Marked point processes as limits of Markovian arrival streams

Søren Asmussen
Institute of Electronic Systems, Aalborg University
Fr. Bajersv. 7, DK-9220 Aalborg, Denmark

Ger Koole
Department of Mathematics and Computer Science, University of Leiden
P.O. Box 9512, 2300 RA Leiden, the Netherlands

Abstract

A Markovian arrival stream is a marked point process generated by the state transitions of a given Markovian environmental process and Poisson arrival rates depending on the environment. It is shown that to a given marked point process $\mathcal{N}$ there is a sequence $\{\mathcal{N}^{(m)}\}$ of such Markovian arrival streams with the property that $\mathcal{N}^{(m)} \to^{\mathcal{D}} \mathcal{N}$ as $m \to \infty$. Various related corollaries (involving stationarity, convergence of moments and ergodicity) and counterexamples are discussed as well.
Published in *Journal of Applied Probability* **30**:365–372, 1993.

1 Introduction and statement of results

This paper is concerned with two main model classes for the input to a queueing system, namely marked point processes (MPP's) and Markovian arrival streams (MAS's), which have been used to get beyond the setting of i.i.d. input and introduce genuine dependence. Our main result states that the MAS setting is more general than it may appear at a first sight: any (stationary) MPP is the weak limit of a sequence of (stationary) MAS's.

As usual in Palm theory for point processes (Franken *et al.* [6]), we may consider a MPP either at an arbitrary point of time or at an epoch. In queueing terms, this amounts to distinguishing between the physical time scale and the customer time scale, and this terminology will be used throughout in the paper (thus e.g. we distinguish between time- and customer–stationarity). In the customer time scale, a MPP can be represented as a sequence $\mathcal{S} = \{(T_n, Y_n)\}_{n=0,1,2,\ldots}$ where the T_n represent interarrival times ($0 < T_n < \infty$) and the marks Y_n take values in some Polish space (the mark space); in view of the queueing interpretation, where Y_n is the service time of the nth customer, we shall for simplicity assume that the mark space is $(0, \infty)$ (however, our analysis easily extends to the general case). In the physical time scale, a MPP may be viewed as a point process $\mathcal{N}$ on $[0, \infty) \times (0, \infty)$. In order to be able to speak about weak convergence of MPP's, we need a topology on the state space for a MPP, and this is the sequence space topology (product topology) for the $\mathcal{S}$, and the vague topology in the space of counting measures for the $\mathcal{N}$.

A MAS is defined in terms of a Markov jump process $\{J_t\}$ with finite state space E and an intensity matrix written in the form $\mathbf{C} + \mathbf{D}$, and a family $(B_{ij})_{i,j\in E}$ of distributions on $(0, \infty)$. Here the ijth element d_{ij} of $\mathbf{D}$ denotes the intensity of transitions $i \to j$ accompanied by an arrival with the mark being distributed according to B_{ij} (the case $i = j$ is included), whereas for $i \neq j$ c_{ij} denotes the intensity of the remaining transitions $i \to j$; the diagonal elements c_{ii} of $\mathbf{C}$ are chosen to make the row sums of $\mathbf{C}+\mathbf{D}$ zero. Note that if all B_{ij} are degenerate at 1, we have an ordinary point process. The abbreviation MAP (for Markovian Arrival Process) is used in much of the literature for this version without marks, and the model was first introduced by Rudemo [19] in this setting. Neuts [15] developed computational results for the MAP, and more recently, there has been considerable interest in the MAP or MAS as input to a queue, see Ramaswami [16], Hordijk & Koole [8], [9], [10] (in which the marks are used as customer class numbers), Lucantoni [13], Lucantoni *et al.* [14], Sengupta [21] and Asmussen & Perry [2]. If $d_{ij}=0$ for all $i\neq j$, we have Markov–modulated Poisson arrivals (see e.g. Burman & Smith [4] or Regterschot & de Smit [17]), and another important special case is phase–type renewal arrivals. In the setting of MPP's in the customer time scale, we have to start the MAS with an arrival of type ij (say) at time zero; then $J_0=j$. The interevent times T_n are defined the obvious way.

