J. Appl. Prob. **45,** 846–878 (2008)
Printed in England

IMPACT OF ROUTEING ON CORRELATION STRENGTH IN STATIONARY QUEUEING NETWORK PROCESSES

HANS DADUNA,* *Hamburg University*
RYSZARD SZEKLI,** *Wrocław University*

Abstract

For exponential open and closed queueing networks, we investigate the internal dependence structure, compare the internal dependence for different networks, and discuss the relation of correlation formulae to the existence of spectral gaps and comparison of asymptotic variances. A central prerequisite for the derived theorems is stochastic monotonicity of the networks. The dependence structure of network processes is described by concordance order with respect to various classes of functions. Different networks with the same first-order characteristics are compared with respect to their second-order properties. If a network is perturbed by changing the routeing in a way which holds the routeing equilibrium fixed, the resulting perturbations of the network processes are evaluated.

Keywords: Product form network; concordance ordering; space–time correlation; spectral gap; asymptotic variance; stochastic monotonicity; MCMC; Peskun ordering

2000 Mathematics Subject Classification: Primary 60K25; 60J25

1. Introduction

We revisit classical stochastic networks of the Jackson- and Gordon–Newell-type and investigate the internal dependence structure of the networks, compare the internal dependence for different networks, and discuss further some closely related topics. Dependence will be evaluated as generalized correlation over time and space of the multidimensional network processes, described by concordance order with respect to convex cones of functions of the multidimensional marginals.

The theory of dependence order via integral orders for finite-dimensional vectors is well established, surveys can be found in [12], [18, Chapter 3], and [21, Section 3.4]. In recent years this theory and its applications were extended to dependence ordering of stochastic processes; for examples with state spaces $\mathbb{R}^n$ or subsets thereof, see [7] and [17], and for a more general approach to Markov processes in discrete and continuous time with general partially ordered state space, see [4].

The general theory for comparison of Markov processes with respect to their internal dependence structure revealed that sometimes there is a complicated interplay of monotonicity properties with some generalized correlation structure of the processes. Such a monotonicity requirement is not unexpected if we recall that the theory of association in time for Markovian processes is mainly developed for monotone Markov processes; for a review, see Chapter II

Received 17 April 2008; revision received 25 July 2008.
* Postal address: Department of Mathematics, Hamburg University, Bundesstrasse 55, 20146 Hamburg, Germany. Email address: daduna@math.uni-hamburg.de
** Postal address: Mathematical Institute, Wrocław University, pl. Grunwaldzki 2/4, 50-384 Wrocław, Poland.

of [14]. Proving association in time for a process means that we compare its internal dependence, defined by generalized correlations with the *independent version* of this process. Association is a powerful tool in obtaining probability bounds, e.g. in the realm of interacting processes of attractive particle systems. (A system is called attractive if it exhibits (strong) stochastic monotonicity.)

In the context of stochastic networks it turns out that similar connections between monotonicity and correlation are fundamental, but, owing to the more complex structure of the processes, we usually cannot hope to utilize the strong stochastic order, as required for association, or in the development in [4] and [7].

In the theory of stochastic orders and especially in specific applications, a well-established procedure is to tailor suitable classes of functions that, via integrals over these functions, extract the required properties of the models under consideration. The most well-known example is the class of integrals over convex functions which describes the volatility of processes and, therefore, the risks connected with the process.

Similar ideas will guide our investigations of network processes $\boldsymbol{X} = (X_t : t \geq 0)$ and $\boldsymbol{Y} = (Y_t : t \geq 0)$. These are comparable in the concordance ordering, $\boldsymbol{X} \prec_{\mathrm{cc}} \boldsymbol{Y}$, if, for each pair $(X_{t_1}, \dots, X_{t_n})$ and $(Y_{t_1}, \dots, Y_{t_n})$,

$$\mathrm{E}\biggl[\prod_{i=1}^{n} f_i(X_{t_i})\biggr] \leq \mathrm{E}\biggl[\prod_{i=1}^{n} f_i(Y_{t_i})\biggr] \tag{1.1}$$

holds for all nonnegative increasing functions f_i and all nonnegative decreasing functions as well (i.e. for all comonotone functions). It is our task to identify subclasses $\mathcal{F}$ of functions such that (1.1) holds for all comonotone functions in $\mathcal{F}$ and that, additionally, $\boldsymbol{X}$ and $\boldsymbol{Y}$ fulfill the corresponding stochastic monotonicity properties with respect to the integral order defined via $\mathcal{F}$.

For applications, it is most important to find sufficient conditions by reducing requirement (1.1) to the case in which $n = 2$, and, moreover, in the continuous-time setting to an $n = 2$ analogue for infinitesimal generator inequalities. Tailoring such $\mathcal{F}$-based kernel or generator inequalities for pairs of network processes and combining these with the needed monotonicity structures is the main idea of this paper.

The pairs of network processes in our investigations are always related by some structural similarities; we can usually think of one network being obtained from the other by some structural perturbation. The perturbations we are mainly interested in are due to the perturbations of the routeing of individual customers. We will always give a precise meaning of what the perturbations are and of the resulting structural properties.

In the general theory of concordance order, the set (1.1) of inequalities implies that $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same marginals and that standard covariances $\operatorname{cov}(f(X_s), g(X_t)) \leq \operatorname{cov}(f(Y_s), g(Y_t))$ are ordered for comonotone f and g. If $\mathcal{F}$ is sufficiently rich, these properties will be maintained. Nevertheless, we assume from scratch that only stationary processes are considered and, moreover, that $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same equilibrium.

Our investigation will show that the conditions that determine comparability of dependence, i.e. second-order properties of processes having the same first-order behavior (i.e. the same steady state), are closely connected with further properties of the asymptotic behavior of the processes: the asymptotic variance of certain functionals (performance measures and cost functions) of the network processes and the speed of convergence to stationarity via comparison of the spectral gap. A similar observation in a general setting has already been made in [4].

We will continue this discussion in Section 4 in connection with rather general network comparison formulae which govern the ($n = 2$) infinitesimal generator inequalities. This will especially show that our results resemble those obtained in the construction of optimal Markov chain Monte Carlo (MCMC) methods in simulation. The Peskun order which is used there in connection with reversibility can be used in our framework as well and, moreover, occurs as a special case of dependence ordering with monotonicity requirements, as developed in Section 7.

In Section 4 we then study the correlation-type inequalities for network processes in more detail. We investigate the internal dependence behavior of network processes in steady state under different routeing regimes while the individual nodes' behavior remains unchanged and, for open networks, the total arrival rate at the network is fixed, such that the first-order state characteristics remain invariant for the considered networks after perturbation.

Given a prescribed network in equilibrium, our conjecture is that, if we perturb the routeing process (which governs the movements of the customers after being served at any node) so as to make it more dependent in a specified way, it is possible to show that the joint queue length process after perturbation will be more dependent in some (possibly differently) specified way.

We especially investigate two ways in which the routeing process is perturbed. The first way is by making routeing more chaotic, which is an approach used in statistical mechanics. There exists a well-established method to ordering of the *chaotic behavior of a random walker* if his itinerary is governed by doubly stochastic routeing matrices; see [1, Chapter 1]. We will prove that if the routeing is becoming more chaotic in this sense then the joint queue length process will show less internal dependency.

While the perturbation of the routeing in this case is not connected with any order (numbering) of the nodes, the second way of perturbing the routeing is connected to some preassigned order of the nodes, which is expressed by a graph structure. Assuming that the routeing of customers is compatible with this graph structure, we perturb it by shifting the probability mass in the routeing kernel along paths that are determined by the graph. We will prove that if we shift the masses in a way that routeing becomes more positive dependent then the internal dependence of the joint queue length process will increase.

After having derived the required correlation related inequalities, we investigate in Section 5, in a general framework, the interplay of these correlation inequalities with stochastic monotonicity. The central notion for a pair of Markov processes is a symmetric monotonicity for the processes and their time reversals.

In Section 6 we exploit these principles for specific networks, showing that a delicate balance is necessary between monotonicity and correlation inequalities. Furthermore, we show that it is possible to apply the general principles that are expressed for partially ordered state spaces to different order structures for the network processes.

The paper is opened with a short description of network processes and their steady-state behavior in Section 3, and closed with a discussion of the relation of our results to the methods for constructing optimal MCMC transition kernels in simulation in Section 7.

