Commuting matrices in the sojourn time analysis of MAP/MAP/1 queues

G. Horváth
Budapest University of
Technology and Economics
1521 Budapest, Hungary
ghorvath@hit.bme.hu

M. Telek
MTA-BME Information
Systems Research Group
1521 Budapest, Hungary
telek@hit.bme.hu

B. Van Houdt
University of Antwerp, iMinds
B-2020 Antwerpen, Belgium
benny.vanhoudt@ua.ac.be

ABSTRACT

Queues with Markovian arrival and service processes, i.e., MAP/MAP/1 queues, have been useful in the analysis of computer and communication systems and different representations for their sojourn time distribution have been derived. More specifically, the class of MAP/MAP/1 queues lies at the intersection of the class of QBD queues and the class of semi-Markovian queues. While QBD queues have an order N^2 matrix exponential representation for their sojourn time distribution, where N is the size of the background continuous time Markov chain, the sojourn time distribution of the latter class allows for a more compact representation of order N.

In this paper we unify these two results and show that the key step exists in establishing the commutativity of some fundamental matrices involved in the analysis of the MAP/MAP/1 queue. We prove, using two different approaches, that the required matrices do commute and identify several other sets of commuting matrices. Finally, we generalize some of the results to queueing systems with batch arrivals and services.

Keywords: QBD, MAP/MAP/1 queue, sojourn time distribution, commuting matrices.

1. INTRODUCTION

The class of MAP/MAP/1 queues is a versatile and well-studied class of queueing systems used to model computer and communication systems [5, 6]. Its effectiveness lies in the generality of the Markovian arrival process (MAP) which can be used to fit very different arrival patterns with highly correlated inter-arrival times [11, 7, 15]. The MAP process can also be used to model the service process whenever significant correlation exists in the service times of consecutive customers, e.g. [2], and some authors therefore refer to it as the Markovian service process (MSP). The MAP process has also

been extended and analyzed to allow for batch arrivals and multiple customer types [9, 3].

The queue length distribution of the MAP/MAP/1 queue is well-known to be matrix exponential of order N, where N is the product of the number of states of the arrival and service MAP, as its evolution can be captured by means of a Quasi-Birth-Death Markov chain [10]. The sojourn time distribution of the MAP/MAP/1 queue on the other hand can be obtained as a special case of a class of semi-Markovian queues studied by Sengupta [13, 14] and therefore has a matrix exponential form of order N as well. This result was later generalized in [4] for queues with multitype MAP arrivals.

On a different line of research Ozawa studied the sojourn time distribution of a class of so-called Quasi-Birth-Death (QBD) queues [12] and proved that it has a matrix exponential representation of order N^2, where N is the size of the background continuous time Markov chain. As the class of MAP/MAP/1 queues forms a subclass of the set of QBD queues, the result of Ozawa gives rise to an order N^2 representation for the sojourn time distribution of a MAP/MAP/1 queue.

In this paper we unify these two different representations for the sojourn time distribution in a MAP/MAP/1 queue. It turns out that the key feature is the commutativity of some characteristic matrices that appear in the analysis of the queue length and sojourn time distribution of the MAP/MAP/1 queue. Apart from unifying both results and proving the required commutativity property, we also identify some other sets of commuting matrices that have played a fundamental role in the analysis of the MAP/MAP/1 queue. To prove the latter results we introduce two different approaches: an "algebraic" approach that relies on spectral decomposition arguments and a "stochastic" approach that relies on the stochastic interpretation of the matrices involved. These two approaches have cross fertilized one another several times during the evolution of the field.

The paper is structured as follows. Section 2 reintroduces the QBD queue, while Section 3 summarizes the main results on its sojourn time distribution. In Section 4 we show how an order N representation for the sojourn time distribution of the MAP/MAP/1 queue can be obtained from the QBD queue provided that some funda-

mental matrices commute. In Section 5 we present two different approaches to prove that the required commutativity result holds and also identify other sets of commuting matrices. Finally, in Section 6 we argue that the results of Section 5 can be generalized to queues with BMAP arrivals and services as well.

2. THE QUASI-BIRTH-DEATH QUEUE

In a QBD queue the arrivals and the services are modulated by a common continuous time background Markov chain $\mathcal{Z}(t)$. A set of transitions of the background process are accompanied by an arrival (the associated matrix is denoted by $\boldsymbol{F}$), other transitions of the background process are accompanied by a service completion, assuming that there is at least a customer in the system (given by matrix $\boldsymbol{B}$). There may be transitions by which neither an arrival, nor a service completion occurs (given by matrices $\boldsymbol{L}$ or $\boldsymbol{L}'$ depending on whether the system is busy or empty, respectively). When there is at least one customer in the system the generator of the background process is denoted by $\boldsymbol{Q} = \{q_{ij}, \ i,j = 1, \dots, N\}$. When there is no customer in the queue the generator of the background process might be different and is denoted by $\boldsymbol{Q}' = \{q'_{ij}, \ i,j = 1, \dots, N\}$. Note that $\boldsymbol{Q} = \boldsymbol{B} + \boldsymbol{L} + \boldsymbol{F}$ and $\boldsymbol{Q}' = \boldsymbol{L}' + \boldsymbol{F}$. The stochastic process that keeps track of the number of customers in the system is denoted by $\mathcal{X}(t)$.

With a lexicographical numbering of the states the two-dimensional process $\{\mathcal{X}(t), \mathcal{Z}(t), t > 0\}$ is a QBD Markov chain [1], with its generator given by

$$\boldsymbol{\Pi} = \begin{bmatrix} \boldsymbol{L}' & \boldsymbol{F} & & \\ \boldsymbol{B} & \boldsymbol{L} & \boldsymbol{F} & \\ & \boldsymbol{B} & \boldsymbol{L} & \boldsymbol{F} \\ & & \ddots & \ddots & \ddots \end{bmatrix}. \tag{1}$$

The sojourn time in a QBD queue, $\mathcal{V}$, is defined as the time between an arrival event and the corresponding service instant in steady state assuming a first-come first-served (FCFS) service discipline.

Provided that the QBD Markov chain with transition matrix $\boldsymbol{\Pi}$ is irreducible and positive recurrent, denote its stationary distribution by $\pi = (\pi_0, \pi_1, \dots)$. The j-th entry of the vector π_k corresponds to the steady state probability that there are k customers in the queue while the background process $\mathcal{Z}(t)$ is in state j. As the steady state distribution of a QBD Markov chain is known to have a matrix geometric form [1], π_k can be written as

$$\pi_k = \pi_0 \boldsymbol{R}^k, \quad k > 0, \tag{2}$$

where $\boldsymbol{R}$ is the minimal non-negative solution of the quadratic matrix equation

$$\boldsymbol{0} = \boldsymbol{F} + \boldsymbol{R}\boldsymbol{L} + \boldsymbol{R}^2\boldsymbol{B}, \tag{3}$$

and vector π_0 is the unique solution of the following set of linear equations:

$$0 = \pi_0 \left(\boldsymbol{L}' + \boldsymbol{R}\boldsymbol{B} \right),$$
$$1 = \pi_0 \left(\boldsymbol{I} - \boldsymbol{R} \right)^{-1} \mathbb{1}.$$

For later use we also introduce the matrix $\boldsymbol{U}$ and $\boldsymbol{G}$ as the smallest non-negative solution of

$$\boldsymbol{U} = \boldsymbol{L} + \boldsymbol{F}(-\boldsymbol{U})^{-1}\boldsymbol{B}, \tag{4}$$

$$\boldsymbol{0} = \boldsymbol{B} + \boldsymbol{L}\boldsymbol{G} + \boldsymbol{F}\boldsymbol{G}^2, \tag{5}$$

respectively. The matrices $\boldsymbol{R}$, $\boldsymbol{U}$ and $\boldsymbol{G}$ are all defined by $\boldsymbol{B}, \boldsymbol{L}, \boldsymbol{F}$ and they are related such that $\boldsymbol{R} = \boldsymbol{F}(-\boldsymbol{U})^{-1}$ and $\boldsymbol{G} = (-\boldsymbol{U})^{-1}\boldsymbol{B}$ [1]. The mean arrival rate λ of a QBD queue is given by

$$\lambda = \sum_{i=0}^{\infty} \pi_i \boldsymbol{F} \mathbb{1}.$$

If the arrival and service times are controlled by independent Markov chains $\mathcal{Z}^{(in)}(t)$ and $\mathcal{Z}^{(out)}(t)$, the QBD queue simplifies to a MAP/MAP/1 queue. This is an important special case, as the independence of the arrival and service processes often holds in practice, and the sojourn time distribution can be obtained in a more efficient way (see Section 4.2). By denoting the matrices of the MAP that generates the arrivals by $\boldsymbol{D_0}$ and $\boldsymbol{D_1}$ $(\boldsymbol{D_0} + \boldsymbol{D_1} = \boldsymbol{D}, \boldsymbol{D} = \{d_{ij}, \ i,j = 1, \dots, N^{(in)}\})$ and the matrices of the MAP generating the service events by $\boldsymbol{S_0}$ and $\boldsymbol{S_1}$ $(\boldsymbol{S_0} + \boldsymbol{S_1} = \boldsymbol{S}, \boldsymbol{S} = \{s_{ij}, \ i,j = 1, \dots, N^{(out)}\})$ the blocks of the QBD Markov chain can be expressed as

$$\begin{aligned} \boldsymbol{F} &= \boldsymbol{D_1} \otimes \boldsymbol{I}, \\ \boldsymbol{L} &= \boldsymbol{D_0} \oplus \boldsymbol{S_0}, \\ \boldsymbol{B} &= \boldsymbol{I} \otimes \boldsymbol{S_1}, \\ \boldsymbol{L}' &= \boldsymbol{D_0} \otimes \boldsymbol{I}. \end{aligned} \tag{6}$$

3. SOJOURN TIME IN THE QBD QUEUE

To determine the distribution of the sojourn time it suffices to know the distribution of the queue length at arrival instants and the distribution of the time taken by the QBD queue to generate k service events, for $k \geq 1$.

