Transient analysis of piecewise homogeneous QBD process*

Salah Al-Deen Almousa[1], Gábor Horváth[1] and Miklós Telek[1,2]

[1] Department of Networked Systems and Services,
Technical University of Budapest, Budapest, Hungary
[2] MTA-BME Information Systems Research Group, Budapest, Hungary

e-mail: {ghorvath,almousa,telek}@hit.bme.hu

Abstract

The paper presents a numerical analysis approach for the transient solution of a piecewise homogeneous quasi-birth-death (QBD) process. The proposed approach computes the transient probabilities based on the linear combination of matrix geometric series in Laplace transform domain, and builds on the availability of an efficient numerical inverse Laplace transformation method.

Keywords: piecewise homogeneous QBD process, transient analysis, Laplace transform, numerical analysis, matrix geometric series.

1 Introduction

A continuous time quasi-birth-death (QBD) process is a continuous time Markov chain with a regular block structure in the generator matrix [8, 6]. QBD processes are efficiently used in modelling and analysis of various telecommunication systems and queueing models [5]. Most of the analysis results and application examples of QBD processes focus on the stationary behaviour. The transient analysis of QBD processes has also been considered for a long time [8], but received much less attention. One of the reasons for this moderate attention is the relative complexity of the transient analysis methods compared to the stationary ones. A natural methodology for transient analysis is to express the transient measures in Laplace domain and to inverse transform the result. However, this methodology has not gained significant popularity due to the shortcomings of numerical inverse transformation methods available. Motivated by the availability of a recently introduced enhanced numerical inverse Laplace transformation

*This work is partially supported by the OTKA K-123914 and the TUDFO/51757/2019-ITM grants

(LT) method [4], we reconsidered the Laplace transform-based transient analysis of the QBD process and provide an efficient analysis method for a fairly general set of QBD models.

As an essential ingredient of the proposed numerical procedure we provide the transient description of QBD processes between two boundaries and the two boundaries extension of the $\mathbf{R}(s)$ and $\mathbf{G}(s)$ matrices. To the best of the authors' knowledge, Theorem 1 and 2 and Appendix A provides the most comprehensive discussion of such measures.

Most of this paper is devoted to the transform domain description of the performance measures of interest. Based on that, we can simply apply the concentrated matrix exponential distribution (CME) based numerical inverse LT method [4] to obtain time domain results. The core of the numerical inverse LT method is to evaluate the LT function calculated in this work at some points.

In order to improve the readability of the paper, we first present the analysis for a simple set of piecewise homogeneous QBD processes and later we generalize the considered set of models. Our Mathematica implementation [1] solves the general set of models.

The rest of the paper is organized as follows. Section 2 discusses the preliminaries and the related results in the literature. Section 3 presents the considered piecewise homogeneous QBD process and the performance measures of interest. The principal matrices of homogeneous QBD processes required for the subsequent analysis are provided in Section 4. The main part of the analysis is presented in Section 5 and the summary of the proposed numerical methods in Section 6. Section 7 explains how to apply the procedure for more general QBD models and finally, Section 8 demonstrates the numerical properties of the proposed method.

2 Related works

Some of the transient measures have been considered in early works on QBD processes. E.g., the analysis of the busy period of a level homogeneous infinite QBD is discussed in [8]. In that work, the transform-domain description of the busy period is used to compute the mean busy period, by the derivative of the LT domain description at $s = 0$. In spite of the availability of transform-domain descriptions, to compute transient measures at some time points, time-domain analysis methods (numerical solution of matrix equations with convolution integral, or differential equations) are used.

Some of the few works where transient measures are computed by inverse Laplace transformation of the transform-domain description are [12, 13]. In these works a special variant of the matrix geometric solution is computed for homogeneous infinite QBD processes, where the behavior at level zero differs from the one at higher levels. In [12], it is assumed that the process starts from level zero. In [13], the QBD process can start from arbitrary initial level, but the provided solution is tailored to a specific model of a voice/data integration network. With respect to the numerical accuracy of their method the authors

claim that "the inverse LT operation is an inherently unstable operation" [12, p. 485].

This general belief turned the attention towards time-domain methods for computing transient measures, which resulted in quite sophisticated time-domain analysis methods over the decades. For example, the stationary solution of some appropriately defined stochastic system was used to compute the transient of a QBD process in [3, 10, 11]. A different approach of this trend was to apply a variant of the folding method [7].

In this paper we proposed a general transient analysis method, which allows finite and infinite, piecewise level homogeneous QBD processes with potentially different number of phases in each homogeneous regime without any restriction on the initial and final level.

Throughout the paper, matrices are always denoted by bold letters, the "hat" on the top refers to a reversed quantity with respect to the same quantity without "hat", and underlined letters indicate time-domain expressions.

3 Model definition

This paper aims to calculate the transient distribution of piecewise homogeneous quasi birth-death (PHQBD) processes in an efficient way. PHQBDs are two dimensional continuous time Markov processes $\{\mathcal{X}(t), \mathcal{J}(t), t > 0\}$, where $\mathcal{X}(t) \in \{0, 1, \ldots, M\}$ is the so-called level process and $\mathcal{J}(t) \in \{1, 2, \ldots, N\}$ is the phase process. There are K regimes, and the regime boundaries (also referred to as *thresholds* in the sequel) are denoted by $0 = T_0 < T_1 < \cdots < T_K = M$. The transition rates are spatially homogeneous in every regime of the system. In regime k, matrix $\mathbf{B_k}$ holds the rates of the level backward transitions, $\mathbf{F_k}$ the rates of the level forward transitions, and $\mathbf{L_k}$ the ones of the local transitions that are not accompanied by the change of the level. At the thresholds the behavior of the local transitions can differ from the regular ones. At T_k the local transition matrix is $\mathbf{L'_{k+1}}$,

The generator matrix $\mathbf{Q}$ of the Markov chain has the following block tridiagonal structure

$$
\mathbf{Q} = \begin{array}{c}
\begin{matrix} T_0 & \quad & \quad & T_1 & \quad & \quad & T_2 & \quad & \ldots & \quad & T_K \end{matrix} \\
\left(\begin{array}{ccccccccccccc}
\mathbf{L_1'} & \mathbf{F_1} & & & & & & & & & & & \\
\mathbf{B_1} & \mathbf{L_1} & \mathbf{F_1} & & & & & & & & & & \\
 & \ddots & \ddots & \ddots & & & & & & & & & \\
 & & \mathbf{B_1} & \mathbf{L_1} & \mathbf{F_1} & & & & & & & & \\
 & & & \mathbf{B_1} & \mathbf{L_2'} & \mathbf{F_2} & & & & & & & \\
 & & & & \mathbf{B_2} & \mathbf{L_2} & \mathbf{F_2} & & & & & & \\
 & & & & & \ddots & \ddots & \ddots & & & & & \\
 & & & & & & \mathbf{B_2} & \mathbf{L_2} & \mathbf{F_2} & & & & \\
 & & & & & & & \mathbf{B_2} & \mathbf{L_3'} & \mathbf{F_3} & & & \\
 & & & & & & & & \mathbf{B_3} & \mathbf{L_3} & \mathbf{F_3} & & \\
 & & & & & & & & & \ddots & \ddots & \ddots & \\
 & & & & & & & & & & \mathbf{B_K} & \mathbf{L_K} & \mathbf{F_K} \\
 & & & & & & & & & & & \mathbf{B_K} & \mathbf{L_{K+1}'}
\end{array}\right)
\begin{array}{c} T_0 \\ \\ \\ \\ T_1 \\ \\ \\ \\ T_2 \\ \\ \vdots \\ \\ T_K \end{array}
\end{array}.
$$

Our target is to calculate the transition probabilities of the PHQBD from arbitrary initial to arbitrary final level in time t. That is, we would like to compute $\underline{\mathbf{V}}(t,n,m)$, whose elements, for $0 \le n, m \le T_K$ and $1 \le i, j \le N$, are defined by

$$
[\underline{\mathbf{V}}(t,n,m)]_{i,j} = Pr(\mathcal{X}(t) = m, \mathcal{J}(t) = j | \mathcal{X}(0) = n, \mathcal{J}(0) = i). \tag{1}
$$

The LT of $\underline{\mathbf{V}}(t,n,m)$ is given by $\mathbf{V}(s,n,m) = \int_0^\infty \underline{\mathbf{V}}(t,n,m)\, e^{-st}\, dt$. Additionally, we intend to compute the mean of the level process at time t. That is, for $0 \le n \le T_K$ and $1 \le i, j \le N$, we look for

$$
[\underline{\mathbf{S}}(t,n)]_{i,j} = \sum_{m=0}^{T_K} m [\underline{\mathbf{V}}(t,n,m)]_{i,j}, \tag{2}
$$

whose LT is $\mathbf{S}(s,n) = \int_0^\infty \underline{\mathbf{S}}(t,n)\, e^{-st}\, dt$.

4 Principal quantities with homogeneous transition structure

In this section, we assume the special case when $K = 1$ and we omit the index referring to the regime. That is, we write $\mathbf{B}, \mathbf{L}, \mathbf{F}$ instead of $\mathbf{B_1}, \mathbf{L_1}, \mathbf{F_1}$.

4.1 Level visit and first hitting probabilities with one boundary

In the matrix-analytic methods, the stationary solution of QBDs relies on specific matrices, including matrix $\mathbf{R}$, whose elements provide the mean time spent at level n, starting with a jump at level $n-1$, before the first return to level $n-1$, and matrix $\mathbf{G}$, consisting of the phase transition probabilities at the first

return to level $n-1$ starting from level n [8, 6]. In the definition of these matrices, level $n-1$ can be considered as a boundary, when the QBD reaches this level, we stop observing it.

