Probability in the Engineering and Informational Sciences, **16**, 2002, 241–270. Printed in the U.S.A.

ON THE STRUCTURE OF THE SPACE OF GEOMETRIC PRODUCT-FORM MODELS

NIMROD BAYER
Department of Electrical and Computer Engineering
Ben-Gurion University of the Negev
Beer-Sheva 84105, Israel
E-mail: bayer@fuji.ee.bgu.ac.il

RICHARD J. BOUCHERIE
Faculty of Mathematical Sciences
University of Twente
7500 AE Enschede, The Netherlands

This article deals with Markovian models defined on a finite-dimensional discrete state space and possess a stationary state distribution of a product-form. We view the space of such models as a mathematical object and explore its structure. We focus on models on an orthant $\mathcal{Z}_+^n$, which are homogeneous within subsets of $\mathcal{Z}_+^n$ called walls, and permit only state transitions whose $\|\ \|_\infty$-length is 1. The main finding is that the space of such models exhibits a *decoupling principle*: In order to produce a given product-form distribution, the transition rates on distinct walls of the same dimension can be selected without mutual interference. This principle holds also for state spaces with multiple corners (e.g., bounded boxes in $\mathcal{Z}_+^n$).

In addition, we consider models which are homogeneous throughout a finite-dimensional grid $\mathcal{Z}^n$, now without a fixed restriction on the length of the transitions. We characterize the collection of product-form measures which are invariant for a model of this kind. For such models with bounded transitions, we prove, using Choquet's theorem, that the only possible invariant measures are product-form measures and their combinations.

1. INTRODUCTION

1.1. Background and Motivation

Product-form models of networks and systems have been known since the days of Erlang [5] (see also Kelly [20, p. 321]). However, the startling discovery in 1957, by

 0269-9648/02 $12.50

Jackson [15], of the first product-form queuing network boosted the discovery of further and ever more general classes of product-form models. Some notable landmarks include the following: the generalization made by Jackson himself in 1963 [16], of which a special case, closed networks, was rediscovered by Gordon and Newell in 1967 [13]; queuing networks with multiple customer classes and with various service disciplines (Kelly [17], Baskett et al. [1]); substantiation of insensitivity results for queueing networks (Schassberger [26]); networks with negative customers (Gelenbe [12]); networks with customer batches (e.g., Henderson and Taylor [14] and Boucherie and van Dijk [3]); and queuing networks with signals (Chao and Pinedo [9]).

Along with the discovery of new classes of product-form models, attempts have been made to understand the basis for the product-form property: It has not been clear whether the models discovered have been rare contingencies or samples from a broad space of product-form Markovian models. It has also not been clear what is required from a model to possess the product-form property. Chao and Miyazawa [6], following Kelly [19], point out that two explanations have been given, belonging to two different aspects. The first explanation, introduced by Whittle in 1967 and 1968 [29,30], was that the product-form models are those exhibiting *local balance* or *partial balance*. This means that the global balance equation can be decomposed into sets of partial balance equations in an appropriate way. The second explanation, apparently introduced by Muntz in 1972 [22], was that models of product-form queuing networks are those exhibiting *quasireversibility*: an input–output relation, holding separately for each node of the network. Quasireversibility is different from *reversibility*; the latter property, which is rather special, indeed leads to product-form via local balance (see Kelly [18] or Whittle [31]).

The discovery of Gelenbe networks shattered the partial balance explanation in the original sense of Whittle, because these networks do not possess such partial balance. However, the notion of partial balance was restored, and even assumed new forms, when Boucherie and van Dijk identified a type of partial balance in Gelenbe networks in 1994 [4] and when Chao and Miyazawa discovered *biased local balance* in 1998 [6].

In these attempts at explaining the basis for the product-form property, there were two trends. One trend, represented by van Dijk's 1993 monograph [11], has been to gain insight regarding the product-form property from a system point of view. The opposing or complementary trend has been to seek an ever more general setting, to such an extent that a fairly abstract Markovian framework emerges while some of the tangible system attributes are lost. A characterization of product-form networks, in this more abstract approach, was given by Chao, Miyazawa, Serfozo, and Takada in [8]; this characterization is also available in [27, Thm. 8.8, p. 213] and in [7, Thm. 11.3, p. 315]. (Here, we deal only with a discrete state space, but a characterization of product-form for stochastic networks on a continuous state space is given in Williams [32, Thm. 3.5].) It turns out that in such an abstract framework, neither quasireversibility nor partial balance, even biased, are a necessary condition for product-form.

In this article, we pursue the abstract approach further and strip down the model completely from any system or network concepts such as "arrival" or "departure." Instead, we focus on the Markovian transition structure that appears in models. Here, we study the simplest Markovian transition structures: those with space homogeneity, as appear in queuing networks with single servers, on a state space which is a product space. Furthermore, we allow a certain flexibility for the transition structure on the boundary of the state space. In adopting this approach, we are inspired by studies which looked at models with transition rates modified at the boundary (e.g., Miyazawa and Taylor [21]). Under this approach, the behavior of the Markov chain in the interior of its state space is interpreted as representing the "generic behavior" of a network or a system. The modified model (i.e., the model with the modified transition rates on the boundary) can be used for purposes of approximation and bounding. Work on approximation and bounding of performance criteria using altered models have been done by van Dijk and others; see, for example, Taylor and van Dijk [28]. Characterizing the space of possible modifications may serve as the starting point for model design, possibly in the form of searching for the most suitable product-form approximation for a given original model that is not of a product-form. Practical stochastic network models not of a product-form do exist: See the work of Bayer and Kogan [2] on branching/queuing networks; branching/queuing networks fall outside of the above-mentioned characterization framework set in [8]. Virtually all non-product-form Markovian models of stochastic networks, or related abstracted models such as random walks, do not yield explicit and exact analytical results. Their stationary state distributions are difficult to see. This is the motivation for clinging to product-form models for the purpose of direct modeling or approximation and for searching the space of such models.

Our main objects of study are indeed spaces of models, rather than individual models. Hence, the main finding cannot be expressed at the level of a fixed, individual model. In this article, the word *model* designates some setup of parameters directly determining the transition rates, or *q-matrix*, of a continuous-time Markov chain. Models will be rendered in this article as functions, as function arrays, or as finite vectors. We restrict ourselves to models possessing the property of *space homogeneity* or, simply, *homogeneity*. This property prevails when parallel transitions within the state space $\mathcal{S}$, or within subsets of $\mathcal{S}$ referred to as *walls*, necessarily have the same rate. The Markov chain associated with such a model has the random-walk structure. Accordingly, the product-form distributions considered here are geometric; that is, of the type

$$\pi(\vec{\mathbf{a}}) = c \prod_{i=1}^{n} q_i^{a_i}, \qquad \vec{\mathbf{a}} = \langle a_1, \ldots, a_n \rangle \in \mathcal{S}, \tag{1}$$

where c and the q_i's are constants.

The role of such product-form measures is investigated from two different perspectives. First, it is shown that when the transition rates in the interior of $\mathcal{S}$ are interpreted as representing the generic behavior of the Markov chain, and when this

behavior is extended to a boundaryless state space, product-form measures emerge as the fundamental invariant measures. Second, the design point of view is taken. It is shown that for every product-form measure over the nonnegative orthant, a broad variety of q-matrices can be constructed which have this product-form measure as their invariant measure. The space of such q-matrices corresponding to a given product-form measure is characterized precisely and its structure is described. This can be viewed in another way: Given the generic behavior of the chain and given any product-form measure associated with this behavior (out of a set which we also characterize and describe), boundary rates can be constructed in a broad variety of ways such that the product-form solution is valid on the boundary too. In Section 1.2, we describe our findings in more detail.

1.2. Summary of the Results

The generic behavior of a Markov chain with a random-walk structure, as represented by its transition rates in the interior of the state space, may lead to various product-form invariant measures. To study the role of these product-form measures, within the set of all invariant measures of the chain, the space $\mathcal{M}_n$ of models corresponding to space-homogeneous continuous-time Markov chains on the n-dimensional integer grid $\mathcal{Z}^n$ is considered first. Although the notion of a stationary state distribution is not relevant for a model $\varphi \in \mathcal{M}_n$, because the corresponding chain is not positive recurrent, there may exist infinite measures which are invariant to the transition operators associated with φ. Such measures are said to be invariant for φ. Only product-form measures that are invariant for φ can serve as candidates for the stationary state distribution of a product-form model on the orthant $\mathcal{Z}_+^n$ whose interior transition rates coincide with those of φ. The question arises: Which models in $\mathcal{M}_n$ ($n \geq 2$) have product-form invariant measures? The answer is that every $\varphi \in \mathcal{M}_n$ whose drift is finite but nonzero has quite a set of invariant product-form measures: The corresponding set of vectors $\vec{\mathbf{q}} = \langle q_1, \ldots, q_n \rangle$ (playing a role as in (1)) is a smooth $(n-1)$-dimensional manifold Q_φ, which is the boundary of a bounded and convex set in $(0,\infty)^n$. When φ has bounded transitions, we are able to prove that the product-form measures corresponding to Q_φ are φ's fundamental invariant measures, in the following sense: Any other measure invariant for φ cannot be but a mixture

$$\int_{Q_\varphi} \pi_{\vec{\mathbf{q}}} \, d\zeta(\vec{\mathbf{q}})$$

of these product-form measures $\pi_{\vec{\mathbf{q}}}$. The proof applies Choquet's theorem.

Observing the fundamental role of product-form measures, the issue of model design arises and may take the following form: Given a model φ that represents the transition structure at the interior of the state space, what are the possibilities and degrees of freedom, if any, to construct transition rates at the boundary, such that the invariant measure is some $\pi_{\vec{\mathbf{q}}}$, with $\vec{\mathbf{q}}$ selected from Q_φ? To answer this question, a further space $\mathbb{MI}_n$ of models on the nonnegative orthant of the n-dimensional grid is

introduced. Its definition relies on partitioning the orthant into *walls*. Most walls are parts of the boundary, but the interior is also referred to as a "wall" for notational convenience. The wall to which some state $\vec{\mathbf{a}} = \langle a_1, \ldots, a_n \rangle$ belongs is determined by the coordinates i where $a_i = 0$. The walls are attributed with dimensions, which range from 0 to n. The only wall of dimension 0 is the *corner*, which contains the single point $\langle 0, \ldots, 0 \rangle$. The homogeneity property postulated for $\mathbb{MI}_n$ is weaker than for $\mathcal{M}_n$: Homogeneity prevails within each wall, but parallel transitions belonging to different walls may be assigned different rates. Apart from homogeneity, the models in $\mathbb{MI}_n$ are assumed in this article to comply with a further restrictive assumption: A transition from $\vec{\mathbf{a}} = \langle a_1, \ldots, a_n \rangle$ to $\vec{\mathbf{b}} = \langle b_1, \ldots, b_n \rangle$ is possible only if this transition is "short," in the sense that $|a_i - b_i| \leq 1$, $i = 1, \ldots, n$.

Let $\mathbb{MI}_n$'s subspace of models with a product-form stationary state distribution be denoted by $\mathbb{P}_n$. The structure of $\mathbb{P}_n$ can be described through a model-selection procedure. This procedure may start from the selection of the interior transition rates, which is tantamount to a selection of an arbitrary model φ with short transitions from $\mathcal{M}_n$. The next step may then be the selection of a vector $\vec{\mathbf{q}} = \langle q_1, \ldots, q_n \rangle$ from the $(n-1)$-dimensional manifold Q_φ. An alternative way is to start from an arbitrary $\vec{\mathbf{q}}$ with $0 < q_i < 1$, $i = 1, \ldots, n$. (The restriction $q_i < 1$ is needed if the product-form measure is to be normalizable; actually, we could allow $q_i = 1$ to hold for any proper subset of the coordinates, but this would have incurred a subtlety which we avoid for simplicity.) For every such $\vec{\mathbf{q}}$, let $\mathbb{P}_{n,\vec{\mathbf{q}}}$ denote the models in $\mathbb{P}_n$ for which $\pi_{\vec{\mathbf{q}}}$ is the invariant product-form distribution. To select a model from $\mathbb{P}_{n,\vec{\mathbf{q}}}$, perform the following procedure. First, select the interior transition rates. This selection is subject to a single linear constraint, so the number of degrees of freedom is 1 less than the number of variables. Next, select the transition rates within the walls of dimension $n-1$. These selections are decoupled from each other (i.e., they are not coupled by any joint constraint). The same rule regarding the number of degrees of freedom applies again, for each of these walls. The procedure so proceeds to walls of ever smaller dimension, until the walls of dimension 1 are reached and the selection is exhausted. The same rules hold throughout. While the selection for a wall depends on earlier selections for neighboring walls with higher dimensions, this selection is decoupled from other walls of the same dimension. Thus, we have the *decoupling principle* and, along with it, a broad wealth of product-form models.

What makes this procedure valid is the fact that the selections so taken are guaranteed not to interfere with each other at the corner. The corner constraint is shown to be redundant. The proof that we bring for this redundancy is one that allows the generalization of the decoupling principle for state spaces with multiple corners (e.g., n-dimensional bounded boxes). Such state spaces arise, for example, in connection with stochastic networks with finite buffers [11].

1.3. Organization of This Article

This article proceeds with a section of preliminaries, Section 2, that states some conventions and introduces model spaces and related objects. Then, the first section

of results (Sect. 3) is dedicated to $\mathcal{M}_n$, the space of models on the grid $\mathcal{Z}^n$. The section that follows (Sect. 4) is dedicated to $\mathbb{M}_n$, the space of models on the orthant $\mathcal{Z}_+^n$. All proofs are concentrated in Section 5. The article concludes with a basic example of model design (Sect. 6).

