Stochastic Analysis and Applications, 25: 763–779, 2007

ISSN 0736-2994 print/1532-9356 online
DOI: 10.1080/07362990701419946

On the Solution to QBD Processes with Finite State Space

Essia H. Elhafsi and Mart Molle
Department of Computer Science and Engineering,
University of California, Riverside, California, USA

Abstract: In this article, we present a solution to a class of Quasi-Birth-and-Death processes with finite state space and show that the stationary probability vector has a matrix geometric representation. We show that such models have a level-dependent rate matrix. The corresponding rate matrix is given explicitly in terms of the model parameters. The resulting closed-form expression is proposed as a basis for efficient calculation of the stationary probabilities. The method proposed in this article can be applied to several queueing systems.

Keywords: Closed-form solution; Computation complexity; Markov chain; Matrix geometric.

AMS Subject Classification: 60J22; 65Y20; 68U01.

1. INTRODUCTION

Matrix geometric methods have been widely used in computer performance modeling. Many computer models have a repetitive structure, which leads to Markov models that fit within the matrix geometric framework. Models that satisfy this condition include certain open queueing systems consisting of an infinite capacity queue, processes where the service and interarrival times are given by phase-type distributions [2], and quasi-birth-death (QBD) processes [12].

Received March 24, 2006; Accepted January 22, 2007
Address correspondence to Essia H. Elhafsi, Department of Computer Science and Engineering, University of California, Riverside, CA 92521, USA; E-mail: essia@cs.ucr.edu

Many models can be made approximately matrix geometric by truncating certain portions of the state space and assuming that subsequent state transitions repeat to satisfy the desired form [3, 14, 15].

In this work, we are interested in QBD processes. Such processes arise in the modeling of a wide variety of applications such as telecommunication, computer performance and inventory control. QBD processes have been used as models in queues with phase-type arrivals and services [9], in the shorter queue problem [11], and in the machine repair problem [10]. A QBD process is a Markov process on state space $E = \{(\xi, i); \xi \geq 0, 1 < i < r\}$ with an infinitesimal generator matrix of the following form:

$$Q = \begin{bmatrix} \widehat{L} & \widehat{F} & 0 & 0 & 0 & \dots \\ \widehat{B} & L & F & 0 & 0 & \dots \\ 0 & B & L & F & 0 & \dots \\ 0 & 0 & B & L & F & \dots \\ 0 & 0 & 0 & B & L & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

The block entries B, L, F, and $\widehat{L}$ are square sub-matrices, which satisfy the equilibrium conditions $\widehat{L}e + \widehat{F}e = \widehat{B}e + Le + Fe = (B + L + F)e = 0$; e is a column vector of ones of suitable dimension.

The state space is generally partitioned into a block of boundary states S_0 and the remaining blocks of states S_ξ, $\xi \geq 0$ that represent the repetitive portion of the Markov chain. We use the letters "L", "F" and "B" to describe "local", "forward" and "backward" transition rates respectively with relation to a block of states S_ξ, $\xi \geq 1$ and "$\hat{\cdot}$" for matrices related to states in S_0.

QBD processes have interesting structural properties which can be used to simplify the computation of the stationary probabilities. Two matrices usually denoted by R and G play a major role in the general theory. These matrices have important probabilistic interpretations. An entry (k, j) in G expresses the conditional probability of the process first entering $S_{\xi-1}$ through state j given that it starts from state i of S_ξ [12]. An entry (k, j) in R is the expected time spent in state j of S_ξ before the first visit to $S_{\xi-1}$ given the starting state k in $S_{\xi-1}$ [9].

Both matrices R and G are minimal nonnegative solutions of two nonlinear matrix equations (i.e., $B + LG + FG^2 = 0$ and $F + RL + R^2B = 0$). Moreover for a QBD process, R can be expressed as $R = -F(L + FG)^{-1}$. Although QBD processes can be solved by either computing R or G, the first one is usually used.

Let $\pi = (\pi_0, \pi_1, \dots)$ be the stationary probability vector of a QBD process with infinite state space where π_ξ is the subvector of stationary probabilities for states in block ξ. Under general assumptions, the elements of this vector have the following matrix geometric property [9],

$$\pi_{\xi+1} = \pi_{\xi} \cdot R, \quad \xi \geq 0 \tag{1}$$

For QBD processes with finite state space, however, the situation is quite different due to the presence of additional boundary states. Thus, it is not possible to guarantee in general that the stationary distribution has a matrix geometric structure of the form of Equation (1). Several methods have been proposed in the literature to solve for the stationary probabilities of QBD processes with a finite state space. We discussed and compare some of the methods in Section 4.

The aim of this article is to provide an efficient solution to the stationary probabilities of a QBD process with finite state space. In Section 2, we show that the proposed solution yields a closed-form expression of the stationary probability vector π_{ξ} that is similar to Equation (1). However, we show that the rate matrix R is level-dependent and can be solved using a simple procedure that depends on the system parameters. Moreover, matrix R_{ξ} is computed without the need of solving a matrix quadratic equation, which is generally the case in the algorithmic approaches suggested by Neuts [9] and Hajek [7]. Under certain assumptions we show in Section 3 that matrix R_{ξ} can be written as a simple linear combination of $R_{\xi+1}$ and $R_{\xi+2}$, or of $R_{\xi-1}$ and $R_{\xi-2}$ where ξ is the process level. We apply this solution technique to solve a multiserver finite buffer queuing model. In Section 4, we compare this method to existing methods in terms of constraints and computation complexity.