For the transient analysis we use the time-dependent counterparts of these matrices [6, 5]

$$[\underline{\mathbf{U}}(t)]_{i,j} \triangleq Pr(\mathcal{X}(t)=n, \mathcal{J}(t)=j, \gamma_{n-1}>t \,|\, \mathcal{X}(0)=n, \mathcal{J}(0)=i), \tag{3}$$

$$\begin{aligned}[\underline{\mathbf{R}}(t)]_{i,j} &\triangleq [\mathbf{F}\underline{\mathbf{U}}(t)]_{i,j} \\ &= \lim_{\Delta\to 0}\frac{1}{\Delta} Pr\Big(\mathcal{X}(t)=n, \mathcal{J}(t)=j, \gamma_{n-1}>t, \text{transition in } (-\Delta,0) \\ &\qquad\qquad \Big| \; \mathcal{X}(-\Delta)=n-1, \mathcal{J}(-\Delta)=i\Big)\end{aligned} \tag{4}$$

$$\begin{aligned}[\underline{\mathbf{G}}(t)]_{i,j} &\triangleq [\int_{\tau=0}^{t} \underline{\mathbf{U}}(\tau)d\tau\mathbf{B}]_{i,j} \\ &= Pr(\mathcal{J}(\gamma_{n-1})=j, \gamma_{n-1}<t \,|\, \mathcal{X}(0)=n, \mathcal{J}(0)=i),\end{aligned} \tag{5}$$

where $\gamma_n = \min(t|t>0, \mathcal{X}(t)=n)$ is the time of the first visit to level n. In (4), $\gamma_{n-1}>t$ has an important implicit meaning, namely that $\mathcal{X}(0)=n$ (or $\mathcal{X}(0)>n$ with a negligibly small probability). The Laplace-Stieltjes transform (LST) of matrix $\underline{\mathbf{G}}(t)$ is $\mathbf{G}(s)=\int_0^\infty e^{-st}\,d\underline{\mathbf{G}}(t)$, while the LT of $\underline{\mathbf{R}}(t)$ is given by $\mathbf{R}(s)=\int_0^\infty \underline{\mathbf{R}}(t)\,e^{-st}\,dt$. Note that $\mathbf{G}(0)=\mathbf{G}$ and $\mathbf{R}(0)=\mathbf{R}$ hold.

For a homogeneous QBD (with level forward and level backward transitions being independent of the level) and s with positive real part, i.e., $\Re(s)>0$, the solution of the matrix quadratic equation

$$s\mathbf{G}(s) = \mathbf{B} + \mathbf{L}\mathbf{G}(s) + \mathbf{F}\mathbf{G}^2(s) \tag{6}$$

with eigenvalues inside the unit disk provides $\mathbf{G}(s)$ [6] (see also Appendix B). The powers of matrix $\mathbf{G}(s)$ have the following important probabilistic interpretations

$$[\mathbf{G}^n(s)]_{i,j} = \int_{t=0}^{\infty} e^{-st}\; dPr(\mathcal{J}(\gamma_0)=j, \gamma_0<t \,|\, \mathcal{X}(0)=n, \mathcal{J}(0)=i). \tag{7}$$

Similarly, for a homogeneous QBD and $\Re(s)>0$, the solution of the matrix quadratic equation

$$s\mathbf{R}(s) = \mathbf{F} + \mathbf{R}(s)\mathbf{L} + \mathbf{R}^2(s)\mathbf{B} \tag{8}$$

with eigenvalues inside the unit disk provides $\mathbf{R}(s)$ [9]. The powers of the matrix $\mathbf{R}(s)$ have a similar important probabilistic interpretations

$$\begin{aligned}[\mathbf{R}^m(s)]_{i,j} = \int_0^\infty \lim_{\Delta\to 0}\frac{1}{\Delta} Pr\Big(&\mathcal{X}(t)=m, \mathcal{J}(t)=j, \gamma_0>t, \\ &\text{transition in } (-\Delta,0) \;\Big|\; \mathcal{X}(-\Delta)=0, \mathcal{J}(-\Delta)=i\Big)\cdot e^{-st}\,dt.\end{aligned} \tag{9}$$

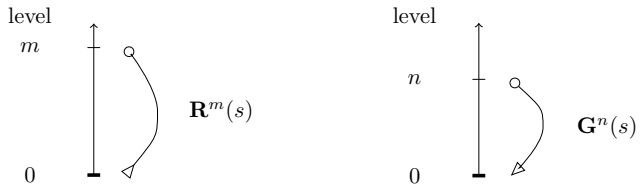

Figure 1: Principal quantities with a lower boundary at 0 (arrow refers to jump, circle refers to sojourn)

That is, $G^n(s)$ is the LST transform with respect to time of the distribution of $\gamma_0, J(\gamma_0)$, given that the process is in level n at time 0. $R^m(s)$ is the LT with respect to t of the probability that the process is in level m at time t and that $\gamma_0 > t$, given that it jumps from level 0 to level 1 at time 0. $\mathbf{G}^n(s)$ is concluded by a downward state transition while $\mathbf{R}^m(s)$ starts with an upward state transition. See Figure 1 for a visual representation of these principal quantities.

In the sequel we are going to make use of the $\mathbf{G}(s)$ and $\mathbf{R}(s)$ matrices of the level-reversed QBD, too (where the roles of the level forward and level backward transitions are swapped), which are denoted by $\hat{\mathbf{G}}(s)$ and $\hat{\mathbf{R}}(s)$ and satisfy the quadratic equations

$$s\hat{\mathbf{G}}(s) = \mathbf{F} + \mathbf{L}\hat{\mathbf{G}}(s) + \mathbf{B}\hat{\mathbf{G}}^2(s), \tag{10}$$

$$s\hat{\mathbf{R}}(s) = \mathbf{B} + \hat{\mathbf{R}}(s)\mathbf{L} + \hat{\mathbf{R}}^2(s)\mathbf{F}. \tag{11}$$

Having $\mathbf{G}(s)$ and $\hat{\mathbf{G}}(s)$ from (6) and (10), $\mathbf{R}(s)$ and $\hat{\mathbf{R}}(s)$ can also be computed [5] from

$$\mathbf{R}(s) = \mathbf{F}(s\mathbf{I} - \mathbf{L} - \mathbf{F}\mathbf{G}(s))^{-1} \tag{12}$$

$$\hat{\mathbf{R}}(s) = \mathbf{B}(s\mathbf{I} - \mathbf{L} - \mathbf{B}\hat{\mathbf{G}}(s))^{-1}. \tag{13}$$

4.2 Level visit probabilities with two boundaries

In this section we still investigate the level-homogeneous (level independent) case, but with two boundaries, 0 and b ($b > 0$), instead of just one boundary. We are going to introduce counterpart measures of $\mathbf{R}(s)$ and $\mathbf{G}(s)$ with two boundaries. The probability of visiting level m ($m \in (0, b)$) at time t, starting from level 0 or b with a state transition, given that the level process of the QBD remains in $(0, b)$ in the $(0, t)$ time interval is

$$[\underline{\mathbf{Z}}^{(b)}(t, m)]_{i,j} = \lim_{\Delta \to 0} \frac{1}{\Delta} Pr\Big(\mathcal{X}(t) = m, \mathcal{J}(t) = j, \xi_{0,b} > t, \\ \text{transition in } (-\Delta, 0) \;\Big|\; \mathcal{X}(-\Delta) = 0, \mathcal{J}(-\Delta) = i \Big), \tag{14}$$

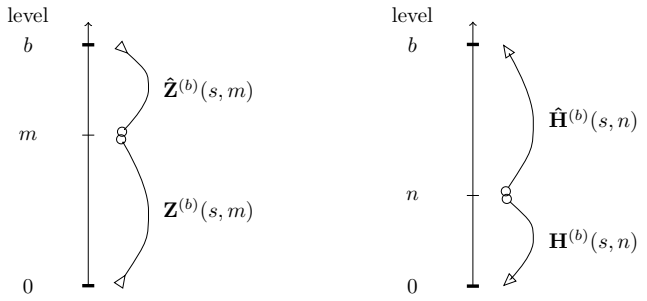

Figure 2: Principal quantities with two boundaries 0 and b (arrow refers to jump, circle refers to sojourn)

$$[\hat{\underline{\mathbf{Z}}}^{(b)}(t,m)]_{i,j} = \lim_{\Delta\to 0}\frac{1}{\Delta}Pr\Big(\mathcal{X}(t)=m,\mathcal{J}(t)=j,\xi_{0,b}>t, \\ \text{transition in } (-\Delta,0) \;\Big|\; \mathcal{X}(-\Delta)=b,\mathcal{J}(-\Delta)=i\Big), \quad (15)$$

where $(0,b)=[0,b]\setminus\{0,b\}$ and $\xi_{0,b}=\min(t|\mathcal{X}(t)\notin(0,b))$ is the first time when the level process leaves the $(0,b)$ interval. See Figure 2 for a visual representation of these principal quantities.

The LT of $\underline{\mathbf{Z}}^{(b)}(t,m)$ and $\hat{\underline{\mathbf{Z}}}^{(b)}(t,m)$ are $\mathbf{Z}^{(b)}(s,m)=\int_0^\infty e^{-st}\underline{\mathbf{Z}}^{(b)}(t,m)dt$ and $\hat{\mathbf{Z}}^{(b)}(s,m)=\int_0^\infty e^{-st}\hat{\underline{\mathbf{Z}}}^{(b)}(t,m)dt$.

Theorem 1. *$\mathbf{Z}^{(b)}(s,m)$ and $\hat{\mathbf{Z}}^{(b)}(s,m)$ can be computed from*

$$\mathbf{Z}^{(b)}(s,m)=\mathbf{M}(s,b)\mathbf{R}^m(s)-\hat{\mathbf{M}}(s,b)\mathbf{R}^b(s)\hat{\mathbf{R}}^{b-m}(s), \quad (16)$$

$$\hat{\mathbf{Z}}^{(b)}(s,m)=\hat{\mathbf{M}}(s,b)\hat{\mathbf{R}}^{b-m}(s)-\mathbf{M}(s,b)\hat{\mathbf{R}}^b(s)\mathbf{R}^m(s), \quad (17)$$

where $\mathbf{M}(s,b)=\left(\mathbf{I}-\mathbf{R}^b(s)\hat{\mathbf{R}}^b(s)\right)^{-1}$ *and* $\hat{\mathbf{M}}(s,b)=\left(\mathbf{I}-\hat{\mathbf{R}}^b(s)\mathbf{R}^b(s)\right)^{-1}$.

With respect to the combination of two matrix exponential/geometric terms, a similar relation is used in the analysis of Markov fluid models [2, Equation (4)], but we were not aware of this relation used for QBD processes previously.

