Product-form solutions for a class of structured multi-dimensional Markov processes

Jori Selen*[†], Ivo J.B.F. Adan*[†], and Johan S.H. van Leeuwaarden[†]

April 3, 2022

Abstract

Motivated by queueing systems with heterogeneous parallel servers, we consider a class of structured multi-dimensional Markov processes whose state space can be partitioned into two parts: a finite set of boundary states and a structured multi-dimensional set of states, exactly one dimension of which is infinite. Using *separation of variables*, we show that the equilibrium distribution, typically of the queue length, can be represented as a linear combination of product forms. For an important subclass of queueing systems, we characterize explicitly the waiting time distribution in terms of mixtures of exponentials.

1 Introduction

Motivated by queueing systems with heterogeneous parallel servers, we introduce a class of structured multi-dimensional Markov processes with one infinite dimension. While most research on multi-server queueing systems focuses on identical, or homogeneous servers, our class includes multi-server queueing systems with nonidentical, or heterogeneous servers.

Among the key examples of queueing systems included in this class are heterogeneous parallel servers with batch arrivals and service times consisting of two exponential phases (e.g. Erlang-2 [17], Coxian-2 [7], hyperexponential-2 [18, 19]), multi-server queueing systems with service interruptions [15], setup times [6, 11] or batch service [4], and a single exponential server fed by a superposition of independent, interrupted Poisson processes. Such queueing systems naturally arise in manufacturing, where jobs may arrive in batches, to be subsequently processed on multiple parallel batch machines. These machines can be nonidentical, differing in processing speed and processing variability and may require setup time after having been idle. For an overview of stochastic models, and queueing models in particular, for manufacturing systems we refer to [8].

The class of structured multi-dimensional Markov processes presented in this paper is an extension of the one in [3]. Whereas [3] deals with identical parallel servers, we allow for nonidentical parallel servers and batch arrivals. For this extended class we will determine the equilibrium distribution using a *separation of variables* technique. This technique allows us to express the equilibrium distribution, typically of the queue length, as a linear combination of product forms (in terms of eigenvalues, and eigenvectors of the finite dimensions). Alternative methods to determine the equilibrium distribution for structured multi-dimensional Markov

---

*Department of Mechanical Engineering, Eindhoven University of Technology, The Netherlands

[†]Department of Mathematics and Computer Science, Eindhoven University of Technology, The Netherlands
E-mail address: j.selen@tue.nl

processes are the matrix-geometric method [16] and the spectral expansion method [14]. The matrix-geometric method expresses the equilibrium distribution in terms of the rate matrix R, which is the minimal nonnegative solution of a nonlinear matrix equation. When R is diagonalizable, the parameters of the product forms of the separation of variables technique are the eigenvalues and left eigenvectors of R. The spectral expansion method treats the equilibrium equations as a homogeneous vector difference equation with constant coefficients, and by substituting product forms, reduces to a single characteristic matrix polynomial for all eigenvalues.

By employing separation of variables, however, the single equation for all eigenvalues can be *decomposed* into a system of equations for (much) fewer eigenvalues. Further, we can express the equilibrium distribution as a linear combination of product forms, the parameters of which can be obtained explicitly. Moreover, for an important subclass of queueing systems, we are able to obtain explicitly the waiting time distribution in terms of mixtures of exponentials.

1.1 Motivating examples

We present three motivating examples that fall within the class of processes considered in this paper. The first two examples are extensions of examples in [3]. The third example serves as a running example throughout the paper.

Example 1.1. (*Parallel heterogeneous servers with Erlang-2 services*) This example is an extension of the class of Markov processes presented in [3] by allowing different service behavior for each server. Jobs arrive according to a Poisson process with rate λ. Jobs are served FCFS by c heterogeneous servers. The service time of server i consists of two exponential phases, both with mean $1/\mu_i$.

Example 1.2. (*The $M^X/M/c$ queue with service interruptions*) This example considers the same class as in [3], but now allowing for batch arrivals. Jobs arrive in batches of sizes that range from 1 to K. Batches of size k arrive according to a Poisson process with rate λ_k. Jobs are served FCFS by c identical parallel servers. Each server, when operative, serves jobs with rate μ. However, servers are subject to breakdowns. The times that servers are operative are exponentially distributed with parameter θ. The repair times are exponentially distributed with parameter ν.

Example 1.3. (*Parallel heterogeneous servers with hypoexponential-2 service times and Poisson batch arrivals*) This example serves as a running example throughout the paper. We consider a queueing model that incorporates both extensions mentioned above. Jobs arrive in batches of sizes that range from 1 to K. Batches of size k arrive according to a Poisson process with rate λ_k. Jobs are served FCFS by c parallel heterogeneous servers. The service time of server i consists of two exponential phases with mean $1/\mu_{1,i}$ and $1/\mu_{2,i}$, respectively.

1.2 Description of the class of processes

We consider a class of irreducible multi-dimensional Markov processes on the state space $V \cup W$ with V a finite set and W defined by all states $\mathbf{n} = (n_0, n_1, \ldots, n_c)$ with $n_0 \in \mathbb{N}_0$ and $n_i \in \{0,1\}$, $i = 1, \ldots, c$. The set V serves to describe possibly less structured boundary behavior. For all three motivating examples of Section 1.1, the set V consists of the states with one or more idle servers. In this paper we primarily focus on the analysis of the structured infinite set of states W. The determination of the probabilities in the finite set V is of minor importance and not discussed here.
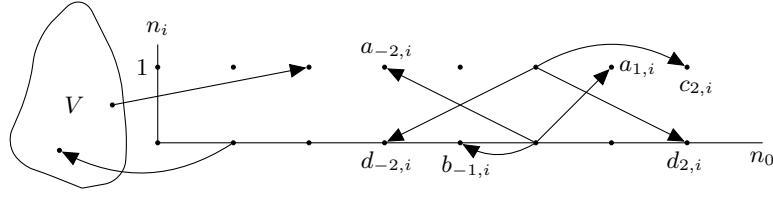
Figure 1: Transition rate diagram of the process in the (n_0, n_i)-plane.

In the n_0-dimension, we allow for jumps in the negative direction of any size, whereas jumps in the positive direction are of maximum size K with K some finite positive integer. We call the states in V and the states with $n_0 < K$ the *boundary states*. For the states in V we assume that no transitions to states $\mathbf{n}$ with $n_0 \geq K$ are possible, i.e., the only way to enter the states $\mathbf{n}$ with $n_0 \geq K$ from the set V is via the states $\mathbf{n}$ with $n_0 < K$.

Transitions in the (n_0, n_i)-plane with $n_0 \geq 0$ are assumed to be independent of transitions in other (n_0, n_j)-planes, $j \neq i$. This allows us to describe in isolation the transition structure on this two-dimensional (n_0, n_i)-plane. More specifically, for the states with $n_0 \geq 0$, the following transitions in the (n_0, n_i)-plane are possible, with $k = -n_0, -n_0 + 1, \ldots, K$:

- From $(n_0, 0)$ to $(n_0 + k, 1)$ with rate $a_{k,i}$.

- From $(n_0, 0)$ to $(n_0 + k, 0)$ with rate $b_{k,i}$.

- From $(n_0, 1)$ to $(n_0 + k, 1)$ with rate $c_{k,i}$.

- From $(n_0, 1)$ to $(n_0 + k, 0)$ with rate $d_{k,i}$.

- From (n_0, n_i) to the states in V with total rate $\sum_{k=-\infty}^{-n_0-1} \big((1-n_i)(a_{k,i}+b_{k,i})+n_i(c_{k,i}+d_{k,i})\big)$.

Hence, transitions from $(n_0, \cdot)$ are only possible to states $(n, \cdot)$ with $n \leq n_0 + K$. The transition rate diagram of the (n_0, n_i)-plane is shown in Figure 1.

We assume that the Markov process is ergodic and denote by $p(\mathbf{n})$ the equilibrium probability of being in state $\mathbf{n}$. One of the main results obtained in this paper is the following expression for the equilibrium distribution in terms of a linear combination of $2^c K$ product forms:

$$p(\mathbf{n}) = \sum_{j=1}^{2^c K} \alpha_j \beta_{0,j}^{n_0} \beta_{1,j}^{n_1} \cdots \beta_{c,j}^{n_c}, \quad \mathbf{n} \in W, \tag{1.1}$$

where $\beta_{i,j}$ and α_j are possibly complex-valued constants. We are able to establish (1.1) for the case $K = 1$, and also for $K \geq 1$ provided the Markov process is *symmetric*, by which we mean that the transition rate diagram of each (n_0, n_i)-plane is identical. More specifically, by symmetry we mean that, for given k all $a_{k,i}$, $i = 1, \ldots, c$, are equal and similarly for the other rates. For $K \geq 2$ and nonsymmetric Markov processes (all Markov processes that are not symmetric), the proposed method might still work, as is supported by numerical evidence (not shown in this paper). However, proving that (1.1) is valid in this case is much more involved. In order to establish (1.1), we employ a separation of variables technique, which is well known from the theory of partial differential equations [12] and has been applied before to an $E_k/E_r/c$ queueing model [5]. Using this technique, a relation between $\beta_{i,j}$, $i = 1, \ldots, c$, and $\beta_{0,j}$ is derived. Instrumental for this technique is to identify the constants $\beta_{0,j}$ as the roots within the unit circle of a certain equation.

