LIGHT-TRAFFIC APPROXIMATIONS FOR MARKOV-MODULATED MULTI-SERVER QUEUES

Bartłomiej BŁASZCZYSZYN
Mathematical Institute, University of Wrocław
pl. Grunwaldzki 2/4, 50-384 Wrocław, Poland

Andreas FREY
Department of Stochastics, University of Ulm
Helmholtzstr. 18, D-89069 Ulm, Germany

Volker SCHMIDT
Department of Stochastics, University of Ulm
Helmholtzstr. 18, D-89069 Ulm, Germany

Abstract

A general concept is considered of expanding the expectation of a wide class of functionals of marked point processes, which expresses this expectation by a sum of integrals over higher-order factorial moment measures of the underlying point process. The idea of factorial moment expansion is applied in order to derive approximation formulas for stationary characteristics of multi-server queues with Markov-modulated arrival process and with the first-come-first-served queueing discipline. Besides real-valued queueing characteristics like waiting time and total work load, we also give approximations for the Kiefer-Wolfowitz work-load vector. A boundedness condition on the service time distributions is given which ensures that the components of the expected stationary work-load vector are analytic functions of the arrival intensity in a neighborhood of zero. If the service times have phase-type distributions, the factorial moment expansion provides a useful computational technique for approximations of moments of the stationary work-load vector. Some numerical examples are given which show how the algorithm works in light traffic.

Keywords: queueing, point processes, moment measures, Palm distributions, Markov-modulated input, phase-type distributions, MM/PH/c, *work-load vector, expansion, light traffic*

AMS 1991 Subject Classification: 60K25, 60G55

1 INTRODUCTION

Stationary queueing systems with more than one server are hard to analyze by closed analytic formulas unless the stationary joint queue length distribution is of some product-form type. Multi-server queues with the first-come-first-served (FCFS) queueing discipline are a class of queueing systems for which such a factorization usually fails and, consequently, analytic formulas are the exception rather than normality. This situation motivated us to consider approximate solutions of stationary multi-server queues. In connection with this we interpret queueing characteristics, like waiting time or work load (or functions of them), as functionals of a stationary marked point process which models the input of the system. Starting from a general representation formula for the expectation of such functionals, which expresses this expectation by a sum of integrals over higher-order factorial moment measures of the underlying marked point process of input data (like in Błaszczyszyn [7]), we arrive at a Taylor-series expansion with respect to the arrival intensity provided that the arrival times form a stationary Markov-modulated Poisson process where the distribution of service times may depend on the actual state of the Markov-modulated environment. If the service times have phase-type distributions, we are able to calculate the coefficients in this expansion up to a certain degree. In this way, a useful computational technique is given for light-traffic approximations of several stationary characteristics of Markov-modulated multi-server queues. Namely, besides approximation formulas for real-valued characteristics like waiting time and total work load, we also give approximations for the Kiefer-Wolfowitz work-load vector. Some numerical experiments show that, under light traffic, the deviations between the results obtained by using our approximation formulas and corresponding exact or simulated results are rather small.

In queueing theory, several approaches have been investigated in order to establish approximation formulas for stationary characteristics of multi-server queues. Most of the earlier papers assume that the arrival times form a homogeneous Poisson process (see e.g. Boxma et al. [9], Cosmetatos [11] [12], Nozaki and Ross [26], Takahashi [33], Tijms et al. [34]). Comprehensive surveys on general algorithms and approximations for multi-server queueing systems are given in Kimura [21], Neuts [24], Stoyan [32] and Tijms [35].

Some papers present specific approximations which work particularly well when, besides the assumption of Poisson arrival times, the traffic load satisfies an additional condition, e.g. the heavy-traffic approximations by the diffusion models considered in Kimura [20] and Yao [37] or the light-traffic approximations considered in Blanc [6], Burman and Smith [10], Wang and Wolff [36]. In Daley and Rolski [13], light-traffic approximations are given for multi-server queues with independent, but not necessarily exponentially distributed interarrival times (i.e. the arrival times form a recurrent point process which need not to be Poisson).

Light-traffic approximation formulas for a wide class of functionals driven by a Poisson arrival process or by a recurrent arrival process with phase-type interarrival distribution, including characteristics of Markovian queueing networks, have been derived in Reiman and Simon [27] [28] and in Simon [31]. Their investigations are quite similar to our approach. But, they do not use the concept of higher-order moment measures and Palm distributions of general marked point processes which turns out to be very useful, in particular for getting higher-order light-traffic approximations for characteristics of Markov-modulated multi-server queues.

Applications of this approach to single-server queues and risk processes with Markov-modulated input and with Poisson input, respectively, can be found in Błaszczyszyn and Rolski [8], Frey and Schmidt [15], Kroese and Schmidt [23]. For first-order light-traffic approximations of the (total) queue length and work load in queues with a general stationary input we refer to Kroese and Schmidt [23], Sigman [30].

The present paper is organized as follows. In Section 2 we introduce some basic notions of marked point processes on the real line, like factorial moment measures, Campbell measures and Palm distributions. We state a general factorial moment expansion for functionals of marked point processes which will be an important tool for deriving higher-order light-traffic approximations. This expansion is a sum of integrals of so-called factorial-moment-expansion (FME) kernels over higher-order factorial moment measures of the underlying point process. It is particularly useful for functionals of Markov-modulated Poisson processes which are also introduced in this section. In Section 3 we recall the definitions of some characteristics of multi-server queues presenting them as functionals of marked point processes and computing FME kernels of these characteristics. The general results stated in the preceding sections are then applied in Section 4 in order to derive light-traffic approximation formulas for stationary characteristics of MM/G/c queues with Markov-modulated arrival process. Moreover, a boundedness condition on the service time distributions is given which ensures that the components of the expected stationary work-load vector are analytic functions of the arrival intensity in a neighborhood of zero. In Section 5 we consider MM/PH/c queues with service time distributions of phase type. Then, we are able to determine the integrals more explicitly which appear in the general factorial moment expansions derived in Section 4. This provides a computational technique for approximations of moments of the stationary work-load vector in MM/PH/c queues. Some numerical examples which show how the algorithm works are also given in Section 5.

2 PRELIMINARIES

2.1 Marked Point Processes

In this paper we consider marked point processes on the real line $\mathbb{R}$ with marks also in $\mathbb{R}$. Our approach will be via point measures on $\mathbb{R} \times \mathbb{R}$ defined on the Borel σ-algebra $\mathcal{B}(\mathbb{R} \times \mathbb{R})$. Let $\mathcal{M}$ be the space of realizations of marked point processes under consideration, i.e. the family of measures finite on every set of the form $A \times \mathbb{R}$, where A is bounded, taking values in $\overline{\mathbb{N}} = \{0, 1, \ldots\} \cup \{\infty\}$. The elements of $\mathcal{M}$ are called *point measures*. Let $\mathcal{M}_s$ be the set of measures $\mu \in \mathcal{M}$ such that $\mu(\{x\} \times \mathbb{R})$ is either 0 or 1 for each x, and call the elements of $\mathcal{M}_s$ *simple point measures*. By

$\varepsilon_{(x,s)}$ we denote the point mass at $(x,s) \in \mathbb{R}^2$, which satisfies

$$\varepsilon_{(x,s)}(A \times L) = \begin{cases} 1 & \text{for } (x,s) \in A \times L \\ 0 & \text{for } (x,s) \notin A \times L. \end{cases}$$

Thus every element μ of $\mathcal{M}_s$ possesses the representation

$$\mu = \sum_{n=-k_1}^{k_2} \varepsilon_{(t_n,m_n)} \qquad \ldots t_{-2} < t_{-1} < 0 \leq t_0 < t_1 \ldots; \quad (0 \leq k_1, k_2 \leq \infty) \tag{2.1}$$

where $\{t_i\}$ is a sequence in $\mathbb{R}$ without accumulation points and $m_i \in \mathbb{R}$. Let $\mathcal{B}(\mathcal{M})$ be the smallest σ-algebra of subsets of $\mathcal{M}$, which makes the functions $\mathcal{M} \ni \mu \to \mu(A)$ measurable for all bounded $A \in \mathcal{B}(\mathbb{R}^2)$. A *marked point process* (m.p.p.) is meant as the identity transformation $\mathcal{M} \ni \mu \to N(\mu) = \mu$ and every probability measure $\mathbf{P}$ on $(\mathcal{M}, \mathcal{B}(\mathcal{M}))$ defines a *stochastic marked point process* $(N, \mathbf{P})$. Sometimes it is useful to assume that the probability measure $\mathbf{P}$ is given on a somewhat more general measurable space (see Section 2.3). We say that a stochastic m.p.p. $(N, \mathbf{P})$ is simple if $\mathbf{P}(\, N \in \mathcal{M}_s\,) = 1$. We assume from now on that all stochastic m.p.p.'s which we consider in this paper are simple.

It is convenient to use the notation $(\boldsymbol{x}, \boldsymbol{s}) = (x_1, s_1, \ldots, x_k, s_k)$ for $\mathbb{R}^{2k}$-valued vectors.

Definition 2.1 The factorial moment measure $M^{(k)}$ of $(N, \mathbf{P})$ of order k is a measure on $\mathbb{R}^{2k}$ given by

$$M^{(k)}(d(\boldsymbol{x}, \boldsymbol{s})) = \mathbf{E}\left[N^{(k)}(d(\boldsymbol{x}, \boldsymbol{s}))\right], \qquad k = 1, 2, \ldots,$$

where

$$\begin{aligned} & N^{(k)}(d(\boldsymbol{x}, \boldsymbol{s})) \\ = \;& N(d(x_1, s_1)) \left(N - \varepsilon_{(x_1,s_1)}\right) (d(x_2, s_2)) \ldots \left(N - \sum_{i=1}^{k-1} \varepsilon_{(x_i,s_i)}\right) (d(x_k, s_k)). \end{aligned} \tag{2.2}$$

Definition 2.2 For each k, the measure $C^{(k)}$ on $\mathbb{R}^{2k} \times \mathcal{M}$ given by

$$C^{(k)}(A \times \Gamma) = \mathbf{E}\left[\int_A \mathbf{1}\left(N - \sum_{i=1}^{k} \varepsilon_{(x_i,s_i)} \in \Gamma\right) N^{(k)}(d(\boldsymbol{x}, \boldsymbol{s}))\right] \tag{2.3}$$

is said to be the kth order (reduced) *Campbell measure* of $(N, \mathbf{P})$.