Proof. According to (9), $[\underline{\mathbf{R}}^m(t)]_{ij}$ is the probability that starting with a jump from level 0 phase i the QBD process stays in level m phase j at time t such that level 0 is not reached before t. This can happen in two different ways. Either the level process remains in levels $[1,b-1]$ during the whole $(0,t)$ interval or it reaches level b before time t. In the second case, there is an epoch τ in $(0,t)$ when the process leaves level b and remains in $(0,b)$ during the interval of time (τ,t). Based on these two different cases $[\underline{\mathbf{R}}^m(t)]_{ij}$ can be written as

$$[\underline{\mathbf{R}}^m(t)]_{ij}=[\underline{\mathbf{Z}}^{(b)}(t,m)]_{ij}+\sum_k\int_{\tau=0}^t[\underline{\mathbf{R}}^b(\tau)]_{ik}[\hat{\underline{\mathbf{Z}}}^{(b)}(t-\tau,m)]_{kj}d\tau, \quad (18)$$

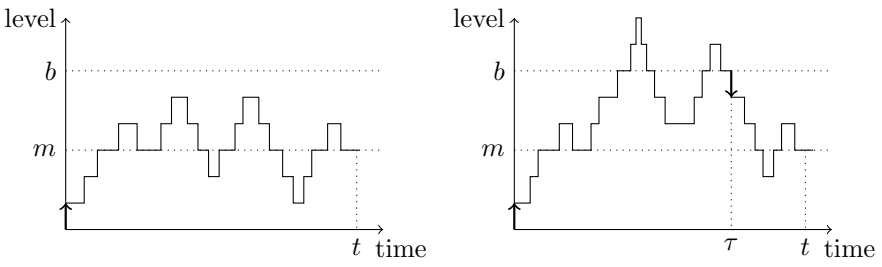

Figure 3: Two trajectories to reach level m at time t

where the second term of the right hand side means that the process stays at level b phase k right before time τ. It has a jump to level $b-1$ at τ and in the whole (τ, t) interval it remains in levels $[1, b-1]$, such that at time t it stays at level m phase j. Figure 3 depicts the possible trajectories represented by (18). The matrix form of the Laplace transform of (18) is

$$\mathbf{R}^m(s) = \mathbf{Z}^{(b)}(s,m) + \mathbf{R}^b(s)\hat{\mathbf{Z}}^{(b)}(s,m). \tag{19}$$

Along similar arguments we can derive the same relation for the level-reversed QBD, leading to

$$\hat{\mathbf{R}}^{b-m}(s) = \hat{\mathbf{Z}}^{(b)}(s,m) + \hat{\mathbf{R}}^b(s)\mathbf{Z}^{(b)}(s,m). \tag{20}$$

Merging (19) and (20) into a single matrix equation gives

$$\begin{bmatrix} \mathbf{R}^m(s) \\ \hat{\mathbf{R}}^{b-m}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{R}^b(s) \\ \hat{\mathbf{R}}^b(s) & \mathbf{I} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{Z}^{(b)}(s,m) \\ \hat{\mathbf{Z}}^{(b)}(s,m) \end{bmatrix}. \tag{21}$$

Multiplying both sides of (21) by

$$\begin{bmatrix} \mathbf{I} & \mathbf{R}^b(s) \\ \hat{\mathbf{R}}^b(s) & \mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} \left(\mathbf{I} - \mathbf{R}^b(s)\hat{\mathbf{R}}^b(s)\right)^{-1} & -\left(\mathbf{I} - \mathbf{R}^b(s)\hat{\mathbf{R}}^b(s)\right)^{-1}\mathbf{R}^b(s) \\ -\left(\mathbf{I} - \hat{\mathbf{R}}^b(s)\mathbf{R}^b(s)\right)^{-1}\hat{\mathbf{R}}^b(s) & \left(\mathbf{I} - \hat{\mathbf{R}}^b(s)\mathbf{R}^b(s)\right)^{-1} \end{bmatrix}$$

from the left we obtain (16) and (17). □

The matrix geometric form in (16) and (17) allows to compute various sums

of $\mathbf{Z}^{(b)}(s,m)$ and $\hat{\mathbf{Z}}^{(b)}(s,m)$ in closed form:

$$\overline{\mathbf{Z}}(s,b) \triangleq \sum_{m=1}^{b-1} \mathbf{Z}^{(b)}(s,m) = \mathbf{M}(s,b)\overline{\mathbf{R}}(s,b) - \hat{\mathbf{M}}(s,b)\mathbf{R}^b(s)\overline{\hat{\mathbf{R}}}(s,b),$$

$$\overrightarrow{\mathbf{Z}}(s,b) \triangleq \sum_{m=1}^{b-1} m\mathbf{Z}^{(b)}(s,m) = \mathbf{M}(s,b)\overrightarrow{\mathbf{R}}(s,b) - \hat{\mathbf{M}}(s,b)\mathbf{R}^b(s)\overleftarrow{\hat{\mathbf{R}}}(s,b),$$

$$\overline{\hat{\mathbf{Z}}}(s,b) \triangleq \sum_{m=1}^{b-1} \hat{\mathbf{Z}}^{(b)}(s,b-m) = \hat{\mathbf{M}}(s,b)\overline{\hat{\mathbf{R}}}(s,b) - \mathbf{M}(s,b)\hat{\mathbf{R}}^b(s)\overline{\mathbf{R}}(s,b),$$

$$\overleftarrow{\hat{\mathbf{Z}}}(s,b) \triangleq \sum_{m=1}^{b-1} m\hat{\mathbf{Z}}^{(b)}(s,b-m) = \hat{\mathbf{M}}(s,b)\overrightarrow{\hat{\mathbf{R}}}(s,b) - \mathbf{M}(s,b)\hat{\mathbf{R}}^b(s)\overleftarrow{\mathbf{R}}(s,b),$$

where

$$\overline{\mathbf{R}}(s,b) \triangleq \sum_{m=1}^{b-1} \mathbf{R}^m(s) = \sum_{m=1}^{b-1} \mathbf{R}^{b-m}(s) = \left(\mathbf{R}(s) - \mathbf{R}^b(s)\right)\left(\mathbf{I} - \mathbf{R}(s)\right)^{-1},$$

$$\overrightarrow{\mathbf{R}}(s,b) \triangleq \sum_{m=1}^{b-1} m\mathbf{R}^m(s) = \sum_{m=1}^{b-1} (b-m)\mathbf{R}^{b-m}(s)$$

$$= \left(\mathbf{I} - \mathbf{R}^b(s)\right)\left(\mathbf{I} - \mathbf{R}(s)\right)^{-2} - \left(\mathbf{I} + (b-1)\mathbf{R}^b(s)\right)\left(\mathbf{I} - \mathbf{R}(s)\right)^{-1},$$

$$\overleftarrow{\mathbf{R}}(s,b) \triangleq \sum_{m=1}^{b-1} m\mathbf{R}^{b-m}(s) = \sum_{m=1}^{b-1} (b-m)\mathbf{R}^m(s) = b\overline{\mathbf{R}}(s,b) - \overrightarrow{\mathbf{R}}(s,b),$$

and $\overline{\hat{\mathbf{R}}}(s,b)$, $\overrightarrow{\hat{\mathbf{R}}}(s,b)$ and $\overleftarrow{\hat{\mathbf{R}}}(s,b)$ are defined similarly based on $\hat{\mathbf{R}}(s)$. We note that for $\Re(s) > 0$, the spectral radius of both, $\mathbf{R}(s)$ and $\hat{\mathbf{R}}(s)$, is less than one.

4.3 Level hitting probabilities with two boundaries

Having two boundaries 0 and $b > 1$, the phase transition probability matrices for hitting either of the boundaries 0 or b are defined by

$$[\underline{\mathbf{H}}^{(b)}(t,n)]_{i,j} = Pr(\mathcal{J}(\gamma_0) = j, \gamma_0 < t, \gamma_0 < \gamma_b | \mathcal{X}(0) = n, \mathcal{J}(0) = i), \tag{22}$$

$$[\underline{\hat{\mathbf{H}}}^{(b)}(t,n)]_{i,j} = Pr(\mathcal{J}(\gamma_b) = j, \gamma_b < t, \gamma_b < \gamma_0 | \mathcal{X}(0) = n, \mathcal{J}(0) = i), \tag{23}$$

for $n \in (0,b)$, with LST defined by $\mathbf{H}^{(b)}(s,n) = \int_{t=0}^{\infty} e^{-st} d\underline{\mathbf{H}}^{(b)}(t,n)$ and $\hat{\mathbf{H}}^{(b)}(s,n) = \int_{t=0}^{\infty} e^{-st} d\underline{\hat{\mathbf{H}}}^{(b)}(t,n)$.

Theorem 2. *$\mathbf{H}^{(b)}(s,n)$ and $\hat{\mathbf{H}}^{(b)}(s,n)$ satisfy*

$$\mathbf{H}^{(b)}(s,n) = \mathbf{G}^n(s)\mathbf{N}(s,b) - \hat{\mathbf{G}}^{b-n}(s)\hat{\mathbf{N}}(s,b)\hat{\mathbf{G}}^b(s), \tag{24}$$

$$\hat{\mathbf{H}}^{(b)}(s,n) = \hat{\mathbf{G}}^{b-n}(s)\hat{\mathbf{N}}(s,b) - \mathbf{G}^n(s)\mathbf{N}(s,b)\mathbf{G}^b(s), \tag{25}$$

where $\mathbf{N}(s,b) = \left(\mathbf{I} - \mathbf{G}^b(s)\hat{\mathbf{G}}^b(s)\right)^{-1}$ *and* $\hat{\mathbf{N}}(s,b) = \left(\mathbf{I} - \hat{\mathbf{G}}^b(s)\mathbf{G}^b(s)\right)^{-1}$.

Proof. The proof follows a similar reasoning as the proof of Theorem 1. According to (7), $[\underline{\mathbf{G}}^n(t)]_{ij}$ is the probability that starting from level n phase i the QBD process reaches level 0 in phase j before time t. This can happen in two different ways. Either the process reaches level 0 before level b or it reaches level b before reaching level 0. In the second case, let θ denote the *first* instance when the process visits level b. According to these two different cases $[\underline{\mathbf{G}}^n(t)]_{ij}$ can be written as

$$[\underline{\mathbf{G}}^n(t)]_{ij} = [\underline{\mathbf{H}}^{(b)}(t,n)]_{ij} + \sum_k \int_{\theta=0}^{t} [\underline{\hat{\mathbf{H}}}^{(b)}(\theta,n)]_{ik} [\underline{\mathbf{G}}^b(t-\theta)]_{kj} d\theta, \tag{26}$$

where the second term of the right hand side means that the process reaches level b phase k at time θ and starting from level b phase k it reaches level 0 phase j within the remaining $t-\theta$ interval. The matrix form of the Laplace-Stieltjes transform of (26) is

$$\mathbf{G}^n(s) = \mathbf{H}^{(b)}(s,n) + \hat{\mathbf{H}}^{(b)}(s,n)\mathbf{G}^b(s). \tag{27}$$

Similarly, for the level-reversed QBD we get

$$\hat{\mathbf{G}}^{b-n}(s) = \hat{\mathbf{H}}^{(b)}(s,n) + \mathbf{H}^{(b)}(s,n)\hat{\mathbf{G}}^b(s). \tag{28}$$

Equations (27) and (28) can be merged to

$$\begin{bmatrix} \mathbf{G}^n(s) & \hat{\mathbf{G}}^{b-n}(s) \end{bmatrix} = \begin{bmatrix} \mathbf{H}^{(b)}(s,n) & \hat{\mathbf{H}}^{(b)}(s,n) \end{bmatrix} \cdot \begin{bmatrix} \mathbf{I} & \hat{\mathbf{G}}^b(s) \\ \mathbf{G}^b(s) & \mathbf{I} \end{bmatrix}. \tag{29}$$

Multiplying both sides of (29) from the right by

$$\begin{bmatrix} \mathbf{I} & \mathbf{G}^b(s) \\ \hat{\mathbf{G}}^b(s) & \mathbf{I} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{N}(s,b) & -\mathbf{N}(s,b)\mathbf{G}^b(s) \\ -\hat{\mathbf{N}}(s,b)\hat{\mathbf{G}}^b(s) & \hat{\mathbf{N}}(s,b) \end{bmatrix},$$

we obtain the explicit solution (24) and (25). □

With respect to the combination of two matrix geometric terms, a result similar to Theorem 2 appeared recently in [3, Proposition 3.5 (with $l = 0$ and $l = C$)]. Appendix A discusses the relation of the $\mathbf{R}^{(b)}(s,m)$ and $\mathbf{H}^{(b)}(s,n)$ type of measures. For $n \in \{0, b\}$ we define $\mathbf{H}^{(b)}(s,0) = \hat{\mathbf{H}}^{(b)}(s,b) = \mathbf{I}$ and $\hat{\mathbf{H}}^{(b)}(s,0) = \mathbf{H}^{(b)}(s,b) = \mathbf{0}$.