2. PRELIMINARIES

The main concern of this section is the definition of models and model spaces. This definition requires a preliminary discussion of state spaces and their walls, which is given after introducing some general conventions. Furthermore, the transition structure needs to be defined. Finally, invariant measures are introduced, thus providing the starting point for the analysis in Section 3.

2.1. General Conventions

Let $\mathcal{R}$, $\mathcal{R}_+$, $\mathcal{Z}$, $\mathcal{Z}_+$, and $\mathcal{N}$ denote the real numbers, the nonnegative reals, the integers, the nonnegative integers, and the positive integers, respectively. Define $\mathcal{B} \triangleq \{0,1\}$ and $\mathcal{T} \triangleq \{-1,0,1\}$; these sets are used in the definition of walls and in the definition of the transition structure.

The symbols $\vec{\mathbf{1}}$ and $\vec{\mathbf{0}}$ stand for vectors of all 1's and all 0's, with their dimension implied by the context. Suppose that $\vec{\mathbf{x}} = \langle x_1, \ldots, x_k \rangle$ and $\vec{\mathbf{y}} = \langle y_1, \ldots, y_k \rangle$ are two vectors and that A is a set of vectors of the same dimension. Define $\vec{\mathbf{x}}\vec{\mathbf{y}}$ to be the vector $\langle x_1 y_1, \ldots, x_k y_k \rangle$, define $\vec{\mathbf{x}}A$ to be the set $\{\vec{\mathbf{x}}\vec{\mathbf{y}} / \vec{\mathbf{y}} \in A\}$, and define $\vec{\mathbf{x}}^{\vec{\mathbf{y}}}$ to be the scalar $\prod_{i=1}^k x_i^{y_i}$. Interpret $|\vec{\mathbf{x}}|$ as $\langle |x_1|, \ldots, |x_k| \rangle$. The norms $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are defined as usual: $\|\vec{\mathbf{x}}\|_1 \triangleq \sum_{i=1}^k |x_i|$ and $\|\vec{\mathbf{x}}\|_\infty \triangleq \max_{i=1,\ldots,k} |x_i|$. Relations such as $\vec{\mathbf{x}} \leq \vec{\mathbf{y}}$ or $\vec{\mathbf{x}} < \vec{\mathbf{y}}$ are interpreted in the componentwise sense. Here, a nonnegative vector $\vec{\mathbf{x}}$ is said to be *majorized* by another nonnegative vector $\vec{\mathbf{y}}$, if $\vec{\mathbf{x}} \leq \vec{\mathbf{y}}$ as well as $\|\vec{\mathbf{x}}\|_1 < \|\vec{\mathbf{y}}\|_1$ hold; we write $\vec{\mathbf{x}} \prec \vec{\mathbf{y}}$.

2.2. State Spaces and Walls

Our state space $\mathcal{S}$ will be either an n-dimensional grid $\mathcal{Z}^n$, or its nonnegative orthant $\mathcal{Z}_+^n$. For subsets of $\mathcal{S}$, we use the following notion of dimension, applying in this discrete context only.

DEFINITION 2.1: *The dimension of $A \subset \mathcal{Z}^n$ is less than or equal to k if there exist some $\vec{\mathbf{x}}_1, \ldots, \vec{\mathbf{x}}_k \in \mathcal{Z}^n$ and an $\vec{\mathbf{a}} \in A$ such that every $\vec{\mathbf{b}} \in A$ admits a representation $\vec{\mathbf{b}} = \vec{\mathbf{a}} + \sum_{i=1}^k m_i \vec{\mathbf{x}}_i$ with $m_1, \ldots, m_k \in \mathcal{Z}$.*

We now introduce the partitioning of $\mathcal{Z}_+^n$ into walls. Define

$$\mathcal{W}_{n,\vec{\mathbf{w}}} \triangleq \vec{\mathbf{w}}\mathcal{N}^n, \qquad \vec{\mathbf{w}} \in \mathcal{B}^n, \tag{2}$$

applying the "$\vec{\mathbf{x}}A$" convention from the previous subsection. The vector $\vec{\mathbf{w}} \in \mathcal{B}^n$ uniquely determines the wall $\mathcal{W}_{n,\vec{\mathbf{w}}}$. Walls are obviously disjoint, and we have the following.

Observation 2.1: $\mathcal{Z}_+^n$ is the disjoint union $\bigcup_{\vec{\mathbf{w}}\in\mathcal{B}^n} \mathcal{W}_{n,\vec{\mathbf{w}}}$.

The reason for referring to the $\mathcal{W}_{n,\vec{\mathbf{w}}}$'s as the walls of $\mathcal{Z}_+^n$ is that all of them except $\mathcal{W}_{n,\vec{\mathbf{1}}}$ are parts of its boundary. The exception $\mathcal{W}_{n,\vec{\mathbf{1}}} = \mathcal{N}^n$ constitutes the interior of $\mathcal{Z}_+^n$ and is referred to as a "wall" for notational convenience. The dimension of $\mathcal{W}_{n,\vec{\mathbf{w}}}$ is clearly $\|\vec{\mathbf{w}}\|_1$. Thus, the walls have various dimensions. For example, the walls of $\mathcal{Z}_+^3$ include the zero-dimensional wall $\mathcal{W}_{3,\langle 0,0,0\rangle}$ (the corner), the one-dimensional walls $\mathcal{W}_{3,\langle 1,0,0\rangle}$, $\mathcal{W}_{3,\langle 0,1,0\rangle}$, and $\mathcal{W}_{3,\langle 0,0,1\rangle}$, the two-dimensional walls $\mathcal{W}_{3,\langle 0,1,1\rangle}$, $\mathcal{W}_{3,\langle 1,0,1\rangle}$, and $\mathcal{W}_{3,\langle 1,1,0\rangle}$, and the three-dimensional interior "wall" $\mathcal{W}_{3,\langle 1,1,1\rangle}$. See Figure 1.

2.3. Classes of State Transitions

Our definition of walls leads us to single out some classes of transitions between state space points. The class of all possible transitions in $\mathcal{Z}^n$ is

$$\mathcal{D}_n \triangleq \{\vec{\mathbf{b}} - \vec{\mathbf{a}}/\vec{\mathbf{a}}, \vec{\mathbf{b}} \in \mathcal{Z}^n, \vec{\mathbf{a}} \neq \vec{\mathbf{b}}\} = \mathcal{Z}^n \backslash \{\vec{\mathbf{0}}\}.$$

The classes of short transitions within the walls of $\mathcal{Z}_+^n$ are

$$\mathcal{D}_{n,\vec{\mathbf{w}}} \triangleq \{\vec{\mathbf{b}} - \vec{\mathbf{a}}/\vec{\mathbf{a}}, \vec{\mathbf{b}} \in \mathcal{W}_{n,\vec{\mathbf{w}}}, \|\vec{\mathbf{b}} - \vec{\mathbf{a}}\|_\infty = 1\}, \qquad \vec{\mathbf{w}} \in \mathcal{B}^n. \tag{3}$$

An illustration of the transition classes $\mathcal{D}_{2,\vec{\mathbf{w}}}$ is presented in Figure 2. The transition classes defined in (3) are required for the definition of models. We also would like to count the number of short transitions. We have the following:

Observation 2.2: $\mathcal{D}_{n,\vec{\mathbf{w}}} = \vec{\mathbf{w}}\mathcal{T}^n \backslash \{\vec{\mathbf{0}}\}$ holds, so the element count $|\mathcal{D}_{n,\vec{\mathbf{w}}}|$ is $3^{\|\vec{\mathbf{w}}\|_1} - 1$. The overall count $\sum_{\vec{\mathbf{w}}\in\mathcal{B}^n} |\mathcal{D}_{n,\vec{\mathbf{w}}}|$ is $4^n - 2^n$.

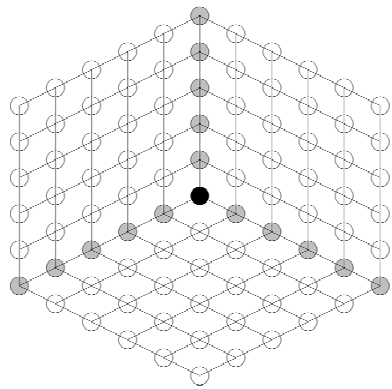

FIGURE 1. The boundary walls of $\mathcal{Z}_+^3$. The corner, the one-dimensional walls, and the two-dimensional walls are represented by the dark ball, the gray balls, and the white balls, respectively.

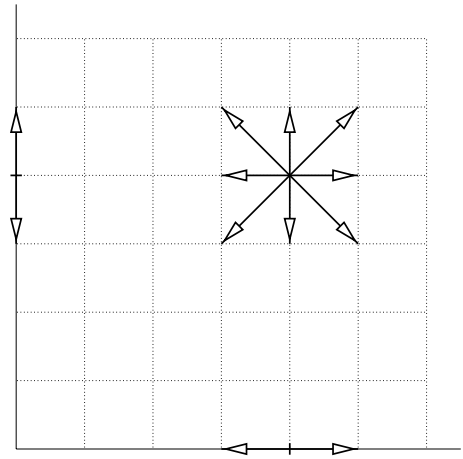

FIGURE 2. The transition classes $\mathcal{D}_{2,\langle 1,1\rangle}$, $\mathcal{D}_{2,\langle 0,1\rangle}$, and $\mathcal{D}_{2,\langle 1,0\rangle}$ ($\mathcal{D}_{2,\langle 0,0\rangle}$ is empty).

2.4. Models and Model Spaces

We will first introduce models on $\mathcal{Z}^n$, required for studying the role of product-form invariant measures, and then proceed to models on $\mathcal{Z}_+^n$. Models will be viewed both as functions on the set of pairs of states and as functions (or function arrays) on the classes of transitions introduced in Section 2.3. Let $\overline{\mathcal{S}^2}$ denote the set of pairs of different state space points, namely $\overline{\mathcal{S}^2} \triangleq \mathcal{S}^2 \backslash \{(\vec{\mathbf{a}}, \vec{\mathbf{a}}) / \vec{\mathbf{a}} \in \mathcal{S}\}$. A model, in our context, is a function $\varphi^* : \overline{\mathcal{S}^2} \mapsto \mathcal{R}_+$ satisfying the following:

1. "Communicativity": For every $(\vec{\mathbf{a}}, \vec{\mathbf{b}}) \in \overline{\mathcal{S}^2}$, there exists a finite sequence $\vec{\mathbf{a}} = \vec{\mathbf{a}}_1, \ldots, \vec{\mathbf{a}}_k = \vec{\mathbf{b}}$ of states such that $\prod_{i=1}^{k-1} \varphi^*(\vec{\mathbf{a}}_i, \vec{\mathbf{a}}_{i+1}) > 0$.
2. "Noninstantaneity": For every $\vec{\mathbf{a}} \in \mathcal{S}$, the sum $\sum_{\vec{\mathbf{b}} \in \mathcal{S} \backslash \{\vec{\mathbf{a}}\}} \varphi^*(\vec{\mathbf{a}}, \vec{\mathbf{b}})$ is finite.

A value $\varphi^*(\vec{\mathbf{a}}, \vec{\mathbf{b}})$ represents the transition rate from $\vec{\mathbf{a}}$ to $\vec{\mathbf{b}}$ of a communicative and noninstantaneous continuous-time Markov chain. Let $\mathcal{M}_n^*$ be the space of all models on $\mathcal{S} = \mathcal{Z}^n$ which possess the following homogeneity property:

$$\vec{\mathbf{b}}_1 - \vec{\mathbf{a}}_1 = \vec{\mathbf{b}}_2 - \vec{\mathbf{a}}_2 \Rightarrow \varphi^*(\vec{\mathbf{a}}_1, \vec{\mathbf{b}}_1) = \varphi^*(\vec{\mathbf{a}}_2, \vec{\mathbf{b}}_2), \qquad \vec{\mathbf{a}}_1, \vec{\mathbf{b}}_1, \vec{\mathbf{a}}_2, \vec{\mathbf{b}}_2 \in \mathcal{Z}^n. \tag{4}$$

Given some $\varphi^* \in \mathcal{M}_n^*$, the homogeneity property allows us to define the function $\varphi : \mathcal{D}_n \mapsto \mathcal{R}_+$, that corresponds to φ^* in the obvious way:

$$\varphi(\vec{\mathbf{d}}) = \varphi^*(\vec{\mathbf{0}}, \vec{\mathbf{d}}), \qquad \vec{\mathbf{d}} \in \mathcal{D}_n;$$

that is, $\varphi(\vec{\mathbf{d}})$ represents the transition rate in the direction $\vec{\mathbf{d}}$. The inverse mapping sets $\varphi^*(\vec{\mathbf{a}}, \vec{\mathbf{b}})$ to $\varphi(\vec{\mathbf{b}} - \vec{\mathbf{a}})$. Let $\mathcal{M}_n$ be the space of all such φ's which correspond to members of $\mathcal{M}_n^*$. Let the subclass $\overline{\mathcal{M}}_n$ contain those $\varphi \in \mathcal{M}_n$ whose support in $\mathcal{D}_n$ is finite, namely the models in $\mathcal{M}_n$ with bounded transitions.