2. THE MODEL AND ITS GENERAL SOLUTION

2.1. Model

Consider a QBD process with finite state space E. For simplicity and as an illustration we define E as follows:

$$E = \begin{cases} (0, i) & \text{if } \xi = 0 \text{ and } 1 \leq i \leq s \\ (\xi, i) & \text{if } 1 \leq \xi \leq \Psi \text{ and } 1 \leq i \leq r \end{cases}$$

where coordinate ξ denotes the "level" and i the "phase" of state (ξ, i); Ψ represents the number of levels.

The generator matrix Q of this process has the following block-tridiagonal structure:

$$Q = \begin{bmatrix} \widehat{L} & \widehat{F} & 0 & 0 & 0 & 0 & 0 & \dots \\ \widehat{B} & L & F & 0 & 0 & 0 & 0 & \dots \\ 0 & B & L & F & 0 & 0 & 0 & \dots \\ 0 & 0 & B & L & F & 0 & 0 & \dots \\ 0 & 0 & 0 & B & L & F & 0 & \dots \\ \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \dots & \dots & \dots & \dots & \dots & \dots & L & F_{\Psi-1} \\ \dots & \dots & \dots & \dots & \dots & \dots & B_{\Psi} & L_{\Psi} \end{bmatrix} \tag{2}$$

The upper boundary block entries $\widehat{L}$, $\widehat{F}$, and $\widehat{B}$ are matrices with respective dimensions $(s \times s)$, $(s \times r)$, and $(r \times s)$. The nonboundary block entries B, L and F and the lower boundary block entries B_{Ψ}, $F_{\Psi-1}$, and L_{Ψ} are all square matrices of dimension $(r \times r)$.

If the underlying Markov chain with generator matrix Q is irreducible, then the matrices L, $\widehat{L}$, and L_{Ψ} along the diagonal can be shown to be nonsingular [9].

2.2. General Solution Technique

1. Global Balance Equations

The stationary probability vector π for Q is generally partitioned as $\pi = [\pi_0, \pi_1, \pi_2, \dots, \pi_{\Psi}]$, where the subvectors π_0 and π_{ξ}, $(1 \le \xi \le \Psi)$ are of dimension s and r, respectively.

Solving $\pi Q = 0$ along with the normalizing equation $\pi e = 1$, where e is a column vector of ones, yields the following set of equations in matrix form:

$$\pi_0 \widehat{L} + \pi_1 \widehat{B} = 0 \tag{3}$$

$$\pi_0 \widehat{F} + \pi_1 L + \pi_2 B = 0 \tag{4}$$

$$\pi_{\xi-1} F + \pi_{\xi} L + \pi_{\xi+1} B = 0, \quad 2 \le \xi < \Psi - 1 \tag{5}$$

$$\pi_{\Psi-2} F + \pi_{\Psi-1} L + \pi_{\Psi} B_{\Psi} = 0, \tag{6}$$

$$\pi_{\Psi-1} F_{\Psi-1} + \pi_{\Psi} L_{\Psi} = 0. \tag{7}$$

2. Computation of the Rate Matrices

Here, we assume that Equation (8) holds among the stationary probability vectors π_{ξ} for states in set S_{ξ}, and R_{ξ} is a square matrix of order r,

$$\pi_{\xi} = \pi_{\xi-1} R_{\xi}, \quad \xi \ge 1 \tag{8}$$

By simple algebraic manipulation of the global balance equations we obtain R_ξ's as follows.

- From Equation (3) and assuming that $\widehat{L}$ is nonsingular, we get,

$$\pi_0 = -\pi_1 \widehat{B}\widehat{L}^{-1} \equiv \pi_1 R_0 \tag{9}$$

- From Equation (7) we obtain the following expression for π_Ψ and R_Ψ where L_Ψ is required to be nonsingular as,

$$\pi_\Psi = -\pi_{\Psi-1} F_{\Psi-1} L_\Psi^{-1} \equiv \pi_{\Psi-1} R_\Psi$$

- Equation (6) leads to the following expression of $\pi_{\Psi-1}$ and $R_{\Psi-1}$,

$$\begin{aligned}\pi_{\Psi-1} &= -\pi_{\Psi-2} F(L - F L_\Psi^{-1} B_\Psi)^{-1} \\ &= -\pi_{\Psi-2} F(L + R_\Psi B_\Psi)^{-1} \\ &\equiv \pi_{\Psi-2} R_{\Psi-1}\end{aligned}$$

- Finally, from Equation (5) we obtain a general relation between $\pi_{\Psi-1}$, π_ξ, and R_Ψ,

$$\begin{aligned}\pi_\xi &= -\pi_{\xi-1} F(L + R_{\xi+1} B)^{-1} \\ &\equiv \pi_{\xi-1} R_\xi, \quad 2 \le \xi < \Psi - 1\end{aligned}$$

  R_ξ can be computed using Algorithm 1.