3

For symmetric Markov processes we can use an aggregated state space described by (n_0, m) with $m = \sum_{i=1}^{c} n_i$, and which forms a semi-infinite strip of states for which the equilibrium probabilities can be expressed as

$$p(n_0, m) = \sum_{j=1}^{K(c+1)} \gamma_j \beta_{0,j}^{n_0} \omega_j(m), \quad n_0 \in \mathbb{N}_0, \ m = 0, \ldots, c, \tag{1.2}$$

where γ_j are possibly complex-valued constants and ω_j is a row vector, the entries of which are given by $\omega_j(m) = \sum_{n_1 + \cdots + n_c = m} \beta_{1,j}^{n_1} \cdots \beta_{c,j}^{n_c}$ with $\beta_{i,j}$ as in (1.1).

In [3] the same class of symmetric Markov processes is studied for the special case $K = 1$, and a generating function technique is used to determine the equilibrium distribution. This generating function technique, however, does not seem applicable to the nonsymmetric processes studied in the this paper, which is why we use separation of variables instead.

We do use a generating function technique, together with (1.2) to describe the distribution of the first passage time to the set V, given that at time $t = 0$, the process starts in state (n_0, m) and there are no transitions $a_k, \ldots, d_k$ with $k > 0$ for $t \geq 0$. For queueing systems, this first passage time is typically tantamount to the waiting time distribution (see e.g. [1, 9]), which is also the case in this paper.

1.3   Parameters of the motivating examples

We now briefly indicate how the motivating examples of Section 1.1 fit in the class of Markov processes.

Example 1.1. (*Continued*) Here, n_0 represents the number of jobs waiting in the queue and $n_i, \ i = 1, \ldots, c$ is the number of finished phases at server i. We have $a_{0,i} = \mu_i$, $b_{1,i} = c_{1,i} = \lambda_i$ with $\sum_i \lambda_i = \lambda$, $d_{-1,i} = \mu_i$ and all other rates are equal to zero. Note that the rates $\lambda_i > 0$ can be chosen arbitrarily as long as they add up to λ.

Example 1.2. (*Continued*) In this case, n_0 represents the number of jobs waiting in the queue and $n_i, \ i = 1, \ldots, c$ states whether server i is operative. As the servers are identical we can omit the index i to obtain the rates $a_0 = \nu$, $c_{-1} = \mu$, $b_k = c_k = \lambda_k/c, \ (k = 1, \ldots, K)$, $d_0 = \theta$ and all other rates are equal to zero.

Example 1.3. (*Continued*) For this model, n_0 represents the number of jobs waiting in the queue and $n_i, \ i = 1, \ldots, c$ is the number of finished phases at server i. We have $a_{0,i} = \mu_{1,i}$, $b_{k,i} = c_{k,i} = \lambda_{k,i} \ (k = 1, \ldots, K)$ with $\sum_i \lambda_{k,i} = \lambda_k$ the arrival rate of batches of size k, $d_{-1,i} = \mu_{2,i}$ and all other rates are equal to zero.

1.4   Structure of the paper

The remainder of the paper is organized as follows. In Section 2 we introduce some further modeling assumptions and derive the ergodicity condition of the multi-dimensional Markov process. Section 3 concerns the analysis of the equilibrium equations to find the equilibrium distribution for jumps of maximum size $K = 1$. Symmetric processes with $K \geq 1$ are studied in Section 4. The aggregated state concept is introduced in Section 5. In Section 6 we study the first passage time to the set V. Finally, in Section 7 we present some conclusions and directions for further research.

2 Modeling assumptions

In this section we impose some conditions on the rates of the Markov process introduced in Section 1 and we determine the ergodicity condition of the process. First, let us introduce the generating functions

$$A_i(z) := \sum_{k=-\infty}^{K} a_{k,i} z^{K-k}, \quad B_i(z) := \sum_{k=-\infty}^{K} b_{k,i} z^{K-k},$$
$$C_i(z) := \sum_{k=-\infty}^{K} c_{k,i} z^{K-k}, \quad D_i(z) := \sum_{k=-\infty}^{K} d_{k,i} z^{K-k}. \tag{2.1}$$

We impose the following assumptions on the rates of the Markov process:

Assumption 2.1. For $i = 1, 2, \ldots, c$,

(i) $A_i(1), B_i(1), C_i(1), D_i(1) < \infty$.

(ii) $A_i(1), D_i(1) > 0$.

(iii) $A_i'(1), B_i'(1), C_i'(1), D_i'(1) < \infty$.

(iv) $a_{K,i} = 0$ or $d_{K,i} = 0$.

(v) $b_{K,i} = c_{K,i} \neq 0$.

The assumptions (i) and (iii) are intuitively clear: both the total outgoing rate and expected jump size should be finite. Assumption (ii) is imposed to exclude cases in which vertical transitions in the (n_0, n_i)-plane are only possible in one direction. Finally, assumptions (iv) and (v) are used to exclude exceptional cases. In Section 3 we show how these assumptions are used. Note that assumptions (iv) and (v) are satisfied for the motivating examples presented in Section 1.1.

We next present the ergodicity condition.

Lemma 2.2. *The Markov process is ergodic if and only if*

$$0 < \sum_{i=1}^{c} \frac{1}{A_i(1) + D_i(1)} \Big(D_i(1)(A_i'(1) - KA_i(1) + B_i'(1) - KB_i(1)) $$
$$+ A_i(1)(C_i'(1) - KC_i(1) + D_i'(1) - KD_i(1)) \Big). \tag{2.2}$$

Proof. We require that the mean aggregate drift in the negative n_0-direction is larger than the mean aggregate drift in the positive n_0-direction; see Neuts' mean drift condition [16]. This gives

$$\sum_{i=1}^{c} \sum_{k=-\infty}^{-1} k\pi_i \begin{pmatrix} b_{k,i} & a_{k,i} \\ d_{k,i} & c_{k,i} \end{pmatrix} \mathbf{1} > \sum_{i=1}^{c} \sum_{k=1}^{K} k\pi_i \begin{pmatrix} b_{k,i} & a_{k,i} \\ d_{k,i} & c_{k,i} \end{pmatrix} \mathbf{1}, \tag{2.3}$$

where $\mathbf{1} = \begin{pmatrix} 1 & 1 \end{pmatrix}^T$ and $\pi_i = \begin{pmatrix} \pi_i(0) & \pi_i(1) \end{pmatrix}$ is the equilibrium distribution of the n_i-dimension in the (n_0, n_i)-plane, which is the solution to

$$\pi_i \begin{pmatrix} -A_i(1) & A_i(1) \\ D_i(1) & -D_i(1) \end{pmatrix} = 0. \tag{2.4}$$

5

Clearly, $\pi_i = \left(\frac{D_i(1)}{A_i(1)+D_i(1)} \quad \frac{A_i(1)}{A_i(1)+D_i(1)} \right)$, and rewriting (2.3) gives

$$\sum_{i=1}^{c} \frac{1}{A_i(1) + D_i(1)} \left(D_i(1) \sum_{k=-\infty}^{K} k(a_{k,i} + b_{k,i}) + A_i(1) \sum_{k=-\infty}^{K} k(c_{k,i} + d_{k,i}) \right) < 0. \qquad (2.5)$$

Substituting $\sum_{k=-\infty}^{K} ka_{k,i} = KA_i(1) - A_i'(1)$ and similarly for the sums of $b_{k,i}$, $c_{k,i}$ and $d_{k,i}$ yields (2.2). $\qquad \square$

We henceforth assume that condition (2.2) holds.

Remark 2.3. (*Ergodicity condition for symmetric processes*) For Markov process with identical rates in each (n_0, n_i)-plane, we can omit the index i in (2.2), so that (2.2) simplifies to

$$0 < D(1)(A'(1) - KA(1) + B'(1) - KB(1)) + A(1)(C'(1) - KC(1) + D'(1) - KD(1)). \quad (2.6)$$

3 Analysis of the equilibrium equations

We now derive the equilibrium distribution $p(\mathbf{n})$ of the Markov process described in Section 1.2. We use a separation of variables technique and exploit the structure of the equilibrium equations for (non-boundary) states in the interior of the state space W.

Let $\mathbf{e}_i$ denote a row vector of zeros of length $c + 1$ with a 1 at position i with $0 \leq i \leq c$. By equating the rate out of and the rate into state $\mathbf{n}$ we obtain

$$\sum_{i=1}^{c} \sum_{k=-\infty}^{K} \Big((1 - n_i)(a_{k,i} + b_{k,i}) + n_i(c_{k,i} + d_{k,i}) \Big) p(\mathbf{n})$$
$$= \sum_{i=1}^{c} \sum_{k=-\infty}^{K} \Big((1 - n_i)\big(b_{k,i} p(\mathbf{n} - k\mathbf{e}_0) + d_{k,i} p(\mathbf{n} - k\mathbf{e}_0 + \mathbf{e}_i) \big)$$
$$+ n_i\big(a_{k,i} p(\mathbf{n} - k\mathbf{e}_0 - \mathbf{e}_i) + c_{k,i} p(\mathbf{n} - k\mathbf{e}_0) \big) \Big), \qquad (3.1)$$

which is valid for all states $\mathbf{n}$ with $n_0 \geq K$ and $n_i \in \{0, 1\}$, $i = 1, \ldots, c$. The equations (3.1) form the *inner equations*, while the equilibrium equations for states with $n_0 < K$ and the set V form the *boundary equations*. As it turns out, the precise form of the boundary equations is not relevant for the first part of the analysis and these equations are therefore not presented. The boundary equations will be used to determine coefficients of the linear combination in Theorems 3.6, 4.8 and 5.1. Note that the number of boundary equations is equal to $|V| + 2^c K$.

