The Annals of Applied Probability
1998, Vol. 8, No. 3, 793–821

MARKOV NETWORK PROCESSES WITH STRING TRANSITIONS[1]

BY RICHARD F. SERFOZO AND BINGYI YANG

Georgia Institute of Technology and CAPS Logistics, Inc.

This study introduces a Markov network process called a string-net. Its state is the vector of quantities of customers or units that move among the nodes, and a transition of the network consists of a string of instantaneous vector increments in the state. The rate of such a string transition is a product of a transition-initiation rate and a string-generation rate. The main result characterizes the stationary distribution of a string-net. Key parameters in this distribution satisfy certain "polynomial traffic equations" involving the string-generation rates. We identify sufficient conditions for the existence of a solution of the polynomial equations, and we relate these equations to a partial balance property and throughputs of the network. Other results describe the stationary behavior of a large class of string-nets in which the vectors in the strings are unit vectors and a string-generation rate is a product of Markov routing probabilities. This class includes recently studied open networks with Jackson-type transitions augmented by transitions in which a signal (or negative customer) deletes units at nodes in one or two stages. The family of string-nets contains essentially all Markov queueing network processes, aside from reversible networks, that have known formulas for their stationary distributions. We discuss old and new variations of Jackson networks with batch services, concurrent or multiple-unit movements of units, state-dependent routings and multiple types of units and routes.

1. Introduction. There are two major classes of Markov queueing network processes that have known formulas for their stationary or equilibrium distributions. One class consists of classical Jackson–Gordon–Newell–Whittle processes; BCMP-Kelly processes with multiple types of units and routes; and processes with state-dependent routings, batch services and concurrent or multiple-unit movements [2, 3, 17, 22–25, 27, 29, 31, 32]. The other class is the family of reversible Markov network processes [25, 32], including those with multiple-unit movements [1, 28].

Recently, several more examples of Markov queueing network processes with tractable stationary distributions have been identified [4, 8–14, 16, 18, 20, 21]. The novelty of these processes is that, in addition to the usual movements of units after service completions as in Jackson networks, there are signals (sometimes called negative customers) that periodically trigger auxiliary departures of units from the network. These departures are single- or double-stage departures of single units (or vectors) randomly selected by Markov probabilities. Although the resulting stationary distributions are similar in

Received May 1996; revised October 1997.

[1]Supported in part by Air Force Office of Scientific Research Contract 91-0013 and NSF Grants DDM-92-24520, DDM-92-15233 and DMI-94-57336.

AMS 1991 *subject classifications*. Primary 60F05; secondary 60K25.

Key words and phrases. Markov processes, queueing processes, stochastic networks, stationary distributions, batch arrivals and services.

form to those of Jackson processes, the main parameters in the distribution are obtained as solutions of a *traffic equation* that differs from the "linear" Jackson traffic equations. The first example in this series is an open Jackson network with single-server nodes and auxiliary deletions of size 1 or 2 (see Example 18). This model was the subject of [11]–[14] and [16]. Then [8] extended this model to more general services and deletion mechanisms (see Examples 17 and 19) and gave insights on quasi-reversibility. The model was extended in [21] to one or two-stage "batch" deletions and [6] and [7] gave further insights into similar quasi-reversible networks. Other related references are [2, 3, 5, 19, 22, 26].

These studies suggest that there is a broader class of Markov queueing network processes with tractable stationary distributions that contains these new examples as well as the Jackson network processes. We introduce such a class called *Markov network processes with string transitions* or *string-nets* for short. In a string-net, a transition of the network is determined by a string of vectors representing multistage subtractions or additions of vector quantities at the nodes and all of this is done instantaneously in a transition. The strings are randomly selected from an arbitrary family of variable length strings by a general probability measure. Jackson networks have only "one-stage transitions" represented by strings whose lengths are exactly one. The other newer examples mentioned above have strings of length 1 or 2, and one example has infinite strings of unit vectors selected by Markov probabilities. String-nets cover a variety of networks with batch or concurrent movements of units where the batch and related probabilities are determined by a string of information. Additionally, unit movements need not be independent as in classical networks. For instance, in "time-warp" parallel simulation programs, when the processing gets out of sync, a message is passed among the processors (nodes) to instantaneously delete superfluous data (units) they might have generated.

In a m-node string-net, the state of the network is a vector $x = (x_1, \dots, x_m)$ that denotes the numbers of units at the nodes (Remark 27 explains that our results also apply to multiple types of units). A transition of the network from x to y has a rate of the form $\sum_{sa} \lambda_{sa} r_{sa}(x, y)$. The λ_{sa} is the rate (or probability or propensity) of selecting a string of vectors s and another add-on vector a as the increments in the state. The $r_{sa}(x, y)$ is a system-dependent transition-initiation rate that may represent service rates at the nodes plus other transition information. With a slight abuse of notation, we call both of these quantities "rates" even though they are only parts of the compound rate $\lambda_{sa} r_{sa}(x, y)$. This compounding or weak coupling of the string-generation and transition-initiation information also arises in the invariant measures (or stationary distributions) of the process. Namely, Theorem 1 gives necessary and sufficient conditions for a string-net to have an invariant measure of the form $\Phi(x) \prod_{j=1}^{m} w_j^{x_j}$. The Φ is determined solely by the transition-initiation rates $r_{sa}(x, y)$, and the parameters w_j are a solution to certain polynomial equations involving the string-generation rates λ_{sa}. We give sufficient conditions for the existence of a solution to these so-called traffic equations and show that the equations are equivalent to equalities of certain average flows in the network (a partial balance property). We also give a complete description of the stationary behavior of a large subclass of string-nets in which the vectors in the strings are unit vectors generated by Markov probabilities.

The organization of our study is as follows. After defining string-nets in Section 2, we characterize their invariant measures in Section 3 and discuss their traffic equations in Section 4, where we also derive formulas for throughputs. Section 5 shows how the results apply to string-nets whose strings consist of only unit vectors; the traffic equations simplify considerably. Section 6 discusses string-nets in which all the strings are of length one; Jackson processes are of this type. Sections 7 and 8 cover string-nets with unit-vector strings selected by Markov probabilities, including multiple types of string selections. The multiplicative structure of the probability of selecting a string is similar to the multiplicative or compounding structure of reversible networks with multiple-unit movements [28]. Section 8 also relates earlier examples in the literature mentioned above to string-nets. Section 9 covers string-nets in which at most two nodes are affected by a transition. Examples are batch-assembly models and a "Jackson network with periodic clearing" in which a transition might involve clearing out all units at a node. In Section 10, we discuss one-dimensional Markov processes with string transitions; they are of interest by themselves and may also be used to represent single nodes that form parts of a network. We end by discussing how the results in this study apply to multiple types of units by simply appending another parameter to the network data and formulas.

2. Definition of string-nets. The Markov network process with string transitions that we shall study is defined as follows.

Consider a network in which discrete units (or customers) move among m nodes (or service stations) where they are processed. The network may be open or closed and several units may move at the same time. We will study the stochastic process $X = \{X(t): t \ge 0\}$ that represents the numbers of units at the respective nodes. The state space is a set E of m-dimensional vectors $x = (x_1, \dots, x_m)$, where x_j denotes the number of units at node j. We place no further assumptions on the form of E, and so our results apply to a variety of network types, including the standard ones that are closed ($|x| = x_1 + \cdots + x_m = N$), open with finite capacity ($|x| \le N$), or open with infinite capacity ($|x| < \infty$). Negative numbers of units at nodes are also permissible. All of our results also apply to multiple types of units that may change type as they move; see Remark 27.

We assume that X is a Markov jump process and hence its behavior is characterized by its transition rates

$$q(x, y) \equiv \lim_{t \to 0} t^{-1} P\{X(t) = y | X(0) = x\}, \qquad x \ne y \in E.$$

The $q(x, y)$ is the rate of a transition $x \to y$ (from x to y). The sojourn time in state x is exponentially distributed with rate $\sum_{y \ne x} q(x, y)$ and at the end of the sojourn, the process jumps to some state $y \in E$ with probability $q(x, y) / \sum_{y \ne x} q(x, y)$.

The dynamics of the network and the form of its transition rates are as follows. Whenever the process is in state x, a typical transition will be to some state of the form $x - (s^1 + \cdots + s^l) + a$ or $x - (s^1 + \cdots + s^k)$, $1 \le k < l$, where the increment vectors a and s^i are in a set A and the string $s = (s^1, \ldots, s^l)$ is in a set $\mathscr{S}$. The A is a finite set of m-dimensional vectors with negative or nonnegative integer entries and A contains the zero vector 0. Positive and negative entries in an increment vector stand for additions or deletions of units (e.g., entries of 5 and -3 in positions j and j' of an increment vector mean that 5 units arrive to node j and 3 units depart from j'). The $\mathscr{S}$ is a countable set of *strings* $s = (s^1, \ldots, s^l)$, where $s^i \in A \backslash \{0\}$ and $l \equiv l(s)$ denotes the string length. Let $L \le \infty$ denote the supremum of these string lengths. Assume that $\mathscr{S}$ contains the empty or zero string, denoted by 0, whose length is zero.

Associated with each string $s \in \mathscr{S}$ are its kth partial sum vectors

$$s(k) = \sum_{i=1}^{k} s^i, \qquad 0 \le k \le l,$$

where $s(0) = 0$ for the zero string. Denote the set of all partial sums of the strings by $S = \{s(k)\colon 1 \le k \le l,\ s \in \mathscr{S}\}$. Think of A as the set of allowable increment vectors and $\mathscr{S}$ as the set of feasible strings of vectors from A that can be subtracted in a transition. Then S (which contains A) is the entire set of network increment vectors. For each $x \in E$ and $d, a \in S$, we define the vector $T_{da}x = x - d + a$, which may or may not be in E. A transition $x \to T_{da}x$ means that the vectors a and d are added and subtracted from x.

In terms of this notation, the transitions of the process X are as follows. Whenever the process X is in state x, a transition is determined by a pair sa in $\mathscr{S} \times A$ that results in one of the following l possibilities.

A *complete sa-transition*: $x \to T_{s(l)a}x = x - (s^1 + \cdots + s^l) + a$.
A *kth partial sa-transition*: $x \to T_{s(k)0}x = x - (s^1 + \cdots + s^k)$, $0 \le k < l$.

Keep in mind that l, with s suppressed, is the length of the string s. Note that the complete sa-transition uses a as well as the whole string s, but the kth partial sa-transition uses only the part $s^1, \ldots, s^k$ of s. Some of these transitions may be infeasible as discussed below. Under the preceding assumptions, each state $x \in E$ is a linear combinations of vectors in A. Assume that the standard m-dimensional unit vectors form a basis that generates the vectors in A and E. This is not a restriction, since one can always represent these vectors by a basis and the form of the basis is not important here.

We assume the rates of these string transitions are as follows.

Type of transition	Rate
complete sa-transition $x \to T_{s(l)a}x$	$\lambda_{sa}\phi_{s(l)}(x)$,
kth partial sa-transition $x \to T_{s(k)0}x$	$\lambda_{sa}(\phi_{s(k)}(x) - \phi_{s(k+1)}(x))$.

