The Annals of Applied Probability
1995, Vol. 5, No. 1, 78–96

INSENSITIVITY IN DISCRETE-TIME GENERALIZED SEMI-MARKOV PROCESSES ALLOWING MULTIPLE EVENTS AND PROBABILISTIC SERVICE SCHEDULING[1]

BY W. HENDERSON, C. E. M. PEARCE, P. G. TAYLOR AND N. M. VAN DIJK

University of Adelaide, University of Adelaide, University of Adelaide and Free University

In this paper we treat a discrete-time generalized semi-Markov process in which simultaneous deaths of more than one lifetime and simultaneous arrivals are permitted and service scheduling is probabilistic. Necessary and sufficient conditions for insensitivity are derived and a simple algorithmic procedure provided whereby the equilibrium probabilities can be furnished when insensitivity holds. Some particular cases of special interest are examined in detail.

1. Introduction. Traditionally, queueing systems have usually been analysed with continuous-time models. In most models a convenient consequence is that simultaneous events occur with probability zero. However with the increasing use of slotted systems, particularly in telecommunications and computer systems, there is a growing demand for discrete-time modelling. In such systems there are natural time points, defined by the time taken for the system to process a single slot, at which to make observations. Moreover it is natural to define loads upon the system, such as requested amounts of service, in discrete amounts corresponding to integral numbers of slots.

Although they can arise via uniformization [see, e.g., Keilson (1979)], such models are not always able to be treated as embeddings in suitable continuous-time processes. Indeed, even a simple two-state Markov chain in discrete time cannot be an embedding in a continuous-time two-state Markov process unless the trace of the one-step transition matrix is at least unity. Discrete-time theory thus does not even in principle automatically lift from that for continuous time.

The most obvious analytical complication for discrete-time queueing models that does not arise in the corresponding continuous-time models is that transitions involving multiple movements may occur. For example, it might easily happen that many arrivals occur during the processing of a single slot.

Received March 1990; revised May 1994.
[1] Supported by the Australian Research Council Grant A69112151.
AMS 1991 *subject classifications.* Primary 60K15; secondary 60K20.
Key words and phrases. Insensitivity, generalized semi-Markov process, product form.

For this reason it is often much more difficult to obtain expressions for performance measures of discrete-time systems than for related continuous-time systems.

In continuous models, differential service rates are imposed through the device of speeds [see, e.g., König and Jansen (1974) and Schassberger (1978b)]. In discrete time, this translates to probabilistic choices of the service elements to be worked on in each time slot. Slotted-time services with probabilistic choice can be useful in ATM switching systems when one wishes to give a differential grade of service to many connections, but it is impracticable to keep track of how much service has been given in the past to each connection. In a military context, it is often impracticable to do connection setup, and one may want to be able to accept traffic on a connectionless basis but with a bias in favour of packets on the basis of their priority category. Probabilistic choices also allow the priority implementation to depend, for example, on a bias factor both for a class and for the number of customers (packets) of that class present.

We shall note in the following section that deterministically sequenced servicing can be recovered from our model.

In cases where demands for numbers of slots in a system can have an arbitrary distribution concentrated on the positive integers, the need for manageable analyses poses the question of insensitivity: When does the equilibrium distribution of the process depend on the demand only through the mean number of slots requested? In this article we consider this question. We employ a discrete-time generalized semi-Markov process (GSMP) framework analogous to that treated in investigations of insensitivity in continuous-time processes [see, e.g., Schassberger (1977, 1978a, b)] except that now multiple events (simultaneous service commencements and/or completions) are possible.

Criteria for insensitivity in continuous-time GSMPs and similar structures have been given by König and Jansen (1974), Schassberger (1977, 1978a, b 1986), Kelly (1979), Burman (1981), Franken, König, Arndt and Schmidt (1982), Henderson (1983), Whittle (1985, 1986), Taylor (1989), Fakinos (1990) and Miyazawa (1991, 1993). Such criteria usually take the form of partial-balance equations that must be satisfied by the GSMP when all its lifetimes are taken to be exponential. Insensitivity is associated also with the existence of a product-form supplemented distribution. We shall see that the corresponding balance criteria in discrete-time GSMPs are more stringent, because of the possibility of transitions involving multiple lifetimes. As an example, we shall see in Section 5 that in contradistinction to the continuous-time case, the discrete-time version of the $M/G/N/N$ loss system is not insensitive. However, there are cases where the criteria are satisfied, leading to insensitive discrete-time models.

The subject of insensitivity in discrete-time processes appears to have been discussed by only a few authors. In Daduna and Schassberger (1981) it was shown that a round-robin queue with general service times has a product-form stationary distribution. Daduna and Schassberger (1983) proved that a discrete-time network of queues with a first-come–first-served discipline at nodes with geometric service times and so-called doubly stochastic disciplines at nodes with type-dependent deterministic service times has a product-form supplemented equilibrium distribution. In both papers the queues operated according to quite restrictive rules, principally disallowing the possibility of multiple simultaneous departures. A similar condition exists in the recent investigation of insensitivity in single-server queues with a moving server [Henderson and Taylor (1992)]. Arrivals were assumed to occur singly, and because only a single server is involved, it is impossible for more than one customer to depart at a time.

Employing a model for a discrete-time GSMP that allows multiple events and probabilistic servicing, we derive necessary and sufficient conditions for there to exist an invariant measure with product form. A useful practical feature of the conditions is that they may be partitioned into a small set of equations, which provides a simple algorithmic derivation for the equilibrium probabilities when insensitivity does hold, and a larger set that acts as a collection of consistency conditions needing to be satisfied for insensitivity to hold. As will become clear, the consistency conditions are not satisfied automatically, nor do they admit simple immediate flux interpretations as in single-movement continuous-time analysis. The model is set up in Section 2 and the basic results are proved in Sections 3 and 4.

In some special cases of practical interest the algorithmic procedure leads to an expression in closed form for the equilibrium probabilities and to simple forms for the consistency conditions. Some particular cases of practical importance are discussed in Sections 5 to 8. The first of these treats the case when all lifetimes are served simultaneously. The following two sections deal with probabilistic servicing when only one lifetime receives service at a time. The final section addresses a model of Daduna and Schassberger (1981, 1983) in which, when multiple lifetimes die simultaneously, all are restarted. We shall assume irreducibility and ergodicity throughout without further comment.

2. The model. Consider a discrete-time process that moves on a set of states G. Incorporated in each state $\mathbf{g} \in G$ are active elements, or lifetimes, from a set $\mathbf{S}$. With a customary abuse of notation we use $\mathbf{g}$ to denote also the set of active elements associated with the state $\mathbf{g}$, that is, we assume implicitly that the set of active elements uniquely specifies the state.

This is done only to keep the notation as simple as possible. In some applications it is convenient to allow an active set $\mathbf{g}$ of service lifetimes to correspond to a number of states $\mathbf{g}_\alpha$, the index α specifying the lifetime(s) undergoing service in the current time slot. We shall remark just before the close of Section 4 on the modifications of our key results necessary to embrace the extension to the general case.

The primary application of systems in which more than one state has the same active elements is in the modelling of round-robin and related service disciplines, in which the plurality of states is to keep track of the sequence of servicing in consecutive time slots. As only one lifetime dies at a time in such models, the present level of generality is not required. For a treatment of such queues, see Henderson and Taylor (1991).

When the element σ becomes alive, it is allocated a nominal lifetime sampled from a discrete distribution $F_\sigma(\cdot)$ with probability mass function $f_\sigma(\cdot)$ and finite mean μ_σ, independently of any other lifetimes that may become alive at that time. At each time point, with probability $c(\mathbf{s}, \mathbf{g})$, each element of a subset $\mathbf{s} \subseteq \mathbf{g}$ has one unit worked off its residual lifetime and the lifetimes of the elements of $\mathbf{g} - \mathbf{s}$ are unaffected. (Here and subsequently $A - B$ is used to denote a set difference $A \setminus B$ when B is a subset of A. We shall also use $A + B$ to denote a disjoint set union.) If the residual service requirement of one or more elements is reduced to zero at this time, then those elements will die and the process will make a transition to a state $\mathbf{h}$. If the active elements of the subset $\mathbf{s}_2 \subseteq \mathbf{s}$ die, the process moves to the state $\mathbf{h} \supseteq \mathbf{g} - \mathbf{s}_2$ with probability $p(\mathbf{g}, \mathbf{s}_2, \mathbf{s}, \mathbf{h})$ with the set $\mathbf{s}_1 = \mathbf{h} - (\mathbf{g} - \mathbf{s}_2)$ of elements being activated. We write $\mathbf{s}_0 = \mathbf{s} - \mathbf{s}_2$ for the set of elements that have a unit worked off their residual lifetimes without these thereby being reduced to zero.

The possibilities $\mathbf{g} = \mathbf{h}$, $\mathbf{g} = \varnothing$ and $\mathbf{s}_2 = \varnothing$ are not excluded in the transitions described above. In particular, allowing the choice $\mathbf{s}_2 = \varnothing$ allows greater economy in modelling than is offered by the standard continuous-time GSMP model [see Schassberger (1978b)]. In the standard model the set $\mathbf{S}$ of lifetimes is usually decomposed into a set $\mathbf{S}'$ of exponentially distributed lifetimes and a set $\mathbf{S}^*$ of generally distributed lifetimes. In many modelling situations the former are used as "triggers" to activate the latter. For example, to model a Poisson arrival to a queue that activates a generally distributed service time, it is necessary to have an exponentially distributed lifetime for the interarrival time and a generally distributed lifetime for the service time. This arises because, in the standard continuous-time GSMP formulation, transitions occur only when a lifetime dies. By contrast, our discrete-time formulation allows lifetimes to be created at a time point at which none has died. Such a device is equivalent to having a set $\mathbf{S}'$ of geometrically distributed lifetimes triggering these transitions, but is simpler notationally and conceptually.