Besides the cited references, there are some related papers available where perturbation of a network process is different from the principles described here. In these investigations the speed of service for the nodes is changed by the same factor as the arrival intensities are changed. It follows that the steady-state probabilities and the other first-order characteristics are unchanged (as we also require), although it can be shown that the internal dependencies of the processes can considerably differ [2], [22]. Such additional perturbations can be considered in our framework as well.

2. Notation

For probability spaces $(\mathbb{E}, \mathcal{E}, \pi)$ and functions $f, g: (\mathbb{E}, \mathcal{E}) \to (\mathbb{R}, \mathbb{B})$, we define the inner product of f and g with respect to π as

$$\langle f, g\rangle_\pi = \int_{\mathbb{E}} f(x)g(x)\,\pi(\mathrm{d}x).$$

We denote by $L_2(\mathbb{E}, \pi)$ the space of square-integrable functions with respect to π, and $\|f\|_\pi = (\langle f, f\rangle_\pi)^{1/2}$.

Let $\mathbb{E}$ be a topological space with a closed partial order '$\prec$'. The Borel σ-algebra, generated by the topology, is denoted by $\mathcal{E}$. We denote by $\mathcal{I}^*(\mathbb{E})$ the set of all real-valued increasing measurable bounded functions on $\mathbb{E}$, by $\mathcal{I}^*_+(\mathbb{E})$ the set of such functions that are nonnegative, and by $\mathcal{I}(\mathbb{E})$ the set of all increasing Borel sets (i.e. sets whose indicator functions are increasing). The decreasing analogues are denoted by $\mathcal{D}^*(\mathbb{E})$, $\mathcal{D}^*_+(\mathbb{E})$, and $\mathcal{D}(\mathbb{E})$, respectively.

Product spaces will be considered with product topology. Unless otherwise specified, on the product space $\mathbb{E}^n$ we use the coordinatewise ordering '$\prec^n$', $n \in \mathbb{N}$.

We denote the Kronecker delta by

$$\delta_{i,j} = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j, \end{cases}$$

and, for any real-valued vector $\xi = (\xi_i : 0 \le i \le J)$, we define the diagonal matrix with entries from ξ by

$$\operatorname{diag}(\xi) = (\delta_{i,j}\xi_i : 0 \le i, j \le J).$$

For $k = 1, \dots, J$, the kth J-dimensional unit (row) vector is $e_k := (\delta_{jk} : j = 1, \dots, J)$.

For vector $(\alpha_1, \dots, \alpha_J) \in \mathbb{R}^J$, the rank statistic $\mathcal{R}(\alpha) = (\mathcal{R}_1(\alpha), \dots, \mathcal{R}_J(\alpha)) \in \mathbb{N}^J$ is defined by enumeration of the indices of the $\alpha_{(\cdot)}$ in decreasing order of their associated $\alpha_{(\cdot)}$-values, i.e.

$$\alpha_{\mathcal{R}_i(\alpha)} \ge \alpha_{\mathcal{R}_{i+1}(\alpha)}, \qquad i = 1, \dots, J-1,$$

and ties are resolved according to the natural order of the indices. The vector $A\mathcal{R}(\alpha) = (A\mathcal{R}_1(\alpha), \dots, A\mathcal{R}_J(\alpha)) \in \mathbb{N}^J$ of antiranks of α is defined by $A\mathcal{R}_j(\alpha) = \mathcal{R}_{J+1-j}(\alpha)$, and so yields an enumeration of the indices of α in increasing order of their associated $\alpha_{(\cdot)}$-values.

3. Stochastic network models

A Jackson network [10] consists of J nodes numbered $1, \dots, J$, where customers arrive in independent external Poisson streams at node j with finite intensity $\lambda_j \ge 0$. We set $\Lambda = (\lambda_1, \dots, \lambda_J)$ and $\lambda = \lambda_1 + \cdots + \lambda_J > 0$. Customers are indistinguishable and follow the same rules. Requests for service are exponentially distributed with mean 1 at all nodes, and constitute an independent family of variables which is independent of the arrival streams.

Nodes are exponential single servers with state-dependent service rates and an infinite waiting room under a first-come–first-served (FCFS) regime. If at node j there are $n_j > 0$ customers present, either in service or waiting, then service is provided there at rate $\mu_j(n_j) > 0$. (Therefore, in general, the obtained service time is not exponential 1.) We assume that $\sup\{\mu_j(k) : j \in \{1, \dots, J\},\ k \in \mathbb{N}\} < \infty$.

Routeing is Markovian, i.e. a customer departing from node i immediately proceeds to node j with probability $r_{ij} \ge 0$ and departs from the network with probability r_{j0}. Taking

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

$r_{0i} = \lambda_i/\lambda$ and $r_{00} = 0$, we assume that the routeing matrix $R = [r_{ij}]_{i,j=0,\dots,J}$ is irreducible. This ensures that the traffic equations,

$$\eta_j = \lambda_j + \sum_{i=1}^{J} \eta_i r_{ij}, \qquad j = 1, \dots, J, \tag{3.1}$$

have a unique solution which we denote by $\eta = (\eta_j : j = 1, \dots, J)$. We extend the traffic equation, (3.1), to a steady-state equation for a routeing Markov chain by

$$\eta_j = \sum_{i=0}^{J} \eta_i r_{ij}, \qquad j = 0, 1, \dots, J, \tag{3.2}$$

which has the solution $\eta = (\eta_j : j = 0, 1, \dots, J)$, where $\eta_0 := \lambda$, and we use η_j given in (3.1). We use η in both meanings and refer to (3.2) as the *extended traffic solution* η. Usually, η is not a stochastic vector, and we define the unique stochastic solution of (3.2) by

$$\xi = (\xi_j : j = 0, 1, \dots, J).$$

Let $\boldsymbol{X} = (X_t : t \geq 0)$ denote the vector process recording the joint queue lengths in the network for time t. Here $X_t = (X_1(t), \dots, X_J(t)) \in \mathbb{N}^J$ means that at time t there are $X_j(t)$ customers present at node j, either in service or waiting. The assumptions put on the system imply that $\boldsymbol{X}$ is a strong Markov process on the state space $\mathbb{N}^J$ with generator $Q^{\boldsymbol{X}} = (Q^{\boldsymbol{X}}(\underline{n}, \underline{m}) : \underline{m}, \underline{n} \in \mathbb{N}^J)$, which is, for $g : \mathbb{N}^J \to \mathbb{R}$,

$$\begin{aligned}(Q^{\boldsymbol{X}} g)(\underline{n}) = {} & \sum_{j=1}^{J} \lambda_j (g(\underline{n} + e_j) - g(\underline{n})) + \sum_{j=1}^{J} (1 - \delta_{0n_j}) \mu_j(n_j) r_{j0} (g(\underline{n} - e_j) - g(\underline{n})) \\ & + \sum_{j=1}^{J} (1 - \delta_{0n_j}) \mu_j(n_j) \sum_{i=1}^{J} r_{ji} (g(\underline{n} - e_j + e_i) - g(\underline{n})).\end{aligned}$$

From $\sup\{\mu_j(k) : j \in \{1, \dots, J\},\ k \in \mathbb{N}\} < \infty$, it follows that $Q^{\boldsymbol{X}}$ is a bounded operator, i.e. $\inf_{\underline{n} \in \mathbb{N}^J} Q^{\boldsymbol{X}}(\underline{n}, \underline{n}) > -\infty$. We assume throughout that the network process $\boldsymbol{X}$ is ergodic.

For an ergodic network process $\boldsymbol{X}$, Jackson's theorem [10] states that the unique steady state and limiting distribution π on $\mathbb{N}^J$ is

$$\pi(\underline{n}) = \pi((n_1, \dots, n_J)) = \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)} C(j)^{-1} \tag{3.3}$$

with normalizing constants $C(j)$ for marginal distributions of $\boldsymbol{X}$.

A Gordon–Newell network [5], [11] is defined with the same set of J nodes as the Jackson network and a similar independence assumption with respect to the routeing of I customers, who cycle forever in the network according to a Markovian routeing matrix $R = [r_{ij}]_{i,j=1,\dots,J}$. We assume that R is irreducible, which implies that the traffic equations,

$$\eta_j = \sum_{i=1}^{J} \eta_i r_{ij}, \qquad j = 1, \dots, J, \tag{3.4}$$

have a unique probability solution that we denote by $\eta = (\eta_j : j = 1, \dots, J)$.