Definition 2.3 The *reduced k-fold Palm distributions* of $(N, \mathbf{P})$ are the transitions functions $P_{\cdot}^{(k)}(\cdot) : \mathbb{R}^{2k} \times \mathcal{B}(\mathcal{M}) \to [0,1]$ defined by

$$\mathbf{P}^{(k)}_{\boldsymbol{x},\boldsymbol{s}}(d\mu) = \frac{dC^{(k)}(\cdot \times d\mu)}{dC^{(k)}(\cdot \times \mathcal{M})}(\boldsymbol{x}, \boldsymbol{s}). \tag{2.4}$$

The construction of the k-fold Palm distributions is fundamentally simple: for each $\Gamma \in \mathcal{B}(\mathcal{M})$, we have $C^{(k)}(\cdot \times \Gamma) \ll C^{(k)}(\cdot \times \mathcal{M})$, so there exists the Radon-Nikodym derivative $\mathbf{P}_{\boldsymbol{x},\boldsymbol{s}}(\Gamma)$. One must then ensure that it is a probability measure for $M^{(k)}$-almost every $\boldsymbol{x}, \boldsymbol{s}$; see e.g. Kallenberg [17], Karr [18], König and Schmidt [22] for details of the idea which goes back to Ryll-Nardzewski [29].

2.2 Factorial Moment Expansion

Let $c \in \mathbb{N} = \{1, 2, \ldots\}$ be an arbitrary natural number, and let ψ be an $\mathbb{R}^c$-valued functional of a point process, i.e. a $(\mathcal{B}(\mathcal{M}), \mathcal{B}(\mathbb{R}^c))$-measurable function $\psi : \mathcal{M} \mapsto \mathbb{R}^c$, and assume that it is integrable with respect to the distribution of some stochastic point process $(N, \mathbf{P})$. Following the general method of *factorial moment expansion* (FME) in this section we present an expansion of the expectation $\mathbf{E}\psi(N)$ in terms of factorial moment measures of $(N, \mathbf{P})$.

For this we define, for each $k \in \mathbb{N}$, the *k-th FME kernel* of the functional ψ on $\mathcal{M}_s$ as a functional $\psi_{(k)}$ on $\mathbb{R}^{2k} \times \mathcal{M}_s$ by

$$\psi_{(k)}^{\boldsymbol{x},\boldsymbol{s}}(\mu) = \begin{cases} \sum_{j=0}^{k} (-1)^{k-j} \sum_{\pi \in \{\binom{k}{j}\}} \psi\left(\mu|^{x_k} + \sum_{i \in \pi} \varepsilon_{(x_i, s_i)}\right) & \text{for } x_1 < \ldots < x_k, \\ \mathbf{0} & \text{otherwise}, \end{cases} \tag{2.5}$$

where $\{\binom{k}{j}\}$ denotes the collection of all subsets of $\{1, \ldots, k\}$ consisting of j elements and $\mu|^x$ is given by

$$\mu|^x(B) = \mu(B \cap ((x, \infty) \times \mathbb{R})); \qquad B \in \mathcal{B}(\mathbb{R}^2).$$

We call the functional $\psi_{(k)}(o)$ taken on the null measure $o \in \mathcal{M}_s$ (i.e. $o(B) = 0$; $B \in \mathcal{B}(\mathbb{R}^2)$) the *k th null FME kernel*. The function $\psi_{(k)}^{\boldsymbol{x},\boldsymbol{s}}(o)$ appeared for the first time in Reiman and Simon [28].

A functional ψ on $\mathcal{M}_s$ is said to be *continuous at* $\pm\infty$ if

$$\lim_{x \to \infty} \psi(\mu|^x + \nu) = \psi(\nu), \qquad \lim_{x \to -\infty} \psi(\mu|^x) = \psi(\mu), \tag{2.6}$$

for every $\mu, \nu \in \mathcal{M}_s$, such that $\nu(\mathbb{R}^2) < \infty$. The following theorem is a straightforward extension of Theorem 3.2 in Błaszczyszyn [7] (see also Błaszczyszyn and Rolski [8], Kroese and Schmidt [23]).

Theorem 2.1 *Let ψ be a continuous at $\pm\infty$ functional of a simple m.p.p. $(N, \mathbf{P})$. If*

$$\int_{\mathbb{R}^{2i}} \int_{\mathcal{M}_s} \left|\psi_{(i)}^{\boldsymbol{x},\boldsymbol{s}}(\mu)\right| \mathbf{P}_{\boldsymbol{x},\boldsymbol{s}}^{(i)}(d\mu) M^{(i)}(d(\boldsymbol{x}, \boldsymbol{s})) < \infty \tag{2.7}$$

for all $i = 1, \ldots, n$ then

$$\begin{aligned} \mathbf{E}\psi(N) \;\; &= \;\; \psi(o) + \sum_{i=1}^{n-1} \int_{\mathbb{R}^{2i}} \psi_{(i)}^{\boldsymbol{x},\boldsymbol{s}}(o) M^{(i)}(d(\boldsymbol{x}, \boldsymbol{s})) \\ &\quad + \int_{\mathbb{R}^{2n}} \int_{\mathcal{M}_s} \psi_{(n)}^{\boldsymbol{x},\boldsymbol{s}}(\mu)\, \mathbf{P}_{\boldsymbol{x},\boldsymbol{s}}^{(n)}(d\mu) M^{(n)}(d(\boldsymbol{x}, \boldsymbol{s})). \end{aligned} \tag{2.8}$$

The *proof* of this theorem is sketched in the Appendix.

2.3 Markov-Modulated Point Processes

The following Markov-modulated marked point process is considered in this paper. Let $X = \{X(t)\}$ be a homogeneous Markov process assuming values in $\{1, \ldots, \ell\}$, and let the dynamics of X be governed by the intensity matrix $\boldsymbol{\Lambda}$. We say that the process X defines an environment. We assume that X is irreducible, so, since the state space is finite, there exists a uniquely determined stationary initial distribution of X denoted by $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_\ell)$. Note that realizations of X can be taken piecewise constant. Now let $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_\ell)$ be a vector of nonnegative numbers and $\boldsymbol{B} = (B_1, \ldots, B_\ell)$ a vector which components are distributions on $\mathbb{R}$. Assume that if the environment is in state i, points appear according to a stationary Poisson process with (constant) rate β_i and if there is a point it gets a mark drawn independently of everything else with distribution B_i. We call such a stochastic m.p.p. *Markov-modulated.* Formally $(N, X, \mathbf{P})$, where $\mathbf{P}$ denotes the distribution of (N, X) on a certain probability space, is a *Cox process*, i.e. a mixture of (non-homogeneous) marked Poisson processes, with stochastic intensity measure on $\mathbb{R} \times \mathbb{R}$ of the form

$$\beta_{X(t)}\, B_{X(t)}(ds)dt.$$

The advantage of Markov-modulated m.p.p.'s is that they form a rather large class of point processes whose factorial moment measures can still be expressed by relatively simple formulas. In Błaszczyszyn and Rolski [8] it has been shown how the following result can be deduced from the general theory of Cox processes and from the definitions of factorial moment measures and reduced Palm distributions (see also Reiman and Simon [27]).

Theorem 2.2 *Assume that X is stationary with initial distribution $\boldsymbol{\pi}$. The k-th factorial moment measure $M^{(k)}$ of $(N, \mathbf{P})$ is given by*

$$M^{(k)}(d(\boldsymbol{x}, \boldsymbol{s})) = \boldsymbol{\pi} \boldsymbol{S}(ds_{(1)}) e^{\boldsymbol{\Lambda}(x_{(2)} - x_{(1)})} \boldsymbol{S}(ds_{(2)}) e^{\boldsymbol{\Lambda}(x_{(3)} - x_{(2)})} \tag{2.9}$$
$$\ldots \boldsymbol{S}(ds_{(k-1)}) e^{\boldsymbol{\Lambda}(x_{(k)} - x_{(k-1)})} \boldsymbol{S}(ds_{(k)}) \boldsymbol{e}\, dx_1 dx_2 \ldots dx_k$$

where

$$\boldsymbol{S}(A) = \mathrm{diag}(\beta_i B_i(A)),$$

$\boldsymbol{e}$ is a column vector with all entries equal to 1 and by $(x_{(1)}, \ldots, x_{(k)})$ we mean a permutation of $(x_1, \ldots, x_k)$ which components $x_{(1)}, x_{(2)}, \ldots, x_{(k)}$ are arranged in ascending order. Furthermore $e^{\boldsymbol{\Lambda}x}$ is the matrix exponential given by

$$e^{\boldsymbol{\Lambda}x} = \sum_{i=0}^{\infty} \frac{(\boldsymbol{\Lambda}x)^i}{i!}. \tag{2.10}$$

In order to derive (2.9) it suffices to use the fact that $(N, \mathbf{P})$ conditioned on $X(\cdot) = f(\cdot)$, where $f(\cdot)$ is a piecewise constant function taking values in $\{1, \ldots, \ell\}$, is the Poisson point process on $\mathbb{R} \times \mathbb{R}_+$ with the intensity measure

$$\beta_{f(t)} B_{f(t)}(ds)\, dt,$$

and, consequently, thus its kth order factorial moment measure is the k-fold product of this intensity measure. From the formula for the finite-dimensional distributions of a homogeneous continuous-time Markov chain we have

$$
\begin{aligned}
M^{(k)}(d(\boldsymbol{x},\boldsymbol{s})) &= \mathbf{E}\Big[\prod_{i=1}^{k} \beta_{X(x_i)} B_{X(x_i)}(ds_i)dx_i \Big| X\Big] \\
&= \sum_{j_1,\dots,j_k=1}^{\ell} \pi_{j_1}\beta_{j_1}B_{j_1}(ds_{(1)}) \prod_{i=2}^{k} \Big(e^{\boldsymbol{\Lambda}(x_{(i)}-x_{(i-1)})}\Big)_{j_{i-1}j_i} \beta_{j_i}\, B_{j_i}(ds_{(i)})\, d\boldsymbol{x},
\end{aligned}
$$

the matrix notation of which gives (2.9). In a similar manner the following result is proved in [8]; it is useful for estimating the remainder term in (2.8) and for checking the validity of (2.7).

Theorem 2.3 *The reduced k-fold Palm distribution $\mathbf{P}^{(k)}_{\boldsymbol{x},\boldsymbol{s}}$ of $(N,\mathbf{P})$ is again the distribution of a Cox process with the stochastic intensity measure*

$$\beta_{\tilde{X}(t)}\, B_{\tilde{X}(t)}(dm)dt$$

for some stochastic process $\tilde{X}=\{\tilde{X}(t)\}$ taking values in $\{1,\dots,\ell\}$.

3 MULTI-SERVER QUEUES

In this section we recall the definitions of some characteristics of G/G/c queues presenting them as functionals of marked point processes. We also compute FME kernels of these characteristics. In general, for such an approach, points of the point process describe arrival times of customers while marks represent their service requirement.