Supplement the states $\mathbf{g}$ with a vector of lifetimes $\mathbf{x}$, where $x_\sigma \geq 1$ is the residual lifetime of an active element $\sigma \in \mathbf{g}$ and $x_\sigma = 0$ for $\sigma \in \mathbf{S} - \mathbf{g}$. Then the process with supplemented states is a Markov chain. A transition from state $(\mathbf{g}, \mathbf{x})$ to state $(\mathbf{h}, \mathbf{y})$ is possible only if $\mathbf{h}$ is of the form

$$\mathbf{h} = (\mathbf{g} - \mathbf{s}_2) + \mathbf{s}_1 \tag{2.1}$$

with the set of residual lifetimes $\mathbf{y}$ of the form

$$\mathbf{y} = \mathbf{x} - \mathbf{e}(\mathbf{s}) + \mathbf{y}_1(\mathbf{s}_1) \tag{2.2}$$

for some $\mathbf{s}_1 \subseteq \mathbf{S} - \mathbf{s}_0$. Here $\mathbf{e}(\mathbf{s})$ is the vector with unity in each position $\sigma \in \mathbf{s}$ and zeros elsewhere, and $\mathbf{y}_1$ is the vector of freshly created lifetimes corresponding to $\mathbf{s}_1$.

The equilibrium equations for the Markov chains are then

$$\pi(\mathbf{h},\mathbf{y}) = \sum_{\mathbf{s}_0,\mathbf{s}_1,\mathbf{s}_2} \pi(\mathbf{g},\mathbf{x})c(\mathbf{s},\mathbf{g})p(\mathbf{g},\mathbf{s}_2,\mathbf{s},\mathbf{h}) \prod_{\sigma\in\mathbf{s}_1} f_\sigma(y_\sigma), \tag{2.3}$$

where $\mathbf{g}$ and $\mathbf{x}$ are related to $\mathbf{h}$ and $\mathbf{y}$ by (2.1) and (2.2).

We are interested in the situation in which there exists a solution to (2.3) with product form over the supplementary variables, that is, when

$$\pi(\mathbf{h},\mathbf{y}) = \pi(\mathbf{h}) \prod_{\sigma\in\mathbf{h}} \frac{r_\sigma(y_\sigma)}{\mu_\sigma}, \qquad \mathbf{h}\in G, \tag{2.4}$$

where $r_\sigma(y) = F_\sigma^*(y) := \sum_{n\geq y} f_\sigma(n)$. Before examining this situation, we derive Proposition 1 below, which conveniently encapsulates the notion of insensitivity. To this end, we define first, for each $\sigma\in\mathbf{S}$, the set of distributions

$$\Phi_\sigma := \left\{ f_\sigma(\cdot) \colon \sum_{l=1}^{\infty} l f_\sigma(l) = \mu_\sigma \right\}$$

and for each state $\mathbf{h}$, the set of lifetimes

$$\mathbf{H}(\mathbf{h}) := \{\sigma\in\mathbf{h} \colon \mu_\sigma = 1\}.$$

Where there is no ambiguity, the argument of $\mathbf{H}$ is suppressed.

PROPOSITION 1. *Let* $(\mathbf{h},\mathbf{y})$ *be a supplemented state, with* $y_\sigma > 1\,\forall\sigma\in\mathbf{h}-\mathbf{H}$ *and* $y_\sigma = 1\,\forall\sigma\in\mathbf{H}$. *Then in order that* $(a_\mathbf{k};\mathbf{k}\subseteq\mathbf{h})$ *be a solution to*

$$\sum_{\mathbf{k}\subseteq\mathbf{h}} a_\mathbf{k}\left[\prod_{\sigma\in\mathbf{k}} f_\sigma(y_\sigma)\right]\prod_{\sigma\in\mathbf{h}-\mathbf{k}} r_\sigma(y_\sigma+1)=0 \tag{2.5}$$

$$\forall f_\sigma\in\Phi_\sigma,\ \sigma\in\mathbf{h},\ \forall\mathbf{k}\subseteq\mathbf{h},$$

with $(a_\mathbf{k})$ *not depending on* $\{f_\sigma\}$ *except possibly through the means* $\{\mu_\sigma;\sigma\in\mathbf{h}\}$, *it is necessary and sufficient that*

$$a_\mathbf{k} = 0 \quad \textit{for all } \mathbf{k} \textit{ satisfying } \mathbf{H}\subseteq\mathbf{k}\subseteq\mathbf{h}. \tag{2.6}$$

3. Proof of Proposition 1. The constraint that f_σ is a probability mass function with mean μ_σ can be expressed as the conditions that r_σ is a nonnegative and monotone nonincreasing function satisfying

$$r_\sigma(1) = 1$$

and

$$\sum_{l\geq 1} r_\sigma(l) = \mu_\sigma. \tag{3.1}$$

If $\mu_\sigma = 1$, this necessarily means that

$$r_\sigma(l) = 0 \qquad \forall l>1,\ \sigma\in\mathbf{H}. \tag{3.2}$$

On account of (3.2), we have if $\mathbf{H} \neq \varnothing$, that

(3.3) $$\prod_{\sigma \in \mathbf{h}-\mathbf{k}} r_\sigma(y_\sigma + 1) = 0 \quad \text{for all } \mathbf{k} \not\supseteq \mathbf{H}.$$

Condition (3.3) holds also (vacuously) if $\mathbf{H} = \varnothing$. Sufficiency follows immediately.

For necessity, observe that (2.5) may be written as

(3.4) $$\sum_{\mathbf{H} \subseteq \mathbf{k} \subseteq \mathbf{h}} a_{\mathbf{k}} \left[\prod_{\sigma \in \mathbf{k}} \{ r_\sigma(y_\sigma) - r_\sigma(y_\sigma + 1) \} \right] \prod_{\sigma \in \mathbf{h}-\mathbf{k}} r_\sigma(y_\sigma + 1) = 0, \quad \mathbf{k} \subseteq \mathbf{h}.$$

Because $\mu_\sigma > 1$ for all $\sigma \in \mathbf{h} - \mathbf{H}$, there exists for each such σ a net $(f_{\sigma,t}; t \in I) \subseteq \Phi_\sigma$ such that:

(a) For each t,

$$r_{\sigma,t}(l - 1) > r_{\sigma,t}(l) > r_{\sigma,t}(l + 1) > r_{\sigma,t}(l + 2)$$

for $l = y_\sigma$ and also for some $l = y'_\sigma > y_\sigma + 1$.

(b) By varying the choice of t, the values $r_{\sigma,t}(l)$ and $r_{\sigma,t}(l + 1)$ can be varied continuously and independently over nontrivial intervals J_σ, J^*_σ, with the mean constraint (3.1) being maintained through simultaneous variations of $r_{\sigma,t}(y'_\sigma)$, $r_{\sigma,t}(y'_\sigma + 1)$.

It follows from the elementary properties of analytic functions in several complex variables that the coefficient of each distinct multinomial

$$\left(\prod_{\sigma \in \mathbf{n}} z_\sigma \right) \prod_{\sigma \in \mathbf{h}-\mathbf{n}} z'_\sigma$$

in the multilinear form

$$\sum_{\mathbf{H} \subseteq \mathbf{k} \subseteq \mathbf{h}} a_{\mathbf{k}} \left[\prod_{\sigma \in \kappa} (z_\sigma - z'_\sigma) \right] \prod_{\sigma \in \mathbf{h}-\mathbf{k}} z'_\sigma$$

must vanish, that is,

(3.5) $$\sum_{\mathbf{k} \subseteq \mathbf{n} \subseteq \mathbf{h}} a_{\mathbf{n}} (-1)^{|\mathbf{n}-\mathbf{k}|} = 0 \qquad \forall \mathbf{k}: \mathbf{H} \subseteq \mathbf{k} \subseteq \mathbf{h}.$$

An induction on $|\mathbf{h} - \mathbf{k}|$ establishes that each $a_{\mathbf{k}} = 0$. The basis, with $\mathbf{k} = \mathbf{h}$, is trivial. For the inductive step, the assumption $a_{\mathbf{n}} = 0$ for all $\mathbf{n} \supset \mathbf{k}$ gives $a_{\mathbf{k}} = 0$ immediately from (3.5). □

The proposition can be paraphrased as saying that the functionals

$$\left[\prod_{\sigma \in \mathbf{k}} f_\sigma(y_\sigma) \right] \prod_{\sigma \in \mathbf{h}-\mathbf{k}} r_\sigma(y_\sigma + 1)$$

for $\mathbf{H} \subseteq \mathbf{k} \subseteq \mathbf{h}$ are linearly independent.