Let entry j of the vector $\hat{\pi}_k$ denote the probability that the QBD queue is at level k just after the arrival epoch, while the background process is in state j. Further, let entry (i, j) of the matrix $\boldsymbol{N}(k, t)$ denote the probability that exactly k service events occur in a non-idle interval of length t, while the phase of the underlying process is i and j at the start and end of the interval, respectively, that is

$$[\boldsymbol{N}(k, t)]_{i,j} =$$
$$P(\mathcal{X}_s(t) = 1, \mathcal{Z}(t) = j | \mathcal{X}_s(0) = k + 1, \mathcal{Z}(0) = i),$$

where $\mathcal{X}_s(t)$ corresponds to the level of the two-dimensional Markov chain $\{\mathcal{X}_s(t), \mathcal{Z}(t), t > 0\}$ with its generator given by

$$\boldsymbol{\Pi} = \begin{bmatrix} \boldsymbol{L}' + \boldsymbol{F} & & & \\ \boldsymbol{B} & \boldsymbol{L} + \boldsymbol{F} & & \\ & \boldsymbol{B} & \boldsymbol{L} + \boldsymbol{F} & \\ & & \ddots & \ddots \end{bmatrix}. \tag{7}$$

$\boldsymbol{N}(k, t)$ is determined by the following set of differential

equations [1]:

$$\frac{\partial}{\partial t}\boldsymbol{N}(0,t) = \boldsymbol{N}(0,t)(\boldsymbol{L}+\boldsymbol{F}), \qquad (8)$$

$$\frac{\partial}{\partial t}\boldsymbol{N}(k,t) = \boldsymbol{N}(k,t)(\boldsymbol{L}+\boldsymbol{F}) + \boldsymbol{N}(k-1,t)\boldsymbol{B}, \qquad (9)$$

for $k=1,\ldots,\infty$ with boundary conditions $\boldsymbol{N}(0,0)=\boldsymbol{I}$ and $\boldsymbol{N}(k,0)=\boldsymbol{0}$ for $k>0$. The generating function of the departure events is defined by $\boldsymbol{N}^*(z,t)=\sum_{k=0}^{\infty} z^k \boldsymbol{N}(k,t)$. Multiplying (8) and (9) by z^k and summing up for $k=0,1,\ldots$ gives

$$\frac{\partial}{\partial t}\boldsymbol{N}^*(z,t) = \boldsymbol{N}(z,t)(\boldsymbol{L}+\boldsymbol{F}+z\boldsymbol{B}), \qquad (10)$$

with initial condition $\boldsymbol{N}^*(z,0)=\boldsymbol{I}$. Its solution is given by

$$\boldsymbol{N}^*(z,t) = e^{(\boldsymbol{L}+\boldsymbol{F}+z\boldsymbol{B})t}. \qquad (11)$$

Ozawa [12] established the following two theorems, where the second theorem shows that the sojourn time distribution has a matrix exponential form of order N^2:

THEOREM 1. *(Theorem 1 in [12]) The vectors $\hat{\pi}_k$ are given by*

$$\hat{\pi}_1 = \frac{1}{\lambda}\pi_0 \boldsymbol{F},$$
$$\hat{\pi}_k = \hat{\pi}_1 \hat{\boldsymbol{R}}^{k-1}, \quad k=2,\ldots,\infty, \qquad (12)$$

with $\hat{\boldsymbol{R}}$ given by

$$\hat{\boldsymbol{R}} = (-\boldsymbol{U})^{-1}\boldsymbol{F}. \qquad (13)$$

THEOREM 2. *(Theorem 2 in [12]) The distribution of the sojourn time is given by*

$$P(\mathcal{V}<t) = 1 - (\mathbb{1}^T \otimes \hat{\eta})e^{((\boldsymbol{L}+\boldsymbol{F})^T \otimes \boldsymbol{I})+(\boldsymbol{B}^T \otimes \hat{\boldsymbol{R}}))t}vec\langle\boldsymbol{I}\rangle, \qquad (14)$$

where $\hat{\eta}$ is the stationary phase distribution at arrivals

$$\hat{\eta} = \hat{\pi}_1\left(\boldsymbol{I}-\hat{\boldsymbol{R}}\right)^{-1}, \qquad (15)$$

and $vec\langle\rangle$ denotes the column-stacking operator.

Remark 1: Theorem 1 was proven using probabilistic arguments in [12], but can also be proven easily in an algebraic manner as

$$\hat{\pi}_k = \frac{\pi_{k-1}\boldsymbol{F}}{\sum_{i=0}^{\infty}\pi_i \boldsymbol{F}\mathbb{1}} = \frac{1}{\lambda}\pi_{k-1}\boldsymbol{F} = \frac{1}{\lambda}\pi_0 \boldsymbol{R}^{k-1}\boldsymbol{F}$$
$$= \frac{1}{\lambda}\pi_0\left(\boldsymbol{F}(-\boldsymbol{U})^{-1}\right)^{k-1}\boldsymbol{F} = \frac{1}{\lambda}\pi_0\boldsymbol{F}\left((-\boldsymbol{U})^{-1}\boldsymbol{F}\right)^{k-1}$$
$$= \frac{1}{\lambda}\pi_0\boldsymbol{F}\hat{\boldsymbol{R}}^{k-1}. \qquad (16)$$

Remark 2: It was also noted in [12, Remark 1] that the sojourn time distribution can also be expressed as $P(\mathcal{V}>t) = \hat{\eta}\boldsymbol{W}(t)\mathbb{1}$, where

$$\boldsymbol{W}(t) = \sum_{k=0}^{\infty}\hat{\boldsymbol{R}}^k \boldsymbol{N}(k,t), \qquad (17)$$

and that $\boldsymbol{W}(t)$ is the solution of the differential equation

$$\frac{d}{dt}\boldsymbol{W}(t) = \boldsymbol{W}(t)(\boldsymbol{L}+\boldsymbol{F}) + \hat{\boldsymbol{R}}\boldsymbol{W}(t)\boldsymbol{B}. \qquad (18)$$

with $\boldsymbol{W}(0)=\boldsymbol{I}$. Note if $\hat{\boldsymbol{R}}$ and $\boldsymbol{W}(t)$ were to commute, this differential equation immediately leads to a matrix exponential distribution for the sojourn time of order N. Ozawa [12] notes that $\hat{\boldsymbol{R}}$ and $\boldsymbol{W}(t)$ commute for the M/PH/1 queue, but not in general for the QBD queue. In fact, even for the MAP/M/1 queue $\hat{\boldsymbol{R}}$ and $\boldsymbol{W}(t)$ do not commute in general, meaning (18) does not give immediate rise to an order N representation. More specifically, for the MAP/M/1 queue we can easily see that $\boldsymbol{W}(t)$ can be expressed as

$$\boldsymbol{W}(t) = \sum_{k=0}^{\infty}\hat{\boldsymbol{R}}^k e^{\boldsymbol{D}t}\frac{(\mu t)^k}{k!}e^{-\mu t} = e^{\hat{\boldsymbol{R}}\mu t}e^{(\boldsymbol{D}-\mu\boldsymbol{I})t}. \qquad (19)$$

Thus $\hat{\boldsymbol{R}}$ and $\boldsymbol{W}(t)$ only commute if $\hat{\boldsymbol{R}}$ and $e^{(\boldsymbol{D}-\mu\boldsymbol{I})t}$ commute, which only holds in some special cases.

4. SOJOURN TIME IN A MAP/MAP/1 QUEUE

4.1 The QBD queue approach

In this section, we will make a slight modification to $\boldsymbol{W}(t)$ for the MAP/MAP/1 queue such that we obtain a differential equation where the modified $\boldsymbol{W}(t)$, denoted as $\tilde{\boldsymbol{W}}(t)$, and $\hat{\boldsymbol{R}}$ always commute.