These transition rates can be viewed as the compounding of two rates as follows. The nonnegative λ_{sa} is the rate at which an sa-transition occurs, where

$\lambda_{00} = 0$. In addition, within an sa-transition, $\phi_{s(l)}(x)$ is the nonnegative rate of subtracting the complete vector $s(l)$ from x and adding a; and $\phi_{s(k)}(x) - \phi_{s(k+1)}(x)$ is the rate of subtracting exactly $s(k)$ (the kth partial of s) from x, where $0 \le k < l$. A compounding of these two rates yields the transition rates above. The motivation for these types of rates will become clearer as we cover examples. For each $d \in S$, the ϕ_d is a nonnegative function on the space E. For convenience, we extend its definition to all integer-valued m-dimensional vectors by setting $\phi_d(x) = 0$, for $x \notin E$.

The preceding description says that whenever the Markov process X is in state x, the times to the next complete sa-transition and kth partial sa-transition are independent, exponentially distributed with rates shown in the table above. Then the time to the next $x \to y$ transition is exponentially distributed and its rate $q(x, y)$ is the sum of appropriate sa-transition rates. That is, the transition rates of the process X are

$$q(x, y) = \sum_{s,\, a} \lambda_{sa} r_{sa}(x, y), \qquad y \neq x \text{ in } E, \tag{1}$$

where

$$\begin{aligned} r_{sa}(x, y) &= \phi_{s(l)}(x) 1(y = T_{s(l)a} x) \\ &\quad + \sum_{k=0}^{l-1} [\phi_{s(k)}(x) - \phi_{s(k+1)}(x)] 1(y = T_{s(k)0} x), \qquad x, y \in E. \end{aligned} \tag{2}$$

Here 1(statement) denotes the indicator function that is 1 or 0 according as the "statement" is true or false. All sums on s, a herein are for $s \in \mathscr{S}$ and $a \in A$, unless specified otherwise and $\sum_{k=0}^{-1} = 0$. Since a transition $x \to y$ is possible under several combinations of subtractions and additions, its rate $q(x, y)$ is a sum of rates, some of which may be 0 due to the λ_{sa}, ϕ_d or the indicator functions being 0. The rate functions λ_{sa} and ϕ_d as well as the sets A and $\mathscr{S}$ can have a variety of forms depending on the routing and service rules of the network. For instance, for a closed network, the rate λ_{sa} can be positive only if $|s(k)| = |a - s(l)| = 0$, for $1 \le k < l$.

Note that the rate of the exponential sojourn time in state x is

$$\sum_{y \neq x} q(x, y) = \sum_{s,\, a} \lambda_{sa} [\phi_0(x) - r_{sa}(x, x)]. \tag{3}$$

This follows since

$$\sum_{y \in E} r_{sa}(x, y) = \phi_0(x), \tag{4}$$

which is due to the telescoping series in (2) and the fact that the sum of the indicators over y is 1.

To complete the definition of the process, more assumptions are in order. First, we assume that $\sum_s \lambda_{sa} < \infty$ for each $a \in A$. This and the finiteness of A ensure that the rate (3) is finite. Second, we adopt the standard assumption that the process is irreducible on the space E (otherwise, we could let E denote a closed communicating class). Finally, we assume the family of rate functions ϕ_d, $d \in S$, satisfy the following properties.

Φ-Reversibility. There exists a positive function Φ on E such that

$$\Phi(T_{ad}x)\phi_d(T_{ad}x) = \Phi(x)\phi_a(x), \qquad x \in E,\ d \in S \quad \text{and} \quad a \in A \text{ with } \sum_s \lambda_{sa} > 0. \tag{5}$$

Domination of ϕ_0. If $L \geq 2$, then $\phi_0(x) \geq \phi_a(x)$, for each $x \in E$ and $a \in A$.

Note that the ϕ_d's are Φ-reversible if they are of the form

$$\phi_d(x) = \gamma(x-d)/\Phi(x), \qquad d \in S,\ x \in E, \tag{6}$$

for some function γ that is nonnegative on E and is 0 outside of E. This form, which has been used in many of the studies mentioned in the introduction, can also be used to incorporate state-dependent routing probabilities [27]. These ϕ_d's also satisfy the ϕ_0-dominance assumption when γ is nonincreasing and each vector $a \in A$ is nonnegative.

We adopt the name "Φ-reversibility" because if the ϕ_d are Φ-reversible, then a Markov jump process with transition rates $\tilde{q}(x, y) = \sum_{d,\, a \in A} \phi_d(x) 1(y = T_{da}x)$ is reversible with respect to Φ [i.e., $\Phi(x)\tilde{q}(x, y) = \Phi(y)\tilde{q}(y, x)$]. Note that (5) implies $\phi_a(x) = 0$ when $T_{ad}x \notin E$ for some $d \in S$, because $\phi_d(x') = 0$ when $x' \notin E$. This says that an sa-transition in state x is not feasible or is blocked if any one of the possible new states resulting from a complete or partial transition is not in E.

From the Φ-reversibility and ϕ_0-dominance, it follows that

$$\begin{aligned}\Phi(T_{0d}x)\phi_{d+a}(T_{0d}x) &= \Phi(x)\phi_a(x) \leq \Phi(x)\phi_0(x)\\ &= \Phi(T_{0d}x)\phi_d(T_{0d}x), \qquad d \in S,\ a \in A.\end{aligned}$$

Thus $\phi_d(x) \geq \phi_{d+a}(x)$. This ensures that the rates in the second sum in (1) are not negative.

We say that the process X defined above with rates (1) that satisfy the preceding assumptions is a *Markov network process with string transitions* or a *string-net*. The data for this process are E, A, $\mathscr{S}$, L, $\{\lambda_{sa}: s \in \mathscr{S}, a \in A\}$ and $\{\phi_d(\cdot): d \in S\}$. To model an actual network with this process, one would specify this data from the operational features of the network.

3. Invariant measures of string-nets. In this section, we characterize invariant measures of the process X when the polynomial "traffic equations" (8) have a solution. These are equations of certain traffic flows, and conditions under which they have a solution are given in the next section.

In addition to the notation above, we denote the rate of all string transitions with s as the initial segment by

$$\Lambda_s = \sum_{s',\, a} \lambda_{(ss')a}, \qquad s \in \mathscr{S}, \tag{7}$$

where the string (ss') denotes the concatenation of the strings s and s'. We sometimes use $\Lambda_{(sa)}$ for $sa \in \mathscr{S}$, where $\Lambda_{(0a)} \equiv \Lambda_a$. We are now ready for our main result.

THEOREM 1. *If there exist positive numbers $w_1, \ldots, w_m$ that satisfy*

$$\eta(a) \sum_s \prod_{k=1}^{l} \eta(s^k)\Lambda_{(sa)} = \sum_s \prod_{k=1}^{l} \eta(s^k)\lambda_{sa}, \qquad a \in A_0 \equiv A\backslash\{0\}, \tag{8}$$

where $\eta(x) = \prod_{j=1}^m w_j^{x_j}$ and the sums in (8) *are finite, then $\pi(x) = \Phi(x)\eta(x)$, $x \in E$, is an invariant measure for the network process X. Furthermore, a necessary and sufficient condition for the process to have an invariant measure of this form is*

$$\sum_{a \in A_0} D(a)[\phi_0(x) - \phi_a(x)\eta(a)^{-1}] = 0, \qquad x \in E, \tag{9}$$

where $D(a)$ denotes the right side of (8) *minus its left side.*

In the two-part measure $\pi(x) = \Phi(x)\eta(x)$, the Φ is determined only by the ϕ_a's, and η is determined only by the λ_{sa}'s. This product or "weak coupling"of Φ and η comes from the product of the ϕ_a's and λ_{sa}'s in the transition rates q and the Φ-reversibility of the ϕ_a's. This weakly coupled distribution π is sometimes called a product form distribution because η has a geometric product form. Note that π is indeed a product form if Φ is, which is the case for the classical Jackson network. However, π may be any measure on E [e.g., the network with $\phi_d(x) = \pi(x-d)/\pi(x)$ and $\lambda_{sa} \equiv 1$, has π as an invariant measure].

From a key identity (11) in the proof below, it follows that the summation in (9) times $\pi(x)$ is the difference between the two sides of the balance equations for the process X (this should be 0 for the balance equations to be satisfied). Note that the summation is a weighted average of the differences $D(a)$ of the two sides of the traffic equations (8). The weights $\phi_0(x) - \phi_a(x)\eta(a)^{-1}$, which arise in (11), do not seem to have any special meaning.

PROOF. The balance equations that an invariant measure π must satisfy are

$$\pi(x) \sum_{y \in E} q(x, y) = \sum_{y \in E} \pi(y) q(y, x), \qquad x \in E. \tag{10}$$

The usual convention is that $q(x, x) = 0$, but here we define $q(x, x) = \sum_{s,a} \lambda_{sa} r_{sa}(x, x)$. This does not affect the equality and it simplifies some expressions.

Let $L(x)$ and $R(x)$ denote the left and right sides of (10), respectively, and suppose $\pi(x) = \Phi(x)\eta(x)$. The proof will proceed as follows. A short calculation yields $L(x) = \pi_0(x)\Lambda_0$, and more complicated analysis of $R(x)$ yields the identity

$$L(x) = R(x) + \pi(x) \sum_{a \neq 0} D(a)[\phi_0(x) - \phi_a(x)\eta(a)^{-1}]. \tag{11}$$

From this it follows that if $D(a) = 0$, $a \in A_0$, then $L(x) = R(x)$, $x \in E$, and hence π is an invariant measure of the process. This proves the first assertion of the theorem. Also, π is an invariant measure if and only if the last summation in (11) is 0. This proves the second assertion of the theorem.

It remains to prove (11). Using the transition rate formulas (1), (2) and the property $x = T_{dd'}y$ if and only if $y = T_{d'd}x$, it follows that the right side of (10) is

$$\begin{aligned} R(x) &= \sum_{y\in E} \pi(y) \sum_{s,a} \lambda_{sa} r_{sa}(y, x) = \sum_{s,a} \pi(T_{as(l)}x)\lambda_{sa} r_{sa}(T_{as(l)}x, x) \\ &= \sum_{s,a} \pi(T_{as(l)}x)\lambda_{sa}\phi_{s(l)}(T_{as(l)}x) \\ &\quad + \sum_{s,a}\sum_{k=0}^{l-1} \pi(T_{0s(k)}x)\lambda_{sa}[\phi_{s(k)}(T_{0s(k)}x) - \phi_{s(k+1)}(T_{s^{k+1}s(k+1)}x)]. \end{aligned} \tag{12}$$

Here we also use our convention that the functions π and ϕ_d are defined to be zero outside of E and that $T_{0s(k)}x = T_{s^{k+1},\, s(k+1)}x$. Now, the Φ-reversibility assumption and $\pi(x) = \Phi(x)\eta(x)$ and $\eta(x+y) = \eta(x)\eta(y)$, ensure that

$$\pi(T_{ad}x)\phi_d(T_{ad}x) = \pi(x)\phi_a(x)\eta(d)\eta(a)^{-1}, \qquad x \in E,\ a \in A,\ d \in S. \tag{13}$$

Applying this to (12), we obtain

$$\begin{aligned} R(x) &= \pi(x)\sum_{s,a} \eta(s(l))\eta(a)^{-1}\lambda_{sa}\phi_a(x) \\ &\quad + \pi(x)\sum_{s,a}\lambda_{sa}\sum_{k=0}^{l-1}\eta(s(k))[\phi_0(x) - \phi_{s^{k+1}}(x)]. \end{aligned} \tag{14}$$