Let $\mathbb{MI}_n^*$ be the space of all models on $\mathcal{S} = \mathcal{Z}_+^n$ possessing the following two properties. The first property is homogeneity, although weaker than for $\mathcal{M}_n^*$: The implication in (4) applies here only when both $\vec{\mathbf{a}}_1$ and $\vec{\mathbf{a}}_2$, or both $\vec{\mathbf{b}}_1$ and $\vec{\mathbf{b}}_2$, belong to the same wall. Thus, homogeneity is not required for parallel transitions within different walls. For two parallel transitions within the same wall, or at least with two endpoints within the same wall, homogeneity remains a requirement. The second property is permitting short transitions only:

$$\|\vec{\mathbf{b}} - \vec{\mathbf{a}}\|_\infty \neq 1 \Rightarrow \varphi^*(\vec{\mathbf{a}}, \vec{\mathbf{b}}) = 0, \qquad \vec{\mathbf{a}}, \vec{\mathbf{b}} \in \mathcal{Z}_+^n.$$

Given some $\varphi^* \in \mathbb{MI}_n^*$, define the function array

$$\varphi = \{\varphi_{\vec{\mathbf{w}}} : \mathcal{D}_{n,\vec{\mathbf{w}}} \mapsto \mathcal{R}_+\}_{\vec{\mathbf{w}} \in \mathcal{B}^n \setminus \{\vec{\mathbf{0}}\}}, \tag{5}$$

where $\varphi_{\vec{\mathbf{w}}}$ gives the transition rates within the wall $\mathcal{W}_{n,\vec{\mathbf{w}}}$, in the following way: For an arbitrary $\vec{\mathbf{w}} \in \mathcal{B}^n \setminus \{\vec{\mathbf{0}}\}$ and an arbitrary $\vec{\mathbf{d}} \in \mathcal{D}_{n,\vec{\mathbf{w}}}$, set $\varphi_{\vec{\mathbf{w}}}(\vec{\mathbf{d}})$ to be the value of any $\varphi^*(\vec{\mathbf{a}}, \vec{\mathbf{b}})$ with $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ in $\mathcal{W}_{n,\vec{\mathbf{w}}}$ and satisfying $\vec{\mathbf{b}} - \vec{\mathbf{a}} = \vec{\mathbf{d}}$. Having stated the mapping $\varphi^* \mapsto \varphi$, we shall also state the inverse mapping $\varphi \mapsto \varphi^*$; it will not be difficult to see, with the aid of two examples, that the two mappings are proper and that one is indeed the inverse of the other. For arbitrary $\vec{\mathbf{a}}, \vec{\mathbf{b}} \in \mathcal{Z}_+^n$ satisfying $\|\vec{\mathbf{b}} - \vec{\mathbf{a}}\|_\infty = 1$, with $\vec{\mathbf{a}}$ belonging, say, to $\mathcal{W}_{n,\vec{\mathbf{w}}}$ and $\vec{\mathbf{b}}$ belonging, say, to $\mathcal{W}_{n,\vec{\mathbf{v}}}$, set

$$\varphi^*(\vec{\mathbf{a}}, \vec{\mathbf{b}}) = \varphi_{\max\{\vec{\mathbf{w}}, \vec{\mathbf{v}}\}}(\vec{\mathbf{b}} - \vec{\mathbf{a}});$$

the maximum is taken componentwise. Let $\mathbb{MI}_n$ be the space of all function arrays of the type (5), which so correspond to members of $\mathbb{MI}_n^*$.

Example 2.1: Let us demonstrate the construction by computing $\varphi^*(\vec{\mathbf{a}}, \vec{\mathbf{b}})$, with $\varphi^* \in \mathbb{MI}_2^*$, $\vec{\mathbf{a}} = \langle 1,0 \rangle$, and $\vec{\mathbf{b}} = \langle 0,1 \rangle$, from the corresponding $\varphi \in \mathbb{MI}_2$. The indices of the walls $\mathcal{W}_{2,\vec{\mathbf{w}}}$ and $\mathcal{W}_{2,\vec{\mathbf{v}}}$ to which the points $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ belong happen to be $\vec{\mathbf{w}} = \vec{\mathbf{a}}$ and $\vec{\mathbf{v}} = \vec{\mathbf{b}}$. That is because $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ have been chosen adjacent to the corner. We have $\max\{\vec{\mathbf{w}}, \vec{\mathbf{v}}\} = \langle 1,1 \rangle$ and $\vec{\mathbf{b}} - \vec{\mathbf{a}} = \langle -1,1 \rangle$, so

$$\varphi^*(\vec{\mathbf{a}}, \vec{\mathbf{b}}) = \varphi_{\langle 1,1 \rangle}(\langle -1,1 \rangle). \tag{6}$$

Thus, in spite of the fact that both $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ belong to the boundary of the state space, $\varphi^*(\vec{\mathbf{a}}, \vec{\mathbf{b}})$ is computed from $\varphi_{\langle 1,1 \rangle}$, which expresses the transition rates in the interior. The reason is that $\vec{\mathbf{a}}$ and $\vec{\mathbf{b}}$ belong to distinct walls, and the transition between them can be regarded as passing through the interior. This transition is parallel, for instance, to the transition from $\langle 2,0 \rangle$ to $\langle 1,1 \rangle$, and the latter is parallel, say, to the transition from $\langle 2,1 \rangle$ to $\langle 1,2 \rangle$, whose both ends are interior points. Equation (6) is thus mandated by the (weak) homogeneity assumption.

Example 2.2: Let us remain in $\mathbb{MI}_2^*$ and in its counterpart $\mathbb{MI}_2$. The transition from $\vec{\mathbf{a}} = \langle 5,0 \rangle$ to $\vec{\mathbf{b}} = \langle 6,0 \rangle$ is parallel to the transition from $\vec{\mathbf{a}}' = \langle 5,3 \rangle$ to $\vec{\mathbf{b}}' = \langle 6,3 \rangle$. How-

ever, since the former lies within the wall $\mathcal{W}_{2,\langle 1,0\rangle}$ and the latter lies within the wall $\mathcal{W}_{2,\langle 1,1\rangle}$, homogeneity does not apply. The values $\varphi^*(\vec{\mathbf{a}},\vec{\mathbf{b}})$ and $\varphi^*(\vec{\mathbf{a}}',\vec{\mathbf{b}}')$ are not forced to be equal. These values are given by $\varphi_{\langle 1,0\rangle}(\langle 1,0\rangle)$ and $\varphi_{\langle 1,1\rangle}(\langle 1,0\rangle)$, respectively.

In view of Observation 2.2, $\mathbb{M}\mathbb{I}_n$ is essentially $\mathcal{R}_+^{4^n-2^n}$, excluding those singular elements of $\mathcal{R}_+^{4^n-2^n}$ which do not correspond to communicative chains. The members of $\boldsymbol{m}_n$ and $\mathbb{M}\mathbb{I}_n$ too, like those of $\boldsymbol{m}_n^*$ and $\mathbb{M}\mathbb{I}_n^*$, will be called models.

2.5. State Space Measures

When speaking of a measure, say μ, we always mean, unless explicitly stating otherwise, that μ is a measure (unsigned) on $(\mathcal{S},2^{\mathcal{S}})$, with $\mu(\mathcal{S})>0$. The state space $\mathcal{S}$ can be either $\mathcal{Z}^n$ or $\mathcal{Z}_+^n$. Such a measure is specified through singletons. We write $\mu(\vec{\mathbf{a}})$ as a shorthand for $\mu(\{\vec{\mathbf{a}}\})$. A measure μ is said to be of a *geometric product-form* if there exists a vector $\vec{\mathbf{q}}\in(0,\infty)^n$ satisfying

$$\mu(\vec{\mathbf{a}}) = \mu(\vec{\mathbf{0}})\cdot\vec{\mathbf{q}}^{\vec{\mathbf{a}}}, \qquad \vec{\mathbf{a}}\in\mathcal{S};$$

recall the "$\vec{\mathbf{x}}^{\vec{\mathbf{y}}}$" convention from Section 2.1. For every $\vec{\mathbf{q}}\in(0,\infty)^n$, let $\pi_{\vec{\mathbf{q}}}$ denote the corresponding geometric product-form measure with $\pi_{\vec{\mathbf{q}}}(\vec{\mathbf{0}})=1$. A measure μ is said to be *invariant* for a model φ from $\boldsymbol{m}_n$ or from $\mathbb{M}\mathbb{I}_n$ if it satisfies

$$\mu(\vec{\mathbf{a}})\sum_{\vec{\mathbf{b}}\in\mathcal{S}\setminus\{\vec{\mathbf{a}}\}}\varphi^*(\vec{\mathbf{a}},\vec{\mathbf{b}}) = \sum_{\vec{\mathbf{b}}\in\mathcal{S}\setminus\{\vec{\mathbf{a}}\}}\mu(\vec{\mathbf{b}})\varphi^*(\vec{\mathbf{b}},\vec{\mathbf{a}}), \qquad \vec{\mathbf{a}}\in\mathcal{S}, \tag{7}$$

where φ^* is the $\boldsymbol{m}_n^*$ or $\mathbb{M}\mathbb{I}_n^*$ counterpart of φ. Under the Markov chain semantics of φ^*, (7) is the steady state version of Kolmogorov's forward equation, but allowing solutions with $\mu(\mathcal{S})=\infty$. This equation is also known as the *global balance equation*. The communicativity postulate implies the following:

Observation 2.3: If μ is invariant for some model, then $\mu(\vec{\mathbf{a}})$ is positive for every $\vec{\mathbf{a}}\in\mathcal{S}$.

3. MEASURES INVARIANT FOR MODELS IN $\boldsymbol{m}_n$

This section considers the measures that are invariant for a model $\varphi\in\boldsymbol{m}_n$, namely for a homogeneous continuous-time Markov chain over $\mathcal{S}=\mathcal{Z}^n$. Such measures cannot be finite. We will show that the only possible invariant measures are product-form measures and their combinations. This fact expresses the fundamental role of product-form measures for space-homogeneous Markov chains. The proof of this result is lengthy, and is deferred to Section 5, where all of the proofs are concentrated.

First, we will characterize the set of vectors $\vec{\mathbf{q}}$ such that $\pi_{\vec{\mathbf{q}}}$ is invariant for a given model. In addition, we are interested in the subset of such $\vec{\mathbf{q}}$'s that satisfy $\vec{\mathbf{q}}<\vec{\mathbf{1}}$, since their $\pi_{\vec{\mathbf{q}}}$ may be used as the stationary state distribution of a model

on the orthant $\mathcal{Z}_+^n$ with the same interior behavior. Given a model $\varphi \in \mathcal{M}_n$, define the generating function $\tilde{\varphi}: (0,\infty)^n \mapsto \mathcal{R}_+ \cup \{\infty\}$ through

$$\tilde{\varphi}(\vec{\mathbf{s}}) \triangleq \sum_{\vec{\mathbf{d}} \in \mathcal{D}_n} \varphi(-\vec{\mathbf{d}}) \vec{\mathbf{s}}^{\vec{\mathbf{d}}}, \qquad \vec{\mathbf{s}} \in (0,\infty)^n.$$

Let Q_φ denote the set of vectors $\vec{\mathbf{q}} \in (0,\infty)^n$ such that $\pi_{\vec{\mathbf{q}}}$ is invariant for φ. By rewriting (7) in terms of φ itself, we reach the following:

Observation 3.1 (Q_φ*'s Identification):* Q_φ is the set of $\vec{\mathbf{q}}$'s solving the equation

$$\tilde{\varphi}(\vec{\mathbf{q}}) = \tilde{\varphi}(\vec{\mathbf{1}}).$$

With $\tilde{\varphi}$ already introduced, we are ready for the following.

DEFINITION 3.1: *The quantity* $\sum_{\vec{\mathbf{d}} \in \mathcal{D}_n} \varphi(\vec{\mathbf{d}}) \vec{\mathbf{d}}$*, which when convergent is equal to* $-\nabla \tilde{\varphi}(\vec{\mathbf{1}})$*, is called the drift of* φ*.*

In order to study Q_φ, let us list a few properties of $\tilde{\varphi}$. Its domain of convergence, dom $\tilde{\varphi}$, includes the point $\vec{\mathbf{s}} = \vec{\mathbf{1}}$ due to the noninstantaneity property. Being a sum of convex functions, $\tilde{\varphi}$ is convex. Moreover, it is strictly convex: By the communicativity property, for every $i = 1, \ldots, n$, there exists at least one $\vec{\mathbf{d}} \in \mathcal{D}_n$ with $d_i < 0$ such that $\varphi(\vec{\mathbf{d}}) > 0$. The factor $s_i^{d_i}$ of $\vec{\mathbf{s}}^{\vec{\mathbf{d}}}$ is strictly convex on $(0,\infty)$. Therefore, $\tilde{\varphi}$ is strictly convex on $(0,\infty)^n$, being a sum of convex functions, of which at least one comprises such a strictly convex factor, for every i. Communicativity implies also that for every $i = 1, \ldots, n$, there exists at least one $\vec{\mathbf{d}} \in \mathcal{D}_n$ with $d_i > 0$ such that $\varphi(\vec{\mathbf{d}}) > 0$. Thus, $\tilde{\varphi}$ has a unique minimum. When letting $\vec{\mathbf{s}}$ follow any straight line away from this minimum, including in a direction toward the boundary of $(0,\infty)^n$, the value of $\tilde{\varphi}(\vec{\mathbf{s}})$ goes to infinity. If dom $\tilde{\varphi}$ has a nonempty interior, which happens when the drift is convergent, then the gradient $\nabla \tilde{\varphi}$ is defined and is finite throughout this interior. When letting $\vec{\mathbf{s}}$ approach a boundary point of dom $\tilde{\varphi}$, along any path in the interior of dom $\tilde{\varphi}$, the value of $\|\nabla \tilde{\varphi}(\vec{\mathbf{s}})\|_1$ goes to infinity. All these facts lead to the following.

Observation 3.2 (Q_φ*'s Properties):* Q_φ is the boundary of a bounded and convex level set of $\tilde{\varphi}$. The point $\vec{\mathbf{1}}$ is always in Q_φ. It is the sole point iff every component of the drift of φ is either zero or nonconvergent. When $n \geq 2$ and the drift is convergent, Q_φ is an $(n-1)$-dimensional smooth manifold in $(0,\infty)^n$, and every point of Q_φ is an extreme point.

