Working Paper

Analysing $E_k|E_r|c$ Queues

I.J.B.F. Adan
W.A. van de Waarsenburg
J. Wessels

WP-92-047
July 1992

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Analysing $E_k|E_r|c$ Queues

I.J.B.F. Adan
W.A. van de Waarsenburg
J. Wessels

WP-92-047
July 1992

IIASA International Institute for Applied Systems Analysis □ A-2361 Laxenburg □ Austria
Telephone: +43 2236 715210 □ Telex: 079 137 iiasa a □ Telefax: +43 2236 71313

Analysing $E_k|E_r|c$ Queues

I.J.B.F. Adan*
W.A. van de Waarsenburg*
J. Wessels†*

Abstract

In this paper we study a system consisting of c parallel identical servers and a common queue. The service times are Erlang-r distributed and the interarrival times are Erlang-k distributed. The service discipline is first-come first-served. The waiting process may be characterised by $(n_{-1}, n_0, n_1, \ldots, n_c)$ where n_{-1} represents the number of remaining arrival stages, n_0 the number of waiting jobs and n_i, $i = 1, \ldots, c$, the number of remaining service stages for server i. Bertsimas has proved that the equilibrium probability for a saturated state (all $n_i > 0, i = 1, \ldots, c$) can be written as a linear combination of geometric terms with n_0 as exponent. In the present paper it is shown that the coefficients also have a geometric form with respect to n_{-1}, n_1, $\ldots$, n_c. It is also shown how the factors may be found efficiently. The present paper uses a direct approach for solving the equilibrium equations rather than a generating function approach as Bertsimas does. The direct approach has been inspired by previous work of two of the authors on the shortest queue problem in particular and the two-dimensional random walk more generally. Although the paper extends results of Bertsimas it is self-contained.

1 Introduction

The $E_k|E_r|c$ queueing system is a typical example of an elegantly modelled and seemingly simple system which, nevertheless, has never been analyzed satisfactorily. Apparently, the system behaviour is more complex than the simple formulation suggests. A generating-function approach may be found in Poyntz and Jackson [9], where it has been worked out a little bit for $c = 2$ and $c = 3$. This approach however, doesn't yield exact expressions for the equilibrium probabilities, but merely leads to involving numerical procedures. In the present paper a rather direct approach (without generating functions) will be presented for solving the equilibrium equations. The inspiration for using this approach came from its relative success in solving the equilibrium equations for the shortest queue problem (cf. [1], [2]) and its generalization to the two-dimensional random walk on the integer grid in the first quadrant (cf. [1], [3]). In these cases the essential idea was to avoid the integration of the equations for boundary states with the equations for states in the inner region by constructing the usual functional equations for the generating functions. This integration could be avoided by first constructing a sufficiently rich solution base for the equilibrium equations in inner points and then use this base for finding a linear combination which also satisfies the equations for the states on the boundary. In [1], [2] and [3] the linear combination is found by a compensation procedure.

To some extent, a similar approach is followed by Bertsimas [4] for the $C_k|C_r|c$ problem, which is more general that the $E_k|E_r|c$ problem. Bertsimas proves in this way that the equilibrium

*Eindhoven University of Technology, Department of Mathematics and Computer Science, p.o. Box 513, 5600 MB - Eindhoven, The Netherlands.
†International Institute for Applied Systems Analysis, A-2361 Laxenburg, Austria.

probabilities for saturated states can be written as a linear combination of terms which are geometric in the number of waiting customers. In the present paper it is shown that Bertsimas' representation can be refined considerably for the $E_k|E_r|c$ problem. This refined representation gives more insight in the behaviour of the processes and is also very useful for computational purposes. The refinement is obtained by a complete direct analysis of the equilibrium equations, contrary to the approach of Bertsimas, who also separates the equilibrium equations for boundary and inner states, but still treats the latter by generating functions. Extensions of the present approach to more complicated distributions are possible, however, in such cases the results are correspondingly more complex. Therefore, it seems sensible to introduce our approach for $E_k|E_r|c$ problems separately.

This paper is organised as follows. In section 2 the model is introduced and the relevant equilibrium equations are formulated. In section 3 the products of powers are constructed satisfying the equations in the states with all servers busy. Section 4 presents the main theorem, stating that the equilibrium probabilities can be expressed as a linear combination of products of powers. In section 5 the symmetry of the model is exploited and in section 6 two special cases are worked out in more detail. The final section is devoted to comments and conclusions. Two technical lemmas are proved in the Appendices A and B.

2 Model and equations

Consider a system with c parallel identical servers and a common queue. The service times are Erlang-r distributed with mean r/μ and the interarrival times are Erlang-k distributed with mean k/λ. An Erlang-l distributed time with mean l/ν is interpreted as to be composed of l stages, each with a negative-exponentially distributed length (parameter value ν). The service discipline is first-come first-served. An arriving job, who finds any free servers, chooses each free server with equal probability. This queueing system can be modelled as a continuous-time Markov process with a state space consisting of the vectors $\overline{n} = (n_{-1}, n_0, n_1, \ldots, n_c)$, where n_{-1} is the number of remaining arrival stages, n_0 is the number of waiting jobs and n_i is the number of remaining service stages for server i, $i = 1, \ldots, c$. Below we formulate the equilibrium equations for the states $\overline{n}$ with all servers busy, i.e. $n_i > 0$ for $i = 1, \ldots, c$. Let $\overline{e}_i = (0, \ldots, 0, 1, 0, \ldots, 0)$ have $c+2$ components, with the one at the same place as n_i in $\overline{n}$ and let $\nu_{\overline{n}}$ denote $\{i \geq 1 \mid n_i = r\}$. By equating the rate out of and the rate into state $\overline{n}$ we obtain

$$\begin{aligned} p(\overline{n})(\lambda + c\mu) &= p(\overline{n} + \overline{e}_{-1})\lambda + \sum_{i \in \nu_{\overline{n}}} p(\overline{n} + \overline{e}_0 - (r-1)\overline{e}_i)\mu + \\ &\quad + \sum_{i \notin \nu_{\overline{n}}} p(\overline{n} + \overline{e}_i)\mu, \qquad n_{-1} < k, n_i > 0, i = 1, \ldots, c; \end{aligned} \tag{1}$$