The quantities defined in this section assume a completely homogeneous transition structure of the QBD process with backward, forward, and local transitions defined by matrix $\mathbf{B}$, $\mathbf{F}$, and $\mathbf{L}$. For the case of piecewise homogeneous QBD processes these quantities can be computed for each regime based on $\mathbf{B_k}$, $\mathbf{F_k}$, and $\mathbf{L_k}$ for $k = 1, \ldots, K$. The quantities specific to regime k are indicated by subscript k in the sequel.

5 The transient analysis of piecewise homogeneous QBDs

To express the LT of the transient distribution, we start with computing the first return probability matrices $\mathbf{Y}(s,n)$ and $\hat{\mathbf{Y}}(s,n)$, then the transient distribution at some specific levels (including the thresholds), and finally, from these auxiliary quantities we obtain the target quantity $\mathbf{V}(s,n,m)$.

The i,jth elements of the first return probability matrices are defined by

$$[\mathbf{Y}(s,n)]_{i,j} = \int_{t=0}^{\infty} e^{-st}\, dPr(\mathcal{J}(\gamma_n) = j, \gamma_n < t \,|\, \mathcal{X}(0) = n+1, \mathcal{J}(0) = i), \quad (30)$$

$$[\hat{\mathbf{Y}}(s,n)]_{i,j} = \int_{t=0}^{\infty} e^{-st}\, dPr(\mathcal{J}(\gamma_n) = j, \gamma_n < t \,|\, \mathcal{X}(0) = n-1, \mathcal{J}(0) = i). \quad (31)$$

The definition of $\mathbf{Y}(s,n)$ and $\hat{\mathbf{Y}}(s,n)$ would be identical to those of $\mathbf{G}(s)$ and $\hat{\mathbf{G}}(s)$, if the QBD was infinite and completely homogeneous. However, in finite and PHQBD processes the time till the first return to level n (and the corresponding phase transitions) depends on n due to the presence of multiple boundaries.

Theorem 3. *For level $T_{k-1} \le n < T_k, 1 \le k \le K$, matrices $\mathbf{Y}(s,n)$ can be obtained recursively as*

$$\mathbf{Y}(s,n) = \mathbf{H}_k^{(T_k-n)}(s,1) + \hat{\mathbf{H}}_k^{(T_k-n)}(s,1)\Big(s\mathbf{I} - \mathbf{L}'_{\mathbf{k+1}} - \mathcal{I}_{\{k<K\}}\mathbf{F_{k+1}}\mathbf{Y}(s,T_k) - \mathbf{B_k}\hat{\mathbf{H}}_k^{(T_k-n)}(s,T_k-n-1)\Big)^{-1}\mathbf{B_k}\mathbf{H}_k^{(T_k-n)}(s,T_k-n-1), \quad (32)$$

where $\mathcal{I}_{\{\bullet\}}$ denotes the indicator of $\bullet$. Similarly, for level $T_k < n \le T_{k+1}, k \ge 0$, matrices $\hat{\mathbf{Y}}(s,n)$ can be obtained recursively as

$$\hat{\mathbf{Y}}(s,n) = \hat{\mathbf{H}}_{k+1}^{(n-T_k)}(s,n-T_k-1) + \mathbf{H}_{k+1}^{(n-T_k)}(s,n-T_k-1)\Big(s\mathbf{I} - \mathbf{L}'_{\mathbf{k+1}} - \mathbf{F_{k+1}}\mathbf{H}_{k+1}^{(n-T_k)}(s,1) - \mathcal{I}_{\{k>0\}}\mathbf{B_k}\hat{\mathbf{Y}}(s,T_k)\Big)^{-1}\mathbf{F_{k+1}}\hat{\mathbf{H}}_{k+1}^{(n-T_k)}(s,1). \quad (33)$$

Proof. To prove the theorem, for $n, m > u$, we introduce the so called taboo extension [6] of the transient probability defined in (1)

$$[{}_u\underline{\mathbf{V}}(t,n,m)]_{i,j} = Pr(\mathcal{X}(t) = m, \mathcal{J}(t) = j, \gamma_u > t | \mathcal{X}(0) = n, \mathcal{J}(0) = i), \quad (34)$$

and its LT ${}_u\mathbf{V}(s,n,m) = \int_0^{\infty} {}_u\underline{\mathbf{V}}(t,n,m)\, e^{-st}\, dt$. In this taboo extension, it is required that level u is not visited before time t.

To show (32) we study the time and the phase transitions till the first return to level n starting from level $n+1$. Two cases need to be considered. With probability $\mathbf{H}_k^{(T_k-n)}(s,1)$ the level process returns to level n before hitting boundary T_k, or, with probability $\hat{\mathbf{H}}_k^{(T_k-n)}(s,1)$ the T_k boundary is reached before the first return to n. In the later case, conditioning on the last visit to level T_k we have

$$\mathbf{Y}(s,n) = \mathbf{H}_k^{(T_k-n)}(s,1) + \hat{\mathbf{H}}_k^{(T_k-n)}(s,1)\, {}_n\mathbf{V}(s,T_k,T_k)\, \mathbf{B_k}\, \mathbf{H}_k^{(T_k-n)}(s,T_k-n-1),$$

where the second term represents that after reaching level T_k (with probability $\hat{\mathbf{H}}_k^{(T_k-n)}(s,1)$), the process re-visits level T_k without reaching level n ($_n\mathbf{V}(s,T_k,T_k)$), and finally T_k is left backwards ($\mathbf{B_k}$), and level n is reached before returning to T_k again ($\mathbf{H}_k^{(T_k-n)}(s,T_k-n-1)$). For the missing taboo matrix we write

$$
\begin{aligned}
&{}_n\mathbf{V}(s,T_k,T_k) = \Big(s\mathbf{I}-\mathbf{L}'_{\mathbf{k+1}}\Big)^{-1} + \\
&\Big(s\mathbf{I}-\mathbf{L}'_{\mathbf{k+1}}\Big)^{-1}\Big(\mathcal{I}_{\{k<K\}}\mathbf{F_{k+1}}\mathbf{Y}(s,T_k)+\mathbf{B_k}\hat{\mathbf{H}}_k^{(T_k-n)}(s,T_k-n-1)\Big){}_n\mathbf{V}(s,T_k,T_k).
\end{aligned}
$$

The first term of the left hand side describes the case that the process stays at level T_k for the whole time period. The second term describes the case that the process spends some time at level T_k, after that it either moves forward ($\mathbf{F_{k+1}}$ if $k<K$) and returns to level T_k ($\mathbf{Y}(s,T_k)$) or it moves backward ($\mathbf{B_k}$) and returns to level T_k without reaching level n ($\hat{\mathbf{H}}_k^{(T_k-n)}(s,T_k-n-1)$). In both of these cases, a new taboo matrix needs to be considered after returning to level T_k. Multiplying both sides by $(s\mathbf{I}-\mathbf{L}'_{\mathbf{k+1}})$ gives

$$
\begin{aligned}
&{}_n\mathbf{V}(s,T_k,T_k) = \\
&\qquad\Big(s\mathbf{I}-\mathbf{L}'_{\mathbf{k+1}}-\mathcal{I}_{\{k<K\}}\mathbf{F_{k+1}}\mathbf{Y}(s,T_k)-\mathbf{B_k}\hat{\mathbf{H}}_k^{(T_k-n)}(s,T_k-n-1)\Big)^{-1} \qquad (35)
\end{aligned}
$$

from which (32) comes.

The recursive relations for matrix $\hat{\mathbf{Y}}(s,n)$, (33), giving the time and the phase transitions till the first return to level n starting from level $n-1$, can be proved similarly. □

The next theorem provides the LT of the transient distribution given that the initial and the target levels are both thresholds.

Theorem 4. *The transient probability matrix, given that both, the initial and the target level are the same threshold T_k, $k=0,\ldots,K$, can be expressed by*

$$
\mathbf{V}(s,T_k,T_k) = \big(s\mathbf{I}-\mathbf{L}'_{\mathbf{k+1}}-\mathcal{I}_{\{k<K\}}\mathbf{F_{k+1}}\mathbf{Y}(s,T_k)-\mathcal{I}_{\{k>0\}}\mathbf{B_k}\hat{\mathbf{Y}}(s,T_k)\big)^{-1}, \qquad (36)
$$

in LT domain. When the initial level, T_k, and the target level, T_ℓ, are different and $k<\ell$, the transient probability matrix is given by

$$
\begin{aligned}
\mathbf{V}(s,T_k,T_\ell) = &\big(s\mathbf{I}-\mathbf{L}'_{\mathbf{k+1}}-\mathbf{F_{k+1}}\mathbf{H}_{k+1}^{(T_{k+1}-T_k)}(s,1)-\mathcal{I}_{\{k>0\}}\mathbf{B_k}\hat{\mathbf{Y}}(s,T_k)\big)^{-1} \\
&\cdot\mathbf{F_{k+1}}\hat{\mathbf{H}}_{k+1}^{(T_{k+1}-T_k)}(s,1)\mathbf{V}(s,T_{k+1},T_\ell). \qquad (37)
\end{aligned}
$$

If $k>\ell$ hold, the transient probability matrix satisfies

$$
\begin{aligned}
&\mathbf{V}(s,T_k,T_\ell) \\
&= \big(s\mathbf{I}-\mathbf{L}'_{\mathbf{k+1}}-\mathcal{I}_{\{k<K\}}\mathbf{F_{k+1}}\mathbf{Y}(s,T_k)-\mathbf{B_k}\hat{\mathbf{H}}_k^{(T_k-T_{k-1})}(s,T_k-T_{k-1}-1)\big)^{-1} \\
&\quad\cdot\mathbf{B_k}\mathbf{H}_k^{(T_k-T_{k-1})}(s,T_k-T_{k-1}-1)\mathbf{V}(s,T_{k-1},T_\ell). \qquad (38)
\end{aligned}
$$

Proof. (36) can be obtained in the same way as ${}_n\mathbf{V}(s, T_k, T_k)$ in (35).