The joint queue length vector process $\boldsymbol{X} = (X(t) := (X_j(t): j = 1, \dots, J): t \geq 0)$ of the Gordon–Newell network is an ergodic Markov process with state space $S(I, J) = \{\underline{n} = (n_1, \dots, n_J) \in \mathbb{N}^J : n_1 + \cdots + n_J = I\}$ and generator $Q^{\boldsymbol{X}}$, which is, for $g\colon S(I, J) \to \mathbb{R}$,

$$(Q^{\boldsymbol{X}} g)(\underline{n}) = \sum_{j=1}^{J} (1 - \delta_{0 n_j}) \mu_j(n_j) \sum_{i=1}^{J} r_{ji} (g(\underline{n} - e_j + e_i) - g(\underline{n})).$$

The unique stationary and limiting distribution $\pi = \pi(I, J)$ of $\boldsymbol{X}$ on $S(I, J)$ was obtained in [5] and [11] as (with $G(I, J)$ being the overall normalizing constant)

$$\pi(I, J)(n_1, \dots, n_J) = G(I, J)^{-1} \prod_{j=1}^{J} \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)}, \qquad (n_1, \dots, n_J) \in S(I, J).$$

4. Correlation inequalities via generators

For a queue length network process $\boldsymbol{X}$ with generator $Q^{\boldsymbol{X}}$ and stationary distribution π, we are interested in *one-step* correlation expressions:

$$\langle f, Q^{\boldsymbol{X}} g \rangle_\pi. \tag{4.1}$$

If $f = g$ then (4.1) is (the negative of) a quadratic form, because $-Q^{\boldsymbol{X}}$ is positive definite. Equation (4.1) occurs in the definition of Cheeger's constant, which is helpful to bound the second largest eigenvalue of $Q^{\boldsymbol{X}}$ (because division of (4.1) by $\langle f, f \rangle_\pi$ yields Rayleigh quotients), which essentially governs the speed of convergence of $\boldsymbol{X}$ to its equilibrium.

Equation (4.1) can be utilized to determine the asymptotic variance of costs or performance measures associated with Markovian processes (network processes) and to compare the asymptotic variances of two such related processes.

In a natural way, the correlations occur when comparing the dependence structure of $\boldsymbol{X}$ with that of a related process $\widetilde{\boldsymbol{X}}$ with the same stationary distribution π, where we evaluate

$$\langle f, Q^{\boldsymbol{X}} g \rangle_\pi - \langle f, Q^{\widetilde{\boldsymbol{X}}} g \rangle_\pi; \tag{4.2}$$

see, e.g. (iv) and (v) of Theorem 5.2, below.

Because we are dealing with processes having bounded generators, properties connected with (4.1) can be turned into properties of

$$\langle f, I + \varepsilon Q^{\boldsymbol{X}} g \rangle_\pi = \mathrm{E}_\pi[f(X_0) g(X_\tau)], \tag{4.3}$$

where I is the identity operator, $\varepsilon > 0$ is sufficiently small such that $I + \varepsilon Q^{\boldsymbol{X}}$ is a stochastic matrix, and $\tau \sim \exp(\varepsilon)$ (exponentially distributed). This enables us to directly apply discrete-time methods to characterize properties of continuous-time processes in the range of problems sketched above.

We begin this section with important new expressions that connect, for continuous-time processes, the differences (4.2) of covariances for related network processes with some covariances for the corresponding routeing matrices. The idea behind these expressions is that the original network with routeing matrix R is subject to some perturbation, which is realized by a perturbation of the routeing scheme, that yields a new routeing matrix $\tilde{R}$ having the same solution η of the traffic equation, but showing different second-order properties, and a perturbed network process $\tilde{\boldsymbol{X}}$.

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

We then discuss how to utilize our findings to (i) comparisons of asymptotic variances in central limit theorems for performance functionals, (ii) comparisons of spectral gaps when determining the speed of convergence for network processes, and (iii) comparisons of dependencies for related Markov chains.

Proposition 4.1. *Suppose that $\mathbf{X}$ is an ergodic Jackson network process with routeing matrix R and that $\widetilde{\mathbf{X}}$ is the Jackson network process having the same arrival and service intensities but routeing matrix $\widetilde{R} = [\widetilde{r}_{ij}]$ such that the extended traffic solutions η of the traffic equation for R and $\widetilde{R}$ coincide. Then, for arbitrary real functions f and g,*

$$\langle f, Q^{\mathbf{X}} g\rangle_\pi - \langle f, Q^{\widetilde{\mathbf{X}}} g\rangle_\pi = \frac{\lambda}{\xi_0} \mathrm{E}_\pi[\mathrm{tr}(W^{g,f}(X_t)\mathrm{diag}(\xi)(R - \widetilde{R}))],$$

where ξ is the probability solution of the extended traffic equation, (3.2), $e_0 = (0, \dots, 0)$, and

$$W^{g,f}(\underline{n}) = [g(\underline{n} + e_i) f(\underline{n} + e_j)]_{i,j=0,1,\dots,J}.$$

Proof. We first compute $\langle f, Q^{\mathbf{X}} g\rangle_\pi$, which is, from the definition (recall that $\mu_j(0) = 0$ for all j),

$$\begin{aligned}
&\sum_{\underline{n}\in\mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) \sum_{m\in\mathbb{N}^J} Q^{\mathbf{X}}(\underline{n}, \underline{m}) g(\underline{m}) \\
&\quad = \sum_{\underline{n}\in\mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) \biggl(\sum_{j=1}^J \lambda_j g(\underline{n} + e_j) + \sum_{j=1}^J \mu_j(n_j) r_{j0} g(\underline{n} - e_j) \\
&\qquad\qquad + \sum_{j=1}^J \mu_j(n_j) \sum_{\substack{i=1\\ i\neq j}}^J r_{ji} g(\underline{n} - e_j + e_i) \\
&\qquad\qquad - \biggl(\sum_{j=1}^J \lambda_j + \sum_{j=1}^J \mu_j(n_j)(1 - r_{jj}) \biggr) g(\underline{n}) \biggr) \\
&\quad = G^{-1} \sum_{\underline{n}\in\mathbb{N}^J} \prod_{j=1}^J \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)} \biggl(\sum_{j=1}^J f(\underline{n}) g(\underline{n} + e_j) \lambda r_{0j} + \sum_{j=1}^J f(\underline{n}) g(\underline{n} - e_j) \mu_j(n_j) r_{j0} \\
&\qquad\qquad + \sum_{j=1}^J \sum_{i=1}^J f(\underline{n}) g(\underline{n} - e_j + e_i) \mu_j(n_j) r_{ji} \biggr) \\
&\qquad - \sum_{\underline{n}\in\mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^J (\lambda_j + \mu_j(n_j)) \\
&\quad = G^{-1} \sum_{\underline{n}\in\mathbb{N}^J} \prod_{j=1}^J \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)} \sum_{j=1}^J \lambda f(\underline{n}) g(\underline{n} + e_j) r_{0j} \\
&\qquad + G^{-1} \sum_{j=1}^J \sum_{\underline{n}\in\mathbb{N}^J} \prod_{j=1}^J \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)} \mu_j(n_j) \sum_{j=1}^J f(\underline{n}) g(\underline{n} - e_j) \mu_j(n_j) r_{j0}
\end{aligned}$$

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

$$
\begin{aligned}
&+ G^{-1} \sum_{j=1}^{J} \sum_{i=1}^{J} \sum_{\underline{n} \in \mathbb{N}^J} \prod_{j=1}^{J} \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)} \mu_j(n_j) f(\underline{n}) g(\underline{n} - e_j + e_i) r_{ji} \\
&- \sum_{\underline{n} \in \mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^{J} (\lambda_j + \mu_j(n_j)) \\
= &\sum_{\underline{n} \in \mathbb{N}^J} \pi(\underline{n}) \bigg(\sum_{i=1}^{J} \eta_0 f(\underline{n}) g(\underline{n} + e_i) r_{0i} + \sum_{j=1}^{J} \eta_j f(\underline{n} + e_j) g(\underline{n}) r_{j0} \\
&\qquad + \sum_{j=1}^{J} \sum_{i=1}^{J} \eta_j f(\underline{n} + e_j +) g(\underline{n} + e_i) r_{ji} \bigg) \\
&- \sum_{\underline{n} \in \mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^{J} (\lambda_j + \mu_j(n_j)) \\
= &\frac{\lambda}{\xi_0} \sum_{\underline{n} \in \mathbb{N}^J} \pi(\underline{n}) \sum_{i=0}^{J} \sum_{j=0}^{J} \xi_j f_j(\underline{n}) g_i(\underline{n}) r_{ji} - \sum_{\underline{n} \in \mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^{J} (\lambda_j + \mu_j(n_j)) \\
= &\frac{\lambda}{\xi_0} \mathrm{E}_\pi [\mathrm{tr}(W^{f,g} \mathrm{diag}(\xi) R)] - \sum_{\underline{n} \in \mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^{J} (\lambda_j + \mu_j(n_j)).
\end{aligned}
$$

In the second equality we used $r_{0j} = \lambda_j / \lambda$ and collected terms. The next block displays the most explicit form of our relevant equation, which enables us to cancel and shift the summation indices in the fourth equality. In the fifth equality we used $\eta_0 = \lambda$ and $\eta_j = \lambda \xi_j / \xi_0$, and applied the definitions of $f_j(\cdot)$ and $g_i(\cdot)$. In the sixth equality we collected definitions of the respective matrices and applied the trace operator.