Our main result is the following:

Theorem 1 *The class of MAS's is dense in the class of MPP's in both the customer–time scale and the physical time scale. That is:*
(a) *for a given MPP $\mathcal{S}$ in the customer–time scale there is a sequence $\{\mathcal{S}^{(m)}\}$ of MAS's such that $\mathcal{S}^{(m)}\to^{\mathcal{D}}\mathcal{S}$ as $m\to\infty$;*
(b) *for a given MPP $\mathcal{N}$ in the physical time scale there is a sequence $\{\mathcal{N}^{(m)}\}$ of MAS's such that $\mathcal{N}^{(m)}\to^{\mathcal{D}}\mathcal{N}$ as $m\to\infty$.*

Theorem 1 may be viewed as a parallel to the classical result of Schassberger [20] stating that phase–type distributions are (weakly) dense in the space of all distributions on $[0,\infty)$, and its implications for the practical worker are similar: one may argue that in many cases the loss of generality by restricting attention to a MAS matters less than the fact that typically models involving MAS's are computationally tractable, while those involving (stationary) MPP's in their full generality are not. A further similar application is insensitivity, cf. e.g. Franken *et al.* [6] Ch. 6: If one can prove a particular result for queues with a MAS input, it will hold also for queues having a stationary MPP as input, provided one can verify the relevant continuity conditions. For example, such an argument would yield the optimality of the shortest queue policy in the case of MPP input, using the results on MAS input in Hordijk & Koole [8], [9], [10] (however, the verification of the continuity conditions is sufficiently complicated to be the possible topic of a separate paper). Thus, we feel that Theorem 1 and its corollaries below may have considerable impact from a modeling point of view (however, mathematically Theorem 1 is hardly deep and may also be expected from a point process result of Hermann [7]).

In many applications, one needs to restrict oneself to stationary MPP's in order to obtain non–trivial results. This motivates the following result:

Corollary 2 *If, in the setting of Theorem 1(a), $\mathcal{S}$ is stationary, one can choose all $\mathcal{S}^{(m)}$ to be stationary as well, and similarly for $\mathcal{N}$ and the $\mathcal{N}^{(m)}$ in (b).*

We can also extend Theorem 1 to convergence of moments:

Corollary 3 *In the setting of Theorem 1(a), one can can choose the $\mathcal{S}^{(m)}$ such that $\mathrm{E}\left(T_n^{(m)}\right)^p \to \mathrm{E}T_n^p$, $\mathrm{E}\left(Y_n^{(m)}\right)^p\to \mathrm{E}Y_n^p$ for all $p<\infty$.*

This result is of particular importance in connection with continuity questions for queues, see e.g. Borovkov [3] Ch. 1.11, Franken *et al.* [6] Ch. 3 and Stoyan [23] Ch. 8. For example, the following result provides the justification of a MAS approximation in some simple cases:

Corollary 4 *Assume that $\mathcal{S}$ is an ergodic stationary MPP in the customer–time scale and that* $\mathrm{E}(Y_n - T_n) < 0$. *Then there exists a sequence of stable single–server FIFO queues with MAS input such that $W^{(m)} \to^{\mathcal{D}} W$ where $W^{(m)}$ is the stationary actual waiting time of the mth queue with MAS input and W the stationary actual waiting time of the queue with MPP input.*

Algorithms providing the steady–state characteristics of a single–server FIFO queue with MAS input are given in Sengupta [21] and Asmussen & Perry [2].

Section 2 contains the proofs of Theorem 1 and Corollaries 1–3 (also some remarks on ergodicity are given), and in Section 3 we give a counterexample showing that the equivalent of Theorem 1 does not hold for Markov–modulated Poisson arrivals.

2 Proofs

First we give some preliminaries; for further facts used in the following, we refer to standard texts like Daley & Vere–Jones [5], Franken *et al.* [6], Kallenberg [12], Rolski [18] and Serfozo [22].

By standard characterizations of weak convergence in product spaces, we have

$$\mathcal{S}^{(m)} = \{(T_n^{(m)}, Y_n^{(m)})\}_{n\in\mathbb{N}} \xrightarrow{\mathcal{D}} \mathcal{S} = \{(T_n, Y_n)\}_{n\in\mathbb{N}}, \quad m \to \infty, \tag{1}$$

if and only if

$$\{(T_n^{(m)}, Y_n^{(m)})\}_{n\le N} \xrightarrow{\mathcal{D}} \{(T_n, Y_n)\}_{n\le N}, \quad m \to \infty, \tag{2}$$

for all finite N.