In models on $\mathcal{S} = \mathcal{Z}_+^n$, for the measure $\pi_{\vec{\mathbf{q}}}$ to be finite, the vector $\vec{\mathbf{q}}$ must belong to the unit cube $\mathcal{C} \triangleq (0,1]^n \backslash \{\vec{\mathbf{1}}\}$. The relation between Q_φ and $\mathcal{C}$ has to do with the drift of φ.

PROPOSITION 3.1 (The Relation between Q_φ and $\mathcal{C}$): *The following three cases may hold when* $n \geq 2$ *and the drift of* φ *is convergent (see Fig. 3):*

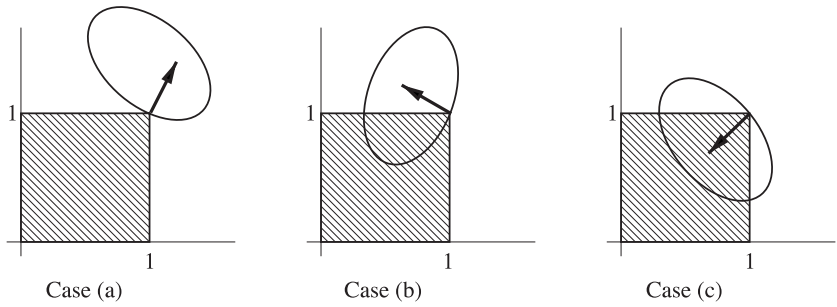

FIGURE 3. A schematic illustration for Proposition 3.1. The oval represents Q_φ, the box represents $\mathcal{C}$, and the arrow represents the drift.

Case (a): The drift is nonnegative (componentwise). In this case, $Q_\varphi \cap \mathcal{C}$ is empty.

Case (b): The drift is neither nonnegative nor negative. In this case, $Q_\varphi \cap \mathcal{C}$ is nonempty and has a zero Euclidean distance from $\vec{\mathbf{1}}$.

Case (c): The drift is negative. In this case, $Q_\varphi \cap \mathcal{C}$ is nonempty and has a positive distance from $\vec{\mathbf{1}}$.

Mixtures of product-form measures from $\{\pi_{\vec{\mathbf{q}}}\}_{\vec{\mathbf{q}} \in Q_\varphi}$ also satisfy (7), of course, and are thus invariant for φ as well. The converse statement would have said that every measure invariant for φ is either itself a member of $\{\pi_{\vec{\mathbf{q}}}\}_{\vec{\mathbf{q}} \in Q_\varphi}$ or can be represented as such a mixture. The following theorem, whose proof is fairly lengthy, states this claim for models in $\overline{\mathcal{M}}_n$; this is the subclass of models with bounded transitions introduced in Section 2.4.

THEOREM 3.1 (Representation of Measures Invariant for Models in $\overline{\mathcal{M}}_n$): *Let $\varphi \in \overline{\mathcal{M}}_n$ and let μ be a measure invariant for φ. Then, there exists a unique Borel probability measure ζ on $(0,\infty)^n$ such that*

$$[\mu(\vec{\mathbf{0}})^{-1}]\mu = \int_{Q_\varphi} \pi_{\vec{\mathbf{q}}}\, d\zeta(\vec{\mathbf{q}}).$$

4. CHARACTERIZATION OF $\mathbb{P}_n$

Recall from Section 1 that $\mathbb{P}_{n,\vec{\mathbf{q}}}$, with $\vec{\mathbf{q}} \in (0,1)^n$, is the subspace of models from $\mathbb{M}_n$ for which $\pi_{\vec{\mathbf{q}}}$ is invariant. (Note that for every such model, $\pi_{\vec{\mathbf{q}}}$ is the *unique* invariant probability measure.) The $\vec{\mathbf{q}}$'s in $\mathcal{C}\backslash(0,1)^n$, namely those with at least one component equal to 1, are avoided here. As will turn out later, their $\mathbb{P}_{n,\vec{\mathbf{q}}}$ are singular. Recall also that $\mathbb{M}_n$ and thus $\mathbb{P}_{n,\vec{\mathbf{q}}}$, can be viewed as subsets of $\mathcal{R}_+^{4^n-2^n}$. We will characterize $\mathbb{P}_{n,\vec{\mathbf{q}}}$

as the intersection between $\mathbb{MI}_n$ and the solution space of a homogeneous linear system

$$\mathbf{A}\vec{\mathbf{x}} = \vec{\mathbf{0}}, \tag{8}$$

with $\mathbf{A}$ having $4^n - 2^n$ columns. The sequel gives this characterization while concentrating on the special features of $\mathbf{A}$ which lead to the decoupling principle. The existence of nonnegative solutions, necessary for the intersection with $\mathbb{MI}_n$ to be nonempty, is addressed immediately after the characterization. First we need the following.

DEFINITION 4.1 ("Hierarchically Partitioned Matrix"): *An $m \times k$ real matrix $\mathbf{A} = (a_{i,j})$, with $m \leq k$, will be referred to as a "hierarchically partitioned matrix" if there exists a partial order "$\ll$" on $\{1,\ldots,m\}$ and a partitioning of $\{1,\ldots,k\}$ into m nonempty sets $P_1,\ldots,P_m$, such that*

1. *A component $a_{i,j}$, with $j \in P_\ell$, say, can be nonzero only if $i = \ell$ or if $i \ll \ell$.*
2. *For every $i = 1,\ldots,m$, there exists at least one $j \in P_i$ such that $a_{i,j} \neq 0$.*

The structure suggested by this definition is block triangular, up to a permutation of the columns, yet possibly with greater sparsity resulting from the order relation being partial. Observe that if the $\mathbf{A}$ of (8) is hierarchically partitioned, then the solution space of (8) admits the following recursive characterization: For every $i = 1,\ldots,m$, the portion $\langle x_j \rangle_{j \in P_i}$ of the vectors $\vec{\mathbf{x}}$ in the solution space is the hyperplane

$$\sum_{j \in P_i} a_{i,j} x_j = - \sum_{j \in \bigcup_{\{\ell / i \ll \ell\}} P_\ell} a_{i,j} x_j; \tag{9}$$

here, the entire right-hand side is regarded as a constant, adopting a point of view which defines the lower portions of $\vec{\mathbf{x}}$ in terms of the higher ones. Thus, the dimension of a portion $\langle x_j \rangle_{j \in P_i}$, conditional on all higher portions, is $|P_i| - 1$. Suppose that i_1 and i_2 are such that neither $i_1 \ll i_2$ nor $i_2 \ll i_1$ holds. Then, conditional on all portions higher than any of them, the portions P_{i_1} and P_{i_2} of $\vec{\mathbf{x}}$ are decoupled from each other. In the context of the characterization of $\mathbb{P}_{n,\vec{\mathbf{q}}}$, the last property will be referred to as the decoupling principle. We are now ready to give the characterization.

THEOREM 4.1 (Characterization of $\mathbb{P}_{n,\vec{\mathbf{q}}}$): *$\mathbb{P}_{n,\vec{\mathbf{q}}}$ is the intersection between $\mathbb{MI}_n$ and the solution space of a homogeneous linear system of the type (8), with $\mathbf{A}$ being hierarchically partitioned. The $m = 2^n - 1$ rows of the matrix $\mathbf{A}$ are indexed by $\vec{\mathbf{w}}$, $\vec{\mathbf{w}} \in \mathcal{B}^n \backslash \{\vec{\mathbf{0}}\}$, and correspond to the walls $\mathcal{W}_{n,\vec{\mathbf{w}}}$ of $\mathcal{Z}^n_+$, excluding the corner $\mathcal{W}_{n,\vec{\mathbf{0}}}$. The $k = 4^n - 2^n$ columns are indexed by $(\vec{\mathbf{v}}, \vec{\mathbf{d}})$, $\vec{\mathbf{v}} \in \mathcal{B}^n \backslash \{\vec{\mathbf{0}}\}$, $\vec{\mathbf{d}} \in \mathcal{D}_{n,\vec{\mathbf{v}}}$. The columns of $\mathbf{A}$ and the variables [i.e., the model elements $\varphi_{\vec{\mathbf{v}}}(\vec{\mathbf{d}})$] are partitioned according to the subscript $\vec{\mathbf{v}}$. The partial order among partitions is the majorization $\prec$ (recall Sect. 2.1). The matrix element $a_{\vec{\mathbf{w}},(\vec{\mathbf{v}},\vec{\mathbf{d}})}$, serving as the coefficient of $\varphi_{\vec{\mathbf{v}}}(\vec{\mathbf{d}})$ in the*

row contributed by $\mathcal{W}_{n,\vec{\mathbf{w}}}$, *with* $\vec{\mathbf{w}}$ *being equal to or majorized by* $\vec{\mathbf{v}}$, *is expressed as follows using indicators of conditions:*

$$a_{\vec{\mathbf{w}},(\vec{\mathbf{v}},\vec{\mathbf{d}})} = 1_{\{-\vec{\mathbf{d}}\le\vec{\mathbf{1}}-2(\vec{\mathbf{v}}-\vec{\mathbf{w}})\}} - \vec{\mathbf{q}}^{-\vec{\mathbf{d}}} 1_{\{\vec{\mathbf{d}}\le\vec{\mathbf{1}}-2(\vec{\mathbf{v}}-\vec{\mathbf{w}})\}}. \tag{10}$$

Having given the characterization, we now address the existence of nonnegative solutions. In a recursive representation of $\mathbb{P}_{n,\vec{\mathbf{q}}}$, of the type discussed in connection with (9), the coefficients on the left-hand side are derived from (10) with $\vec{\mathbf{w}} = \vec{\mathbf{v}}$ holding. These coefficients appear in pairs $(1 - \vec{\mathbf{q}}^{-\vec{\mathbf{d}}}, 1 - \vec{\mathbf{q}}^{\vec{\mathbf{d}}})$, due to the symmetry of the transition classes $\mathcal{D}_{n,\vec{\mathbf{w}}}$ [see (3)]. The restriction $q_i \neq 1$, $i = 1,\ldots,n$, is essential to ensure the existence of at least one such pair whose members are nonzero, for each hyperplane. (When $q_i = 1$, no such pair exists at least for the one-dimensional wall $\mathcal{W}_{n,\langle 0,\ldots,0,1,0,\ldots,0\rangle}$ with the 1 at the ith place, so there is no corresponding hyperplane. This is a singularity that would violate the characterization.) The opposite signs of the members within these pairs imply that each hyperplane, representing a portion of $\vec{\mathbf{x}}$, has an unbounded intersection with the nonnegative orthant of the Euclidean subspace to which it (the hyperplane) belongs; the explanation for this is rooted in the following elementary principle, which can be checked readily.

Lemma 4.1 (Elementary; Serves the Explanation on Nonnegative Solutions): *Let* $\sum_{i=1}^{2j} c_i x_i = d$ *be a hyperplane in* $\mathcal{R}^{2j}$ *(note that this statement implies that not all of the coefficients* c_i *are zero). Suppose that for every* $i = 1,\ldots,j$, *the coefficients* c_i *and* c_{i+j} *are either both zero or both nonzero and have opposite signs. Then, the hyperplane has a nonempty intersection with* $\mathcal{R}_+^{2j}$. *Moreover, this intersection is not confined to any box of the type* $J_1 \times \cdots \times J_{2j}$, *where the* $J_1,\ldots,J_{2j}$ *are intervals in* $\mathcal{R}_+$ *of which at least one is finite.*

The linear system taking part in Theorem 4.1 is essentially an adaptation of the global balance equation (7) for the $\mathbb{M}|_n$ framework, with the transition rates, rather than the measure, playing the role of the unknown. However, a key feature of the special structure that emerges is that although each of the walls $\mathcal{W}_{n,\vec{\mathbf{w}}}$ with $\vec{\mathbf{w}} \in \mathcal{B}^n \backslash \{\vec{\mathbf{0}}\}$ contributes one equation (i.e., one row to the matrix $\mathbf{A}$), no equation is contributed by the corner $\mathcal{W}_{n,\vec{\mathbf{0}}}$. The absence of such an equation is what enables the matrix to comply with the requirements of Definition 4.1. That is because the row contributed by $\mathcal{W}_{n,\vec{\mathbf{0}}}$ could not have been associated with any nonempty partition of variables, since $\mathcal{D}_{n,\vec{\mathbf{0}}}$ is empty. Yet, the corner does induce a valid equation, which is eliminated due to redundancy. It is this redundancy which is responsible for the richness of $\mathbb{P}_n$ via the decoupling principle. The redundancy of the corner equation could have been deduced from the following elementary fact: The global balance equation, regardless of any special structure, always has a single redundancy when the chain is uniformizable; that is, when

$$\sup_{\vec{\mathbf{a}}\in\mathcal{S}} \sum_{\vec{\mathbf{b}}\in\mathcal{S}\backslash\{\vec{\mathbf{a}}\}} \varphi^*(\vec{\mathbf{a}},\vec{\mathbf{b}}) < \infty,$$

and when the invariant measure is known to be finite. However, this kind of deduction would not have captured the deeper essence of the redundancy here. In Section 5, we bring a fairly intricate proof for Theorem 4.1, which does rely on the special structure of $\mathbb{MI}_n$ and on the product-form property and, on the other hand, does not rely on the finiteness of the measure. As discussed at the end of the proof, this intricacy is essential for inferring

Remark 4.1 (State Spaces with Multiple Corners): The decoupling principle can be generalized for state spaces of the form $\mathcal{S} = Z_1 \times Z_2 \times \cdots \times Z_n$, where every Z_i is a finite or an infinite succession of integers. Such state spaces arise, for example, in connection with stochastic networks with finite buffers [11]. The basis for this generalization is discussed at the end of the proof of Theorem 4.1 in Section 5.

Another type of generalization is provided by the following.