$$\begin{aligned} p(\overline{n})(\lambda + c\mu) &= p(\overline{n} - \overline{e}_0 - (k-1)\overline{e}_{-1})\lambda + \sum_{i \in \nu_{\overline{n}}} p(\overline{n} + \overline{e}_0 - (r-1)\overline{e}_i)\mu + \\ &\quad + \sum_{i \notin \nu_{\overline{n}}} p(\overline{n} + \overline{e}_i)\mu, \qquad n_{-1} = k, n_0 > 0, n_i > 0, i = 1, \ldots, c; \end{aligned} \tag{2}$$

$$\begin{aligned} p(\overline{n})(\lambda + c\mu) &= \sum_{i \in \nu_{\overline{n}}} p(\overline{n} - (k-1)\overline{e}_{-1} - r\overline{e}_i)\lambda + \sum_{i \in \nu_{\overline{n}}} p(\overline{n} + \overline{e}_0 - (r-1)\overline{e}_i)\mu + \\ &\quad + \sum_{i \notin \nu_{\overline{n}}} p(\overline{n} + \overline{e}_i)\mu, \qquad n_{-1} = k, n_0 = 0, n_i > 0, i = 1, \ldots, c. \end{aligned} \tag{3}$$

Bertsimas essentially proves in [4] that $p(\overline{n})$ may be written as

$$p(\overline{n}) = \sum_{j=1}^{a(c,r)} D_{n_{-1},j} R_{(n_1,\ldots,n_c),j} w_j^{n_0} \ , \qquad n_i > 0, i = 1, \ldots, c$$

where for each j the factor w_j is characterized by a system of nonlinear equations involving the Laplace transforms of the interarrival and service time distributions and the coefficients $D_{n_{-1},j}$ and $R_{(n_1,\ldots,n_c),j}$ have to be solved from a system of k respectively $a(c,r) = \binom{c+r-1}{r-1}$ linear homogeneous equations. In fact, he uses a slightly aggregated state concept, which exploits the fact that servers are identical.

In section 3 we will develop a more detailed representation in which also the $D-$ and $R-$ factors are replaced by geometric forms. Although we assume to have identical servers, we don't use this assumption explicitly and the results of sections 3 and 4 can easily be extended to the case of non-identical servers.

3 Analysis of the equations for nonboundary states

The geometric forms we will investigate as solutions for (1) and (2) are $\beta_{-1}^{n_{-1}} \beta_0^{n_0} \ldots \beta_c^{n_c}$. We will first characterize the products which satisfy (1), (2) and then we will use these products to construct a linear combination also satisfying the boundary conditions. The boundary conditions are formed by (3) and by the equilibrium equations for states with $n_i = 0$ for some $i \in \{1, \ldots, c\}$.

Inserting $p(\overline{n}) = \beta_{-1}^{n_{-1}} \beta_0^{n_0} \ldots \beta_c^{n_c}$ in (1)-(2) and then dividing both sides of the resulting equations by the common powers leads to the following characterization:

Lemma 1: *The product $\beta_{-1}^{n_{-1}} \beta_0^{n_0} \ldots \beta_c^{n_c}$ satisfies (1)-(2) iff $\beta_{-1}, \beta_0, \ldots, \beta_c$ satisfy*

$$\lambda + c\mu = \beta_{-1}\lambda + \sum_{i \in \mathcal{V}} \frac{\beta_0 \mu}{\beta_i^{r-1}} + \sum_{i \notin \mathcal{V}} \beta_i \mu \tag{4}$$

$$\lambda + c\mu = \frac{\beta_{-1}\lambda}{\beta_0 \beta_{-1}^k} + \sum_{i \in \mathcal{V}} \frac{\beta_0 \mu}{\beta_i^{r-1}} + \sum_{i \notin \mathcal{V}} \beta_i \mu, \tag{5}$$

for each $\mathcal{V} \subset \{1, 2, \ldots, c\}$.

This lemma states that the parameters $\beta_{-1}, \beta_0, \ldots, \beta_c$ are characterised by a set of 2^{c+1} equations. Luckily, it can be shown that most of these equations can be expressed by a linear combination of a set of $c+2$ *basic equations*, which are formulated below.

Lemma 2: *The product $\beta_{-1}^{n_{-1}} \beta_0^{n_0} \ldots \beta_c^{n_c}$ satisfies (1)-(2) iff $\beta_{-1}, \beta_0, \ldots, \beta_c$ satisfy*

$$\lambda + c\mu = \beta_{-1}\lambda + \sum_{i=1}^{c} \beta_i \mu, \tag{6}$$

$$\lambda + c\mu = \beta_{-1}\lambda + \frac{\beta_0 \mu}{\beta_j^{r-1}} + \sum_{\substack{i=1 \\ i \neq j}}^{c} \beta_i \mu, \qquad j = 1, \ldots, c, \tag{7}$$

$$\lambda + c\mu = \frac{\beta_{-1}\lambda}{\beta_0\beta_{-1}^k} + \sum_{i=1}^{c}\beta_i\mu. \tag{8}$$

Proof: The equations (6)-(8) are a subset of the equations (4)-(5). So it remains to prove that each equation in the set (4)-(5) can be found as a linear combination of the equations (6)-(8). Let $\mathcal{V} \subset \{1,\ldots,c\}$. Then equation (4) is obtained by addition of equation (7) for all $j \in \mathcal{V}$ and then subtracting $|\mathcal{V}| - 1$ times equation (6). Equation (5) follows by addition of equation (7) for all $j \in \mathcal{V}$ and equation (8) and then subtracting $|\mathcal{V}|$ times equation (6). □