We can write any MPP $\mathcal{N}$ as $\mathcal{N} = \sum_{n=1}^{\infty} \delta_{(V_n,Y_n)}$ where V_n is the time of the nth arrival and Y_n its mark. We then let $T_0 = V_0$, $T_n = V_n - V_{n-1}$, $n = 1, 2, \ldots$ and $\mathcal{S} = \mathcal{S}(\mathcal{N}) = \{(T_n, Y_n)\}$. The following result is basic by allowing us to work mainly in the customer time scale:

Proposition 5 *Let $\mathcal{N}$, $\mathcal{N}^{(m)}$ be MPP's such that $\mathcal{S}(\mathcal{N}^{(m)}) \to^{\mathcal{D}} \mathcal{S}(\mathcal{N})$. Then $\mathcal{N}^{(m)} \to^{\mathcal{D}} \mathcal{N}$.*

Proof The proposition is essentially well known, but since we could not find an immediate reference, we shall give the proof.

For $f : \mathbb{R}^2 \to \mathbb{R}$ we define $\mathcal{N}(f) = \sum_{n=0}^{\infty} f(V_n, Y_n)$. It suffices to show $\mathcal{N}^{(m)}(f) \to^{\mathcal{D}} \mathcal{N}(f)$ for f continuous with compact support B (Theorem 6.1 of [22]).

Define $f^N : \mathbb{R}^{2N+2} \to \mathbb{R}$ by $f^N(x_0, \ldots, x_{2N+1}) = \sum_{n=0}^{N} f(x_0 + x_2 + \cdots + x_{2n}, x_{2n+1})$ (note that if $x_0, x_2, \ldots, x_{2n}$ are sojourn times, then $x_0 + x_2 + \cdots + x_{2n}$ is an arrival time). Then f^N is bounded and continuous, and hence

$$\sum_{n=0}^{N} f(V_n^{(m)}, Y_n^{(m)}) \xrightarrow{\mathcal{D}} \sum_{n=0}^{N} f(V_n, Y_n) \tag{3}$$

for all N.

Now let $\epsilon > 0$ be given and choose some number $b < \infty$ and some N such that $B \subseteq [0, b] \times (0, \infty)$ and $\mathbb{P}(V_N \le b) < \epsilon$. Then also $\mathbb{P}(V_N^{(m)} \le b) < \epsilon$ for all large m and hence

$$\mathbb{P}\left(\mathcal{N}^{(m)}(f) \ne \sum_{n=0}^{N} f(V_n^{(m)}, Y_n^{(m)})\right) < \epsilon, \quad \mathbb{P}\left(\mathcal{N}(f) \ne \sum_{n=0}^{N} f(V_n, Y_n)\right) < \epsilon.$$

Since ϵ is arbitrary, $\mathcal{N}^{(m)}(f) \to^{\mathcal{D}} \mathcal{N}(f)$ follows by combining with (3).□

Proof of Theorem 1 In view of Proposition 1, it suffices to prove part (a). Let $\mathcal{S}$ be a given MPP. For any h, let H be the integer part of h^{-1} and define the operator $\tau = \tau^{(h)}$ by $\tau^{(h)}(y) = 1$

when $0 < y \le 2h$, $= k$ when $kh < y \le (k+1)h$ with $2 \le k \le H^2 - 1$, and $= H^2$ when $y > H^2$. Let $\mathcal{S}^{(h)}$ be the MPP that we obtain from $\mathcal{S} = \{(T_n, Y_n)\}$ by replacing Y_n by $Y_n^{(h)} = \tau(Y_n)h$ and T_n by an Erlang distributed random variable $T_n^{(h)}$ with $\tau(T_n)$ stages and intensity parameter h^{-1}; representing $T_n^{(h)}$ as a sum of $\tau(T_n)$ exponentially distributed r.v.'s, we assume these to be mutually independent and independent of $\mathcal{N}$. Then, using Lemma 6.1 of Hordijk & Schassberger [11] and (2), $\mathcal{S}^{(h)} \to^{\mathcal{D}} \mathcal{S}$ as $h \downarrow 0$. Thus, to complete the proof, it suffices to find a sequence $\mathcal{S}^{(h;m)}$ such that $\mathcal{S}^{(h;m)} \to^{\mathcal{D}} \mathcal{S}^{(h)}$ as $m \to \infty$.

