second largest eigenvalue, in modulus (recall that the largest is $\lambda_1^{(P)} = 1$). A rule of thumb is choosing q such that $|\lambda_m^{(P)}|^k > \epsilon$, for some small positive ϵ. In the numerical experiments of section 7 it was used $\epsilon = \epsilon_m$, where ϵ_m is the smallest positive floating-point number such that $1 + \epsilon_m > 1$. From (36) we have that $\lambda_m^{(P)} = 1 + \frac{\lambda_m^{(Q)}}{q}$, where $\lambda_m^{(Q)}$ is the second smallest eigenvalue of $\mathbf{Q}$, in modulus (the smallest is $\lambda_m^{(Q)} = 0$). Thus, we have that it must be: $q > |\lambda_m^{(Q)}|/(1 - \epsilon^{1/k})$. Putting all together, the uniformization parameter used in the numerical experiments of section 7 has been chosen has:

$$q = \max\left\{\max_i |(\mathbf{Q})_{ii}|, |\min_l \Re\lambda_l^{(Q)}|, \frac{|\lambda_m^{(Q)}|}{1 - \epsilon^{1/k}}\right\} \quad (43)$$

7 Numerical Experiments

In this section three different methods to compute the transient solution of a CTMC with infinitesimal generator $\mathbf{Q}$ are compared. The following notation is used:

- *qevec*: Eigenvectors method (see section 5).
- *qvand*: Vandermonde method, (see section 5.1).
- *uvand*: Uniformized Vandermonde method, (see section 6).

The results have been computed using the R numerical tool [8] version 2.11.1, with its internal lapack and blas libraries. The experiments were done in a PC with a 64 bits Intel Xeon Dual-Core 2.3 GHz, and 12 GB of RAM.

The experiments have been done using an M/M/1/N queue, which is one of the few queues for which there exist a close formula for the transient solution. $N \geq 1$ is the system capacity. The state of the Markov chain is the number in the system (so, thorough this section $\pi_j(t)$, $0 \leq j \leq N$ is the probability of the queue having j customers in the system at time t), and we assume that the initial state is $j = 0$. The system capacity (N), has been varied between 10 and 10^4 (note that the number of states of the Markov chain is $N + 1$). In the interval $[10, 100]$ all values of N were computed, since each computation took few seconds. In the interval $(100, 10^4]$ there were taken 20 points evenly spaced in log scale (10 per decade). The service rate has been set to $\mu = 1$ and the arrival rate have been set to three different values: $\lambda = 1$, 10^{-3} and 10^{-6}.

The methods are compared against the formula proposed by Sharma and Tarabia in [10]. For the sake of completeness, the formula has been included in appendix D. Numerical evaluation of Sharma's formula gives very accurate results for values of t up to 100 s approximately (see appendix D).

In order to estimate the error for each of the methods $m = \{$*qevec, qvand, uvand*$\}$ it has been proceed as follows. $\pi_0^{(m)}(t)$ have been evaluated at 60 values of t evenly spaced in log scale in the range $(10^{-1}, 10^2]$ (20 points per decade). All values $\pi_0^{(m)}(t) < 0$ and $\pi_0^{(m)}(t) > 1$ where considered as failed. If the number of failed points was larger than 12 (more than 20% of failed points), it was considered that the method failed. For the non-failed points, n, it was computed:

$$\text{error}^{(m)} = \frac{1}{n}\sum_{i=1}^{n} \frac{(\pi_0^{(m)}(t_i) - \pi_0^{(sh)}(t_i))^2}{\pi_0^{(sh)}(t_i)} \quad (44)$$

where $\pi_0^{(sh)}$ is the probability obtained with Sharma's formula.

Figure 2 shows the error obtained in each scenario. The failed points, or those points where the solver failed to compute the UC, are marked with a dot at error=1. Figure 3 shows the condition number of the matrix used to solve the UC: The Eigenvectors matrix for $qevec$, and the Vandermonde matrices $\mathbf{V}^{(Q)}$ and $\mathbf{V}^{(P)}$ for $qvand$ and $uvand$ respectively (see section 6). In figure 3 the points located at $y = 1$ mark the scenarios where the solver failed to compute the UC (because rounding errors made the numerical tool finding the matrix singular). Both figures are in log-log scale.