More specifically we introduce the matrix $\tilde{\boldsymbol{W}}(t)$ similar to (17) as

$$\tilde{\boldsymbol{W}}(t) = \sum_{k=0}^{\infty}\hat{\boldsymbol{R}}^k \tilde{\boldsymbol{N}}(k,t), \qquad (20)$$

where $\tilde{\boldsymbol{N}}(k,t)$ is defined as the solution to the differential equation

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{N}}(0,t) = \tilde{\boldsymbol{N}}(0,t)(\boldsymbol{I}\otimes \boldsymbol{S_0}), \qquad (21)$$

$$\frac{\partial}{\partial t}\tilde{\boldsymbol{N}}(k,t) = \tilde{\boldsymbol{N}}(k,t)(\boldsymbol{I}\otimes \boldsymbol{S_0}) + \tilde{\boldsymbol{N}}(k-1,t)(\boldsymbol{I}\otimes \boldsymbol{S_1}), \qquad (22)$$

for $k=1,\ldots,\infty$ with $\tilde{\boldsymbol{N}}(0,0)=\boldsymbol{I}$ and $\tilde{\boldsymbol{N}}(k,0)=\boldsymbol{0}$ for $k>0$. Observe that the definition of $\tilde{\boldsymbol{N}}(k,t)$ differs from $\boldsymbol{N}(k,t)$ in that $\tilde{\boldsymbol{N}}(k,t)$ does not follow the evolution of the arrival process, more precisely the phase of the arrival process remains fixed. This slight difference will turn out to be essential in the subsequent discussion.

We can now establish the following theorem, the proof of which is similar in nature to the one of Theorem 2 in [12] and is included for completeness:

THEOREM 3. *The sojourn time distribution in a MAP/MAP/1 queue can be expressed as $P(\mathcal{V}>t) = \hat{\eta}\tilde{\boldsymbol{W}}(t)\mathbb{1}$, where $\tilde{\boldsymbol{W}}(t)$ is the solution to the differential equation*

$$\frac{d}{dt}\tilde{\boldsymbol{W}}(t) = \tilde{\boldsymbol{W}}(t)(\boldsymbol{I}\otimes \boldsymbol{S_0}) + \hat{\boldsymbol{R}}\tilde{\boldsymbol{W}}(t)(\boldsymbol{I}\otimes \boldsymbol{S_1}). \qquad (23)$$

with $\tilde{\boldsymbol{W}}(0) = \boldsymbol{I}$.

PROOF. The probability that the sojourn time of an arriving customer is greater than t equals the probability that the number of service events generated up to time t is less than the number of customers the arriving customer found in the system (including itself). Hence, we have

$$
\begin{aligned}
P(\mathcal{V} > t) &= \sum_{n=1}^{\infty} \hat{\pi}_n \sum_{k=0}^{n-1} \tilde{\boldsymbol{N}}(k,t)\mathbb{1} \\
&= \sum_{n=1}^{\infty} \hat{\pi}_1 \hat{\boldsymbol{R}}^{n-1} \sum_{k=0}^{n-1} \tilde{\boldsymbol{N}}(k,t)\mathbb{1} \qquad (24) \\
&= \sum_{k=0}^{\infty} \hat{\eta} \hat{\boldsymbol{R}}^k \tilde{\boldsymbol{N}}(k,t)\mathbb{1} = \hat{\eta}\tilde{\boldsymbol{W}}(t)\mathbb{1},
\end{aligned}
$$

where $\hat{\eta} = \sum_{k=1}^{\infty} \hat{\pi}_1 \hat{\boldsymbol{R}}^{k-1}$ has a closed form given by (15). To obtain the differential equation in (23) for $\tilde{\boldsymbol{W}}(t)$, it suffices to sum (21) and (22) after left-multiplying them by $\hat{\boldsymbol{R}}^k$. $\square$

Remark 3: Making use of the $\mathrm{vec}\langle\rangle$ operator and utilizing its properties, Theorem 3 yields

$$
\begin{aligned}
\frac{d}{dt}\mathrm{vec}\langle\tilde{\boldsymbol{W}}(t)\rangle &= ((\boldsymbol{I} \otimes \boldsymbol{S_0})^T \otimes \boldsymbol{I})\mathrm{vec}\langle\tilde{\boldsymbol{W}}(t)\rangle \\
&\quad + ((\boldsymbol{I} \otimes \boldsymbol{S_1})^T \otimes \hat{\boldsymbol{R}})\mathrm{vec}\langle\tilde{\boldsymbol{W}}(t)\rangle,
\end{aligned}
$$

for which the closed form solution is

$$
\mathrm{vec}\langle\tilde{\boldsymbol{W}}(t)\rangle = e^{((\boldsymbol{I} \otimes \boldsymbol{S_0})^T \otimes \boldsymbol{I}) + ((\boldsymbol{I} \otimes \boldsymbol{S_1})^T \otimes \hat{\boldsymbol{R}}))t}\mathrm{vec}\langle\boldsymbol{I}\rangle, \quad (25)
$$

by noting that $\tilde{\boldsymbol{W}}(0) = \boldsymbol{I}$. Thus the distribution of the sojourn time in a MAP/MAP/1 queue can also be expressed as

$$
\begin{aligned}
P(\mathcal{V} < t) &= 1 - \hat{\eta}\tilde{\boldsymbol{W}}(t)\mathbb{1} \\
&= 1 - (\mathbb{1}^T \otimes \hat{\eta})e^{((\boldsymbol{I} \otimes \boldsymbol{S_0})^T \otimes \boldsymbol{I}) + ((\boldsymbol{I} \otimes \boldsymbol{S_1})^T \otimes \hat{\boldsymbol{R}}))t}\mathrm{vec}\langle\boldsymbol{I}\rangle.
\end{aligned}
$$

This distribution is a matrix exponential distribution of order N^2 and is therefore of little interest. Theorem 3 is however interesting as we will now show that $\tilde{\boldsymbol{W}}(t)$ and $\hat{\boldsymbol{R}}$ do commute in general. To prove this formally we introduce the following assumption:

ASSUMPTION 1. *The matrices* $\hat{\boldsymbol{R}}$ *and* $(\boldsymbol{I} \otimes \boldsymbol{S_0}) + \hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1})$ *can be diagonalized.*

A matrix $\boldsymbol{A}$ of size $N \times N$ can be diagonalized if there exists a non-singular matrix $\boldsymbol{P}$ such that $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ is a diagonal matrix, or equivalently there exists a set of N (not necessarily distinct) eigenvalues for $\boldsymbol{A}$ with N linearly independent corresponding eigenvectors. Two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ are said to be simultaneously diagonalizable if there exists a non-singular matrix $\boldsymbol{P}$ such that both $\boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ and $\boldsymbol{P}^{-1}\boldsymbol{B}\boldsymbol{P}$ are diagonal matrices, or equivalently both $\boldsymbol{A}$ and $\boldsymbol{B}$ are diagonalizable and they share the same set of eigenvectors.

Note, we do not assume that $\hat{\boldsymbol{R}}$ and $(\boldsymbol{I} \otimes \boldsymbol{S_0}) + \hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1})$ can be simultaneously diagonalized. Further on we will

prove that $\hat{\boldsymbol{R}}$ and $(\boldsymbol{I} \otimes \boldsymbol{S_0}) + \hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1})$ commute, which will imply that they can in fact be simultaneously diagonalized, because a set of diagonalizable matrices commutes if and only if the set is simultaneously diagonalizable. We feel that Assumption 1 is probably not required for $\tilde{\boldsymbol{W}}(t)$ and $\hat{\boldsymbol{R}}$ to commute, but did not manage to come up with a formal proof without this assumption.

THEOREM 4. *Under Assumption 1 the matrices* $\tilde{\boldsymbol{W}}(t)$ *and* $\hat{\boldsymbol{R}}$ *commute.*

PROOF. Denote $\lambda_1, \ldots, \lambda_N$ as the N eigenvalues of $\hat{\boldsymbol{R}}$ with N linearly independent corresponding (left) eigenvectors $\phi_1, \ldots, \phi_N$. We prove that $\tilde{\boldsymbol{W}}(t)$ and $\hat{\boldsymbol{R}}$ can be simultaneously diagonalized, which suffices to prove that they commute. As such it is sufficient to show that ϕ_i is an eigenvector of $\tilde{\boldsymbol{W}}(t)$ as well, for $i = 1, \ldots, N$. As ϕ_i is a left-eigenvector of $\hat{\boldsymbol{R}}$ we have

$$
\begin{aligned}
\phi_i\tilde{\boldsymbol{W}}(t) &= \phi_i \sum_{k=0}^{\infty} \hat{\boldsymbol{R}}^k \tilde{\boldsymbol{N}}(k,t) = \phi_i \sum_{k=0}^{\infty} \lambda_i^k \tilde{\boldsymbol{N}}(k,t) \\
&= \phi_i\tilde{\boldsymbol{N}}^*(z,t)|_{z=\lambda_i} = \phi_i e^{((\boldsymbol{I} \otimes \boldsymbol{S_0}) + \lambda_i(\boldsymbol{I} \otimes \boldsymbol{S_1}))t},
\end{aligned}
\quad (26)
$$

where $\tilde{\boldsymbol{N}}^*(z,t) = \sum_{k=0}^{\infty} z^k \tilde{\boldsymbol{N}}(k,t)$ is the generating function based on (21) and (22) and similar to (11) we have

$$
\tilde{\boldsymbol{N}}^*(z,t) = e^{((\boldsymbol{I} \otimes \boldsymbol{S_0}) + z(\boldsymbol{I} \otimes \boldsymbol{S_1}))t}. \quad (27)
$$