We use the equivalent representation $\{(S_n^{(h)}, Y_n^{(h)})\}$ of $\mathcal{N}^{(h)}$ where $S_n^{(h)}$ is the number of stages for $T_n^{(h)}$. For any m, we define the state space $E_{h;m}$ for $\mathcal{S}^{(h;m)}$ as

$$E_{h;m} = \{(s_0, k_0, y_0), (s_1, y_1), \ldots, (s_m, y_m) : 1 \le s_n \le H, 1 \le y_n \le H, 1 \le k_0 \le s_0\}.$$

Transitions

$$(s_0, k_0, y_0), (s_1, y_1), \ldots, (s_m, y_m) \to (s_0, k_0 - 1, y_0), (s_1, y_1), \ldots, (s_m, y_m), \quad k_0 \ge 2, \tag{4}$$

for $\{J_t^{(h;m)}\}$ have intensity h^{-1}, whereas the intensity of the remaining transitions can be chosen arbitrary for Theorem 1 except that the sum over (s_{m+1}, y_{m+1}) of the intensities for all transitions of the form

$$(s_0, k_0, 1), (s_1, y_1), \ldots, (s_m, y_m) \to (s_1, s_1, y_1), \ldots, (s_m, y_m), (s_{m+1}, y_{m+1}) \tag{5}$$

should be h^{-1} as well. Note that in terms of the matrices $\mathbf{C}$, $\mathbf{D}$ defining a MAS, the transitions of the form (4) are those corresponding to the off-diagonal elements of $\mathbf{C}$, while those of the form (5) corresponds to $\mathbf{D}$. At each transition of form (5) an arrival with mark s_1 occurs. Thus, if we start the MAS by having $\{J_t^{(h;m)}\}$ just entered state $((S_0, S_0, Y_0), (S_1, Y_1), \ldots, (S_m, Y_m))$, we have

$$\{(T_n^{(h;m)}, Y_n^{(h;m)})\}_{n \le m} \stackrel{\mathcal{D}}{=} \{(T_n^{(h)}, Y_n^{(h)})\}_{n \le m}$$

so that by (2) $\mathcal{S}^{(h;m)} \to^{\mathcal{D}} \mathcal{S}^{(h)}$. □

Proof of Corollary 1 Consider first part (a) (customer–stationarity). We take the intensity of a transition of the form (5) to be

$$\frac{1}{h} \cdot \mathbb{P}\Big((S_1^{(h)}, Y_1^{(h)}), \ldots, (S_m^{(h)}, Y_m^{(h)}), (S_{m+1}^{(h)}, Y_{m+1}^{(h)}) = (s_1, y_1), \ldots, (s_m, y_m), (s_{m+1}, y_{m+1}) \;\Big|\; (S_0^{(h)}, Y_0^{(h)}), \ldots, (S_m^{(h)}, Y_m^{(h)}) = (s_0, y_0), \ldots, (s_m, y_m)\Big). \tag{6}$$

All transition rates not corresponding to (4) or (5) are zero. Then, since

$$J_0^{(h;m)} = \Big((S_0^{(h)}, S_0^{(h)}, Y_0^{(h)}), (S_1^{(h)}, Y_1^{(h)}), \ldots, (S_m^{(h)}, Y_m^{(h)})\Big), \tag{7}$$

it follows immediately that the distribution of $\{J_t^{(h;m)}\}$ just after the first jump of type (5), that is, at time $T_0^{(h;m)}$, is the same as the distribution of

$$\Big((S_1^{(h)}, S_1^{(h)}, Y_1^{(h)}), (S_2^{(h)}, Y_2^{(h)}), \ldots, (S_{m+1}^{(h)}, Y_{m+1}^{(h)})\Big),$$

which by stationarity has the same distribution as in (7). Therefore the initial condition for the MAP after arrival 1 is the same as for the MAP after arrival 0, which shows the stationarity.