We now solve the equations (6)-(8). Subtracting (6) from (8) yields that $\beta_j^r = \beta_0$ for $j = 1,\ldots,c$. Hence, by introducing the parameters x_i satisfying $x_i^r = 1$ for $i = 1,\ldots,c$, we may write

$$\beta_0 = y^r, \quad \beta_i = x_i y, \quad i = 1,\ldots,c, \tag{9}$$

for some y. Subtracting (6) from (8) yields that $\beta_{-1}^k = 1/\beta_0$ and thus $\beta_{-1}^k = 1/y^r$ by (9). Hence, if we set

$$y = \alpha^k, \tag{10}$$

then we obtain that $\beta_{-1}^k = 1/\alpha^{rk}$. So, by introducing x_{-1} satisfying $x_{-1}^k = 1$, we have

$$\beta_{-1} = x_{-1}/\alpha^r. \tag{11}$$

Insertion of (9)-(11) in equation (6) leads to the following equation for α:

$$\alpha^r(\lambda + c\mu) = x_{-1}\lambda + \alpha^{r+k}\sum_{i=1}^{c} x_i\mu. \tag{12}$$

On the other hand, it is easily verified that for any root α of this equation the product $\beta_{-1}^{n_{-1}}\beta_0^{n_0}\ldots\beta_c^{n_c}$ with $\beta_{-1},\beta_0,\ldots,\beta_c$ satisfying the specifications (9)-(11), is a solution of the equations (6)-(8). These findings are summarized in the following lemma.

Lemma 3: *The product $\beta_{-1}^{n_{-1}}\beta_0^{n_0}\ldots\beta_c^{n_c}$ satisfies (1)-(2) iff $\beta_{-1},\beta_0,\ldots,\beta_c$ satisfy*

$$\beta_{-1} = x_{-1}/\alpha^r, \quad \beta_0 = \alpha^{kr}, \quad \beta_i = x_i\alpha^k, \quad i = 1,\ldots,c,$$

where $x_{-1}^k = x_i^r = 1$ for $i = 1,\ldots,c$ and α is a root of equation (12).

This lemma characterizes the set of products $\beta_{-1}^{n_{-1}}\beta_0^{n_0}\ldots\beta_c^{n_c}$ satisfying (1)-(2). Clearly, only products which can be normalised, i.e. whose sum over all states $\overline{n}$ converges absolutely, are useful. This implies that $|\beta_0| < 1$, or equivalently $|\alpha| < 1$. The next lemma, which is proved in appendix A, states how many roots of equation (12) lay inside the unit circle. The condition in that lemma states that the offered workload per unit time may not exceed the maximal service capacity.

Lemma 4: *Provided $\frac{\lambda r}{k\mu} < c$, equation (12) has exactly r simple roots inside the unit circle for each set of parameters x_i satisfying $x_{-1}^k = x_i^r = 1$ for $i = 1,\ldots,c$.*

We henceforth assume that the utilisation condition in lemma 4 is satisfied.

Assumption: $\frac{\lambda r}{k\mu} < c$.

According to the lemmas 3 and 4, there are r products $\beta_{-1}^{n-1}\beta_0^{n_0}\dots\beta_c^{n_c}$ with $|\beta_0| < 1$ satisfying (1)-(2) for each feasible choice of $x_{-1}, x_1, \dots, x_c$. So we find rkr^c products. However, below we show that some of them are identical.

We may write

$$r = \rho d, \quad k = \kappa d,$$

where $d = \gcd(r, k)$ and $\gcd(\rho, \kappa) = 1$. For each α satisfying (12) it follows that $u\alpha$ with $u^d = 1$ also satisfies (12) and both roots lead to the same products, since the factors β_i depend only upon α^d. Therefore we set $\gamma = \alpha^d$. Then equation (12) reduces to

$$\gamma^\rho(\lambda + c\mu) = x_{-1}\lambda + \gamma^{\rho+\kappa}\sum_{i=1}^{c} x_i\mu. \tag{13}$$

and lemma 3 can be restated as follows.

Lemma 5: *The product $\beta_{-1}^{n-1}\beta_0^{n_0}\dots\beta_c^{n_c}$ satisfies (1)-(2) iff $\beta_{-1}, \beta_0, \dots, \beta_c$ satisfy*

$$\beta_{-1} = x_{-1}/\gamma^\rho, \quad \beta_0 = \gamma^{\kappa r}, \quad \beta_i = x_i\gamma^\kappa, \quad i = 1, \dots, c,$$

where $x_{-1}^k = x_i^r = 1$ for $i = 1, \dots, c$ and γ is a root of equation (13).

For two roots γ_1 and γ_2 of (13) inside the unit circle, it is easily verified that $\gamma_1^\kappa \neq \gamma_2^\kappa$ (see also lemma 7), and so γ_1 and γ_2 lead to different values for β_i. Hence, from lemma 4 (with r, k replaced by ρ, κ) it follows that for each feasible choice of the parameters x_i there are ρ products $\beta_{-1}^{n-1}\beta_0^{n_0}\dots\beta_c^{n_c}$ with $|\beta_0| < 1$ satisfying (1)-(2). So, by letting run the parameters x_i through all feasible values, we find ρkr^c products. However, there still are duplicates among these products, since the two sets of parameters $x_{-1}/u^{\rho/\kappa}, ux_1, \dots, ux_c$ with $u^r = 1$ and $vx_{-1}, x_1, \dots, x_c$ with $v^\kappa = 1$ lead to exactly the same products as the original set $x_{-1}, x_1, \dots, x_c$. This follows by observing that if γ is a root of (13) for the original parameter set, then $\gamma/u^{1/\kappa}$ and $w\gamma$ with w satisfying $w^\kappa = 1$ and $w^\rho = v$, are roots of (13) for the first, respectively second parameter set, and the three roots lead to the same product. Hence, there are $r\kappa$ copies of each product. To avoid these copies we arbitrarily decide to restrict the feasible values for x_1 to $x_1 = 1$ and the feasible values for x_{-1} to the first d roots of the equation $x_{-1}^k = 1$, i.e. $x_{-1} = e^{\frac{2\pi i}{k}n}$ for some $n = 0, 1, \dots, d-1$. Now, by letting run the parameters x_i through this restricted set of feasible values, we find $\rho dr^{c-1} = r^c$ products and it is readily verified that there are no duplicates among them. We label these products $\beta_{-1,j}^{n-1}\beta_{0,j}^{n_0}\dots\beta_{c,j}^{n_c}$, $j = 1, 2, \dots, r^c$. In this way we have characterised the set of product forms which satisfy the equilibrium equations (1) and (2). In the next section, we will prove our main result which states that $p(\overline{n})$ can be expressed as a linear combination of the products $\beta_{-1,j}^{n-1}\beta_{0,j}^{n_0}\dots\beta_{c,j}^{n_c}$, $j = 1, 2, \dots, r^c$ for states $\overline{n}$ with all servers busy.