To proceed, we need a convenient expression for the last sum on s, a, k. Note that $s = 0$ has no contribution to the sum, and hence we ignore it. Also, any $s \neq 0$ can be written as the concatenation $s = (s'as'')$ for some $s', s'' \in \mathscr{S}$ and $a \in A_0$. Now, make the change-of-variables $s^{k+1} = a$ and $sa = (s'as'')a'$ and reverse the order of the summations and recall the definition of Λ_s. Then the last sum in (14) becomes

$$\begin{aligned} &\sum_{s', a\neq 0} \eta(s'(l')) \sum_{s'', a'} \lambda_{(s'as'')a'}[\phi_0(x) - \phi_a(x)] \\ &\quad = \sum_s \eta(s(l)) \sum_{a\neq 0} \Lambda_{(sa)}[\phi_0(x) - \phi_a(x)]. \end{aligned} \tag{15}$$

Substituting this in (14) and recalling that $D(a)$ equals the right side of (8) minus its left side, we arrive at

$$\begin{aligned} R(x) &= \pi(x)\sum_{a\neq 0}\phi_a(x)\eta(a)^{-1}D(a) \\ &\quad + \pi(x)\phi_0(x)\sum_s \eta(s(l))\Big[\lambda_{s0} + \sum_{a\neq 0}\Lambda_{(sa)}\Big]. \end{aligned} \tag{16}$$

Next, note that the left side of the balance equation (10), in light of (4), is

$$
\begin{aligned}
L(x) &= \pi(x)\sum_{s,a}\lambda_{sa}\sum_{y\in E} r_{sa}(x,y)\\
&= \pi(x)\sum_{s,a}\lambda_{sa}\phi_0(x) = \pi(x)\phi_0(x)\Lambda_0.
\end{aligned}
\tag{17}
$$

Now, using the fact that $s' \neq 0$ can be expressed as $s' = (sas'')$ for some $s, s'' \in S$ and $a \in A_0$, we have the identity

$$
\sum_s \eta(s(l))\Lambda_s = \Lambda_0 + \sum_{s\neq 0}\eta(s(l))\Lambda_s = \Lambda_0 + \sum_{s,\, a\neq 0}\eta(s(l))\eta(a)\Lambda_{(sa)}. \tag{18}
$$

Also, by its definition, $\Lambda_s = \lambda_{s0} + \sum_{a\neq 0}(\lambda_{sa} + \Lambda_{(sa)})$. Substituting this in the left side of (18) and using terms from its right side yields

$$
\Lambda_0 = \sum_{a\neq 0} D(a) + \sum_s \eta(s(l))\Bigg[\lambda_{s0} + \sum_{a\neq 0}\Lambda_{(sa)}\Bigg]. \tag{19}
$$

Finally, substituting this in (17) and using (16) yields the identity (11). □

4. Traffic equations, partial balance and throughputs. This section begins with insights into the existence of solutions to the traffic equations. Next, we show that the traffic equations are equalities of certain average flows in the network (a partial balance property). The section ends with an expression for throughputs at the nodes.

Note that the hypothesis of Theorem 1 (the first sentence) is actually two hypotheses.

1. There are positive γ_a, $a \in A_0$, that satisfy

$$
\gamma_a \sum_s \prod_{k=1}^{l} \gamma_{s^k}\Lambda_{(sa)} = \sum_s \prod_{k=1}^{l}\gamma_{s^k}\lambda_{sa}, \qquad a \in A_0, \tag{20}
$$

   where $\gamma_0 = 1$ and these sums are finite.
2. $\gamma_a = \prod_{k=1}^{m} w_j^{a_j}$, $a \in A_0$, for some positive numbers $w_1, \dots, w_m$.

Let us first consider hypothesis 1. With a slight abuse of notation, interpret A_0 as an ordered set and view $\gamma \equiv (\gamma_a,\ a \in A_0)$ as a vector. Write (20) as $\gamma_a g_a(\gamma) = h_a(\gamma)$, where $g_a(\gamma)$ and $h_a(\gamma)$ denote the summations on the left and right sides of (20) as functions of γ. In other words, (20) is the same as $\gamma = f(\gamma)$, where $f(\gamma) = \{f_a(\gamma)\colon a \in A_0\}$ is the vector-valued function defined by $f_a(\gamma) = h_a(\gamma)/g_a(\gamma)$ for γ in the region where the numerator is finite and the denominator is not zero. Here a vector inequality $\gamma \leq \bar{\gamma}$ means $\gamma_a \leq \bar{\gamma}_a$ for each $a \in A_0$ and 0 and 1 are the vectors of all zeros and all ones.

From the preceding observations, it follows that the set of solutions to (20) is equal to the set of fixed points of f. Here is a general criterion for the existence of a solution to (20) [i.e., a sufficient condition for hypothesis 1].

THEOREM 2. *If there are vectors* $0 \le \underline{\gamma} < \overline{\gamma}$ *such that* $0 < g_a(\underline{\gamma}) < \infty$,

$$\underline{\gamma}_a g_a(\overline{\gamma}) < h_a(\underline{\gamma}) \quad \text{and} \quad h_a(\overline{\gamma}) < \overline{\gamma}_a g_a(\underline{\gamma}), \qquad a \in A_0, \tag{21}$$

then there exists a vector γ *that satisfies* (20) *and* $\underline{\gamma} < \gamma < \overline{\gamma}$.

PROOF. Let $C = \{\gamma: \underline{\gamma} \le \gamma \le \overline{\gamma}\}$. Since $g_a(\gamma)$ and $h_a(\gamma)$ are increasing in γ, it follows that all the terms in (21) are finite and

$$\underline{\gamma}_a < h_a(\underline{\gamma})/g_a(\overline{\gamma}) \le f_a(\gamma) \le h_a(\overline{\gamma})/g_a(\underline{\gamma}) < \overline{\gamma}_a, \qquad \gamma \in C,\ a \in A_0.$$

Thus, f maps C into C. Also, f is clearly continuous. Then f has a fixed point $\gamma \in C$ by Brouwer's fixed-point theorem. Furthermore, $\underline{\gamma} < \gamma < \overline{\gamma}$, because of the strict inequalities in the preceding display. □

Theorem 2 is a framework for obtaining specific conditions for a solution to (20) in terms of the structure of the λ_{sa}'s. Examples are in the following sections. The next result is a simpler version of Theorem 2 when $g_a(0) > 0$ and the existence of $\underline{\gamma}$ is guaranteed.

COROLLARY 3. *There exists a positive solution to the traffic equations* (20) *if the following conditions hold*:

(a) *The set* $A^* = \{a \in A_0: \lambda_{0a} > 0\}$ *is not empty and, for each* $a \in A_0 \backslash A^*$, *there is a* $s \in S$ *such that* $\lambda_{sa} > 0$ *and* $s^k \in A^*$, *for* $1 \le k \le l$.
(b) *There is a positive vector* $\overline{\gamma}$ *such that* $g_a(\overline{\gamma}) < \infty$ *and* $h_a(\overline{\gamma}) < \overline{\gamma}_a g_a(0)$.

[Open networks typically satisfy (a) and $\overline{\gamma} = 1$ is often adequate for (b).]

PROOF. Let $\underline{\gamma} = \{\underline{\gamma}_a: a \in A_0\}$ be a vector in $(0, \overline{\gamma})$ such that $\underline{\gamma}_a < \lambda_{0a}/g_a(\overline{\gamma})$, for $a \in A^*$ and

$$\underline{\gamma}_a < \sum_s \prod_{k=1}^{l} \underline{\gamma}_{s^k} \lambda_{sa} 1(s^k \in A^*,\ 1 \le k \le l)/g_a(\overline{\gamma}), \qquad a \in A_0 \backslash A^*.$$

Assumptions (a) and (b) ensure that $0 < g_a(\underline{\gamma}) < \infty$. Since $g_a(\gamma)$ and $h_a(\gamma)$ are increasing in γ, it follows that (21) holds. Thus, the assertion follows by Theorem 2. □

Now, consider the hypothesis (2) that the solution γ to (20) has the geometric form $\gamma_a = \prod_{k=1}^m w_j^{a_j}$, $a \in A_0$, for some positive $w_1, \dots, w_m$. This hypothesis is satisfied for the large class of networks discussed in the next sections with strings composed of unit vectors. For the general case, we have the following observation. The problem is to determine when there are positive $w_1, \dots, w_m$ that satisfy the linear equations

$$\log \gamma_a = \sum_{j=1}^{m} a_j \log w_j, \qquad a \in A_0, \tag{22}$$

for known γ_a's. From a standard property of linear algebra, we have the following result.

REMARK 4 (Geometric solutions). Let M denote the matrix whose rows are vectors in A_0 and let M' denote the matrix M augmented by the column $(\log \gamma_a)_{a\in A_0}$. Then there is a solution $w_1, \dots, w_m$ to (22) if and only if M and M' have the same rank, which is at most m. If they have the same rank and it is less than m, then there are an infinite number of solutions. Uniqueness is not important for our purposes. However, the solution is unique if M and M' have the same rank m, which is true when A_0 consists of m linearly independent vectors. $\square$

We now justify that equations (8) are traffic equations. Throughout the rest of this section, we assume that the network process X is ergodic with stationary distribution $\pi(x) = c\Phi(x)\eta(x)$, where $\eta(x) = \prod_{j=1}^m w_j^{x_j}$ and the w_j's satisfy (8). Equations (8) for $a \neq 0$ along with the identity (19) imply

$$\sum_{a\neq 0} \lambda_{0a} = \sum_{s\neq 0} \eta(s(l))\left[\lambda_{s0} + \sum_{a\neq 0} \Lambda_{(sa)}\right], \quad x \in E. \tag{23}$$

This is the *traffic equation for $a = 0$*.

Recall that by the ergodic theorem for Markov processes, the quantity $\sum_{(x,\, y)\in C} \pi(x)q(x, y)$ is the average number of $x \to y$ transitions of X per unit time, where $(x, y) \in C$. This average number of *C-transitions*, which is a limiting average, is also the expected number of C-transitions in a unit time interval when the process is stationary.

We shall consider two types of transitions related to the traffic equations. For $a \in A$ and $x \in E$, let $\tilde{\lambda}_a(x)$ denote the average number of transitions of X per unit time in which the vector a is added to the state x such that the transition leads to the new state $x + a$. We call $\tilde{\lambda}_a(x)$ the *rate of exits from x via an a-addition*. Similarly, let $\hat{\lambda}_a(x)$ denote the *rate of entrances into x via an a-subtraction*: the average number of transitions of X per unit time in which the process enters state x (during a transition) from a subtraction of the vector a. Here are expressions for these rates.

PROPOSITION 5 (Partial balance). *For each $x \in E$,*

$$\tilde{\lambda}_a(x) = \begin{cases} \pi(x)\phi_0(x) \sum_{s} \eta(s(l))\lambda_{sa}, & \text{if } a \neq 0, \\ \pi(x)\phi_0(x) \sum_{s\neq 0} \eta(s(l))\left[\lambda_{s0} + \sum_{a\neq 0} \Lambda_{(sa)}\right], & \text{if } a = 0, \end{cases} \tag{24}$$

$$\hat{\lambda}_a(x) = \begin{cases} \pi(x)\phi_0(x)\eta(a) \sum_{s} \eta(s(l))\Lambda_{(sa)}, & \text{if } a \neq 0, \\ \pi(x)\phi_0(x) \sum_{a\neq 0} \lambda_{0a}, & \text{if } a = 0. \end{cases} \tag{25}$$

Hence, the traffic equations (8), (23) *are equivalent to*

$$\hat{\lambda}_a(x) = \tilde{\lambda}_a(x), \qquad a \in A,\ x \in E. \tag{26}$$

Expression (26) says the average number of entrances into x via an a-subtraction is equal to (or balanced with) the average number of exits out of x via an a-addition. The equivalence between this balance (26) of traffic flows and the equations (8), (23) is the reason why we call the latter traffic equations. The equality (26) is sometimes called a *partial balance property* for the process since it is the balance equation (10) for only a part of the summation. Note that (26) also implies that, for each fixed state x, the average number of entrances into x via any a-subtraction equals the average number of exits from x via any a-addition [namely $\sum_{a\in A} \hat{\lambda}_a(x) = \sum_{a\in A} \tilde{\lambda}_a(x)$]. A similar sum on x, for a fixed a, says the average number of a-subtractions is equal to the average number of a-additions, regardless of the state x.