At this point we need to utilize that some essential matrices of MAP/MAP/1 queues commute, which is discussed in detail below. According to Theorem 12 in Section 5.2 the matrices $\hat{\boldsymbol{R}}$ and $(\boldsymbol{I} \otimes \boldsymbol{S_0}) + \hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1})$ commute, thus the eigenvectors $\phi_1, \ldots, \phi_N$ of $\hat{\boldsymbol{R}}$ are also eigenvectors of $(\boldsymbol{I} \otimes \boldsymbol{S_0}) + \hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1})$, as a set of diagonalizable matrices commutes if and only if the set is simultaneously diagonalizable. Let γ_i be the eigenvalue of $(\boldsymbol{I} \otimes \boldsymbol{S_0}) + \hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1})$ associated with ϕ_i, then

$$
\begin{aligned}
\gamma_i\phi_i &= \phi_i \left((\boldsymbol{I} \otimes \boldsymbol{S_0}) + \hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1})\right) \\
&= \phi_i(\boldsymbol{I} \otimes \boldsymbol{S_0}) + \phi_i\hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1}) \\
&= \phi_i(\boldsymbol{I} \otimes \boldsymbol{S_0}) + \lambda_i\phi_i(\boldsymbol{I} \otimes \boldsymbol{S_1}) \\
&= \phi_i \left((\boldsymbol{I} \otimes \boldsymbol{S_0}) + \lambda_i(\boldsymbol{I} \otimes \boldsymbol{S_1})\right).
\end{aligned}
$$

Consequently, ϕ_i is also an eigenvector of $(\boldsymbol{I} \otimes \boldsymbol{S_0}) + \lambda_i(\boldsymbol{I} \otimes \boldsymbol{S_1})$ and $e^{((\boldsymbol{I} \otimes \boldsymbol{S_0}) + \lambda_i(\boldsymbol{I} \otimes \boldsymbol{S_1}))t}$ as required. $\square$

Remark 5: For the MAP/M/1 queue we can easily see that $\tilde{\boldsymbol{W}}(t)$ is found as

$$
\tilde{\boldsymbol{W}}(t) = \sum_{k=0}^{\infty} \hat{\boldsymbol{R}}^k \frac{(\mu t)^k}{k!}e^{-\mu t} = e^{-\mu t}\,e^{\hat{\boldsymbol{R}}\mu t}, \quad (28)
$$

meaning $\hat{\boldsymbol{R}}$ and $\tilde{\boldsymbol{W}}(t)$ commute even without Assumption 1. For the M/MAP/1 queue on the other hand, $\tilde{\boldsymbol{N}}(k,t)$ and $\boldsymbol{N}(k,t)$ are the same, meaning $\tilde{\boldsymbol{W}}(t) = \boldsymbol{W}(t)$ and $\boldsymbol{W}(t)$ therefore commutes with $\hat{\boldsymbol{R}}$, which generalizes the observation of Ozawa that they commute for the M/PH/1 queue.

COROLLARY 1. *Under Assumption 1 the sojourn time distribution of a MAP/MAP/1 queue has an order N matrix exponential representation given by*

$$P(\mathcal{V} < t) = 1 - \hat{\eta} e^{((\boldsymbol{I} \otimes \boldsymbol{S_0}) + \hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1}))t} \mathbb{1}. \qquad (29)$$

PROOF. Utilizing Theorem 4 makes the solution of (23) more efficient, as we now have

$$\frac{d}{dt} \tilde{\boldsymbol{W}}(t) = \tilde{\boldsymbol{W}}(t)[(\boldsymbol{I} \otimes \boldsymbol{S_0}) + \hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1})], \qquad (30)$$

from which (29) follows. $\square$

4.2 The age process approach

The MAP/MAP/1 queue is also a special case of a class of semi-Markovian queues studied by Sengupta [14], where the sojourn time of a semi-Markovian queue was shown to be representable as an order N matrix exponential distribution. We start by presenting this result customized to the MAP/MAP/1 queue.

The analysis of semi-Markovian queues is based on the analysis of the age process. Although the age process can be defined at all time epochs t, it suffices to define it when the server is busy (by censoring out the idle periods). Define the age process $\{\mathcal{A}(t), \mathcal{Z}(t), t \geq 0\}$ as follows. $\mathcal{A}(t) \geq 0$ represents the age of the customer in service at time t, that is, $t - \mathcal{A}(t)$ represents the time of arrival of the customer in service. $\mathcal{Z}(t)$ keeps track of the phase of the service process at time t and the phase of the arrival process at time $t - \mathcal{A}(t)$.

According to [14] the stationary distribution of the age process is matrix exponential as

$$\alpha(x) = \alpha(0) e^{\boldsymbol{T}x}, \qquad (31)$$

where matrix $\boldsymbol{T}$ is closely related to matrix $\hat{\boldsymbol{R}}$ as ([14], Equation (15))

$$\boldsymbol{T} = (\boldsymbol{I} \otimes \boldsymbol{S_0}) + \hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1}), \qquad (32)$$

and the vector $\alpha(0)$ is given by

$$\alpha(0) = (\theta \otimes \beta)(-\boldsymbol{T}), \qquad (33)$$

where the vectors β and θ are the solutions of $\beta(\boldsymbol{S_0} + \boldsymbol{S_1}) = 0, \beta \mathbb{1} = 1$ and $\theta(-\boldsymbol{D_0})^{-1} \boldsymbol{D_1} = \theta, \theta \mathbb{1} = 1$, respectively.

As the sojourn time of a customer is its age at its service completion, we have that

$$P(\mathcal{V} < t) = \frac{\int_0^t \alpha(x)(\boldsymbol{I} \otimes \boldsymbol{S_1}) \mathbb{1} \, dx}{\int_0^\infty \alpha(x)(\boldsymbol{I} \otimes \boldsymbol{S_1}) \mathbb{1} \, dx}$$
$$= 1 - \frac{1}{c}(\theta \otimes \beta) e^{\boldsymbol{T}t}(\boldsymbol{I} \otimes \boldsymbol{S_1}) \mathbb{1}, \qquad (34)$$

where c is a normalization constant $c = (\theta \otimes \beta)(\boldsymbol{I} \otimes \boldsymbol{S_1}) \mathbb{1}$, that equals the mean service rate $\mu = \beta \boldsymbol{S_1} \mathbb{1}$.

The two expressions for the distribution of the sojourn time in a MAP/MAP/1 queue given by (34) and (29) can be proven to be equal in a direct manner.

To this end, we first express the probability vector corresponding to an arrival to the empty queue by two different ways.

- Based on the queue process this probability vector equals to $\hat{\pi}_1$.

- We can express this probability vector based on the age process as well, it is the probability that the next arrival occurs later than the sojourn time of a customer. Hence we get

$$\frac{\int_0^\infty \alpha(x)(\boldsymbol{I} \otimes \boldsymbol{S_1}) e^{(\boldsymbol{D_0} \otimes \boldsymbol{I})x}((-\boldsymbol{D_0})^{-1} \boldsymbol{D_1} \otimes \boldsymbol{I}) dx}{\int_0^\infty \alpha(x)(\boldsymbol{I} \otimes \boldsymbol{S_1}) \mathbb{1} \, dx}, \qquad (35)$$

where the denominator equals to μ (see above) and the numerator is $\alpha(0)$ due to Lemma 2.4 in [13].

Thus, we can conclude that $\hat{\pi}_1 = \alpha(0)/\mu = (\theta \otimes \beta)(-\boldsymbol{T})/\mu$ holds. Applying this equality in (34) yields

$$P(\mathcal{V} > t) = \hat{\eta} e^{\boldsymbol{T}t} \mathbb{1} = \hat{\pi}_1(\boldsymbol{I} - \hat{\boldsymbol{R}})^{-1} e^{\boldsymbol{T}t} \mathbb{1}$$
$$= \frac{1}{\mu}(\theta \otimes \beta)(-\boldsymbol{T})(\boldsymbol{I} - \hat{\boldsymbol{R}})^{-1} e^{\boldsymbol{T}t} \mathbb{1}.$$

Exploiting the fact that the matrices $\hat{\boldsymbol{R}}, \boldsymbol{T}$ and $e^{\boldsymbol{T}t}$ commute (due to Theorem 12 and (32)) yields

$$P(\mathcal{V} > t) = \frac{1}{\mu}(\theta \otimes \beta)(-\boldsymbol{T})(\boldsymbol{I} - \hat{\boldsymbol{R}})^{-1} e^{\boldsymbol{T}t} \mathbb{1}$$
$$= \frac{1}{\mu}(\theta \otimes \beta) e^{\boldsymbol{T}t}(\boldsymbol{I} - \hat{\boldsymbol{R}})^{-1}(-\boldsymbol{T}) \mathbb{1}$$
$$= \frac{1}{c}(\theta \otimes \beta) e^{\boldsymbol{T}t}(\boldsymbol{I} \otimes \boldsymbol{S_1}) \mathbb{1},$$

where in the last step we utilized that $(\boldsymbol{I} - \hat{\boldsymbol{R}})^{-1}(-\boldsymbol{T}) \mathbb{1} = (\boldsymbol{I} \otimes \boldsymbol{S_1}) \mathbb{1}$ which can be proven as follows: (32) clearly implies that

$$-\boldsymbol{T} + (\boldsymbol{I} \otimes \boldsymbol{S_1}) = -(\boldsymbol{I} \otimes \boldsymbol{S_0}) + (\boldsymbol{I} - \hat{\boldsymbol{R}})(\boldsymbol{I} \otimes \boldsymbol{S_1}),$$

which yields

$$(\boldsymbol{I} - \hat{\boldsymbol{R}})^{-1}(-\boldsymbol{T}) = (\boldsymbol{I} \otimes \boldsymbol{S_1}) - (\boldsymbol{I} - \hat{\boldsymbol{R}})^{-1}(\boldsymbol{I} \otimes (\boldsymbol{S_0} + \boldsymbol{S_1})),$$

and the equality follows by post-multiplying it with $\mathbb{1}$ as $(\boldsymbol{S_0} + \boldsymbol{S_1}) \mathbb{1} = \boldsymbol{0}$.