4. The balance conditions. Define

$$\theta(\mathbf{h}) := \pi(\mathbf{h}) \Big/ \prod_{\sigma \in \mathbf{h}} \mu_\sigma ,$$

with the denominator of the right-hand side taken as unity when $\mathbf{h} = \varnothing$, and recalling that $\mathbf{s} = \mathbf{s}_0 + \mathbf{s}_2$, write

$$\psi(\mathbf{m}, \mathbf{s}_2, \mathbf{u}, \mathbf{h}) := \sum_{\mathbf{s}_0 \subseteq \mathbf{m}} c(\mathbf{s}, \mathbf{u}) p(\mathbf{u}, \mathbf{s}_2, \mathbf{s}, \mathbf{h}) \tag{4.1}$$

for the total transition probability from $\mathbf{u}$ to $\mathbf{h}$ with at least the lifetimes in $\mathbf{h} - \mathbf{m}$ not having a segment processed, conditional on the lifetimes in $\mathbf{s}_2$ expiring. We have the following result.

THEOREM 1. *A necessary and sufficient condition for the supplemented, discrete-time GSMP to have an invariant measure of the form* (2.4) *is that* $\theta(\mathbf{h})$ *satisfy the equations*

$$\theta(\mathbf{h}) = \sum_{\mathbf{m} \subseteq \mathbf{t} \subseteq \mathbf{h}} \sum_{\mathbf{s}_2 \subseteq \mathbf{S} - \mathbf{t}} \theta(\mathbf{t} + \mathbf{s}_2) \psi(\mathbf{m}, \mathbf{s}_2, \mathbf{t} + \mathbf{s}_2, \mathbf{h}) \tag{4.2}$$

taken over all $\mathbf{h}$ *and* $\varnothing \subseteq \mathbf{m} \subseteq \mathbf{h} - \mathbf{H}$.

PROOF. Consider a state $(\mathbf{h}, \mathbf{y})$ with $y_\sigma > 1$ for all $\sigma \in \mathbf{h} - \mathbf{H}$. Of necessity we have $y_\sigma = 1$ for $\sigma \in \mathbf{H}$. On substitution of the product form (2.4) into (2.3) for such states, the right-hand side becomes

$$\begin{aligned} &\sum_{\mathbf{s}_0 + \mathbf{s}_1 \subseteq \mathbf{h}} \sum_{\mathbf{s}_2 \subseteq (\mathbf{S} - \mathbf{h}) + \mathbf{s}_1} \theta(\mathbf{g}) c(\mathbf{s}, \mathbf{g}) p(\mathbf{g}, \mathbf{s}_2, \mathbf{s}, \mathbf{h}) \\ &\quad \times \left[\prod_{\sigma \in \mathbf{s}_1} f_\sigma(y_\sigma) \right] \left[\prod_{\sigma \in \mathbf{s}_0} r_\sigma(y_\sigma + 1) \right] \prod_{\sigma \in \mathbf{h} - (\mathbf{s}_1 + \mathbf{s}_0)} r_\sigma(y_\sigma), \end{aligned} \tag{4.3}$$

because $\prod_{\sigma \in \mathbf{s}_2} r_\sigma(1) = 1$. Because

$$\begin{aligned} \prod_{\sigma \in \mathbf{u}} r_\sigma(y_\sigma) &= \prod_{\sigma \in \mathbf{u}} [r_\sigma(y_\sigma + 1) + f_\sigma(y_\sigma)] \\ &= \sum_{\mathbf{w} \subseteq \mathbf{u}} \left[\prod_{\sigma \in \mathbf{w}} f_\sigma(y_\sigma) \right] \prod_{\sigma \in \mathbf{u} - \mathbf{w}} r_\sigma(y_\sigma + 1), \end{aligned} \tag{4.4}$$

(4.3) can be rewritten as

$$\begin{aligned} &\sum_{\mathbf{s}_0 + \mathbf{s}_1 + \mathbf{w} \subseteq \mathbf{h}} \sum_{\mathbf{s}_2 \subseteq (\mathbf{S} - \mathbf{h}) + \mathbf{s}_1} \theta(\mathbf{g}) c(\mathbf{s}, \mathbf{g}) p(\mathbf{g}, \mathbf{s}_2, \mathbf{s}, \mathbf{h}) \\ &\quad \times \left[\prod_{\sigma \in \mathbf{s}_1 + \mathbf{w}} f_\sigma(y_\sigma) \right] \prod_{\sigma \in \mathbf{h} - (\mathbf{s}_1 + \mathbf{w})} r_\sigma(y_\sigma + 1). \end{aligned}$$

The substitution $\mathbf{k} = \mathbf{s}_1 + \mathbf{w}$ provides

$$\begin{aligned} &\sum_{\mathbf{s}_0 + \mathbf{k} \subseteq \mathbf{h}} \sum_{\mathbf{s}_1 \subseteq \mathbf{k}} \sum_{\mathbf{s}_2 \subseteq (\mathbf{S} - \mathbf{h}) + \mathbf{s}_1} \theta(\mathbf{g}) c(\mathbf{s}, \mathbf{g}) p(\mathbf{g}, \mathbf{s}_2, \mathbf{s}, \mathbf{h}) \\ &\quad \times \left[\prod_{\sigma \in \mathbf{k}} f_\sigma(y_\sigma) \right] \prod_{\sigma \in \mathbf{h} - \mathbf{k}} r_\sigma(y_\sigma + 1). \end{aligned} \tag{4.5}$$

An expansion of the left-hand side of (2.3) with (2.4) substituted, gives by virtue of (4.4) and the definition of $\theta(\mathbf{h})$,

$$\theta(\mathbf{h}) \sum_{\mathbf{k}\subseteq\mathbf{h}} \left[\prod_{\sigma\in\mathbf{k}} f_\sigma(y_\sigma) \right] \prod_{\sigma\in\mathbf{h}-\mathbf{k}} r_\sigma(y_\sigma+1). \tag{4.6}$$

By Proposition 1 a necessary and sufficient condition for (4.5) to equal (4.6) for all possible vectors of distributions with $f_\sigma(\cdot)$ selected from Φ_σ is that the corresponding coefficients of $[\prod_{\sigma\in\mathbf{k}} f_\sigma(y_\sigma)]\prod_{\sigma\in\mathbf{h}-\mathbf{k}} r_\sigma(y_\sigma+1)$ are equal for $\mathbf{k}\supseteq\mathbf{H}$. The balance equations

$$\theta(\mathbf{h}) = \sum_{\mathbf{s}_1\subseteq\mathbf{k}} \sum_{\mathbf{s}_2\subseteq(\mathbf{S}-\mathbf{h})+\mathbf{s}_1} \theta(\mathbf{g})\psi(\mathbf{h}-\mathbf{k},\mathbf{s}_2,\mathbf{g},\mathbf{h}) \tag{4.7}$$

follow. The alternative form (4.2), which obviates the need for (2.1) to give $\mathbf{g}$, arises under a change to relative complement sets $\mathbf{m}=\mathbf{h}-\mathbf{k}$ and $\mathbf{t}=\mathbf{h}-\mathbf{s}_1$. For states with $y_\sigma=1$ for some $\sigma\in\mathbf{h}-\mathbf{H}$ the sufficiency part of the proof of Proposition 1 goes through and so (2.4) is still satisfied. □

Remark. Based on experience with continuous-time GSMPs it is tempting to try to explain the balance relations (4.2) in terms of fluxes experienced by the process when all lifetimes are taken to be geometric. For this process a lifetime σ, which receives a segment of service, dies with probability $1/\mu_\sigma$ and remains alive with probability $1-1/\mu_\sigma$. Write (4.2) as

$$\frac{\pi(\mathbf{h})}{\prod_{\sigma\in\mathbf{h}-\mathbf{m}}\mu_\sigma} = \sum_{\mathbf{m}\subseteq\mathbf{t}\subseteq\mathbf{h}} \sum_{\mathbf{s}_2\subseteq\mathbf{S}-\mathbf{t}} \frac{\pi(\mathbf{t}+\mathbf{s}_2)}{\prod_{\sigma\in(\mathbf{t}+\mathbf{s}_2)-\mathbf{m}}\mu_\sigma}\psi(\mathbf{m},\mathbf{s}_2,\mathbf{t}+\mathbf{s}_2,\mathbf{h}). \tag{4.8}$$

The numerator terms on the right-hand side suggest an interpretation in terms of the total flux into $\mathbf{h}$ from states containing $\mathbf{m}$ due to subsets $\mathbf{s}_0$ of $\mathbf{m}$ receiving service but not dying and subsets $\mathbf{s}_2$ of lifetimes not in $\mathbf{m}$ receiving service and dying. The left-hand side invites an interpretation in terms of the flux leaving $\mathbf{h}$ due to deaths of all lifetimes not in $\mathbf{m}$. However, for this interpretation to be correct, we would expect a factor of $\prod_{\sigma\in\mathbf{s}_0}(1-1/\mu_\sigma)$ on the right-hand side and we would not expect the factor $\prod_{\sigma\in\mathbf{t}-\mathbf{m}}1/\mu_\sigma$. Also, on the left-hand side no account is taken of which lifetimes actually receive a segment of service. Thus, in general, there seems to be no flux interpretation for the balance condition. Later we shall see that certain special cases of (4.8) can be given a flux interpretation.

Corollary. *Let* $\theta(\mathbf{h})=\pi(\mathbf{h})/\prod_{\sigma\in\mathbf{h}-\mathbf{m}}\mu_\sigma$ *be a solution to* (4.2) *and let* $(\pi(\mathbf{h},\mathbf{y}))$ *be the supplemented solution to* (2.3) *when the* σ *component of the vector of distributions is selected from* Φ_σ. *Then*

$$\sum_{\mathbf{y}} \pi(\mathbf{h},\mathbf{y}) = \pi(\mathbf{h}), \qquad \mathbf{h}\in G.$$

PROOF. The result follows from Theorem 1 on summation of (2.4) over $\mathbf{y}$. □

The corollary gives us a sufficient condition for insensitivity of the discrete-time GSMP. If a solution can be found to (4.2) that corresponds to a probability distribution $\pi(\mathbf{h})$, then the marginal distribution that the process is in state $\mathbf{h}$ remains invariant provided each distribution f_σ is selected from Φ_σ.