Similarly, we compute

$$
\begin{aligned}
\langle f, Q^{\widetilde{X}} g \rangle_\pi &= \sum_{n \in \mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) \sum_{\underline{m} \in \mathbb{N}^J} Q^{\widetilde{X}}(\underline{n}, \underline{m}) g(\underline{m}) \\
&= \frac{\lambda}{\xi_0} \mathrm{E}_\pi [\mathrm{tr}(W^{f,g} \mathrm{diag}(\xi) \widetilde{R})] - \sum_{\underline{n} \in \mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^{J} (\widetilde{\lambda}_j + \mu_j(n_j)),
\end{aligned}
$$

which immediately yields the statement of the proposition. From the assumption that the extended traffic solutions η (and $\widetilde{\eta}$) of the traffic equation for R and $\widetilde{R}$ coincide and from

$$
\sum_{j=1}^{J} \lambda_j = \eta_0 = \widetilde{\eta}_0 = \sum_{j=1}^{J} \widetilde{\lambda}_j,
$$

it follows that the remainder terms in both expressions are the same. This completes the proof.

Proposition 4.2. *Suppose that* $\boldsymbol{X}$ *is an ergodic Gordon–Newell network process with routeing matrix* R *and that* $\widetilde{\boldsymbol{X}}$ *is the Gordon–Newell network process having the same service intensities but routeing matrix* $\widetilde{R} = [\widetilde{r}_{ij}]$ *such that the stochastic traffic solutions* η *of the traffic equation* *for R and $\widetilde{R}$ coincide. Then, for arbitrary real functions f and g,*

$$\langle f, Q^X g\rangle_\pi - \langle f, Q^{\tilde{X}} g\rangle_\pi = \frac{G(I-1,J)}{G(I,J)} \mathrm{E}_{\pi^{I-1,J}}[\mathrm{tr}(W^{g,f}(X_t)\mathrm{diag}(\eta)(R-\widetilde{R}))],$$

where η is the probability solution of the traffic equation, (3.4), $e_0 = (0,\dots,0)$, and

$$W^{g,f}(\underline{n}) = [g(\underline{n}+e_i)f(\underline{n}+e_j)]_{i,j=1,\dots,J}.$$

Proof. The proof is similar to the proof of Proposition 4.1 and, therefore, we sketch only the main points. From the definition we have

$$\begin{aligned}
&\sum_{\underline{n}\in\mathbb{S}(I,J)} \pi^{(I,J)}(\underline{n}) f(\underline{n}) \sum_{\underline{m}\in\mathbb{S}(I,J)} Q^X(\underline{n},\underline{m}) g(\underline{m})\\
&= G^{-1} \sum_{\underline{n}\in S(I,J)} \prod_{\ell=1}^{J}\prod_{k=1}^{n_\ell} \frac{\eta_\ell}{\mu_\ell(k)} \sum_{j=1}^{J}\sum_{i=1}^{J} f(\underline{n}) g(\underline{n}-e_j+e_i)\mu_j(n_j) r_{ji}\\
&\quad - \sum_{\underline{n}\in S(I,J)} \pi^{(I,J)}(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^{J} \mu_j(n_j)\\
&= G(I,J)^{-1} \sum_{j=1}^{J}\sum_{i=1}^{J} r_{ji} \sum_{\substack{\underline{n}\in S(I,J)\\ n_j>0}} \prod_{\ell=1}^{J}\prod_{k=1}^{n_\ell} \frac{\eta_\ell}{\mu_\ell(k)} \mu_j(n_j) f(\underline{n}) g(\underline{n}-e_j+e_i)\\
&\quad - \sum_{\underline{n}\in S(I,J)} \pi^{(I,J)}(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^{J} \mu_j(n_j)\\
&= G(I,J)^{-1} \sum_{j=1}^{J} \eta_j \sum_{i=1}^{J} r_{ji} \sum_{\underline{n}\in S(I-1,J)} \prod_{\ell=1}^{J}\prod_{k=1}^{n_\ell} \frac{\eta_\ell}{\mu_\ell(k)} f(\underline{n}+e_j) g(\underline{n}+e_i)\\
&\quad - \sum_{\underline{n}\in S(I,J)} \pi^{(I,J)}(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^{J} \mu_j(n_j)\\
&= \frac{G(I-1,J)}{G(I,J)} \sum_{\underline{n}\in S(I-1,J)} \pi^{(I-1,J)}(\underline{n}) \sum_{j=1}^{J} \eta_j \sum_{i=1}^{J} r_{ji} f_j(\underline{n}) g_i(\underline{n})\\
&\quad - \sum_{\underline{n}\in S(I,J)} \pi^{(I,J)}(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^{J} \mu_j(n_j)\\
&= \frac{G(I-1,J)}{G(I,J)} \mathrm{E}_{\pi^{I-1,J}}[\mathrm{tr}(W^{f,g}\mathrm{diag}(\xi)R)] - \sum_{\underline{n}\in S(I,J)} \pi^{(I,J)}(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^{J} \mu_j(n_j).
\end{aligned}$$

In the first equality we collected terms. The next block in the second equality displays the most explicit form of our relevant equation, which enables us to cancel and shift the summation indices in the third equality. In the fourth equality we used the product form of the Gordon–Newell network equilibrium and applied the definitions of $f_j(\cdot)$ and $g_i(\cdot)$. In the fifth equality we collected definitions of the respective matrices and applied the trace operator. Similarly, we compute

$$\begin{aligned}\langle f, Q^{\tilde{X}} g\rangle_\pi &= \sum_{\underline{n}\in S(I,J)} \pi(\underline{n}) f(\underline{n}) \sum_{\underline{m}\in S(I,J)} Q^{\tilde{X}}(\underline{n},\underline{m}) g(\underline{m}) \\ &= \frac{G(I-1,J)}{G(I,J)} \mathrm{E}_{\pi^{I-1,J}}[\mathrm{tr}(W^{f,g}\mathrm{diag}(\xi)\widetilde{R})] \\ &\quad - \sum_{\underline{n}\in S(I,J)} \pi^{(I,J)}(\underline{n}) f(\underline{n}) g(\underline{n}) \sum_{j=1}^{J} \mu_j(n_j),\end{aligned}$$

which immediately yields the statement of the proposition. From the assumption that the stochastic traffic solutions η (and $\widetilde{\eta}$) of the traffic equation for R and $\widetilde{R}$ coincide, it follows that the remainder terms in both expressions are the same. This completes the proof.

We can reformulate the results of Propositions 4.1 and 4.2 in a form which is of independent interest, because it immediately relates our results to methods dealt with in optimizing MCMC simulation. For convenience, introducing the notation $H^f(\underline{n}, i) := f(\underline{n} + e_i)$, which in our framework occurs as $H^f(X_t, i) := f(X_t + e_i)$ (and similarly for g), we obtain the following corollary.