3.1 Input and Work-Load Vector

In the following we assume that the m.p.p. $(N,\mathbf{P})$ is stationary, i.e.

$$N=\sum_i \varepsilon_{(t_i,m_i)} \stackrel{\mathbf{P}}{=} \sum_i \varepsilon_{(t_i-x,m_i)} \quad \text{for each } x\in\mathbb{R}.$$

Then define the traffic intensity ρ as

$$\rho=\frac{1}{c}\int_{\mathbb{R}} s\, M^{(1)}((0,1)\times ds),$$

where c is the number of available servers, i.e. the maximal service capacity. Furthermore, we assume that $(N,\mathbf{P})$ is ergodic. This means that the distribution $\mathbf{P}$ of N cannot be represented as a non-trivial mixture of stationary point-process distributions.

Next we study, for the G/G/c queue with the FCFS queueing discipline, the work-load vector $\boldsymbol{V}=(V_1,\dots,V_c)$ introduced by Kiefer and Wolfowitz [19]. For $\mu\in\mathcal{M}$ as given in (2.1), for each $\boldsymbol{y}\in\mathbb{R}^c_+$ and $n\in\{1,2,\dots,k_1\}$ define

$$
\begin{aligned}
&\boldsymbol{\vartheta}(\boldsymbol{y};n,\mu) \qquad\qquad (3.1)\\
&= \Re\Big(\dots\Re\Big(\Re(\Re(\boldsymbol{y})+\boldsymbol{i}m_{-n}+\bar{\boldsymbol{e}}(t_{-n}-t_{1-n}))_+ +\boldsymbol{i}m_{1-n}+\bar{\boldsymbol{e}}(t_{1-n}-t_{2-n})\Big)_+ \\
&\qquad \dots+\boldsymbol{i}m_{-1}+\bar{\boldsymbol{e}}t_{-1}\Big)_+,
\end{aligned}
$$

where $\boldsymbol{i}, \bar{\boldsymbol{e}} \in \mathbb{R}_+^c$ with $\boldsymbol{i} = (1, 0, \ldots, 0)$ and $\bar{\boldsymbol{e}} = (1, 1, \ldots, 1)$, $\mathbb{R}_+^c \ni \boldsymbol{y}_+ = (y_{1+}, \ldots, y_{c+})$ with $y_{i+} = \max(0, y_i)$, and the function $\Re : \mathbb{R}_+^c \to \mathbb{R}_+^c$ rearranges the components of its c-dimensional vector argument in ascending order. Note that $\boldsymbol{\vartheta}(\boldsymbol{y}; n, \mu)$ can be seen as the work-load vector at time 0, given that only n customers arriving before 0 are taken into consideration and that the $(-n)$th arriving customer finds the system in state $\Re(\boldsymbol{y})$. We define the *work-load vector* in G/G/c as

$$\boldsymbol{V}(\boldsymbol{y}; \mu) = \sup_{n>0} \boldsymbol{\vartheta}(\boldsymbol{y}; n, \mu). \tag{3.2}$$

By using the standard arguments (see e.g. Baccelli and Brémaud [3] or Franken, König, Arndt and Schmidt [14]) it is easy to see that for a stationary ergodic input $(N, \mathbf{P})$ with $\rho < 1$, the random vector $\boldsymbol{V}(\boldsymbol{y}; N)$ is well-defined for all $\boldsymbol{y}$. Moreover if $\rho < 1/c$ then $\boldsymbol{V}(\boldsymbol{y}; N)$ does not depend on $\boldsymbol{y}$, i.e.

$$\boldsymbol{V}(N) \equiv \boldsymbol{V}(\boldsymbol{y}; N) \quad \mathbf{P}\text{−a.s.} \tag{3.3}$$

We also define the *total work load* $V(\boldsymbol{y}; \mu)$ as the sum of the components of $\boldsymbol{V}(\boldsymbol{y}; \mu)$; i.e.,

$$V(\boldsymbol{y}; \mu) = V_1(\boldsymbol{y}; \mu) + \ldots + V_c(\boldsymbol{y}; \mu) \tag{3.4}$$

and, for $\rho < 1/c$,

$$V(N) = V_1(N) + \ldots + V_c(N). \tag{3.5}$$

Note that the smallest component $V_1(N)$ of $\boldsymbol{V}(N)$ is the *waiting time* which a hypothetical customer would have to wait until service beginning if he arrived at time 0.

3.2 FME Kernels

We now calculate FME kernels (defined by (2.5)) of the queueing characteristics considered above. Here and in the sequel let $\boldsymbol{s} = (s_1, \ldots, s_k) \in \mathbb{R}_+^k$, $\boldsymbol{x} = (x_1, \ldots, x_k) \in \mathbb{R}_-^k$ and $x_1 < \ldots < x_k < 0$ which will be denoted by $(\boldsymbol{x}, \boldsymbol{s}) \in (\mathbb{R}_- \times \mathbb{R}_+)_{\text{ord}}^k$ for short. Thus for the first FME kernel of the work-load vector $\psi = \boldsymbol{V}$ we have

$$\begin{aligned} \boldsymbol{V}_{(1)}^{x,s}(\mu) \;=\; & \boldsymbol{\vartheta}\Big(\big(0, \ldots, 0, (s + x - t_{-\mu((x,0)\times\mathbb{R})})_+\big); \mu((x,0) \times \mathbb{R}), \mu\Big) \\ & - \boldsymbol{\vartheta}\Big((0, \ldots, 0, 0); \mu((x,0) \times \mathbb{R}), \mu\Big), \end{aligned} \tag{3.6}$$

where μ is as in (2.1), $x < 0$, $s \geq 0$ and we put $t_{-0} = 0$ and $\vartheta(\boldsymbol{y}; 0, \mu) = \boldsymbol{y}$. Then the first null kernel of the work-load vector is

$$\boldsymbol{V}_{(1)}^{x,s}(o) = (0, \ldots, 0, (s + x)_+). \tag{3.7}$$

In general, the kth FME kernel of $\boldsymbol{V}$, for $1 \leq k \leq c$, has the form

$$\begin{aligned} \boldsymbol{V}_{(k)}^{\boldsymbol{x},\boldsymbol{s}}(\mu) \;=\; & \sum_{j=0}^{k} (-1)^{k-j} \sum_{\pi \in \{\binom{k}{j}\}} \\ & \boldsymbol{\vartheta}\Big(\big(\overbrace{0, \ldots, 0}^{c-j}, (s_i + x_i - t_{-\mu((x_k,0)\times\mathbb{R})})_+ : \; i \in \pi\big); \mu((x_k, 0) \times \mathbb{R}), \mu\Big) \end{aligned} \tag{3.8}$$

and consequently the kth null kernel for $1 \le k \le c$ is

$$\boldsymbol{V}_{(k)}^{\boldsymbol{x},\boldsymbol{s}}(o) = \left(\overbrace{0,\ldots,0}^{c-k}, \xi_{(k)}^{\boldsymbol{x},\boldsymbol{s}}, -\binom{k-1}{1}\xi_{(k)}^{\boldsymbol{x},\boldsymbol{s}}, \ldots, (-1)^{k-2}\binom{k-1}{k-2}\xi_{(k)}^{\boldsymbol{x},\boldsymbol{s}}, (-1)^{k-1}\xi_{(k)}^{\boldsymbol{x},\boldsymbol{s}}\right) \tag{3.9}$$

with

$$\xi_{(k)}^{\boldsymbol{x},\boldsymbol{s}} = \min_{1\le i\le k}(s_i + x_i)_+. \tag{3.10}$$

Note that (3.9) follows from (3.1) and (3.8) by (3.7) and by induction with respect to k. Furthermore for the null kernels of the total work load we have

$$V_{(1)}^{x,s}(o) \quad = \quad (s+x)_+ \tag{3.11}$$

and

$$V_{(k)}^{\boldsymbol{x},\boldsymbol{s}}(o) \quad \equiv \quad 0 \qquad (k = 0, 2, 3, \ldots, c). \tag{3.12}$$

We also present the FME kernel of order $c+1$ for the work-load vector. For this we define

$$\boldsymbol{\eta}_{(c+1)}^{\boldsymbol{x},\boldsymbol{s}} = \Re\Big(\Re\big((s_1 + x_1 - x_{c+1})_+, \ldots, (s_c + x_c - x_{c+1})_+\big) + \boldsymbol{i}s_{c+1}\Big). \tag{3.13}$$

Then we have

$$\begin{aligned}
\boldsymbol{V}_{(c+1)}^{\boldsymbol{x},\boldsymbol{s}}(\mu) \quad &= \quad \boldsymbol{\vartheta}\Big(\big(\,\boldsymbol{\eta}_{(c+1)}^{\boldsymbol{x},\boldsymbol{s}} + \bar{\boldsymbol{e}}(x_{c+1} - t_{-\mu((x_{c+1},0)\times\mathbb{R})})\big)_+; \mu((x_{c+1},0)\times\mathbb{R}), \mu\Big) \\
&\quad + \sum_{j=0}^{c}(-1)^{c+1-j} \sum_{\pi\in\{\binom{c+1}{j}\}} \\
&\quad \boldsymbol{\vartheta}\Big(\Re\big(\overbrace{0,\ldots,0}^{c-j}, (s_i + x_i - t_{-\mu((x_{c+1},0)\times\mathbb{R})})_+ :\ i \in \pi\big); \mu((x_{c+1},0)\times\mathbb{R}), \mu\Big).
\end{aligned} \tag{3.14}$$

Finally, from (3.9) and (3.14) we get that the $(c+1)$th null kernel of total work load is given by

$$\begin{aligned}
&V_{(c+1)}^{\boldsymbol{x},\boldsymbol{s}}(o) \\
&= \left(\min_{1\le i\le c}(s_i + x_i - x_{c+1})_+ + s_{c+1} + x_{c+1}\right)_+ - \min_{1\le i\le c}(s_i + x_i)_+ - (s_{c+1} + x_{c+1})_+ .
\end{aligned} \tag{3.15}$$

4 LIGHT-TRAFFIC APPROXIMATIONS

Suppose now that the input of the considered multi-server queue is a stationary ergodic Markov-modulated process (see Section 2.3) with arrival rates β_i, such that $\sum_{i=1}^{\ell}\pi_i\beta_i = 1$, with factorial moment measures $M^{(k)}$ and with mark distributions B_i. In this case we use the notation MM/G/c. Next, instead of the arrival rates β_i we consider the rescaled arrival rates $\beta_i^* = \beta^*\beta_i$. We denote the resulting rescaled Markov-modulated point process by $(N, \mathbf{P}^{\beta^*})$. Furthermore denote by $\mathbf{E}^{\beta^*}$

the expectation taken with respect to $\mathbf{P}^{\beta^*}$ and note that β^* is the average arrival intensity.