Figure 2 shows that $qevec$ is very accurate for $\lambda = 1$. This is because $\mathbf{Q}$ is symmetric for $\lambda = 1$, and it is known that Eigenvectors method works very well in this case (see e.g. [7]). In fact figure 3 shows that the condition number for $\lambda = 1$ in $qevec$ is constant and equal to 1. For $\lambda \neq 1$ figure 3 shows that the condition number for $qevec$ increases rapidly with increasing N. In fact, the lower is λ, the less symmetric is $\mathbf{Q}$, and the smaller are the values of N for which $qevec$ is able to solve the UC. Regarding the method $qvand$, figure 3 shows that the condition number increases rapidly with N for $\lambda = 1$ and 10^{-3}, explaining the bad results observed for this method in figure 2 for these values of λ. This is because the eigenvalues of $\mathbf{Q}$ are out of the unit circle, and thus, the norm of $\mathbf{V}^{(Q)}$ increases rapidly with increasing N. This can be observed in figure 4. This figure shows a complex-plane plot of the eigenvalues obtained with the numerical tool for the matrix $\mathbf{Q}$, and the uniformized matrix $\mathbf{P}$, for $N = 1000$.

Finally, figure 2 shows that $uvand$ is the only method that succeeds to solve for all values of λ and N. Figure 3 shows that the condition number of $\mathbf{V}^{(P)}$ is between 10^{20} and 10^{25} for most of the values of N. This is a large condition number, however, figure 2 shows that the error is very

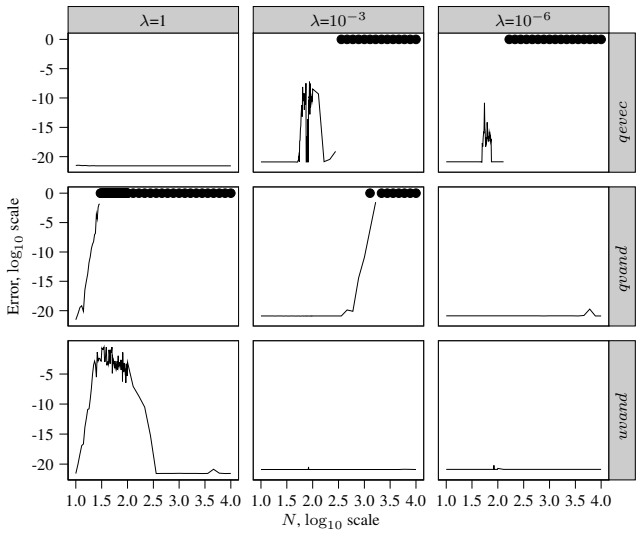

Figure 2: Error.

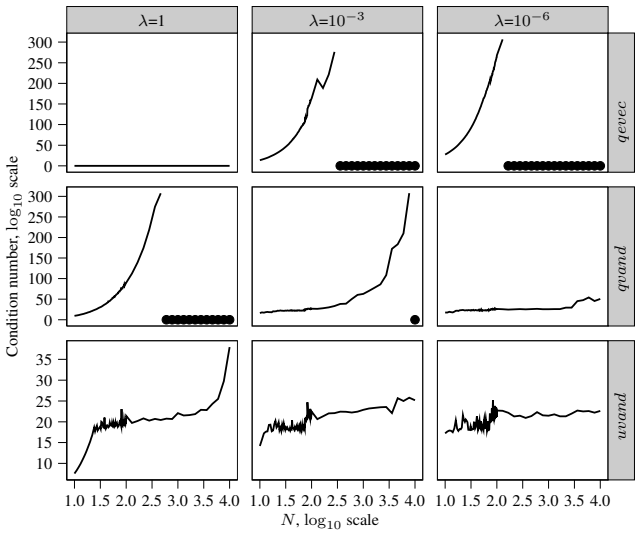

Figure 3: Condition number.