When $k < \ell$, conditioning on the first time reaching level T_{k+1} leads to (37). The indicator in Equation (37) ensures that there is no backward transition at level 0. Similarly, for $k > \ell$, to obtain (38), we have to condition on the first time when the level process reaches level T_{k-1} for the first time. Again, level T_K needs different treatment since there are no forward transitions at that level. □

Next, we go one step further with the derivation of the transient distribution by allowing any initial level between the boundaries, but the target level still has to be a threshold.

Theorem 5. *The transient probability matrix with initial level $T_{k-1} < n < T_k$ and final level T_ℓ, with $k \le \ell$, can be computed by*

$$\begin{aligned}\mathbf{V}(s, n, T_\ell) = \big(s\mathbf{I} - \mathbf{L_k} - \mathbf{B_k}\hat{\mathbf{Y}}(s, n) - \mathbf{F_k}\mathbf{H}_k^{(T_k-n)}(s, 1)\big)^{-1} \\ \cdot \mathbf{F_k}\hat{\mathbf{H}}_k^{(T_k-n)}(s, 1)\mathbf{V}(s, T_k, T_\ell),\end{aligned} \tag{39}$$

and for $k > \ell$ it can be computed by

$$\begin{aligned}\mathbf{V}(s, n, T_\ell) = \big(s\mathbf{I} - \mathbf{L_k} - \mathbf{B_k}\hat{\mathbf{H}}_k^{(n-T_{k-1})}(s, n - T_{k-1} - 1) - \mathbf{F_k}\mathbf{Y}(s, n)\big)^{-1} \\ \cdot \mathbf{B_k}\mathbf{H}_k^{(n-T_{k-1})}(s, n - T_{k-1} - 1)\mathbf{V}(s, T_{k-1}, T_\ell).\end{aligned} \tag{40}$$

Proof. Conditioning on the first hitting time of level T_k in the first case and of level T_{k-1} in the second case yields the theorem. □

Theorem 6. *For any $n \in (T_{k-1}, T_k), k = 1, \ldots, K$, matrix $\mathbf{V}(s, n, n)$ is given by*

$$\mathbf{V}(s, n, n) = \big(s\mathbf{I} - \mathbf{L_k} - \mathbf{B_k}\hat{\mathbf{Y}}(s, n) - \mathbf{F_k}\mathbf{Y}(s, n)\big)^{-1}. \tag{41}$$

Proof. Equation (41) can be obtained in the same way as ${}_n\mathbf{V}(s, T_k, T_k)$ in (35). □

Finally, from the matrices defined and calculated above, we can obtain the missing ingredients of the transient distribution as follows.

Theorem 7. *Assume that $n \ne m$, $n \in (T_{k-1}, T_k)$ and $m \in (T_{\ell-1}, T_\ell)$. If $k \ne \ell$, the transient probability matrix satisfies*

$$\begin{aligned}\mathbf{V}(s, n, m) = \mathbf{V}(s, n, T_{\ell-1})\mathbf{Z}^{(T_\ell - T_{\ell-1})}(s, m - T_{\ell-1}) \\ + \mathbf{V}(s, n, T_\ell)\hat{\mathbf{Z}}^{(T_\ell - T_{\ell-1})}(s, T_\ell - m).\end{aligned} \tag{42}$$

Furthermore, if $k = \ell$, it is given by

$$\mathbf{V}(s, n, m) = \begin{cases} \mathbf{V}(s, n, n)\mathbf{Z}^{(T_k-n)}(s, m - n) & \\ \quad + \mathbf{V}(s, n, T_k)\hat{\mathbf{Z}}^{(T_k-n)}(s, T_k - m), & \text{for } n < m, \\ \mathbf{V}(s, n, T_{k-1})\mathbf{Z}^{(n-T_{k-1})}(s, m - T_{k-1}) & \\ \quad + \mathbf{V}(s, n, n)\hat{\mathbf{Z}}^{(n-T_{k-1})}(s, n - m), & \text{for } n > m. \end{cases} \tag{43}$$

Proof. To obtain (42), we condition on the last instant when the process leaves a boundary before visiting level m at time t. The first term on the right hand side of (42) describes the case when the last visited boundary was $T_{\ell-1}$ and the second term describes when it was T_ℓ.

Equation (43) follows a similar pattern as (42), but with different boundaries. In case of $n < m$ the two considered boundaries are n and T_k, while for $n > m$ the boundaries are T_{k-1} and n. □

It can be noticed that, due to the structure of the $\mathbf{Z}^{(b)}(s,m)$ and $\hat{\mathbf{Z}}^{(b)}(s,m)$ matrices in (16) and (17), the solution is a combination of two matrix-geometric terms. This fact is widely known in case of the stationary solution, see e.g. [6], but it has not been discussed for the transient solution yet.

The following theorem presents the mean of the level process in Laplace domain.

Theorem 8. *The matrix* $\mathbf{S}(s,n)$ *can be computed as*

$$\mathbf{S}(s,n) = \sum_{\ell=1}^{K} T_\ell \mathbf{V}(s,n,T_\ell) + \sum_{\ell=1}^{K} \overrightarrow{\mathbf{V}_\ell}(s,n),$$

where, for $n \notin (T_{\ell-1}, T_\ell)$

$$\begin{aligned}\overrightarrow{\mathbf{V}_\ell}(s,n) =& \mathbf{V}(s,n,T_{\ell-1}) \left(T_{\ell-1}\overline{\mathbf{Z}_\ell}(s,T_\ell - T_{\ell-1}) + \overrightarrow{\mathbf{Z}_\ell}(s,T_\ell - T_{\ell-1})\right) \\ &+ \mathbf{V}(s,n,T_\ell) \left(T_{\ell-1}\overline{\hat{\mathbf{Z}}_\ell}(s,T_\ell - T_{\ell-1}) + \overleftarrow{\hat{\mathbf{Z}}_\ell}(s,T_\ell - T_{\ell-1})\right). \end{aligned} \tag{44}$$

and for $n \in (T_{\ell-1}, T_\ell)$

$$\begin{aligned}\overrightarrow{\mathbf{V}_\ell}(s,n) =& \mathbf{V}(s,n,T_{\ell-1}) \left(T_{\ell-1}\overline{\mathbf{Z}_\ell}(s,n - T_{\ell-1}) + \overrightarrow{\mathbf{Z}_\ell}(s,n - T_{\ell-1})\right) \\ &+ \mathbf{V}(s,n,n) \Big(T_{\ell-1}\overline{\hat{\mathbf{Z}}_\ell}(s,n - T_{\ell-1}) + \overleftarrow{\hat{\mathbf{Z}}_\ell}(s,n - T_{\ell-1}) \\ &\qquad + n + n\overline{\mathbf{Z}_\ell}(s,T_\ell - n) + \overrightarrow{\mathbf{Z}_\ell}(s,T_\ell - n)\Big) \\ &+ \mathbf{V}(s,n,T_\ell) \left(n\overline{\hat{\mathbf{Z}}_\ell}(s,T_\ell - n) + \overleftarrow{\hat{\mathbf{Z}}_\ell}(s,T_\ell - n)\right) \end{aligned} \tag{45}$$

Proof. We have

$$\begin{aligned}\mathbf{S}(s,n) &= \sum_{m=1}^{T_K} m\mathbf{V}(s,n,m) = \sum_{\ell=1}^{K} T_\ell \mathbf{V}(s,n,T_\ell) + \sum_{\ell=1}^{K} \sum_{m=T_{\ell-1}+1}^{T_\ell - 1} m\mathbf{V}(s,n,m) \\ &= \sum_{\ell=1}^{K} T_\ell \mathbf{V}(s,n,T_\ell) + \overrightarrow{\mathbf{V}_\ell}(s,n),\end{aligned}$$

where $\overrightarrow{\mathbf{V}_\ell}(s,n) = \sum_{m=T_{\ell-1}+1}^{T_\ell - 1} m\mathbf{V}(s,n,m)$. For $n \notin (T_{\ell-1}, T_\ell)$, $\overrightarrow{\mathbf{V}_\ell}(s,n)$ is computed based on (42), which gives (44), while for $n \in (T_{\ell-1}, T_\ell)$ it is computed based on (43), which gives (45). □

6 Summary of the algorithm

The computation of $\mathbf{V}(s,n,m)$ can be broken down to the following steps.

For each regime $k=1,\ldots,K$ where $T_k - T_{k-1} > 1$, calculate

1. $\mathbf{G_k}(s)$ and $\hat{\mathbf{G}}_\mathbf{k}(s)$ from (6) and (10),
2. $\mathbf{R_k}(s)$, and $\hat{\mathbf{R}}_\mathbf{k}(s)$ from (12) and (13),
3. $\mathbf{H}_k^{T_k-T_{k-1}}(s,1)$, $\hat{\mathbf{H}}_k^{T_k-T_{k-1}}(s,1)$ and $\mathbf{H}_k^{T_k-T_{k-1}}(s,T_k-T_{k-1}-1)$, $\hat{\mathbf{H}}_k^{T_k-T_{k-1}}(s,T_k-T_{k-1}-1)$ based on Section 4.3,

For each regime $k=1,\ldots,K$, calculate

4. $\mathbf{Y}(s,T_{k-1})$ and $\hat{\mathbf{Y}}(s,T_k)$ based on Theorem 3,
5. $\mathbf{V}(s,T_k,T_\ell)$ for $\ell=0,\ldots,K$ based on Theorem 4.

Finally,

6. If $n \in (T_{k-1},T_k)$ and $n \le m$ calculate $\mathbf{H}_k^{(T_k-n)}(s,1)$ and $\hat{\mathbf{H}}_k^{(T_k-n)}(s,1)$ based on Section 4.3.
7. If $n \in (T_{k-1},T_k)$ and $n \ge m$ calculate $\mathbf{H}_k^{(n-T_{k-1})}(s,n-T_{k-1}-1)$ and $\hat{\mathbf{H}}_k^{(n-T_{k-1})}(s,n-T_{k-1}-1)$.
8. If $n \in (T_{k-1},T_k)$ and $n = m$ calculate $\mathbf{Y}(s,n)$ and $\hat{\mathbf{Y}}(s,n)$ based on Theorem 3.
9. If $n \in (T_{k-1},T_k)$, $m \in (T_{\ell-1},T_\ell)$ and $k \neq \ell$ calculate the sojourn probability matrices $\mathbf{Z}^{(T_\ell-T_{\ell-1})}(s,m-T_{\ell-1})$ and $\hat{\mathbf{Z}}^{(T_\ell-T_{\ell-1})}(s,T_\ell-m)$ based on Section 4.2.
10. If $n,m \in (T_{k-1},T_k)$ and $n < m$ calculate the sojourn probability matrices $\mathbf{Z}^{(T_k-n)}(s,m-n)$ and $\hat{\mathbf{Z}}^{(T_k-n)}(s,T_k-m)$.
11. If $n,m \in (T_{k-1},T_k)$ and $n > m$ calculate the sojourn probability matrices $\mathbf{Z}^{(n-T_{k-1})}(s,m-T_{k-1})$ and $\hat{\mathbf{Z}}^{(n-T_{k-1})}(s,n-m)$.
12. Calculate the transition probability matrix $\mathbf{V}(s,n,m)$ based on Theorem 5, Theorem 6 or Theorem 7.
13. If $\mathbf{S}(s,n)$ is also needed, compute $\overrightarrow{\mathbf{V}_k}(s,n)$ for $k=1,\ldots,K$ and $\mathbf{S}(s,n)$ based on Theorem 8.