PROOF. First consider the case $a \neq 0$. A transition of X in which the vector a is added to a state x such that the transition leads to the new state $x + a$ is necessarily a complete sa-transition that starts from $x + s(l)$ and lands in $x + a$, for any $s \in \mathscr{S}$. Then by the comment above on the ergodic theorem for Markov processes,

$$\tilde{\lambda}_a(x) = \sum_s \pi(x + s(l))\lambda_{sa}\phi_{s(l)}(x + s(l)).$$

This reduces, in light of (13), to the first line in (24). Next, note that a transition of X in which it enters state x (during a transition) due to a subtraction of the vector a can only happen when the process is in state $x + s(l) + a$ and a $(sas')a'$-string transition occurs causing the process to enter state x at the stage in which a is subtracted. Arguing as above,

$$\begin{aligned}\hat{\lambda}_a(x) &= \sum_s \pi(x + s(l) + a) \sum_{s',\,a'} \lambda_{(sas')a'}\phi_{s(l)+a}(x + s(l) + a)\\ &= \sum_s \pi(x + s(l) + a)\Lambda_{(sa)}\phi_{s(l)+a}(x + s(l) + a),\end{aligned}$$

and this reduces to the first line in (25).

Now, consider the case $a = 0$. Since 0-additions involve complete $s0$-transitions and other partial transitions as well, we have

$$\tilde{\lambda}_0(x) = \sum_{s\neq 0} \pi(x + s(l))\phi_{s(l)}(x + s(l))\left[\lambda_{s0} + \sum_{a\neq 0}\sum_{s',\,a'} \lambda_{(sas')a'}\right]$$

and this reduces to the second line in (24). Also, the second line in (25) clearly follows since 0-subtractions only involve complete $0a$-transitions. Finally, a glance at (24), (25) and the traffic equations (8) and (23) verifies that the traffic equations are equivalent to (26). □

A measure of a network's performance is its *throughput vector* $(\tilde{\lambda}_1, \dots, \tilde{\lambda}_m)$, where $\tilde{\lambda}_j$ denotes the average number of units per unit time that enter node

j. Since the process is ergodic, $\tilde{\lambda}_j$ is also the average number of departures per unit time from j.

PROPOSITION 6 (Throughputs at nodes). *The throughput at node j is*

$$\tilde{\lambda}_j = \sum_{s,\,a} \alpha_a \lambda_{sa} \left(a_j^+ + \sum_{i=1}^{l} (s_j^i)^- \right) + \sum_{s,\,a} \lambda_{sa} \sum_{k=0}^{l-1} (\alpha_0 - \alpha_{s^{k+1}}) \sum_{i=1}^{k} (s_j^i)^-, \tag{27}$$

where $\alpha_a = \sum_x \pi(x)\phi_a(x)$, $y^+ = \max\{0, y\}$ *and* $y^- = -\min\{0, y\}$.

PROOF. By the ergodic theorem for Markov processes, $\tilde{\lambda}_j = \sum_{x,\,y} \pi(x) \times q(x, y) f(x, y)$, where $f(x, y)$ describes the number of arrivals to j in an $x \to y$ transition. That is,

$$\begin{aligned}\tilde{\lambda}_j = &\sum_x \pi(x) \sum_{s,\,a} \lambda_{sa} \phi_{s(l)}(x) 1(x - s(l) + a \in E) \left[a_j^+ + \sum_{i=1}^{l} (s_j^i)^- \right] \\ &+ \sum_x \pi(x) \sum_{s,\,a} \lambda_{sa} \sum_{k=0}^{l-1} [\phi_{s(k)}(x) - \phi_{s(k+1)}(x)] 1(x - s(k) \in E) \sum_{i=1}^{k} (s_j^i)^-.\end{aligned}$$

Now, by two uses of Φ-reversibility and the structure of π, we have

$$\begin{aligned}\sum_x \pi(x)\phi_{s(k)}(x) 1(x - s(k) + a \in E) &= \sum_x \pi(T_{as(k)}x)\phi_{s(k)}(T_{as(k)}x)\eta(a)\eta(s(k))^{-1} \\ &= \sum_x \pi(x)\phi_a(x) = \alpha_a.\end{aligned}$$

Similarly,

$$\sum_x \pi(x)\phi_{s(k+1)}(x) 1(x - s(k+1) - s^{k+1} \in E) = \alpha_{s^{k+1}}.$$

Then applying these equalities to the two sums on x in the preceding display yields (27). □

5. Processes with unit-vector string transitions. This section describes the results above when the increments in the network are unit vectors instead of general vectors.

Suppose X is an open network process and its increment vectors consist of the m-dimensional unit vectors $e_1, \dots, e_m$ and $e_0 = 0$. We say that the process has *unit-vector string transitions*. In this case, the unit vectors are associated with the node numbers and it is convenient to let $A = \{0, 1, \dots, m\}$ denote the node numbers instead of the vectors. Then $s = (s^1, \dots, s^l)$ is a string of node numbers, $s(k) = \sum_{i=1}^k e_{s^i}$ and the rates are λ_{sj}, $j \in A$. The transition rates and results above are the same, aside from the change in notation from vectors to node numbers. For instance, the traffic equations (8) are

$$w_j \sum_s \prod_{k=1}^{l} w_{s^k} \sum_{s',\,j'} \lambda_{(sjs')j'} = \sum_s \prod_{k=1}^{l} w_{s^k} \lambda_{sj}, \qquad 1 \le j \le m. \tag{28}$$

The following is a combination of Theorems 1 and 2 for unit-vector string transitions. Consider (28) written as $w_j g_j(w) = h_j(w)$, where $g_j(w)$ and $h_j(w)$ denote the summations on the left and right sides of (28) as functions of $w = (w_1, \dots, w_m)$.

THEOREM 7. *Suppose there there are vectors* $0 \le \underline{w} < \overline{w}$ *such that* $0 < g_j(\overline{w}) < \infty$,

$$\underline{w}_j g_j(\overline{w}) < h_j(\underline{w}) \quad \text{and} \quad h_j(\overline{w}) < \overline{w}_j g_j(\underline{w}), \qquad 1 \le j \le m.$$

Then there is a vector w *that satisfies* (28) *and* $\underline{w} < w < \overline{w}$. *Moreover*, $\pi(x) = \Phi(x) \prod_{j=1}^m w_j^{x_j}$, $x \in E$, *is an invariant measure for the network process* X.

Note that this result is simpler than Theorems 1 and 2 because w_j plays the role of γ_a in Theorem 2 and hence there is no issue of verifying that γ_a is a product.

For closed networks, unit-vector transitions make sense only for the case of one-stage transitions ($L = 1$). One can also define analogous unit-vector transitions when the set A of increment vectors consists of only the negative unit vectors or it consists of a combination of negative and positive unit vectors.

We now derive expressions for throughputs and service rates. For the rest of this section, assume that the network process X is ergodic and denote its stationary distribution by $\pi(x) = c\Phi(x) \prod_{j=1}^m w_j^{x_j}$. Let $\tilde{\mu}_j$ denote the average number of departures per unit time from node j when the node is not empty. The $\tilde{\lambda}_j$ and $\tilde{\mu}_j$ are often called the *effective arrival and service rates* for node j and the ratio $\tilde{\lambda}_j / \tilde{\mu}_j$ is the *traffic intensity*.

PROPOSITION 8. *For the network process* X *with unit-vector string transitions*,

$$\tilde{\lambda}_j = \sum_x \pi(x) \phi_j(x) \sum_s \lambda_{sj} \quad \text{and} \quad \tilde{\mu}_j = \tilde{\lambda}_j \Big/ \sum_x 1(e_j \le x) \pi(x).$$

PROOF. The first expression is an obvious special case of (27). By the strong law of large numbers for Markov processes, the effective departure rate is

$$\tilde{\mu}_j = \sum_x \hat{\lambda}_j(x) \Big/ \sum_x 1(e_j \le x) \pi(x),$$

which is the average number of departures from j per unit time divided by the portion of time j is nonempty. Then $\sum_x \hat{\lambda}_j(x) = \tilde{\lambda}_j$ by Proposition 5. □

Another important performance measure of the network is the average sojourn or waiting time of a unit in a node j of the network. This average is defined by $W_j \equiv \lim_{n \to \infty} n^{-1} \sum_{\nu=1}^n W_j(\nu)$, where $W_j(\nu)$ is the waiting time in j of the νth unit to enter j. We will use the average number of units in j, which is $L_j \equiv \sum_x x_j \pi(x)$. The Little law for Markovian systems (Theorem 25) in [30] yields the following result.

PROPOSITION 9 (Little law for waiting times). *If the network has unit-vector string transitions and the state space contains a vector x with $x_j = 0$, then W_j exists and $L_j = \tilde{\lambda}_j W_j$. Here $\tilde{\lambda}_j$ is necessarily finite, but W_j and L_j may both be finite or infinite. This assertion for averages also holds for expected values when the system is stationary* (*or in equilibrium*). *In this case, L_j is the expected number of units in j, the $\tilde{\lambda}_j$ is the expected number of units that enter j in a unit time interval and W_j is the expected sojourn time for an arbitrary unit in j under the Palm probability that a unit enters j at time* 0.

Similar Little laws apply to batch arrivals into j, but more information is needed on how the "order of units" in a batch affect their individual service times. In these cases, one can state a law for all units labeled as the kth unit within a batch arriving into j—the W_j and L_j would be the average waiting times and queue lengths for these kth arrivals and $\tilde{\lambda}_j$ would be the arrival rate of batches into j of size k or more. The expected waiting time in a sector (subset of nodes) in a Jackson network is described in [30]. To obtain similar results for vector-transitions, one would need more information on where each unit in a batch actually moves; the net number of movements is not adequate to describe waiting times as it is under single-unit movements.

The computation of throughputs, average waiting times and other performance parameters, even for a Jackson network, is difficult for a moderate-sized network. However, since there is a closed-form expression for the stationary distribution of the network, one can compute these parameters by Monte Carlo simulation as discussed in [29] (the Metropolis Markov chain or other reversible chains are natural choices for the simulation vehicle).

6. Networks with one-stage transitions. This section shows how the results above apply to string-nets in which all of the strings are exactly of length 1: each transition involves only one pair of addition–subtraction vectors. A Jackson network is a special case in which the increment vectors are unit vectors.