4 Satisfying the boundary conditions

Theorem 1: *For all states $\overline{n}$ with $n_i > 0$ for $i = 1, \dots, c$, it holds that*

$$p(\overline{n}) = \sum_{j=1}^{r^c} a_j\beta_{-1,j}^{n-1}\beta_{0,j}^{n_0}\dots\beta_{c,j}^{n_c} \tag{14}$$

for suitably chosen coefficients a_j.

Proof: For each choice of the coefficients a_j the linear combination

$$p(\overline{n}) = \sum_{j=1}^{r^c} a_j \beta_{-1,j}^{n_{-1}} \beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c}, \qquad n_i > 0, i = 1, \ldots, c$$

satisfies the equations (1)-(2). The remaining equations are the equations (3) and the equilibrium equations for states with at least one server free. These equations form a linear, homogeneous system for the unknowns a_j and the unknown quantities $p(\overline{n})$ in states with at least one server free. The number of equations is equal to the number of unknowns. Hence, by first omitting the equation in $\overline{0}$, the reduced system has a nonnull solution. The equation in $\overline{0}$ is automatically satisfied, since inserting the solution $p(\overline{n})$ in the equations in states $\overline{n} \neq \overline{0}$ and then summing over these equations and changing summations exactly yields the desired equation. Changing summations is allowed, since the sum of $p(\overline{n})$ over all states absolutely converges. This follows from the fact that $|\beta_{0,j}| < 1$ for all j. Hence, $p(\overline{n})$ is an absolutely convergent solution of all equilibrium equations. It remains to show that $p(\overline{n})$ is a nonnull-solution. This follows from the next result, which is proved in the appendix.

Lemma 6: *The products $\beta_{-1,j}^{n_{-1}} \beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c}$, $j = 1, \ldots, r^c$ are linearly independent on the set of states $\overline{n}$ with all servers busy, i.e.*

$$\sum_{j=1}^{r^c} a_j \beta_{-1,j}^{n_{-1}} \beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c} = 0 \qquad (0 < n_{-1} \leq k, n_0 \geq 0, 0 < n_i \leq r, i = 1, \ldots, c) \tag{15}$$

iff $a_j \equiv 0$.

From a result of Foster [5] we may now conclude that the Markov process is ergodic and normalization of the $p(\overline{n})$ produces the equilibrium probabilities. □

5 Exploiting the symmetry

Since we have not used the fact that all servers are identical, the results in the previous sections are still valid in case server i works with rate μ_i, where the rates μ_i are not necessarily identical. In this section, however, we exploit the fact that all servers are identical and we show that the number of coefficients a_j to be determined, can be reduced from r^c to $\binom{c+r-1}{r-1}$. The same result could have been obtained directly by chosing the same aggregated state concept as in [4].

By letting the parameters x_i run through all feasible values, we find $\rho k r^c$ products satisfying (1)-(2) (including, of course, exactly κr copies of each product). Since equation (13) is invariant under permutations of $x_1, \ldots, x_c$, it follows that for each $\beta_{-1}^{n_{-1}} \beta_0^{n_0} \ldots \beta_c^{n_c}$ satisfying (1)-(2), all products obtained by permuting the factors $\beta_1, \ldots, \beta_c$ also satisfy (1)-(2). This suggests to split up the set of $\rho k r^c$ products into subsets of products, which are equal up to a permutation of the last c factors. The number of such subsets is $\rho k \binom{c+r-1}{r-1}$. However, since each product has exactly κr copies, each subset also has exactly κr copies. So there are $\binom{c+r-1}{r-1}$ different subsets. Let Π_l be the set of labels of the products $\beta_{-1,j}^{n_{-1}} \beta_{0,j}^{n_0} \ldots \beta_{c,j}^{n_c}$ in the l-th subset. Then expression (14) may be written as

$$p(\overline{n}) = \sum_{l=1}^{\binom{c+r-1}{r-1}} \sum_{j \in \Pi_l} a_j \beta_{-1,j}^{n_{-1}} \beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c}. \tag{16}$$

By symmetry

$$p(n_{-1}, n_0, n_1, \dots, n_c) = p(n_{-1}, n_0, n_{\pi(1)}, \dots, n_{\pi(c)})$$

for each permutation π of $\{1, \dots, c\}$. Hence, for all states $\overline{n}$ with $n_i > 0$ for $i = 1, \dots, c$

$$\sum_{l=1}^{\binom{c+r-1}{r-1}} \sum_{j \in \Pi_l} a_j \beta_{-1,j}^{n_{-1}} \beta_{0,j}^{n_0} \beta_{1,j}^{n_1} \cdots \beta_{c,j}^{n_c} = \sum_{l=1}^{\binom{c+r-1}{r-1}} \sum_{j \in \Pi_l} a_j \beta_{-1,j}^{n_{-1}} \beta_{0,j}^{n_0} \beta_{\pi(1),j}^{n_1} \cdots \beta_{\pi(c),j}^{n_c}.$$