5. COMMUTING MATRICES IN MAP/MAP/1 QUEUES

In this section we will identify four sets of commuting matrices related to the MAP/MAP/1 queue. We will provide two different approaches to prove these relations. The first will be based on the spectral decomposition of the matrices involved and as such will require some mild assumptions on the existence of the eigenvectors involved. The second approach will not require any assumptions and is based on a relationship derived from the age process, as such we refer to it as the queueing based approach.

5.1 Spectral decomposition approach

We start by establishing some relations for the matrices $\boldsymbol{R}$ and $\boldsymbol{G}$ and will rely on the following assumption:

ASSUMPTION 2. *The matrix $\boldsymbol{G}$ can be diagonalized and inverted. Denote its eigenvalues as $\bar{\lambda}_1, \ldots, \bar{\lambda}_N$ and its corresponding right eigenvectors as $u_1, \ldots, u_N$. We further assume that the matrices $\boldsymbol{D_0} + \bar{\lambda}_i \boldsymbol{D_1}$ and $\boldsymbol{S_1} + \bar{\lambda}_i \boldsymbol{S_0}$ can be diagonalized for $i = 1, \ldots, N$.*

By the definition of $\bar{\lambda}_i$ and (5) we have

$$\det \left[(\boldsymbol{I} \otimes \boldsymbol{S_1}) + (\boldsymbol{D_0} \oplus \boldsymbol{S_0})\bar{\lambda}_i + (\boldsymbol{D_1} \otimes \boldsymbol{I})\bar{\lambda}_i^2 \right] = 0 \quad (36)$$

on the unit disk (that is $|\bar{\lambda}_i| \leq 1$) and u_i is the solution of

$$\left[(\boldsymbol{I} \otimes \boldsymbol{S_1}) + (\boldsymbol{D_0} \oplus \boldsymbol{S_0})\bar{\lambda}_i + (\boldsymbol{D_1} \otimes \boldsymbol{I})\bar{\lambda}_i^2 \right] u_i = 0, \quad (37)$$

which can be written as

$$\det \left[(\boldsymbol{D_0}\bar{\lambda}_i + \boldsymbol{D_1}\bar{\lambda}_i^2) \oplus (\boldsymbol{S_1} + \boldsymbol{S_0}\bar{\lambda}_i) \right] = 0 \quad (38)$$

and

$$\left[(\boldsymbol{D_0}\bar{\lambda}_i + \boldsymbol{D_1}\bar{\lambda}_i^2) \oplus (\boldsymbol{S_1} + \boldsymbol{S_0}\bar{\lambda}_i) \right] u_i = 0 \quad (39)$$

Now, if $\boldsymbol{A}$ and $\boldsymbol{B}$ can both be diagonalized, then so can $\boldsymbol{A} \oplus \boldsymbol{B}$ and it can be directly verified [8, Theorem 13.16] that all of its eigenvalues and eigenvectors are the sums and Kronecker products of the eigenvalues and eigenvectors of $\boldsymbol{A}$ and $\boldsymbol{B}$, respectively.

Thus, for a given $\bar{\lambda}_i$ let δ_{ij}, for $j = 1, \ldots, N^{(in)}$, and σ_{ik}, for $k = 1, \ldots, N^{(out)}$, be the eigenvalues of $(\boldsymbol{D_0}\bar{\lambda}_i + \boldsymbol{D_1}\bar{\lambda}_i^2)$ and $(\boldsymbol{S_1} + \boldsymbol{S_0}\bar{\lambda}_i)$, respectively. Further, denote the right eigenvector associated with δ_{ij} and σ_{ik} as $u_{ij}^{(D)}$ and $u_{ik}^{(S)}$, respectively. The eigenvalues of $(\boldsymbol{D_0}\bar{\lambda}_i + \boldsymbol{D_1}\bar{\lambda}_i^2) \oplus (\boldsymbol{S_1} + \boldsymbol{S_0}\bar{\lambda}_i)$ are $\delta_{ij} + \sigma_{ik}$ and as u_i is an eigenvector with eigenvalue 0, we have by Assumption 2

$$u_i = u_{ij}^{(D)} \otimes u_{ik}^{(S)} \quad (40)$$

for some j and k such that $\delta_{ij} = -\sigma_{ik}$.

LEMMA 1. *Under Assumption 2 the vector u_i is an eigenvector of the matrices $(\boldsymbol{D_0}\bar{\lambda}_i + \boldsymbol{D_1}\bar{\lambda}_i^2) \otimes \boldsymbol{I}$ and $\boldsymbol{I} \otimes (\boldsymbol{S_1} + \boldsymbol{S_0}\bar{\lambda}_i)$.*

PROOF. By (40) we have

$$\left((\boldsymbol{D_0}\bar{\lambda}_i + \boldsymbol{D_1}\bar{\lambda}_i^2) \otimes \boldsymbol{I} \right) u_i =$$
$$\left((\boldsymbol{D_0}\bar{\lambda}_i + \boldsymbol{D_1}\bar{\lambda}_i^2) \otimes \boldsymbol{I} \right) (u_{ij}^{(D)} \otimes u_{ik}^{(S)}) =$$
$$(\boldsymbol{D_0}\bar{\lambda}_i + \boldsymbol{D_1}\bar{\lambda}_i^2) u_{ij}^{(D)} \otimes u_{ik}^{(S)} =$$
$$\delta_{ij} u_{ij}^{(D)} \otimes u_{ik}^{(S)} = \delta_{ij} u_i .$$

The same derivation applied to $\boldsymbol{I} \otimes (\boldsymbol{S_1} + \boldsymbol{S_0}\bar{\lambda}_i)$ results in

$$\left(\boldsymbol{I} \otimes (\boldsymbol{S_1} + \boldsymbol{S_0}\bar{\lambda}_i) \right) u_i = \sigma_{ik} u_i.$$

$\square$

THEOREM 5. *Under Assumption 2 the matrices $\boldsymbol{G}$, $\boldsymbol{D_0} \otimes \boldsymbol{I} + (\boldsymbol{D_1} \otimes \boldsymbol{I})\boldsymbol{G}$ and $\boldsymbol{I} \otimes \boldsymbol{S_1} + (\boldsymbol{I} \otimes \boldsymbol{S_0})\boldsymbol{G}$ commute.*

PROOF. It suffices to show that u_i is a right eigenvector of $\boldsymbol{D_0} \otimes \boldsymbol{I} + (\boldsymbol{D_1} \otimes \boldsymbol{I})\boldsymbol{G}$ and $\boldsymbol{I} \otimes \boldsymbol{S_1} + (\boldsymbol{I} \otimes \boldsymbol{S_0})\boldsymbol{G}$ as this shows that these matrices can be simultaneously diagonalized. Further,

$$(\boldsymbol{I} \otimes \boldsymbol{S_1} + (\boldsymbol{I} \otimes \boldsymbol{S_0})\boldsymbol{G}) u_i =$$
$$\left(\boldsymbol{I} \otimes \boldsymbol{S_1} + (\boldsymbol{I} \otimes \boldsymbol{S_0})\bar{\lambda}_i \right) u_i = \sigma_{ik} u_i,$$

by Lemma 1. A similar argument can be used for $\boldsymbol{D_0} \otimes \boldsymbol{I} + (\boldsymbol{D_1} \otimes \boldsymbol{I})\boldsymbol{G}$, but we need the additional requirement that $\bar{\lambda}_i \neq 0$ to conclude that u_i is an eigenvector of $(\boldsymbol{D_0} + \boldsymbol{D_1}\bar{\lambda}_i) \otimes \boldsymbol{I}$ from Lemma 1. $\square$

Completely analogue we can establish the following:

ASSUMPTION 3. *The matrix $\boldsymbol{R}$ can be diagonalized and inverted. Denote its eigenvalues as $\hat{\lambda}_1, \ldots, \hat{\lambda}_N$ and its corresponding left eigenvectors as $v_1, \ldots, v_N$. We further assume that the matrices $\boldsymbol{D_1} + \hat{\lambda}_i \boldsymbol{D_0}$ and $\boldsymbol{S_0} + \hat{\lambda}_i \boldsymbol{S_1}$ can be diagonalized for $i = 1, \ldots, N$.*