Corollary 4.1. (a) *For Jackson network processes* $\boldsymbol{X}$ *and* $\tilde{\boldsymbol{X}}$ *as in Proposition 4.1 with* ξ *the probability solution of the extended traffic equation, (3.2), we have*

$$\langle f, Q^{\boldsymbol{X}} g\rangle_\pi - \langle f, Q^{\tilde{\boldsymbol{X}}} g\rangle_\pi = \frac{\lambda}{\xi_0} \mathrm{E}_\pi \langle H^f(X_t, \cdot), (R - \tilde{R}) H^g(X_t, \cdot)\rangle_\xi. \tag{4.4}$$

(b) *For Gordon–Newell network processes* $\boldsymbol{X}$ *and* $\tilde{\boldsymbol{X}}$ *as in Proposition 4.2 with* η *the probability solution of the traffic equation, we have*

$$\langle f, Q^{\boldsymbol{X}} g\rangle_\pi - \langle f, Q^{\tilde{\boldsymbol{X}}} g\rangle_\pi = \frac{G(I-1,J)}{G(I,J)} \mathrm{E}_{\pi^{I-1,J}} \langle H^f(X_t, \cdot), (R - \tilde{R}) H^g(X_t, \cdot)\rangle_\eta. \tag{4.5}$$

There are several appealing interpretations of (4.4) and (4.5) which will guide some of our forthcoming arguments. We discuss the closed network case, (4.5).

The inner product $\langle H^f(X_t, \cdot), \big(R - \tilde{R}\big) H^g(X_t, \cdot)\rangle_\eta$ can be evaluated pathwise for any elementary event and, whenever, e.g. the difference $R - \widetilde{R}$ is positive definite, the integral $\mathrm{E}_{\pi^{I-1,J}}(\cdot)$ (across Ω) is over nonnegative functions. Recalling that η is stationary for R and $\widetilde{R}$, we obtain

$$\begin{aligned}&\langle H^f(X_t, \cdot), (R - \tilde{R}) H^g(X_t, \cdot)\rangle_\eta \\ &\qquad = \mathrm{E}_\eta[H^f(X_t, V_0) H^g(X_t, V_1)] - \mathrm{E}_\eta[H^f(X_t, \widetilde{V}_0) H^g(X_t, \widetilde{V}_1)],\end{aligned}$$

where $\boldsymbol{V} = (V_n : n \in \mathbb{N})$ and $\widetilde{\boldsymbol{V}} = (\widetilde{V}_n : n \in \mathbb{N})$ are Markov (routeing) chains with common steady state η and different transition matrices R and $\widetilde{R}$. If we consider formally a network process $\boldsymbol{X}$ and Markov chains $\boldsymbol{V}$ and $\widetilde{\boldsymbol{V}}$ that are independent of $\boldsymbol{X}$, we obtain

$$\begin{aligned}\langle f, Q^{\boldsymbol{X}} g\rangle_\pi - \langle f, Q^{\widetilde{\boldsymbol{X}}} g\rangle_\pi &= \frac{G(I-1,J)}{G(I,J)} (\mathrm{E}_{\pi^{I-1,J}}[\mathrm{E}_\eta[H^f(X_t, V_0) H^g(X_t, V_1)]] \\ &\qquad - \mathrm{E}_{\pi^{I-1,J}}[\mathrm{E}_\eta[H^f(X_t, \widetilde{V}_0) H^g(X_t, \widetilde{V}_1)]]) \\ &= \frac{G(I-1,J)}{G(I,J)} (\mathrm{E}_\eta[\mathrm{E}_{\pi^{I-1,J}}[H^f(X_t, V_0) H^g(X_t, V_1)]] \\ &\qquad - \mathrm{E}_\eta[\mathrm{E}_{\pi^{I-1,J}}[H^f(X_t, \widetilde{V}_0) H^g(X_t, \widetilde{V}_1)]]),\end{aligned}$$

where the latter equality follows by Fubini. The last expression is a representation through stochastically ordered processes, and applies whenever we can show that the difference between the covariances is nonnegative or nonpositive throughout.

Corollary 4.1 points out the relevance of the following orderings for transition matrices which are well known in the theory of optimal selection of transition kernels for MCMC simulation. In our investigations these orders will be utilized to compare routeing processes via their transition matrices.

Definition 4.1. Let $R = [r_{ij}]$ and $\widetilde{R} = [\widetilde{r}_{ij}]$ be transition matrices on a finite set $\mathbb{E}$ such that $\eta R = \eta \widetilde{R} = \eta$.

We say that $\widetilde{R}$ is smaller than R in the positive definite order, $\widetilde{R} \prec_{\mathrm{pd}} R$, if their difference $R - \widetilde{R}$ is positive definite on $L_2(\mathbb{E}, \eta)$.

We say that $\widetilde{R}$ is smaller than R in the Peskun order, $\widetilde{R} \prec_{\mathrm{P}} R$, if, for all $j, i \in \mathbb{E}$ with $i \neq j$, $\widetilde{r}_{ji} \leq r_{ji}$ holds; see [20].

Peskun used the latter order to compare reversible transition matrices with the same stationary distribution and their asymptotic variance, and Tierney [23] showed that the main property used in the proof of Peskun, namely that $R \prec_{\mathrm{P}} \widetilde{R}$ implies that $\widetilde{R} \prec_{\mathrm{pd}} R$, holds without reversibility assumptions.

4.1. Applications

4.1.1. *Asymptotic variance.* Peskun and Tierney derived comparison theorems for the asymptotic variance of Markov chains for application to optimal selection of MCMC transition kernels in discrete time. These asymptotic variances occur as variance in the limiting distribution of central limit theorems (CLTs) for the MCMC estimators.

In the setting of queueing networks, performance measures of interest are usually steady-state mean values of performance indices, $\pi(f) = \mathrm{E}_\pi[f(X_t)]$, which can be estimated as time averages, justified by the ergodic theorem for Markov processes, i.e. in discrete time we have, for large n,

$$\mathrm{E}_\pi[f(X_t)] \sim \frac{1}{n} \sum_{k=1}^{n} f(X_k).$$

Under some regularity conditions on a homogeneous Markov chain with one-step transition kernel K, there is a CLT of the form

$$\sqrt{n}\left(\frac{1}{n} \sum_{k=1}^{n} f(X_k) - \mathrm{E}_\pi[f(X_t)]\right) \xrightarrow{\mathrm{w}} N(0, v(f, K))$$

(where '$\xrightarrow{\mathrm{w}}$' denotes weak convergence), where the asymptotic variance is

$$v(f, K) = \langle f, f \rangle_\pi - \pi(f) + 2 \sum_{k=1}^{\infty} \langle f, K^k f \rangle_\pi.$$

To arrange a discrete-time framework for our network processes $\boldsymbol{X}$, we consider the Markov chains with transition matrices

$$K = I + \varepsilon Q^{\boldsymbol{X}}$$

(with sufficiently small $\varepsilon > 0$) that occur in the compound Poisson representation of the transition probabilities of the network processes.

Proposition 4.3. (a) *Consider ergodic Jackson networks with the same arrival and service intensities, and with queue length processes $\boldsymbol{X}$ and $\widetilde{\boldsymbol{X}}$. Assume that the* extended *routeing matrices R and $\widetilde{R}$ are reversible with respect to ξ.*

If R and $\widetilde{R}$ are ordered in the Peskun order, $\widetilde{R} \prec_{\mathrm{P}} R$, then, for any real function f, we have

$$v(f, I + \varepsilon Q^{\tilde{\boldsymbol{X}}}) \geq v(f, I + \varepsilon Q^{\boldsymbol{X}}).$$

(b) *Consider ergodic Gordon–Newell networks with the same service intensities and with queue length processes $\boldsymbol{X}$ and $\widetilde{\boldsymbol{X}}$. Assume that the routeing matrices R and $\widetilde{R}$ are reversible with respect to η.*

If R and $\widetilde{R}$ are ordered in the Peskun order, $\widetilde{R} \prec_{\mathrm{P}} R$, then, for any real function f, we have

$$v(f, I + \varepsilon Q^{\tilde{\boldsymbol{X}}}) \geq v(f, I + \varepsilon Q^{\boldsymbol{X}}).$$

Proof. As the proofs of (a) and (b) follow the same lines, we sketch only (a). Because the routeing matrices are reversible, the network processes are reversible as well. The local balance equations with respect to π are therefore fulfilled by $Q^{\tilde{\boldsymbol{X}}}$ and $Q^{\boldsymbol{X}}$. This immediately yields the fact that the transition matrices $I + \varepsilon Q^{\boldsymbol{X}}$ and $I + \varepsilon Q^{\tilde{\boldsymbol{X}}}$ are reversible with respect to π. We can therefore apply Theorem 4 of [23] for (a). For the Gordon–Newell network with finite state space, we apply Peskun's theorem [20]. This completes the proof.