We search for linearly independent candidate solutions $p(\mathbf{n})$ of (3.1) for which $\sum_{\mathbf{n}} p(\mathbf{n})$ absolutely converges. We have the following lemma.

Lemma 3.1. *If there are $2^c K$ linearly independent and absolutely convergent solutions of (3.1), labeled $p_j(\mathbf{n})$, $j = 1, \ldots, 2^c K$, then the equilibrium distribution $p(\mathbf{n})$ can be expressed as*

$$p(\mathbf{n}) = \sum_{j=1}^{2^c K} \alpha_j p_j(\mathbf{n}), \quad \mathbf{n} \in W, \qquad (3.2)$$

where the (possibly complex-valued) constants α_j are uniquely determined from the boundary equations and the normalization condition.

Proof. The proof is along similar lines as the proof of [5, Theorem 4.1]. For each choice of α_j, the sequence $p(\mathbf{n})$ given by (3.2) satisfies the equations (3.1). The remaining equations are the boundary equations. These equations form a linear, homogeneous system for the unknown constants α_j and the boundary probabilities in V. The number of equations is equal to the number of unknowns, namely $|V| + 2^c K$. By first omitting the equilibrium equation in state $\mathbf{n} = \mathbf{0}$, say, the reduced homogeneous system of equilibrium equations obviously has a nonzero solution. The equilibrium equation in $\mathbf{n} = \mathbf{0}$ is automatically satisfied, since inserting $p(\mathbf{n})$ in the equations in all the other states and changing summations exactly yields the desired equation. Change of summations is allowed, since the sum of $p(\mathbf{n})$ over all states converges absolutely. It remains to show that $p(\mathbf{n})$ is a nonzero solution. This readily follows if at least one of the boundary probabilities is nonzero. If all these probabilities are zero, then at least one of the constants α_j must be nonzero. In this case, the property that the solutions $p_j(\mathbf{n})$, $j = 1, \ldots, 2^c K$ are linearly independent implies that $p(\mathbf{n})$ is a nonzero solution. Linear independence means that a linear combination of the solutions $\sum_{j=1}^{2^c K} \alpha_j p_j(\mathbf{n})$ is equal to zero if and only if $\alpha_j = 0$ for all j. Using [10, Theorem 1], we can conclude that the Markov process is ergodic and that a normalized version of $p(\mathbf{n})$ produces the equilibrium distribution. The uniqueness of the constants α_j follows from the uniqueness of the equilibrium distribution and the independence of the solutions $p_j(\mathbf{n})$, $j = 1, \ldots, 2^c K$. $\qquad \square$

From Lemma 3.1 and the uniqueness of the equilibrium distribution we obtain the following result.

Corollary 3.2. *The dimension of the space of absolutely convergent solutions of* (3.1) *is at most* $2^c K$.

Proof. Say we find $2^c K + 1$ linearly independent, absolutely convergent solutions. Let us label these solutions as $p_j(\mathbf{n})$, $j = 1, \ldots, 2^c K + 1$. According to Lemma 3.1 we can express $p(\mathbf{n})$ as a linear combination of the first $2^c K$ products. We know that at least one of the coefficients α_j has to be nonzero (since $p(\mathbf{n}) > 0$). Without loss of generality we assume that $\alpha_1 \neq 0$. Let us now construct $p(\mathbf{n})$ with the last $2^c K$ solutions and label the coefficients as ξ_j, thus ignoring the first product. As both sets of $2^c K$ linearly independent, absolutely convergent solutions represent the same equilibrium probability, taking the difference yields

$$\alpha_1 p_1(\mathbf{n}) + \sum_{j=2}^{2^c K} (\alpha_j - \xi_j) p_j(\mathbf{n}) - \xi_{2^c K+1} \; p_{2^c K+1}(\mathbf{n}) = 0. \tag{3.3}$$

The solutions $p_j(\mathbf{n})$, $j = 1, \ldots, 2^c K + 1$ are linearly independent and therefore the only way to construct the zero solution is by settings all coefficients equal to 0. As $\alpha_1 \neq 0$, we get a contradiction, which implies that the corollary holds. $\qquad \square$

Lemma 3.1 reduces the problem of determining the equilibrium distribution to that of finding $2^c K$ linearly independent solutions of (3.1). We next seek solutions satisfying the inner conditions (3.1) of the special form

$$p(\mathbf{n}) = \beta_0^{n_0} \beta_1^{n_1} \cdots \beta_c^{n_c}. \tag{3.4}$$

Clearly, only products that can be normalized, i.e. for which the sum over all states converges absolutely, are useful. This implies that $|\beta_0| < 1$.

Substituting (3.4) into the equilibrium equations, dividing by $\beta_0^{n_0-K}\beta_1^{n_1}\cdots\beta_c^{n_c}$ and bringing all terms to one side, gives

$$0 = \sum_{i=1}^{c}\left((1-n_i)\Big(B_i(\beta_0)+D_i(\beta_0)\beta_i\Big)+n_i\Big(A_i(\beta_0)\frac{1}{\beta_i}+C_i(\beta_0)\Big)\right)$$
$$-\beta_0^K\sum_{i=1}^{c}\big((1-n_i)(A_i(1)+B_i(1))+n_i(C_i(1)+D_i(1))\big). \tag{3.5}$$

Since the above equation holds for all states $\mathbf{n}$ with $n_0 \geq K$ and the right-hand side of this equation must be zero, we require that the coefficients of n_i vanish, which amounts to

$$0 = A_i(\beta_0)\frac{1}{\beta_i}-B_i(\beta_0)+C_i(\beta_0)-D_i(\beta_0)\beta_i+\beta_0^K(A_i(1)+B_i(1)-C_i(1)-D_i(1)). \tag{3.6}$$

This is a quadratic equation in β_i, with solutions

$$\beta_i = \frac{F_i(\beta_0)}{2D_i(\beta_0)}+x_i\frac{\sqrt{F_i(\beta_0)^2+4A_i(\beta_0)D_i(\beta_0)}}{2D_i(\beta_0)}, \quad i=1,\ldots,c, \tag{3.7}$$

where $F_i(\beta_0)=\beta_0^K(A_i(1)+B_i(1)-C_i(1)-D_i(1))-B_i(\beta_0)+C_i(\beta_0)$ and $x_i^2=1$, so that x_i is either -1 or 1. As we cannot divide by zero, we need $D_i(\beta_0) > 0$, which is always valid if $\beta_0 \in (0,1)$ by Assumption 2.1(ii). If $\beta_0 \in (0,1)$, then $F_i(\beta_0)^2+4A_i(\beta_0)D_i(\beta_0) > 0$ and (3.7) depends on x_i. Note that β_0 is still to be determined.

We can now substitute (3.7) into (3.5) to find that

$$0 = \sum_{i=1}^{c}\Big(x_i\sqrt{F_i(\beta_0)^2+4A_i(\beta_0)D_i(\beta_0)}+(B_i(\beta_0)+C_i(\beta_0))$$
$$-\beta_0^K(A_i(1)+B_i(1)+C_i(1)+D_i(1))\Big). \tag{3.8}$$

This leads to 2^c equations for β_0 due to all possible combinations of x_i's.

Lemma 3.3. *For each combination of x_i's, and $K=1$, equation (3.8) has at least one root $\beta_0 \in (0,1)$.*

Proof. Let $h_i(\beta_0)$ denote the function inside the summation in (3.8). Assumptions 2.1(iv) and (v) imply that $F_i(\beta_0)^2+4A_i(\beta_0)D_i(\beta_0)$ is equal to zero for $\beta_0=0$, so that

$$h_i(0)=B_i(0)+C_i(0), \quad h_i(1)=(x_i-1)(A_i(1)+D_i(1)). \tag{3.9}$$

Thus,

$$\sum_{i=1}^{c}h_i(0)>0, \quad \sum_{i=1}^{c}h_i(1)\leq 0. \tag{3.10}$$

Note that $\sum_{i=1}^{c}h_i(1)=0$ only occurs for $x_i=1$, $i=1,\ldots,c$, in which case the derivative $\sum_{i=1}^{c}h_i'(1)$ is equal to the right-hand side of (2.2) and is positive. This proves that there is at least one zero in $(0,1)$. For all other combinations of x_i's, we conclude there is at least one zero as well, since $\sum_{i=1}^{c}h_i(\beta_0)$ is continuous. $\square$

By combining Lemma 3.3 and (3.7) we then find the product $\beta_0^{n_0}\cdots\beta_c^{n_c}$ for each feasible combination of x_i's, so we obtain at least 2^c products. Let us select one root β_0 obtained from each equation and label these products as $\beta_{0,j}^{n_0}\cdots\beta_{c,j}^{n_c}$, $j=1,\ldots,2^c$. In this way we have characterized the product forms satisfying the equilibrium equations (3.1).