THEOREM 6. *Under Assumption 3 the matrices $\boldsymbol{R}$, $\boldsymbol{D_1} \otimes \boldsymbol{I} + \boldsymbol{R}(\boldsymbol{D_0} \otimes \boldsymbol{I})$ and $\boldsymbol{I} \otimes \boldsymbol{S_0} + \boldsymbol{R}(\boldsymbol{I} \otimes \boldsymbol{S_1})$ commute.*

We now derive some results for the matrices $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{G}}$, where $\hat{\boldsymbol{G}}$ is defined as

$$\hat{\boldsymbol{G}} = \boldsymbol{B}(-\boldsymbol{U})^{-1}. \quad (41)$$

The matrix $\boldsymbol{U}$ defined by (4) satisfies $\boldsymbol{U} = \boldsymbol{L} + \boldsymbol{F}\boldsymbol{G} = \boldsymbol{L} + \boldsymbol{R}\boldsymbol{B}$, it is non-singular, $\boldsymbol{G} = (-\boldsymbol{U})^{-1}\boldsymbol{B}$ and $\boldsymbol{R} = \boldsymbol{F}(-\boldsymbol{U})^{-1}$. Further, recall that $\hat{\boldsymbol{R}} = (-\boldsymbol{U})^{-1}\boldsymbol{F}$. It is worth noting that $\hat{\boldsymbol{R}}$ is the R-matrix of the GI/M/1-type Markov chain if we observe the MAP/MAP/1 queue only at arrival instants, while $\hat{\boldsymbol{G}}$ is not the G-matrix of the M/G/1-type Markov chain if we only observe at service completion instants since this G-matrix is the same as the one of the QBD Markov chain.

THEOREM 7. *Under Assumption 2 the matrices $\hat{\boldsymbol{G}}$ and $\boldsymbol{D_0} \otimes \boldsymbol{I} + (\boldsymbol{D_1} \otimes \boldsymbol{I})\hat{\boldsymbol{G}}$ commute.*

PROOF. By the definition of $\hat{\boldsymbol{G}}$ we have $\hat{\boldsymbol{G}} = \boldsymbol{U}\boldsymbol{G}\boldsymbol{U}^{-1}$, which implies that the eigenvalues of $\hat{\boldsymbol{G}}$ and $\boldsymbol{G}$ are the same and the right eigenvector of $\hat{\boldsymbol{G}}$ associated with $\bar{\lambda}_i$ is $\hat{u}_i = \boldsymbol{U}u_i$, where u_i is the right eigenvector of $\boldsymbol{G}$ associated with $\bar{\lambda}_i$.

We are going to show that $\hat{u}_i$ is also a right eigenvector of $(\bar{\lambda}_i \boldsymbol{D_0} + \bar{\lambda}_i^2 \boldsymbol{D_1}) \otimes \boldsymbol{I}$. To this end we are going to rely on the following three equalities

- $u_i = u_{ij}^{(D)} \otimes u_{ik}^{(S)}$,

- $(\boldsymbol{D_0}\bar{\lambda}_i + \boldsymbol{D_1}\bar{\lambda}_i^2)u_{ij}^{(D)} = \delta_{ij} u_{ij}^{(D)}$,

- $\hat{u}_i = -1/\bar{\lambda}_i(\boldsymbol{I} \otimes \boldsymbol{S_1})u_i,$

where the latter equality comes from the fact that $\hat{u}_i$ is the right eigenvector of $\hat{\boldsymbol{G}}$ associated with $\bar{\lambda}_i$ that is

$$\bar{\lambda}_i \hat{u}_i = \hat{\boldsymbol{G}}\hat{u}_i = \underbrace{\boldsymbol{B}(-\boldsymbol{U})^{-1}}_{\hat{\boldsymbol{G}}}\underbrace{\boldsymbol{U}u_i}_{\hat{u}_i} = -\underbrace{(\boldsymbol{I} \otimes \boldsymbol{S_1})}_{\boldsymbol{B}}u_i.$$

We have

$$\begin{aligned}
&[(\bar{\lambda}_i \boldsymbol{D_0} + \bar{\lambda}_i^2 \boldsymbol{D_1}) \otimes \boldsymbol{I}] \ \hat{u}_i \\
&= -1/\bar{\lambda}_i \ [(\bar{\lambda}_i \boldsymbol{D_0} + \bar{\lambda}_i^2 \boldsymbol{D_1}) \otimes \boldsymbol{I}](\boldsymbol{I} \otimes \boldsymbol{S_1})u_i \\
&= -1/\bar{\lambda}_i \ [(\bar{\lambda}_i \boldsymbol{D_0} + \bar{\lambda}_i^2 \boldsymbol{D_1})u_{ij}^{(D)} \otimes \boldsymbol{S_1}u_{ik}^{(S)}] \\
&= -1/\bar{\lambda}_i \ [\delta_{ij}u_{ij}^{(D)} \otimes \boldsymbol{S_1}u_{ik}^{(S)}] \\
&= -\delta_{ij}/\bar{\lambda}_i \ [\boldsymbol{I} \otimes \boldsymbol{S_1}](u_{ij}^{(D)} \otimes u_{ik}^{(S)}) \\
&= \delta_{ij}/\bar{\lambda}_i \ [\boldsymbol{I} \otimes \boldsymbol{S_1}](-\boldsymbol{U})^{-1}\boldsymbol{U}u_i \\
&= \delta_{ij}/\bar{\lambda}_i \ \hat{\boldsymbol{G}} \ \hat{u}_i = \delta_{ij} \ \hat{u}_i \ .
\end{aligned}$$

The remainder of the proof now exists in repeating the argument used to prove Theorem 5. $\square$

Remark 6: Given Theorem 5 it seems natural to assume that $\hat{\boldsymbol{G}}$ and $\boldsymbol{I} \otimes \boldsymbol{S_1} + (\boldsymbol{I} \otimes \boldsymbol{S_0})\hat{\boldsymbol{G}}$ also commute, but numerical experiments show that this is not the case. This is because $\hat{u}_i$ is not a right eigenvector of $\boldsymbol{I} \otimes \boldsymbol{S_1} + (\boldsymbol{I} \otimes \boldsymbol{S_0})\bar{\lambda}_i$.

THEOREM 8. *Under Assumption 3 the matrices $\hat{\boldsymbol{R}}$ and $\boldsymbol{I} \otimes \boldsymbol{S_0} + \hat{\boldsymbol{R}}(\boldsymbol{I} \otimes \boldsymbol{S_1})$ commute.*

PROOF. Following the pattern of the proof of Theorem 7 for the matrices $\boldsymbol{R}$ and $\hat{\boldsymbol{R}}$ we obtain the statement. The key observation is that $\hat{\boldsymbol{R}}$ and $\boldsymbol{R}$ have the same eigenvalues as $\hat{\boldsymbol{R}} = \boldsymbol{U}^{-1}\boldsymbol{R}\boldsymbol{U}$ and the left eigenvector $\hat{v}_i$ of $\hat{\boldsymbol{R}}$ corresponding to $\hat{\lambda}_i$ is given by $v_i\boldsymbol{U}$, which allows one to obtain the identity $\hat{\lambda}_i\hat{v}_i = -\hat{v}_i(\boldsymbol{D_1} \otimes \boldsymbol{I})$. $\square$

5.2   Queueing based approach
In this section we basically prove Theorem 5 to 8 using a different approach such that Assumptions 2 and 3 are not required. We start by showing that

$$(-\boldsymbol{U})^{-1} = \int_{u=0}^{\infty} e^{\boldsymbol{T}u}(e^{\boldsymbol{D_0}u} \otimes \boldsymbol{I})du, \qquad (42)$$

using the stochastic interpretation of $(-\boldsymbol{U})^{-1}$ and $e^{\boldsymbol{T}u}$. More specifically, entry (i,j), with $i = (i_1, i_2)$ and $j = (j_1, j_2)$, of $(-\boldsymbol{U})^{-1}$ holds the expected amount of time that the arrival and service processes spend in state j_1 and j_2, respectively, while there is a single customer in the queue during a busy period that was initiated while the arrival and service process were in state i_1 and i_2, respectively. Next, consider the probabilistic interpretation of entry (i,k) of $e^{\boldsymbol{T}u}$ with $k = (k_1, k_2)$ [13]: it is the expected number of times during a busy period that the age of the customer c in service equals u, the current service state equals k_2 and the state of the arrival process was k_1 when customer c arrived, given

that the busy period was initiated in state $i = (i_1, i_2)$. Thus, each of these visits contributes to entry (i, j) of $(-\boldsymbol{U})^{-1}$ if $j_2 = k_2$ and there are no arrivals in an interval of length u after customer c arrived and the state of the arrival process is k_1 at the start and j_1 at the end of the interval, which is given by entry (k_1, j_1) of the matrix $e^{\boldsymbol{D_0}u}$. This establishes (42).