Consider the process X with strings exactly of length 1, which implies $\mathscr{S} = S = A$. Then all transitions are of the form $x \to T_{da}x$, for $d, a \in A$, and the transition rates (1) become

$$q(x, y) = \sum_{d, a \in A} \lambda_{da}\phi_d(x)1(y = T_{da}x), \qquad x \neq y \text{ in } E. \tag{29}$$

In other words, whenever the process is in state x, the time to its next potential move to $T_{da}x$ via a da-transition is exponentially distributed with rate $\lambda_{da}\phi_d(x)$. The Φ-reversibility assumption implies that $\phi_a(x) = 0$, if $T_{ad}x \notin E$ for some $d \in A$ with $\lambda_{da} > 0$. Note that the ϕ_0-dominance assumption is not relevant since $L = 1$. We call X a network process with *one-stage batch transitions*—the vectors in A are the allowable batch increments in the process.

The following is Theorem 1 for the network process X with one-stage batch transitions and transition rates (29), where ϕ_d are Φ-reversible.

THEOREM 10. *If* $\gamma = (\gamma_a\colon a \in A)$ *is a positive vector, with* $\gamma_0 = 1$, *that satisfies*

$$\gamma_a \sum_{d\in A_0} \lambda_{ad} = \sum_{d\in A_0} \gamma_d \lambda_{da}, \qquad a \in A_0, \tag{30}$$

and $\gamma_a = \prod_{j=1}^m w_j^{a_j}$, $a \in A_0$, *for some positive* $w_1, \dots, w_m$, *then* $\pi(x) = \Phi(x)\prod_{j=1}^m w_j^{x_j}$, $x \in E$, *is an invariant measure for the process* X. *Furthermore, a necessary and sufficient condition for an invariant measure of this form is*

$$\sum_{a\in A_0} D(a)\left[\phi_0(x) - \phi_a(x)\prod_{j=1}^m w_j^{-a_j}\right] = 0, \qquad x \in E,$$

where $D(a)$ *denotes the right side of* (30) *minus its left side.*

In this case, the traffic equations (8) reduce to (30) because in $\Lambda_{(sa)}$, the s must be 0 since $L = 1$, and $\Lambda_{(0a)} = \sum_{d\in A} \lambda_{ad}$.

Note that (30) is a balance equation for a Markov process on the finite set A_0 with transition rates λ_{ad} and hence there exists a positive solution γ to the equation. The solution is a geometric product form under the criterion in Remark 4.

According to Proposition 5, the measure π in Theorem 10 satisfies the partial balance property

$$\pi(x)\sum_{d\in A_0} q(x, T_{ad}x) = \sum_{d\in A_0} \pi(T_{ad}x)q(T_{ad}x, x)1(T_{ad}x \in E), \qquad a \in A_0,\ x \in E. \tag{31}$$

Here are two examples of the preceding theorem.

EXAMPLE 11 (Jackson network processes). Consider the special case of the process X with one-stage transitions and unit-vector increments (the set of allowable increments are the m unit vectors plus 0). Its transition rates are

$$q(x, T_{jk}x) = \lambda_{jk}\phi_j(x), \qquad j, k \in A \equiv \{0, 1, \dots, m\},$$

where $T_{jk}x = x - e_j + e_k$. Such a transition is usually viewed as a single unit moving from node j to node k, although this can also represent many units moving at once as long as the "net movement" in the transition results in one less unit at j and one more unit at k. The $\phi_j(x)$ is called the departure rate (or service rate) at node j when the network is in state x and λ_{jk} is the rate of a unit moving from j to k. We call X a *Jackson network process* with system-dependent service rates. In the classical case where $\phi_j(x_j)$ is a function of only x_j, these functions are automatically Φ-reversible with $\Phi(x) = \prod_{j=1}^m \prod_{n=1}^{x_j} \phi_j(n)^{-1}$. The assumption that the process is irreducible is equivalent to λ_{jk} being an irreducible matrix. Then the traffic equations, which are

$$w_j \sum_{k\in A} \lambda_{jk} = \sum_{k\in A} w_k \lambda_{kj}, \qquad 1 \le j \le m,$$

have a positive solution $w_1, \dots, w_m$. Therefore, Theorem 10 contains the well-known result that $\pi(x) = \Phi(x) \prod_{j=1}^m w_j^{x_j}$ is an invariant measure. The process also satisfies the partial balance condition (31).

EXAMPLE 12 (Independent concurrent movements of units). Henderson and Taylor [23] introduced the following network process (they used slightly different transition rates that are normalized to be probabilities). Let X denote the network process with one-stage, batch transitions with rates (29), where A denotes a set of m-dimensional vectors. For simplicity, assume the network is closed (the open case is similar). In a da-transition in state x, think of $\phi_d(x)$ as the rate at which the batch d is released from the network and λ_{da} as the rate in which d is changed into the addition batch a. To describe the units in these vectors by their node locations, we define

$$\mathscr{I}(d) = \left\{ \mathbf{i} = (i_1, \dots, i_{|d|})\colon \sum_{n=1}^{|d|} 1(i_n = j) = d_j,\ 1 \le j \le m \right\},$$

which is the set of possible node indices that "represent" d.

Assume the units in the batch d move concurrently such that r_{jk} is the rate (probability or propensity) for a single unit in the batch to move from j to k in the node set $\{1, \dots, m\}$ and that $r_{j_1,k_1} \cdots r_{j_{|d|},k_{|d|}}$ is the rate that the released batch $\mathbf{j} \in \mathscr{I}(d)$ results in the batch addition $\mathbf{k} \in \mathscr{I}(a)$, where $|d| = |a|$. This rate is a compounding of the single-unit rates. Then the rate of a da-transition in state x is $\lambda_{da}\phi_d(x)$, where

$$\lambda_{da} = \sum_{\mathbf{j}\in\mathscr{I}(d)} \sum_{\mathbf{k}\in\mathscr{I}(a)} r_{j_1,k_1} \cdots r_{j_{|d|},k_{|d|}}, \qquad d, a \in A \text{ with } |d| = |a|.$$

Note that the probability of d and a being generated is $\lambda_{da}/\sum_a \lambda_{da}$, and if the r_{jk}'s are probabilities with $\sum_{k=1}^m r_{jk} = 1$ for each j, then $\sum_a \lambda_{da} = |d|!/d_1! \cdots d_m!$.

Assume the rates r_{jk} are irreducible in the sense that there are positive $w_1, \dots, w_m$, that satisfy

$$w_j \sum_{k=1}^m r_{jk} = \sum_{k=1}^m w_k r_{kj}, \qquad 1 \le j \le m.$$

Define $\eta(x) = \prod_{j=1}^m w_j^{x_j}$. Because the w_j's satisfy the preceding equations, we have

$$\begin{aligned}\eta(a) \sum_{d\in A} \lambda_{ad} &= \sum_{\mathbf{j}\in\mathscr{I}(a)} \sum_{k_1,\dots,k_{|a|}} w_{j_1} r_{j_1,k_1} \cdots w_{j_{|a|}} r_{j_{|a|},k_{|a|}} \\ &= \sum_{\mathbf{j}\in\mathscr{I}(a)} \sum_{k_1,\dots,k_{|a|}} w_{k_1} r_{k_1,j_1} \cdots w_{k_{|a|}} r_{k_{|a|},j_{|a|}} = \sum_{d\in A} \eta(d)\lambda_{da}, \qquad a \in A_0.\end{aligned}$$

Therefore, by Theorem 10 it follows that $\pi(x) = \Phi(x)\eta(x)$, $x \in E$, is an invariant measure for the process.

7. Networks with compound-rate string transitions. A large class of network processes with string transitions are those in which the rate λ_{sa} of a sa-string is a product or compounding of several rates representing micro features of the network. This section illustrates this class with an example of a network in which a string is generated by a Markov chain mechanism. The ideas here readily extend to a variety of networks with compound-rate string transitions.

Consider the m-node network that operates as follows. The network is open and its state space consists of all m-dimensional vectors with nonnegative integer-valued entries. Units enter the network at the nodes according to independent Poisson processes with respective rates $\lambda_1, \dots, \lambda_m$; a zero rate for a node means it has no external arrivals. The services at each node j are independent and exponentially distributed with rate μ_j. The results below also apply, with minor modifications, to general Φ-reversible service rates and closed or open networks with other types of state spaces.

A transition of the network is triggered by the movement of a single unit. An external arrival to a node just adds one unit to the node and no other units move. On the other hand, a service completion at a node may trigger a transition in which single units are successively deleted from a string of nodes $s_1, \dots, s_\nu$ and, at the end, one unit might be added to some node j in $A \equiv \{0, 1, \dots, m\}$. All of this occurs instantaneously and the number of deletions $\nu \le L$ is a stopping index that may be random.

The procedure for such a transition triggered by a service completion is as follows. Whenever a normal service completes at some node $s_1 \in A_0 \equiv A \backslash \{0\}$ (with rate μ_{s_1}), then with probability $Q_{s_1 j}$ one unit moves to some node $j \in A$ and the procedure stops, or with probability $P_{s_1 s_2}$, one unit exits the network from node s_1 and a signal goes to node $s_2 \in A_0$ to delete a unit there provided that node is not empty [$\sum_j (P_{ij} + Q_{ij}) = 1$ for each i]. If node s_2 is empty or if $L = 1$, the procedure stops. Otherwise, the preceding events are repeated until stopping. That is, for each $k \ge 1$, the departure from node s_k, with probability $Q_{s_k j}$, adds one unit to node j and stops the procedure; with probability $P_{s_k s_{k+1}}$, it triggers another departure from node s_{k+1} provided this node is nonempty and, if node s_{k+1} is empty or $k = L$, the procedure stops. Think of P_{ij} as probabilities of "propagating new departures" and Q_{ij} as probabilities of "quitting" (or stopping) the string deletions.

In summary of the preceding description, typical transitions of the network are as follows:

1. An arrival into node j from outside the network: $x \to x + e_j$.
2. String deletions stopped because node s_{k+1} is empty or $k = L$: $x \to x - e_{s_1} - \cdots - e_{s_k}$.
3. String deletions stopped by the quitting probability $Q_{s_k j}$: $x \to x - e_{s_1} - \cdots - e_{s_k} + e_j$.

As in the previous sections, we let X denote the stochastic process representing the numbers of units at the nodes. The data for this process are the arrival rates $\lambda = (\lambda_1, \dots, \lambda_m)$, service rates $\mu = (\mu_1, \dots, \mu_m)$, maximum string length L, and propagating and quitting probabilities P_{ij}, Q_{ij}. Define $P = (P_{ij})$ and $Q = (Q_{ij})$ for $i, j \in A_0$. We also assume that the inverse of the matrix $I - Q$ exists, where I denotes the identity matrix. We will frequently use the vector

$$\tilde{\lambda} = \lambda(I - Q)^{-1},$$

whose entries are effective arrival rates, as we will soon see.

We first justify that this network process is a string-net.

PROPOSITION 13. *Under the preceding assumptions, X is a Markov network process with unit-vector string transitions and its associated traffic equation* (28) *in matrix form is*

$$\mu \sum_{n=0}^{L-1} (WP)^n W = \lambda + \mu \sum_{n=0}^{L-1} (WP)^n WQ, \tag{32}$$

where W is a diagonal matrix with diagonal entries $w_1, \dots, w_m$.