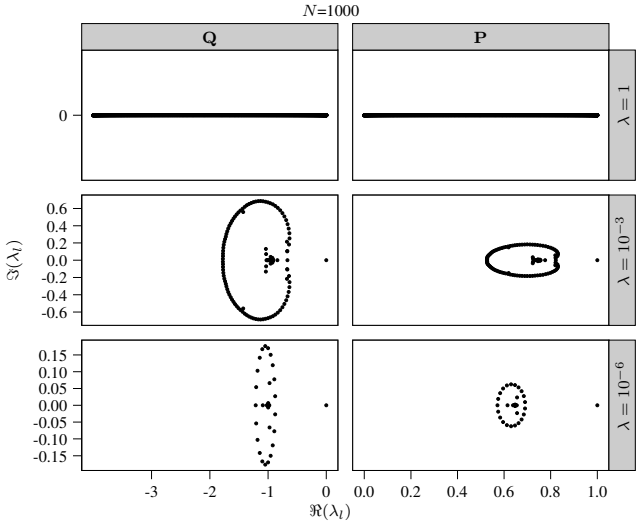

Figure 4: Complex-plane plot of the eigenvalues of the matrices $\mathbf{Q}$ and $\mathbf{P}$ for $N = 1000$.

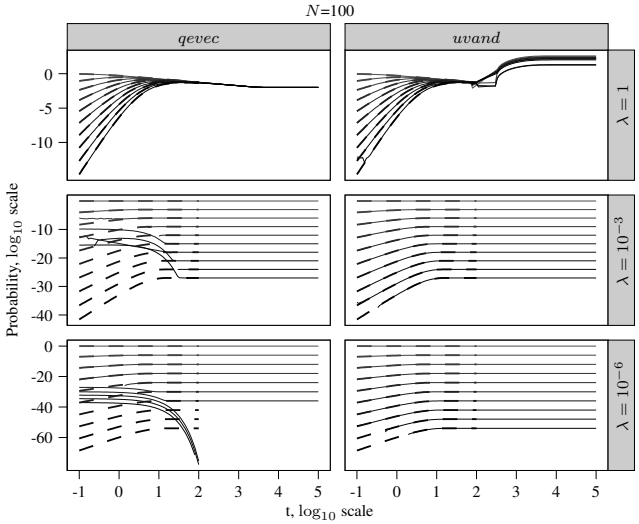

Figure 5: Probabilities $\pi_j(t)$, $j = \{0, \cdots 9\}$ for the $qevec$ and $uvand$ methods for $N = 100$. Values obtained with Sharma's formula are plotted with dashed lines.

low for $\lambda \neq 1$.

To explain what happens with $uvand$ when $\lambda = 1$, figure 5 depicts the probabilities $\pi_j^{(m)}(t)$, $j = 0, \cdots 9$ for the scenarios $m = \{$*qevec, uvand*$\}$ and $\lambda = \{1,\ 10^{-3},\ 10^{-6}\}$ when $N = 100$. In the interval $t \in [10^{-1},\ 10^2]$ the values obtained with Sharma's formula are plotted with dashed lines (note that the error depicted in figure 2 was computed in this interval). Figure 5 shows that $\pi_j^{(uvand)}(t)$ starts diverging when t is approximately 100 s. In fact, the error observed in figure 2 in $uvand$ method for N between 10 and $10^{2.5}$, is because there were taken some samples of $\pi_j^{(uvand)}(t)$ in the region where it diverges. For $N > 10^{2.5}$ the error is negligible because the divergence of $\pi_j^{(uvand)}(t)$ starts after the interval considered in the error estimation.