Remark 4.2 (Two-Dimensional State Spaces with Slanted Walls): State spaces with slanted walls are easy to describe in the two-dimensional case. See Figure 4. The decoupling principle can be generalized for the two non-right-corner types shown in the figure. The proof is at the same time a generalization and a specialization of the proof of Theorem 4.1 and is omitted.

In view of the last two remarks, what is exposed by the proof of Theorem 4.1 brought in Section 5 is the following: The decoupling principle is an inherent property of product-form model spaces with homogeneity.

5. PROOFS

5.1. Proof of Proposition 3.1

Exclude the case $-\nabla\tilde{\varphi}(\vec{\mathbf{1}}) = \vec{\mathbf{0}}$, which has been covered earlier in Observation 3.2. The essence of the proposition is expressed in Lemma 5.1. The lemma is elementary and is given without a proof. The association between the proposition and the lemma

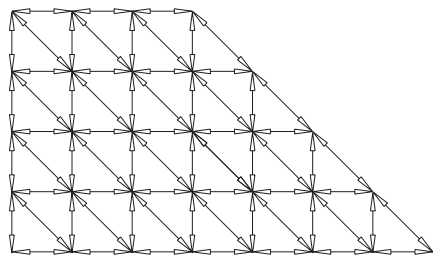

FIGURE 4. A two-dimensional state space with a slanted wall, creating a blunt corner and a sharp corner.

is drawn after stating the lemma. The definitions of a *cone* and of a *supporting half-space* are available, for example, in Rockafellar [23, pp. 13 and 99].

LEMMA 5.1: *Let $A \subset \mathcal{R}^k$ be nonempty, bounded in $\|\cdot\|_1$, convex, containing the point $\vec{\mathbf{0}}$ in its boundary ∂A, and having a unique supporting half-space H with $\vec{\mathbf{0}} \in \partial H$. Also, let $K \subset \mathcal{R}^k$ be a nonempty and convex cone not containing $\vec{\mathbf{0}}$. The following three cases may hold (see a suggestive illustration in Fig. 5):*

Case (a): $H \cap K$ *is empty. In this case,* $\partial A \cap K$ *is empty.*

Case (b): $H \cap K$ *is nonempty, but* K *is not contained in the interior of* H*. In this case,* $K \cap \partial A$ *is nonempty and has a zero Euclidean distance from* $\vec{\mathbf{0}}$*.*

Case (c): K *is contained in the interior of* H*. In this case,* $K \cap \partial A$ *is nonempty and has a positive distance from* $\vec{\mathbf{0}}$*.*

Proposition 3.1 reduces into Lemma 5.1 via the following association. The role of ∂A is played by $Q_\varphi - \vec{\mathbf{1}}$ (defined as $\{\vec{\mathbf{q}} - \vec{\mathbf{1}} / \vec{\mathbf{q}} \in Q_\varphi\}$). The role of K is played by the nonpositive orthant, excluding $\vec{\mathbf{0}}$. This cone can be viewed as an extension of $\mathcal{C} - \vec{\mathbf{1}}$. In fact, the extension does not have any influence, since Q_φ is confined to $(0,\infty)^n$. The set Q_φ has a unique supporting half-space H' with $\vec{\mathbf{1}} \in \partial H'$, due to Observation 3.2. The role of H is played by $H' - \vec{\mathbf{1}}$. It is known that

$$H' = \{\vec{\mathbf{x}} \in \mathcal{R}^n / (\vec{\mathbf{x}} - \vec{\mathbf{1}}) \cdot \nabla \tilde{\varphi}(\vec{\mathbf{1}}) \leq 0\}.$$

Hence the translation of the conditions on the drift into conditions on $H \cap K$.

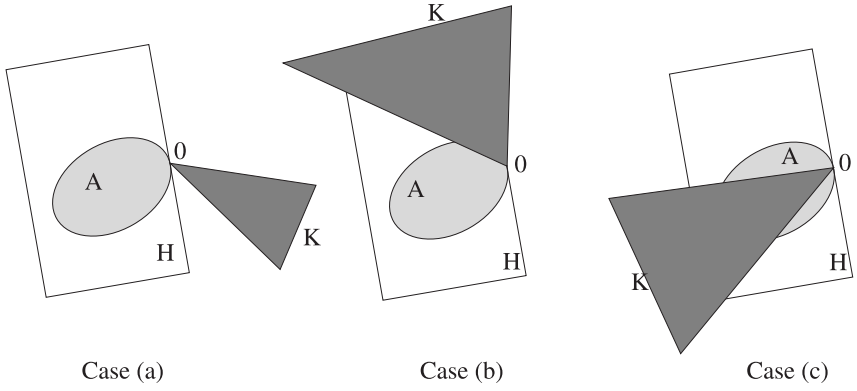

FIGURE 5. A suggestive illustration for Lemma 5.1.

5.2. Proof of Theorem 3.1

The uniqueness of ζ is evident from the nature of the $\pi_{\vec{\mathbf{q}}}$: Different combinations of product-form measures cannot give identical aggregated measures. The proof of its existence calls for applying Choquet's theorem. This is proved to be possible, with the theorem applied for a normed space of measures on $(\mathcal{Z}^n, 2^{\mathcal{Z}^n})$. The norm is ℓ_1, augmented by geometrical weights. The course of the proof is as follows. First, a suitable version of Choquet's theorem is formulated. Then, the major part of the proof is dedicated to establishing the setting for applying the theorem. Finally, Choquet's theorem is invoked and the conclusion is adjusted to fit the original setting of Theorem 3.1.

THEOREM 5.1 (Choquet's Theorem, Adapted for a Separable Normed Space): *Let $(Y, \|\ \|)$ be a separable normed vector space. Let $K \subset Y$ be compact and convex. Then, for every $x \in K$, there exists a Borel probability measure γ, concentrated on the set of extreme points $E \triangleq \operatorname{ext} K$, such that $x = \int_E y\, d\gamma(y)$.*

In a more general formulation of the theorem, Y is an abstract topological vector space in which the dual space Y^* separates points, and γ is supported on the closure of E. See, for example, Rudin [25, p. 85, Exercise 25]. However, when Y is separable and metric, such γ exists on E itself [25, p. 376, Solution of Exercise 25]. Also, when Y is a normed space, then it is locally convex. In such a case, it is guaranteed that Y^* separates points [25, p. 59, Corollary].

We now set out to establish the setting for the application of Theorem 5.1. Let Υ denote the vector space of *signed* measures on $(\mathcal{Z}^n, 2^{\mathcal{Z}^n})$. For every $p > 0$, let the function $\eta_p : \Upsilon \mapsto \mathcal{R}_+ \cup \{\infty\}$ be defined through

$$\eta_p(\nu) \triangleq \sum_{\vec{\mathbf{a}} \in \mathcal{Z}^n} p^{-\|\vec{\mathbf{a}}\|_1} |\nu(\vec{\mathbf{a}})|, \qquad \nu \in \Upsilon,$$

and let

$$\Upsilon_p \triangleq \{\nu \in \Upsilon / \eta_p(\nu) < \infty\}.$$

We shall employ normed vector spaces of the type (Υ_p, η_p). Observe that they are separable. We shall now introduce a further type of subsets of Υ, with members that are normalized in some sense, and have "bounded growth." For every $r \geq 1$, let $\Delta_r \subset \Upsilon$ contain those ν which satisfy the following two conditions:

1. $\nu(\vec{\mathbf{0}}) = 1$.
2. $\|\vec{\mathbf{a}} - \vec{\mathbf{b}}\|_1 = 1 \Rightarrow r^{-1} \leq \nu(\vec{\mathbf{a}})/\nu(\vec{\mathbf{b}}) \leq r$.

LEMMA 5.2: *If $p > r$, then Δ_r is compact in (Υ_p, η_p).*

PROOF: First, observe that $p > r \Rightarrow \Delta_r \subset \Upsilon_p$. The compactness claim can be reduced into a sequential compactness claim, due to the Borel–Lebesgue theorem (see, e.g., Royden [24, p. 155]). Moreover, the sequential compactness claim can

be further reduced to the following: Every sequence in Δ_r contains a subsequence which converges in the pointwise sense (i.e., at every point of $\mathcal{Z}^n$). The pointwise convergence would imply convergence in the norm η_p, as can be inferred using the inequality

$$\eta_p(\nu_1 - \nu_2) \leq \sum_{\{\vec{\mathbf{a}} \in \mathcal{Z}^n / \|\vec{\mathbf{a}}\|_1 \leq y\}} p^{-\|\vec{\mathbf{a}}\|_1} |\nu_1(\vec{\mathbf{a}}) - \nu_2(\vec{\mathbf{a}})| + 2 \sum_{\{\vec{\mathbf{a}} \in \mathcal{Z}^n / \|\vec{\mathbf{a}}\|_1 > y\}} \left(\frac{r}{p}\right)^{\|\vec{\mathbf{a}}\|_1}, \qquad \nu_1, \nu_2 \in \Delta_r, y \in \mathcal{R}_+,$$

plus standard "ε-arguments." Indeed, the requested subsequence that converges in the pointwise sense can be extracted by diagonalization; see, for example, Chung [10, p. 84]. ■

We now turn our attention to measures which are invariant for φ. Let $\mathfrak{I}_\varphi$ denote the set of state space measures ν (of the type of Sect. 2.5, not signed measures) which are invariant for φ and satisfy $\nu(\vec{\mathbf{0}}) = 1$. It can be seen directly that $\mathfrak{I}_\varphi$ is convex. We would like to express the invariance for φ in terms of operators on Υ. For every $\vec{\mathbf{d}} \in \mathcal{D}_n$, define the *shift operator* $\mathbb{S}_{\vec{\mathbf{d}}} : \Upsilon \mapsto \Upsilon$ through

$$(\mathbb{S}_{\vec{\mathbf{d}}} \nu)(\vec{\mathbf{a}}) = \nu(\vec{\mathbf{a}} - \vec{\mathbf{d}}), \qquad \nu \in \Upsilon, \vec{\mathbf{a}} \in \mathcal{Z}^n.$$

By considering the global balance equation (7), as rewritten in terms of φ itself and divided by $\sum_{\vec{\mathbf{d}} \in \mathcal{D}_n} \varphi(\vec{\mathbf{d}})$, we arrive at the following.

Observation 5.1:

(a) The invariance of the elements of $\mathfrak{I}_\varphi$ for φ is tantamount to an invariance for an operator of the form $\sum_{\vec{\mathbf{d}} \in \mathcal{D}_n} c_{\vec{\mathbf{d}}} \mathbb{S}_{\vec{\mathbf{d}}}$, with the coefficients $c_{\vec{\mathbf{d}}}$ taking values in $[0,1)$; the sum of these coefficients is 1, and since $\varphi \in \overline{\mathcal{M}}_n$, only finitely many of them are nonzero.

(b) Moreover, the elements of $\mathfrak{I}_\varphi$ are invariant to all powers of the operator $\sum_{\vec{\mathbf{d}} \in \mathcal{D}_n} c_{\vec{\mathbf{d}}} \mathbb{S}_{\vec{\mathbf{d}}}$ mentioned in part a. These powers have the same form as the original operator, and the communicativity implies that for every $\vec{\mathbf{d}} \in \mathcal{D}_n$, there exists a power with a positive $c_{\vec{\mathbf{d}}}$.

This leads to the following key fact.

LEMMA 5.3: *$\mathfrak{I}_\varphi$ is contained in some Δ_r.*

PROOF: For every $\vec{\mathbf{d}} \in \mathcal{D}_n$, single out an operator of the type indicated in part b of Observation 5.1, one for which the coefficient $c_{\vec{\mathbf{d}}}$ is positive. Note that for every $\nu \in \mathfrak{I}_\varphi$, there holds the inequality $\nu \geq c_{\vec{\mathbf{d}}} \mathbb{S}_{\vec{\mathbf{d}}} \nu$. Now, it is not difficult to see that the claim holds true for every

$$r \geq \max_{\{\vec{\mathbf{d}} \in \mathcal{D}_n / \|\vec{\mathbf{d}}\|_1 = 1\}} c_{\vec{\mathbf{d}}}^{-1}.$$

■

LEMMA 5.4: *$\Im_\varphi$ is compact in some* (Υ_p, η_p).

PROOF: Choose some r such that $\Im_\varphi \subset \Delta_r$ (see Lemma 5.3) and some $p > r$. In view of Lemma 5.2, the compactness of $\Im_\varphi$ in (Υ_p, η_p) will be established if we verify that $\Im_\varphi$ is closed in (Υ_p, η_p). The following is to be verified: Let all the elements of a sequence $\{\nu_i\}_{i=1}^{\infty} \subset \Upsilon_p$ satisfy $\nu_i(\vec{\mathbf{0}}) = 1$ and be invariant to an operator $\sum_{\vec{\mathbf{d}} \in \mathcal{D}_n} c_{\vec{\mathbf{d}}} \mathcal{S}_{\vec{\mathbf{d}}}$ of the type of Observation 5.1. Suppose that the sequence converges in η_p to some $\nu \in \Upsilon_p$ [i.e., $\eta_p(\nu - \nu_i) \to 0$]. Then, ν must also satisfy $\nu(\vec{\mathbf{0}}) = 1$ and be invariant to the same operator—so far the target. The fulfillment of the requirement $\nu(\vec{\mathbf{0}}) = 1$ follows from the fact that convergence in η_p obviously implies pointwise convergence. To verify the invariance, we check that $\eta_p(\nu - \sum_{\vec{\mathbf{d}} \in \mathcal{D}_n} c_{\vec{\mathbf{d}}} \mathcal{S}_{\vec{\mathbf{d}}} \nu) = 0$. That is accomplished by applying η_p, and then $\limsup_{i \to \infty}$, on

$$\nu - \sum_{\vec{\mathbf{d}} \in \mathcal{D}_n} c_{\vec{\mathbf{d}}} \mathcal{S}_{\vec{\mathbf{d}}} \nu = (\nu - \nu_i) - \sum_{\vec{\mathbf{d}} \in \mathcal{D}_n} c_{\vec{\mathbf{d}}} \mathcal{S}_{\vec{\mathbf{d}}} (\nu - \nu_i), \qquad i = 1, 2 \ldots,$$

while using

$$\eta_p\big(\mathcal{S}_{\vec{\mathbf{d}}}(\nu - \nu_i)\big) \le (p\vec{\mathbf{1}})^{|\vec{\mathbf{d}}|} \eta_p(\nu - \nu_i). \qquad \blacksquare$$

Consider now the extreme points of $\Im_\varphi$.