In Algorithm 1, I represents the identity matrix of dimension $(r \times r)$. We also assume that $-\widehat{L}^{-1}$ and $-L_\Psi^{-1}$ are nonnegative matrices. Moreover, if $(L + R_\xi B)$ is stable then it is nonsingular. Note that the rate matrices R_ξ, $\xi = 1, \ldots, \Psi$ are obtained through purely algebraic manipulations starting from the global balance conditions; they have no probabilistic interpretation, and, therefore, do not coincide in general with the rate matrix introduced by Neuts. Moreover, the rate matrices R_ξ introduced here are not always positive and this could lead to some numerical instabilities.

It is worth noting that the solution to the rate matrices can be generalized to the infinite state space solution given in [9]. For infinitely large number of blocks ($\Psi \to \infty$), the rate matrices R_ξ will converge to R which is the minimal nonnegative solution to

$$R = -F(L + RB)^{-1}$$

This condition is equivalent to the following nonlinear equation introduced in [9]

$$F + RL + R^2 B = 0$$

3. Stationary Probabilities

Theorem 1. *For any QBD process with finite state space, having an infinitesimal generator matrix given by Equation* (2)*, the stationary probabilities are given in matrix-geometric form by*

$$\pi_\xi = \pi_1 R_\xi^*$$

where $R_\xi^* = \prod_{j=1}^{\xi} R_j$ *and* R_j *is computed using Algorithm* 1.

Algorithm 1 Compute $R(\xi)$

$R_\Psi \leftarrow -F L_\Psi^{-1}$
if $\xi > 1 \rightarrow$
  for $j = \Psi - 1 \rightarrow \xi + 1$ **do**
    $R_j \leftarrow -F(L + R_{j+1}B)^{-1}$
  od
  return $R_\xi \leftarrow -F(L + R_{\xi+1}B)^{-1}$
fi
if $\xi == 0 \rightarrow$
  return $R_0 \leftarrow -\widehat{B}\widehat{L}^{-1}$
fi
if $\xi == 1 \rightarrow$
  return $R_1 \leftarrow I$
fi

Solving Equations (3) and (4) for π_1 leads to,

$$\pi_1(R_0\widehat{F} + L + R_2 B) = 0 \tag{10}$$

Thus, after substitution and mathematical manipulation, Equation (11) follows from the normalizing condition $\pi_0 e_0 + \sum_{\xi=1}^{\Psi} \pi_\xi \ \ e = 1$ and Equation (9), where e_0 is a column vector of ones with suitable dimension.

$$\pi_1\left(R_0 e_0 + \sum_{\xi=1}^{\Psi} R_\xi^* e\right) = 1 \tag{11}$$

Solving the system of linear equations given by Equations (10) and (11), we solve for π_1, and from Equation (8), we obtain π_ξ as,

$$\begin{aligned}\pi_\xi &= \pi_{\xi-1}R_\xi \\ &= \pi_{\xi-2}R_{\xi-1}R_\xi \\ &= \pi_{\xi-2}R_{\xi-2}R_{\xi-1}R_\xi \\ &\ \ \vdots\end{aligned} \tag{12}$$

$$= \pi_1 R_1 \dots R_{\xi-2} R_{\xi-1} R_\xi$$

$$= \pi_1 \prod_{j=1}^{\xi} R_j = \pi_1 R_\xi^*$$

Algorithm 2 is used to compute the stationary probabilities π.

Algorithm 2 Compute π

1: **for** $j = 1 \to \xi$ **do**
2: $R_j \leftarrow compute\ R(j)$
3: $\pi_j \leftarrow \pi_1 R_j$
4: **od**

4. Average Number in the System

Using first principles, we compute the number of customers in the system N (including the ones in service if any) as,

$$N = \pi_0 n_0 + \sum_{\xi=1}^{\Psi} \pi_1 \left(\prod_{j=1}^{\xi} R_j \right) n_\xi$$

$$= \pi_1 \left((R_0 n_0 + N_1) + \sum_{\xi=2}^{\Psi} R_\xi^* n_\xi \right)$$

where n_ξ is a column vector of length r and its jth entry gives the number of customers in the system when its current state is at level ξ with phase i. Similarly, n_0 is the corresponding number of customers in the system for the boundary block.

3. SPECIAL CASES SOLUTION

We show in this section that under certain conditions the solution to the stationary probabilities of a QBD process can be obtained from the continuous time Markov chain (CTMC) with at most one matrix inversion. Moreover, under such assumptions we obtain expressions for the stationary probabilities with simpler solutions than in the general case.

We will use the following example, which is inspired from [4, 8], to illustrate our solution method.