Steps 1-5 are independent of the initial and final levels, n and m, hence they have to be performed only once when transient distribution is needed at several n and m points.

Computational complexity

Each of the matrices to compute is of size $N \times N$, and the computational complexity of obtaining one of such matrices is $\mathcal{O}(N^3)$. In each of Step 1 - Step 4, $\mathcal{O}(K)$ matrices are computed, while in Step 5 $\mathcal{O}(K^2)$ matrices are computed. The rest of the steps have lower order complexity, hence the overall computational complexity of the procedure is $\mathcal{O}(K^2N^3)$. The related memory complexity is $\mathcal{O}(K^2N^2)$.

The provided computational complexity still hide the computational cost to compute the fundamental matrices at the required points. If the CME based numerical inverse Laplace transform method [4] is used with order V and fundamental matrices are computed in Z iterations, then the computational complexity is $\mathcal{O}(K^2N^3VZ)$.

7 Model extensions

7.1 PHQBD with regime dependent number of phases

In the model defined in Section 3, the number of phases is N in each level of each regime. This restriction strongly limits the practical applicability of the model, since many practically interesting queueing models violate this assumption. Fortunately, the presented analytical treatment remains applicable also when the number of phases changes in different regimes.

When $T_{k-1} < T_k - 1$ the $\mathbf{B_k}$, $\mathbf{L_k}$, $\mathbf{F_k}$, and $\mathbf{L'_k}$, matrices should be square in order to obtain dimensionally correct matrix operations in the above detailed numerical procedure. But for $T_{k-1} = T_k - 1$, the number of phases on level T_{k-1} and T_k (denoted by N_{k-1} and N_k) can differ. In this case, matrices $\mathbf{B_k}$, $\mathbf{L_k}$, $\mathbf{F_k}$, and $\mathbf{L'_k}$ of size $N_k \times N_{k-1}$, $N_k \times N_k$, $N_{k-1} \times N_k$ and $N_{k-1} \times N_{k-1}$, respectively, still result in dimensionally correct matrix operations.

This way a PHQBD with regime dependent number of phases can be transformed to a PHQBD where the number of phases change only in regimes of size one. For example, if $N_{T_0} \neq N_{T_1}$ then the following PHQBD description can be applied:

$$
\begin{array}{c}
\begin{array}{cccccc} T_0 & & T_1 & \dots & T_K & \end{array} \\
\begin{pmatrix}
\mathbf{L'_1} & \mathbf{F_1} & & & & \\
\mathbf{B_1} & \mathbf{L_1} & \mathbf{F_1} & & & \\
& \ddots & \ddots & \ddots & & \\
& & \mathbf{B_1} & \mathbf{L_1} & \mathbf{F'_1} & \\
& & & \mathbf{B'_1} & \mathbf{L'_2} & \mathbf{F_2} \\
& & & & \mathbf{B_2} & \mathbf{L_2} \\
& & & & & \ddots
\end{pmatrix}
\begin{array}{c} T_0 \\ \\ \\ \\ T_1 \\ \\ \end{array} \\
\begin{array}{cccc} & & & T_K \end{array}
\end{array}
\Longrightarrow
\begin{array}{c}
\begin{array}{cccccc} T_0 & & T_{1'} & T_{2'} & \dots & T_{K'} \end{array} \\
\begin{pmatrix}
\mathbf{L'_1} & \mathbf{F_1} & & & & \\
\mathbf{B_1} & \mathbf{L_1} & \mathbf{F_1} & & & \\
& \ddots & \ddots & \ddots & & \\
& & \mathbf{B_1} & \mathbf{L'_2} & \mathbf{F_2} & \\
& & & \mathbf{B_2} & \mathbf{L'_3} & \mathbf{F_3} \\
& & & & \mathbf{B_3} & \mathbf{L_3} \\
& & & & & \ddots
\end{pmatrix}
\begin{array}{c} T_0 \\ \\ \\ T_{1'} \\ T_{2'} \\ \\ \end{array} \\
\begin{array}{cccc} & & & T_{K'} \end{array}
\end{array}
$$

A completely level dependent finite QBD with arbitrary finite number of phases on each level can be described by a PHQBD with $T_k = k$.

7.2 Infinite PHQBD

The case when the PHQBD is infinite with a finite number of regimes (i.e., K is finite) is characterized by the following generator

$$\mathbf{Q}_{infinite} = \begin{array}{c} \begin{array}{ccccccccccccc} T_0 & & & T_1 & & & T_2 & \dots & & T_K & & & \\ \end{array} \\ \begin{pmatrix} \mathbf{L_1'} & \mathbf{F_1} & & & & & & & & & & & \\ \mathbf{B_1} & \mathbf{L_1} & \mathbf{F_1} & & & & & & & & & & \\ & \ddots & \ddots & \ddots & & & & & & & & & \\ & & \mathbf{B_1} & \mathbf{L_1} & \mathbf{F_1} & & & & & & & & \\ & & & \mathbf{B_1} & \mathbf{L_2'} & \mathbf{F_2} & & & & & & & \\ & & & & \mathbf{B_2} & \mathbf{L_2} & \mathbf{F_2} & & & & & & \\ & & & & & \ddots & \ddots & \ddots & & & & & \\ & & & & & & \mathbf{B_2} & \mathbf{L_2} & \mathbf{F_2} & & & & \\ & & & & & & & \mathbf{B_2} & \mathbf{L_3'} & \mathbf{F_3} & & & \\ & & & & & & & & \mathbf{B_3} & \mathbf{L_3} & \mathbf{F_3} & & \\ & & & & & & & & & \ddots & \ddots & \ddots & & \\ & & & & & & & & & & \mathbf{B_K} & \mathbf{L_K} & \mathbf{F_K} & \\ & & & & & & & & & & & \mathbf{B_K} & \mathbf{L_{K+1}'} & \mathbf{F_{K+1}} \\ & & & & & & & & & & & & \mathbf{B_{K+1}} & \mathbf{L_{K+1}} & \ddots \\ & & & & & & & & & & & & & \ddots & \ddots \end{pmatrix} \end{array} \begin{array}{c} T_0 \\ \\ \\ \\ T_1 \\ \\ \\ \\ T_2 \\ \\ \\ \\ T_K \\ \\ \\ \end{array}.$$

The analysis of this case requires the following slight modifications of the above detailed procedure:

- Compute $\mathbf{G}_{K+1}(s)$ and $\mathbf{R}_{K+1}(s)$ and based on them, the required $\mathbf{H}_{K+1}^{(\bullet)}(s,\bullet)$, $\hat{\mathbf{H}}_{K+1}^{(\bullet)}(s,\bullet)$, $\mathbf{Z}_{K+1}^{(\bullet)}(s,\bullet)$ and $\hat{\mathbf{Z}}_{K+1}^{(\bullet)}(s,\bullet)$ matrices.
- For $n \geq T_K$, $\mathbf{Y}(s,n) = \mathbf{G}_{K+1}(s)$ and (33) remains valid for $k = K+1$.
- Remove $\mathcal{I}_{\{k<K\}}$ from all expressions, e.g., (32), (35), (36) and (38).
- For $n > T_K$, (40) and (41) remain valid with $k = K+1$.
- Theorem 7 remains valid for $\max(n,m) \leq T_K$. The cases when $\max(n,m) > T_K$ can be computed based on

$$\mathbf{V}(s,n,m) = \begin{cases} \mathbf{V}(s,n,z_n)\mathbf{R}_{K+1}^{m-z_n}(s) & \text{if } n < m, \\ \mathbf{G}_{K+1}^{n-z_m}(s)\mathbf{V}(s,z_m,m) & \text{if } n > m, \end{cases} \tag{46}$$

  where $z_n = \max(n, T_K)$ and $z_m = \max(m, T_K)$.

- For $\mathbf{S}(s,n)$ compute

$$\mathbf{S}(s,n) = \overrightarrow{\mathbf{V}_{K+1}}(s,n) + \sum_{\ell=1}^{K} T_\ell \mathbf{V}(s,n,T_\ell) + \sum_{\ell=1}^{K} \overrightarrow{\mathbf{V}_\ell}(s,n),$$

where, for $\ell \leq K$, $\overrightarrow{\mathbf{V}_\ell}(s,n)$ can be computed according to Theorem 8 and

$$\overrightarrow{\mathbf{V}_{K+1}}(s,n) =$$

$$\begin{cases} \mathbf{V}(s,n,T_K)\left(T_K\overline{\mathbf{R}_{K+1}}(s,\infty)+\overrightarrow{\mathbf{R}_{K+1}}(s,\infty)\right), & \text{if } n \leq T_K, \\ \mathbf{V}(s,n,T_K)\left(T_K\overline{\mathbf{Z}_{K+1}}(s,n-T_K)+\overrightarrow{\mathbf{Z}_{K+1}}(s,n-T_K)\right) & \\ +\mathbf{V}(s,n,n)\Big(T_K\overline{\hat{\mathbf{Z}}_{K+1}}(s,n-T_K)+\overleftarrow{\hat{\mathbf{Z}}_{K+1}}(s,n-T_K) & \\ \qquad +n+n\overline{\mathbf{R}_{K+1}}(s,\infty)+\overrightarrow{\mathbf{R}_{K+1}}(s,\infty)\Big), & \text{if } n > T_K, \end{cases}$$

with $\overline{\mathbf{R}}(s,\infty) = \mathbf{R}(s)\left(\mathbf{I}-\mathbf{R}(s)\right)^{-1}$ and $\overrightarrow{\mathbf{R}}(s,\infty) = \mathbf{R}(s)\left(\mathbf{I}-\mathbf{R}(s)\right)^{-2}$.