4.1.2. *Comparison of spectral gaps.* Let $\boldsymbol{X}$ be a continuous-time homogeneous ergodic Markov process with stationary probability π and generator $Q^{\boldsymbol{X}}$. The spectral gap of $Q^{\boldsymbol{X}}$ is

$$\operatorname{gap}(Q^{\boldsymbol{X}}) = \inf\{\langle f, -Q^{\boldsymbol{X}} f\rangle_{\pi} : f \in L_2(\mathbb{E}, \pi),\ \pi(f) = 0, \langle f, f\rangle\pi = 1\}.$$

The spectral gap determines, for $\boldsymbol{X}$, the speed of convergence to equilibrium π in the $L_2(\mathbb{E}, \pi)$-norm $\|\cdot\|_{\pi}$: $\operatorname{gap}(Q^{\boldsymbol{X}})$ is the largest number Δ such that, for the transition semigroup $P = (P_t : t \geq 0)$ of $\boldsymbol{X}$,

$$\|P_t f - \pi(f)\|_{\pi} \leq \mathrm{e}^{-\Delta t}\|f - \pi(f)\|_{\pi} \quad \text{for all } f \in L_2(\mathbb{E}, \pi)$$

holds.

For Gordon–Newell networks, the spectral gap is always greater than 0, while, for Jackson networks, the situation is more delicate: zero gaps and nonzero gaps can occur. Iscoe and McDonald [8], [9] and Lorek [15] proved, under some natural assumptions, necessary and sufficient conditions for nonzero spectral gaps in Jackson networks. The case of positive gaps is proved by using an attached vector of independent birth–death processes to bound the gap away from 0.

We show that, for some classes of Jackson networks, we can even strictly bound the gap of the queue length network process $\boldsymbol{X}$ from below by the gap of some multidimensional birth–death process, which in the next proposition will be the network process $\widetilde{\boldsymbol{X}}$. Because we focus on the intuitive but rather strong Peskun ordering of the routeing matrices, we need additional assumptions on the routeing. These assumptions constitute a detailed balance which determines an additional internal structure of a Markov chain and its global balance equation (equilibrium equation). Such detailed balance equations are prevalent in many networks with (nearly) product-form steady states, and often open the way to solve the global balance equation for the steady state. Equation (4.6), below, equalizes the routeing flow from any node into the (inner) network to the flow out of the (inner) network to that node.

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

Proposition 4.4. *Consider an ergodic Jackson network process* $\boldsymbol{X}$ *with state-dependent service intensities and with positive external arrival rates* $\lambda_i > 0$ *at all nodes* $i = 1, \dots, J$. *Assume that the extended routeing matrix* $R = [r_{ij}]_{i,j=0,1,\dots,J}$ *has strict positive departure probabilities* $r_{i0} > 0$ *from every node* $i = 1, \dots, J$.

Furthermore, assume that the routeing of $\boldsymbol{X}$ *fulfills overall balance for all network nodes with respect to the solution* η_i, $i = 1, \dots, J$, *of the traffic equation, (3.1), i.e.*

$$\eta_j \sum_{i=1}^{J} r_{j,i} = \sum_{i=1}^{J} \eta_i r_{i,j} \quad \textit{for all } j = 1, \dots, J. \tag{4.6}$$

Then there exists an ergodic Jackson network process $\widetilde{\boldsymbol{X}}$ *of independent birth–death processes, the nodes of which have the same service intensities and external arrival rates* $\widetilde{\lambda}_i = \lambda_i$, *such that* $\operatorname{gap}(Q^{\boldsymbol{X}}) \geq \operatorname{gap}(Q^{\widetilde{\boldsymbol{X}}})$.

Proof. We define the ith birth–death process by $\widetilde{\lambda}_i = \lambda_i/\lambda$, $\widetilde{r}_{i0} = r_{i0}$, and $\widetilde{r}_{ii} = 1 - r_{i0}$.

Obviously, we have $\widetilde{R} \prec_{\mathrm{P}} R$ and, therefore, $R \prec_{\mathrm{pd}} \widetilde{R}$ from [23, Lemma 3]. Thus, for real functions $f \in L_2(\mathbb{E}, \pi)$ with $\pi(f) = 0$ and $\langle f, f \rangle_\pi = 1$, it follows from Corollary 4.1(a) that

$$0 \leq \langle f, -Q^{\widetilde{\boldsymbol{X}}} f \rangle_\pi \leq \langle f, -Q^{\boldsymbol{X}} f \rangle_\pi$$

whenever R and $\widetilde{R}$ have the same solution ξ of the extended traffic equation. Given this, the infima on both sides are ordered as well; so, by the definition of the spectral gaps, they are ordered in the same direction.

It remains to show that $\xi \widetilde{R} = \xi$, which can be seen directly: for $j \in \{1, \dots, J\}$, we have $\widetilde{\eta}_j = \widetilde{\lambda}_j + \widetilde{\eta}_j \widetilde{r}_{jj}$, and the solution of this system is uniquely defined. But, from (4.6) we find, via

$$\eta_j = \lambda_j + \sum_{i=1}^{J} \eta_i r_{i,j} = \lambda_j + \eta_j \sum_{i=1}^{J} r_{j,i} = \lambda_j + \eta_j (1 - r_{j0}) \quad \text{for all } j = 1, \dots, J$$

with $1 - r_{j0} = \widetilde{r}_{ii}$, that $\widetilde{\eta}_j = \eta_j$ for all $j = 1, \dots, J$ holds, and from the definition we have $\widetilde{\eta}_0 = \lambda = \eta_0$. This completes the proof.

Extending this proposition to a more general setting we immediately obtain, from (4.4) and (4.5), correlation inequalities which bound (4.2). So, we can immediately conclude for some networks that $\operatorname{gap}(Q^{\widetilde{\boldsymbol{X}}}) \leq \operatorname{gap}(Q^{\boldsymbol{X}})$ holds. A consequence of the fact that *Peskun yields positive definiteness* is that if we perturb routeing of customers in the networks by shifting mass from nondiagonal entries to the diagonal (leaving the routeing equilibrium fixed), then the speed of convergence of the perturbed process can only decrease. This is just what was intended in the optimization of MCMCs, and Peskun gave conditions for this. Similarly, we obtain the following proposition.

Proposition 4.5. (a) *Consider ergodic Jackson networks with the same arrival and service intensities, and with state processes* $\boldsymbol{X}$ *and* $\widetilde{\boldsymbol{X}}$. *Assume that, for the extended routeing matrices* R *and* $\widetilde{R}$, *the stochastic solutions* ξ *of the traffic equation coincide. If* R *and* $\widetilde{R}$ *are ordered in the positive definite order,* $R \prec_{\mathrm{pd}} \widetilde{R}$, *then, for any real function* f, *we have*

$$\langle f, Q^{\widetilde{\boldsymbol{X}}} f \rangle_\pi \geq \langle f, Q^{\boldsymbol{X}} f \rangle_\pi \quad \textit{and} \quad \operatorname{gap}(Q^{\widetilde{\boldsymbol{X}}}) \leq \operatorname{gap}(Q^{\boldsymbol{X}}).$$

(b) *Consider ergodic Gordon–Newell networks with the same service intensities and state processes* $\boldsymbol{X}$ *and* $\widetilde{\boldsymbol{X}}$. *Assume that, for the routeing matrices* R *and* $\widetilde{R}$, *the stochastic solutions* η *of the traffic equation coincide. If* R *and* $\widetilde{R}$ *are ordered in the positive definite order,* $R \prec_{\mathrm{pd}} \widetilde{R}$, *then, for any real function* f, *we have*

$$\langle f, Q^{\tilde{X}} f\rangle_\pi \geq \langle f, Q^{X} f\rangle_\pi \quad \textit{and} \quad \mathrm{gap}(Q^{\widetilde{X}}) \leq \mathrm{gap}(Q^{X}).$$

4.1.3. *Comparison of dependencies.* Expression (4.1) for continuous-time Markov processes is transformed via the embedded uniformization chain, (4.3), to a true covariance and via (4.2) to a comparison of covariance functionals for two Markov processes and their at Poissonian times embedded chains, i.e. with $\tau \sim \exp(\eta)$ we obtain

$$\mathrm{E}_\pi[f(X_0)g(X_\tau)] = \langle f, (I + \eta Q^{X})g\rangle_\pi \leq \langle f, (I + \eta Q^{\widetilde{X}})g\rangle_\pi = \mathrm{E}_\pi[f(\widetilde{X}_0)g(\widetilde{X}_\tau)].$$

Transforming this into statements in the continuous-time setting will need in general additional monotonicity properties of the processes. It turns out that monotonicity is in some cases a direct substitute for the strong reversibility assumption which is needed to prove Peskun's theorem; see Corollary 7.1 and Theorem 7.1, below.