3.1. Example 1

We consider a service center where packets arrive according to a Poisson process of rate λ. The service center has one input and two outputs that

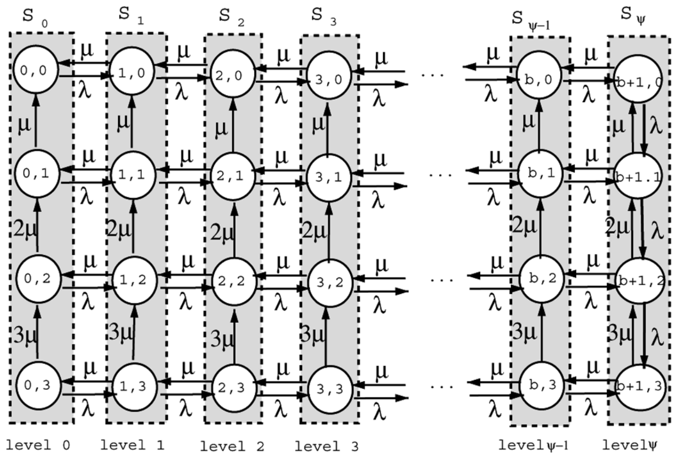

Figure 1. Example 1 CTMC: Shared queue with finite storage when $m_1 = 3$ and $m_2 = 1$.

lead to the same destination via two alternate paths. There is one finite-capacity queue for waiting packets, which is shared by both outputs. We assume that the first hop on path ℓ, $\ell = 1, 2$ consists of a multiserver *link group*, where there are m_ℓ identical servers running in parallel serving output port ℓ. However, we also assume that mean network transit delay is different for the two paths, and without loss of generality that the delay is larger for path 1. Therefore, all packets are routed to path 2 unless the number of waiting packets exceeds a threshold b. To simplify the example, we will assume that the service center can store at most $C = m_1 + m_2 + b$ packets (both waiting and in service).

Let $\mathbf{Q}(t)$ denote the system state at time t with $\mathbf{Q}(t) := [k, j]$, where $0 \leq k \leq m_1$ denotes the number of packets currently being transmitted on link group 1, $j \in \{0, 1, 2, \ldots\}$ denotes the remaining number of packets currently in the system and $v_2(j) \equiv \min(j, m_2)$ denotes the number of packets are currently being transmitted on link group 2. Thus, $j - v_2(j)$ packets are just waiting in the queue. $\{\mathbf{Q}(t)\}_{t \geq 0}$ can be represented as a continuous time Markov chain. The CTMC for the system and the transition rates when $m_1 = 3$ and $m_2 = 1$ are shown in Figure 1 and the corresponding block matrices are given as:

$$B = \begin{bmatrix} \mu & 0 & 0 & 0 \\ 0 & \mu & 0 & 0 \\ 0 & 0 & \mu & 0 \\ 0 & 0 & 0 & \mu \end{bmatrix}, \quad F = \begin{bmatrix} \lambda & 0 & 0 & 0 \\ 0 & \lambda & 0 & 0 \\ 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & \lambda \end{bmatrix},$$

$$L = \begin{bmatrix} -(\lambda+\mu) & 0 & 0 & 0 \\ \mu & -(\lambda+2\mu) & 0 & 0 \\ 0 & \mu & -(\lambda+2\mu) & 0 \\ 0 & 0 & \mu & -(\lambda+2\mu) \end{bmatrix},$$

$$\widehat{L} = \begin{bmatrix} -\lambda & 0 & 0 & 0 \\ \mu & -(\lambda+\mu) & 0 & 0 \\ 0 & \mu & -(\lambda+\mu) & 0 \\ 0 & 0 & \mu & -(\lambda+\mu) \end{bmatrix},$$

$$\widehat{L}_\Psi = \begin{bmatrix} -(\lambda+\mu) & \lambda & 0 & 0 \\ \mu & -(\lambda+2\mu) & \lambda & 0 \\ 0 & \mu & -(\lambda+2\mu) & \lambda \\ 0 & 0 & \mu & -2\mu \end{bmatrix},$$

$$\widehat{B} = B_\Psi = B, \quad \widehat{F} = F_{\Psi-1} = F.$$

Note that matrices B, B_Ψ, $\widehat{B}$, F, F_Ψ, and $\widehat{F}$ are diagonal and can be expressed as

$$B = B_\Psi = \widehat{B} = \mu I$$
$$F = F_{\Psi-1} = \widehat{F} = \lambda I.$$

This is very fortunate, since having B and F be diagonal matrices means that their inverse can be trivially computed by forming the scalar inverse of their diagonal elements. For the rest of the analysis, we assume that $m_2 = 1$ and $m_1 \geq 1$. The block matrices corresponding to the generator matrix Q are therefore, dependent on m_1. Therefore, the larger m_1 is, the larger are the sizes of the matrices.

3.2. Case 1 Analysis

Here, we make the following assumption:

Assumption. B, B_Ψ, $\widehat{B}$, F, F_Ψ, and $\widehat{F}$ are nonsingular diagonal matrices.

Solution. We can solve the global balance equations given by Equations (13–15) and obtain a unique solution to the stationary probabilities $\{\pi_\xi\}_{0\leq\xi\leq\Psi}$.