The main subject of the present section is to investigate the behavior of the work-load vector when $\beta^* \to 0$. In particular, we assume that β^* is sufficiently small to fulfill

$$\frac{1-\epsilon}{\beta^*\beta} > \int_0^\infty s\, B(ds) \tag{4.1}$$

for some $\epsilon > 0$, where B is the distribution given by

$$B([0,s)) = \min_{1\le i\le \ell} B_i([0,s)), \quad \text{and} \quad \beta = \max_{1\le i\le \ell} \beta_i \, .$$

Denote by $\overline{\mathbf{P}}^{\beta^*}$ the distribution of an independently marked Poisson p.p. with intensity measure $\beta^*\beta(1-\epsilon)^{-1}\, B(ds)dx$. By $\overline{\mathbf{E}}^{\beta^*}$ we denote the corresponding expectation. Furthermore, we consider the quantity

$$W = W(\mu) = \sup_{x<0} \int_{\mathbb{R}_+} s\, \mu((x,0)\times ds) + x, \tag{4.2}$$

which can be interpreted as a work load in a single-server queue. We will utilize condition (4.1) to obtain the following approximations.

Theorem 4.1 *Let $f : \mathbb{R}_+ \to \mathbb{R}_+$ be a nondecreasing function, such that $f(0) = 0$ and $f(x) \le C + Dx^\alpha$ for some $C, D, \alpha \ge 0$. If $\int_0^\infty s^{\alpha+\max(k,2)}\, B(ds) < \infty$ then for the d-th component of the stationary work-load vector in the* MM/G/c *queue we have*

$$\begin{aligned}
&\mathbf{E}^{\beta^*} f(V_d(N)) \\
&= \sum_{i=c+1-d}^{k} (\beta^*)^i (-1)^{i+d-c-1} \binom{i-1}{i+d-c-1} \int_{(\mathbb{R}_-\times\mathbb{R}_+)^i_{\text{ord}}} f(\xi_{(i)}^{\boldsymbol{x},\boldsymbol{s}})\, M^{(i)}(d(\boldsymbol{x},\boldsymbol{s})) \\
&\qquad + \mathrm{o}((\beta^*)^k)
\end{aligned} \tag{4.3}$$

for each $k \in \{c+1-d,\dots,c\}$, $d \in \{1,\dots,c\}$, where $\xi_{(i)}$, $M^{(i)}$ are given by (3.10), (2.9).

Remark. For the GI/G/c queue the $(c+1-d)$th term of (4.3) has been derived in Daley and Rolski [13]. Observe that by summing up both sides of (4.3) over all $d \in \{1,\dots,c\}$ we get for the total work load in MM/G/c

$$\begin{aligned}
\mathbf{E}^{\beta^*} V(N) &= \beta^* \int_{\mathbb{R}_-\times\mathbb{R}_+} (s+x)_+\, M^{(1)}(d(x,s)) + \mathrm{o}((\beta^*)^c) \\
&= \beta^* \sum_{i=1}^{\ell} \pi_i \beta_i \int_0^\infty \frac{s^2}{2}\, B_i(ds) + \mathrm{o}((\beta^*)^c)
\end{aligned}$$

provided that $\int_0^\infty s^{\max(c,2)+1}\, B(ds) < \infty$.

Under a somewhat stronger condition, a refinement of this light-traffic approximation is given in the following theorem.

Theorem 4.2 *For the stationary total work load in the* MM/G/c *queue we have*

$$\begin{aligned}\mathbf{E}^{\beta^*} V(N) &= \beta^* \sum_{i=1}^{\ell} \pi_i \beta_i \int_0^\infty \frac{s^2}{2}\, B_i(ds) && (4.4)\\ &\quad + (\beta^*)^{c+1} \int_{(\mathbb{R}_-\times\mathbb{R}_+)^{c+1}_{\text{ord}}} V^{\boldsymbol{x},\boldsymbol{s}}_{(c+1)}(o)\, M^{(c+1)}(d(\boldsymbol{x},\boldsymbol{s})) + \mathrm{o}((\beta^*)^{c+1})\end{aligned}$$

provided that $\int_0^\infty s^{c+2}\, B(ds) < \infty$, *where* $V^{\boldsymbol{x},\boldsymbol{s}}_{(c+1)}(o)$ *and* $M^{(c+1)}$ *are given by* (3.15), (2.9) *respectively.*

Now under a much stronger condition on the distribution B, which in particular implies finiteness of its all moments, we show that the expected value of each component of the workload-vector in the MM/G/c queue is an analytic function of the arrival intensity. Concerning a related result for finite-horizon functionals driven by a Poisson process we refer to Zazanis [38].

Theorem 4.3 *The expected value of each component of the stationary work-load vector in the* MM/G/c *queue is an analytic function of the arrival intensity* β^* *at* $\beta^* = 0+$, *and its power series is convergent for each* $\beta^* < 1/(\beta \int_0^\infty s\, B(ds))$ *whenever*

$$\lim_{n\to\infty} \frac{(2\beta^*)^{n-1}}{(n-1)!} \int_{\mathbb{R}^n} \Big(\sum_{i=1}^{n} s_i\Big)^{n+1} B(ds_1)\dots B(ds_n) = 0. \qquad (4.5)$$

Proof of Theorem 4.1. By Theorem 2.1 we have to check the finiteness condition (2.7) for each $i \in \{1,\dots,k\}$ and to prove that

$$\begin{aligned}&\lim_{\beta^*\to 0} \int_{(\mathbb{R}_-\times\mathbb{R}_+)^k_{\text{ord}}} \int_{\mathcal{M}} \Big(f(V_d)\Big)^{\boldsymbol{x},\boldsymbol{s}}_{(k)}(\mu)\, (\mathbf{P}^{\beta^*})^{(k)}_{\boldsymbol{x},\boldsymbol{s}}(d\mu) M^{(k)}(d(\boldsymbol{x},\boldsymbol{s})) && (4.6)\\ &= (-1)^{k+d-c-1} \binom{k-1}{k+d-c-1} \int_{(\mathbb{R}_-\times\mathbb{R}_+)^k_{\text{ord}}} f(\xi^{\boldsymbol{x},\boldsymbol{s}}_{(k)})\, M^{(k)}(d(\boldsymbol{x},\boldsymbol{s})).\end{aligned}$$

Let $i \geq 1$. Observe by (3.1), (3.8) and (3.2), (3.3) that

$$\begin{aligned}\left|\Big(f(V_d)\Big)^{\boldsymbol{x},\boldsymbol{s}}_{(i)}(\mu)\right| &\leq 2^i f\Big(W(\mu) + \sum_{l=1}^{i} s_l\Big) \mathbf{1}\Big(x_i - \sum_{l=1}^{i} s_l < x_1 < x_2 \ldots < x_i < 0\Big)\\ &\quad \times \mathbf{1}\Big(\sum_{l=1}^{i} s_l + \int_{\mathbb{R}_+} s\, \mu((x_i,0)\times ds) + x_i > 0\Big) && (4.7)\end{aligned}$$

where we have used the fact that $\sum\limits_{j=0}^{k} \binom{k}{j} = 2^k$. Moreover, the last indicator in (4.7) is bounded by

$$\mathbf{1}\Big(\sum_{l=1}^{i} s_l + \sup_{x<0}\Big\{\int_{\mathbb{R}_+} s\, \mu((x,0)\times ds) + (1-\epsilon)x\Big\} > -\epsilon x_i\Big).$$

Observe that the right-hand side (RHS) of (4.7) is increasing in μ, i.e. with respect to adding points and enlarging marks in μ. Thus for its expectation we have

$$\int_{\mathcal{M}_s} (\ldots) \mathbf{P}^{\beta^*}{}^{(i)}_{\boldsymbol{x},\boldsymbol{s}}(d\mu) \le \int_{\mathcal{M}_s} (\ldots) \overline{\mathbf{P}}^{\beta^*}(d\mu), \tag{4.8}$$

because, by Theorem 2.3 the distribution $(\mathbf{P}^{\beta^*})^{(i)}_{\boldsymbol{x},\boldsymbol{s}}$ generates neither more points nor larger marks than $\overline{\mathbf{P}}^{\beta^*}$. Now changing the order of integration, using the representation formulae (2.9) for the factorial moment measure $M^{(i)}$ and integrating with respect to $dx_1 \ldots dx_{i-1}$ we see that for (2.7) it suffices to provide finiteness of

$$\overline{\mathbf{E}}^{\beta^*} \int_{\mathbb{R}^i_+} f\Big(W + \sum_{l=1}^{i} s_l\Big)\Big(\sum_{l=1}^{i} s_l\Big)^{i-1} \int_{\mathbb{R}_-} \mathbf{1}\Big(\sum_{l=1}^{i} s_l + W > -\epsilon x_i\Big)\, dx_i\, B(ds_1) \ldots B(ds_i)$$

for which the assumptions of Theorem 4.1 are enough. To justify (4.6), having (4.7), (4.8) and the finiteness conditions, apply Lebesgue's dominated convergence theorem for passing with the limit under the outer integral. Next apply Levi's monotone convergence theorem for the integral with respect to $(\mathbb{P}^{\beta^*})^{(i)}_{\boldsymbol{x},\boldsymbol{s}}$ with $\beta^* \to 0$ (formally using the fact that $(\mathbb{P}^{\beta^*})^{(i)}_{\boldsymbol{x},\boldsymbol{s}}$, conditioned on $\tilde{X}(\cdot)$ of Theorem 2.3, can be seen as the distribution of an independent thinning of a Poisson process, with retention probability β^*, and that one can construct a probability space on which the convergence holds with probability 1). Then (4.6) follows by (3.9). □

Remark. Note that a sufficient condition for the approximation (4.3) with a general increasing function f such that $f(0) = 0$ is a somewhat cumbersome condition

$$\int_{\mathbb{R}^i_+} \overline{\mathbf{E}}^{\beta^*}\Big[f\Big(W + \sum_{l=1}^{i} s_l\Big)\Big(\sum_{l=1}^{i} s_l\Big)^{i-1}\Big(\sum_{l=1}^{i} s_l + W\Big)\Big]\, B(ds_1) \ldots B(ds_i) < \infty.$$

Proof of Theorem 4.2. Use (3.14), (3.15) and apply the same arguments as in the proof of Theorem 4.1. □