LEMMA 5.5: *$\operatorname{ext} \Im_\varphi \subset \{\pi_{\vec{\mathbf{q}}}\}_{\vec{\mathbf{q}} \in \mathcal{Q}_\varphi}$ holds.*

PROOF: Pick some $\nu \in \operatorname{ext} \Im_\varphi$. The claim that ν has a geometric product-form will follow by verifying that $\mathcal{S}_{\vec{\mathbf{d}}} \nu$ and ν are equal, up to a multiplicative factor, for every $\vec{\mathbf{d}} \in \mathcal{D}_n$. Observation 5.1 implies that there exists a coefficient $c_{\vec{\mathbf{d}}} > 0$ and a nonnegative $\nu' \in \Upsilon$ such that

$$\nu = c_{\vec{\mathbf{d}}} \mathcal{S}_{\vec{\mathbf{d}}} \nu + \nu'. \tag{11}$$

However, $\mathcal{S}_{\vec{\mathbf{d}}} \nu$ too is invariant for φ: Apply on it the operator to which ν is invariant and use the commutativity of the shifts. Therefore, ν', by being a difference, is invariant for φ as well. This implies that the right-hand side of (11) can be rendered, through appropriate renormalization, as a convex combination of two elements of $\Im_\varphi$. However, both of them, including the one proportional to $\mathcal{S}_{\vec{\mathbf{d}}} \nu$, in which we are interested, must be equal to ν by the hypothesis that $\nu \in \operatorname{ext} \Im_\varphi$. $\blacksquare$

For every p, let σ_p denote the Borel σ-algebra on Υ_p induced by η_p. We are ready to invoke Theorem 5.1. From the theorem, we draw the following conclusion: There exists a probability measure ζ' on some (Υ_p, σ_p), such that

$$[\mu(\vec{\mathbf{0}})^{-1}]\mu = \int_{\{\pi_{\vec{\mathbf{q}}}\}_{\vec{\mathbf{q}} \in \mathcal{Q}_\varphi}} \pi_{\vec{\mathbf{q}}} \, d\zeta'(\pi_{\vec{\mathbf{q}}});$$

first take the integral over $\operatorname{ext} \Im_\varphi$, but then observe that $\operatorname{ext} \Im_\varphi$ cannot be smaller than $\{\pi_{\vec{\mathbf{q}}}\}_{\vec{\mathbf{q}} \in \mathcal{Q}_\varphi}$. This conclusion needs a slight adjustment to the original setting of Theo-

rem 3.1, where the integration is performed on Q_φ itself. The validity of this adjustment would follow by the measurability from the Borel σ-algebra on $(0,\infty)^n$ to the restriction of σ_p to $\{\pi_{\vec{\mathbf{q}}}\}_{\vec{\mathbf{q}}\in Q_\varphi}$ of the mapping $\vec{\mathbf{q}} \mapsto \pi_{\vec{\mathbf{q}}}$. From the definition of η_p, it is clear that this measurability holds.

5.3. Proof of Theorem 4.1

This proof should comprise two ingredients:

1. Verification that the matrix whose elements are defined in the theorem indeed satisfies the requirements of Definition 4.1.
2. Verification that $\mathbb{P}_{n,\vec{\mathbf{q}}}$ is, indeed, the intersection between $\mathbb{M}\mathbb{I}_n$ and the solution space of (8), with the matrix defined in the theorem.

The first ingredient is addressed in the very formulation of the theorem and in the ensuing discussion about nonnegative solutions. The second ingredient will be fulfilled by first showing the validity for a modified matrix, consisting of the declared $\mathbf{A}$ plus an additional row, and then verifying that the additional row is, in fact, redundant.

The global balance equation (7), with a fixed μ, becomes an equation in φ^*. The space $\mathbb{P}_{n,\vec{\mathbf{q}}}$ is the set of all $\varphi \in \mathbb{M}\mathbb{I}_n$ whose corresponding φ^* satisfy (7) with $\mu = \pi_{\vec{\mathbf{q}}}$. This equation system (we now consider each contribution by some $\vec{\mathbf{a}} \in \mathcal{S}$ as one equation) should be rewritten in terms of φ itself. Due to the space homogeneity, the collection of distinct equations corresponds to the collection of walls. In order to write down these equations in φ, some further state transition classes should be introduced. Let

$$\mathcal{D}_{n,\vec{\mathbf{w}}|\vec{\mathbf{v}}} \triangleq \{\vec{\mathbf{b}} - \vec{\mathbf{a}} / \vec{\mathbf{a}} \in \mathcal{W}_{n,\vec{\mathbf{v}}}, \vec{\mathbf{b}} \in \mathcal{W}_{n,\vec{\mathbf{w}}}, \|\vec{\mathbf{b}} - \vec{\mathbf{a}}\|_\infty = 1\}, \qquad \vec{\mathbf{w}}, \vec{\mathbf{v}} \in \mathcal{B}^n$$

[compare with (3)]. See an illustration of the classes $\mathcal{D}_{2,\vec{\mathbf{w}}|\vec{\mathbf{1}}}$ in Figure 6. Observe that the overall set of short transitions into any state of $\mathcal{W}_{n,\vec{\mathbf{w}}}$ is

$$\bigcup_{\{\vec{\mathbf{v}}\in\mathcal{B}^n / \vec{\mathbf{v}} \geq \vec{\mathbf{w}}\}} \mathcal{D}_{n,\vec{\mathbf{w}}|\vec{\mathbf{v}}}.$$

Moreover, the above union is disjoint. Observe also that the following characterization holds:

$$\vec{\mathbf{v}} \geq \vec{\mathbf{w}} \Rightarrow \mathcal{D}_{n,\vec{\mathbf{w}}|\vec{\mathbf{v}}} = \{\vec{\mathbf{d}} \in \mathcal{D}_{n,\vec{\mathbf{v}}} / \vec{\mathbf{d}} \leq \vec{\mathbf{1}} - 2(\vec{\mathbf{v}} - \vec{\mathbf{w}})\}; \tag{12}$$

this characterization succinctly says that when $\vec{\mathbf{v}} \geq \vec{\mathbf{w}}$ and $\vec{\mathbf{d}} \in \mathcal{D}_{n,\vec{\mathbf{w}}|\vec{\mathbf{v}}}$, then for every $i = 1,\ldots,n$ where $v_i = w_i = 1$, the value of d_i can be any value in $\mathcal{T}$, for every i where $v_i = 1$ and $w_i = 0$, the value of d_i must be -1, and for every i where $v_i = w_i = 0$, the value of d_i must be 0 (by the very fact that $\vec{\mathbf{d}}$ is also an element of $\mathcal{D}_{n,\vec{\mathbf{v}}}$). The rewriting of (7) in terms of the function array (5) is based on the correspondence between φ^* and φ, as defined in Section 2.4. In the process, both sides of (7) are divided by $\vec{\mathbf{q}}^{\vec{\mathbf{a}}}$, and terms from both sides are collected according to state transitions. The following equation system results:

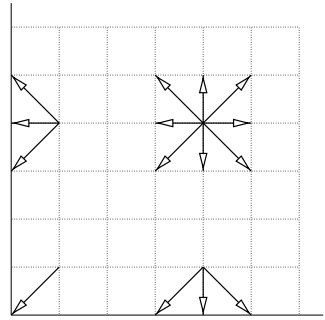

FIGURE 6. The transition classes $\mathcal{D}_{2,\langle 1,1\rangle|\langle 1,1\rangle}$ $(=\mathcal{D}_{2,\langle 1,1\rangle})$, $\mathcal{D}_{2,\langle 1,0\rangle|\langle 1,1\rangle}$, $\mathcal{D}_{2,\langle 0,1\rangle|\langle 1,1\rangle}$, and $\mathcal{D}_{2,\langle 0,0\rangle|\langle 1,1\rangle}$ (compare with Fig. 2).

$$\sum_{\vec{\mathbf{v}}\geq\vec{\mathbf{w}}}\ \sum_{\vec{\mathbf{d}}\in\mathcal{D}_{n,\vec{\mathbf{w}}|\vec{\mathbf{v}}}}[\varphi_{\vec{\mathbf{v}}}(-\vec{\mathbf{d}})-\vec{\mathbf{q}}^{-\vec{\mathbf{d}}}\varphi_{\vec{\mathbf{v}}}(\vec{\mathbf{d}})]=0,\qquad \vec{\mathbf{w}}\in\mathcal{B}^n. \tag{13}$$

In view of (12), the equation system (13) almost matches the description in the theorem. The only remaining discrepancy is in the presence of an equation for $\vec{\mathbf{w}}=\vec{\mathbf{0}}$ in (13). (Note that the equation with $\vec{\mathbf{w}}=\vec{\mathbf{0}}$ is proper: The undefined $\varphi_{\vec{\mathbf{0}}}$ does not actually appear, since $\mathcal{D}_{n,\vec{\mathbf{0}}|\vec{\mathbf{0}}}$ is empty.) The redundancy of this equation is to be verified. To this end, we explicitly give real numbers $\{g_{\vec{\mathbf{w}}}\}_{\vec{\mathbf{w}}\in\mathcal{B}^n}$, with $g_{\vec{\mathbf{0}}}=1$, such that the weighted sum of the equations in (13), with these numbers serving as the weights, is zero. The numbers we use are $g_{\vec{\mathbf{w}}}=\vec{\mathbf{r}}^{-\vec{\mathbf{w}}}$, with $\vec{\mathbf{r}}=\langle r_1,\ldots,r_n\rangle$ given through

$$r_i=q_i^{-1}-1,\qquad i=1,\ldots,n. \tag{14}$$

Let $h_{\vec{\mathbf{w}},\vec{\mathbf{v}},\vec{\mathbf{d}}}$ denote the coefficient belonging to the variable $\varphi_{\vec{\mathbf{v}}}(\vec{\mathbf{d}})$ in the equation contributed by $\vec{\mathbf{w}}$. From (12) and (13), we have

$$h_{\vec{\mathbf{w}},\vec{\mathbf{v}},\vec{\mathbf{d}}}=\begin{cases}1_{\{-\vec{\mathbf{d}}\leq\vec{\mathbf{1}}-2(\vec{\mathbf{v}}-\vec{\mathbf{w}})\}}-\vec{\mathbf{q}}^{-\vec{\mathbf{d}}}1_{\{\vec{\mathbf{d}}\leq\vec{\mathbf{1}}-2(\vec{\mathbf{v}}-\vec{\mathbf{w}})\}} & \text{if } \vec{\mathbf{w}}\leq\vec{\mathbf{v}},\\ 0 & \text{otherwise,}\end{cases}\qquad \vec{\mathbf{w}}\in\mathcal{B}^n,\ \vec{\mathbf{v}}\in\mathcal{B}^n\backslash\{\vec{\mathbf{0}}\},\ \vec{\mathbf{d}}\in\vec{\mathbf{v}}\mathcal{T}^n\backslash\{\vec{\mathbf{0}}\}.$$

Our target is to verify that for each $\varphi_{\vec{\mathbf{v}}}(\vec{\mathbf{d}})$, the weighted sum of coefficients is zero; namely we have to verify that

$$\sum_{\vec{\mathbf{w}}\in\mathcal{B}^n}g_{\vec{\mathbf{w}}}\,h_{\vec{\mathbf{w}},\vec{\mathbf{v}},\vec{\mathbf{d}}}=0,\qquad \vec{\mathbf{v}}\in\mathcal{B}^n\backslash\{\vec{\mathbf{0}}\},\ \vec{\mathbf{d}}\in\vec{\mathbf{v}}\mathcal{T}^n\backslash\{\vec{\mathbf{0}}\}.$$

The above target equation is converted, by substitution of the values and a slight manipulation, into

$$\vec{\mathbf{q}}^{-\vec{\mathbf{d}}}\sum_{\{\vec{\mathbf{w}}\in\mathcal{B}^n/\vec{\mathbf{v}}-(1/2)(\vec{\mathbf{1}}-\vec{\mathbf{d}})\leq\vec{\mathbf{w}}\leq\vec{\mathbf{v}}\}}\vec{\mathbf{r}}^{-\vec{\mathbf{w}}}=\sum_{\{\vec{\mathbf{w}}\in\mathcal{B}^n/\vec{\mathbf{v}}-(1/2)(\vec{\mathbf{1}}+\vec{\mathbf{d}})\leq\vec{\mathbf{w}}\leq\vec{\mathbf{v}}\}}\vec{\mathbf{r}}^{-\vec{\mathbf{w}}},$$
$$\vec{\mathbf{v}}\in\mathcal{B}^n\backslash\{\vec{\mathbf{0}}\},\ \vec{\mathbf{d}}\in\vec{\mathbf{v}}\mathcal{T}^n\backslash\{\vec{\mathbf{0}}\}.$$