Comparison of dependencies is the central point of Sections 5 and 6. These will strongly utilize the subsequent parts of this section.

4.2. Doubly stochastic routeing: increasing chaos and correlation inequalities via generators

In this section the perturbation of the network process is due to the routeing of the customers becoming more chaotic. In statistical physics there is a well-established method to express *chaotic* behavior of a random walker, if his itinerary is governed by doubly stochastic routeing matrices. Alberti and Uhlmann [1, Chapter 1] provided an indepth study of stochasticity and partial order that elaborates on these methods. Following their ideas in this section, we consider (mainly) Gordon–Newell networks with doubly stochastic routeing matrices.

The method used to transform doubly stochastic routeing and, thus, to classify derived perturbed routeing processes with respect to the amount of chaoticity is as follows.

Consider an arbitrary row $r(i) := (r_{ij} : j = 1, 2, \ldots, J)$ of the Gordon–Newell network's routeing matrix R and a doubly stochastic matrix $T = [t_{ij}]_{i,j=1,\ldots,J}$. Then the ith row vector of the product RT is smaller than $r(i)$ in the sense of the majorization ordering; see [13, p. 18]. This means that the probability mass is more equally distributed in each row after multiplication. The routeing scheme is then more equally distributed too. Nevertheless, the solution of the traffic equation for RT and, therefore, the steady state of the network under the RT regime is the same as under R, namely, the normalized solution of the traffic equation, (3.4), is in both cases the uniform distribution on $\{1, 2, \ldots, J\}$.

An extremal situation is when all the rows of R are identically distributed, which corresponds to 'independent routeing'. Moreover, if the rows are uniformly distributed, this reflects the *most chaotic* routeing behavior.

It is tempting to conclude that more chaotic routeing leads to less internal dependencies over time of the individual routeing chains of the customers and will, therefore, lead to less internal dependence over time of the joint queue length process. This should be visible by the occurrence of inequalities for (4.2) and will be exploited below. Let

$$\mathcal{L} = \left\{ f : S(I, J) \to \mathbb{R}_+ : f(n_1, \ldots, n_J) = a + \sum_{i=1}^{J} \alpha_i n_i,\ \alpha_i \in \mathbb{R},\ i = 1, \ldots, J,\ a \in \mathbb{R}_+ \right\}$$

denote the convex cone of nonnegative affine-linear functions on $S(I, J)$. The following theorem is a prototype of the correlation theorems we are interested in.

Theorem 4.1. (Linear service rates.) *Consider two ergodic Gordon–Newell network processes with common stationary distribution* π*:* X *with a doubly stochastic routeing matrix* R *and* $\widetilde{X}$ *with the routeing matrix* $\widetilde{R} = [\widetilde{r}_{ij}] = RT$ *for a doubly stochastic matrix* $T = [t_{ij} : i, j = 1, \dots, J]$. *All other parameters of the networks are assumed to be the same.*

Consider pairs of nonnegative affine-linear functions:

$$f : S(I, J) \to \mathbb{R}_+ : f(n_1, \dots, n_J) = a + \sum_{i=1}^{J} \alpha_i n_i \in \mathcal{L}$$

$$\text{and} \quad g : S(I, J) \to \mathbb{R}_+ : g(n_1, \dots, n_J) = b + \sum_{i=1}^{J} \beta_i n_i \in \mathcal{L}$$

with

$$\mathcal{R}(\alpha_1, \dots, \alpha_J) = \mathcal{R}(\beta_1, \dots, \beta_J).$$

Then, for all such pairs of functions with $f, g \in \mathcal{I}_+^*(\mathbb{N}^J) \cap \mathcal{L}$ *and* $f, g \in \mathcal{D}_+^*(\mathbb{N}^J) \cap \mathcal{L}$,

$$\langle f, Q^{\tilde{X}} g \rangle_\pi \le \langle f, Q^{X} g \rangle_\pi$$

holds.

Proof. The proof uses properties of majorization ordering for the coefficient vectors and needs some technical requisites.

1. We first assume that the coefficients of g are positive, i.e. $\beta_j > 0,\ j = 1, \dots, J$. The proof of Proposition 4.2 shows that, for f and g, we have to evaluate

$$\begin{aligned}
&\langle f, Q^{X} g \rangle_\pi - \langle f, Q^{\tilde{X}} g \rangle_\pi \\
&\quad = \frac{G(I-1, J)}{G(I, J)} \sum_{\underline{n} \in S(I-1, J)} \pi^{(I-1,J)}(\underline{n}) \biggl(\sum_{i=1}^{J} \sum_{j=1}^{J} \eta_j r_{ji} f_j(\underline{n}) g_j(\underline{n}) \\
&\qquad\qquad - \sum_{i=1}^{J} \sum_{j=1}^{J} \eta_j \widetilde{r}_{ji} f_j(\underline{n}) g_j(\underline{n}) \biggr).
\end{aligned}$$

For $f(\underline{n}) = a + \sum_{i=1}^{J} \alpha_i n_i$ and $g(\underline{n}) = b + \sum_{i=1}^{J} \beta_i n_i$, we have

$$\begin{aligned}
&\sum_{\underline{n} \in S(I-1, J)} \pi^{(I-1,J)}(\underline{n}) \sum_{i=1}^{J} \sum_{j=1}^{J} \eta_j r_{ji} f_j(\underline{n}) g_j(\underline{n}) \\
&\quad = \sum_{i=1}^{J} \sum_{j=1}^{J} \eta_j r_{ji} \, \mathrm{E}_{\pi^{(I-1,J)}}[(f + \alpha_j)(g + \beta_i)] \\
&\quad = \sum_{i=1}^{J} \sum_{j=1}^{J} \eta_j r_{ji} (\mathrm{E}_{\pi^{(I-1,J)}}[fg] + \beta_i \, \mathrm{E}_{\pi^{(I-1,J)}}[f] + \alpha_j \, \mathrm{E}_{\pi^{(I-1,J)}}[g] + \alpha_j \beta_i),
\end{aligned}$$

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

and computing the parallel expression for $\widetilde{X}$ we obtain

$$\langle f, Q^X g\rangle_\pi - \langle f, Q^{\widetilde{X}} g\rangle_\pi = \frac{G(I-1,J)}{G(I,J)}\Bigg(\sum_{j=1}^J \eta_j \sum_{i=1}^J r_{ji}\alpha_j\beta_i - \sum_{j=1}^J \eta_j \sum_{i=1}^J \bigg(\sum_{k=1}^J r_{jk}t_{ki}\bigg)\alpha_j\beta_i\Bigg). \tag{4.7}$$

To show that (4.7) is nonnegative, we first observe that $\eta_j = J^{-1},\ j=1,\dots,J$, such that we have to show that

$$\sum_{j=1}^J \alpha_j \sum_{i=1}^J r_{ji}\beta_i - \sum_{j=1}^J \alpha_j \sum_{i=1}^J \bigg(\sum_{k=1}^J r_{jk}t_{ki}\bigg)\beta_i \geq 0. \tag{4.8}$$

Denoting the transpose of a matrix M by $M^\top$, we rewrite (4.8) as

$$\sum_{j=1}^J \alpha_j(\beta R^\top)_j - \sum_{j=1}^J \alpha_j(\beta T^\top R^\top)_j \geq 0,$$

and observe that, with respect to the majorization order '$\prec$', we have (see [18, Theorem 1.5.34]) $(\beta T^\top) \prec \beta$, and, therefore,

$$(\beta T^\top R^\top) \prec \beta R^\top.$$

The rank vectors of the vectors $\alpha = (\alpha_1,\dots,\alpha_J)$ and $\beta = (\beta_1,\dots,\beta_J)$ are the same from the assumptions and, by the principle of equalizing mass transfer, by applying doubly stochastic transformations, it follows that the order statistics of $(\beta T^\top)$ and $(\beta T^\top R^\top)$ are the same as those of β and, therefore, of α. From this, it follows, for the vector $(\alpha_{[1]},\dots,\alpha_{[J]})$, which is the decreasing rearrangement of the vector α, that

$$\sum_{j=1}^J \alpha_j(\beta R^\top)_j = \sum_{j=1}^J \alpha_{[j]}(\beta R^\top)_{[j]} \tag{4.9}$$

and

$$\sum_{j=1}^J \alpha_j(\beta T^\top R^\top)_j = \sum_{j=1}^J \alpha_{[j]}(\beta T^\top R^\top)_{[j]}. \tag{4.10}$$

The right-hand sides of (4.9) and (4.10) are integrals of the *decreasing function* $i \to \alpha_{[i]},\ i = 1,\dots,J$, with respect to the counting densities

$$i \to (\beta R^\top)_{[i]},\quad i=1,\dots,J,\quad \text{and, respectively,}\quad i \to (\beta T^\top R^\top)_{[i]},\quad i=1,\dots,J. \tag{4.11}$$

(Here we need the positivity of β to perform the next simple step, the positivity of α is not required.)