Proof of Theorem 4.3. By Theorem 2.1 we have to verify the finiteness condition (2.7) for all $i \in \mathbb{N}$ and show that the remainder term in (2.8) tends to 0 when $n \to \infty$. Since now, in comparison to Theorem 4.1, we consider the special case $f(x) = x$, we are able to give a little more careful bound than (4.7). Namely, for each component d of the work-load vector $\boldsymbol{V}$ we have

$$\begin{aligned} \Big|(V_d)^{\boldsymbol{x},\boldsymbol{s}}_{(n)}(\mu)\Big| \;\le\; & 2^{n-1} s_1 \mathbf{1}\Big(x_n - \sum_{l=1}^{n} s_l < x_1 < x_2 \ldots < x_n < 0\Big) \\ & \times \mathbf{1}\Big(\sum_{l=1}^{n} s_l + \sup_{x<0}\Big\{\int_{\mathbb{R}_+} s\, \mu((x,0) \times ds) + (1-\epsilon)x\Big\} > -\epsilon x_n\Big). \end{aligned} \tag{4.9}$$

Now using analogous arguments as in the proof of Theorem 4.1 we get the inequality

$$\int_{(\mathbb{R}_- \times \mathbb{R}_+)^n_{\text{ord}}} \int_{\mathcal{M}} \Big|(V_d)^{\boldsymbol{x},\boldsymbol{s}}_{(n)}(\mu)\Big|\, \mathbf{P}^{\beta^*}{}^{(n)}_{\boldsymbol{x},\boldsymbol{s}}(d\mu) M^{(n)}(d(\boldsymbol{x},\boldsymbol{s})) \tag{4.10}$$

$$\leq \frac{2^n}{\epsilon(n-1)!} \int_{\mathbb{R}^n} \Big(\sum_{l=1}^n s_l\Big)^n \Big(\sum_{l=1}^n s_l + \overline{\mathbf{E}}^{\beta^*} W\Big)\, B(ds_1) \dots B(ds_n).$$

We see that for (2.7) it is enough to prove the the finiteness of the RHS of (4.10) for all n. Then the remainder term in (2.8) vanishes if the RHS of (4.10) times $(\beta^*)^n$ tends to 0 when $n \to \infty$. $\square$

5 COMPUTATIONAL RESULTS

5.1 Phase-Type Distributions

We apply now the expansions derived in the preceding section together with the method of phase-type distributions (see Neuts [24], [25]) in order show how numerical results can be obtained for MM/PH/c queues in light traffic. A distribution B on $\mathbb{R}_+$ is said to be of *phase-type* (PH) with representation $(\boldsymbol{\alpha}, \boldsymbol{T}, d)$ if there is a homogeneous Markov process on the state space $\{0, 1, \dots, d\}$, with intensity matrix

$$\left(\begin{array}{c|c} 0 & \mathbf{0} \\ \hline \boldsymbol{t}_0 & \boldsymbol{T} \end{array}\right) = \left(\begin{array}{c|cccc} 0 & 0 & 0 & \dots & 0 \\ \hline t_{10} & t_{11} & t_{12} & \dots & t_{1d} \\ t_{20} & t_{21} & t_{22} & \dots & t_{2d} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ t_{d0} & t_{d1} & t_{d2} & \dots & t_{dd} \end{array}\right),$$

and with 0 as the absorbing state, such that B is the distribution of the time to absorption in 0 given the initial distribution

$$\Big(1 - \sum_{i=1}^d \alpha_i, \boldsymbol{\alpha}\Big), \qquad \text{where } \boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_d).$$

We assume for $\boldsymbol{T}$ that the states $1, \dots, d$ are all transient, so that the absorption into the state 0 from any initial state is certain. Note that the states $1, \dots, d$ are transient if and only if the $(d \times d)$-matrix $\boldsymbol{T}$ is non-singular. Fundamental identities are

$$1 - B(s) = \boldsymbol{\alpha} e^{\boldsymbol{T} s} \boldsymbol{e}, \tag{5.1}$$

$$\int_0^\infty s^n\, B(ds) = (-1)^n n!\, \boldsymbol{\alpha} \boldsymbol{T}^{-n} \boldsymbol{e}, \tag{5.2}$$

where $e^{\boldsymbol{T}}$ is the matrix exponential of $\boldsymbol{T}$ (see (2.10) in Section 2.3) and $\boldsymbol{e}$ is the column-vector of ones of appropriate size (see Section 2.2 of Neuts [24]).

Besides the standard matrix operations as sum $\boldsymbol{A} + \boldsymbol{B}$ and product $\boldsymbol{AB}$ (in the following matrices are denoted in bold capital letters) we consider the Kronecker product of a $(k_1 \times m_1)$-matrix $\boldsymbol{A}$ and a $(k_2 \times m_2)$-matrix $\boldsymbol{B}$ which is a $(k_1 k_2) \times (m_1 m_2)$-matrix, denoted by $\boldsymbol{A} \otimes \boldsymbol{B}$, with

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \dots & a_{1m_1}\boldsymbol{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k_1 1}\boldsymbol{B} & a_{k_1 2}\boldsymbol{B} & \dots & a_{k_1 m_1}\boldsymbol{B} \end{pmatrix}.$$

Furthermore if $\boldsymbol{A}$ and $\boldsymbol{B}$ are any matrices the Kronecker sum is given by

$$\boldsymbol{A} \oplus \boldsymbol{B} = \boldsymbol{A} \otimes \boldsymbol{I}_{\boldsymbol{B}} + \boldsymbol{I}_{\boldsymbol{A}} \otimes \boldsymbol{B},$$

where $\boldsymbol{I}_{\boldsymbol{A}}, \boldsymbol{I}_{\boldsymbol{B}}$ is the identity matrix of size $\boldsymbol{A}$ and $\boldsymbol{B}$ respectively.

In the following we assume that the distributions B_i of service times in the MM/G/c queue considered in Section 4 are PH-distributions with representations $(\boldsymbol{\alpha}_i, \boldsymbol{T}_i, d_i)$. Note that the crucial part of the light-traffic approximation derived in Section 4 is the relatively simple form of the function $\xi_{(k)}^{(\boldsymbol{x},\boldsymbol{s})}$ given by (3.10). In (4.3) this function is integrated with respect to the factorial moment measure $M^{(k)}$. When $\boldsymbol{x}$ is fixed then (see Theorem 2.2) $M^{(k)}$ is a certain mixture of k-products of B_i $(i = 1, \ldots, \ell)$. In connection to this we will use the following auxiliary result (for the special case $\boldsymbol{x} = 0$, cf. also Corollary 2.1 in Burman and Smith [10]).

Lemma 5.1 *Fix $\boldsymbol{x} \in \mathbb{R}_-^k$ and suppose that the random variables S_i $(i = 1, \ldots, k)$ are independent and have PH distributions with representations $(\boldsymbol{\alpha}_{j_i}, \boldsymbol{T}_{j_i}, d_{j_i})$ where $j_i \in \{1, \ldots, \ell\}$. Then*

$$\xi_{(k)}^{\boldsymbol{x},\boldsymbol{S}} = \min_{1 \le i \le k} (S_i + x_i)_+$$

has the PH-distribution with representation $(\boldsymbol{\alpha}_{\boldsymbol{j}}(\boldsymbol{x}), \boldsymbol{T}_{\boldsymbol{j}}, d_{\boldsymbol{j}})$ given by

$$\begin{aligned}
\boldsymbol{\alpha}_{\boldsymbol{j}}(\boldsymbol{x}) &= \bigotimes_{i=1}^{k} \boldsymbol{\alpha}_{j_i} e^{-\boldsymbol{T}_{j_i} x_i}, \\
\boldsymbol{T}_{\boldsymbol{j}} &= \bigoplus_{i=1}^{k} \boldsymbol{T}_{j_i}, \\
d_{\boldsymbol{j}} &= \prod_{i=1}^{k} d_{j_i},
\end{aligned}$$

where $\boldsymbol{j} = (j_1, \ldots, j_k)$.

Proof. We use the probabilistic interpretation of a PH-distribution as the distribution of the absorption time in a certain Markov jump process. Namely we consider the vector of Markov jump processes corresponding to $S_1, \ldots, S_k$. We start running these processes at $x_1, \ldots, x_k < 0$ respectively, and follow their phase states on a common product space

$$\bigtimes_{i=1}^{k} \{1, \ldots, d_{j_i}\}$$

completed with one absorbing state. Note that $\boldsymbol{\alpha}_{\boldsymbol{j}}(\boldsymbol{x})$ is a defective probability distribution which gives the probabilities of reaching time 0 with certain configuration of phase states. We treat $\boldsymbol{\alpha}_{\boldsymbol{j}}(\boldsymbol{x})$ as the (defective) initial distribution for the PH-representation of $\xi_{(k)}^{\boldsymbol{x},\boldsymbol{S}}$. Similarly, since each entry of $\boldsymbol{T}_{\boldsymbol{j}}$ (except the diagonal) is the intensity of some state transition of certain (one) S_i we interpret $\boldsymbol{T}_{\boldsymbol{j}}$ as the sub-intensity matrix of $\xi_{(k)}^{\boldsymbol{x},\boldsymbol{S}}$. Obviously $\xi_{(k)}^{\boldsymbol{x},\boldsymbol{S}}$ is the time to the first absorption after time 0 among $S_1, \ldots, S_k$ and, consequently the assertion follows. $\square$

Now we can determine the coefficients in the factorial moment expansion (4.3) for the special case $f(x) = x^n$, which can be used to compute approximations of all moments of the stationary work-load vector in multi-server queues. For making our arguments more transparent we first consider the M/PH/c queue with an independently marked Poisson input.

Theorem 5.1 *If $B_1 = \ldots = B_\ell = B$ with representation $(\boldsymbol{\alpha}, \boldsymbol{T}, d)$ and $\beta_1 = \ldots = \beta_\ell = 1$ then*

$$\begin{aligned}&\int_{(\mathbb{R}_-\times\mathbb{R}_+)^k_{\mathrm{ord}}} \left(\xi_{(k)}^{\boldsymbol{x},\boldsymbol{s}}\right)^n M^{(k)}(d(\boldsymbol{x},\boldsymbol{s}))\\ &= (-1)^{n+k} n! \Big(\Big(\ldots(\boldsymbol{\alpha}\boldsymbol{T}^{-1}\otimes\boldsymbol{\alpha})(\boldsymbol{T}\oplus\boldsymbol{T})^{-1}\ldots\otimes\boldsymbol{\alpha}\Big)\Big(\bigoplus_{i=1}^{k-1}\boldsymbol{T}\Big)^{-1}\otimes\boldsymbol{\alpha}\Big)\Big(\bigoplus_{i=1}^{k}\boldsymbol{T}\Big)^{-1-n}\boldsymbol{e}\end{aligned}$$

for $n \in \mathbb{N}$ where $\boldsymbol{e}$ is the column vector of ones of size d^k.