Equation (42) implies that

$$\boldsymbol{T}(-\boldsymbol{U})^{-1} + (-\boldsymbol{U})^{-1}(\boldsymbol{D_0} \otimes \boldsymbol{I}) = -\boldsymbol{I}, \qquad (43)$$

as $\boldsymbol{X} = -\int_{u=0}^{\infty} e^{\boldsymbol{A}u}\boldsymbol{C}e^{\boldsymbol{B}u}du$ is the unique solution of $\boldsymbol{A}\boldsymbol{X} + \boldsymbol{X}\boldsymbol{B} = \boldsymbol{C}$ if both $\boldsymbol{A}$ and $\boldsymbol{B}$ are stable matrices [8, Theorem 13.19] (that is, the real parts of the eigenvalues of $\boldsymbol{A}$ and $\boldsymbol{B}$ are negative). It is well known that the matrix $\boldsymbol{D_0}$ is stable, while $\boldsymbol{T}$ is stable due to Lemma 2.4(b) in [14].

THEOREM 9. *The matrices $\boldsymbol{G}$, $(\boldsymbol{D_0} \otimes \boldsymbol{I}) + (\boldsymbol{D_1} \otimes \boldsymbol{I})\boldsymbol{G}$ and $(\boldsymbol{I} \otimes \boldsymbol{S_1}) + (\boldsymbol{I} \otimes \boldsymbol{S_0})\boldsymbol{G}$ commute.*

PROOF. To simplify the notation we introduce $\mathcal{D}\boldsymbol{G} = (\boldsymbol{D_0} \otimes \boldsymbol{I}) + (\boldsymbol{D_1} \otimes \boldsymbol{I})\boldsymbol{G}$ and $\mathcal{S}\boldsymbol{G} = (\boldsymbol{I} \otimes \boldsymbol{S_1}) + (\boldsymbol{I} \otimes \boldsymbol{S_0})\boldsymbol{G}$. First, post-multiply (43) by $(\boldsymbol{I} \otimes \boldsymbol{S_1})$ and use the fact that $\boldsymbol{G} = (-\boldsymbol{U})^{-1}(\boldsymbol{I} \otimes \boldsymbol{S_1})$ to obtain

$$\boldsymbol{T}\boldsymbol{G} + \boldsymbol{G}(\boldsymbol{D_0} \otimes \boldsymbol{I}) = -(\boldsymbol{I} \otimes \boldsymbol{S_1}),$$

where we also used the fact that $(\boldsymbol{I} \otimes \boldsymbol{S_1})$ and $(\boldsymbol{D_0} \otimes \boldsymbol{I})$ commute. Using (32) and $\hat{\boldsymbol{R}} = (-\boldsymbol{U})^{-1}(\boldsymbol{D_1} \otimes \boldsymbol{I})$ yields

$$(\boldsymbol{I} \otimes \boldsymbol{S_0})\boldsymbol{G} + \boldsymbol{G}(\boldsymbol{D_1} \otimes \boldsymbol{I})\boldsymbol{G} + \boldsymbol{G}(\boldsymbol{D_0} \otimes \boldsymbol{I}) = -(\boldsymbol{I} \otimes \boldsymbol{S_1}).$$

In other words,

$$\boldsymbol{G} \ \mathcal{D}\boldsymbol{G} \ = -(\boldsymbol{I} \otimes \boldsymbol{S_0})\boldsymbol{G} - (\boldsymbol{I} \otimes \boldsymbol{S_1}). \qquad (44)$$

From the quadratic equation (5) for $\boldsymbol{G}$ we find

$$\mathcal{D}\boldsymbol{G} \ \boldsymbol{G} = -(\boldsymbol{I} \otimes \boldsymbol{S_0})\boldsymbol{G} - (\boldsymbol{I} \otimes \boldsymbol{S_1}), \qquad (45)$$

meaning $\mathcal{D}\boldsymbol{G} \ \boldsymbol{G} = \boldsymbol{G} \ \mathcal{D}\boldsymbol{G}$. By (44)

$$\mathcal{S}\boldsymbol{G} \ \boldsymbol{G} = -\boldsymbol{G} \ \mathcal{D}\boldsymbol{G} \ \boldsymbol{G},$$

while by (45), we have

$$\boldsymbol{G} \ \mathcal{S}\boldsymbol{G} \ = -\boldsymbol{G} \ \mathcal{D}\boldsymbol{G} \ \boldsymbol{G},$$

which yields $\boldsymbol{G} \ \mathcal{S}\boldsymbol{G} \ = \ \mathcal{S}\boldsymbol{G} \ \boldsymbol{G}$. Finally, if $\boldsymbol{G}$ commutes with $\mathcal{D}\boldsymbol{G}$ and $\mathcal{S}\boldsymbol{G}$, then

$$\begin{aligned}
\mathcal{S}\boldsymbol{G} \ \mathcal{D}\boldsymbol{G} &= (\boldsymbol{I} \otimes \boldsymbol{S_1}) \ \mathcal{D}\boldsymbol{G} \ + (\boldsymbol{I} \otimes \boldsymbol{S_0}) \ \mathcal{D}\boldsymbol{G} \ \boldsymbol{G} \\
&= (\boldsymbol{I} \otimes \boldsymbol{D_0}) \ \mathcal{S}\boldsymbol{G} \ + (\boldsymbol{I} \otimes \boldsymbol{D_1}) \ \mathcal{S}\boldsymbol{G} \ \boldsymbol{G} = \ \mathcal{D}\boldsymbol{G} \ \mathcal{S}\boldsymbol{G}.
\end{aligned}$$

$\square$

THEOREM 10. *The matrices $\boldsymbol{R}$, $(\boldsymbol{I} \otimes \boldsymbol{S_0}) + \boldsymbol{R}(\boldsymbol{I} \otimes \boldsymbol{S_1})$ and $(\boldsymbol{D_1} \otimes \boldsymbol{I}) + \boldsymbol{R}(\boldsymbol{D_0} \otimes \boldsymbol{I})$ commute.*

PROOF. Introduce $\mathcal{S}\boldsymbol{R} = (\boldsymbol{I} \otimes \boldsymbol{S_0}) + \boldsymbol{R}(\boldsymbol{I} \otimes \boldsymbol{S_1})$ and $\mathcal{D}\boldsymbol{R} = \boldsymbol{D_1} \otimes \boldsymbol{I} + \boldsymbol{R}(\boldsymbol{D_0} \otimes \boldsymbol{I})$. By pre-multiplying (43) with $(\boldsymbol{D_1} \otimes \boldsymbol{I})$ one finds

$$(\boldsymbol{D_1} \otimes \boldsymbol{I})\boldsymbol{T}(-\boldsymbol{U})^{-1} + \boldsymbol{R}(\boldsymbol{D_0} \otimes \boldsymbol{I}) = -(\boldsymbol{D_1} \otimes \boldsymbol{I}).$$

Using the expression for T and $\tilde{R}$ shows that

$$(I \otimes S_0)R + R(I \otimes S_1)R = -(D_1 \otimes I) - R(D_0 \otimes I),$$

that is,

$$\mathcal{SR}\, R = -(D_1 \otimes I) - R(D_0 \otimes I). \qquad (46)$$

The fact that R and $\mathcal{SR}$ commute now follows from the fact that quadratic equation (3) for R can be written as

$$R\, \mathcal{SR} = -(D_1 \otimes I) - R(D_0 \otimes I). \qquad (47)$$

Equation (46) implies

$$R\, \mathcal{DR} = -R\, \mathcal{SR}\, R,$$

while (47) yields

$$\mathcal{DR}\, R = -R\, \mathcal{SR}\, R,$$

meaning R and $\mathcal{DR}$ commute. Similar to the G matrix, as R commutes with $\mathcal{SR}$ and $\mathcal{DR}$, $\mathcal{DR}$ and $\mathcal{SR}$ also commute. $\square$

THEOREM 11. *The matrices $\hat{G}$ and $D_0 \otimes I + (D_1 \otimes I)\hat{G}$ commute.*

PROOF. Let $\mathcal{D\hat{G}} = D_0 \otimes I + (D_1 \otimes I)\hat{G}$ and $\mathcal{S\hat{R}} = I \otimes S_0 + \hat{R}(I \otimes S_1)$. Pre-multiplying (43) with $(I \otimes S_1)$ gives

$$\hat{G}(D_0 \otimes I) = (I \otimes S_1)[TU^{-1} - I],$$

which indicates that

$$\hat{G}\, \mathcal{D\hat{G}}$$
$$= (I \otimes S_1)[TU^{-1} - I] + \hat{G}(D_1 \otimes I)(I \otimes S_1)(-U)^{-1}$$
$$= (I \otimes S_1)[TU^{-1} - I + \hat{R}(I \otimes S_1)(-U)^{-1}].$$

Using the expression $T = \mathcal{S\hat{R}}$ yields

$$\hat{G}\, \mathcal{D\hat{G}} = (I \otimes S_1)[(I \otimes S_0)U^{-1} - I]. \qquad (48)$$

Further, by definition of $\mathcal{D\hat{G}}$ and the fact that $\hat{G} = (I \otimes S_1)(-U)^{-1}$ and $\hat{G}^2 = (I \otimes S_1)G(-U)^{-1}$, we have

$$\mathcal{D\hat{G}}\, \hat{G} = (I \otimes S_1)[(D_0 \otimes I) + (D_1 \otimes I)G](-U)^{-1}.$$

As $U = (I \otimes S_0) + \mathcal{DG}$, we get

$$\mathcal{D\hat{G}}\, \hat{G} = (I \otimes S_1)[(I \otimes S_0)U^{-1} - I]. \qquad (49)$$