Since the products $\beta_{-1,j}^{n_{-1}} \beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c}$ are linearly independent on this set of states, it follows that for each l the coefficients a_l with $j \in \Pi_l$ are identical. So, denoting some arbitrarily chosen index in Π_l by j_l, we obtain from (16) that

$$p(\overline{n}) = \sum_{l=1}^{\binom{c+r-1}{r-1}} a_{j_l} \sum_{j \in \Pi_l} \beta_{-1,j}^{n_{-1}} \beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c}.$$

For each state $\overline{n}$ we denote by $s_i(\overline{n})$ the number of servers with i remaining service stages. Then it is readily seen that

$$\sum_{j \in \Pi_l} \beta_{-1,j}^{n_{-1}} \beta_{0,j}^{n_0} \cdots \beta_{c,j}^{n_c} = \beta_{-1,j_l}^{n_{-1}} \beta_{0,j_l}^{n_0} B_l(\overline{n}) \tag{17}$$

where $B_l(\overline{n})$ is the coefficient of $z_1^{s_1(\overline{n})} \dots z_r^{s_r(\overline{n})}$ in the polynomial

$$\prod_{i=1}^{c} (\beta_{i,j_l} z_1 + \beta_{i,j_l}^2 z_2 + \dots + \beta_{i,j_l}^r z_r).$$

Since $\beta_{i,j_l} = x_{i,j_l} \beta_{1,j_l}$ with $x_{i,j_l}^r = 1$ we may write

$$B_l(\overline{n}) = \beta_{1,j_l}^{n_1 + \dots + n_c} C_l(\overline{n})$$

where $C_l(\overline{n})$ is the coefficient of $z_1^{s_1(\overline{n})} \dots z_r^{s_r(\overline{n})}$ in the polynomial

$$\prod_{i=1}^{c} (x_{i,j_l} z_1 + x_{i,j_l}^2 z_2 + \dots + x_{i,j_l}^r z_r).$$

These findings are summarized in the following theorem.

Theorem 2: *For all states $\overline{n}$ with $n_i > 0$ for $i = 1, \dots, c$, it holds that*

$$p(\overline{n}) = \sum_{l=1}^{\binom{c+r-1}{r-1}} a_{j_l} \beta_{-1,j_l}^{n_{-1}} \beta_{0,j_l}^{n_0} \beta_{1,j_l}^{n_1 + \dots + n_c} C_l(\overline{n})$$

for suitably chosen coefficients a_{j_l}.

6 Two special cases

Of course, for particular cases, the results can be made more explicit. In this section we work out in more detail the results for the $E_k|E_2|c$ queue and the $E_k|E_r|1$ queue.

Let us first consider the $E_k|E_2|c$ queue. For k being odd, we can restrict the feasible values for x_{-1} to $x_{-1} = 1$, so equation (13) reduces to

$$\gamma^2(\lambda + c\mu) = \lambda + \gamma^{2+k}(c - 2l)\mu \tag{18}$$

where l is the number of parameters x_i with $x_i = -1$. We now define γ_{2l} and γ_{2l+1} as the roots inside the unit circle of (18) for $l = 0, 1, \ldots, \lfloor (c-1)/2 \rfloor$ and, if c is even, then γ_c is the positive root of (18) for $l = c/2$. From theorem 2 it easily follows that for states $\overline{n}$ with all servers busy

$$p(\overline{n}) = \sum_{l=0}^{c} a_l \beta_{-1,l}^{n_{-1}} \beta_{0,l}^{n_0} \beta_{1,l}^{n_1+\ldots+n_c} C_l(\overline{n})$$

for suitably chosen coefficients a_l, where $\beta_{-1,l} = \gamma_l^2$, $\beta_{0,l} = \gamma_l^{2k}$ and $\beta_{1,l} = \gamma_l$ and $C_l(\overline{n})$ is the coefficient of $z_1^{s_1(\overline{n})} z_2^{s_2(\overline{n})} = z_1^{s_1(\overline{n})} z_2^{c-s_1(\overline{n})}$ in the polynomial

$$\begin{aligned}
(z_1 + z_2)^{c-l}(-z_1 + z_2)^l &= \sum_{i=0}^{c-l} \binom{c-l}{i} z_1^i z_2^{c-l-i} \sum_{j=0}^{l} \binom{l}{j} (-z_1)^j z_2^{l-j} \\
&= \sum_{i=0}^{c-l} \sum_{j=0}^{l} \binom{c-l}{i} \binom{l}{j} (-1)^j z_1^{i+j} z_2^{c-i-j} \\
&\overset{i+j=m}{=} \sum_{m=0}^{c} \sum_{i=\max(0,m-l)}^{\min(c-l,m)} \binom{c-l}{i} \binom{l}{m-i} (-1)^{m-i} z_1^m z_2^{c-m}.
\end{aligned}$$

For $k = 1$ these results are equivalent to the ones in Shapiro [11]. The case that k is even can be treated similarly.

Let us now consider the $E_k|E_r|1$ queue. From theorem 1 it follows that for states $\overline{n}$ with a busy server

$$p(\overline{n}) = p(n_{-1}, n_0, n_1) = \sum_{j=1}^{r} a_j \beta_{-1,j}^{n_{-1}} \beta_{0,j}^{n_0} \beta_{1,j}^{n_1}. \tag{19}$$

In general the coefficients a_j have to be solved numerically from the boundary equations. However, in this case, they can be solved explicitly from the equations (3) stating that

$$p(k, 0, n_1)(\lambda + \mu) = p(k, 0, n_1 + 1)\mu, \quad n_1 = 1, 2, \ldots, r - 1; \tag{20}$$

$$p(k, 0, r)(\lambda + \mu) = p(1, 0, 0)\lambda + p(k, 1, 1)\mu. \tag{21}$$

Substitution of expression (19) in equation (20) and using (8) yields

$$\sum_{j=1}^{r} a_j \beta_{-1,j}^{k} \beta_{0,j}^{-1} \beta_{1,j}^{n_1} = 0, \quad n_1 = 1, 2, \ldots, r - 1.$$