Fix some $\vec{\mathbf{v}} = \langle v_1, \ldots, v_n \rangle \in \mathcal{B}^n \backslash \{\vec{\mathbf{0}}\}$, and some $\vec{\mathbf{d}} = \langle d_1, \ldots, d_n \rangle \in \vec{\mathbf{v}}\mathcal{T}^n \backslash \{\vec{\mathbf{0}}\}$. Suppose that $v_i = 0$ holds for some $i \in \{1, \ldots, n\}$. Then, d_i must also be zero. Likewise, w_i must be zero for every $\vec{\mathbf{w}} = \langle w_1, \ldots, w_n \rangle$ participating in any of the two summations. A coordinate i with $v_i = 0$ can thus be ignored. Hence, no generality will be lost if we focus on $\vec{\mathbf{v}} = \vec{\mathbf{1}}$. The target now reduces into verifying that

$$\vec{\mathbf{q}}^{-\vec{\mathbf{d}}} \sum_{\{\vec{\mathbf{w}} \in \mathcal{B}^n / \vec{\mathbf{w}} \geq (1/2)(\vec{\mathbf{1}} + \vec{\mathbf{d}})\}} \vec{\mathbf{r}}^{-\vec{\mathbf{w}}} = \sum_{\{\vec{\mathbf{w}} \in \mathcal{B}^n / \vec{\mathbf{w}} \geq (1/2)(\vec{\mathbf{1}} - \vec{\mathbf{d}})\}} \vec{\mathbf{r}}^{-\vec{\mathbf{w}}}, \qquad \vec{\mathbf{d}} \in \mathcal{T}^n \backslash \{\vec{\mathbf{0}}\}. \tag{15}$$

Fix again an arbitrary $\vec{\mathbf{d}} = \langle d_1, \ldots, d_n \rangle$, this time from $\mathcal{T}^n \backslash \{\vec{\mathbf{0}}\}$. Designate the index sets

$$I_t \triangleq \{i = 1, \ldots, n / d_i = t\}, \qquad t \in \mathcal{T}.$$

Adopt the following convention: Given a vector $\vec{\mathbf{x}} = \langle x_1, \ldots, x_k \rangle$ and a partial index set $I \subset \{1, \ldots, k\}$, let $\vec{\mathbf{x}}_I$ be the vector of dimension $|I|$ obtained by restriction. A member $\vec{\mathbf{w}}$ of the summation set at the left-hand side of (15) admits the following form: $\vec{\mathbf{w}}_{I_{-1}}$ can take any value in $\mathcal{B}^{|I_{-1}|}$, and $\vec{\mathbf{w}}_{I_0 \cup I_1}$ must be $\vec{\mathbf{1}}$. Similarly, the form of a member $\vec{\mathbf{w}}$ of the summation set at the right-hand side is as follows: $\vec{\mathbf{w}}_{I_{-1} \cup I_0}$ must be $\vec{\mathbf{1}}$, and $\vec{\mathbf{w}}_{I_1}$ can take any value in $\mathcal{B}^{|I_1|}$. By decomposing all the vectors involved in (15) into their I_{-1}, I_0, and I_1 parts and performing a slight rearrangement, (15) further reduces to

$$\vec{\mathbf{q}}_{I_{-1}}^{\vec{\mathbf{1}}} \sum_{\vec{\mathbf{w}} \in \mathcal{B}^{|I_{-1}|}} \vec{\mathbf{r}}_{I_{-1}}^{\vec{\mathbf{1}} - \vec{\mathbf{w}}} = \vec{\mathbf{q}}_{I_1}^{\vec{\mathbf{1}}} \sum_{\vec{\mathbf{w}} \in \mathcal{B}^{|I_1|}} \vec{\mathbf{r}}_{I_1}^{\vec{\mathbf{1}} - \vec{\mathbf{w}}}.$$

The last target equation indeed holds true, as both sides are equal to 1. Recall (14), and apply the following identity, whose verification by induction on k is immediate:

$$\sum_{\vec{\mathbf{w}} \in \mathcal{B}^k} \vec{\mathbf{y}}^{\vec{\mathbf{w}}} = (\vec{\mathbf{y}} + \vec{\mathbf{1}})^{\vec{\mathbf{1}}}, \qquad \vec{\mathbf{y}} \in \mathcal{R}^k.$$

The case where I_{-1} or I_1 are empty requires some attention, but does not require separate treatment if the following convention is adhered to: Let $\mathcal{B}^0$ contain a single element—the "zero-dimension vector." When raised to the power of itself, this "vector" gives 1—the conventional value of an empty product. The proof is complete.

The above proof stays valid if $q_i > 1$ for some or all of the coordinates i (although cases with $q_i = 1$ are singularities); that is, the proof does not rely on the finiteness of the product-form measure. This provides the basis for Remark 4.1: When there are multiple corners, the model with its product-form measure, as viewed from different corners, is isomorphic to a model with a product-form measure on the orthant, but with a measure that is not necessarily finite; the masses seem to grow in some coordinates when looking from corners other than the original one. The need not to rely on the finiteness of the measure means that the proof could not have been inferred from the elementary fact concerning the single redundancy of the global balance equation of a general Markov chain; rather, the proof must have relied on the special properties of the models studied.

6. A BASIC EXAMPLE OF MODEL DESIGN

The purpose of this section is to give a taste of model design, using Theorem 4.1 and the decoupling principle. For simplicity, we give an example of dimension $n = 2$, although these tools are handy for any n. Consider the service system in Figure 7. The system is defined by the following parameters (the index i takes the values 1 and 2):

λ_i: Rate of external arrival to station i

θ_i: Service rate at station i

$\rho_i^{\text{depart}}, \rho_i^{\text{transfer}}, \rho_i^{\text{branch}}$: Probabilities of the three types of events that may occur upon completing a service at station i (see description in the caption).

All of the service durations are assumed to be exponentially distributed and independent from each other and from previous eventualities. Whether the customer departs, is transferred to the other station, or branches also does not depend on any previous observable eventuality. The design problem is as follows: Minimize the

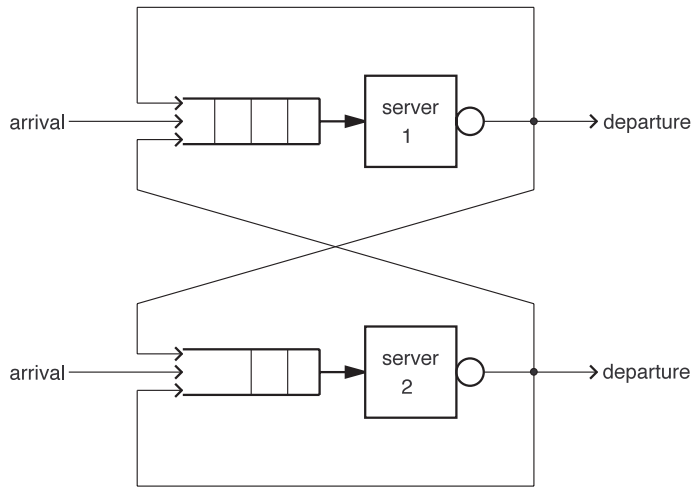

FIGURE 7. A sample service system. Upon completing a service in one of the two service stations, the customer can depart from the system, be transferred to the other station, or branch into three descendants. In the latter case, two of the descendants return to the queue of the same station and the third goes to the other station. The branching phenomenon is symbolized by the circles at the outputs. Note that when branching occurs, the number of customers at each of the two stations is incremented by 1.

total service effort $\theta_1 + \theta_2$ while complying with a performance requirement of the type

$$\begin{pmatrix}\text{the average number of}\\ \text{customers at station } i\end{pmatrix} \leq \begin{pmatrix}\text{given}\\ \text{threshold}\end{pmatrix}, \qquad i = 1,2,$$

or of the type

$$\begin{pmatrix}\text{the probability that the number of customers}\\ \text{at station } i \text{ exceeds a given threshold}\end{pmatrix} \leq \begin{pmatrix}\text{given}\\ \text{probability}\end{pmatrix}, \qquad i = 1,2.$$

If the stationary state distribution were of a geometric product-form $\pi_{\vec{\mathbf{q}}}$, then it would have been easy to translate such requirements into a required value of $\vec{\mathbf{q}} = \langle q_1, q_2 \rangle$. However, although the model clearly belongs to $\mathbb{MI}_2$, it does not necessarily belong to $\mathbb{P}_2$, namely to $\mathbb{MI}_2$'s subspace of product-form models. This model does not fall in any of the familiar classes of models that are known to be contained in $\mathbb{P}_2$, or more generally in $\mathbb{P}_n$, such as Jackson networks or Gelenbe networks. Nevertheless, for the purpose of designing this model, we will stay in $\mathbb{P}_{2,\langle q_1,q_2\rangle}$, with the $\langle q_1, q_2 \rangle$ calculated from the performance requirements. This space of models transcends way beyond the instances of the familiar classes of product-form models: The theory tells us that $\mathbb{P}_{2,\langle q_1,q_2\rangle}$ is a broad and well-characterized space, spread out within $\mathbb{MI}_2$. Moreover, the theory suggests that the geometric product-form stationary state distributions are fundamental for models with space homogeneity; stationary state distributions of non-$\mathbb{P}_2$ models in $\mathbb{MI}_2$ embody some intractable deformation due to the boundary. $\mathbb{P}_2$, or more generally $\mathbb{P}_n$, is a safe haven where the stationary state distribution is clear and explicit and where the treatment boils down to linear algebra and to linear programming.

Hence, we perform the minimization with regard to models in $\mathbb{P}_{2,\langle q_1,q_2\rangle}$ in which, instead of λ_1 and λ_2, there are fictitious values $\hat{\lambda}_1$ and $\hat{\lambda}_2$ in the interior and other fictitious values $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$ on the boundary, satisfying the constraints

$$\hat{\lambda}_1, \tilde{\lambda}_1 \geq \lambda_1, \qquad \hat{\lambda}_2, \tilde{\lambda}_2 \geq \lambda_2;$$

see Figure 8. For the selected performance requirements, the performance of the true model will be at least as good as the performance of a model with such fictitious arrival rates and with the same service rates, so the minimization will yield an upper bound on the needed service rates. Moreover, in order to achieve greater flexibility and to approach $\mathbb{P}_{2,\langle q_1,q_2\rangle}$ at a spot as close as possible to the true model, we could play with fictitious values for *any* nonnegative transition ($\vec{\mathbf{d}} \geq \vec{\mathbf{0}}$) as well as for any nonpositive transition ($\vec{\mathbf{d}} \leq \vec{\mathbf{0}}$). For the former transitions, the fictitious values should be constrained to be higher than the true values, and vice versa for the latter transitions. However, limiting ourselves to playing with the arrival rates only has the following merit: One does not have to rely on the monotonicity of the performance with respect to the transition rates, which may need a difficult rigorous proof; it is

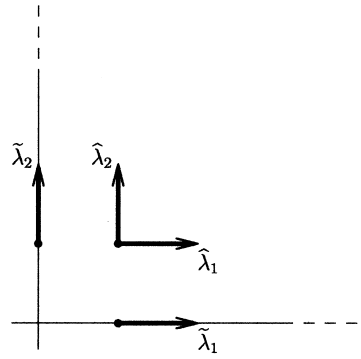

FIGURE 8. Replacing the true arrival rates λ_i with fictitious values: $\hat{\lambda}_i$ in the interior and $\tilde{\lambda}_i$ on the boundary.

possible to physically mimic the fictitious arrival rates in the real system, by artificially generating extra arrivals.

Theorem 4.1 states that $\mathbb{P}_{2,\langle q_1,q_2\rangle}$ is the intersection between $\mathbb{MI}_2$ and the solution space of the following homogeneous linear system (the partitioning due to walls is emphasized):

$$\left[\begin{array}{cccccccc|cc|cc}
1-q_1q_2 & 1-q_2 & 1-\frac{q_2}{q_1} & 1-q_1 & 1-\frac{1}{q_1} & 1-\frac{q_1}{q_2} & 1-\frac{1}{q_2} & 1-\frac{1}{q_1q_2} & 0 & 0 & 0 & 0 \\
\hline
-q_1q_2 & -q_2 & -\frac{q_2}{q_1} & 0 & 0 & 1 & 1 & 1 & 1-q_1 & 1-\frac{1}{q_1} & 0 & 0 \\
\hline
-q_1q_2 & 0 & 1 & -q_1 & 1 & -\frac{q_1}{q_2} & 0 & 1 & 0 & 0 & 1-q_2 & 1-\frac{1}{q_2}
\end{array}\right]$$

$$\times \left[\begin{array}{c}
\varphi_{\langle 1,1\rangle}(\langle -1,-1\rangle) \\
\varphi_{\langle 1,1\rangle}(\langle 0,-1\rangle) \\
\varphi_{\langle 1,1\rangle}(\langle 1,-1\rangle) \\
\varphi_{\langle 1,1\rangle}(\langle -1,0\rangle) \\
\varphi_{\langle 1,1\rangle}(\langle 1,0\rangle) \\
\varphi_{\langle 1,1\rangle}(\langle -1,1\rangle) \\
\varphi_{\langle 1,1\rangle}(\langle 0,1\rangle) \\
\varphi_{\langle 1,1\rangle}(\langle 1,1\rangle) \\
\hline
\varphi_{\langle 1,0\rangle}(\langle -1,0\rangle) \\
\varphi_{\langle 1,0\rangle}(\langle 1,0\rangle) \\
\hline
\varphi_{\langle 0,1\rangle}(\langle 0,-1\rangle) \\
\varphi_{\langle 0,1\rangle}(\langle 0,1\rangle)
\end{array}\right] = \left[\begin{array}{c} 0 \\ 0 \\ 0 \end{array}\right]. \tag{16}$$