THEOREM 2. *When product form occurs, the scaled probabilities* $\theta(\mathbf{h})$ *may be determined recursively from the formula*

$$\theta(\mathbf{h}) = \sum_{\tau \supset \mathbf{h}} \theta(\tau)\Omega(\tau, \mathbf{h}),$$

where

$$\Omega(\tau, \mathbf{h}) := \frac{\sum_{\mathbf{h}-\mathbf{H} \subseteq \mathbf{t} \subseteq \mathbf{h}} \psi(\mathbf{h}-\mathbf{H}, \tau-\mathbf{t}, \tau, \mathbf{h})}{1 - \sum_{\mathbf{h}-\mathbf{H} \subseteq \mathbf{t} \subseteq \mathbf{h}} \psi(\mathbf{h}-\mathbf{H}, \mathbf{h}-\mathbf{t}, \mathbf{h}, \mathbf{h})}, \qquad \varnothing \subseteq \mathbf{h} \subset \tau \subseteq \mathbf{S}.$$

PROOF. Each $\mathbf{t}$ in (4.2) contains no elements σ for which $\mu_\sigma = 1$. Accordingly for $\mathbf{m} = \mathbf{h} - \mathbf{H}$, relation (4.2) can be recast, after a change of the order of summation, as

$$\theta(\mathbf{h}) = \sum_{\tau \supseteq \mathbf{h}} \theta(\tau) \sum_{\mathbf{h}-\mathbf{H} \subseteq \mathbf{t} \subseteq \mathbf{h}} \psi(\mathbf{h}-\mathbf{H}, \tau-\mathbf{t}, \tau, \mathbf{h}),$$

whence we have the stated result. □

We deduce the following.

COROLLARY. *When the supplemented GSMS has product form* (2.4), *the invariant measure is given by*

$$\theta(\mathbf{h}) = \theta(\mathbf{S}) \sum \prod_{i=0}^{n-1} \Omega(\mathbf{h}_i, \mathbf{h}_{i+1}), \qquad \varnothing \subseteq \mathbf{h} \subset \mathbf{S},$$

where the summation is over all chains

$$\mathbf{h} = \mathbf{h}_0 \subset \mathbf{h}_1 \subset \cdots \subset \mathbf{h}_n = \mathbf{S}.$$

When $\mathbf{h} = \mathbf{S}$, relation (4.2) for $\mathbf{m} = \mathbf{S} - \mathbf{H}(\mathbf{S})$ is an identity. The remaining relations (4.2) provide a set of necessary and sufficient conditions for insensitivity.

In a number of special cases the balance equations (4.2) reduce to simple equations. In the following sections we discuss some of these.

REMARK. The foregoing arguments go through with obvious modifications when the same set of active elements is associated with more than one state. Suppose $\mathbf{h}$ is associated with states $\mathbf{h}_\alpha$ and $\mathbf{g}$ with $\mathbf{g}_\beta$, where α, β belong to appropriate ordered index sets. We can form corresponding row vectors $\Theta(\mathbf{h}), \Theta(\mathbf{g})$ with respective entires $\theta_\alpha(\mathbf{h})$, $\theta_\beta(\mathbf{g})$, in an obvious notation. Similarly we may define a matrix $\Psi(\mathbf{m}, \mathbf{s}_2, \mathbf{u}, \mathbf{h})$ with (β, α) entry

$$\psi_{(\beta,\alpha)}(\mathbf{m}, \mathbf{s}_2, \mathbf{u}, \mathbf{h}) = \sum_{\mathbf{s}_0 \subseteq \mathbf{m}} c_\beta(\mathbf{s}, \mathbf{u}) p_{\beta,\alpha}(\mathbf{m}, \mathbf{s}_2, \mathbf{u}, \mathbf{h})$$

relating to a transition from $\mathbf{u}_\beta$ to $\mathbf{h}_\alpha$. Then

$$\Theta(\mathbf{h}) = \sum_{\tau \supseteq \mathbf{h}} \Theta(\tau) \sum_{\mathbf{h} - \mathbf{H} \subseteq \mathbf{t} \subseteq \mathbf{h}} \Psi(\mathbf{h} - \mathbf{H}, \tau - \mathbf{t}, \tau, \mathbf{h})$$

and we have the recursive formula

$$\Theta(\mathbf{h}) = \sum_{\tau \supset \mathbf{h}} \theta(\tau) \Omega(\tau, \mathbf{h}),$$

where $\Omega(\tau, \mathbf{h})$ is given as

$$\sum_{\mathbf{h} - \mathbf{H} \subseteq \mathbf{t} \subseteq \mathbf{h}} \Psi(\mathbf{h} - \mathbf{H}, \tau - \mathbf{t}, \tau, \mathbf{h})$$

$$\times \left[I - \sum_{\mathbf{h} - \mathbf{H} \subseteq \mathbf{t} \subseteq \mathbf{h}} \Psi(\mathbf{h} - \mathbf{H}, \mathbf{h} - \mathbf{t}, \mathbf{h}, \mathbf{h}) \right]^{-1}, \qquad \varnothing \subseteq \mathbf{h} \subset \tau \subseteq \mathbf{S}.$$

The nonsingularity of the matrix in brackets is guaranteed because irreducibility of the process ensures that its second term is substochastic.

For simplicity we assume henceforth for each state $\mathbf{h}$ that

$$\mathbf{H}(\mathbf{h}) = \varnothing. \tag{4.9}$$

5. The case when all lifetimes are served simultaneously. A discrete-time GSMP that operates so that every lifetime that is alive receives a unit of service at each time point can be modelled using the structure of Section 2 by putting

$$c(\mathbf{s}, \mathbf{g}) = \delta_{\mathbf{s},\mathbf{g}},$$

where

$$\delta_{\mathbf{a},\mathbf{b}} = \begin{cases} 1, & \text{if } \mathbf{a} = \mathbf{b}, \\ 0, & \text{if } \mathbf{a} \neq \mathbf{b}. \end{cases}$$

Because

$$p(\mathbf{g}, \mathbf{s}_2, \mathbf{s}, \mathbf{h}) = \delta_{\mathbf{s},\mathbf{g}} p(\mathbf{g}, \mathbf{s}_2, \mathbf{g}, \mathbf{h}),$$

the third argument of the routing probability p becomes redundant and p may be contracted conveniently to $p(\mathbf{g}, \mathbf{s}_2, \mathbf{h})$.

It follows that for $\varnothing \subseteq \mathbf{m} \subseteq \mathbf{t}$,

$$\psi(\mathbf{m}, \mathbf{s}_2, \mathbf{t} + \mathbf{s}_2, \mathbf{h}) = \delta_{\mathbf{m},\mathbf{t}} p(\mathbf{t} + \mathbf{s}_2, \mathbf{s}_2, \mathbf{h}) \tag{5.1}$$

and so (4.2) reduces to

$$\theta(\mathbf{h}) = \sum_{\mathbf{s}_2 \subseteq \mathbf{S} - \mathbf{m}} \theta(\mathbf{m} + \mathbf{s}_2) p(\mathbf{m} + \mathbf{s}_2, \mathbf{s}_2, \mathbf{h}), \tag{5.2}$$

which must be satisfied for all $\mathbf{m}$ for which $\varnothing \subseteq \mathbf{m} \subseteq \mathbf{h}$.

If the term $\prod_{\sigma \in \mathbf{m}} \mu_\sigma^{-1}$ is cancelled from both sides of (5.2), this becomes, on reverting temporarily to the use of $\mathbf{k} = \mathbf{h} - \mathbf{m}$,

$$(5.3)\qquad \begin{aligned} &\frac{\pi(\mathbf{h})}{\prod_{\sigma \in \mathbf{k}} \mu_\sigma} \prod_{\sigma \in \mathbf{h}-\mathbf{k}} \left(1 - \frac{1}{\mu_\sigma}\right) \\ &= \sum_{\mathbf{s}_2 \subseteq \mathbf{S} - (\mathbf{h}-\mathbf{k})} \frac{\pi((\mathbf{h}-\mathbf{k}) + \mathbf{s}_2)}{\prod_{\sigma \in \mathbf{s}_2} \mu_\sigma} \\ &\quad \times \left[\prod_{\sigma \in \mathbf{h}-\mathbf{k}} \left(1 - \frac{1}{\mu_\sigma}\right)\right] p((\mathbf{h}-\mathbf{k}) + \mathbf{s}_2, \mathbf{s}_2, \mathbf{h}). \end{aligned}$$

Readers familiar with insensitivity results in continuous-time GSMPs will recognize a similarity between this equation and the partial-balance equations, necessary and sufficient for product form in standard continuous-time GSMPs. The left-hand side represents the flux out of state $\mathbf{h}$ due to the death of the lifetimes in $\mathbf{k}$ and the right-hand side that into state $\mathbf{h}$ due to the birth of the lifetimes in $\mathbf{k}$.