- The global balance equations are given as follows,

$$\pi_0\widehat{L} + \mu\pi_1 I = 0 \tag{13}$$

$$\lambda\pi_{\xi-1}I + \pi_\xi L + \mu\pi_{\xi+1}I = 0, \quad \xi = 1, \ldots, \Psi - 1 \tag{14}$$

$$\lambda\pi_{\Psi-1}I + \pi_\Psi L_\Psi = 0 \tag{15}$$

- The stationary probabilities are,

$$\pi_1 = -\frac{1}{\mu}\pi_0\widehat{L} \equiv \pi_0 R_1$$

$$\pi_2 = -\frac{1}{\mu}\pi_0(\lambda I + R_1 L) \equiv \pi_0 R_2$$

$$\pi_3 = -\frac{1}{\mu}\pi_0(\lambda R_1 + R_2 L) \equiv \pi_0 R_3$$

$$\vdots$$

$$\pi_{\Psi-1} = -\frac{1}{\mu}\pi_0(\lambda R_{\Psi-3} + R_{\Psi-2}L) \equiv \pi_0 R_{\Psi-1}$$

$$\pi_\Psi = -\frac{1}{\mu}\pi_0(\lambda R_{\Psi-2} + R_{\Psi-1}L) \equiv \pi_0 R_\Psi$$

which can be generalized to the following form:

$$\pi_\xi = \pi_0 R_\xi, \quad \xi = 0, \ldots, \Psi$$

where R_ξ's are computed using the following procedure:

$$\begin{cases} R_0 = I \\ R_1 = -\frac{1}{\mu}\widehat{L} \\ R_\xi = -\frac{1}{\mu}(\lambda R_{\xi-2} + R_{\xi-1}L), \quad \xi = 2, \ldots, \Psi \end{cases}$$

The computation of the rate matrices involves only matrix additions and multiplications. No matrix inversions are required.

Note that Equation (15) can be expressed as,

$$\pi_0(\lambda R_{\Psi-1} + R_\Psi L_\Psi) = 0 \tag{16}$$

Therefore, π_0 can be obtained by solving Equation (16) along with the following normalizing condition:

$$\pi_0 \sum_{i=0}^{\Psi} R_i e = 1$$

3.3. Case 2 Analysis

In this case, we make the following assumption:

Assumption. F and $F_{\Psi-1}$ are nonsingular diagonal matrices.

Another example where such an assumption holds and more specifically B, $\widehat{B}$, and B_{Ψ} are singular, is presented in [12].

Solution. We solve the global balance equations given by Equations (19)–(20) and obtain a unique solution to the stationary probabilities $\{\pi_{\xi}\}_{0\leq\xi\leq\Psi}$.

- The global balance equations are given as follows,

$$\pi_0\widehat{L} + \pi_1\widehat{B} = 0 \tag{17}$$

$$\pi_0\widehat{F} + \pi_1 L + \pi_2 B = 0 \tag{18}$$

$$\lambda\pi_{\xi-1}I + \pi_{\xi}L + \pi_{\xi+1}B = 0, \quad \xi = 1, \ldots, \Psi - 1 \tag{19}$$

$$\lambda\pi_{\Psi-1}I + \pi_{\Psi}L_{\Psi} = 0 \tag{20}$$

- The stationary probabilities are,

$$\begin{aligned}
\pi_{\Psi} &= \pi_{\Psi}I \equiv \pi_{\Psi}R_{\Psi} \\
\pi_{\Psi-1} &= -\frac{1}{\lambda}\pi_{\Psi}L_{\Psi} \equiv \pi_{\Psi}R_{\Psi-1} \\
\pi_{\Psi-2} &= -\frac{1}{\lambda}\pi_{\Psi}(R_{\Psi-1}L + B) \equiv \pi_{\Psi}R_{\Psi-2} \\
\pi_{\Psi-3} &= -\frac{1}{\lambda}\pi_{\Psi}(R_{\Psi-2}L + BR_{\Psi-1}) \equiv \pi_{\Psi}R_{\Psi-3} \\
&\vdots \\
\pi_1 &= -\frac{1}{\lambda}\pi_{\Psi}(R_2 L + BR_3) \equiv \pi_{\Psi}R_1 \\
\pi_0 &= -\frac{1}{\lambda}\pi_{\Psi}(R_2 L + BR_3)\widehat{B}\widehat{L}^{-1} \equiv \pi_{\Psi}R_0
\end{aligned}$$

which can be generalized to the following solution,

$$\pi_{\xi} = \pi_{\Psi}R_{\xi}, \quad \xi = 0, \ldots, \Psi$$

where R_ξ's can be solved using the following procedure

$$\begin{cases} R_\Psi = I \\ R_{\Psi-1} = -\frac{1}{\lambda} L_\Psi \\ R_\xi = -\frac{1}{\lambda}(R_{\xi+1}L + BR_{\xi+2}) \quad \xi = 1, \ldots, \Psi - 2 \\ R_0 = \frac{1}{\lambda}(R_2 L + BR_3)\widehat{B\widehat{L}}^{-1} \end{cases}$$

Here, the computation of the rate matrices R_ξ, $\xi > 0$ involves only matrix additions and multiplications. Only the computation of R_0 requires one matrix inversion ($\widehat{L}^{-1}$).