This set of equations can be solved by using Cramer's rule. Since the determinants involved are *VanderMonde* determinants, we easily obtain

$$\begin{aligned} a_j^{-1} &= C\beta_{-1,j}\beta_{0,j}^{-1}\beta_{1,j}\prod_{i\neq j}(\beta_{1,j}-\beta_{1,i}) \\ &\overset{\beta_{0,j}=\beta_{1,j}^r}{=} C\beta_{-1,j}\prod_{i\neq j}\left(1-\frac{\beta_{1,i}}{\beta_{1,j}}\right), \quad j=1,\ldots,r \end{aligned} \tag{22}$$

for some constant C. To satisfy equation (21) we have to set

$$p(1,0,0)=\sum_{j=1}^{r}a_j\beta_{-1,j}\beta_{0,j}^{-1}\beta_{1,j}^r. \tag{23}$$

Finally, the constant C can be determined from the following equation stating that $1/k$ is the fraction of time there is one remaining arrival stage.

$$p(1,0,0)+\sum_{n_0=0}^{\infty}\sum_{n_1=1}^{r}p(1,n_0,n_1)=\frac{1}{k}. \tag{24}$$

By substituting the expressions (19) and (23) in the left-hand side of (24) it follows that

$$\begin{aligned} p(1,0,0)+\sum_{n_0=0}^{\infty}\sum_{n_1=1}^{r}p(1,n_0,n_1) &= \sum_{j=1}^{r}\frac{a_j\beta_{-1,j}}{1-\beta_{1,j}} \\ &\overset{(22)}{=} \frac{1}{C\prod_j\beta_{1,j}}\sum_{j=1}^{r}\frac{1}{\left(\frac{1}{\beta_{1,j}}-1\right)\prod_{i\neq j}\left(1-\frac{\beta_{1,i}}{\beta_{1,j}}\right)} \\ &= \frac{1}{C\prod_j\beta_{1,j}\prod_j\left(\frac{1}{\beta_{1,j}}-1\right)}\sum_{j=1}^{r}\frac{\prod_{i\neq j}\left(\frac{1}{\beta_{1,i}}-1\right)}{\prod_{i\neq j}\left(\frac{1}{\beta_{1,i}}-\frac{1}{\beta_{1,j}}\right)} \\ &= \frac{1}{C\prod_j(1-\beta_{1,j})}. \end{aligned} \tag{25}$$

The last equality follows by observing that the polynomial

$$P(x)\overset{\text{def}}{=}\sum_{j=1}^{r}\frac{\prod_{i\neq j}\left(\frac{1}{\beta_{1,i}}-x\right)}{\prod_{i\neq j}\left(\frac{1}{\beta_{1,i}}-\frac{1}{\beta_{1,j}}\right)}$$

with degree $r-1$ satisfies $P(1/\beta_{1,j})=1$ for $j=1,\ldots,r$, so $P(x)\equiv 1$ and thus $P(1)=1$. From (24) and (25) we find

$$C=\frac{k}{\prod_j(1-\beta_{1,j})}. \tag{26}$$

Hence, for the $E_k|E_r|1$ queue the probabilities $p(\overline{n})$ are characterised explicitly by the expressions (19), (22) and (26). Similar expressions may be found for the mean queue length and the mean waiting time, compare e.g. the results in Ikeda [7], where an expression for the mean waiting time in terms of a double series is derived.

7 Comments and conclusions

It has been shown that the equilibrium distribution for the $E_k|E_r|c$ queue may be characterised as a finite linear combination of products of powers. In these products each component of the state is represented by one power-factor. These power-factors have been characterised completely. The weights of the linear combination are determined by a set of linear equations. Although the analysis has been worked out for the case of identical servers, the main result (theorem 1) can easily be extended to the case of nonidentical servers. In that case the linear combination contains r^c terms which makes the solution tractable from a computational point of view. If there are some groups of identical servers, then the required number of terms decreases considerably. In the case that all servers are equal, the number decreases to $\binom{c+r-1}{r-1}$. For the case $c = 1$, the equations for the weights can be solved explicitly. Also in other special cases the results can be made somewhat more explicit.

With respect to possible extensions, there are several directions of interest. One interesting direction is constituted by the case of separate queues for the c servers with some sort of allocation of the incoming jobs to the respective servers. For such systems it has been proved that the so-called shortest-delay routing is optimal under some conditions (see Hordijk, Koole [6]). Shortest-delay routing allocates an incoming job to the server with the lowest number of phases in its queue. This situation has been treated to some extent by Adan in [1] (see section 6.2) for the case $c = 2$, $k = 1$. In this situation one gets a linear combination of countably many terms which can be computed according to a tree-like recurrent scheme. The results are complete for $r = 2$. For $r > 2$, however, some properties of the recurrent scheme are not established yet. For $c > 2$, it is unlikely that the approach will work, since it apparently does not work for the shortest queue problem with 3 servers, Poisson arrivals, and exponentially distributed service times.

Another interesting direction of extension consists of replacing E_k and/or E_r by more general distribution classes. One may think here of Coxians (cf. Bertsimas [4]) or of finite mixtures of Erlang distributions with the same scale parameters, but also of general phase-type distributions (cf. Neuts [8]). The class of finite mixtures of Erlang distributions with the same scale parameters is dense in the set of all distributions (cf. Schassberger [10]) and leads to the same state description as the one in the present paper, however, with slightly more complicated transition behaviour. A simplification would be to accept only mixtures of two Erlang distributions with $k-1$ and k phases for the interarrival times and with $r-1$ and r phases for the service times, since (as Tijms [12] points out) this class seems to be sufficiently rich for most practical purposes. Particularly for the two last-mentioned situations, it seems most likely that similar results may be derived as in the present paper.