Hence, $\mathcal{D\hat{G}}\, \hat{G} = \hat{G}\, \mathcal{D\hat{G}}$ due to (48) and (49). $\square$

THEOREM 12. *The matrices $\hat{R}$ and $I \otimes S_0 + \hat{R}(I \otimes S_1)$ commute.*

PROOF. Post-multiplying (43) by $(D_1 \otimes I)$ implies that

$$T(-U)^{-1}(D_1 \otimes I) = (U^{-1}(D_0 \otimes I) - I)(D_1 \otimes I).$$

As noted before $\mathcal{S\hat{R}} = T$ and $\hat{R} = (-U)^{-1}(D_1 \otimes I)$, meaning

$$\mathcal{S\hat{R}}\, \hat{R} = (U^{-1}(D_0 \otimes I) - I)(D_1 \otimes I).$$

Since $U = (D_0 \otimes I) + \mathcal{SR}$, we therefore get

$$\mathcal{S\hat{R}}\, \hat{R} = -[U^{-1}(I \otimes S_0) + U^{-1}R(I \otimes S_1)](D_1 \otimes I).$$

As $R = (D_1 \otimes I)(-U)^{-1}$, $\hat{R} = (-U)^{-1}(D_1 \otimes I)$ and $(D_1 \otimes I)$ commutes with $(I \otimes S_0)$ and $(I \otimes S_1)$, this implies

$$\mathcal{S\hat{R}}\, \hat{R} = \hat{R}(I \otimes S_0) + \hat{R}^2(I \otimes S_1) = \hat{R}\, \mathcal{S\hat{R}}.$$

$\square$

6. BMAP ARRIVALS AND SERVICES

In this section we indicate that Theorems 5 to 8 can be generalized to the case where either the arrival or service process is a batch Markovian arrival process (BMAP). When we establish these results we will restrict ourselves to the spectral decomposition approach. It is also possible to derive these results by extending the queueing based approach of Section 5.2.

6.1 BMAP/MAP/1 queue

The M/G/1-type Markov chain describing the behavior of a BMAP/MAP/1 queue has the following structure [11]

$$\begin{bmatrix} D_0 \otimes I & D_1 \otimes I & D_2 \otimes I & \dots \\ I \otimes S_1 & D_0 \oplus S_0 & D_1 \otimes I & D_2 \otimes I & \dots \\ & I \otimes S_1 & D_0 \oplus S_0 & D_1 \otimes I & \dots \\ & & \ddots & \ddots & \ddots \end{bmatrix},$$

where the matrices D_k, for $k \geq 0$, characterize the BMAP. The characteristic matrix G of this M/G/1-type Markov chain is the minimal non-negative solution of

$$0 = (I \otimes S_1) + (D_0 \oplus S_0)G + \sum_{i=1}^{\infty}(D_i \otimes I)G^{i+1}$$

and matrices U and $\hat{G}$ can be defined as

$$U = (D_0 \oplus S_0) + \sum_{i=1}^{\infty}(D_i \otimes I)G^i,$$

$$\hat{G} = (I \otimes S_1)(-U)^{-1}.$$

ASSUMPTION 4. *The matrix G can be diagonalized and inverted. Denote its eigenvalues as $\bar{\lambda}_1, \dots, \bar{\lambda}_N$ and its corresponding right eigenvectors as $u_1, \dots, u_N$. We further assume that the matrices $\sum_{i=0}^{\infty}\bar{\lambda}_i^k D_k$ and $S_1 + \bar{\lambda}_i S_0$ can be diagonalized for $i = 1, \dots, N$.*

The proofs of the following two theorems follow the same pattern as the one of Theorem 5 and 7, respectively. A similar commutativity of G and $\sum_{i=0}^{\infty}D_i G^i$ is mentioned by Lucantoni in [9, Corollary 1] for the BMAP/G/1 queue.

THEOREM 13. *Under Assumption 4 the matrices G, $\sum_{i=0}^{\infty}(D_i \otimes I)G^i$ and $I \otimes S_1 + (I \otimes S_0)G$ commute.*

THEOREM 14. *Under Assumption 4 the matrices $\hat{G}$ and $\sum_{i=0}^{\infty}(D_i \otimes I)\hat{G}^i$ commute.*

6.2 MAP/BMAP/1 queue

The structure of the transition matrix of the GI/M/1-type Markov chain describing the behavior of an MAP/BMAP/1 queue is [10]

$$
\begin{bmatrix}
D_0 \otimes I & D_1 \otimes I & & & \\
I \otimes \hat{S}_1 & D_0 \oplus S_0 & D_1 \otimes I & & \\
I \otimes \hat{S}_2 & I \otimes S_1 & D_0 \oplus S_0 & D_1 \otimes I & \\
I \otimes \hat{S}_3 & I \otimes S_2 & I \otimes S_1 & D_0 \oplus S_0 & D_1 \otimes I \\
\vdots & \ddots & \ddots & \ddots & \ddots
\end{bmatrix}
$$

where $\hat{S}_i = \sum_{j=i}^{\infty} S_j$ and the matrices S_k, for $k \geq 0$, characterize the BMAP.

The characteristic matrix R of this GI/M/1-type Markov chain is the minimal non-negative solution of

$$
0 = (D_1 \otimes I) + R(D_0 \oplus S_0) + \sum_{i=1}^{\infty} R^{i+1}(I \otimes S_i). \tag{50}
$$

Reordering (50) results in

$$
R = (D_1 \otimes I)\left(-(D_0 \oplus S_0) - \sum_{i=1}^{\infty} R^i(I \otimes S_i)\right)^{-1}.
$$

Based on this decomposition we introduce the matrices U and $\hat{R}$ as follows

$$
U = (D_0 \oplus S_0) + \sum_{i=1}^{\infty} R^i(I \otimes S_i),
$$

$$
\hat{R} = (-U)^{-1}(D_1 \otimes I).
$$

ASSUMPTION 5. *The matrix R can be diagonalized and inverted. Denote its eigenvalues as $\hat{\lambda}_1, \ldots, \hat{\lambda}_N$ and its corresponding left eigenvectors as $v_1, \ldots, v_N$. We further assume that the matrices $D_1 + \hat{\lambda}_i D_0$ and $\sum_{k=0}^{\infty} \lambda_i^k S_k$ can be diagonalized for $i = 1, \ldots, N$.*

The following two theorems can be proven similar to Theorem 6 and 8, respectively:

THEOREM 15. *Under Assumption 5 the matrices R, $D_1 \otimes I + R(D_0 \otimes I)$ and $\sum_{i=0}^{\infty} R^i(I \otimes S_i)$ commute.*

THEOREM 16. *Under Assumption 5 the matrices $\hat{R}$ and $\sum_{i=0}^{\infty} \hat{R}^i(I \otimes S_i)$ commute.*

Acknowledgment

7. REFERENCES

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

[2] V. Gupta, M. Burroughs, and M. Harchol-Balter. Analysis of scheduling policies under correlated job sizes. *Perform. Eval.*, 67(11):996–1013, November 2010.

[3] Q. He. The versatility of the MMAP[K] and the MMAP[K]/G[K]/1 queue. *Queueing Systems*, 38:397–418, 2001.

[4] Q. He. Analysis of a continuous time MMAP[K]/PH[K]/1/FCFS queue: Age process, sojourn times, and queue lengths. *Journal of Systems Science and Complexity*, 25:133–155, 2012.

[5] A. Heindl, Q. Zhang, and E. Smirni. ETAQA truncation models for the MAP/MAP/1 departure process. In *QEST '04: Proceedings of the The Quantitative Evaluation of Systems, First International Conference*, pages 100–109, Washington, DC, USA, 2004.

[6] A. Horváth, G. Horváth, and M. Telek. A joint moments based analysis of networks of MAP/MAP/1 queues. *Performance Evaluation*, 67:759–778, 2010.

[7] G. Horváth, P. Buchholz, and M. Telek. A MAP fitting approach with independent approximation of the inter-arrival time distribution and the lag correlation. In *Proc of QEST 2005*. IEEE Computer Society, 2005.

[8] A. J. Laub. *Matrix Analysis For Scientists And Engineers.* Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2004.

[9] D. M. Lucantoni. New results on the single server queue with a batch Markovian arrival process. *Stochastic models*, 7(1):1–46, 1991.

[10] M. F. Neuts. *Matrix-Geometric Solutions in Stochastic Models.* John Hopkins University Press, Baltimore, MD, USA, 1981.

[11] M. F. Neuts. *Structured Stochastic Matrices of M/G/1-type and their Applications.* Marcel Dekker, New York, NY, 1989.

[12] T. Ozawa. Sojourn time distributions in the queue defined by a general QBD process. *Queueing Syst. Theory Appl.*, 53(4):203–211, August 2006.

[13] B. Sengupta. Markov processes whose steady state distribution is matrix exponential with an application to the GI/PH/1 queue. *Adv. in Appl. Probab.*, 21:159–180, 1989.

[14] B. Sengupta. The semi-Markovian queue: theory and applications. *Stochastic Models*, 6(3):383–413, 1990.

[15] M. Telek and G. Horváth. A minimal representation of markov arrival processes and a moments matching method. *Perform. Eval.*, 64(9-12):1153–1168, October 2007.