Lemma 3.4. *The products* $\beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c}$, $j = 1, \ldots, 2^c$ *are linearly independent on* $\mathbf{n} \in W$.

Proof. The proof is based on a property of the Vandermonde matrix. Let the numbers $y_1, \ldots, y_M$ satisfy $y_i \neq y_j$, $i \neq j$. Then

$$\sum_{j=1}^{M} \phi_j y_j^n = 0, \ n = 0, 1, \ldots, M - 1 \iff \phi_j = 0, \ j = 1, \ldots, M. \tag{3.11}$$

To prove this property, note that the left-hand side of the if and only if relation in (3.11) can be written as $Y\Phi = 0$ with $\Phi = \begin{pmatrix} \phi_1 & \cdots & \phi_M \end{pmatrix}^T$ and

$$Y = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ y_1 & y_2 & \cdots & y_M \\ y_1^2 & y_2^2 & \cdots & y_M^2 \\ \vdots & \vdots & & \vdots \\ y_1^{M-1} & y_2^{M-1} & \cdots & y_M^{M-1} \end{pmatrix}. \tag{3.12}$$

Since Y is a Vandermonde matrix, its determinant is given by $\prod_{1 \leq i < j \leq M}(y_i - y_j)$, which is nonzero, since $y_i \neq y_j$, if $i \neq j$. Thus property (3.11) holds.

We want to show that the following property holds to prove linear independence:

$$\sum_{j=1}^{2^c} \alpha_j \beta_{0,j}^{n_0} \beta_{1,j}^{n_1} \cdots \beta_{c,j}^{n_c} = 0, \ \mathbf{n} \in W \iff \alpha_j = 0, \ j = 1, \ldots, 2^c. \tag{3.13}$$

Suppose the number of different $\beta_{0,j}$'s is M. For any (fixed) choice of n_i, $i = 1, \ldots, c$, we then rewrite the left-hand side of (3.13) as

$$\sum_i \beta_{0,i}^{n_0} \sum_{\beta_{0,j} = \beta_{0,i}} \alpha_j \beta_{1,j}^{n_1} \cdots \beta_{c,j}^{n_c} = 0, \quad n_0 = 0, 1, \ldots, M - 1. \tag{3.14}$$

Hence, by (3.11),

$$\sum_{\beta_{0,j} = \beta_{0,i}} \alpha_j \beta_{1,j}^{n_1} \cdots \beta_{c,j}^{n_c} = 0, \quad n_i \in \{0, 1\}, \ i = 1, \ldots, c. \tag{3.15}$$

For $n_1 \in \{0, 1\}$ and any choice of n_i, $i = 2, \ldots, c$, we rewrite the above equation as

$$\sum_{\substack{\beta_{0,j} = \beta_{0,i} \\ x_1 = -1}} \alpha_j \beta_{2,j}^{n_2} \cdots \beta_{c,j}^{n_c} \left(\frac{F_1(\beta_{0,i})}{2D_1(\beta_{0,i})} - \frac{\sqrt{F_1(\beta_{0,i})^2 + 4A_1(\beta_{0,i})D_1(\beta_{0,i})}}{2D_1(\beta_{0,i})} \right)^{n_1}$$

$$+ \sum_{\substack{\beta_{0,j} = \beta_{0,i} \\ x_1 = 1}} \alpha_j \beta_{2,j}^{n_2} \cdots \beta_{c,j}^{n_c} \left(\frac{F_1(\beta_{0,i})}{2D_1(\beta_{0,i})} + \frac{\sqrt{F_1(\beta_{0,i})^2 + 4A_1(\beta_{0,i})D_1(\beta_{0,i})}}{2D_1(\beta_{0,i})} \right)^{n_1} = 0. \tag{3.16}$$

Therefore, again by (3.11), for $n_i \in \{0, 1\}$, $i = 2, \ldots, c$,

$$\sum_{\substack{\beta_{0,j} = \beta_{0,i} \\ x_1 = -1}} \alpha_j \beta_{2,j}^{n_2} \cdots \beta_{c,j}^{n_c} = 0, \quad \text{and} \quad \sum_{\substack{\beta_{0,j} = \beta_{0,i} \\ x_1 = 1}} \alpha_j \beta_{2,j}^{n_2} \cdots \beta_{c,j}^{n_c} = 0. \tag{3.17}$$

Repeating the steps in (3.16) and (3.17) for $n_2, n_3, \ldots, n_c$, we finally get $\alpha_j = 0$ for all j in accordance with (3.13) and thus we conclude that the products are linearly independent. $\square$

Suppose we could find an additional root β_0 in one of the equations in (3.8), leading to $2^c + 1$ roots. Then we could use the same strategy as for the proof of Lemma 3.4 to prove that these $2^c + 1$ products are linearly independent. From Corollary 3.2 we deduce that finding an additional root is not possible and we formulate the following corollary.

Corollary 3.5. *For each combination of x_i's, and $K = 1$, equation* (3.8) *has exactly one root $\beta_0 \in (0, 1)$, leading to a total of 2^c roots.*

The following theorem states that the equilibrium distribution $p(\mathbf{n})$, $\mathbf{n} \in W$ for the case $K = 1$ can be expressed as a linear combination of product forms.

Theorem 3.6. *For the multi-dimensional Markov processes with $K = 1$, the equilibrium distribution takes the form*

$$p(\mathbf{n}) = \sum_{j=1}^{2^c} \alpha_j \beta_{0,j}^{n_0} \beta_{1,j}^{n_1} \cdots \beta_{c,j}^{n_c}, \quad \mathbf{n} \in W, \tag{3.18}$$

where $\beta_{0,j}$ are the roots of (3.8) *in the interval $(0, 1)$ for each combination of x_i's and $\beta_{i,j}$, $i = 1, \ldots, c$ are the corresponding constants found from* (3.7). *The real-valued constants α_j are determined from the boundary equations and the normalization condition.*

Proof. Corollary 3.5 says that we find exactly 2^c solutions β_0. Using (3.7) we find 2^c products $\beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c}$, $j = 1, \ldots, 2^c$. These products are absolutely convergent, as $|\beta_0| < 1$, and linearly independent, according to Lemma 3.4. Then, Lemma 3.1 shows that the 2^c linearly independent and absolutely convergent solutions can be used to express the equilibrium distribution as a linear combination of these solutions. $\square$

Corollary 3.7. *For the case $K = 1$, the maximal root $\beta_{0,j}$, which governs the tail behavior of the distribution, is always attained for all $x_i = 1$.*

Proof. This proof uses the notation of the proof of Lemma 3.3. Let $\hat{h}_i(\beta_0)$, $i = 1, \ldots, c$ be the functions with all $x_i = 1$. Let $\hat{\beta}_0$ be the root of $\sum_{i=1}^{c} \hat{h}_i(\beta_0) = 0$. Then, for $x_1, \ldots, x_c$ with at least one $x_i = -1$, we have that $\sum_{i=1}^{c} h_i(0) > 0$ and $\sum_{i=1}^{c} h_i(\hat{\beta}_0) < \sum_{i=1}^{c} \hat{h}_i(\hat{\beta}_0) = 0$. Hence, together with Corollary 3.5, we conclude that the root of $\sum_{i=1}^{c} h_i(\beta_0) = 0$ is in $(0, \hat{\beta}_0)$. $\square$

Let us show how to obtain the equilibrium distribution for an example.

Example 3.8. Consider Example 1.3 with $K = 1$ and 2 heterogeneous servers. There are four different combinations of x_1 and x_2. For each combination we can solve (3.8) to obtain one root $\beta_0 \in (0, 1)$. This leads to four roots β_0. For each root and corresponding combination of x_1 and x_2 we obtain β_1 and β_2 from (3.7) to finally find four product forms. The coefficients of the linear combination α_j can be found from the boundary equations and normalization condition. In doing so, we have uniquely determined the equilibrium distribution $p(\mathbf{n})$ as shown in Theorem 3.6.

Processes with larger jumps in the positive n_0-direction ($K \geq 2$) are harder to study than their counterparts $K = 1$. The number of roots of (3.8) may not be the same for each combination of x_i's as is the case for the Markov processes that are skip-free in the positive n_0-direction. The proposed method might still work for processes with $K \geq 2$, which is supported by numerical experiments (not shown here). However, proving that there are $2^c K$ roots is much more involved. We will show this in the next section, where we study *symmetric* Markov processes with $K \geq 1$.

4 Symmetric processes

Remember that we refer by symmetric processes to the processes described in Section 1.2 with the additional assumption that the rates in the (n_0, n_i)-planes are identical. The queueing systems mentioned in Section 1.1 give rise to symmetric processes when the servers are identical (homogeneous). We exploit this symmetry to prove that (3.8) has $2^c K$ roots with $|\beta_0| < 1$. For symmetric processes we omit the indices i of the rates $a_{k,i}, \ldots, d_{k,i}$.