The intuitive explanation of the divergence of $\pi_j^{(uvand)}(t)$ when $\lambda = 1$ is the following. The Uniformized Vandermonde method can be interpreted as a fitting of the points of the uniformized probabilities, $\pi_j^{(P)}(n)$, to the samples $\mathbf{P}^n$. In the interval where these samples are taken, the corresponding probabilities obtained for $\pi_j^{(Q)}(t)$ are very accurate (recall from 6.1 that $\pi_j^{(P)}(n)$ converges to $\pi_j^{(Q)}(t)$ at the points $n = q\,t$). If $\pi_j^{(P)}(n)$ reaches the stationary distribution within the *sampling interval* $n = 0, 1, \cdots k-1$, then $uvand$ is very accurate for all values of t. This fact holds for $\lambda = \{10^{-3},\ 10^{-6}\}$, but for $\lambda = 1$ the probabilities $\pi_j^{(Q)}(t)$ converge to the stationary regime very slowly.

The duration of of the transient regime of $\pi_j^{(Q)}(t)$ is related with the second smallest eigenvalue, in modulus: $\lambda_l^* = \min_{l,\lambda_l \neq 0} |\lambda_l^{(Q)}|$. If λ_l^* is close to 0, then $\mathrm{e}^{-\lambda_l^*\, t}$ will vanish for large t, thus, with a long *transient* regime. Define

$$n^* = q\, t^* = \frac{0.1\, q}{\min_{l,\lambda_l \neq 0} |\lambda_l^{(Q)}|} \tag{45}$$

Note that $\mathrm{e}^{-\lambda_l^*\, t^*} = \mathrm{e}^{-0.1} \approx 0.9$. Thus, t^* is an estimation of the time where $\pi_j^{(Q)}(t)$ reaches the 10 % of its stationary

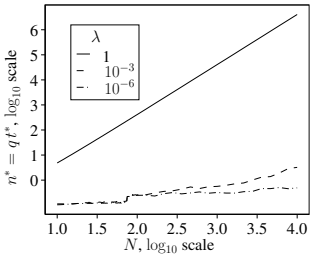

Figure 6: Values $n^* = q\,t^*$.

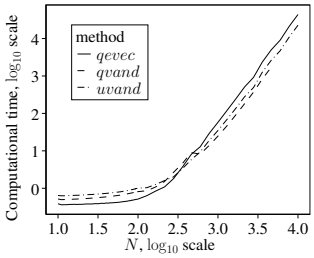

Figure 7: Computational time.

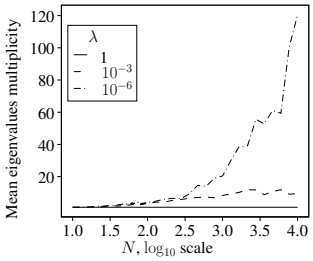

Figure 8: Mean eigenvalues multiplicity.

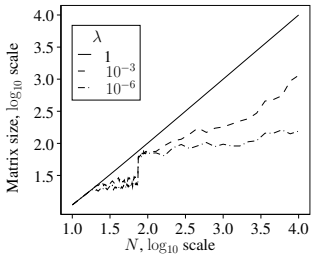

Figure 9: $\mathbf{V}^{(P)}$ matrix size in the $uvand$ method.

regime. Figure 6 depicts the values of $n^* = q\,t^*$ given by equation (45). The figure confirms that $\pi_j^{(Q)}(t)$ converges very fast to its stationary regime for $\lambda = \{10^{-3},\ 10^{-6}\}$, but very slowly for $\lambda = 1$. For instance, when $N = 100$, $n^* = \{413,\ 0.26,\ 0.25\}$ for $\lambda = \{1,\ 10^{-3},\ 10^{-6}\}$, respectively. The fact that for $\lambda = 1$, $\lambda_l^* \to 0$ when $N \to \infty$ explains also the pike of the condition number in $uvand$ observed in figure 3.