In our particular kind of models within $\mathbb{P}_{2,\langle q_1,q_2\rangle}$, on which the minimization is performed, we have

$$\begin{bmatrix} \varphi_{\langle 1,1\rangle}(\langle -1,-1\rangle) \\ \varphi_{\langle 1,1\rangle}(\langle 0,-1\rangle) \\ \varphi_{\langle 1,1\rangle}(\langle 1,-1\rangle) \\ \varphi_{\langle 1,1\rangle}(\langle -1,0\rangle) \\ \varphi_{\langle 1,1\rangle}(\langle 1,0\rangle) \\ \varphi_{\langle 1,1\rangle}(\langle -1,1\rangle) \\ \varphi_{\langle 1,1\rangle}(\langle 0,1\rangle) \\ \varphi_{\langle 1,1\rangle}(\langle 1,1\rangle) \\ \hline \varphi_{\langle 1,0\rangle}(\langle -1,0\rangle) \\ \varphi_{\langle 1,0\rangle}(\langle 1,0\rangle) \\ \hline \varphi_{\langle 0,1\rangle}(\langle 0,-1\rangle) \\ \varphi_{\langle 0,1\rangle}(\langle 0,1\rangle) \end{bmatrix} = \begin{bmatrix} 0 \\ \theta_2 \rho_2^{\text{depart}} \\ \theta_2 \rho_2^{\text{transfer}} \\ \theta_1 \rho_1^{\text{depart}} \\ \hat{\lambda}_1 \\ \theta_1 \rho_1^{\text{transfer}} \\ \hat{\lambda}_2 \\ \theta_1 \rho_1^{\text{branch}} + \theta_2 \rho_2^{\text{branch}} \\ \hline \theta_1 \rho_1^{\text{depart}} \\ \tilde{\lambda}_1 \\ \hline \theta_2 \rho_2^{\text{depart}} \\ \tilde{\lambda}_2 \end{bmatrix}. \tag{17}$$

By substituting the right-hand side of (17) into (16), we obtain three equality constraints on our six variables θ_1, θ_2, $\hat{\lambda}_1$, $\hat{\lambda}_2$, $\tilde{\lambda}_1$, and $\tilde{\lambda}_2$. These equality constraints can be written in the form

$$\begin{aligned} a_1\theta_1 + a_2\theta_2 &= \left(\frac{1}{q_1} - 1\right)\hat{\lambda}_1 + \left(\frac{1}{q_2} - 1\right)\hat{\lambda}_2, \\ b_1\theta_1 + c_2\theta_2 &= \left(\frac{1}{q_1} - 1\right)\tilde{\lambda}_1 - \hat{\lambda}_2, \\ c_1\theta_1 + b_2\theta_2 &= \left(\frac{1}{q_2} - 1\right)\tilde{\lambda}_2 - \hat{\lambda}_1, \end{aligned}$$

where the coefficients a_1, b_1, and c_1 are given through

$$\begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix} \triangleq \begin{bmatrix} 1-q_1 & 1-\dfrac{q_1}{q_2} & 1-\dfrac{1}{q_1 q_2} \\ 1-q_1 & 1 & 1 \\ -q_1 & -\dfrac{q_1}{q_2} & 1 \end{bmatrix} \begin{bmatrix} \rho_1^{\text{depart}} \\ \rho_1^{\text{transfer}} \\ \rho_1^{\text{branch}} \end{bmatrix}; \tag{18}$$

the coefficients a_2, b_2, and c_2 are given by a dual expression, obtained by replacing every subscript 1 in (18) with 2, and vice versa. Due to the decoupling principle, once θ_1, θ_2, $\hat{\lambda}_1$, and $\hat{\lambda}_2$ are fixed, the remaining two variables, namely $\tilde{\lambda}_1$ and $\tilde{\lambda}_2$,

which are associated with two different one-dimensional walls, are decoupled from each other. This facilitates their elimination and the simplification of the equations. There remains the following linear program:

$$\min_{(\theta_1,\theta_2,\hat{\lambda}_1,\hat{\lambda}_2)\in\mathcal{R}_+^4} \theta_1+\theta_2$$

subject to

$$\begin{aligned}
&a_1\theta_1+a_2\theta_2=\left(\frac{1}{q_1}-1\right)\hat{\lambda}_1+\left(\frac{1}{q_2}-1\right)\hat{\lambda}_2,\\
&\hat{\lambda}_1\geq\lambda_1,\qquad \hat{\lambda}_2\geq\lambda_2,\\
&b_1\theta_1+c_2\theta_2+\hat{\lambda}_2\geq\left(\frac{1}{q_1}-1\right)\lambda_1,\\
&c_1\theta_1+b_2\theta_2+\hat{\lambda}_1\geq\left(\frac{1}{q_2}-1\right)\lambda_2. \qquad \textbf{(19)}
\end{aligned}$$

To take a numerical case, suppose for simplicity that the given parameters are equal for the two stations (so the station subscript is omitted); we also add artificial symmetry constraints for the variables (i.e., $\theta_1=\theta_2\triangleq\theta$, $\hat{\lambda}_1=\hat{\lambda}_2\triangleq\hat{\lambda}$, and $\tilde{\lambda}_1=\tilde{\lambda}_2\triangleq\tilde{\lambda}$). Suppose that $\lambda=1$, $\rho^{\text{depart}}=0.6$, $\rho^{\text{transfer}}=0.3$, $\rho^{\text{branch}}=0.1$, and $q=\frac{4}{5}$. Then, the optimal solution, under these artificial symmetry constraints, can be calculated readily and found to be $\theta=3\frac{47}{51}$, $\hat{\lambda}=1$, and $\tilde{\lambda}=1\frac{25}{51}$. Note that, in principle, we could extend the optimization so as to scan values of $\langle q_1,q_2\rangle$ that are smaller than the specific values that we calculated; the question of whether this may affect the optimal solution is open.

The design problem that has been considered so far could also have been handled using the following simple heuristics: Calculate the traffic through the stations and solve the problem as if each station were an $M/M/1$ queue. Applying this for the above numerical case, we find, from the mean traffic flow equation,

$$f=\lambda+f\rho^{\text{transfer}}+3f\rho^{\text{branch}},$$

that the traffic f through each of the stations is 2.5; that is, the traffic has an intensity 2.5 times larger than that of the external arrival stream. To produce a geometric factor of $\frac{4}{5}$ in an $M/M/1$ queue with this traffic, there is a need for a service rate of $\theta=3\frac{1}{8}$. This is a heuristic approximation for the solution, whereas the former value of $\theta=3\frac{47}{51}$ is a rigorous upper bound; better bounds are potentially available by using more flexible constraint schemes.

Because of the possibility of handling the original problem through a heuristics which essentially reduces it into a one-dimensional problem, we would like to present a variant of this problem for which there is no substitute for product-form models of full dimension: Suppose that the two stations share a common buffer space, which is limited; a customer that finds the buffer full is lost. We now require that the proba-

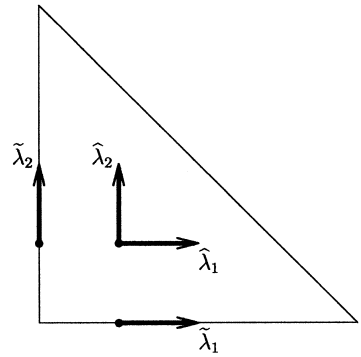

FIGURE 9. The state space in another variant of the original problem, where there is a common and limited buffer space (compare with Fig. 8).

bility that the buffer be full not exceed a given threshold, and this leads, again, to a $\langle q_1, q_2 \rangle$ or to an admissible range of $\langle q_1, q_2 \rangle$'s. In this new variant, the state space appears as in Figure 9. The added slanted wall and two sharp corners would comply with Remark 4.2 unless the transitions in direction $\langle 1,1 \rangle$ were present. See Figure 10. Although the space homogeneity, in its original sense, is disrupted, the transition structure is still space homogeneous in the following sense: The global balance equation is identical for all the states in any given wall (including the interior). The conclusion of Remark 4.2 remains valid. Thus, due to the decoupling principle, all that happens to the linear program (19) in this variant of the problem is that a single

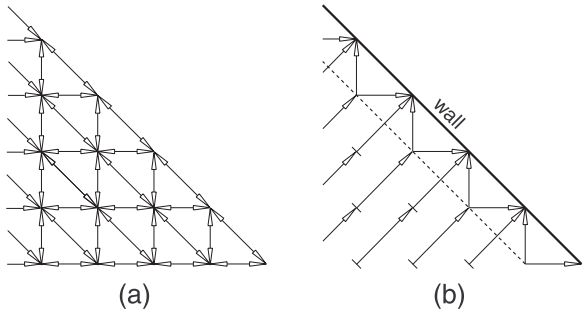

FIGURE 10. The transition rates of the system with the limited buffer (shown near a sharp corner) are a superposition of (a) and (b): (a) gives the "legal" transitions of Figure 4 and (b) gives the transition rates in direction $\langle 1,1 \rangle$, that change their behavior at the diagonal adjacent to the wall. There, at the diagonal, the latter transitions are diverted into the two directions $\langle 1,0 \rangle$ and $\langle 0,1 \rangle$ while preserving their total value.

linear equality constraint in θ_1, θ_2, $\hat{\lambda}_1$, and $\hat{\lambda}_2$ is added, contributed by the slanted wall.

Acknowledgments

We are pleased to acknowledge the help of the referees in improving the paper.

The preliminary stage of this work was supported by European grant BRA-QMIPS of CEC DG XIII, which enabled the research of N. Bayer at CWI Amsterdam. The research of R.J. Boucherie has been made possible by a fellowship from the Royal Netherlands Academy of Arts and Sciences.

References

1. Baskett, F., Chandy, K.M., Muntz, R.R., & Palacios, F.G. (1975). Open, closed and mixed networks of queues with different classes of customers. *Journal of the ACM* 22: 248–260.
2. Bayer, N. & Kogan, Y.A. (1997). Branching/queueing networks: Their introduction and near-decomposability asymptotics. *Queueing Systems* 27: 251–269.
3. Boucherie, R.J. & van Dijk, N.M. (1991). Product forms for queueing networks with state dependent multiple job transitions. *Advances in Applied Probability* 23: 152–187.
4. Boucherie, R.J. & van Dijk, N.M. (1994). Local balance in queueing networks with positive and negative customers. *Annals of Operations Research* 48: 463–492.
5. Brockmeyer E., Halstrom, H.L., & Jensen, A. (1948). *The life and works of A. K. Erlang*. Copenhagen: Academy of Technical Sciences.
6. Chao, X. & Miyazawa, M. (1998). On quasi-reversibility and local balance: An alternative derivation of the product-form results. *Operations Research* 46: 927–933.
7. Chao, X., Miyazawa, M., & Pinedo, M. (1999). *Queueing networks—Customers, signals and product form solutions*. New York: Wiley.
8. Chao, X., Miyazawa, M., Serfozo, R.F., & Takada, H. (1998). Markov network processes with product form stationary distributions. *Queueing Systems* 28(4): 377–401.
9. Chao, X. & Pinedo, M. (1993). On generalized networks of queues with positive and negative arrivals. *Probability in the Engineering and Informational Sciences* 7: 301–334.
10. Chung, K.L. (1974). *A course in probability theory*, 2nd ed. New York: Academic Press.
11. van Dijk, N.M. (1993). *Queueing networks and product forms*. New York: Wiley.
12. Gelenbe, E. (1991). Product-form queueing networks with negative and positive customers. *Journal of Applied Probability* 28: 656–663.
13. Gordon, W.J. & Newell, G.F. (1967). Closed queueing systems with exponential servers. *Operations Research* 15: 254–265.
14. Henderson, W. & Taylor, P.G. (1990). Product form in networks of queues with batch arrivals and batch services. *Queueing Systems* 6: 71–88.
15. Jackson, J.R. (1957). Networks of waiting lines. *Operations Research* 5: 518–521.
16. Jackson, J.R. (1963). Jobshop-like queueing systems. *Management Science* 10: 131–142.
17. Kelly, F.P. (1975). Networks of queues with customers of different types. *Journal of Applied Probability* 12: 542–554.
18. Kelly, F.P. (1979). *Reversibility and stochastic networks*. New York: Wiley.
19. Kelly, F.P. (1982). Networks of quasi-reversible nodes. In I.R.L. Disney & T.J. Ott (eds.), *Applied probability—Computer science: The interface*, Vol. I. Boston: Birkhauser, pp. 3–26.
20. Kelly, F.P. (1991). Loss networks. *The Annals of Applied Probability* 1: 319–378.
21. Miyazawa, M. & Taylor, P.G. (1997). Geometric product-form distribution for a queueing network with nonstandard batch arrivals and batch transfers. *Advances in Applied Probability* 29: 523–544.
22. Muntz, R.R. (1972). Poisson departure processes and queueing networks. Technical Report RC4145, IBM. A shorter version appeared in the Proceedings of the Seventh Annual Conference on Information Science and Systems, Princeton, 1973, pp. 435–440.
23. Rockafellar, T.R. (1970). *Convex analysis*. Princeton, NJ: Princeton University Press.

24. Royden, H.L. (1988). *Real analysis*, 3rd ed. New York: Macmillan.
25. Rudin, W. (1973). *Functional analysis*. New York: McGraw-Hill.
26. Schassberger, R. (1978). The insensitivity of stationary probabilities in networks of queues. *Advances in Applied Probability* 10: 906–912.
27. Serfozo, R. (1999). *Introduction to stochastic networks*. New York: Springer-Verlag.
28. Taylor, P.G. & van Dijk, N.M. (1998). Strong stochastic bounds for the stationary distribution of a class of multicomponent performability models. *Operations Research* 46: 665–674.
29. Whittle, P. (1967). Nonlinear migration processes. *Bulletin of the International Institute of Statistics* 42: 642–647.
30. Whittle, P. (1968). Equilibrium distributions for an open migration process. *Journal of Applied Probability* 5: 567–571.
31. Whittle, P. (1986). *Systems in stochastic equilibrium*. New York: Wiley.
32. Williams, R.J. (1995). Semimartingale reflecting Brownian motions in the orthant. In F.P. Kelly & R.J. Williams (eds), *Stochastic networks*. New York: Springer-Verlag, pp. 125–138.