PROOF. Because of the Poisson arrivals and exponential service times, the network process is clearly Markovian. The rates of its sj-transitions are $\lambda_{0j} = \lambda_j$ and, for $s \neq 0$,

$$\lambda_{sj} = \Lambda_s Q_{s_l j}, \quad j \neq 0 \quad \text{and} \quad \lambda_{s0} = \Lambda_s \left[Q_{s_l 0} + 1(l = L) \sum_i P_{s_l i} \right],$$

where $\Lambda_s = \mu_{s_1} \prod_{k=1}^{l-1} P_{s_k s_{k+1}}$. Now, consistent with its definition in the last section, $\Lambda_0 = \sum_j (\lambda_j + \mu_j)$ and, for $s \neq 0$, $\Lambda_s = \sum_{s', j} \lambda_{(ss')j}$, which is the rate of all string transitions whose first part is s.

Next note that the departure rate functions have the special form $\phi_j(x) = 1(e_j \leq x)$ because they must satisfy

$$\phi_{s(k)}(x) - \phi_{s(k+1)}(x) = 1(0 \leq x - s(k),\ 0 \not\leq x - s(k+1)).$$

Clearly, these ϕ_j's are Φ-reversible with $\Phi(\cdot) = 1$ and the ϕ_0-dominance assumption is satisfied. Under these specifications, X is a Markov network process with unit-vector string transitions.

In this setting, the traffic equations (28) reduce to (32) since

$$w_j \sum_s \eta(s(l)) \Lambda_{(sj)} = \sum_{s \neq 0} \mu_{s_1} \prod_{k=1}^{l-1} w_{s_k} P_{s_k s_{k+1}} w_j 1(s_l = j) = \left(\mu \sum_{n=0}^{L-1} (WP)^n W \right)_j, \tag{33}$$

$$\sum_s \eta(s(l)) \lambda_{sj} = \left(\lambda + \mu \sum_{n=0}^{L-1} (WP)^n WQ \right)_j. \tag{34}$$

□

The following characterization of solutions to the traffic equation is analogous to Theorem 2. We will use $\overline{w}_j = \tilde{\lambda}_j/\mu_j$, $1 \le j \le m$. Recall that $\tilde{\lambda} = \lambda(I-Q)^{-1}$.

THEOREM 14. (a) *Suppose* $L = \infty$ *and* $\sum_{n=0}^{\infty}(PW)^n < \infty$, *where*

$$w_j = \tilde{\lambda}_j/(\mu + \tilde{\lambda}P)_j, \qquad 1 \le j \le m.$$

Then $w_1, \ldots, w_m$ *is the unique solution to the traffic equation* (32).

(b) *Suppose* $L < \infty$ *and* $\sum_{n=0}^{\infty}(P\overline{W})^n < \infty$. *Then there exists a solution* w *to the traffic equation* (32) *in the open rectangle* $(0, \overline{w})$. *Futhermore, let* $\mathbf{w}_n$ *denote a sequence of vectors defined by* $\mathbf{w}_0 = 0$ *and* $\mathbf{w}_{n+1} = h(\mathbf{w}_n)$, *where* $h(w) = (h_1(w), \ldots, h_m(w))$ *and*

$$h_j(w) = (\tilde{\lambda} + \mu(WP)^L W)_j/(\mu + \tilde{\lambda}P)_j, \qquad w \in C \equiv [0, \overline{w}],\ 1 \le j \le m.$$

Then $\mathbf{w}_n$ *is a nondecreasing sequence whose limit is the minimal solution to the traffic equation* (32) (*any other solution is greater than or equal to this limit*).

PROOF. First note that by subtracting the right side of (32) from its left and dividing by $I - Q$, this traffic equation can be written as

$$\mu \sum_{n=0}^{L-1}(WP)^n W = \tilde{\lambda}. \tag{35}$$

Now, assume the assumptions in part (a) hold. Multiplying both sides of (35) on the right by the matrix $(I - PW)$ yields $\mu W = \tilde{\lambda} - \tilde{\lambda}PW$. That is, $(\mu + \tilde{\lambda}P)W = \tilde{\lambda}$, for $w \in C$. This proves the assertion in part (a).

Next, assume $L < \infty$. Note that equation (35) is the same as $f(w) = w$, where $f(w) = (f_1(w), \ldots, f_m(w))$ is defined by

$$f_j(w) = \tilde{\lambda}_j \Big/ \left(\mu \sum_{n=0}^{L-1}(WP)^n\right)_j, \qquad w \in C,\ 1 \le j \le m.$$

Clearly f is positive, continuous, nonincreasing and its range is contained in C since $0 < f(\overline{w}) < \overline{w}$. Then, by Brouwer's fixed point theorem, f has a fixed point in C and hence this point is a solution to the traffic equation (35). Furthermore, this solution is in the open rectangle $(0, \overline{w})$ since this set contains the range of f.

For the rest of the proof, we need another representation of the traffic equation (35). Multiplying both sides of it on the right by the matrix $(I - PW)$ yields

$$\mu[I - (WP)^L]W = \tilde{\lambda} - \tilde{\lambda}PW, \qquad w \in C. \tag{36}$$

Writing this as $(\mu + \tilde{\lambda}P)W = \tilde{\lambda} + \mu(WP)^L W$ and recalling the definition of h in part (b), it is clear that the traffic equation (35) is equivalent to $w = h(w)$, for $w \ge 0$. Hence the solutions to (35) are the same as the fixed points of h. Since h is non-decreasing, it follows by induction that $\mathbf{w}_n$ is a nondecreasing sequence. Then the limit $w^* = \lim_{n\to\infty} \mathbf{w}_n$ exists. Now, as $n \to \infty$ in $\mathbf{w}_{n+1} = h(\mathbf{w}_n)$, the continuity of h ensures that $w^* = h(w^*)$. Thus w^* is a solution to the traffic equation (35). It remains to show that if w' is any solution to the equation, then $w^* \le w'$. To prove this, it suffices to show that $\mathbf{w}_n \le w'$ for each n. But this follows by induction, since $\mathbf{w}_0 = 0 \le w'$ and, assuming $\mathbf{w}_n \le w'$ for some n, then $\mathbf{w}_{n+1} = h(\mathbf{w}_n) \le h(w') = w'$ because h is nondecreasing. □

Obtaining a solution of the traffic equation by successively computing $\mathbf{w}_{n+1} = h(\mathbf{w}_n)$ is very efficient; $\mathbf{w}_n$ converges to its limit very fast.

The next result describes invariant measures for the network process and it says that the process is ergodic if the arrival rates to the nodes are less than the service capacities of the nodes.

THEOREM 15. *Suppose the vector w is a solution to the traffic equation* (32) *and $0 < w < \overline{w}$. Then $\prod_{j=1}^m w_j^{x_j}$, $x \ge 0$, is an invariant measure for the process X. This process is ergodic if and only if $0 < w < 1$. In particular, the process is ergodic if $\tilde{\lambda} < \mu$, and, in case $L = \infty$, the process is ergodic if and only if $\tilde{\lambda} < \mu + \tilde{\lambda} P$. When the process is ergodic, its stationary distribution is*

$$\pi(x) = \prod_{j=1}^{m} (1 - w_j) w_j^{x_j}, \qquad x \ge 0.$$

In addition, $\tilde{\lambda}$ is the throughput vector and the vector of average numbers of units that depart from the respective nodes per unit time when they are busy is

$$\tilde{\mu} = \begin{cases} \mu + \tilde{\lambda} P, & \text{if } L = \infty, \\ \mu \sum_{n=0}^{L} (WP)^n, & \text{if } L < \infty. \end{cases} \tag{37}$$

Also, $w_j = \tilde{\lambda}_j / \tilde{\mu}_j$, which follows by the traffic equation, is the traffic intensity *at node j.*

PROOF. The first two assertions follow by Theorems 1 and 14 [recall that $\Phi(\cdot) = 1$] and the fact that $\prod_{j=1}^m w_j^{x_j}$ is finite if and only if $w_j < 1$ for each j. If $\tilde{\lambda} < \mu$ (i.e., $\overline{w} < 1$), then by Theorem 14 we know that there is a solution w to the traffic equation that satisfies $0 < w < 1$; hence the process X is ergodic. The assertion for the case $L = \infty$ also follows by Theorem 14.

Now, assume that X is ergodic. By Proposition 5, we know that the effective arrival rate to j is $\sum_s \eta(s_l) \lambda_{sj}$. Then this rate equals $\tilde{\lambda}_j$, since the equalities of (33) and (34) along with (35) yield

$$\sum_s \eta(s_l) \lambda_{sj} = \left(\mu \sum_{n=0}^{L-1} (WP)^n W \right)_j = \tilde{\lambda}_j.$$

Next, observe that Proposition 8 says that the effective departure rate from j is

$$\tilde{\lambda}_j \Big/ \sum_x 1(e_j \le x) \pi(x) = \tilde{\lambda}_j / w_j.$$

Then this average equals $\tilde{\mu}_j$ defined by (37) since $\tilde{\mu}_j = w_j^{-1}\tilde{\lambda}_j$ by the traffic equation (32). $\square$

8. Networks with multiple, compound-rate string transitions. In this section, we discuss string-nets with compound-rate string transitions for multiple types of string initiations. We also discuss processes in the literature that are examples of string-nets.

Consider the network process X in the previous section with the following generalizations:

1. There are *multiple types of services* or string initiations indexed by $i \in \mathscr{I}$, and $\mu(i) = (\mu_1(i), \dots, \mu_m(i))$ denotes the type i service rates at the nodes.
2. A type i service completion generates a string of deletions and a possible addition as before, but now the *propagation and quitting probabilities depend on the stage k and index i*. Specifically, a kth departure from node s_k may trigger a departure from node s_{k+1} with probability $P_{s_k s_{k+1}}(i, k)$ or, with probability $Q_{s_k j}(i, k)$, it may add one unit to j and then quit. With no loss in generality, the maximum string length L is independent of i.

The following result is analogous to Theorem 15. The condition (40) for ergodicity is that the arrival rates $\tilde{\lambda}(I)$ into the nodes are less than the service rates $\tilde{\mu}(0)$.

THEOREM 16. *The process X described above is a string-net and its traffic equation is $\tilde{\mu}(W)W = \tilde{\lambda}(W)$, where*

$$\tilde{\lambda}(W) = \lambda + \sum_{i \in \mathscr{I}} \mu(i) \sum_{n=0}^{L-1} \left[\prod_{k=1}^{n} WP(i, k) \right] WQ(i, n+1), \tag{38}$$

$$\tilde{\mu}(W) = \sum_{i \in \mathscr{I}} \mu(i) \sum_{n=0}^{L-1} \left[\prod_{k=1}^{n} WP(i, k) \right]. \tag{39}$$

If $\tilde{\lambda}(\overline{W}) < \tilde{\mu}(0)\overline{W} = \sum_{i \in \mathscr{I}} \mu(i)\overline{W}$ for some $m \times m$ diagonal matrix $\overline{W}$ with positive diagonal entries, then there is a solution w to the traffic equation in $(0, \overline{w})$. In particular, if

$$\tilde{\lambda}(I) < \sum_{i \in \mathscr{I}} \mu(i), \tag{40}$$

then there exists a solution w to the traffic equation in $(0, 1)$, and hence the process X is ergodic and its stationary distribution is $\pi(x) = \prod_{j=1}^{m}(1 - w_j)w_j^{x_j}$, $x \geq 0$, and $\tilde{\lambda}(W)$ and $\tilde{\mu}(W)$ are the effective arrival and service rate vectors.

PROOF. The traffic equation $\tilde{\mu}(W)W = \tilde{\lambda}(W)$ is the obvious analogue of (32). The rest of the proof follows similarly to that of Theorems 14 and 15 because a solution of the traffic equation is a fixed point of the vector-valued function

$$f(w) = (\tilde{\lambda}(W)_1/\tilde{\mu}(W)_1, \dots, \tilde{\lambda}(W)_m/\tilde{\mu}(W)_m), \qquad w \in [0, \overline{w}]. \qquad \square$$

The rest of this section is devoted to examples of the preceding theorem. We start with a case that can be treated as in the last section.