The divergence problem of the Uniformized Vandermonde method when λ_l^* is close to zero can be fixed as follows: Compute n^* given by (45). If $n^* \leq k$ (recall the k is the number of UC to be determined), then build the Vandermonde system taking the samples $n = 0, \cdots k-1$, as it was done to obtain the results of figure 5. By the *samples* we refer to the powers of the eigenvalues used to construct the Vandermonde matrix $\mathbf{V}^{(P)}$, and the corresponding values of $\mathbf{P}^n$ to build the matrix $\mathbf{B}$ (see section 3.2). Otherwise, if $n^* > k$, then build the Vandermonde system taking k samples evenly spaced in log scale in the interval $[0, n^*]$. Additionally, the equation that corresponds to the limiting distribution, $\pi_j^{(P)}(\infty)$, can be added to the system of equations. Note a QR decomposition can be used to solve an over-dimensioned system of equations.

Figure 7 shows the computation time in seconds for each method in the scenario with $\lambda = 1$, for the points where the numerical tool was able to solve the UC. It can be seen that for values of N up to $10^{2.5}$, approximately, $qevec$ is the fastest method. However, for larger values of N the methods $qvand$ and $uvand$ are faster. This is motivated by the cost of computing the eigenvectors in the $qevec$ method.

An additional advantage of the $uvand$ method is the reduction of the required UC in some scenarios. This is shown in figures 8 and 9, which respectively plot the mean multiplicity of the eigenvalues computed by the numerical tool, and the $\mathbf{V}^{(P)}$ matrix size (number of required UC, k, in the $uvand$ method). In the extreme case ($\lambda = 10^{-6}$ and $N = 10^4$) figure 8 shows that the numerical tool yielded a mean multiplicity of the eigenvalues equal to 120.5. Figure 9 shows that only 155 different UC were needed to solve this scenario. This result comes from the fact that the maximum multiplicity was limited to $M = 5$ (see remark 6.2),

8 Conclusions

In this paper are investigated the class of methods based on the *undetermined coefficients* approach to compute the transient solution of discrete and continuous time Markov chains (DTMC and CTMC). Two methods that belong to this class are investigated: The well known Eigenvectors method and the Vandermonde method. Even if both methods share the same principle, only the Eigenvectors method is normally described in the literature. However, Eigenvectors method has the drawback that it can only be used to solve non defective matrices. The Vandermonde method does not have this limitation, thus, giving a more general solution. The paper derives the equations required to use the Vandermonde method for both DTMC and CTMC. Simple examples are given that show the simplicity of the Vandermonde method to solve small systems without the help of a computer.

The Vandermonde method is also exploited to derive interesting relations between a CTMC and its uniformized chain. The equations that allow computing the undetermined coefficients (UC) of a CTMC in terms of the UC of its uniformized chain are derived. Based on them, it is proposed a new approach to compute the transient solution referred to as the *Uniformized Vandermonde* method. The advantage of this method are analyzed through extensive numerical results. The method is simple to implement and numerical results show that it has important advantages: (1) it is not affected by the presence of confluent eigenvalues, (2) it is very general, i.e. it able to give an accurate solution in all the scenarios were it was tested, (3) it allows saving computer memory by merging confluent eigenvalues.

Acknowledgments

This work was supported by Spanish grants TIN2010-21378-C02-01 and 2009-SGR-1167, and by the European NoE EuroNF.

Appendixes

A Multiplicity of the dominant eigenvalue

Assume a finite Markov chain. Define an *irreducible closed set* (ICS) C_k as a set where all states communicate with each other, i.e. $i \leftrightarrow j, \forall\, i, j \in C_k$. In other words, the probability f_{ij} of reaching a state j from a state i is $f_{ij} = 1, f_{ji} = 1,\ \forall\, i, j \in C_k$, and $f_{il} = 0,\ \forall\, i \in C_k,\ l \notin C_k$. Note that an absorbing state form an ICS of only one element, and transient states do not belong to any ICS.

Let refer to the eigenvalue $\lambda^{(P)} = 1$ of an stochastic matrix $\mathbf{P}$, and $\lambda^{(Q)} = 0$ of an infinitesimal generator $\mathbf{Q}$, as the *dominant eigenvalue*.