With $\eta_c := -\frac{1}{c} \sum_{i=1}^c x_i \in [-1, 1]$, equation (3.8) simplifies to

$$\eta_c \sqrt{F(\beta_0)^2 + 4A(\beta_0)D(\beta_0)} = B(\beta_0) + C(\beta_0) - \beta_0^K (A(1) + B(1) + C(1) + D(1)). \qquad (4.1)$$

Note that if there exists a root β_0 of (4.1) with $|\beta_0| < 1$ such that

$$0 = F(\beta_0)^2 + 4A(\beta_0)D(\beta_0), \qquad (4.2)$$

then this root β_0 is a solution of (4.1) for all combinations of x_i's. This is not desired, because if (4.2) holds for some β_0, the corresponding β_i, $i = 1, \ldots, c$, found from (3.7) do not depend on x_i and thus we obtain *dependent* solutions. We therefore make the following assumption.

Assumption 4.1. For all roots β_0 of (4.1) with $|\beta_0| < 1$, assume that $F(\beta_0)^2 + 4A(\beta_0)D(\beta_0) \neq 0$.

Note that Assumption 4.1 is satisfied for the case $K = 1$, because $\beta_0 \in (0, 1)$ and thus the right-hand side of (4.2) is strictly positive, due to Assumptions 2.1(iv) and (v). Using Assumption 4.1 we should mention that the linear independence of the product forms derived in Lemma 3.4 is also applicable for $K \geq 2$, where the number of products is now $2^c K$. In the following example we show that for certain models Assumption 4.1 always holds.

Example 4.2. Consider Example 1.3, but now with $\mu := \mu_{1,i} = \mu_{2,i}$, i.e. parallel identical servers with Erlang-2 service times. The right-hand side of (4.2) then equals $4\mu^2 \beta_0^{2K+1}$, thus (4.2) does not hold for feasible values of β_0 as $\beta_0 = 0$ is not a solution of (4.1).

However, we can construct situations for which Assumption 4.1 is violated. We do so in the following example.

Example 4.3. Consider again Example 1.3, now with 2 parallel identical servers, where batches of size two arrive with rate λ and there are no other arrivals. The ergodicity condition for this system is given by $\lambda < \frac{\mu_1 \mu_2}{\mu_1 + \mu_2}$. Denote a root of (4.2) by $\beta_{0,a}$ and a root for which the right-hand side of (4.1) vanishes by $\beta_{0,b}$, so that

$$\beta_{0,a} = -\frac{(\mu_1 - \mu_2)^2}{4\mu_1 \mu_2}, \quad \beta_{0,b} = -\sqrt{\frac{\lambda}{\lambda + \mu_1 + \mu_2}}. \qquad (4.3)$$

For $\lambda = 1$, $\mu_1 = 6$ and $\mu_2 = 2$, we find that $\beta_{0,a} = \beta_{0,b} = -1/3$ and the ergodicity condition holds. Consequently, we find $\beta_0 = -1/3$ as a root of (4.1) for all combinations of x_i's, and by (3.7) we find that the products corresponding to this root are identical and thus dependent. It is interesting to see what type of products we find from (4.1) for these parameter settings. This is summarized in Table 1. We see that there are only six different products. This is problematic as we require $2^c K = 8$ solutions. A possible solution to obtain the required number of independent solutions, is to try and find products of the form $(an_0 + b)\beta_0^{n_0} \beta_1^{n_1} \beta_2^{n_2}$ with $\beta_0 = -1/3$, and to select a and b values such that the inner conditions (3.1) are satisfied. We leave this for future research.

x_1	x_2	β_0	β_1	β_2
1	1	0.74	3.77	3.77
		-0.33	-3	-3
-1	1	0.33	-1.24	7.24
		-0.33	-3	-3
1	-1	0.33	7.24	-1.24
		-0.33	-3	-3
-1	-1	-0.32	-2.48	-2.48
		0.26	-1.28	-1.28
		-0.33	-3	-3

Table 1: Solutions of (4.1) and corresponding β_i's for Example 4.3.

Recall that in (3.7) we need $D(\beta_0) \neq 0$. As the roots are no longer restricted to the interval $(0,1)$, but more generally should lie within the unit circle, it is possible that $D(\beta_0) = 0$ for some root β_0. We therefore make the following assumption.

Assumption 4.4. For all roots β_0 of (4.1) with $|\beta_0| < 1$, assume that $D(\beta_0) \neq 0$.

We now take squares of both sides of (4.1), thus eliminating the square root, and we will prove, under the ergodicity condition (2.6), the following result.

Lemma 4.5. Let $\eta_c := -\frac{1}{c} \sum_{i=1}^{c} x_i$. For each combination of x_i's, the equation

$$\eta_c^2 \big(F(\beta_0)^2 + 4A(\beta_0)D(\beta_0) \big) = \big(B(\beta_0) + C(\beta_0) - \beta_0^K (A(1) + B(1) + C(1) + D(1)) \big)^2, \quad (4.4)$$

has $2K$ solutions β_0 in the unit circle.

Proof. We bring all terms in (4.4) to one side to obtain

$$\big(B(\beta_0) + C(\beta_0) - \beta_0^K (A(1) + B(1) + C(1) + D(1)) \big)^2 - \eta_c^2 F(\beta_0)^2 - 4\eta_c^2 A(\beta_0)D(\beta_0) = 0. \quad (4.5)$$

To determine the number of roots of (4.5) within the unit circle, we will use Rouché's theorem (see e.g. Theorem 9.2.3 in [13], or more recent work in [2]) twice. For now, let us ignore the $-4\eta_c^2 A(\beta_0)D(\beta_0)$ term. This leaves us with a difference of two squares, which we factor into its two terms. After some rearranging, these two terms can be written as

$$B(\beta_0)(1 - \eta_c) + C(\beta_0)(1 + \eta_c) - \beta_0^K \big((A(1) + B(1))(1 - \eta_c) + (C(1) + D(1))(1 + \eta_c) \big) \quad (4.6)$$

and

$$B(\beta_0)(1 + \eta_c) + C(\beta_0)(1 - \eta_c) - \beta_0^K \big((A(1) + B(1))(1 + \eta_c) + (C(1) + D(1))(1 - \eta_c) \big). \quad (4.7)$$

Let us first focus on (4.6). We define

$$f_1(z) := -z^K \big((A(1) + B(1))(1 - \eta_c) + (C(1) + D(1))(1 + \eta_c) \big), \quad (4.8)$$
$$g_1(z) := B(z)(1 - \eta_c) + C(z)(1 + \eta_c). \quad (4.9)$$

Let $L \subset \mathbb{C}$ denote the closed unit disc and ∂L the unit circle. Clearly, $f_1(z)$ has only one root $z = 0$ in L of multiplicity K. We shall now show that

$$|f_1(z)| > |g_1(z)|, \quad z \in \partial L, \quad (4.10)$$

so that it follows from Rouché's theorem that $f_1(z) + g_1(z)$ also has K zeros in L.

Using that $(1 - \eta_c)$ and $(1 + \eta_c)$ are both nonnegative and evaluating $f_1(z)$ and $g_1(z)$ along ∂L, gives

$$
\begin{aligned}
|f_1(z)| &= |-z^K\big((A(1) + B(1))(1 - \eta_c) + (C(1) + D(1))(1 + \eta_c)\big)| \\
&= (A(1) + B(1))(1 - \eta_c) + (C(1) + D(1))(1 + \eta_c) \\
&> B(1)(1 - \eta_c) + C(1)(1 + \eta_c) = B(|z|)(1 - \eta_c) + C(|z|)(1 + \eta_c) \\
&\geq |B(z)(1 - \eta_c) + C(z)(1 + \eta_c)| = |g_1(z)|.
\end{aligned}
\tag{4.11}
$$

Thus, (4.6) has K roots in L and using a similar argument, (4.7) also has K roots in L. Therefore, the product of (4.6) and (4.7) yields $2K$ solutions within the unit circle.

Now we wish to once again use Rouché's theorem, this time on the complete equation (4.5). Let us define

$$
f_2(z) := \big(B(z) + C(z) - z^K(A(1) + B(1) + C(1) + D(1))\big)^2 - \eta_c^2 F(z)^2, \tag{4.12}
$$
$$
g_2(z) := 4\eta_c^2 A(z)D(z), \tag{4.13}
$$

where $f_2(z)$ is the product of (4.6) and (4.7) and note that (4.5) can be expressed as $f_2(z) - g_2(z) = 0$. We have already determined that $f_2(z)$ has $2K$ zeros in L. We will now work towards the proof that each equation of (4.4) has $2K$ solutions. For readability, we introduce

$$
\tau_1(\eta_c) := (A(1) + B(1))(1 - \eta_c) + (C(1) + D(1))(1 + \eta_c), \tag{4.14}
$$
$$
\tau_2(\eta_c) := (A(1) + B(1))(1 + \eta_c) + (C(1) + D(1))(1 - \eta_c). \tag{4.15}
$$

Let us first establish a relation between $|f_2(z)|$ and $f_2(|z|)$:

$$
\begin{aligned}
|f_2(z)| &= |\big(B(z) + C(z) - z^K(A(1) + B(1) + C(1) + D(1))\big)^2 - \eta_c^2 F(z)^2| \\
&= |B(z)(1 - \eta_c) + C(z)(1 + \eta_c) - z^K \tau_1(\eta_c)| \\
&\quad \times |B(z)(1 + \eta_c) + C(z)(1 - \eta_c) - z^K \tau_2(\eta_c)| \\
&= |-B(z)(1 - \eta_c) - C(z)(1 + \eta_c) + z^K \tau_1(\eta_c)| \\
&\quad \times |-B(z)(1 + \eta_c) - C(z)(1 - \eta_c) + z^K \tau_2(\eta_c)| \\
&\geq \big(-B(|z|)(1 - \eta_c) - C(|z|)(1 + \eta_c) + |z|^K \tau_1(\eta_c)\big) \\
&\quad \times \big(-B(|z|)(1 + \eta_c) - C(|z|)(1 - \eta_c) + |z|^K \tau_2(\eta_c)\big) \\
&= f_2(|z|).
\end{aligned}
\tag{4.16}
$$

Also, because of the triangle inequality, $g_2(|z|) \geq |g_2(z)|$. It thus suffices to show that

$$
f_2(|z|) > g_2(|z|), \quad z \in \partial L, \tag{4.17}
$$

to prove that (4.5) has $2K$ roots for each combination of x_i's.