In a number of situations, as when deaths correspond to departures from the system, p reduces further to a two-parameter form

$$(5.4)\qquad p(\mathbf{u}, \mathbf{v}, \mathbf{w}) = q(\mathbf{u} - \mathbf{v}, \mathbf{w}).$$

When this holds, the insensitivity conditions (5.2) can be expressed as

$$(5.5)\qquad \theta(\mathbf{h}) = q(\mathbf{m}, \mathbf{h}) D(\mathbf{m}), \qquad \varnothing \subseteq \mathbf{m} \subseteq \mathbf{h},$$

where

$$(5.6)\qquad D(\mathbf{m}) := \sum_{\mathbf{m} \subseteq \mathbf{g} \subseteq \mathbf{S}} \theta(\mathbf{g}).$$

This simple relation leads quickly to the following characterization.

THEOREM 3. *Suppose that all active lifetimes are served simultaneously and that* (5.4) *holds. For* $\mathbf{h}_1, \mathbf{h}_2 \in G$, *let* $\mathbf{g}_1, \dots, \mathbf{g}_n$ *be an enumeration of the distinct states* $\mathbf{g}$ *for which* $\mathbf{g} \supseteq \mathbf{h}_1 \cup \mathbf{h}_2$. *A necessary and sufficient condition for insensitivity is that, for each* $\mathbf{h}_1, \mathbf{h}_2 \in G$ *for which* $n > 1$,

$$(5.7)\qquad \begin{aligned} &q(\mathbf{h}_1, \mathbf{g}_1) : q(\mathbf{h}_1, \mathbf{g}_2) : \cdots : q(\mathbf{h}_1, \mathbf{g}_n) \\ &\quad = q(\mathbf{h}_2, \mathbf{g}_1) : q(\mathbf{h}_2, \mathbf{g}_2) : \cdots : q(\mathbf{h}_2, \mathbf{g}_n). \end{aligned}$$

When this condition is satisfied, the equilibrium distribution satisfies

$$(5.8)\qquad \pi(\mathbf{h}) \propto q(\varnothing, \mathbf{h}) \prod_{\sigma \in \mathbf{h}} \mu_\sigma.$$

EXAMPLE. As a particular instance of such a system, consider a heterogeneous two-server loss system with arrival probabilities dependent on the current state. The state space of the system may be written $G = \{0, 1, 2, 3\}$, where 0 denotes the empty system, j the state in which server j only is occupied ($j = 1, 2$) and 3 the state with both servers busy.

Conditions (5.7) reduce to

$$q(0,j):q(0,3) = q(j,j):q(j,3), \qquad j = 1,2,$$

or equivalently,

$$q(0,j) + q(0,3):q(0,3) = q(j,j) + q(j,3):q(j,3) = 1:q(j,3)$$

or

$$(5.9) \qquad \begin{aligned} &[q(0,j) + q(0,3)]/q(2-j,3) \\ &\quad = q(0,3)/[q(1,3)q(2,3)], \qquad j = 1,2. \end{aligned}$$

That is, (a) the ratio of the probability that server j attracts an arrival when both servers are free, to that of attracting an arrival when server j only is free, is the same for $j = 1,2$, and (b) the common value is the right-hand side of (5.9).

The first condition may be regarded as a natural extension (allowing for multiple arrivals) of a well-known continuous-time reversibility example for two servers [see Kelly (1979), page 23]. The second condition is entirely novel.

Suppose now that the number of arrivals during a time slot is Poisson with parameter λ. Condition (a) says that a singleton arrival finding both servers free chooses either with probability one-half. Given this condition is satisfied, the second simplifies to

$$e^{-\lambda} = \frac{1 - \lambda/2}{1 + \lambda/2},$$

which is false for all positive λ. This example shows that the discrete-time $M/G/N/N$ system is not, in general, insensitive.

6. The case when only one lifetime is served at a time. Many models of practical interest (such as the discrete-time version of the processor-sharing queue) have a single server that, at each time point, allocates its service to exactly one of the existing lifetimes according to a probability distribution. Such GSMPs can be modelled in the current framework by putting

$$(6.1) \qquad c(\mathbf{s},\mathbf{g}) = 0 \quad \text{if } |\mathbf{s}| \neq 1 \text{ and } |\mathbf{g}| \geq 1.$$

Of course we must have

$$(6.2) \qquad c(\varnothing,\varnothing) = 1$$

because no lifetimes can be worked on if none is alive. With these restrictions we have that for $\varnothing \subseteq \mathbf{m} \subseteq \mathbf{t}$,

$$\psi(\mathbf{m},\mathbf{s}_2,\mathbf{t}+\mathbf{s}_2,\mathbf{h}) = \begin{cases} c(\mathbf{s}_2,\mathbf{t}+\mathbf{s}_2)p(\mathbf{t}+\mathbf{s}_2,\mathbf{s}_2,\mathbf{s}_2,\mathbf{h}), & |\mathbf{s}_2| = 1, \\ \sum_{\mathbf{s}_0 \in \mathbf{m}} c(\boldsymbol{\sigma}_0,\mathbf{t})p(\mathbf{t},\varnothing,\boldsymbol{\sigma}_0,\mathbf{h}), & \mathbf{s}_2 = \varnothing, \mathbf{m} \neq \varnothing, \\ p(\varnothing,\varnothing,\varnothing,\mathbf{h})\,\delta_{\mathbf{t},\varnothing}, & \mathbf{s}_2 = \varnothing, \mathbf{m} = \varnothing, \\ 0, & \text{otherwise.} \end{cases}$$

Here as in the remainder of the paper we write for notational convenience $\boldsymbol{\sigma} := \{\sigma\}$ (with or without a subscript) for a singleton set associated with an element $\sigma \in S$.

For $\varnothing \subseteq \mathbf{m} \subseteq \mathbf{h}$, (4.2) becomes

$$
\begin{aligned}
\theta(\mathbf{h}) &= \theta(\varnothing) p(\varnothing, \varnothing, \varnothing, \mathbf{h}) I(\mathbf{m} = \varnothing) \\
&\quad + \sum_{\mathbf{m} \subseteq \mathbf{t} \subseteq \mathbf{h}} \theta(\mathbf{t}) \sum_{\sigma_0 \in \mathbf{m}} c(\boldsymbol{\sigma}_0, \mathbf{t}) p(\mathbf{t}, \varnothing, \boldsymbol{\sigma}_0, \mathbf{h}) I(\mathbf{m} \neq \varnothing) \\
&\quad + \sum_{\mathbf{m} \subseteq \mathbf{t} \subseteq \mathbf{h}} \sum_{\sigma_2 \in \mathbf{S} - \mathbf{t}} \theta(\mathbf{t} + \boldsymbol{\sigma}_2) c(\boldsymbol{\sigma}_2, \mathbf{t} + \boldsymbol{\sigma}_2) p(\mathbf{t} + \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_2, \boldsymbol{\sigma}_2, \mathbf{h}),
\end{aligned}
\tag{6.3}
$$

where $I(\cdot)$ denotes the characteristic function.

As in the previous section, an important case is where

$$
p(\mathbf{t}, \mathbf{u}, \mathbf{w}, \mathbf{v}) = q(\mathbf{t} - \mathbf{u}, \mathbf{v}), \qquad \mathbf{u} \subseteq \mathbf{w} \text{ and } |\mathbf{w}| \leq 1. \tag{6.4}
$$

With this assumption, (6.3) reads

$$
\begin{aligned}
\theta(\mathbf{h}) &= \theta(\varnothing) q(\varnothing, \mathbf{h}) I(\mathbf{m} = \varnothing) \\
&\quad + \sum_{\mathbf{m} \subseteq \mathbf{t} \subseteq \mathbf{h}} \theta(\mathbf{t}) \sum_{\sigma_0 \in \mathbf{m}} c(\boldsymbol{\sigma}_0, \mathbf{t}) q(\mathbf{t}, \mathbf{h}) I(\mathbf{m} \neq \varnothing) \\
&\quad + \sum_{\mathbf{m} \subseteq \mathbf{t} \subseteq \mathbf{h}} \sum_{\sigma_2 \in \mathbf{S} - \mathbf{t}} \theta(\mathbf{t} + \boldsymbol{\sigma}_2) c(\boldsymbol{\sigma}_2, \mathbf{t} + \boldsymbol{\sigma}_2) q(\mathbf{t}, \mathbf{h})
\end{aligned}
\tag{6.5}
$$

for $\varnothing \subseteq \mathbf{m} \subseteq \mathbf{h}$.

The conditions for insensitivity are more complicated than in the previous section, including conditions on both the transition probabilities $q(\mathbf{u}, \mathbf{v})$ and the selection probabilities $c(\boldsymbol{\sigma}, \mathbf{t})$. We have the following necessary and sufficient condition for product form.