To solve for π_Ψ, we solve Equation (18) which can be written as,

$$\pi_\Psi(R_0\widehat{F} + R_1 L + BR_2) = 0$$

and the following normalizing condition:

$$\pi_\Psi\left(R_0 e_0 + \sum_{\xi=1}^{\Psi} R_\xi e\right) = 1$$

4. METHODS COMPARISON

In this section, we compare several methods proposed in the literature to solve for the stationary probabilities of an irreducible finite QBD process. The common idea of these methods is to reduce the global balance system to a smaller system and to express the steady state probabilities as a function of the solution to this reduced system. However, the methods differ in both the way the reduced system is obtained and the way the steady state probabilities are computed from its solution. Moreover, the methods operate under different assumptions.

In this section, we compare several existing solution methods with the one proposed in this article. For each method we provide the computational complexity and the constraints that must be satisfied. We assume that all the block submatrices are of dimension $n \times n$ and we express the computational complexity in terms of the number of matrix operations performed: number of matrix multiplications, additions, and inversions that each method requires.

Hajek [7] showed that the stationary probability vector can be written as a sum of two matrix geometric terms plus a linear term. The corresponding computations involve solving two matrix quadratic equations (to compute R and G) which generally involves the use of numerical techniques and then finding the stationary probability distribution on the boundary states along with a normalizing constant. The state space here is obtained by truncating the infinite state space. This method assumes that L is nonsingular diagonal block matrix which holds when the process is irreducible [9].

Gun [5, 6] provides results for finite state QBD processes with an application to finite capacity queues with phase-type servers. A closed-form expression for the stationary probabilities is presented for some matrix R to which a closed-form expression is provided as well. This method applies to finite QBD processes, provided that some matrices are nonsingular ($\widehat{L}$, L_Ψ, and M where M is the solution to a set of matrix linear equations) and some matrix equalities hold. The solution provides relations between the stationary probabilities that cannot be exploited to yield in general a recursive solution for the stationary probability vector components.

An explicit matrix analytic solution for a broad class of finite QBD processes is proposed by Vittoria [1]. A recursive equation is provided for the computation of the nonboundary subvectors. The only assumption of this method is that matrix B is nonsingular. Moreover, the solution technique is extended to a generalized QBD *(GQBD)* process without repetitive structure in the generator matrix. The assumption in this case is that all the lower diagonal matrices ($B_{\xi,\xi-1}$) are nonsingular.

The method proposed by Li [13] assumes that the boundary matrices satisfy the following conditions: $F = F_{\Psi-1}$, $L_\Psi = L + F$, and $B_\Psi = B$ which is a special case of the general solution presented in this article.

The method proposed by Le Boudec [16] provides a recursive equation for the computation of the stationary probabilities. The solution applies to more general processes than finite QBD processes, namely the class of finite state space processes with an upper block Hessenberg[1] generator matrix. No repetitive block structure is necessary. The method assumes that all the submatrices are square blocks of the same size.

Now consider the general method proposed in this article. Here we require that $\widehat{L}$ and L_Ψ are nonsingular matrices. Moreover, if $\widehat{L}$ and L_Ψ are nonnegative the system will be stable (R_ξ positive). This method does not need to solve any quadratic nonlinear equation to obtain the stationary probabilities. We also provide a simpler solution method for the stationary probabilities when B, $\widehat{B}$, B_Ψ, F, $\widehat{F}$, and $F_{\Psi-1}$ are nonsingular matrices and when only F, $\widehat{F}$ and $F_{\Psi-1}$ are nonsingular matrices.

To compare the complexity of the methods we assume that the block size (or number of states in a level) is r and that the addition of two $r \times r$

[1]A finite square matrix A of size K is upper block Hessenberg if the building blocks are such that $A_{m,n} = 0$ for $n < m - 1$.

matrices requires n operations, the multiplication of two $r \times r$ matrices requires r^2 operations, and the inversion of an $r \times r$ matrix requires r^3 operations.

With regard to the general method proposed in this article, to compute the stationary probabilities, we have to compute R_ξ, $\xi = 0, 1, 2, \ldots, \Psi$ (see Algorithm 1), which requires $(\Psi + 1)$ matrix inversions, $(4\Psi + 3)$ matrix multiplications and Ψ matrix additions.

Special case 1 method requires $3(\Psi + 1)$ matrix multiplications and 3Ψ matrix additions. No matrix inversion is required.

Special case 2 method on the other hand, requires one matrix inversion, $(4\Psi + 5)$ matrix multiplications and $(2\Psi + 1)$ matrix additions.

We compare the computation complexity of the methods for various block sizes applied to Example 1 (Figure 1). To vary the block size, we vary the value of $m1$ (recall that $m1$ defines the block size). The general method has the worst performance (see Figure 2) and its performance decays as the block size increases. Case 1 method has the best performance. The same results are obtained for fixed block size (see Figure 3). In all cases, the curves in Figure 2 show that the methods computation complexity is nonlinear in r.