EXAMPLE 17 (Homogeneous propagation and quitting probabilities). Consider the process described above in which all the propagating and quitting probabilities are P_{jk} and Q_{jk}, respectively. Then the traffic equation is the same as that in Proposition 13 with $\mu = \sum_i \mu(i)$. Consequently, the assertions of Proposition 13 and Theorems 14 and 15 apply automatically. This model with $L = \infty$ was studied in [8], but the authors did not establish the existence of the w_j's. We now know by Theorem 14 that the w_j's are given by $w_j = \tilde{\lambda}_j/(\mu + \tilde{\lambda}P)_j$, $1 \le j \le m$. We also now understand the model for $L < \infty$, which had not been considered before.

Another special situation of interest is when the propagating and quitting matrices are homogeneous and equal to P and Q, respectively, after the first stage. Then (38) and (39) reduce to

$$\tilde{\lambda}(W) = \lambda + \sum_{i\in\mathscr{I}} \mu(i)\left[WQ(i,1) + WP(i,1)\sum_{n=0}^{L-2}(WP)^n WQ\right], \tag{41}$$

$$\tilde{\mu}(W) = \sum_{i\in\mathscr{I}} \mu(i)\left[I + WP(i,1)\sum_{n=0}^{L-2}(WP)^n\right]. \tag{42}$$

The following are examples of this case.

EXAMPLE 18 (Regular and negative units with two-stage strings). This example was the subject of [11]–[13]. Consider the process described above in which there are regular and negative units (types 1 and 2) with two-stage strings ($L = 2$) that evolve as follows. Whenever a regular unit finishes a service at node i with rate $\mu(1)_i$, it either enters a node j with probability Q'_{ij} for another service, or the unit becomes a negative unit and enters node j with probability P_{ij}. If this negative unit encounters no units at j, nothing more happens. Otherwise, one unit is deleted from j and one regular unit enters a node k with probability Q_{jk} (entering $k = 0$ means the unit exits the network). In addition, negative units from outside enter the nodes according to independent Poisson processes with rates $\mu(2) = (\mu(2)_1, \dots, \mu(2)_m)$. If a negative unit entering i encounters no units there, then nothing more happens; otherwise, one unit is deleted from i and one regular unit enters a node k with probability Q_{jk}. In terms of the notation above,

$$P(1,1) = P, \qquad Q(1,1) = Q', \qquad Q(1,2) = Q = Q(2,1),$$
$$P(1,2) = P(2,1) = P(2,2) = Q(2,2) = 0.$$

Then Theorem 16 applies with (41), (42) reduced to

$$\tilde{\lambda}(W) = \lambda + \mu(1)[WQ' + WPWQ] + \mu(2)WQ, \qquad \tilde{\mu}(W) = \mu(1) + \mu(2) + \mu(1)WP.$$

The sufficient condition for ergodicity is

$$(\lambda + \mu(1)[Q' + PQ] + \mu(2)Q)_j < (\mu(1) + \mu(2))_j, \qquad 1 \le j \le m.$$

This process was developed in [11], [13], [14]. The model in [11] assumed a negative customer did not generate a regular service [i.e., $Q(2,1) = 0$], and the existence of a solution to the traffic equations was established later in [16]. The authors considered a fixed point for two types of traffic intensities for positive and negative customers in traffic equations that are different from ours but represent essentially the same flow balance. The model in [13] is the one above with different notation and without the sufficient condition for ergodicity. And the model in [14] incorporates the situation in which a signal at a node may trigger a batch (instead of a single unit) to exit the network from that node, similar to the model in the next section. Another extension to batch arrivals as well as batch departures, comparable to those in the next section, is in [21].

EXAMPLE 19 (Regular and negative units with infinite strings). Consider the process related to (41), (42) in which $L = \infty$ and all the propagation and quitting probability matrices are P, Q except for those at the first stage. Clearly $\sum_{n=0}^{\infty}(WP)^n = (I - WP)^{-1}$ exists for $W \le I$, provided that $(I - P)^{-1}$ exists. In this case, Theorem 16 applies and the sufficient condition for ergodicity is

$$\left(\lambda + \sum_{i \in \mathscr{I}} \mu(i)[Q(i,1) + P(i,1)(I - P)^{-1}Q]\right)_j < \sum_{i \in \mathscr{I}} \mu(i)_j, \qquad 1 \le j \le m.$$

Although we know the traffic equations have a solution, we cannot obtain a closed-form expression for it as we did in Example 17. Chao and Pinedo [8] studied the special case of this process with $\mathscr{I} = \{1, 2\}$, and $P(2,1) = P$ and $Q(2,1) = Q$, but they did not establish the existence of a solution to the traffic equations.

9. String-nets with two-node batch transitions. In this section, we discuss string-nets in which a transition involves a batch deletion at a single node and a batch addition at another single node.

Suppose the string-net X is such that each element of A is of the form $a = ne_j$ and each nonzero string in S is of the form $s = (e_i \cdots e_i)$ (l copies of e_i), where $0 \le j \le m$, $n \ge 1$ and $l \le \infty$. This means that for such a pair sa, the complete and partial transitions are $x \to x - le_i + ne_j$ or $x \to x - ke_i$. We say that X has *two-node string transitions* since exactly two nodes are affected in a transition. We write $\lambda_{sa} = \lambda_{li,nj}$. Under an sa-transition with $s = e_i e_i \dots$ and $l = \infty$, all units from node i would be cleared out and we denote its rate simply by $\lambda_{\infty i}$. Such a "clearing" transition might represent a dispatching or assembly of units (or a catastrophe [5]) that clears out all units at i. Here we let $\gamma = (\gamma_{nj}\colon 1 \le j \le m,\ n \ge 1)$ denote the vector γ in Theorem 2 and $\gamma_{0j} = 1$. Also, the summations on l are the conventional ones that do not include a term for $l = \infty$; this term is treated separately.

THEOREM 20. *For the network process X with two-node batch transitions, the traffic equations* (8) *are*

$$\gamma_{nj}\sum_{l=0}^{\infty}\gamma_{lj}[r_{n+l,\,j}+\lambda_{\infty j}]=\sum_{li}\gamma_{li}\lambda_{li,\,nj},\qquad 1\le j\le m,\ n\ge 1, \tag{43}$$

where $r_{k,\,j}\equiv\sum_{l'\ge k}\sum_{n'j'}\lambda_{l'j,\,n'j'}$. If these equations have a solution of the form $\gamma_{nj}=w_j^n$, for some positive $w_1,\dots,w_m$, then $\pi(x)=\Phi(x)\prod_{j=1}^m w_j^{x_j}$, $x\in E$, is an invariant measure for the process.

PROOF. This follows by Theorem 1, where the traffic equations (8) reduce to (43) since, for any $sa=li,\,nj$, the $\Lambda_{(sa)}=r_{n+l,j}1(i=j)$ for $l<\infty$ and $\Lambda_{(sa)}=\lambda_{\infty j}1(i=j)$ for $l=\infty$. □

Here are a few examples.

EXAMPLE 21 (Open Jackson process with periodic clearing). Suppose the process X with two-node batch transitions has strings of only length 1 or ∞ and $A=\{e_1,\dots,e_m\}$. Then all transitions are standard Jackson types ($x\to x-e_i+e_j$) or there is a clearing ($x\to x-x_ie_i$). The rates of these transitions are

$$q(x,x-e_i+e_j)=\phi_j(x)\lambda_{i,\,j},\qquad q(x,x-x_ie_i)=\phi_i(x)\lambda_{\infty i}. \tag{44}$$

We call X a *Jackson network process with periodic clearing*. Without loss in generality, assume that $\lambda_{i,\,j}$ is an irreducible matrix.

THEOREM 22. *Suppose the process X with transition rates* (44) *satisfies $\sum_x\Phi(x)<\infty$. Then it is ergodic and its stationary distribution is $\pi(x)=c\Phi(x)\prod_{j=1}^m w_j^{x_j}$, $x\in E$, where $w_0=1$ and $w_1,\dots,w_m$ in $(0,1)$ satisfy the traffic equations*

$$w_j\left[\sum_i\lambda_{j,\,i}+\lambda_{\infty j}\sum_{\nu=0}^{\infty}w_j^{\nu}\right]=\sum_i w_i\lambda_{i,\,j},\qquad 1\le j\le m. \tag{45}$$

Furthermore, the effective arrival and service rates, $\tilde\lambda_j$ and $\tilde\mu_j$, for node j are given by the sums on the right and left sides of (45), *respectively.*

PROOF. First note that equations (45) are clearly a special case of the traffic equations (43). We will consider (45) written as $w_jg_j(w)=h_j(w)$ and apply Theorem 2 to justify that it has a solution. To this end, let $\overline{w}$ be a vector in $(0,1)$ that satisfies

$$\overline{w}_j\sum_i\lambda_{j,\,i}=\sum_i\overline{w}_i\lambda_{i,\,j},\qquad 1\le j\le m.$$

Define $A^*=\{j\colon\lambda_{0,\,j}>0\}$. This set is not empty because $\lambda_{i,\,j}$ is irreducible. Let $\underline{w}$ be a vector in $(0,\overline{w})$ such that

$$\underline{w}_j<\lambda_{0,\,j}/g_j(\overline{w}),\qquad j\in A^*,$$

$$\underline{w}_j<\sum_{i\in A^*}\underline{w}_i\lambda_{ij}/g_j(\overline{w}),\qquad j\in\{1,\dots,m\}\backslash A^*.$$

From the definition of $\overline{w}$ and these inequalities, it follows that

$$\underline{w}_j g_j(\overline{w}) < h_j(\underline{w}) \quad \text{and} \quad h_j(\overline{w}) = \overline{w}_j \sum_i \lambda_{j,i} < \overline{w}_j g_j(\underline{w}), \qquad 1 \le j \le m.$$

From these inequalities and Theorem 2, it follows that (45) has a solution $w \in (0,1)$. Then the first assertion of the theorem follows by Theorem 1. Also, Proposition 8 justifies, as in the proof of Theorem 15, that (45) is the same as $w_j \tilde{\mu}_j = \tilde{\lambda}_j$. $\square$

EXAMPLE 23 (Assembly networks). Consider the string-net described in Theorem 22 with the following features. Units arrive to the nodes by independent Poisson processes with rates $\lambda_1, \dots, \lambda_m$. Services at node i are exponential with rate μ_i. When a service at i completes, K_i units, if available at i, are assembled into one unit and sent to node j with probability Q_{ij}. If there are less than K_i units at i, then all the units at i are assembled into one defective unit and discarded (sent to node 0). Then the process X that represents the numbers of units at the nodes has two-node batch transitions. Its traffic equations (43) are

$$w_j \mu_j \sum_{l=0}^{K_j-1} w_j^l = \lambda_j + \sum_{i=1}^m \mu_i w_i^{K_i} Q_{ij}, \qquad 1 \le j \le m.$$

Then Theorem 22 applies in this setting under the assumption that

$$\lambda_j + \sum_{i=1}^m \mu_i Q_{ij} < K_j \mu_j,$$

which says that the service capacity at node j is greater than the arrival rate. This model was introduced in [10] but did not establish the existence of a solution to the traffic equation.