Let us now evaluate $f_2(|z|)$ along ∂L, so $|z| = 1$, and

$$
\begin{aligned}
f_2(1) &= (-A(1) - D(1))^2 - \eta_c^2(A(1) - D(1))^2 \\
&= A(1)^2 + 2A(1)B(1) + B(1)^2 - \eta_c^2(A(1)^2 - 2A(1)B(1) + D(1)^2) \\
&= (1 - \eta_c^2)(A(1)^2 + D(1)^2) + 2(1 + \eta_c^2)A(1)D(1) \\
&\geq 4\eta_c^2 A(1)D(1) = g_2(1).
\end{aligned}
\tag{4.18}
$$

Equality between $f_2(1)$ and $g_2(1)$ occurs only when $\eta_c^2 = 1$. We can apply Rouché's theorem for all $\eta_c^2 < 1$.

In order to use the theorem when $\eta_c^2 = 1$, we will essentially evaluate $f_2(|z|)$ and $g_2(|z|)$ along the circle $|z| = 1 - \epsilon$. To accomplish this, we use the Taylor expansion, $f(1 - \epsilon) = f(1) - \epsilon f'(1) + o(\epsilon)$. We want to show that $f_2(1 - \epsilon) > g_2(1 - \epsilon)$ for ϵ sufficiently small. As we have already determined that $f_2(1) = g_2(1)$, this leaves us to prove that

$$f_2'(1) < g_2'(1). \tag{4.19}$$

Using $\eta_c^2 = 1$,

$$f_2'(1) = -4A(1)\big(C'(1) - K(C(1) + D(1))\big) - 4D(1)\big(B'(1) - K(A(1) + B(1))\big) \tag{4.20}$$

and

$$g_2'(1) = 4(A(1)D'(1) + D(1)A'(1)). \tag{4.21}$$

From the ergodicity condition (2.6) we then know that $f_2'(1) < g_2'(1)$. Therefore, Rouché's theorem is applicable and (4.5) has $2K$ roots for each combination of x_i's. $\qquad\square$

Remark 4.6. (*Duplicate roots*) We still need to resolve the issue of having twice as many roots as needed, namely $2^{c+1}K$ roots, which is a result of squaring (4.1). For every possible value of η_c, say η_c^*, $-\eta_c^*$ is also one of the values of η_c. Using η_c^* or $-\eta_c^*$ will yield the same roots β_0, thus half of our roots are duplicates. If $\eta_c = 0$ we find $2K$ roots of which half are duplicates as well. This leaves us with $2^c K$ useful roots within the unit circle.

Remark 4.7. (*Number of unique solutions*) The sum $\sum_{i=1}^c x_i$ can take $c + 1$ unique values, and using Lemma 4.5 and Remark 4.6 we see that there are $K(c + 1)$ unique roots β_0. Notice, however, that this still gives $2^c K$ unique solutions in the form of $\beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c}$, as different values of x_i lead to different products, even though $\beta_{0,j}$ is the same. Here we have used Assumption 4.1 to make sure that the coefficient of x_i in (3.7) is nonzero.

Let us now formulate the following theorem.

Theorem 4.8. *For the symmetric multi-dimensional Markov processes with $K \geq 1$, the equilibrium distribution takes the form*

$$p(\mathbf{n}) = \sum_{j=1}^{2^c K} \alpha_j \beta_{0,j}^{n_0} \beta_{1,j}^{n_1} \cdots \beta_{c,j}^{n_c}, \quad \mathbf{n} \in W, \tag{4.22}$$

where $\beta_{0,j}$ are the roots of (4.1) within the unit circle and $\beta_{i,j}$, $i = 1, \ldots, c$ are the corresponding constants found from (3.7). The (possibly complex-valued) constants α_j are determined from the boundary equations and the normalization condition.

Proof. The proof is similar to the proof of Theorem 3.6. Lemma 4.5 together with Remark 4.6 describe that we find exactly $2^c K$ solutions β_0. Using (3.7) we find $2^c K$ products $\beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c}$, $j = 1, \ldots, 2^c K$. These products are absolutely convergent, as $|\beta_0| < 1$, and linearly independent, according to Lemma 3.4. Then, Lemma 3.1 shows that the $2^c K$ linearly independent and absolutely convergent solutions can be used to express the equilibrium distribution as a linear combination of these solutions. $\qquad\square$

5 Aggregated state concept

We now show how the results from the symmetric multi-dimensional Markov process are related to a Markov process on a semi-infinite strip.
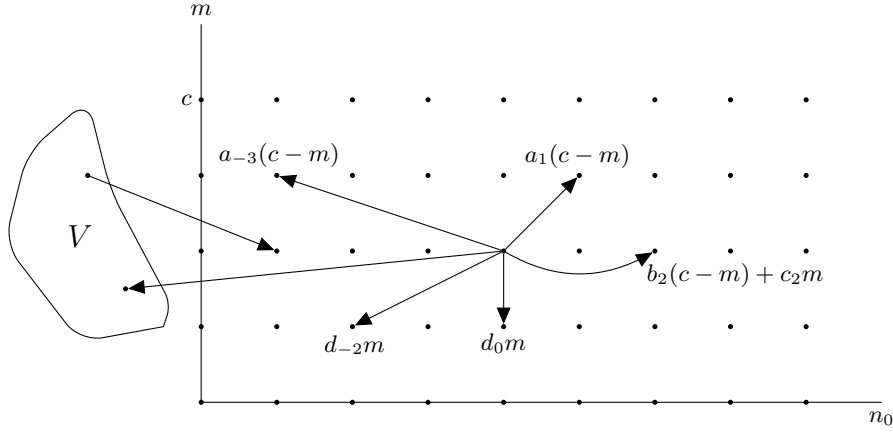
Figure 2: Transition rate diagram for a symmetric Markov process on the aggregated state space.

For symmetric Markov processes, we can exploit the symmetry and transform the Markov process to an aggregated state space characterized by (n_0, m), where $m = \sum_{i=1}^{c} n_i$. This state space is a semi-infinite strip of states. The transition structure induced by the structure of the multi-dimensional Markov process is as follows. For $n_0 \geq K$ and $k = -n_0, -n_0 + 1, \ldots, K$:

- From (n_0, m) to $(n_0 + k, m + 1)$ with rate $a_k(c - m)$.

- From (n_0, m) to $(n_0 + k, m)$ with rate $b_k(c - m) + c_k m$.

- From (n_0, m) to $(n_0 + k, m - 1)$ with rate $d_k m$.

- From (n_0, m) to the states in V with total rate $\sum_{k=-\infty}^{-n_0 - 1} \big((a_k + b_k)(c - m) + (c_k + d_k)m \big)$.

Transitions from (n_0, m) are only possible to states (n, l) with $n \leq n_0 + K$ and $l \in \{m - 1, m, m + 1\}$. Further, the rates a_k and b_k are scaled by a factor $c - m$, and the rates c_k and d_k are scaled by a factor m. Observe the linear structure due to the symmetry in the transitions in the dimensions n_i, $i = 1, \ldots, c$. The transition rate diagram is shown in Figure 2.

Note that the ergodicity condition of the model on the aggregated state space is identical to the one derived for the symmetric multi-dimensional model, i.e. (2.6).

We are interested in the equilibrium distribution of this aggregated state space. We obtain this distribution by using the basis of solutions of the multi-dimensional symmetric Markov process. Translation of the basis solutions $p_j(\mathbf{n}) = \beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c}$ to the aggregated state description yields $p_j(n_0, m) = \beta_{0,j}^{n_0} \omega_j(m)$, where $\omega_j(m)$ is given by

$$\omega_j(m) = \sum_{n_1 + \cdots + n_c = m} \beta_{1,j}^{n_1} \cdots \beta_{c,j}^{n_c}. \tag{5.1}$$

Basis solutions that are equal up to a permutation of the last c factors are translated to exactly the same basis solutions in the new state space description. Hence, we relabel the basis solutions $\beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c}$, $j = 1, \ldots, 2^c K$ such that the first $K(c + 1)$ solutions cannot be obtained from one another by permuting the last c factors. Alternatively, one can only use the $c + 1$ unique values of η_c in (4.1) to obtain $K(c + 1)$ basis solutions. Hence, we obtain the reduced set of solutions $\beta_{0,j}^{n_0} \omega_j(m)$, $j = 1, \ldots, K(c + 1)$. Then, Theorem 5.1 is a translation of Theorem 4.8 in terms of the aggregated state space.