Theorem A.1 The multiplicity of the dominant eigenvalue is equal to the number of irreducible closed sets.

Proof. Assume a stochastic matrix $\mathbf{P}$ with M irreducible closed sets: $C_k,\ k = 1,\ \cdots M$. By properly numbering the states, we can write $\mathbf{P}$ as an M block diagonal matrix with the probabilities of the transient states in the last rows, as shown in (46) (see [2, chapter XV]). The blocks $\mathbf{P}_k,\ k = 1,\ \cdots M$ are irreducible stochastic matrices formed with the probabilities of the irreducible closed sets: $C_k,\ k = 1,\ \cdots M$.

$$\mathbf{P} = \left[\begin{array}{ccc|c} \mathbf{P}_1 & & & \\ & \ddots & & 0 \\ & & \mathbf{P}_M & \\ \hline \geq 0 & & & \mathbf{T} \end{array}\right] \tag{46}$$

Clearly, the characteristic polynomial of $\mathbf{P}$ is equal to the product of the characteristic polynomials of the submatrices $\mathbf{P}_k,\ k = 1,\ \cdots M$ and $\mathbf{T}$. Since $\mathbf{P}_k,\ k = 1,\ \cdots M$ are irreducible stochastic matrices, by the Perron-Frobenius theorem we know that each of them will have a single eigenvalue equal to 1. On the other hand, for the transitorial states it must be $\lim_{n\to\infty} \mathbf{T}^n = \mathbf{0}$. Therefore, all the eigenvalues of $\mathbf{T}$ must be $|\lambda_l^{(T)}| < 1$, and thus, we conclude that the multiplicity of the eigenvalue $\lambda^{(P)} = 1$ must be M.

For an infinitesimal generator $\mathbf{Q}$, take any uniformized matrix $\mathbf{P}$ (equation (29)). Clearly, the number of irreducible closed sets (M) of $\mathbf{Q}$ must be the same as those of its uniformized matrix $\mathbf{P}$. Since the multiplicity of the eigenvalue $\lambda^{(P)} = 1$ must be M, using equation (35) we conclude that so it must be for the eigenvalue $\lambda^{(Q)} = 0$ of $\mathbf{Q}$. □

B Derivation of $\frac{\partial^n \pi_j(0)}{\partial t^n}$

In order to solve the system of equations (26) we need to compute the derivatives of functions of the type:

$$\pi_{k-1}(t) = (a_0 + a_1\, t + \cdots + a_{k-1}\, t^{k-1})\, \mathrm{e}^{\lambda\, t} \tag{47}$$

In this appendix a general formula is obtained. This is necessary to implement a script for solving the system of equations given by (26). Start by noting that:

$$\frac{\partial \pi_{k-1}(t)}{\partial t} = [(\lambda a_0 + a_1) + (\lambda a_1 + 2\, a_2)\, t\, + \cdots$$
$$(\lambda a_{k-2} + (k-1)\, a_{k-1})\, t^{k-2} + \lambda a_{k-1}\, t^{k-1})]\, \mathrm{e}^{\lambda\, t} \tag{48}$$

which is again a polynomial in t of degree $(k-1)$ times an exponential. Let identify the coefficients of the polynomial associated with $\frac{\partial^n \pi_{k-1}(t)}{\partial t^n}$ by $a_m(n), m = 0, \cdots (k-1)$, and define the column vector $\mathbf{a}(n) = \begin{bmatrix} a_0(n) & \cdots & a_{k-1}(n) \end{bmatrix}$. Note that $\mathbf{a}(0) = \begin{bmatrix} a_0 & \cdots & a_{k-1} \end{bmatrix}$. By induction it can be easily obtained that $\mathbf{a}(n)$ obeys the systems of linear difference equations $\mathbf{a}(n+1) = \mathbf{A}\, \mathbf{a}(n)$, where:

$$\mathbf{A}^{k\times k} = \begin{bmatrix} \lambda & 1 & 0 & 0 & \cdots \\ 0 & \lambda & 2 & 0 & \cdots \\ 0 & 0 & \lambda & 3 & \cdots \\ \cdots & & & & \end{bmatrix} \tag{49}$$