Table 1 presents a detailed complexity analysis of the various methods. We plot the complexity of each method for a given state

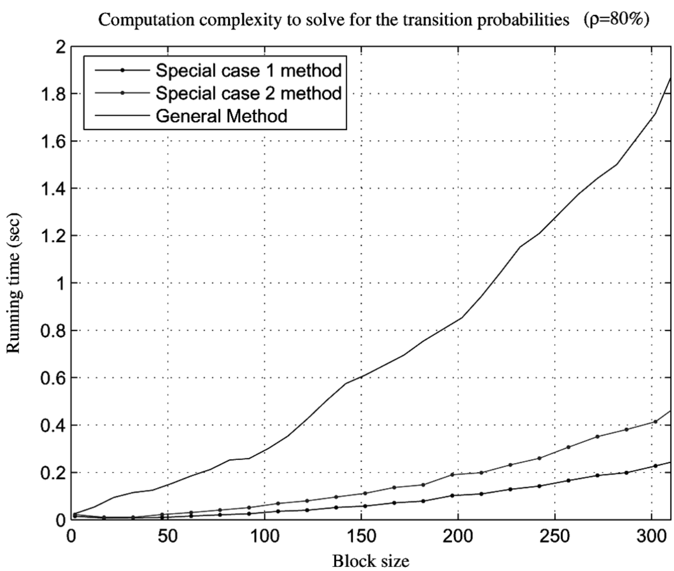

Figure 2. Comparison of the running time to compute the stationary probabilities using the general and the special cases methods for variable block size.

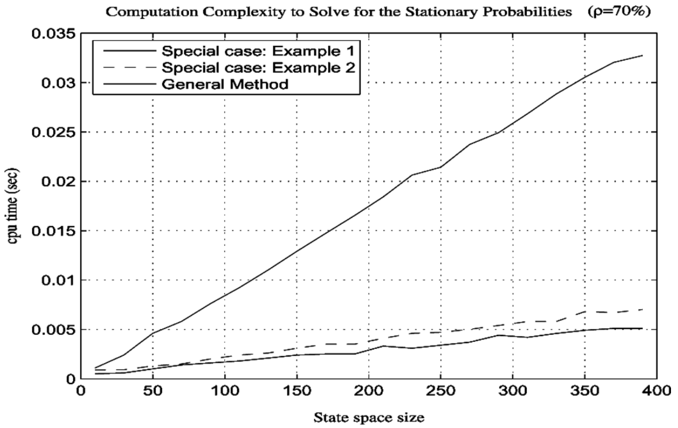

Figure 3. Comparison of the running time to compute the stationary probabilities using the general and the special cases methods for fixed block size of 2.

space size S and various block sizes r (see Figure 4). Note that $S = \Psi r$. Figure 4 shows that case 1 method provides the best computation performance followed by case 2 method at average and large block sizes. They are outperformed by Gun's method at low block sizes. The general method is comparable to the method proposed by Li [13] and the more general method proposed by Le Boudec [16]. It performs worse then Vittoria [1] and Gun [5] because our general method requires more matrix inversions (one inversion for the computation of each R_ξ).

The method proposed by [7], however has the worst performance among all methods. To analyze its performance we assume that for matrix R to converge, β iterations are required.

Our results (Table 1) show that the methods have about the same asymptotic complexity $O(r^3)$ with the exception of the method presented

Table 1. Computation complexity of various methods to compute the stationary probabilities of finite state space QBD processes (S is the size of the state space and r is the block size)

Methods	Computation complexity
Le Boudec [16]	$S(3r + r^2 + 3) + 5r^2 + r^3 + 3r$
Gun et al. [5]	$3Sr + 3r^3 + 32r^2 + 8r$
Vittoria et al. [1]	$7Sr + 4S + 17r^2 - 10r + r^3$
Hajek [7]	$(8 + \beta)r^3 + (17 + 2\beta)r^2 + (6S + 5 + \beta)r + S$
Li et al. [13]	$Sr^2 + 4Sr + S + r^3 + 3r^2 + r$
General method	$Sr^2 + 4Sr + S + r^3 + 3r^2$
Case 1 method	$3S(r + 1) + 3r^2$
Case 2 method	$2S(2r + 1) + r^3 + 6r^2 + r$

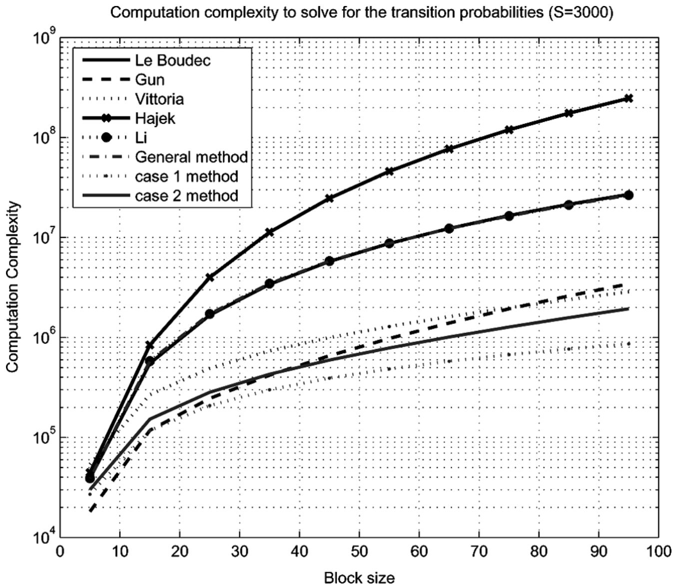

Figure 4. Comparison of the computation complexity of all methods as a function of the block size for a state space size $S = 3000$.