Proof. In the Poisson case $M^{(k)}(d(\boldsymbol{x},\boldsymbol{s})) = B(ds_1)\ldots B(ds_k)\,dx_1\ldots dx_k$. We integrate $\left(\xi_{(k)}^{\boldsymbol{x},\boldsymbol{s}}\right)^n$ according to $B(ds_1)\ldots B(ds_k)$ first and obtain by Lemma 5.1 and (5.2)

$$\int_{\mathbb{R}_+^k} \left(\xi_{(k)}^{\boldsymbol{x},\boldsymbol{s}}\right)^n B(ds_1)\ldots B(dx_k) = (-1)^n n! \bigotimes_{i=1}^{k} \boldsymbol{\alpha} e^{-\boldsymbol{T}x_i}\Big(\bigoplus_{i=1}^{k}\boldsymbol{T}\Big)^{-n}\boldsymbol{e},$$

where $\boldsymbol{x} \in \mathbb{R}_-^k$. Now

$$\begin{aligned}\int_{-\infty}^{x_2} \bigotimes_{i=1}^{k} \boldsymbol{\alpha} e^{-\boldsymbol{T}x_i}\,dx_1 &= \int_{-\infty}^{x_2} \boldsymbol{\alpha} e^{-\boldsymbol{T}x_1}\,dx_1 \otimes \bigotimes_{i=2}^{k} \boldsymbol{\alpha} e^{-\boldsymbol{T}x_i}\\ &= \boldsymbol{\alpha}(-\boldsymbol{T}^{-1})e^{-\boldsymbol{T}x_2} \otimes \bigotimes_{i=2}^{k} \boldsymbol{\alpha} e^{-\boldsymbol{T}x_i}\\ &= -(\boldsymbol{\alpha}\boldsymbol{T}^{-1}\otimes\boldsymbol{\alpha})e^{-(\boldsymbol{T}\oplus\boldsymbol{T})x_2} \otimes \bigotimes_{i=3}^{k} \boldsymbol{\alpha} e^{-\boldsymbol{T}x_i}\end{aligned}$$

Where for the last equality we used the identities $\boldsymbol{C}e^{\boldsymbol{A}} \otimes \boldsymbol{D}e^{\boldsymbol{B}} = (\boldsymbol{C}\otimes\boldsymbol{D})(e^{\boldsymbol{A}} \otimes e^{\boldsymbol{B}}) = (\boldsymbol{C}\otimes\boldsymbol{D})e^{\boldsymbol{A}\oplus\boldsymbol{B}}$. Iteration of this procedure completes the proof. $\square$

It turns out that Theorem 5.1 can be extended to the case of a Markov-modulated input. This provides a computational technique for getting approximations of moments of the stationary work-load vector in general MM/PH/c queues. In the statement of this result we use the symbol $[\boldsymbol{\Lambda}\oplus\boldsymbol{A}]_{jj'}$ to denote the (jj') block of size $\boldsymbol{A}$ of the matrix $\boldsymbol{\Lambda}\oplus\boldsymbol{A}$, i.e. it contains the entries of $\boldsymbol{\Lambda}\oplus\boldsymbol{A}$ with indices from the set $[(j-1)d+1,\ldots,jd] \times [(j'-1)d+1,\ldots,j'd]$ where $\boldsymbol{A}$ is a $d\times d$ matrix.

Theorem 5.2 *Let $M^{(k)}$ be the k-th factorial moment measure of a Markov-modulated m.p.p. with intensities $\beta_1,\ldots,\beta_\ell$ and a random environment given by $(\boldsymbol{\pi},\boldsymbol{\Lambda})$,* *and let the B_i be PH-distributions with representations $(\boldsymbol{\alpha}_i, \boldsymbol{T}_i, d_i)$ $(i = 1, \ldots, \ell)$. Then we have*

$$\begin{aligned}
&\int_{(\mathbb{R}_-\times\mathbb{R}_+)^k_{\mathrm{ord}}} (\xi^{\boldsymbol{x},\boldsymbol{s}}_{(k)})^n \, M^{(k)}(d(\boldsymbol{x},\boldsymbol{s})) \\
&= (-1)^{n+k} n! \sum_{j_1,\ldots,j_k=1}^{\ell} \pi_{j_1} \Big(\prod_{i=1}^{k} \beta_{j_i}\Big) \\
&\qquad \Big(\Big(\ldots(\boldsymbol{\alpha}_{j_1}[(\boldsymbol{\Lambda}\oplus\boldsymbol{T}_{j_1})^{-1}]_{j_1j_2}\otimes\boldsymbol{\alpha}_{j_2})[(\boldsymbol{\Lambda}\oplus\boldsymbol{T}_{j_1}\oplus\boldsymbol{T}_{j_2})^{-1}]_{j_2j_3} \\
&\qquad \ldots\otimes\boldsymbol{\alpha}_{j_{k-1}}\Big)\Big[\Big(\boldsymbol{\Lambda}\oplus\bigoplus_{i=1}^{k-1}\boldsymbol{T}_{j_i}\Big)^{-1}\Big]_{j_{k-1}j_k}\otimes\boldsymbol{\alpha}_{j_k}\Big)\Big(\bigoplus_{i=1}^{k}\boldsymbol{T}_{j_i}\Big)^{-1-n}\boldsymbol{e}_{j_1\cdots j_k},
\end{aligned}$$

where $\boldsymbol{e}_{j_1\cdots j_k}$ is the appropriate column vector of 1's.

Proof. The proof is similar to that of Theorem 5.1. From Theorem 2.2, Lemma 5.1 and (5.2) we get that

$$\begin{aligned}
&\int_{\mathbb{R}^k_+} \Big(\xi^{\boldsymbol{x},\boldsymbol{s}}_{(k)}\Big)^n \, M^{(k)}(d(\boldsymbol{x},\boldsymbol{s})) = (-1)^n n! \sum_{j_1,\ldots,j_k=1}^{\ell} \pi_{j_1}\Big(\prod_{i=1}^{k}\beta_{j_i}\Big) \\
&\quad \int_{-\infty}^{0}\cdots\int_{-\infty}^{x_2}\Big(\bigotimes_{i=1}^{k}\boldsymbol{\alpha}_{j_i}e^{-\boldsymbol{T}_{j_i}x_i}\Big)\prod_{i=1}^{k-1}\Big(e^{\boldsymbol{\Lambda}(x_{i+1}-x_i)}\Big)_{j_ij_{i+1}} dx_1\ldots dx_k\Big(\bigoplus_{i=1}^{k}\boldsymbol{T}_{j_i}\Big)^{-n}\boldsymbol{e}_{j_1\ldots j_k}.
\end{aligned}$$

For the inner integral we have

$$\begin{aligned}
\int_{-\infty}^{x_2}\boldsymbol{\alpha}_{j_1}e^{-\boldsymbol{T}_{j_1}x_1}\Big(e^{\boldsymbol{\Lambda}(x_2-x_1)}\Big)_{j_1j_2}dx_1 &= \Big[\int_{-\infty}^{x_2}\Big(\boldsymbol{I}_{\boldsymbol{\Lambda}}\otimes\boldsymbol{\alpha}_{j_1}\Big)e^{\boldsymbol{\Lambda}(x_2-x_1)\oplus(-\boldsymbol{T}_{j_1}x_1)}\,dx_1\Big]_{j_1j_2} \\
&= -\boldsymbol{\alpha}_{j_1}[(\boldsymbol{\Lambda}\oplus\boldsymbol{T}_{j_1})^{-1}]_{j_1j_2}e^{-\boldsymbol{T}_{j_1}x_2}.
\end{aligned}$$

Again, iteration completes the proof. □

Remark. Note that the expressions derived in Theorems 5.1 and 5.2 depend on n only through the factor $\Big[\Big(\bigoplus_{i=1}^{k}\boldsymbol{T}_{j_i}\Big)^{-1}\Big]^{n+1}$. Thus, the results obtained in the present section yield a convenient technique for computing light-traffic approximations for *all* moments of the work-load vector. Moreover, since $\boldsymbol{\alpha}(\boldsymbol{x})$ is the only term depending on $\boldsymbol{x}$ in the PH-representation given in Lemma 5.1, such approximations can be calculated also for other characteristics of the work-load vector, e.g. for the tail function — using (5.1).

Next we determine the coefficient of $(\beta^*)^{c+1}$ in the expansion (4.4) of the mean stationary total work load in MM/PH/c queues.

Theorem 5.3 *Under the assumptions of Theorem 5.2 we have*

$$\int_{(\mathbb{R}_-\times\mathbb{R}_+)^{c+1}_{\mathrm{ord}}} V^{\boldsymbol{x},\boldsymbol{s}}_{(c+1)}(o)\, M^{(c+1)}(d(\boldsymbol{x},\boldsymbol{s}))$$

$$= \quad (-1)^c \sum_{j_1,\dots,j_{c+1}=1}^{\ell} \pi_{j_1} \Big(\prod_{i=1}^{c+1} \beta_{j_i}\Big) \boldsymbol{\alpha}_{j_{c+1}} (\boldsymbol{T}_{j_{c+1}})^{-1} \boldsymbol{e}_{j_{c+1}}$$

$$\Big(\dots(\boldsymbol{\alpha}_{j_1}[(\boldsymbol{\Lambda} \oplus \boldsymbol{T}_{j_1})^{-1}]_{j_1 j_2} \otimes \boldsymbol{\alpha}_{j_2})[(\boldsymbol{\Lambda} \oplus \boldsymbol{T}_{j_1} \oplus \boldsymbol{T}_{j_2})^{-1}]_{j_2 j_3}$$

$$\dots \otimes \boldsymbol{\alpha}_{j_c}\Big)\Big[\Big(\boldsymbol{\Lambda} \oplus \bigoplus_{i=1}^{c} \boldsymbol{T}_{j_i}\Big)^{-1}\Big]_{j_c j_{c+1}} \Big(\bigoplus_{i=1}^{c} \boldsymbol{T}_{j_i}\Big)^{-1} \boldsymbol{e}_{j_1 \cdots j_c},$$

Proof. The $(c+1)$th null kernel $V^{\boldsymbol{x},\boldsymbol{s}}_{(c+1)}(o)$ given by (3.15) (with $(\boldsymbol{x},\boldsymbol{s}) \in (\mathbb{R}_- \times \mathbb{R}_+)^{c+1}_{\mathrm{ord}}$) can be represented in the following manner:

$$\begin{aligned} V^{\boldsymbol{x},\boldsymbol{s}}_{(c+1)}(o) \quad = \quad & \min_{1\le i\le c} (s_i + s_{c+1} + x_i)_+ \mathbf{1}(x_{c+1} < -s_{c+1}) \\ & + \min_{1\le i\le c} (s_i + x_i - x_{c+1})_+ \mathbf{1}(0 > x_{c+1} \ge -s_{c+1}) \\ & - \min_{1\le i\le c} (s_i + x_i)_+ . \end{aligned} \tag{5.3}$$