This system has the solution $\mathbf{a}(n) = \mathbf{A}^n\, \mathbf{a}(0)$. Writing $\mathbf{A} = \lambda\, \mathbf{I} + \mathbf{U}$, where $\mathbf{U}$ is the matrix:

$$\mathbf{U}^{k\times k} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 2 & 0 & \cdots \\ 0 & 0 & 0 & 3 & \cdots \\ \cdots & & & & \end{bmatrix} \tag{50}$$

we have that:

$$\mathbf{A}^n = \sum_{m=0}^{n} \binom{n}{m} \lambda^{n-m}\, \mathbf{U}^m \tag{51}$$

where $\mathbf{U}^0 = \mathbf{I}$ and $\mathbf{U}^m$ is a matrix with the m upper diagonal equal to:

$$m!, \frac{(m+1)!}{1!}, \frac{(m+2)!}{2!}, \overset{n-k}{\cdots}, \frac{(k-1)!}{(k-m-1)!} \tag{52}$$

for $m = 1, \cdots k-1$, and $\mathbf{U}^m = \mathbf{0}$ for $m \geq k$. Note now that:

$$\frac{\partial^n \pi_{k-1}(0)}{\partial t^n} = a_0(n) \tag{53}$$

which is the first component of the vector $\mathbf{a}(n)$. Thus, using (51) and (52) we obtain:

$$\frac{\partial^n \pi_{k-1}(0)}{\partial t^n} = \sum_{m=0}^{k-1} \binom{n}{m} m!\, \lambda^{n-m}\, a_m = \sum_{m=0}^{k-1} n^{\underline{m}}\, \lambda^{n-m}\, a_m \tag{54}$$

where $n^{\underline{m}} = n\,(n-1)\cdots(n-m+1)$, $n^{\underline{m}} = 0$ for $n < m$.

Let $c_j^{(l,m)}(n)$ be the coefficient of the equation n of the system of equations (26) that multiplies the UC $a_j^{(l,m)}$ of $\pi_j(t)$. Using (54) we conclude that:

$$c_j^{(l,m)}(n) = n^{\underline{m}}\, \lambda_l^{n-m} \tag{55}$$

C Derivation of $q^{(l,m)}(t)$

In equation (32) we need to compute the following summation for the eigenvalues with multiplicity higher than 1:

$$S_m = \sum_{n=0}^{\infty} \frac{(q\,t)^n}{n!} (\lambda_l^{(P)})^n\, n^m = \sum_{n=0}^{\infty} \frac{x^n}{n!}\, n^m,\ m \geq 1 \tag{56}$$

where $x = q\,t\,\lambda_l^{(P)}$. Clearly, $S_0 = \mathrm{e}^x$ (assuming $n^m = 1$, $\forall n$ for $m = 0$) and,

$$\begin{aligned} S_m &= \sum_{n=0}^{\infty} \frac{x^n}{n!}\, n^m = x \sum_{n=1}^{\infty} (n-1+1)^{m-1} \frac{x^{n-1}}{(n-1)!} \\ &= x \sum_{n=1}^{\infty} \sum_{i=0}^{m-1} \binom{m-1}{i} (n-1)^i \frac{x^{n-1}}{(n-1)!} = \\ &\qquad x \sum_{i=0}^{m-1} \binom{m-1}{i} S_i,\ m \geq 1 \end{aligned} \tag{57}$$

From (57) it can be easily derived that S_m are $\mathrm{e}^{q\,t\,\lambda_l^{(P)}}$ times polynomials in t of degree m: $S_m = \mathrm{e}^{q\,t\,\lambda_l^{(P)}}\, q^{(l,m)}(t)$, where:

$$q^{(l,m)}(t) = \sum_{i=1}^{m} q_i^{(m)}\, (q\,\lambda_l^{(P)})^i\, t^i \tag{58}$$

with coefficients $q_i^{(m)}$ given by the recurrence relation:

$$q_i^{(m)} = \begin{cases} 1, & i = 1 \\ \sum_{k=i-1}^{m-1} \binom{m-1}{k} q_{i-1}^{(k)}, & i = 2, \cdots m \end{cases} \tag{59}$$