For part (b), let $\mathcal{N}$ be a given time–stationary MPP and $\mathcal{S}^{(0)} = \mathcal{S}^{(0)}(\mathcal{N})$ the Palm version. Then we just showed the existence of customer–stationary MAS's $\mathcal{S}^{(m)}$ such that $\mathcal{S}^{(m)} \to^{\mathcal{D}} \mathcal{S}^{(0)}$. As in the proof of Corollary 2 below, we can assume that also $\mathrm{E}T_n^{(m)} \to \mathrm{E}T_n^{(0)}$. From each $\mathcal{S}^{(m)}$, we construct the time–stationary version $\tilde{\mathcal{N}}^{(m)}$ in the standard way, which in particular means that

$$\mathrm{E}\left[f(\tilde{T}_0^{(m)})g(\tilde{\mathbf{U}}_N^{(m)})\right] = \frac{1}{\mathrm{E}T_0^{(m)}}\mathrm{E}\left[\int_0^{T_0^{(m)}} f(T_0^{(m)}-t)dt\cdot g(\mathbf{U}_N^{(m)})\right], \tag{8}$$

where $\mathbf{U}_N = (T_1,\dots,T_N,Y_0,\dots,Y_N)$. Taking f,g continuous and bounded, it follows from (8) that

$$\mathrm{E}\left[f(\tilde{T}_0^{(m)})g(\tilde{\mathbf{U}}_N^{(m)})\right] \to \frac{1}{\mathrm{E}T_0^{(0)}}\mathrm{E}\left[\int_0^{T_0^{(0)}} f(T_0^{(0)}-t)dt\cdot g(\mathbf{U}_N^{(0)})\right] = \mathrm{E}\left[f(T_0)g(\mathbf{U}_N)\right].$$

Thus $(\tilde{T}_0^{(m)},\tilde{\mathbf{U}}_N^{(m)}) \to^{\mathcal{D}} (T_0,\mathbf{U}_N)$. I.e., by (2) $\mathcal{S}(\tilde{N}^{(m)}) \to^{\mathcal{D}} \mathcal{S}(\mathcal{N})$, and using Proposition 1 we get $\tilde{N}^{(m)} \to^{\mathcal{D}} \mathcal{N}$.□

Proof of Corollary 2 Since the moments of the $T_n^{(h)}$ are the same as the moments of the $T_n^{(h;m)}$ and the moments of the $Y_n^{(h)}$ are the same as the moments of the $Y_n^{(h;m)}$, we only have to show that it is possible to obtain

$$\mathrm{E}\left(T_n^{(h)}\right)^p \to \mathrm{E}T_n^p, \quad \mathrm{E}\left(Y_n^{(h)}\right)^p \to \mathrm{E}Y_n^p, \qquad h\downarrow 0, \tag{9}$$

in the first step in the proof of Theorem 1. By Fatou's lemma,

$$\liminf_{h\downarrow 0}\mathrm{E}\left(T_n^{(h)}\right)^p \geq \mathrm{E}T_n^p, \quad \liminf_{h\downarrow 0}\mathrm{E}\left(Y_n^{(h)}\right)^p \geq \mathrm{E}Y_n^p,$$

and thus we may assume that $\mathrm{E}T_n^p < \infty$, $\mathrm{E}Y_n^p < \infty$. The case of the $Y_n^{(h)}$ is then immediate since $Y_n^{(h)} \leq h + Y_n$ so that the $Y_n^{(h)p}$ are uniformly integrable. For the $T_n^{(h)}$, we note that $T_n^{(h)} = W_{\tau(T_n)}$ where the r.v. W_k is Erlang distributed with k stages and intensity parameter h^{-1}. By explicit calculus, $\mathrm{E}W_{\tau(t)}^p \to t^p$ for any t and

$$\mathrm{E}W_{\tau(t)}^p \leq \tau(t)^p\Gamma(p+1)h^p \leq (t+h)^p\Gamma(p+1).$$

Writing

$$\mathrm{E}\left(T_n^{(h)}\right)^p = \mathrm{E}\,\mathrm{E}\left[W_{\tau(T_n)}^p | T_n\right],$$

it follows that the integrand converges to T_n^p a.s. and is dominated by $(T_n+h)^p\Gamma(p-1)$. Now appeal to the dominated convergence theorem. □

Corollary 3 is an immediate consequence of Corollary 2, cf. e.g. Borovkov [3] p. 53. Note that the questions of whether also the moments of $W^{(m)}$ converge to those of W appears more intricate than convergence in distribution, and the best results we know of (Asmussen & Johansen [1]) deal only with i.i.d. input.