Now using (2.9) we can integrate first with respect to $dx_1 \dots dx_{c+1}$. Then, for fixed s_{c+1}, by the substitution $x_i \to x_i + s_{c+1}$ $(i = 1, \dots, c+1)$ we see that the integrals of the first and of the third terms in the RHS of (5.3) cancel each other. For the integral of the second term we apply the substitution $x_i \to x_i + x_{c+1}$ $(i = 1, \dots, c)$. Then integrating with respect to $B(ds_1) \dots B(ds_{c+1})$ by (5.2) we get

$$\int_{(\mathbb{R}_-\times\mathbb{R}_+)^{c+1}_{\mathrm{ord}}} V^{\boldsymbol{x},\boldsymbol{s}}_{(c+1)}(o)\, M^{(c+1)}(d(\boldsymbol{x},\boldsymbol{s})) = \sum_{j_1,\dots,j_{c+1}=1}^{\ell} \pi_{j_1} \Big(\prod_{i=1}^{c+1} \beta_{j_i}\Big) \boldsymbol{\alpha}_{j_{c+1}} (\boldsymbol{T}_{j_{c+1}})^{-1} \boldsymbol{e}_{j_{c+1}}$$

$$\int_{-\infty}^{0} \cdots \int_{-\infty}^{x_2} \Big(\bigotimes_{i=1}^{c} \boldsymbol{\alpha}_{j_i} e^{-\boldsymbol{T}_{j_i} x_i}\Big) \prod_{i=1}^{c-1} \Big(e^{\boldsymbol{\Lambda}(x_{i+1}-x_i)}\Big)_{j_i j_{i+1}} \Big(e^{-\boldsymbol{\Lambda} x_c}\Big)_{j_c j_{c+1}}$$

$$dx_1 \dots dx_c \Big(\bigoplus_{i=1}^{c} \boldsymbol{T}_{j_i}\Big)^{-1} e_{j_1 \dots j_c}.$$

Similar arguments like in the case of Theorem 5.2 complete the proof. □

5.2 Numerical Illustrations

Now we illustrate by some numerical examples how the algorithm, which has been developed in the preceding sections of this paper, works. In all these examples, the traffic intensity $\rho = \frac{\beta^*}{c} \sum_{j=1}^{\ell} \pi_j \beta_j \int_0^\infty s\, B_j(ds)$ does not exceed the value $\frac{1}{c}$.

First we consider the mean stationary total work load $\mathbf{E}^{\beta^*} V$ in the M/E$_2$/2 queue with Poisson arrivals. Clearly, in this case, closed formulas are well-known. Using Theorems 4.1 and 5.1 we calculate approximations for $\mathbf{E}^{\beta^*} V$ and compare these values with the exact ones taken from [16]. It turns out that, not only for low traffic intensities, the differences are surprisingly small. Next, we consider expectation of the stationary total work load and second moment of the stationary work-load vector in the MM/E$_2$/2 queue comparing the approximations, obtained by using

Theorems 4.1 and 5.2, with simulation results. Also in this case the differences are rather small, in particular for low traffic intensities.

Example 5.1 Consider the first moment of the stationary total work load in the $M/E_2/2$ queue. Assume that the Poisson arrival process has intensity β^* and let B be the Erlang-distribution of order 2 with expectation 1, i.e. $\boldsymbol{\alpha} = (1, 0)$ and

$$\boldsymbol{T} = \begin{pmatrix} -2 & 2 \\ 0 & -2 \end{pmatrix}$$

Using Theorem 4.1 the total work load $\mathbf{E}^{\beta^*}V$ can be approximated by

$$\mathbf{E}^{\beta^*}V = \beta^* a_1 + (\beta^*)^3 a_3 + \mathrm{o}((\beta^*)^3), \tag{5.4}$$

where a_1 is given by the Remark after Theorem 4.1 and a_3 by Theorem 5.3. Figure 1 shows the exact and approximated results for different values of β^*.

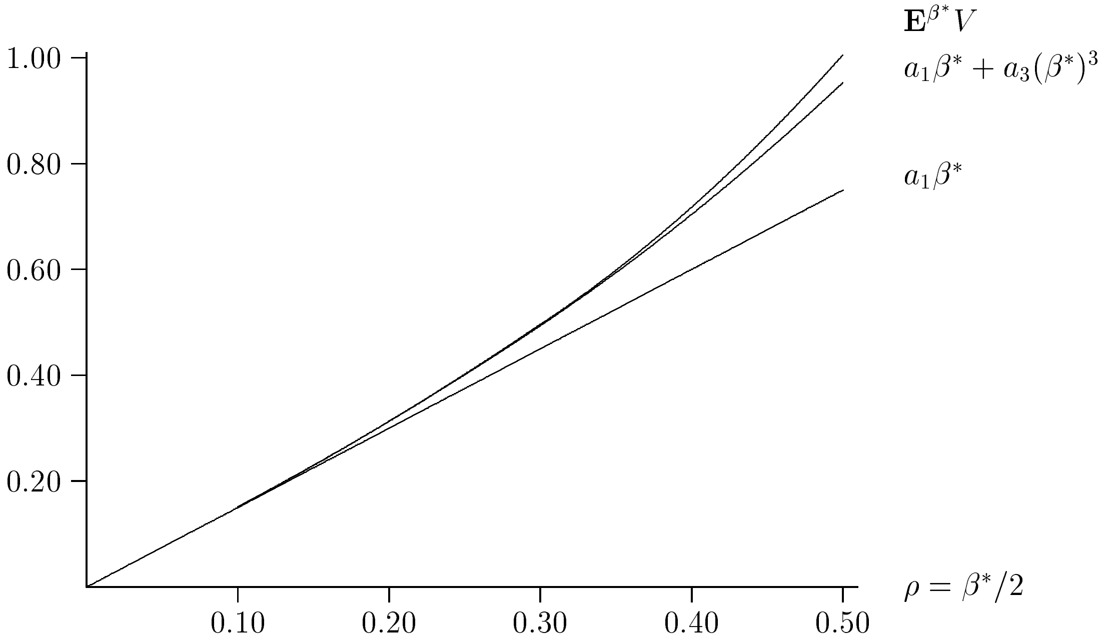

Figure 1: Mean stationary total work load $\mathbf{E}^{\beta^*}V$ *in* $M/E_2/2$.

Example 5.2 Consider the total work load $\mathbf{E}^{\beta^*}V$ in the following $MM/E_2/2$ queue. Let $\boldsymbol{\Lambda} = \begin{pmatrix} -1 & 1 \\ 2 & -2 \end{pmatrix}$ and let the service time distributions B_1 and B_2 be Erlang of order two with expectation $\frac{1}{3}$ and $\frac{2}{9}$, respectively. Thus, we choose $\boldsymbol{T}_1 = \begin{pmatrix} -6 & 6 \\ 0 & -6 \end{pmatrix}$ and $\boldsymbol{T}_2 = \begin{pmatrix} -9 & 9 \\ 0 & -9 \end{pmatrix}$. Using Theorem 4.1 one gets that the total work load $\mathbf{E}^{\beta^*}V$ again can be approximated by (5.4). Figure 2 shows the approximated and simulated values for different values of β^* with $\boldsymbol{\beta} = \left(\frac{1}{2}, 2\right)$.

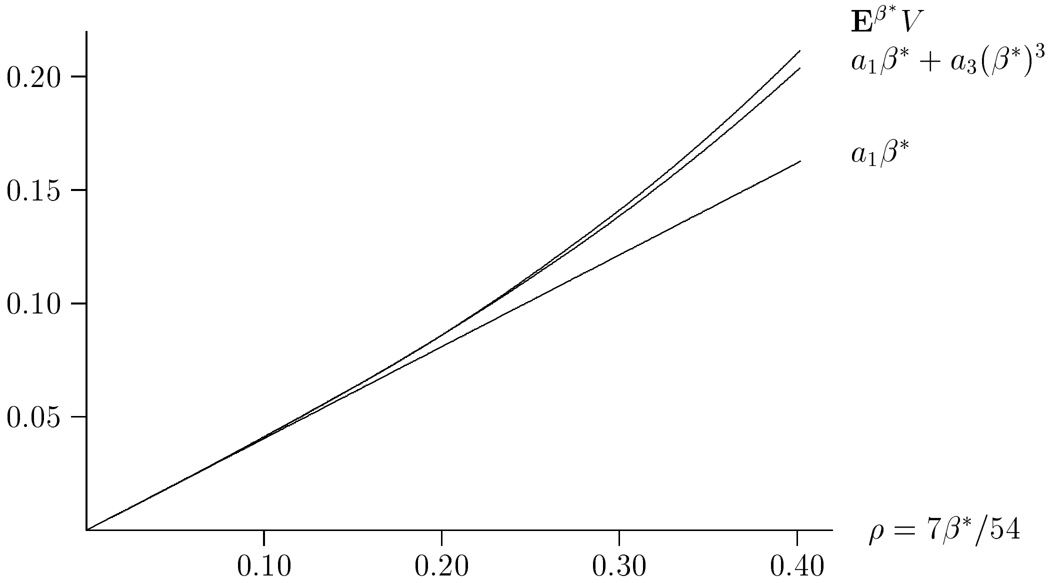

Figure 2: Mean stationary total work load $\mathbf{E}^{\beta^*}V$ *in* MM/E$_2$/2.

Example 5.3 Consider the same setting as in the previous example. Using Theorem 4.1 (with $k = 2$) one gets that the components $\mathbf{E}^{\beta^*}(V_i)^n$ $(i = 1, 2)$ of the nth moment of the work-load vector are given by:

$$\begin{aligned}\mathbf{E}^{\beta^*}(V_1)^n &= (\beta^*)^2 a_2 + \mathrm{o}((\beta^*)^2)\\ \mathbf{E}^{\beta^*}(V_2)^n &= (\beta^*)^1 a_1 - (\beta^*)^2 a_2 + \mathrm{o}((\beta^*)^2),\end{aligned}$$

where a_1 and a_2 are given by Theorem 5.2 (different for various n). The following table shows the second moment of the work-load vector (simulated and approximated values and the relative error) in the MM/E$_2$/2 queue for different values of β^*:

		$\beta^* = 0.5$	$\beta^* = 1.0$	$\beta^* = 1.5$	$\beta^* = 2.0$
	simu.	.00020	.00077	.00235	.00533
$\mathbf{E}^{\beta^*}(V_1)^2$	appr.	.00019	.00078	.00175	.00310
	error (%)	-5.00	1.30	-25.53	-41.84
	simu.	.00976	.01957	.02998	.04306
$\mathbf{E}^{\beta^*}(V_2)^2$	appr.	.00964	.01888	.02774	.03622
	error (%)	-1.23	-3.53	-7.47	-15.88

Table 1: Second moment of the stationary work-load vector $\mathbf{E}^{\beta^*}\boldsymbol{V}$ *in* MM/E$_2$/2; $\rho = \frac{7\beta^*}{54}$.