D Sharma's Formula

This formula is proposed in [10]. Let be an M/M/1/N queue with arrival rate λ, service rate μ, and $\rho = \lambda/\mu$. The probability that there are n units in the system at time t, $\pi_n(t)$, given that the queue is empty at time $t = 0$, can be computed as:

$$\pi_n(t) = \mathrm{e}^{-(1+\rho)t} \rho^n \sum_{m=0}^{\infty} a(m,n)\, \frac{t^m}{m!},\ 0 \leq n \leq N \tag{60}$$

where:

$$a(m,n) = \begin{cases} 0, & m < n \\ 1, & m = n \\ \sum\limits_{r=0}^{\left[\frac{m-n}{2}\right]} \sum\limits_{i=0}^{\left[\frac{r}{N+1}\right]} A(m, r-(N+1)\,i)\, \rho^r - \\ \qquad \sum\limits_{r=\left[\frac{m-n}{2}+1\right]}^{m-n} \sum\limits_{i=0}^{\left[\frac{n-N-r}{N+1}\right]} A(m, r+(N+1)(i+1))\, \rho^r, & m > n \end{cases}$$

and: $A(m,s) = \binom{m}{s} - \binom{m}{s-1}$, taking $\binom{b}{a}$ as zero whenever $b < a$ or $b < 0$. $[x]$ stands for the integral value of x. In [10] it is also given a recurrence relation for the coefficients $a(m,n)$, but it was found faster to compute them using the equation above. To evaluate (60) we need to truncate the summation. Numerical experiments showed that for values of t up to 100 s, approximately, the terms of the summation vanish rapidly. Note that the larger is t, the more terms need to be considered, which becomes numerically unfeasible for the combinatorial coefficients of the equation.

References

[1] Å. Björck and V. Pereyra. "Solution of Vandermonde Systems of Equations." English. In: *Mathematics of Computation* 24.112 (1970), pp. 893–903.

[2] W. Feller. *An Introduction to Probability Theory and Its Applications, Vol. 1, 3rd Edition*. Wiley, 1968.

[3] W. H. Jr, J. Fillmore, and D. Smith. "Matrix Exponentials: Another Approach." In: *SIAM review* 43.4 (2001), pp. 694–706.

[4] S. Karlin and H. Taylor. *A first course in stochastic processes*. Vol. 697. Academic press New York, 1975.

[5] I. Leonard. "The Matrix Exponential." In: *SIAM Rev.* 38.3 (1996), pp. 507–512.

[6] R. A. Marie, A. L. Reibman, and K. S. Trivedi. "Transient analysis of acyclic Markov chains." In: *Performance Evaluation* 7.3 (1987), pp. 175–194.

[7] C. Moler and C. Van Loan. "Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later." In: *SIAM review* 45.1 (2003), pp. 3–49.

[8] R Development Core Team. *R: A Language and Environment for Statistical Computing*. ISBN 3-900051-07-0. R Foundation for Statistical Computing. Vienna, Austria, 2010.

[9] A. L. Reibman and K. S. Trivedi. "Numerical transient analysis of Markov models." In: *Comput. Oper. Res.* 15.1 (1988), pp. 19–36.

[10] O. Sharma and A. Tarabia. "A Simple Transient Analysis of an M/M/1/N Queue." In: *The Indian Journal of Statistics*. A (2 June 2000), pp. 273–281.

[11] G. Stewart. "The Decompositional Approach to Matrix Computation." In: *Computing in Science & Engineering* 2.1 (2002), pp. 50–59.

[12] W. J. Stewart. *Introduction to the numerical solution of Markov chains*. Princeton University Press, 1994.