Theorem 5.1. *For symmetric multi-dimensional Markov processes with the aggregated state space* (n_0, m), *where* $m = \sum_{i=1}^{c} n_i$, *the equilibrium distribution is of the form*

$$p(n_0, m) = \sum_{j=1}^{K(c+1)} \gamma_j \beta_{0,j}^{n_0} \omega_j(m), \quad n_0 \in \mathbb{N}_0, \ m = 0, \ldots, c, \tag{5.2}$$

where $\beta_{0,j}$ *are the unique roots of* (4.1) *within the unit circle and* $\omega_j(m)$ *is given by* (5.1). *The (possibly complex-valued) constants* γ_j *are determined from the boundary equations and normalization condition.*

Remark 5.2. (*Generating function technique*) In [3], a generating function technique is used to obtain the equilibrium distribution for the same symmetric class of processes with $K = 1$, where the generating function is over the finite index m. This technique can also be used to obtain the equilibrium distribution for the symmetric case $K \geq 2$, and in fact, leads to equations for β_0 that are the same as (4.1).

6 First passage times

In this section we study, for symmetric processes on the aggregated state space, the time until first passage to the set V, given that at time $t = 0$, the process starts in state (n_0, m) and for $t \geq 0$ transitions in the positive n_0-direction are not possible (so no transitions with rate $a_k, \ldots, d_k, \ k > 0$ for $t \geq 0$). The motivation for studying the first passage time is that for many queueing systems, this time is exactly the waiting time of a customer arriving in state (n_0, m). Note that this equality holds due to the fact that we only consider the FCFS policy for these queueing systems.

Let $F_{n_0, m}(t)$ denote the probability that the first passage time to the set V exceeds t, given that the process starts in state (n_0, m) at time $t = 0$.

Example 6.1. Consider Example 1.3 with c parallel identical servers and batch arrivals of maximum size K. Let W denote the waiting time of a customer. By conditioning on the state seen on arrival and using that Poisson arrivals see time averages (PASTA, see [20]), we obtain

$$\mathbb{P}(W > t) = \sum_{n_0=0}^{\infty} \sum_{m=0}^{c} p(n_0, m) F_{n_0, m}(t) = \sum_{j=1}^{K(c+1)} \gamma_j \sum_{n_0=0}^{\infty} \sum_{m=0}^{c} \beta_{0,j}^{n_0} \omega_j(m) F_{n_0, m}(t) =: F(t). \tag{6.1}$$

Hence, once $F_{n_0, m}(t)$ is known, we can determine the waiting time distribution.

Let level n_0 denote the set of states $\{(n_0, 0), \ldots, (n_0, c)\}$. We introduce a matrix notation $\Lambda_i, \ i = -n_0, -n_0 + 1, \ldots, K$ for the jumps from level n_0 to level $n_0 + i$, where $n_0 \geq 0$. To exemplify, the (a, b)-th entry of Λ_i describes the rate of going from state (n_0, a) to state $(n_0 + i, b)$. By assuming a structure on the matrices $\Lambda_i, \ i = 1, \ldots, K$ we are able to explicitly derive the first passage time to the set V.

Assumption 6.2. Jumps in the positive n_0-direction are of the form

$$\Lambda_i = \lambda_i I, \quad i = 1, \ldots, K, \tag{6.2}$$

for some $\lambda_i \geq 0$ and I the identity matrix.

Note that for queueing systems, $\lambda_i I, \ i = 1, \ldots, K$, models the Poisson arrival rate of a batch of size i. For the examples presented in Section 1.1 the above assumption holds. Note that the assumption of Poisson arrivals is a natural assumption to make, otherwise (6.1) would

not hold. By introducing a supplementary variable we are still able to analyze the waiting time distribution if there are no Poisson arrivals, see Remark 6.5.

Using the above assumption, we are now able to express $F(t)$ explicitly in terms of a finite sum.

Theorem 6.3. *The function $F(t)$ defined by the right-hand side of* (6.1) *is explicitly given by*

$$F(t) = \sum_{j=1}^{K(c+1)} \frac{\gamma_j \omega_j \mathbf{1}}{1 - \beta_{0,j}} \, \mathrm{e}^{t \sum_{i=1}^{K}(1 - \frac{1}{\beta_{0,j}^i})\lambda_i}, \quad t \geq 0, \tag{6.3}$$

with $\omega_j = \begin{pmatrix} \omega_j(0) & \omega_j(1) & \cdots & \omega_j(c) \end{pmatrix}$.

Proof. We first derive a set of differential equations for the probabilities $F_{n_0,m}(t)$. For small $\Delta t \geq 0$ it holds that

$$F_{n_0}(t + \Delta t) = F_{n_0}(t) + \Delta t \sum_{i=0}^{K} \Lambda_i F_{n_0}(t) + \Delta t \sum_{k=1}^{n_0} \Lambda_{-k} F_{n_0-k}(t) + o(\Delta t), \tag{6.4}$$

where $F_{n_0}(t) = \begin{pmatrix} F_{n_0,0}(t) & F_{n_0,1}(t) & \cdots & F_{n_0,c}(t) \end{pmatrix}^T$. Dividing these equations by Δt and letting Δt tend to zero yields the backward Kolmogorov equations

$$F'_{n_0}(t) = \sum_{i=0}^{K} \Lambda_i F_{n_0}(t) + \sum_{k=1}^{n_0} \Lambda_{-k} F_{n_0-k}(t), \quad n_0 \in \mathbb{N}_0 \tag{6.5}$$

with initial condition $F_{n_0}(0) = \mathbf{1}$, where $\mathbf{1}$ is column vector of size $c+1$ with all ones. To solve these differential equations we use Laplace transforms. Let

$$F^*_{n_0}(s) = \int_0^\infty F_{n_0}(t)\mathrm{e}^{-st} \, \mathrm{d}t, \quad s \geq 0. \tag{6.6}$$

Transforming the differential equations (6.5) for $F_{n_0}(t)$ gives

$$\left(sI - \sum_{i=0}^{K} \Lambda_i \right) F^*_{n_0}(s) = \mathbf{1} + \sum_{k=1}^{n_0} \Lambda_{-k} F^*_{n_0-k}(s). \tag{6.7}$$

From these equations, the Laplace transforms $F^*_{n_0}(s)$ can be solved recursively. Guided by Example 6.1, we define

$$F(t) = \sum_{j=1}^{K(c+1)} \gamma_j \omega_j \sum_{n_0=0}^{\infty} \beta_{0,j}^{n_0} F_{n_0}(t), \quad t \geq 0, \tag{6.8}$$

and the Laplace transforms, for $s \geq 0$,

$$G^*_j(s) = \sum_{n_0=0}^{\infty} \beta_{0,j}^{n_0} F^*_{n_0}(s), \quad j = 1, \ldots, K(c+1), \tag{6.9}$$

$$F^*(s) = \int_0^\infty F(t)\mathrm{e}^{-st} \, \mathrm{d}t = \sum_{j=1}^{K(c+1)} \gamma_j \omega_j G^*_j(s). \tag{6.10}$$

Multiplying (6.7) by $\beta_{0,j}^{n_0}\omega_j$ and summing over all n_0 leads to

$$\omega_j\Big(sI - \sum_{i=0}^{K}\Lambda_i\Big)G_j^*(s) = \frac{\omega_j\mathbf{1}}{1-\beta_{0,j}} + \sum_{n_0=0}^{\infty}\sum_{k=1}^{n_0}\beta_{0,j}^{n_0}\omega_j\Lambda_{-k}F_{n_0-k}^*(s)$$

$$= \frac{\omega_j\mathbf{1}}{1-\beta_{0,j}} + \sum_{k=1}^{\infty}\beta_{0,j}^{k}\omega_j\Lambda_{-k}\sum_{l=0}^{\infty}\beta_{0,j}^{l}F_l^*(s). \tag{6.11}$$

Since

$$0 = \sum_{n_0=0}^{\infty}\beta_{0,j}^{n_0}\omega_j\Lambda_{K-n_0} = \omega_j\sum_{n_0=0}^{K}\beta_{0,j}^{n_0}\Lambda_{K-n_0} + \beta_{0,j}^{K}\omega_j\sum_{n_0=1}^{\infty}\beta_{0,j}^{n_0}\Lambda_{-n_0}, \tag{6.12}$$

we get

$$\omega_j\Big(sI - \sum_{i=0}^{K}\Lambda_i\Big)G_j^*(s) = \frac{\omega_j\mathbf{1}}{1-\beta_{0,j}} - \omega_j\sum_{i=0}^{K}\frac{1}{\beta_{0,j}^{i}}\Lambda_i G_j^*(s), \tag{6.13}$$

and thus

$$\omega_j\Big(sI - \sum_{i=1}^{K}\Big(1 - \frac{1}{\beta_{0,j}^{i}}\Big)\Lambda_i\Big)G_j^*(s) = \frac{\omega_j\mathbf{1}}{1-\beta_{0,j}}. \tag{6.14}$$