THEOREM 4. *Suppose that* (4.9), (6.1), (6.2) *and* (6.4) *apply. Then* (a) *and* (b) *below taken together are necessary and sufficient conditions for product form.*

(a) *For prescribed values of* $\{q(\mathbf{h}, \mathbf{h} + \boldsymbol{\sigma}); \sigma \in \mathbf{S} - \mathbf{h}\}_{\mathbf{h} \subset \mathbf{S}}$, *the quantities* $q(\mathbf{h}, \mathbf{h})$ *are given recursively* (*downward*) *by* $q(\mathbf{S}, \mathbf{S}) = 1$ *and*

$$
q(\mathbf{h}, \mathbf{h}) + \sum_{\sigma \in \mathbf{S} - \mathbf{h}} \frac{q(\mathbf{h}, \mathbf{h} + \boldsymbol{\sigma})}{q(\mathbf{h} + \boldsymbol{\sigma}, \mathbf{h} + \boldsymbol{\sigma})} = 1. \tag{6.6}
$$

The other transition probabilities $q(\mathbf{g}, \mathbf{h})$ *for* $|\mathbf{h} - \mathbf{g}| = n \geq 2$ *are fixed by*

$$
q(\mathbf{g}, \mathbf{h}) = \sum q(\mathbf{g}_{n-1}, \mathbf{h}) \prod_{i=1}^{n-1} \frac{q(\mathbf{g}_{i-1}, \mathbf{g}_i)}{q(\mathbf{g}_i, \mathbf{g}_i)}, \tag{6.7}
$$

where the summation extends over all chains $\mathbf{g} = \mathbf{g}_0 \subset \mathbf{g}_1 \subset \cdots \subset \mathbf{g}_{n-1} \subset \mathbf{h}$.

(b) *The selection probabilities are given by*

$$
c(\boldsymbol{\sigma}, \mathbf{h}) = \frac{\omega(\mathbf{h} - \boldsymbol{\sigma}) q(\mathbf{h} - \boldsymbol{\sigma}, \mathbf{h})}{q(\mathbf{h} - \boldsymbol{\sigma}, \mathbf{h} - \boldsymbol{\sigma}) q(\mathbf{h}, \mathbf{h}) \omega(\mathbf{h})}, \qquad \sigma \in \mathbf{h}, \tag{6.8}
$$

where $\omega(\cdot)$ *is defined on* $\mathbf{S}$ *recursively (upward) by* $\omega(\varnothing) = 1$ *and*

$$\omega(\mathbf{h}) = \sum_{\sigma \in \mathbf{h}} \frac{\omega(\mathbf{h} - \boldsymbol{\sigma}) q(\mathbf{h} - \boldsymbol{\sigma}, \mathbf{h})}{q(\mathbf{h} - \boldsymbol{\sigma}, \mathbf{h} - \boldsymbol{\sigma}) q(\mathbf{h}, \mathbf{h})}, \qquad \varnothing \subset \mathbf{h} \subseteq \mathbf{S}. \tag{6.9}$$

In the event of insensitivity, the equilibrium distribution takes the form

$$\theta(\mathbf{h}) = K\omega(\mathbf{h}), \tag{6.10}$$

where

$$K = 1 \Big/ \sum_{\mathbf{h} \subseteq \mathbf{S}} \omega(\mathbf{h}) \prod_{\sigma \in \mathbf{h}} \mu_\sigma. \tag{6.11}$$

Proof. Establishing the result amounts to showing that, under (6.1) and (6.2), condition (6.5) is equivalent to (6.6)–(6.9) with θ in place of ω. Condition (6.11) is just the normalization

$$1 = \sum_{\mathbf{h} \subseteq \mathbf{S}} \pi(\mathbf{h}) = \sum_{\mathbf{h} \subseteq \mathbf{S}} \theta(\mathbf{h}) \prod_{\sigma \in \mathbf{h}} \mu_\sigma.$$

For $\varnothing \subseteq \mathbf{m} = \mathbf{h}$, (6.5) gives

$$\sum_{\sigma \in \mathbf{S} - \mathbf{h}} \theta(\mathbf{h} + \boldsymbol{\sigma}) c(\boldsymbol{\sigma}, \mathbf{h} + \boldsymbol{\sigma}) = \theta(\mathbf{h}) \frac{1 - q(\mathbf{h}, \mathbf{h})}{q(\mathbf{h}, \mathbf{h})}. \tag{6.12}$$

The right-hand side of this equation, and the corresponding expression obtained when $\mathbf{h}$ is replaced by $\mathbf{h} - \boldsymbol{\sigma}$, may be used to substitute for the corresponding left-hand sides that occur in (6.5) with $\varnothing \subseteq \mathbf{m} = \mathbf{h} - \boldsymbol{\sigma}$, to provide

$$\theta(\mathbf{h}) q(\mathbf{h}, \mathbf{h}) c(\boldsymbol{\sigma}, \mathbf{h}) = \frac{\theta(\mathbf{h} - \boldsymbol{\sigma}) q(\mathbf{h} - \boldsymbol{\sigma}, \mathbf{h})}{q(\mathbf{h} - \boldsymbol{\sigma}, \mathbf{h} - \boldsymbol{\sigma})}, \qquad \sigma \in \mathbf{h}. \tag{6.13}$$

Summation over $\sigma \in \mathbf{h}$ gives

$$\theta(\mathbf{h}) q(\mathbf{h}, \mathbf{h}) = \sum_{\sigma \in \mathbf{h}} \frac{\theta(\mathbf{h} - \boldsymbol{\sigma}) q(\mathbf{h} - \boldsymbol{\sigma}, \mathbf{h})}{q(\mathbf{h} - \boldsymbol{\sigma}, \mathbf{h} - \boldsymbol{\sigma})}, \qquad \varnothing \subset \mathbf{h} \subseteq \mathbf{S}. \tag{6.14}$$

With $\omega(\cdot)$ representing a constant multiple of $\theta(\cdot)$ with the normalization $\omega(\varnothing) = 1$, relations (6.13) and (6.14) are just (6.8) and (6.9).

In the same way, relation (6.12) may be used to simplify the insensitivity condition (6.5) for general $\mathbf{m}$ satisfying $\varnothing \subseteq \mathbf{m} \subset \mathbf{h}$ to

$$\sum_{\mathbf{m} \subset \mathbf{t} \subseteq \mathbf{h}} \theta(\mathbf{t}) q(\mathbf{t}, \mathbf{h}) \sum_{\sigma_0 \in \mathbf{t} - \mathbf{m}} c(\boldsymbol{\sigma}_0, \mathbf{t}) = \sum_{\mathbf{m} \subseteq \mathbf{t} \subset \mathbf{h}} \frac{\theta(\mathbf{t}) q(\mathbf{t}, \mathbf{h})}{q(\mathbf{t}, \mathbf{t})}. \tag{6.15}$$

For $|\mathbf{m}| = |\mathbf{h}| - 1$, this is simply (6.13). Substitution for the selection probabilities from (6.13) into (6.12) reduces the latter to (6.6). Relations (6.6), (6.13) and (6.14) thus give the insensitivity conditions for $|\mathbf{m}| = |\mathbf{h}| - 1, |\mathbf{h}|$. We show now that, given these relations, (6.15) taken over $\varnothing \subseteq \mathbf{m} \subset \mathbf{h}$ is equivalent to

$$q(\mathbf{m}, \mathbf{h}) = \sum_{\sigma \in \mathbf{h} - \mathbf{m}} \frac{q(\mathbf{m} + \boldsymbol{\sigma}, \mathbf{h}) q(\mathbf{m}, \mathbf{m} + \boldsymbol{\sigma})}{q(\mathbf{m} + \boldsymbol{\sigma}, \mathbf{m} + \boldsymbol{\sigma})}, \qquad \varnothing \subseteq \mathbf{m} \subset \mathbf{h}. \tag{6.16}$$

This will suffice to establish the theorem, because (6.16) leads to (6.7) by a straightforward (downward) induction on $|\mathbf{m}|$.

The connection between (6.15) and (6.16) is also established by downward induction on $|\mathbf{m}|$. A basis is provided by $|\mathbf{m}| = |\mathbf{h}| - 1$, when (6.16) is trivial. For the inductive step, we suppose that (6.16) holds for each subset $\mathbf{t}$ of $\mathbf{h}$ properly containing $\mathbf{m}$, that is,

$$(6.17)\quad q(\mathbf{t},\mathbf{h}) = \sum_{\boldsymbol{\sigma}\in\mathbf{h}-\mathbf{t}} \frac{q(\mathbf{t}+\boldsymbol{\sigma},\mathbf{h})q(\mathbf{t},\mathbf{t}+\boldsymbol{\sigma})}{q(\mathbf{t}+\boldsymbol{\sigma},\mathbf{t}+\boldsymbol{\sigma})}, \qquad \varnothing \subseteq \mathbf{m} \subset \mathbf{t} \subset \mathbf{h}.$$

By virtue of (6.13), (6.15) may be rewritten as

$$\sum_{\mathbf{m}\subset\mathbf{t}_1\subseteq\mathbf{h}} \frac{q(\mathbf{t}_1,\mathbf{h})}{q(\mathbf{t}_1,\mathbf{t}_1)} \sum_{\boldsymbol{\sigma}_0\in\mathbf{t}_1-\mathbf{m}} \frac{\theta(\mathbf{t}_1-\boldsymbol{\sigma}_0)q(\mathbf{t}_1-\boldsymbol{\sigma}_0,\mathbf{t}_1)}{q(\mathbf{t}_1-\boldsymbol{\sigma}_0,\mathbf{t}_1-\boldsymbol{\sigma}_0)} = \sum_{\mathbf{m}\subseteq\mathbf{t}\subset\mathbf{h}} \frac{\theta(\mathbf{t})q(\mathbf{t},\mathbf{h})}{q(\mathbf{t},\mathbf{t})}$$

or, on changing to a dummy $\mathbf{t} = \mathbf{t}_1 - \boldsymbol{\sigma}_0$ on the left-hand side,

$$(6.18)\quad \begin{aligned} &\sum_{\mathbf{m}\subseteq\mathbf{t}\subset\mathbf{h}} \frac{\theta(\mathbf{t})}{q(\mathbf{t},\mathbf{t})} \sum_{\boldsymbol{\sigma}_0\in\mathbf{h}-\mathbf{t}} \frac{q(\mathbf{t}+\boldsymbol{\sigma}_0,\mathbf{h})q(\mathbf{t},\mathbf{t}+\boldsymbol{\sigma}_0)}{q(\mathbf{t}+\boldsymbol{\sigma}_0,\mathbf{t}+\boldsymbol{\sigma}_0)} \\ &\quad = \sum_{\mathbf{m}\subseteq\mathbf{t}\subset\mathbf{h}} \frac{\theta(\mathbf{t})q(\mathbf{t},\mathbf{h})}{q(\mathbf{t},\mathbf{t})}. \end{aligned}$$