Appendix A: Proof of Lemma 4

Let

$$\begin{aligned} f(z) &= z^r(\lambda + c\mu) - x_{-1}\lambda, \\ g(z) &= -z^{r+k}\sum_{i=1}^{c} x_i\mu. \end{aligned}$$

Then we have to show that $f(z)+g(z)$ has r simple zeros inside the unit circle. It readily follows that

$$\begin{aligned} |f(z)| &\geq |z|^r(\lambda + c\mu) - \lambda \stackrel{\text{def}}{=} L(|z|), \\ |g(z)| &\leq |z|^{r+k}c\mu \stackrel{\text{def}}{=} l(|z|). \end{aligned}$$

Since $L(1) = l(1)$ and $L'(1) < l'(1)$ by virtue of the condition $\lambda r/k\mu < c$, there exists an $\epsilon > 0$ such that $L(1-\epsilon) > l(1-\epsilon)$. Hence, from Rouché's theorem we can conclude that $f(z) + g(z)$ has r zeros inside the circle $|z| = 1 - \epsilon$. It remains to show that all zeros are simple. Assume that $f(z) + g(z)$ and $f'(z) + g'(z)$ vanish for some z, i.e.

$$z^r(\lambda + c\mu) - x_{-1}\lambda - z^{r+k}\sum_{i=1}^{c} x_i\mu = 0, \tag{27}$$

$$rz^{r-1}(\lambda + c\mu) - (r+k)z^{r+k-1}\sum_{i=1}^{c} x_i\mu = 0. \tag{28}$$

Below we argue that this assumption leads to a contradiction. If $\sum_{i=0}^{c} x_i = 0$, then the equations (27) and (28) have no solution. If $\sum_{i=0}^{c} x_i \neq 0$, it follows from (28) that

$$z^k = \frac{r(\lambda + c\mu)}{(r+k)\sum_{i=1}^{c} x_i\mu} \tag{29}$$

and insertion of (28) in (27) yields

$$z^{r+k} = \frac{rx_{-1}\lambda}{k\sum_{i=1}^{c} x_i\mu}. \tag{30}$$

Taking absolute values in (29) and (30) and combining these two equalities gives

$$h(\lambda) \stackrel{\text{def}}{=} \left(\frac{r(\lambda + c\mu)}{(r+k)|\sum_{i=1}^{c} x_i|\mu}\right)^{\frac{r+k}{k}} - \frac{r\lambda}{k|\sum_{i=1}^{c} x_i|\mu} = 0. \tag{31}$$

However, below it will be shown that $h(\lambda) > 0$ contradicting (31). Then we may conclude that there are no values of z satisfying (27) and (28). First, note that $h(0) > 0$ and $h''(x) > h''(0) > 0$ for all $x > 0$. Then $h'(0) \geq 0$ implies $h(x) > 0$ for $x > 0$ and thus in particular $h(\lambda) > 0$. Otherwise, if $h'(0) < 0$, then $h'(x) = 0$ for a unique $\hat{x} > 0$ given by

$$\hat{x} = \frac{r+k}{r}|\sum_{i=1}^{c} x_i|\mu - c\mu.$$

It is easily verified that

$$h(\hat{x}) = \frac{r}{k}\left(\frac{c}{|\sum_{i=1}^{c} x_i|} - 1\right).$$

If $|\sum_{i=1}^{c} x_i| < c$, then $h(\hat{x}) > 0$ and thus, since $\hat{x}$ is the global minimum of $h(x)$ for $x > 0$, it follows that $h(\lambda) > 0$. Finally, if $|\sum_{i=1}^{c} x_i| = c$, then $h(\hat{x}) = 0$ and $\hat{x}$ simplifies to $\hat{x} = c\mu k/r$. The condition $\lambda r/k\mu < c$ implies that $\lambda < \hat{x}$, so $h(\lambda) > h(\hat{x}) = 0$, completing the proof of lemma 4. □

Appendix B: Proof of Lemma 6

Let us assume that (15) holds. By substituting $\beta_{0,j} = \beta_{1,j}^r$ and $\beta_{i,j} = x_{i,j}\beta_{1,j}$, $i = 2, \ldots, c$ where $x_{i,j}$ are solutions of $x_{i,j}^r = 1$, into equation (15) and introducing the variable m denoting the total number of uncompleted service stages in the system, i.e.

$$m = n_0 r + n_1 + \ldots + n_c,$$

it readily follows that equation (15) implies

$$\sum_{j=1}^{r^c} a_j \beta_{-1,j}^{n_{-1}} \beta_{1,j}^m x_{2,j}^{n_2} \ldots x_{c,j}^{n_c} = 0 \quad (0 < n_{-1} \le k, m \ge cr, 0 < n_i \le r, i = 2, \ldots, c). \tag{32}$$

For each feasible choice of $n_{-1}, m, n_3, \ldots, n_c$ we obtain from (32) that

$$\sum_{n=0}^{r-1} \left(\sum_{x_{2,j} = e^{\frac{2\pi i}{r} n}} a_j \beta_{-1,j}^{n_{-1}} \beta_{1,j}^m x_{3,j}^{n_3} \ldots x_{c,j}^{n_c} \right) \left(e^{\frac{2\pi i}{r} n} \right)^{n_2} = 0 \quad (0 < n_2 \le r).$$

By well known properties of the *VanderMonde matrix* this implies that for $n = 0, \ldots, r-1$

$$\sum_{x_{2,j} = e^{\frac{2\pi i}{r} n}} a_j \beta_{-1,j}^{n_{-1}} \beta_{1,j}^m x_{3,j}^{n_3} \ldots x_{c,j}^{n_c} = 0 \quad (0 < n_{-1} \le k, m \ge cr, 0 < n_i \le r, i = 3, \ldots, c).$$

By repeatedly applying this procedure, equation (32) eventually decomposes into r^{c-1} sets of equations, i.e. for each choice of x_j satisfying $x_j^r = 1$, $j = 2, \ldots, c$ we obtain

$$\sum_{x_{2,j} = x_2, \ldots, x_{c,j} = x_c} a_j \beta_{-1,j}^{n_{-1}} \beta_{1,j}^m = 0 \quad (0 < n_{-1} \le k, m \ge cr). \tag{33}$$

The sum in (33) runs over exactly r products of which the factors $\beta_{-1,j}$ and $\beta_{1,j}$ are given by $\beta_{-1,j} = x_{-1}/\gamma^\rho$ and $\beta_{1,j} = \gamma^\kappa$ with $x_{-1} = e^{\frac{2\pi i}{k} n}$ for some $n = 0, 1, \ldots, d-1$ and γ is a root inside the unit circle of

$$\gamma^\rho(\lambda + c\mu) = x_{-1}\lambda + \gamma^{\rho+\kappa} \sum_{i=1}^{c} x_i \mu. \tag{34}$$

From the next lemma it follows that the products in (33) have different factors $\beta_{1,j}$ implying that $a_j \equiv 0$. Since, on the other hand, $a_j \equiv 0$ trivially implies (15), this completes the proof of lemma 6.