8 Numerical examples

To demonstrate the applicability of the proposed approach, we evaluate examples using our Mathematica implementation [1], which also contains the code of the examples and the figures presented in this section. Our implementation was validated in two steps. In case of finite QBD processes we compared the results of our tool with the results of general Markov chain solvers. In case of infinite QBD processes we compared the results with the ones computed for equivalent finite QBD processes with a sufficiently large buffer size.

In the following subsections, the first two examples are from the literature and, as a fairly general example, we also study the behaviour of the MAP/MAP/1 queue with N policy.

8.1 QBD model of a CSMA/CD communication protocol

In [12], a QBD process with irregular level zero is considered which models a simple, infinite population CSMA/CD (Carrier Sense Multiple Access with Collision Detection) communication protocol. Level n of this model represents that the number of busy or blocked users is n. Each level of this structure (except level zero) has three channel states: the safe state (a), the collision state (b), and the contention state (c). That is, this QBD has one phase on level zero and 3 phases on each higher level and the QBD has homogeneous (level independent) transition structure from level one.

The numerical examples in [12] considered two cases for the transitions rates, such that one of the cases is positive recurrent and one is transient. Figure 4 depicts the transient probability of level zero (i.e., empty buffer) for both cases. The depicted time interval suggests that the transient probabilities converge to a positive stationary in the recurrent case and to zero in the transient case. The limiting value of the recurrent case shows a perfect match with the stationary solution of the model. Figure 5 plots two transient probability curves which are presented also in [12]. In each case the system starts from an empty buffer.

Comparing these figures with the original ones, Figure 4, 5, 7 and 8 in [12], one can recognize some differences. These differences might come from the inaccuracy of the numerical inverse Laplace transformation applied in [12], or by other source of error. Indeed, the comparison with the results in [12] is rather difficult due to the improper axes labels used in [12].

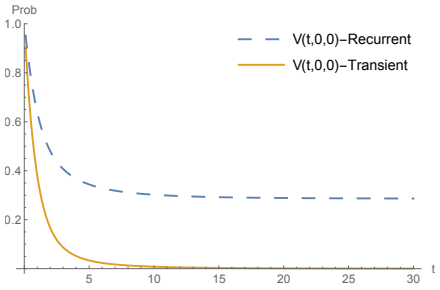

Figure 4: Transient probability of level zero

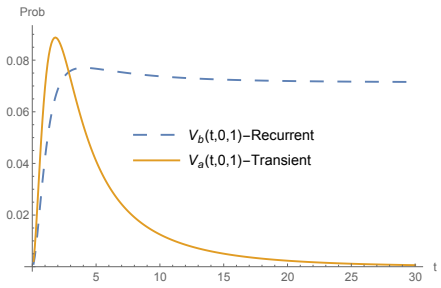

Figure 5: Transient probability of level 1 state (a) and (b)

8.2 QBD model of a TDM voice/data packets integration strategy

In a subsequent work of the same authors, [13], the transient analysis of a QBD process modelling a TDM (time division multiplexing) voice/data packet integration strategy was considered in order to evaluate the time-dependent performance measures such as the expected number of voice packets in service and the voice packet blocking probability. In the considered model, the number of voice packets is restricted to be not greater than 5 and the number of data packets is unbounded. This way the number of voice packets stands for the phase and the number of data packets stands for the level of the QBD process. Various restrictions of data packet arrivals make the zero and the first level different from the higher levels which are level independent. The number of phases is 6 on all levels.

The following performance measures are evaluated in [13]. Starting from level n phase i, the expected number of voice packets in service can be computed as

$$N_V(t,n) \triangleq \underline{e_i}\left[\sum_{j=0}^{\infty}\mathbf{V}(t,n,j)\right]\sum_{k=1}^{6}(k-1)\underline{e_k}^T,$$

where $\underline{e_i}$ is the ith unit row vector, whose only non-zero entry is the ith entry which equals to one. This expression computes the mean number of voice packets utilizing the fact that on each level the number of voice packets in phase k

is $k-1$. The probability of voice packet blocking can be computed as

$$P_B(t,n) \triangleq \underline{e_i}\left[\sum_{j=2}^{\infty}\mathbf{V}(t,n,j)\right](\underline{e_2}^T+\underline{e_4}^T+\underline{e_6}^T) + \underline{e_i}\mathbf{V}(t,n,1)(\underline{e_2}^T+\underline{e_4}^T)+\underline{e_i}\mathbf{V}(t,n,0)\underline{e_2}^T.$$

This expression sums the probabilities of the states where arriving voice packets are blocked. That is, on level 0 phase 2, on level 1 phase 2 and 4, and on the higher levels phase 2, 4 and 6 are voice packet blocking states.

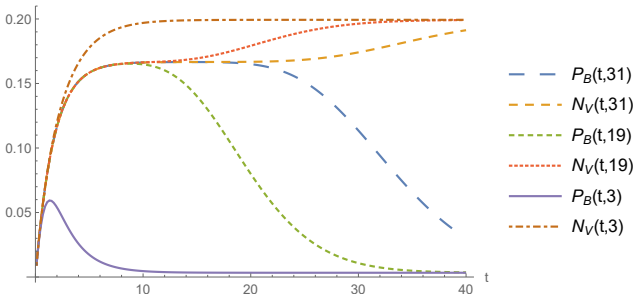

Figure 6: The mean number of voice packets and the data blocking probabilities as a function of time

Figure 6 plots the performance measures as a function of time with various initial levels. The initial phase is $i=1$ in all cases and the initial level is $n=3,19,31$. Also in this example, the figure shows some differences compared to Figure 2 of [13]. All $P_B(t,n)$ measures of the figure converge to a different limit than the one in [13]. Our stationary analysis validated the limit depicted in Figure 6.

8.3 MAP/MAP/1 queue with N policy

To demonstrate applicability of the proposed algorithm in a hysteresis-like queueing model we consider a MAP/MAP/1 queue with N policy, where N policy means that the server is switched off, when the queue gets idle and switched on again, when N customers are accumulated in the queue. The arrival MAP is characterized by matrices $\mathbf{D_0}$ and $\mathbf{D_1}$, while the service MAP by $\mathbf{S_0}$ and $\mathbf{S_1}$. The structure of the Markov chain characterizing the system behaviour is depicted in Figure 7.

To describe the example as a PHQBD we define the regime boundaries as $T_0=0$, $T_1=1$, $T_2=N-1$, $T_3=N$ and it is an infinite PHQBD (discussed in Section 7.2). The associated matrices are

$$\mathbf{L'_1}=\mathbf{L'},\ \mathbf{F_1}=\begin{bmatrix}\mathbf{F} & \mathbf{0}\end{bmatrix},\ \mathbf{B_1}=\begin{bmatrix}\mathbf{0}\\ \mathbf{B}\end{bmatrix},$$

$$\mathbf{L}'_{\mathbf{2}} = \begin{bmatrix} \mathbf{L}' & \mathbf{0} \\ \mathbf{0} & \mathbf{L} \end{bmatrix}, \mathbf{L_2} = \mathbf{L}'_{\mathbf{2}}, \mathbf{F_2} = \begin{bmatrix} \mathbf{F} & \mathbf{0} \\ \mathbf{0} & \mathbf{F} \end{bmatrix}, \mathbf{B_2} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{B} \end{bmatrix},$$
$$\mathbf{L}'_{\mathbf{3}} = \mathbf{L}'_{\mathbf{2}}, \mathbf{F_3} = \begin{bmatrix} \mathbf{F} \\ \mathbf{F} \end{bmatrix}, \mathbf{B_3} = \begin{bmatrix} \mathbf{0} & \mathbf{B} \end{bmatrix},$$
$$\mathbf{L}'_{\mathbf{4}} = \mathbf{L}, \mathbf{L_4} = \mathbf{L}, \mathbf{F_4} = \mathbf{F}, \mathbf{B_4} = \mathbf{B},$$

where $\mathbf{L}' = \mathbf{D_0} \otimes \mathbf{I}$, $\mathbf{L} = \mathbf{D_0} \oplus \mathbf{S_0}$, $\mathbf{F} = \mathbf{D_1} \otimes \mathbf{I}$, $\mathbf{B} = \mathbf{I} \otimes \mathbf{S_1}$. Since $T_0 + 1 = T_1$ and $T_2 + 1 = T_3$, we do not define $\mathbf{L_1}$ and $\mathbf{L_3}$. Furthermore, due to the fact that the sizes of level T_0 and T_1 are different, $\mathbf{F_1}$ and $\mathbf{B_1}$ are non-square and the same property applies for level T_2 and T_3 and matrices $\mathbf{F_3}$ and $\mathbf{B_3}$. In the numerical computations we used $N = 40$ and

$$\mathbf{D_0} = \begin{bmatrix} -8 & 1 & 3 \\ 0 & -6 & 4 \\ 2 & 0 & -3 \end{bmatrix}, \mathbf{D_1} = \begin{bmatrix} 3 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \mathbf{S_0} = \begin{bmatrix} -3 & 1 \\ 6 & -7 \end{bmatrix}, \mathbf{S_1} = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}.$$

The queueing system is overloaded and the PHQBD is transient, because the stationary arrival rate of MAP($\mathbf{D_0}, \mathbf{D_1}$) is 1.875 and the service rate of MAP($\mathbf{S_0}, \mathbf{S_1}$) is 1.7.

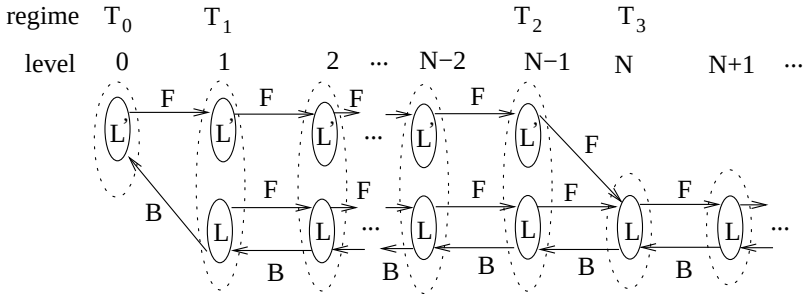

Figure 7: The PHQBD structure of the MAP/MAP/1 queue with N policy, where $\mathbf{L}' = \mathbf{D_0} \otimes \mathbf{I}$, $\mathbf{L} = \mathbf{D_0} \oplus \mathbf{S_0}$, $\mathbf{F} = \mathbf{D_1} \otimes \mathbf{I}$, $\mathbf{B} = \mathbf{I} \otimes \mathbf{S_1}$

Figure 8 plots the transient probabilities of level 40 starting from level 5 and 35 with busy/idle server (lower/upper line of blocks in Figure 7). The initial phase distribution inside the given block is the stationary distribution of the arrival and service MAPs. The figure indicates that the transient probability of staying at level 40 has a dominant peak when starting from a state with idle server. It is because, when the server is idle only arrivals take place and the random fluctuation of the stochastic behavior is not altered by departing customers. The location of the peak is approximately at the sum of $40-5$ $(40-35)$ arrivals when starting from level 5 (35). After the peak the curves start converging toward the stationary distribution of level 40. When the server is busy at the initial state, the arrivals and departures affects the random fluctuation of the stochastic behavior together. It results in much smoother transient probability curves. Also these smooth curves take low values until the probability of reaching level 40 from the initial level is negligible.