APPENDIX

In order to make the paper more self-contained we sketch here the proof of Theorem 2.1, although it is a straightforward extension (to vector-valued functionals of m.p.p.'s) of Theorem 3.2 in Błaszczyszyn [7]. The proof is provided by induction with respect to n and based on the following auxiliary result obtained, in a somewhat different form, in Baccelli and Brémaud [4] (see also Lemma 3.3 in [7]).

Lemma A.1 *For a continuous at $\pm\infty$ functional ψ of a simple m.p.p. $(N, \mathbf{P})$ on* $\mathbb{R}$ *it holds*

$$\mathbf{E}\psi(N) = \psi(o) + \int_{\mathbb{R}^2} \int_{M_s} \psi^{x,s}_{(1)}(\mu)\, \mathbf{P}^{(1)}_{x,s}(d\mu)\, M^{(1)}(d(x,s)) \tag{A.1}$$

provided that

$$\int_{\mathbb{R}^2} \int_{M_s} |\psi^{x,s}_{(1)}(\mu)|\, \mathbf{P}^{(1)}_{x,s}(d\mu) M^{(1)}(d(x,s)) < \infty. \tag{A.2}$$

Note that, for $n = 1$, (2.7) and (2.8) is identical with (A.2) and (A.1) respectively. In order to prove Theorem 2.1 for arbitrary $n \in \mathbb{N}$ observe that the integral at the RHS of (A.1) can be rewritten in the form

$$\int_{\mathbb{R}^2} \int_{M_s} \psi^{x,s}_{(1)}(\mu)\, \mathbf{P}^{(1)}_{x,s}(d\mu)\, M^{(1)}(d(x,s)) = \int_{\mathbb{R}^2} \mathbf{E}^{(1)}_{x,s} \psi^{x,s}_{(1)}(N)\, M^{(1)}(d(x,s)),$$

where $\mathbf{E}^{(1)}_{x,s}$ denotes the expectation taken with respect to $\mathbf{P}^{(1)}_{x,s}$. Thus we can again apply Lemma A.1, now to the expected value $\mathbf{E}^{(1)}_{x,s} \psi^{x,s}_{(1)}(N)$ for any given x, s (with $\psi^{x,s}_{(1)}$ as a functional of the m.p.p. $(N, \mathbf{P}^{(1)}_{x,s})$). Next we use the fact that

$$\psi^{x_1,s_1,\ldots,x_i,s_i}_{(i)}(\mu) = \left(\ldots \psi^{x_1,s_1}_{(1)} \ldots\right)^{x_i,s_i}_{(1)}(\mu)$$

and the following identities

$$\begin{aligned}
\left(\mathbf{P}^{(i-1)}_{x_1,s_1,\ldots,x_{i-1},s_{i-1}}\right)_{x_i,s_i} &= \mathbf{P}^{(i)}_{x_1,s_1,\ldots,x_{i-1},s_{i-1},x_i,s_i} \qquad M^{(i)}\text{-a.e.} \\
M^{(i)}(d(x_1,s_1,\ldots,x_{i-1},s_{i-1},x_i,s_i)) &= M^{(1)}_{\mathbf{P}^{(i-1)}_{x_1,s_1,\ldots,x_{i-1},s_{i-1}}}(d(x_i,s_i)) \\
&\quad \times M^{(i-1)}(d(x_1,s_1,\ldots,x_{i-1},s_{i-1})),
\end{aligned}$$

where $M^{(1)}_{\mathbf{P}^{(i-1)}_{x_1,s_1,\ldots,x_{i-1},s_{i-1}}}$ denotes the first moment measure of $(N, \mathbf{P}^{(i-1)}_{x_1,s_1,\ldots,x_{i-1},s_{i-1}})$ (see e.g. Kallenberg [17], König and Schmidt [22]). This yields (2.8) for $n = 2$ and by induction for arbitrary $n \in \mathbb{N}$, which completes the proof of Theorem 2.1.

ACKNOWLEDGEMENT

This work was performed while the first author was a Visiting Scientist at the Department of Stochastics of the University of Ulm; he was supported by a grant of the Deutsche Forschungsgemeinschaft (DFG) and also by the KBN under grant 2 1023 91 01. We are grateful to Tomasz Rolski for his suggestions, which helped us to improve the presentation of the material in the paper.

References

[1] Asmussen, S. (1987) *Applied Probability and Queues.* J. Wiley & Sons, Chichester.

[2] Asmussen, S. (1991) Ladder heights and the Markov-modulated M/G/1 queue. *Stoch. Proc. Appl* **37**, 313–326.

[3] Baccelli, F. and Brémaud, P. (1994) *Elements of Queueing Theory.* Springer-Verlag, Berlin.

[4] Baccelli, F. and Brémaud, P. (1993). Virtual customers in sensitivity and light traffic analysis via Campbell's formula for point processes. *Adv. Appl. Prob.*, **25** 221–234.

[5] Billingsley, P. (1990) *Probability and Measure.* J. Wiley & Sons, New York.

[6] Blanc, J.P.C. (1992) The power-series algorithm applied to the shortest-queue model. *Operations Res.* **40**, 157-167.

[7] Błaszczyszyn, B. (1995) Factorial moment expansion for stochastic systems. *Stoch. Proc. Appl.* (to appear).

[8] Błaszczyszyn, B. and Rolski, T. (1996) Expansions for Markov-modulated systems and approximations of ruin probability. *J. Appl. Prob.* (to appear).

[9] Boxma, O.J., Cohen, J.W. and Huffels, N. (1979) Approximations of the mean waiting time in an M/G/s queueing system. *Operations Res.* **27**, 1115–1127.

[10] Burman, D.Y. and Smith, D.R. (1983) A light traffic theorem for multi-server queues. *Math. Operat. Res.* **8**, 15–25.

[11] Cosmetatos, G.P. (1975) Approximate explicit formulae for the average queueing time in the processes (M/D/r) and (D/M/r). *INFORS* **13**, 328–331.

[12] Cosmetatos, G.P. (1976) Some approximate equlibrium results for the multi-server queue M/G/r. *Operat. Res. Quart.* **27**, 615–620.

[13] Daley, D.J. and Rolski, T. (1992) Light traffic approximations in many-server queues. *Adv. Appl. Prob.* **24**, 202–218.

[14] Franken, P., König, D., Arndt, U. and Schmidt, V. (1981) *Queues and Point Processes.* Akademie-Verlag, Berlin and J. Wiley & Sons, Chichester.

[15] Frey, A. and Schmidt, V. (1994) Taylor-series expansion for multivariate characteristics of risk processes. Preprint, University of Ulm.

[16] Hiller, F.S. and Yu, O.S. (1981) *Queueing Tables and Graphs.* North Holland, New York.

[17] Kallenberg, O. (1983) *Random Measures* (3rd ed.). Akademie-Verlag, Berlin and Academic Press, New York.

[18] Karr, A.F. (1991) *Point Processes and Their Statistical Inference* (2nd ed.). Marcel Dekker, New York.

[19] Kiefer, J. and Wolfowitz, J. (1955) On the theory of queues with many servers. *Trans. Amer. Math. Soc.* **78**, 1–18.

[20] Kimura, T. (1983) Diffusion approximation for an M/G/m queue. *Operations Res.* **31**, 304–321.

[21] Kimura, T. (1994) Approximations for multi-server queues: system interpolations. *Queueing Systems* **17**, 347-382.

[22] König, D. and Schmidt, V. (1992) *Random Point Processes.* B.G. Teubner, Stuttgart (in Germ.).

[23] Kroese, D.P. and Schmidt V. (1995) Light-traffic analysis for queues with spatially distributed arrivals. *Math. Operat. Res.* (to appear).

[24] Neuts, M.F. (1981) *Matrix-Geometric Solutions in Stochastic Models.* Johns Hopkins University Press, Baltimore, MD.

[25] Neuts, M.F. (1989) *Structured Matrices of the M/G/1 Type and Their Applications.* Marcel Dekker, New York.

[26] Nozaki, S.A. and Ross, S.M. (1978) Approximations in finite capacity multiserver queues with Poisson arrivals. *J. Appl. Prob.* **15**, 826–834.

[27] Reiman, M.I. and Simon, B. (1988) Light traffic limits of sojourn time distributions in Markovian queueing networks. *Stochastic Models* **4**, 191–233.

[28] Reiman, M.I. and Simon, B. (1989) Open queueing systems in light traffic. *Math. Operat. Res.* **14**, 26–59.

[29] Ryll-Nardzewski, C. (1961) Remarks on processes of calls. *Proc. 4th Berkeley Symp. Math. Statist. Probab.* **2**, 455–465.

[30] Sigman, K. (1992) Light traffic for workload in queues. *Queueing Systems* **11**, 429–442.

[31] Simon, B. (1993) Calculating light traffic limits for sojourn times in open Markovian queueing systems. *Stochastic Models* **9**, 213–231.

[32] Stoyan, D. (1983) *Comparison Methods for Queues and Other Stochastic Models.* J. Wiley & Sons, Chichester.

[33] Takahashi, Y. (1977) An approximation formula for the mean waiting time of a M/G/c queue. *J. Operat. Res. Soc. Japan* **20**, 150–163.

[34] Tijms, H.C., Van Hoorn, M.H. and Federgruen, A. (1981) Approximations for the steady-state probabilities in the M/G/c queue. *Adv. Appl. Prob.* **13**, 186–206.

[35] Tijms, H.C. (1986) *Stochastic Modelling and Analysis: A Computational Approach.* J. Wiley & Sons, Chichester.

[36] Wang, C.-L. and Wolff, W. (1993) Light traffic approximations for the M/G/c queue. Preprint, University of California.

[37] Yao, D.D. (1985) Refining the diffusion approximation for the M/G/m queue. *Operations Res.* **33**, 1266–1277.

[38] Zazanis, M.A. (1992) Analyticity of Poisson-driven stochastic systems. *Adv. Appl. Prob.* **24**, 532-541.