If the equation resultant on summation of a multiple $\theta(\mathbf{t})/q(\mathbf{t},\mathbf{t})$ of (6.17) over all $\mathbf{t}$ with $\mathbf{m} \subset \mathbf{t} \subset \mathbf{h}$ is subtracted from (6.18), we obtain

$$\frac{\theta(\mathbf{m})}{q(\mathbf{m},\mathbf{m})} \sum_{\boldsymbol{\sigma}\in\mathbf{h}-\mathbf{m}} \frac{q(\mathbf{m}+\boldsymbol{\sigma},\mathbf{h})q(\mathbf{m},\mathbf{m}+\boldsymbol{\sigma})}{q(\mathbf{m}+\boldsymbol{\sigma},\mathbf{m}+\boldsymbol{\sigma})} = \frac{\theta(\mathbf{m})q(\mathbf{m},\mathbf{h})}{q(\mathbf{m},\mathbf{m})},$$

and the desired relation (6.16) follows at once. □

REMARKS.

1. Because

$$(6.19)\qquad \sum_{\mathbf{h}\supseteq\mathbf{g}} q(\mathbf{g},\mathbf{h}) = 1,$$

(6.6) is implicit in (6.16). To see this, sum over $\mathbf{h}$ in (6.16). We have

$$\begin{aligned} 1 &= q(\mathbf{m},\mathbf{m}) + \sum_{\mathbf{m}\subset\mathbf{h}\subseteq\mathbf{S}} q(\mathbf{m},\mathbf{h}) \\ &= q(\mathbf{m},\mathbf{m}) + \sum_{\mathbf{m}\subset\mathbf{h}\subseteq\mathbf{S}} \sum_{\boldsymbol{\sigma}\in\mathbf{h}-\mathbf{m}} \frac{q(\mathbf{m}+\boldsymbol{\sigma},\mathbf{h})q(\mathbf{m},\mathbf{m}+\boldsymbol{\sigma})}{q(\mathbf{m}+\boldsymbol{\sigma},\mathbf{m}+\boldsymbol{\sigma})} \\ &= q(\mathbf{m},\mathbf{m}) + \sum_{\boldsymbol{\sigma}\in\mathbf{S}-\mathbf{m}} \frac{q(\mathbf{m},\mathbf{m}+\boldsymbol{\sigma})}{q(\mathbf{m}+\boldsymbol{\sigma},\mathbf{m}+\boldsymbol{\sigma})} \sum_{\mathbf{h}\supseteq\mathbf{m}+\boldsymbol{\sigma}} q(\mathbf{m}+\boldsymbol{\sigma},\mathbf{h}) \\ &= q(\mathbf{m},\mathbf{m}) + \sum_{\boldsymbol{\sigma}\in\mathbf{S}-\mathbf{m}} \frac{q(\mathbf{m},\mathbf{m}+\boldsymbol{\sigma})}{q(\mathbf{m}+\boldsymbol{\sigma},\mathbf{m}+\boldsymbol{\sigma})}. \end{aligned}$$

Thus if some parameterization other than that provided in terms of the transition probabilities $\{q(\mathbf{h}, \mathbf{h} + \boldsymbol{\sigma})\}$ is employed, only the constraints (6.7) need to be considered, relation (6.19) taking account of (6.6).

2. If insensitivity is known, then the quantities $\theta(\mathbf{h})$ are found more simply via (6.13) than (6.14).
3. The equilibrium probabilities for singleton states are given from (6.14) as

$$\theta(\boldsymbol{\sigma}) = \frac{\theta(\varnothing)}{q(\varnothing,\varnothing)} \frac{q(\varnothing,\boldsymbol{\sigma})}{q(\boldsymbol{\sigma},\boldsymbol{\sigma})}, \qquad \sigma \in \mathbf{S}. \tag{6.20}$$

7. One lifetime served at a time: random choice. A special case of the previous section of some interest arises when the service facility chooses at random which of the active lifetimes it will work on, that is,

$$c(\boldsymbol{\sigma}, \mathbf{h}) = 1/|\mathbf{h}| \qquad \forall \sigma \in \mathbf{h}, \mathbf{h} \neq \varnothing.$$

Because the choice probabilities are prescribed, relation (6.13) [or (6.8)] then imposes further conditions that must be satisfied for insensitivity to occur. These are essentially consistency conditions.

For $|\mathbf{h}| = n$, (6.13) gives the solution

$$\theta(\mathbf{h}) = \theta(\varnothing)\frac{n!\, q(\mathbf{h}_{n-1}, \mathbf{h})}{q(\varnothing,\varnothing) q(\mathbf{h},\mathbf{h})} \prod_{i=1}^{n-1} \frac{q(\mathbf{h}_{i-1}, \mathbf{h}_i)}{q(\mathbf{h}_i, \mathbf{h}_i)^2}, \qquad \varnothing \subset \mathbf{h} \subseteq \mathbf{S}, \tag{7.1}$$

for any chain of states $\varnothing = \mathbf{h}_0 \subset \mathbf{h}_1 \subset \cdots \subset \mathbf{h}_n = \mathbf{h}$ with $|\mathbf{h}_i - \mathbf{h}_{i-1}| = 1$. For $n = 1$, the empty product is taken as unity. The condition imposed by (6.13) is that, for each given $\mathbf{h}$, the value of $\theta(\mathbf{h})$ be independent of the choice of chain.

Any such chain of states may be associated in a one-to-one manner with a corresponding sequence $\mathbf{h}_1 - \mathbf{h}_0, \mathbf{h}_2 - \mathbf{h}_1, \dots, \mathbf{h}_n - \mathbf{h}_{n-1}$ of distinct active elements of $\mathbf{S}$. Because any ordering of n elements can be converted into any other ordering of the same n elements by a sequence of permutations of adjacent elements, it suffices for consistency that the value of $\theta(\mathbf{h})$ be the same for any two chains of the form

$$\varnothing = \mathbf{h}_0 \subset \mathbf{h}_1 \subset \cdots \subset \mathbf{h}_n = \mathbf{h},$$
$$\varnothing = \mathbf{h}'_0 \subset \mathbf{h}'_1 \subset \cdots \subset \mathbf{h}'_n = \mathbf{h},$$

for which $\mathbf{h}'_i = \mathbf{h}_i$ for all i not equal to some j, $0 < j < n$. If we put

$$\mathbf{h}_j = \mathbf{h}_{j-1} + \boldsymbol{\sigma}, \qquad \mathbf{h}_{j+1} = \mathbf{h}_j + \boldsymbol{\sigma}', \qquad \boldsymbol{\sigma} \neq \boldsymbol{\sigma}',$$

then

$$\mathbf{h}'_j = \mathbf{h}'_{j-1} + \boldsymbol{\sigma}', \qquad \mathbf{h}'_{j+1} = \mathbf{h}_j + \boldsymbol{\sigma}.$$

For consistency, we have from substitution in (7.1) that

$$\frac{q(\mathbf{g}, \mathbf{g} + \boldsymbol{\sigma}) q(\mathbf{g} + \boldsymbol{\sigma}, \mathbf{g} + \boldsymbol{\sigma} + \boldsymbol{\sigma}')}{q(\mathbf{g} + \boldsymbol{\sigma}, \mathbf{g} + \boldsymbol{\sigma})} = \frac{q(\mathbf{g}, \mathbf{g} + \boldsymbol{\sigma}') q(\mathbf{g} + \boldsymbol{\sigma}', \mathbf{g} + \boldsymbol{\sigma} + \boldsymbol{\sigma}')}{q(\mathbf{g} + \boldsymbol{\sigma}', \mathbf{g} + \boldsymbol{\sigma}')}, \tag{7.2}$$

where $\mathbf{g} = \mathbf{h}_{j-1}$. Taken over all chains, the consistency condition is thus that (7.2) holds for all $\mathbf{g}$ with $|\mathbf{g}| < |\mathbf{S}| - 1$ and $\sigma, \sigma' \in \mathbf{S} - \mathbf{g}$ $(\sigma \neq \sigma')$.