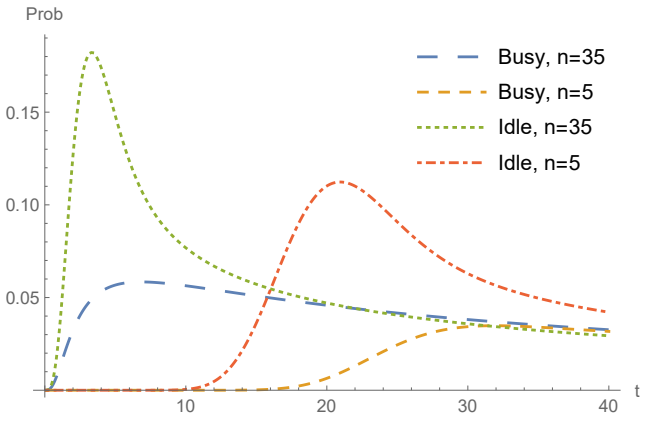

Figure 8: Transient probability of level 40 starting from 4 initial cases

Appendix

A Relation of $\mathbf{Z}^{(b)}(s,m)$ and $\mathbf{H}^{(b)}(s,n)$ type of measures

Equations (3) - (5) present the relation of the transient measures ($\mathbf{U}(s)$, $\mathbf{R}(s)$ and $\mathbf{G}(s)$) with a single (lower) boundary, which allows one to compute any of the three matrix functions from any other one of the three [5]. Similar relations exist between the transient measures with two boundaries. To introduce this relation we define $[\underline{\mathbf{W}}^{(b)}(t,n,m)]_{i,j}$, the probability of visiting level m at time t, starting from level n, given that the level process of the QBD remains in level $(0,b)$ in the $(0,t)$ time interval $(n,m \in (0,b))$, that is

$$[\underline{\mathbf{W}}^{(b)}(t,n,m)]_{i,j} = Pr\Big(\mathcal{X}(t)=m, \mathcal{J}(t)=j, \xi_{0,b} > t \mid \mathcal{X}(0)=n, \mathcal{J}(0)=i\Big)$$

and its Laplace transform is $\mathbf{W}^{(b)}(s,n,m) = \int_0^\infty e^{-st}\underline{\mathbf{W}}^{(b)}(t,n,m)dt$.

In (3) - (5), the process starts at the lower boundary. For the related measures with two boundaries we have

$$\mathbf{Z}^{(b)}(s,m) = \mathbf{F}\mathbf{W}^{(b)}(s,1,m), \quad \mathbf{H}^{(b)}(s,n) = \mathbf{W}^{(b)}(s,n,a+1)\mathbf{B}, \tag{47}$$

and similar relations hold for starting from the upper boundary

$$\hat{\mathbf{Z}}^{(b)}(s,m) = \mathbf{B}\mathbf{W}^{(b)}(s,b-1,m), \quad \hat{\mathbf{H}}^{(b)}(s,n) = \mathbf{W}^{(b)}(s,n,b-1)\mathbf{F}. \tag{48}$$

For $\mathbf{W}^{(b)}(s,n,m)$, we have

$$\begin{aligned}
&\mathbf{W}^{(b)}(s,n,m)\\
&= \begin{cases} \Big(s\mathbf{I}-\mathbf{L}-\mathbf{Z}^{(b-n)}(s,1)\mathbf{F}-\hat{\mathbf{Z}}^{(n)}(s,n-1)\mathbf{B}\Big)^{-1} & \text{if } n=m,\\ \mathbf{W}^{(b)}(s,n,n)\mathbf{Z}^{(b-n)}(s,m-n) & \text{if } n<m,\\ \mathbf{W}^{(b)}(s,n,n)\hat{\mathbf{Z}}^{(n)}(s,n-m) & \text{if } n>m, \end{cases}\\
&= \begin{cases} \Big(s\mathbf{I}-\mathbf{L}-\mathbf{B}\mathbf{H}^{(b-n)}(s,1)-\mathbf{F}\hat{\mathbf{H}}^{(n)}(s,n-1)\Big)^{-1} & \text{if } n=m,\\ \hat{\mathbf{H}}^{(m)}(s,n)\mathbf{W}^{(b)}(s,m,m) & \text{if } n<m,\\ \mathbf{H}^{(b-m)}(s,n-m)\mathbf{W}^{(b)}(s,m,m) & \text{if } n>m, \end{cases}
\end{aligned}$$

where for $n \neq m$, the relations are obtained by conditioning on the last visit to n and the first visit to m. The first equation provides $\mathbf{W}^{(b)}(s,n,m)$ based on Theorem 1, while the second equation based on Theorem 2. Having $\mathbf{W}^{(b)}(s,n,m)$, any of $\mathbf{Z}^{(b)}(s,m)$, $\hat{\mathbf{Z}}^{(b)}(s,m)$, $\mathbf{H}^{(b)}(s,n)$, $\hat{\mathbf{H}}^{(b)}(s,n)$ can be computed using (47) or (48).

B Spectral radius of $\mathbf{G}(s)$ and $\mathbf{R}(s)$

Theorem 9. *If $\Re(s) > 0$ then $sp(\mathbf{G}(s)) < 1$, $sp(\hat{\mathbf{G}}(s)) < 1$, $sp(\mathbf{R}(s)) < 1$ and $sp(\hat{\mathbf{R}}(s)) < 1$, where $sp()$ denotes the spectral radius.*

Proof. The proof is based on the following properties, $\underline{\mathbf{G}}(t)$ and $\underline{\mathbf{R}}(t)$ are non-negative and finite, $\underline{\mathbf{G}}(t)$ is monotone increasing, $\mathbf{G} = \int_{t=0}^{\infty} d\underline{\mathbf{G}}(t)$, and for any matrix $\mathbf{A}$ and matrix norm $\|\cdot\|$ induced by a monotonic vector norm, $sp(\mathbf{A}) \leq \|\mathbf{A}\|$. Using these, we can write

$$\begin{aligned}
sp(\mathbf{G}(s)) \leq \|\mathbf{G}(s)\| &= \left\| \int_{t=0}^{\infty} e^{-st} d\underline{\mathbf{G}}(t) \right\| \leq \left\| \int_{t=0}^{\infty} |e^{-st}| d\underline{\mathbf{G}}(t) \right\| < \left\| \int_{t=0}^{\infty} d\underline{\mathbf{G}}(t) \right\|\\
&= \|\mathbf{G}\| \leq 1.
\end{aligned}$$

For $\mathbf{R}(s)$, we show that the infinite sum $\sum_{m=1}^{\infty} \mathbf{R}^m(s)$ is finite. From (4) and (9), we have

$$\begin{aligned}
&\left| \sum_{m=1}^{\infty} [\mathbf{R}^m(s)]_{i,j} \right|\\
&= \left| \sum_{m=1}^{\infty} \int_0^{\infty} \sum_k e^{-st} F_{ik} Pr\Big(\mathcal{X}(t)=m, \mathcal{J}(t)=j, \gamma_0 > t \;\Big|\; \mathcal{X}(0)=1, \mathcal{J}(0)=k\Big)\, dt \right|\\
&\leq \sum_k F_{ik} \int_0^{\infty} Pr\Big(\mathcal{J}(t)=j, \gamma_0 > t \Big| \mathcal{X}(0)=1, \mathcal{J}(0)=k\Big) \cdot |e^{-st}|\, dt\\
&\leq \sum_k F_{ik} \int_0^{\infty} |e^{-st}|\, dt = \sum_k F_{ik} \frac{1}{\Re(s)} < \infty.
\end{aligned}$$

The same reasoning applies for the level reversed matrices, $\hat{\mathbf{G}}(s)$ and $\hat{\mathbf{R}}(s)$. □

Based on Theorem 9, Theorem 9.3.1 of [6] is applicable and ensures that the solution of (6) and (8) with eigenvalues inside the unit disk is unique.

References

[1] Mathematica implementation of the procedure and the examples. `http://webspn.hit.bme.hu/~telek/techrep/transientQBD.zip`. [Online; accessed 22-Sept-2019].

[2] Ana da Silva Soares and Guy Latouche. Fluid queues with level dependent evolution. *European Journal of Operational Research*, 196:1041–1048, 08 2009.

[3] Sarah Dendievel, Sophie Hautphenne, Guy Latouche, and Peter G. Taylor. The time-dependent expected reward and deviation matrix of a finite QBD process. *Linear Algebra and its Applications*, 570:61 – 92, 2019.

[4] Illés Horváth, Gábor Horváth, Salah Al-Deen Almousa, and Miklós Telek. Numerical inverse laplace transformation by concentrated matrix exponential distributions. *Performance Evaluation*, 2019.

[5] L. Lakatos, L. Szeidl, and M. Telek. *Introduction to Queueing Systems with Telecommunication Applications, 2nd edition*. Springer, 2019.

[6] G. Latouche and V. Ramaswami. *Introduction to Matrix Analytic Methods in Stochastic Modeling*. Society for Industrial and Applied Mathematics, 1999.

[7] San-qi Li and Hong-Dah Sheng. Generalized folding algorithm for transient analysis of finite QBD processes and its queueing applications. In *Computations with Markov Chains*, pages 463–481. Springer, 1995.

[8] M.F. Neuts. *Matrix Geometric Solutions in Stochastic Models*. Johns Hopkins University Press, Baltimore, 1981.

[9] V. Ramaswami. The busy period of queues which have a matrix-geometric steady state probability vector. *Opsearch*, 19, 01 1982.

[10] Benny Van Houdt and Chris Blondia. QBDs with marked time epochs: a framework for transient performance measures. In *Second International Conference on the Quantitative Evaluation of Systems (QEST'05)*, pages 210–219. IEEE, 2005.

[11] Jeroen Van Velthoven, Benny Van Houdt, and Chris Blondia. Simultaneous transient analysis of QBD Markov chains for all initial configurations using a level based recursion. In *Fourth International Conference on the Quantitative Evaluation of Systems (QEST 2007)*, pages 79–90. IEEE, 2007.

[12] Ji Zhang and Edward J Coyle. Transient analysis of quasi-birth-death processes. *Stochastic Models*, 5(3):459–496, 1989.

[13] Ji Zhang and Edward J Coyle. The transient performance analysis of voice/data integrated networks. In *Proceedings. IEEE INFOCOM'90: Ninth Annual Joint Conference of the IEEE Computer and Communications Societies*, pages 963–968. IEEE, 1990.