Lemma 7: *Let γ_1 and γ_2 be roots inside the unit circle of (34) with $x_{-1} = e^{\frac{2\pi i}{k} n_1}$ and $x_{-1} = e^{\frac{2\pi i}{k} n_2}$ respectively, where $0 \le n_1 < n_2 < d$. Then $\gamma_1^\kappa \ne \gamma_2^\kappa$. The same result holds if $n_1 = n_2$ provided that $\gamma_1 \ne \gamma_2$.*

Proof: It is shown that the assumption $\gamma_1^\kappa = \gamma_2^\kappa$ leads to a contradiction. The roots γ_1 and γ_2 satisfy

$$\gamma_1^\rho(\lambda + c\mu) = y_1\lambda + \gamma_1^{\rho+\kappa} \sum_{i=1}^{c} x_i \mu, \tag{35}$$

$$\gamma_2^\rho(\lambda + c\mu) = y_2\lambda + \gamma_2^{\rho+\kappa}\sum_{i=1}^{c} x_i\mu, \tag{36}$$

where $y_1 = e^{\frac{2\pi i}{k}n_1}$ and $y_2 = e^{\frac{2\pi i}{k}n_2}$. Multiplying (35) with y_2 and (36) with y_1 and then subtracting both equations and inserting $\gamma_2^\kappa = \gamma_1^\kappa$ yields

$$(y_2\gamma_1^\rho - y_1\gamma_2^\rho)(\lambda + c\mu) = \gamma_1^\kappa(y_2\gamma_1^\rho - y_1\gamma_2^\rho)\sum_{i=1}^{c} x_i\mu. \tag{37}$$

If we can prove that $y_2\gamma_1^\rho - y_1\gamma_2^\rho \neq 0$, then division of (37) by this term results in

$$\lambda + c\mu = \gamma_1^\kappa \sum_{i=1}^{c} x_i\mu,$$

which contradicts that $|\gamma_1| < 1$ and thus completes the proof of lemma 7. Hence, it remains to prove that $y_2\gamma_1^\rho \neq y_1\gamma_2^\rho$. Let us suppose to the contrary, i.e.

$$\gamma_2^\rho = \frac{y_2}{y_1}\gamma_1^\rho = e^{\frac{2\pi i}{k}(n_2-n_1)}\gamma_1^\rho. \tag{38}$$

On the other hand, since $\gamma_2 \neq \gamma_1$ and $\gamma_2^\kappa = \gamma_1^\kappa$, it follows that $\gamma_2 = e^{\frac{2\pi i}{\kappa}n}\gamma_1$ for some $n = 1, 2, \ldots, \kappa - 1$. Substitution of this relation in (38) yields

$$e^{\frac{2\pi i}{\kappa}n\rho} = e^{\frac{2\pi i}{k}(n_2-n_1)}.$$

This equality implies that (recall that $k = \kappa d$)

$$n\rho = \frac{n_2 - n_1}{d} + m\kappa \tag{39}$$

for some integer m. Since $n_2 - n_1 < d$ it follows that (39) leads to a contradiction if $n_2 > n_1$. Otherwise, if $n_2 = n_1$, then, since $\gcd(\rho, \kappa) = 1$, equality (39) implies that n can be divided by κ contradicting that $0 < n < \kappa$. □

References

[1] I.J.B.F. ADAN, *"A compensation approach for queueing problems."* Ph.D. thesis at Eindhoven University of Technology, November 1991 (to be published as a monograph).

[2] I.J.B.F. ADAN, J. WESSELS, W.H.M. ZIJM, "Analysis of the symmetric shortest queue problem", *Stochastic Models,* vol. 6, pp. 691–713, 1990.

[3] I.J.B.F. ADAN, J. WESSELS, W.H.M. ZIJM, "A compensation approach for two-dimensional Markov processes," Memorandum COSOR 91-28, Eindhoven University of Technology, Department of Mathematics and Computing Science, October 1991 (submitted to a journal).

[4] D. BERTSIMAS, "An analytic approach to a general class of $G|G|s$ queueing systems", *Opns. Res.,* vol. 38, pp. 139–155, 1990.

[5] F.G. Foster, "On the stochastic matrices associated with certain queueing processes", *Ann. Math. Stat.*, vol. 24, pp. 355–360, 1953.

[6] A. Hordijk, G. Koole, "On the optimality of the generalized shortest queue policy", *Probab. Engineer. Inform. Sci.*, 4 (1990), pp. 477–487.

[7] Z. Ikeda, "Mean waiting time of a Gamma/Gamma/1 queue", *Opns. Res. Let.*, vol. 10, pp. 177–181, 1991.

[8] M.F. Neuts, *Matrix-geometric solutions in stochastic models,* Johns Hopkins University Press, Baltimore 1981.

[9] C.D. Poyntz, R.R.P. Jackson, "The steady-state solution for the queueing process $E_k|E_m|r$", *O. R. Quart.*, vol. 24, pp. 615-625, 1973.

[10] R.S. Schassberger, *Warteschlangen,* Springer-Verlag, Berlin 1973.

[11] S. Shapiro, "The *M*-server queue with Poisson input and Gamma-distributed service of order two", *Opns. Res.*, vol. 14, pp. 685–694, 1966.

[12] H.C. Tijms, *Stochastic modelling and analysis: a computational approach.* John Wiley, Chichester 1986.