8. The models of Daduna and Schassberger. Daduna and Schassberger (1981, 1983) considered special cases of GSMPs in which it is not permitted for a transition to occur when the lifetime of more than one element of $\mathbf{h}$ is reduced to zero at the same time point. The mechanism they employed is to assume that such elements are allocated a new lifetime sampled from the original distribution. In our current model this is equivalent to putting

$$p(\mathbf{h}, \mathbf{s}_2, \mathbf{s}, \mathbf{h}') = \delta_{\mathbf{h}, \mathbf{h}'}$$

for all $\mathbf{s}_0$ and $\mathbf{s}_2$ with $|\mathbf{s}_2| > 1$. With this assumption we have from (4.1) that

$$\psi(\mathbf{m}, \mathbf{s}_2, \mathbf{u}, \mathbf{h}) = \sum_{\mathbf{s}_0 \subseteq \mathbf{m}} c(\mathbf{s}, \mathbf{u})\, \delta_{\mathbf{u}, \mathbf{h}} \quad \text{for } |\mathbf{s}_2| > 1,$$

and substitution into the more physical version (4.6) of the insensitivity equations yields

$$\begin{aligned} \theta(\mathbf{h}) = & \sum_{\mathbf{s}_1 \subseteq \mathbf{k}} \sum_{\substack{\mathbf{s}_2 \subseteq (\mathbf{S}-\mathbf{h})+\mathbf{s}_1 \\ |\mathbf{s}_2| \leq 1}} \theta(\mathbf{g})\, \psi(\mathbf{h} - \mathbf{k}, \mathbf{s}_2, \mathbf{g}, \mathbf{h}) \\ & + \sum_{\substack{\mathbf{s}_1 \subseteq \mathbf{k} \\ |\mathbf{s}_1| > 1}} \theta(\mathbf{h}) \sum_{\mathbf{s}_0 \subseteq \mathbf{h}-\mathbf{k}} c(\mathbf{s}_0 + \mathbf{s}_1, \mathbf{h}). \end{aligned} \tag{8.1}$$

Note that, for all $\mathbf{h}$ and $\mathbf{k} \subseteq \mathbf{h}$,

$$\sum_{\mathbf{s}_1 \subseteq \mathbf{k}} \sum_{\mathbf{s}_0 \subseteq \mathbf{h}-\mathbf{k}} c(\mathbf{s}_0 + \mathbf{s}_1, \mathbf{h}) = 1.$$

From this, (8.1) can be rearranged as

$$\theta(\mathbf{h}) \sum_{\substack{\mathbf{s}_1 \subseteq \mathbf{k} \\ |\mathbf{s}_1| \leq 1}} \sum_{\mathbf{s}_0 \subseteq \mathbf{h}-\mathbf{k}} c(\mathbf{s}_0 + \mathbf{s}_1, \mathbf{h}) = \sum_{\mathbf{s}_1 \subseteq \mathbf{k}} \sum_{\substack{\mathbf{s}_2 (\mathbf{S}-\mathbf{h})+\mathbf{s}_1 \\ |\mathbf{s}_2| \leq 1}} \theta(\mathbf{g})\, \psi(\mathbf{h} - \mathbf{k}, \mathbf{s}_2, \mathbf{g}, \mathbf{h}).$$

If we assume, as in Section 5, that $c(\mathbf{s}, \mathbf{g}) = \delta_{\mathbf{s}, \mathbf{g}}$, then the left-hand side of (8.2) is zero whenever $|\mathbf{k}| > 1$. With the notation of Section 5 and use of (5.1), the right-hand side becomes

$$\sum_{\substack{\mathbf{s}_2 \subseteq (\mathbf{S}-\mathbf{h})+\mathbf{k} \\ |\mathbf{s}_2| \leq 1}} \theta((\mathbf{h} - \mathbf{k}) + \mathbf{s}_2) p((\mathbf{h} - \mathbf{k}) + \mathbf{s}_2, \mathbf{s}_2, \mathbf{h})$$

and it follows that the only way that (8.1) can be satisfied is if $p((\mathbf{h} - \mathbf{k}) + \mathbf{s}_2, \mathbf{s}_2, \mathbf{h}) = 0$ whenever $|\mathbf{k}| > 1$. Hence in these circumstances the GSMP can have product form only if, in addition to no transitions occurring when more than one lifetime dies, it is impossible for more than one lifetime to be created at the same time point.

Moreover, for GSMPs satisfying this condition, only the conditions for $|\mathbf{k}| = 0, 1$ are nontrivial. With these two choices (and $\mathbf{k} = \boldsymbol{\sigma}_1$, say, in the latter), (8.2) reduces, respectively, to

$$\pi(\mathbf{h}) = \pi(\mathbf{h})p(\mathbf{h},\varnothing,\mathbf{h}) + \sum_{\boldsymbol{\sigma}_2 \in \mathbf{S}-\mathbf{h}} \frac{\pi(\mathbf{h}+\boldsymbol{\sigma}_2)}{\mu_{\sigma_2}} p(\mathbf{h}+\boldsymbol{\sigma}_2,\boldsymbol{\sigma}_2,\mathbf{h}), \tag{8.3}$$

$$\frac{\pi(\mathbf{h})}{\mu_{\sigma_1}} = \pi(\mathbf{h}-\boldsymbol{\sigma}_1)p(\mathbf{h}-\boldsymbol{\sigma}_1,\varnothing,\mathbf{h}) + \sum_{\boldsymbol{\sigma}_2 \in (\mathbf{S}-\mathbf{h})+\boldsymbol{\sigma}_1} \frac{\pi((\mathbf{h}-\boldsymbol{\sigma}_1)+\boldsymbol{\sigma}_2)}{\mu_{\sigma_2}} p((\mathbf{h}-\boldsymbol{\sigma}_1)+\boldsymbol{\sigma}_2,\boldsymbol{\sigma}_2,\mathbf{h}). \tag{8.4}$$

Equations (8.3) and (8.4) are the familiar balance equations necessary and sufficient for product form and insensitivity in a continuous-time GSMP with unit speeds and a unit rate at which exponential transitions take place in any state [see, e.g., König and Jansen (1974)]. We have thus shown that a discrete-time GSMP, operating according to the rule of Daduna and Schassberger that all lifetimes start again if multiple deaths occur, possesses product form if and only if a related continuous-time GSMP possesses product form. Many examples of the latter are known, yielding many examples of the former.

9. Conclusion. In this paper we have considered discrete-time GSMPs in which multiple lifetimes can die simultaneously. Probabilistic choice of service elements is allowed. We have given necessary and sufficient conditions for such GSMPs to have a product-form supplemented equilibrium distribution of probability, which are thus sufficient for insensitivity. These conditions are seen to be much more stringent than in continuous-time theory.

Specific applications are made to models in which all lifetimes are worked on simultaneously, exactly one lifetime is worked on at a time and to the models of Daduna and Schassberger (1981, 1983).

Acknowledgment. The authors thank the Australian Research Council for supporting this research.

REFERENCES

BURMAN, D. (1981). Insensitivity in queueing systems. *Adv. in Appl. Probab.* **13** 846–859.

DADUNA, H. and SCHASSBERGER, R. (1981). A discrete-time round-robin queue with Bernoulli input and general arithmetic service time distributions. *Acta Inform.* **15** 251–263.

DADUNA, H. and SCHASSBERGER, R. (1983). Networks of queues in discrete time. *Z. Oper. Res.* **27** 159–175.

FAKINOS, D. (1990). On the relation between partial balance and insensitivity. *J. Appl. Probab.* **27** 445–451.

FRANKEN, P., KÖNIG, D., ARNDT, U. and SCHMIDT, V. (1982). *Queues and Point Processes.* Wiley, London.

HENDERSON, W. (1983). Insensitivity and reversed Markov processes. *Adv. in Appl. Probab.* **15** 752–768.

HENDERSON, W. and TAYLOR, P. G. (1991). Insensitivity in discrete time queues with a moving server. *Queueing Systems Theory Appl.* **11** 273–297.

KEILSON, J. (1979). *Markov Chain Models—Rarity and Exponentiality.* Springer, New York.

KELLY, F. (1979). *Reversibility and Stochastic Networks.* Wiley, London.

KÖNIG, D. and JANSEN, U. (1974). Stochastic processes and properties of invariance for queueing systems with speeds and temporary interruptions. In *Transactions of the 7th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes* **A** 335–343. Academia Publishing House of the Czechoslovak Academy of Sciences, Prague.

MIYAZAWA, M. (1991). The characterization of the stationary distribution of the supplemented self-clocking jump process. *Math. Oper. Res.* **16** 547–565.

MIYAZAWA, M. (1993). Insensitivity and product-form decomposability of reallocatable GSMP. *Adv. in Appl. Probab.* **25** 415–437.

SCHASSBERGER, R. (1977). Insensitivity of steady-state distributions of generalized semi-Markov processes. Part I. *Ann. Probab.* **5** 87–89.

SCHASSBERGER, R. (1978a). Insensitivity of steady-state distributions of generalized semi-Markov processes. Part II. *Ann. Probab.* **6** 85–93.

SCHASSBERGER, R. (1978b). Insensitivity of steady-state distributions of generalized semi-Markov processes with speeds. *Adv. in Appl. Probab.* **10** 836–851.

SCHASSBERGER, R. (1986). Two remarks on insensitive stochastic models. *Adv. in Appl. Probab.* **18** 791–814.

TAYLOR, P. G. (1989). Insensitivity in processes with zero speeds. *Adv. in Appl. Probab.* **21** 612–628.

WHITTLE, P. (1985). Partial balance and insensitivity. *J. Appl. Probab.* **22** 168–176.

WHITTLE, P. (1986). Partial balance, insensitivity and weak coupling. *J. Appl. Probab.* **18** 706–723.

W. HENDERSON
C. E. M. PEARCE
P. G. TAYLOR
DEPARTMENT OF APPLIED MATHEMATICS
UNIVERSITY OF ADELAIDE
ADELAIDE
SOUTH AUSTRALIA 5005

N. M. VAN DIJK
FACULTY OF ECONOMICAL SCIENCES
AND ECONOMETRICS
FREE UNIVERSITY
P.O. BOX 716
1007 M.C., AMSTERDAM
THE NETHERLANDS