EXAMPLE 24 (Batch assembly-transfer networks). Consider the preceding assembly network with the following generalizations. Batches of units enter the network at node j according to a Poisson process with rate λ_j, and $a_j(n)$ is the probability that the batch is of size n. Services at node i are exponential with rate μ_i and $b_i(l)$ is the probability that a batch of size l is requested; $Q_{li,nj}$ is the probability that a batch of size l departing from i becomes a batch of n units that enters node j. The resulting process is a string-net with two-node batch transitions and

$$\lambda_{0,nj} = \lambda_j a_j(n), \qquad \lambda_{li,nj} = \mu_i b_i(l) Q_{li,nj}, \qquad \phi_{ne_j}(x) = 1(ne_j \le x_j).$$

By Theorem 20, its traffic equations (43) are

$$\gamma_{nj}\mu_j \sum_{l=0}^{\infty} \gamma_{lj} \sum_{l'=n+l}^{\infty} b_j(l') = \lambda_j a_j(n) + \sum_{i=1}^m \mu_i \sum_{l=1}^{\infty} \gamma_{li} b_i(l) Q_{li,nj},$$

$$1 \le j \le m, \; n \ge 1.$$

This process is the subject of [26] and [7], which gives insights into the process when it does not have a stationary distribution.

10. Further examples. In this section, we describe one-dimensional processes with string transitions. In addition to giving more insight into string transitions, these processes for single nodes can be used as building blocks for networks, comparable to quasi-reversible nodes that are coupled together to form networks. We end by discussing how the results herein apply to multiple types of units.

EXAMPLE 25 (One-dimensional process with string transitions). Consider the Markov network process X for the special case in which the network consists of a single node. For simplicity, assume that the state space is $E = \{0, 1, 2, \dots\}$ and that each increment of a string transition is of unit length. In this setting, a complete sa-transition is $x \to x - s + a$, and a kth partial sa-transition is $x \to x - k$, where $a \in \{0, 1\}$ and the string s now denotes s unit subtractions. Then the transition rates (1) of the process X are $q(x, x+1) = \lambda_{01}\phi_0(x)$ and

$$\begin{aligned} q(x, x-k) &= \sum_{a=0,1} \lambda_{k+a,a}\big[\phi_{k+a}(x) + \phi_k(x) - \phi_{k+1}(x)\big) \\ &= \lambda_{k,0}[2\phi_k(x) - \phi_{k+1}(x)] + \lambda_{k+1,1}\phi_k(x), \qquad 0 < k < x. \end{aligned}$$

Theorem 1 justifies that $\pi(x) = \Phi(x)w^x$ is an invariant measure for the process, where $w > 0$ satisfies

$$\sum_{s=0}^{L-1} w^{s+1}\Lambda_{s+1} = \sum_{s=0}^{L} w^s \lambda_{s1}, \tag{46}$$

and $\Lambda_s = \sum_{s'=s}^{L}(\lambda_{s'0}+\lambda_{s'1})$. Clearly w is the unique positive solution to (46) since this equation is equivalent to $\sum_{s=1}^{L} w^s[\lambda_{s0} + \Lambda_{s+1}] = \lambda_{01}$, which has a unique solution. Note that the process is positive recurrent if and only if $0 < w < 1$. A special case of this model is as follows.

EXAMPLE 26 (A simple production-inventory system). Consider a production system whose cumulative output over time is a Poisson process with rate λ. As the units are produced, they are put in inventory to satisfy random demands. Let $X(t)$ denote the quantity of units in inventory at time t. Whenever there are x units in inventory, the time to the next demand has an exponential distribution with rate μ and the demand is for k units with probability p_k, where $k \le L$. Also, the probability that the demand can be satisfied is $P\{Z \le x - k | Z \le x\}$, where $1 \le k \le \min\{x, L\}$. Think of Z as a nonnegative integer-valued random variable that denotes a feasible inventory level. Then the process X is clearly a Markov process, its nonzero transition rates are $q(x, x+1) = \lambda$ and

$$q(x, x-k) = \mu p_k P\{Z \le x - k | Z \le x\}, \qquad 1 \le k \le \min\{x, L\}.$$

An easy check shows that this process is a special case of the preceding example in which

$$\lambda_{s0} = 0, \qquad \lambda_{01} = \lambda, \qquad \lambda_{k+1,1} = \mu p_k, \qquad \phi_k(x) = P\{Z \le x\}.$$

Therefore, $\pi(x) = \Phi(x)w^x$, $x \ge 0$, is a stationary distribution, where w is the unique solution to $\sum_{s=1}^{L} w^s(p_s + \cdots + p_{L-1}) = \lambda/\mu$.

REMARK 27 (Multiple types of units). Networks with multiple types of units are naturally represented by doubly indexed states $x = (x_{cj}\colon 1 \le j \le m,\ c \in C)$, where x_{cj} denotes the number of units of class or type c at node j. The definition of string transitions extends to this setting by simply using double indices on the increment vectors, strings and rate functions (e.g., $a = (a_{cj})$ and $\lambda_{s,cj}$). Then all of the results herein apply.

For networks with unit-vector strings, a common assumption is that if a c-unit at node j initiates a transition, then the transition rate (such as the ones we have been discussing) is multiplied by the portion $p_{cj}(x) = x_{cj}/\sum_{c'} x_{c'j}$ of c-units at j. These functions are clearly Φ'-reversible, where

$$\Phi'(x) = \prod_{c,\,j} \sum_{c'} x_{c'j}!/x_{cj}!.$$

In this case, the invariant measures for the network process are of the form $\Phi'(x)\Phi(x)\prod_{cj} w_{cj}^{x_{cj}}$. In other words, this multiplication by $\Phi'(x)$ is the only basic difference between the homogeneous unit and multiple unit processes. This explains the presence of the product $\Phi'(x)$ in distributions for multiple units; see for instance [8] or [16].

REFERENCES

[1] BOUCHERIE, R. and VAN DIJK, N. (1990). Spatial birth-death processes with multiple changes and applications to batch service networks and clustering processes. *Adv. in Appl. Probab.* **22** 433–455.

[2] BOUCHERIE, R. and VAN DIJK, N. (1990). Product forms for queueing networks with state dependent multiple job transitions. *Adv. in Appl. Probab.* **23** 152–187.

[3] BOUCHERIE, R. and VAN DIJK, N. (1991). Local balances in queueing networks with positive and negative customers. *Ann. Oper. Res.* **48** 463–492.

[4] CHAO, X. (1995). Networks of queues with customers, signals and arbitrary service time distributions. *Oper. Res.* **43** 537–544.

[5] CHAO, X. (1995). A queueing network model with catastrophes and product form solution. *Oper. Res. Lett.* **18** 75–79.

[6] CHAO, X. and MIYAZAWA, M. (1995). A probabilistic decomposition approach to quasi-reversibility and its applications in coupling of queues. Technical report, New Jersey Institute of Technology and Science Univ. Tokyo.

[7] CHAO, X. and MIYAZAWA, M. (1996). Queueing networks with instantaneous movements: a couping approach by quasi-reversibility. Technical report, New Jersey Institute of Technology and Science Univ. Tokyo.

[8] CHAO, X. and PINEDO, M. (1993). On generalized networks of queues with positive and negative arrivals. *Probab. Engrg. Inform. Sci.* **7** 301–334.

[9] CHAO, X. and PINEDO, M. (1995). Networks of queues with batch services, signals and product form solutions. *Oper. Res. Lett.* **17** 237–242.

[10] CHAO, X., PINEDO, M. and SHAW, D. (1996). An assembly network of queues with product form solution. *J. Appl. Probab.* To appear.

[11] GELENBE, E. (1991). Product-form queueing networks with negative and positive customers. *J. Appl. Probab.* **28** 656–663.

[12] GELENBE, E. (1992). Stability of product form G-networks. *Probab. Engrg. Inform. Sci.* **6** 271–276.

[13] GELENBE, E. (1993a). G-Networks with triggered customer movements. *J. Appl. Probab.* **30** 931–942.

[14] GELENBE, E. (1993b). G-Networks with signals and batch removals. *Probab. Engrg. Inform. Sci.* **7** 335–342.

[15] GELENBE, E., GLYNN, P. and SIGMAN, K. (1991). Queues with negative arrivals. *J. Appl. Probab.* **28** 245–250.

[16] GELENBE, E. and SCHASSBERGER, R. (1992). Stability of G-Networks. *Probab. Engrg. Inform. Sci.* **6** 3271–3276.

[17] GORDON, W. J. and NEWELL, G. F. (1967). Cyclic queueing systems with restricted queue lengths. *Oper. Res.* **15** 266–278.

[18] HENDERSON, W. (1993). Queueing networks with negative customers and negative queueing lengths. *J. Appl. Probab.* **30** 931–942.

[19] HENDERSON, W., NORTHCOTE, B. S. and TAYLOR, P. G. (1994). State-dependent signaling in queueing networks. *Adv. in Appl. Probab.* **24** 934–959.

[20] HENDERSON, W., NORTHCOTE, B. S. and TAYLOR, P. G. (1994). Geometric equilibrium distributions for queues with interactive batch departures. *Ann. Oper. Res.* **48** 463–492.

[21] HENDERSON, W., NORTHCOTE, B. S. and TAYLOR, P. G. (1995). Triggered batch movements in queueing networks. *Queueing Systems Theory Appl.* **21** 125–141.

[22] HENDERSON, W., PEARCE, C. E. M., TAYLOR, P. G. and VAN DIJK, N. M. (1990). Closed queueing networks with batch services. *Queueing Systems Theory Appl.* **6** 59–70.

[23] HENDERSON, W. and TAYLOR, P. G. (1990). Product form in networks of queues with batch arrivals and batch services. *Queueing Systems Theory Appl.* **6** 71–78.

[24] JACKSON, J. R. (1957). Networks of waiting lines. *Oper. Res.* **5** 518–552.

[25] KELLY, F. P. (1979). *Reversibility and Stochastic Networks*. Wiley, London.

[26] MIYAZAWA, M. and TAYLOR, P. G. (1997). A geometric product-form distribution for a queueing network with nonstandard batch arrivals and batch transfers. *Adv. in Appl. Probab.* To appear.

[27] SERFOZO, R. F. (1989). Markovian network processes: congestion-dependent routing and processing. *Queueing Systems Theory Appl.* **5** 5–36.

[28] SERFOZO, R. F. (1991). Reversibility and compound birth-death and migration processes. *Queueing Processes and Related Models*. Oxford Univ. Press.

[29] SERFOZO, R. F. (1993). Queueing networks with dependent nodes and concurrent movements. *Queueing Systems Theory Appl.* **13** 143–182.

[30] SERFOZO, R. F. (1994). Little laws for utility processes and waiting times in queues. *Queueing Systems Theory Appl.* **17** 137–81.

[31] WALRAND, J. (1988). *Introduction to Queueing Networks*. Prentice-Hall, Englewood Cliffs, NJ.

[32] WHITTLE, P. (1986). *Systems in Stochastic Equilibrium*. Wiley, New York.

SCHOOL OF INDUSTRIAL AND SYSTEMS ENGINEERING
GEORGIA INSTITUTE OF TECHNOLOGY
ATLANTA, GEORGIA 30332
E-MAIL: rserfozo@isye.gatech.edu

CAPS LOGISTICS INC.
2575 CUMBERLAND PARKWAY
ATLANTA, GEORGIA 30339
E-MAIL: bingyiy@caps.com