We conclude this section by giving some comments on the obvious question concerning the connection between ergodicity of a MPP $\mathcal{N}$ and its MAS approximations $\mathcal{N}^{(m)}$. However, ergodicity is not preserved by convergence in distribution, and hence there seems to be no simple answer to this. More precisely, note first that for a MAS, a sufficient condition for ergodicity in the terminology of point process theory (the shift invariant σ–field is trivial) is that the environmental process is irreducible. Now let $\mathcal{N}$ be a MAS. Then we can find sequences $\{\mathcal{N}_e^{(m)}\}$, $\{\mathcal{N}_r^{(m)}\}$ of MAS's which are ergodic and non–ergodic, respectively, and have the properties $\mathcal{N}_e^{(m)} \to^{\mathcal{D}} \mathcal{N}$, $\mathcal{N}_r^{(m)} \to^{\mathcal{D}} \mathcal{N}$. Indeed, we may define $\mathcal{N}_e^{(m)}$ by perturbing $\mathcal{N}$ by keeping each transition of the environment w.p. $1-1/m$ and replacing it by a transition to a uniformly chosen state w.p. $1/m$. For $\mathcal{N}_r^{(m)}$, we instead replace the environment E by $E \cup F$, keep the environmental process (on E) w.p. $1-1/m$ and replace it by one moving on F w.p. $1/m$; if we let the added F–component have no arrivals at all (say), $\mathcal{N}_r^{(m)}$ is then non–ergodic. Thus, in all of our approximation theorems it is at our disposal to take the approximating MAS to be either ergodic or non–ergodic.

3 Markov-modulated Poisson arrivals: a counterexample

It seems worth pointing out that Theorem 1 does not hold when the class of MAS's is replaced by the class of Markov–modulated Poisson arrival streams. To see this, consider a stationary point process $\mathcal{N}$ with $\mathbb{P}(N_1 = 0) \le 0.09$, $\mathbb{P}(N_1 \ge 2) \le 0.18$ where N_1 is the number of points in $(0,1)$. Such a process clearly exists, for example the process having points at all $n+U$, where $n = 0, \pm 1, \pm 2 \ldots$ and U is uniform on $(0,1)$. Assume that $\mathcal{N}$ is the weak limit of Markov–modulated Poisson processes (MMPP's) $\mathcal{N}^{(m)}$, and define (in the obvious notation)

$$L^{(m)} = \int_0^1 d^{(m)}_{J_t^{(m)}, J_t^{(m)}}\, dt.$$

Then $L^{(m)}$ is the cumulated intensity of $\mathcal{N}^{(m)}$ over $(0,1)$, hence, as the number of arrivals is independent of the arrival rate, $N_1^{(m)}$ is conditionally Poisson with parameter $L^{(m)}$. Since $N_1^{(m)} \to^{\mathcal{D}} N_1$, we therefore obtain

$$\begin{aligned} \mathbb{P}(N_1^{(m)} = 0) &= \mathrm{E} e^{-L^{(m)}} \to \mathbb{P}(N_1 = 0), \\ \mathbb{P}(N_1^{(m)} \ge 2) &= \mathrm{E}[1 - e^{-L^{(m)}} - L^{(m)} e^{-L^{(m)}}] \to \mathbb{P}(N_1 \ge 2). \end{aligned}$$

If m is so large that $\mathbb{P}(N_1^{(m)} = 0) \le 0.10$, it follows that

$$\begin{aligned} e^{-1}\mathbb{P}(L^{(m)} \le 1) &\le \mathrm{E} e^{-L^{(m)}} \le 0.10, \\ \mathbb{P}(L^{(m)} \le 1) &\le 0.28, \quad \mathbb{P}(L^{(m)} > 1) \ge 0.72, \\ \mathbb{P}(N_1^{(m)} \ge 2) &\ge 0.72[1 - e^{-1} - 1 \cdot e^{-1}] = 0.19, \end{aligned}$$

a contradiction.

The more intuitive motivation for the result is the fact that the variance constant of a MMPP is greater than its mean,

$$\mathbf{Var}\, N_1 = \mathrm{E} N_1^2 - (\mathrm{E} N_1)^2 = \mathrm{E}(L + L^2) - (\mathrm{E} L)^2 \ge \mathrm{E} L = \mathrm{E} N_1,$$

whereas this obviously does not hold for arbitrary stationary MPP's (for example, the $\mathcal{N}$ above has $\mathbf{Var}\, N_1 = 0$, $\mathrm{E} N_1 = 1$). However, since weak convergence of point processes does not automatically entail convergence of moments, this line of thought is more difficult to carry out rigorously.