in special case 1 where the asymptotic complexity is $O(r^2)$ when r is large. For small values of r, their asymptotic complexity is $O(r^2)$ or more precisely $O(Sr)$; the methods presented [13, 16], our general and case 2 method on the other hand are $O(r^3)$, more precisely $O(Sr^2)$. The difference in the methods performance is mainly due to the assumptions (described above) under which the solution was obtained.

5. CONCLUSION

The methods compared in this article for irreducible finite QBD processes have the same asymptotic complexity $O(r^3)$, for large block sizes, considering multiplicative and additive constants except for the special case when B, $\widehat{B}$, B_{Ψ}, F, $\widehat{F}$, $F_{\Psi-1}$ are nonsingular matrices, the asymptotic complexity is only $O(r^2)$. The general method proposed in this article holds an intermediate position in terms of complexity and generality. It is not as general as the method proposed by Le Boudec [16]. However, it is more general than the method proposed by Li [13]. Moreover, it has a lower complexity than several of the methods proposed in the literature. These methods, including the one proposed in this article and with the exception of the one proposed by [16], suffer from numerical stability problems (R is not guaranteed to be nonnegative). We show that under certain assumptions, a simple solution to matrix R and to the stationary probabilities exits.

REFERENCES

1. de Nitto Persone, V., and Grassi, V. 1996. Solution of finite QBD processes. *Journal of Applied Probability* 33(4):1003–1010.
2. El-Rayes, A., Kwiatkowska, M., and Norman, G. 1999. Solving infinite stochastic process algebra models through matrix-geometric methods. In *Proceedings of 7th Process Algebras and Performance Modelling Workshop (PAPM 99)*. University of Zaragoza, 41–62.
3. Elhafsi, E.H., and Molle, M. 2006. The effect of queue structure and asymmetric downstream path delays on the performance of dynamic routing. Technical Report, Computer Science Department, University of California, Riverside, CR-CS-2006-05029, December.
4. Elhafsi, E.H., Molle, M., and Manjunath, D. 2005. Can we use product-form solution techniques in networks with asymmetric paths? In *Proceedings of the International Symposium on Performance Evaluation of Computer and Telecommuniaction Systems* (SPECTS 2005), 1–2.
5. Gun, L., and Makowski, A.M. 1987. Matrix-geometric solution for finite capacity queues with phase-type distributions. *Proceedings of Performance 87*. Brussels, Belgium, 269–282.
6. Gun, L., and Makowski, A.M. 1987. Matrix-geometric solution for two node tandem queueing systems with phase-type servers subject to blocking and failures. *SRC Technical Report 87-210*. Systems Research Center, University of Maryland, College Park, MD.
7. Hajek, B. 1982. Birth-and-death processes on the integers with phases and general boundaries. *Journal of Applied Probability* 19(3):488–499.
8. Lin, W., and Kumar, P.R. 1984. Optimal control of a queueing system with two heterogeneous servers. *IEEE Transactions on Automatic Control* 29:696–703.
9. Neuts, M.F. 1980. *Matrix Geometric Solutions in Stochastic Models*. John Hopkins University Press, Baltimore.
10. Neuts, M.F., and Lacantoni, D.M. 1979. A Markovian queue with N servers subject to breakdowns and repairs. *Management Science* 25:849–861.
11. Rao, B.M., and Posner, M.J.M. 1977. Algorithmic and approximation analyses of the shorter queue model. *Naval Research Logistics* 34:381–398.
12. Riska, A., and Smirni, E. 2002. M/G/1-type markov processes: A tutorial. http://www.cs.wm.edu/riska/paper-MG1-tutorial.pdf.
13. Li, S. 1989. Overload control in a finite message storage buffer. *IEEE Transactions on Communications* 37(12):1330–1338.
14. Tandra, R., Hemachandra, N., and Manjunath, D. 2002. Join minimum cost queue for multi-class customers: Stability and performance bounds. In *Probability in the Engineering and Informational Sciences*. Cambridge University Press, New York, vol. 18, 445–472.
15. van Houtoum, G.J., Adan, I.J.B.F., Wessesl, J., and Zijm, W.H.M. 2000. Performance analysis of parallel identical machines with a generalized shortest queue arrival mechanism. *OR Specktrum* 23:411–428.
16. Le Boudec, J.Y. 1991. An efficient solution method for Markov models of ATM links with loss priorities. *IEEE Journal on Selected Areas in Communications* 9(3):408–417.