We now use Assumption 6.2 to simplify the above expression as

$$\omega_j G_j^*(s) = \frac{\omega_j\mathbf{1}}{1-\beta_{0,j}}\Big(s - \sum_{i=1}^{K}\Big(1 - \frac{1}{\beta_{0,j}^{i}}\Big)\lambda_i\Big)^{-1}. \tag{6.15}$$

Multiplying by γ_j and summing over all j finally yields

$$F^*(s) = \sum_{j=1}^{K(c+1)}\frac{\gamma_j\omega_j\mathbf{1}}{1-\beta_{0,j}}\Big(s - \sum_{i=1}^{K}\Big(1 - \frac{1}{\beta_{0,j}^{i}}\Big)\lambda_i\Big)^{-1}. \tag{6.16}$$

The inverse of the Laplace transform is readily obtained as (6.3). $\qquad\square$

Example 6.4. Again consider Example 1.3 with parallel identical servers with $K = 2$. Let λ_1 be the arrival rate of single jobs and λ_2 the arrival rate of jobs in batches of size two. Since the first passage time to the set V is equal to the waiting time, we have that $F(t) = \mathbb{P}(W > t)$. From Theorem 6.3 we may immediately conclude that the waiting time distribution is the mixture of exponentials

$$\mathbb{P}(W > t) = \sum_{j=1}^{K(c+1)}\frac{\gamma_j\omega_j\mathbf{1}}{1-\beta_{0,j}}\,\mathrm{e}^{t\Big(\big(1-\frac{1}{\beta_{0,j}}\big)\lambda_1 + \big(1-\frac{1}{\beta_{0,j}^{2}}\big)\lambda_2\Big)}, \quad t \geq 0. \tag{6.17}$$

Remark 6.5. ($\sum IPP/M/1$ *queuing systems*) We consider a single exponential server fed by a superposition of independent, interrupted Poisson processes with rate λ, labeled $\sum IPP/M/1$. For an $\sum IPP/M/1$ queueing system Assumption 2.1 does not hold. However, we can obtain the waiting time distribution explicitly. This suggest that the class of models with an explicit first passage time distribution in the form of mixtures of exponentials can possibly be extended beyond the class satisfying Assumption 2.1. Let us denote the probability that an arriving job sees the system in state (n_0, m) by $a(n_0, m)$, where m now reflects the number of 'active' Poisson processes. For the $\sum IPP/M/1$ queuing system this is equal to

$$a(n_0, m) = Np(n_0, m)m\lambda, \tag{6.18}$$

where N is a normalizing constant, given by $N^{-1} = \sum_{n,k} p(n,k)k\lambda$. Hence, by Theorem 5.1, we get that

$$a(n_0, m) = N \sum_{j=1}^{K(c+1)} \gamma_j \beta_{0,j}^{n_0} \omega_j(m) m\lambda, \tag{6.19}$$

where

$$N^{-1} = \sum_{j=1}^{K(c+1)} \sum_{k=0}^{c} \frac{\omega_j(k)k\lambda}{1 - \beta_{0,j}}. \tag{6.20}$$

The waiting time distribution can then be expressed as

$$\mathbb{P}(W > t) = \sum_{n_0=1}^{\infty} \sum_{m=0}^{c} a(n_0, m) F_{n_0, m}(t), \tag{6.21}$$

where $F_{n_0, m}(t)$ is an Erlang-n_0 distribution with parameter μ. Note that for this specific model, $F_{n_0, m}(t)$ does not depend on m and the set V is level 0. Substituting expression (6.19) for $a(n_0, m)$ and the expression for the Erlang-distribution then yields

$$\mathbb{P}(W > t) = N \sum_{j=1}^{K(c+1)} \gamma_j \sum_{m=0}^{c} \omega_j(m) m\lambda \frac{\beta_{0,j} e^{-(1-\beta_{0,j})\mu t}}{1 - \beta_{0,j}}, \tag{6.22}$$

which is a sum of exponentials, similar to (6.3).

7 Conclusion

We have introduced and analyzed a class of structured multi-dimensional Markov processes with one infinite dimension. By employing a separation of variables technique we showed that the equilibrium distribution can be represented as a linear combination of product forms (in terms of eigenvalues, and eigenvectors of the finite dimensions). We have presented an efficient approach to compute the product forms. Using the separation of variables technique, we were able to decompose the single equation for $2^c K$ eigenvalues that arises from the classical spectral expansion method to 2^c equations for fewer eigenvalues per equation, yet still obtaining $2^c K$ eigenvalues. The eigenvectors are then found from the corresponding eigenvalues. The weights of the linear combination were determined by a system of linear equations originating from the boundary equations and normalization condition. For symmetric processes we were able to aggregate the Markov process in terms of a process on a semi-infinite strip of states. The equilibrium distribution of this aggregated process can be expressed in terms of the parameters of the equilibrium distribution of the multi-dimensional Markov process. We have also shown that for these symmetric processes the first passage time to the set V is a mixture of exponentials. For relevant queueing systems this gives the waiting time distribution.

There are at least two extensions of the class presented in this paper that seem worthwhile to explore. The first extension concerns the case of nonsymmetric processes with $K \geq 2$. It is presently unclear how to prove that one obtains $2^c K$ roots from 2^c equations. The other extension is to allow larger finite dimensions, i.e., letting $n_i \in \{0, 1, \ldots, r\}$, $i = 1, \ldots, c$ with $r > 1$. Due to tractability of the results it was opted not to do this for this paper. We believe that similar results can be obtained for these processes, however, one might need to move away from the power terms $\beta_i^{n_i}$ of the finite dimensions, or impose additional structure, such as, for example, present in the $E_k/E_r/c$ queue [5].

19

Acknowledgement

References

[1] I.J.B.F. Adan, T. De Kok, and J.A.C. Resing. A multi-server queueing model with locking. *European Journal of Operational Research*, 116(2):16–26, 2000.

[2] I.J.B.F. Adan, J.S.H. Van Leeuwaarden, and E.M.M. Winands. On the application of Rouché's theorem in queueing theory. *Operations Research Letters*, 34(3):355–360, 2006.

[3] I.J.B.F. Adan and J.A.C. Resing. A class of Markov processes on a semi-infinite strip. In B. Plateau, W.J. Stewart, and M. Silva, editors, *Numerical Solution of Markov Chains*, pages 41–57, 1999.

[4] I.J.B.F. Adan and J.A.C. Resing. Multi-server batch-service systems. *Statistica Neerlandica*, 54(2):202–220, 2000.

[5] I.J.B.F. Adan, W.A. Van De Waarsenburg, and J. Wessels. Analyzing $E_k/E_r/c$ queues. *European Journal of Operational Research*, 92(1):112–124, 1996.

[6] J.R. Artalejo, A. Economou, and M. Lopez-Herrero. Analysis of a multiserver queue with setup times. *Queueing Systems*, 51(1-2):53–76, 2005.

[7] D. Bertsimas and X.A. Papaconstantinou. Analysis of the stationary $E_k/C_2/s$ queueing system. *European Journal of Operational Research*, 37(2):272–287, 1988.

[8] J.A. Buzacott and J.G. Shanthikumar. *Stochastic Models of Manufacturing Systems*, volume 4. Prentice Hall Englewood Cliffs, NJ, 1993.

[9] Bertsimas D. An exact FCFS waiting time analysis for a general class of $G/G/s$ queueing systems. *Queueing Systems*, 3(4):305–320, 1988.

[10] F.G. Foster. On the stochastic matrices associated with certain queuing processes. *The Annals of Mathematical Statistics*, 24(3):355–360, 1953.

[11] A. Gandhi, M. Harchol-Balter, and I.J.B.F. Adan. Server farms with setup costs. *Performance Evaluation*, 67(11):1123–1138, 2010.

[12] P.R. Garabedian. *Partial Differential Equations*. Wiley, 1967.

[13] E. Hille. *Analytic Function Theory*, volume 1. Ginn, 1959.

[14] I. Mitrani and R. Chakka. Spectral expansion solution for a class of Markov models: Application and comparison with the matrix-geometric method. *Performance Evaluation*, 23(3):241–260, 1995.

[15] I.L. Mitrany and B. Avi-Itzhak. A many-server queue with service interruptions. *Operations Research*, 16(3):628–638, 1968.

[16] M.F. Neuts. *Matrix-Geometric Solutions in Stochastic Models: An Algorithmic Approach*. Courier Dover Publications, 1981.

[17] S. Shapiro. The M-server queue with Poisson input and Gamma-distributed service of order two. *Operations Research*, 14(4):685–694, 1966.

[18] J.H.A. De Smit. A numerical solution for the multi-server queue with hyper-exponential service times. *Operations Research Letters*, 2(5):217–224, 1983.

[19] J.H.A. De Smit. The queue $GI/M/s$ with customers of different types or the queue $GI/H_m/s$. *Advances in Applied Probability*, 15(2):392–419, 1983.

[20] R.W. Wolff. Poisson arrivals see time averages. *Operations Research*, 30(2):223–231, 1982.