Also from a practical point of view, the approximation by a MMPP may not be adequate. When we look at the departure process of a queueing system with all arrival and service times exponentially distributed, one often sees that the departure rate changes when a customer leaves. As an example, the departure rate in a M/M/1-queue becomes 0 when the last customer leaves. The change of rate and the occurrence of a point happening simultaneously cannot be modelled with a MMPP. Indeed in Hordijk & Koole [9] the MAS is used to model the first of two queueing centres in tandem.

Acknowledgement We would like to thank Arie Hordijk for stimulating discussions and for initiating the present collaboration.

References

[1] S. Asmussen & H. Johansen (1986) Über eine Stetigkeitsfrage betreffend des Bedienungssystems GI/GI/k. *Elektron. Inf. Kyb. – EIK* **22**, 565–570.

[2] S. Asmussen & D. Perry (1991) On cycle maxima, first passage problems and extreme value theory for queues. *Submitted*.

[3] A.A. Borovkov (1976) *Stochastic Processes in Queueing Theory.* Springer–Verlag, New York, Heidelberg, Berlin.

[4] D.Y. Burman & D.R. Smith (1986) An asymptotic analysis of a queueing system with Makov–modulated arrivals. *Oper. Res.* **34**, 105–119.

[5] D.J. Daley & D. Vere–Jones (1988) *An Introduction to the Theory of Point Processes.* Springer–Verlag, New York.

[6] P. Franken, D. König, U. Arndt & V. Schmidt (1981) *Queues and Point Processes.* John Wiley & Sons, Chichester New York.

[7] U. Herrmann (1965) Ein Approximationssatz für Verteilungen stationärer zufälliger Punktfolgen. *Math. Nachrichten* **30**, 377–381. Springer, Berlin.

[8] A. Hordijk & G. Koole (1990) On the optimality of the generalized shortest queue policy. *Probability in the Engineering and Informational Sciences* **4**, 477–487.

[9] A. Hordijk & G. Koole (1990) On the shortest queue policy for the tandem parallel queue. *Technical Report* **TW–90–08**, University of Leiden.

[10] A. Hordijk & G. Koole (1991) Optimal policies in two stochastic scheduling problems. Working paper.

[11] A. Hordijk & R. Schassberger (1982) Weak convergence for generalized semi-Markov processes. *Stoch. Processes Appl.* **12**, 271–291.

[12] O. Kallenberg (1983) *Random Measures.* Academic Press, New York.

[13] D. Lucantoni (1991) New results on the single server queue with a batch Markovian arrival process. *Stoch. Models*

[14] D. Lucantoni, K.S. Meier–Hellstern & M.F. Neuts (1990) A single server queue with server vacations and a class of non–renewal arrival processes. *Adv. Appl. Probab.* **22**, 676-705.

[15] M.F. Neuts (1979) A versatile Markovian point process. *J. Appl. Probab.* **16**, 764-779.

[16] V. Ramaswami (1980) The N/G/1 queue and its detailed analysis. *Adv. Appl. Probab.* **12**, 222-261.

[17] G.J.K. Regterschot & J.H.A. de Smit (1986) The queue M/G/1 with Markov–modulated arrivals and services. *Math. Oper. Res.* **11**, 456–483.

[18] T. Rolski (1981) *Stationary Random Processes Associated with Point Processes.* Lecture Notes in Statistics **5**. Springer–Verlag, New York.

[19] M. Rudemo (1973) Point processes generated by transitions of Markov chains. *Adv. Appl. Probab.* **5**, 262–286.

[20] R. Schassberger (1973) *Warteschlangen*. Springer, Wien.

[21] B. Sengupta (1990) The semi-Markov queue: Theory and Applications. *Stochastic Models* **6**, 383–413.

[22] R.F. Serfozo (1990) Point processes. *Handbooks in Operations Research and Management Science. Vol 2: Stochastic Models* (D.P. Heyman & M.J. Sobel eds), 1–93. North Holland, Amsterdam.

[23] D. Stoyan (1983) *Comparison Methods for Queues and Other Stochastic Models*. John Wiley & Sons, Chichester, New York.