From majorization ordering, these counting densities fulfill

$$\sum_{i=1}^k (\beta R^\top)_{[i]} \geq \sum_{i=1}^k (\beta T^\top R^\top)_{[i]} \quad \text{for all } k=1,\dots,J-1$$

$$\text{and}\quad \sum_{i=1}^J (\beta R^\top)_{[i]} = \sum_{i=1}^J (\beta T^\top R^\top)_{[i]},$$

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

which says that the associated (suitably normalized) probabilities are strongly stochastically ordered. It follows that

$$\sum_{j=1}^{J}\alpha_j(\beta R^\top)_j=\sum_{j=1}^{J}\alpha_{[j]}(\beta R^\top)_{[j]}\geq\sum_{j=1}^{J}\alpha_{[j]}(\beta T^\top R^\top)_{[j]}=\sum_{j=1}^{J}\alpha_j(\beta T^\top R^\top)_j,$$

which is (4.8).

2. If β has negative components, we add, componentwise, some $K>0$ such that $\beta_j^\top=\beta_j+K>0,\ j=1,\dots,J$. Then the conclusion in (4.11) holds with $\beta^\top$. But it is obvious that the transformation $\beta\to\beta^\top$ leaves the essential equation (4.8) invariant. This completes the proof.

In Theorem 4.1, for $f=g$, the rank condition is trivially fulfilled. This yields the following corollary.

Corollary 4.2. *Under the assumptions of Theorem 4.1, for all* $f\in\mathcal{I}_+^*(\mathbb{N}^J)\cap\mathcal{L}$ *and* $f\in\mathcal{D}_+^*(\mathbb{N}^J)\cap\mathcal{L}$, *it holds that*

$$\langle f,Q^{\widetilde{X}}f\rangle_\pi\leq\langle f,Q^Xf\rangle_\pi.$$

Note that, for $f=g$, the proof of Theorem 4.1 shows that $R(I-T)$ is nonnegative definite.

For the Jackson networks in this section, we assume that the extended routeing matrix R is doubly stochastic. Then the probability solution of the extended traffic equation $xR=x$, (3.2), is the uniform distribution on $\{0,1,\dots,J\}$. It follows that the traffic equation, (3.1), has the unique solution $\eta(i)=\lambda,\ i\in\{1,2,\dots,J\}$. We further assume that the doubly stochastic transformation matrix $T=[t_{ij}]_{i,j=0,1,\dots,J}$ additionally fulfills the condition $t_{00}=1$ in order to keep the solution of the modified traffic equation. The traffic equation for the network under the $\widetilde{R}=RT$ regime is

$$\widetilde{\eta}_j=\sum_{k=1}^{J}\lambda_kt_{kj}+\sum_{i=1}^{J}\widetilde{\eta}_i\sum_{k=1}^{J}r_{ik}t_{kj},\qquad j=1,\dots,J.$$

The external arrival rates are changed to $\widetilde{\lambda}_j:=\sum_{k=1}^{J}\lambda_kt_{kj},\ j=1,\dots,J$, and the total arrival rate to the system therefore remains the same. Moreover, $\widetilde{\Lambda}=(\widetilde{\lambda}_1,\dots,\widetilde{\lambda}_J)$ is smaller in the majorization order than Λ (i.e. it is more equally distributed). Furthermore, the departure probabilities remain unchanged: $r_{j0}=\widetilde{r}_{j0},\ j=1,\dots,J$.

The next theorem contributes to calculating performance measures which depend only on the total population size of the network, which is relevant, for example, in busy period analysis. The result does not seem to be intuitive because it reveals a surprising insensitivity property of the networks. The proof goes by a direct but lengthy computation.

Theorem 4.2. *Consider two ergodic Jackson network processes with common stationary distribution* π: $\boldsymbol{X}$ *with a doubly stochastic routeing matrix* R *and* $\widetilde{\boldsymbol{X}}$ *with the routeing matrix* $\widetilde{R}=[\widetilde{r}_{ij}]=RT$ *for a doubly stochastic matrix* T *such that* $t_{00}=1$. *All other parameters of the networks are assumed to be the same. Let*

$$\mathcal{F}=\{f:\mathbb{N}^J\to\mathbb{R}\colon f(n_1,\dots,n_J)=\hat{f}(n_1+\cdots+n_J)\text{ for some }\hat{f}\colon\mathbb{R}\to\mathbb{R}_+\}$$

be the set of nonnegative real-valued functions on $\mathbb{N}^J$, *which depend on the arguments only through their sum.*

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

Then, for all such pairs of functions $f, g \in \mathcal{I}_+^*(\mathbb{N}^J) \cap \mathcal{F}$ and $f, g \in \mathcal{D}_+^*(\mathbb{N}^J) \cap \mathcal{F}$, it holds that

$$\langle f, Q^{\widetilde{X}} g\rangle_\pi = \langle f, Q^X g\rangle_\pi .$$

4.3. Robin-Hood transforms: increasing concordance and correlation inequalities via generators

In Subsection 4.2 we considered different degrees of chaoticity for the routeing of customers and some consequences thereof. We did not ask for details of the structure of the random walk performed by the traveling customers. This is justified in many cases; in many networks, however, these random walks are more structured and, clearly, we should incorporate such prior knowledge into the performance assessment of the system. For example, a linear tandem network is an extreme case where an order of the nodes is prescribed and is of importance for the movements of the customers.

Order structures of random walks can usually be described by an underlying (directed or undirected) graph, which is often connected with a partial order on the set of nodes. In the tandem network we have a total order on the node set, which completely determines the customers' feasible movements.

If the node set is equipped with a partial order, which is relevant for the customers' migration, then it is tempting to consider perturbations of the routeing processes that are in line with this order. To be more precise, we have an up–down relation between the nodes and the question is how the steady-state performance reacts on routeing more up or down.

The construction of Corollary 2.1 and Example 3.1 of [4], which is sometimes called the Robin-Hood transform because in a certain sense it equalizes the frequencies of the random walker to visit the different nodes, yields a change of routeing such that it is more or less dependent in a well-defined way. The construction is as follows.

Consider some homogeneous Markov chain $\boldsymbol{X}$ on the ordered state space $(\mathbb{E}, \prec)$ with transition matrix p in equilibrium with the steady state π.

Assume that, for $a, b, c, d \in \mathbb{E}$, we have $a \prec c$ and $b \prec d$ such that $(a, d) \in \mathbb{E}^2$ and $(c, b) \in \mathbb{E}^2$ are not comparable with respect to the product order, and that $\mathrm{P}^{(X_0,X_1)}(a, d) \geq \alpha$ and $\mathrm{P}^{(X_0,X_1)}(c, b) \geq \alpha$.

Construct the distribution $\mathrm{P}^{(Y_0,Y_1)}$ of a random vector (Y_0, Y_1) from $\mathrm{P}^{(X_0,X_1)}$ by

$$\begin{aligned} \mathrm{P}^{(Y_0,Y_1)}(a, b) &= \mathrm{P}^{(X_0,X_1)}(a, b) + \alpha, \qquad & \mathrm{P}^{(Y_0,Y_1)}(c, d) &= \mathrm{P}^{(X_0,X_1)}(c, d) + \alpha, \\ \mathrm{P}^{(Y_0,Y_1)}(a, d) &= \mathrm{P}^{(X_0,X_1)}(a, d) - \alpha, \qquad & \mathrm{P}^{(Y_0,Y_1)}(c, b) &= \mathrm{P}^{(X_0,X_1)}(c, b) - \alpha, \end{aligned}$$
$$\mathrm{P}^{(Y_0,Y_1)}(u, v) = \mathrm{P}^{(X_0,X_1)}(u, v) \quad \text{for all other } (u, v) \in \mathbb{E}^2.$$

(This is the Robin-Hood transform.)

The one-dimensional marginals of both (X_0, X_1) and (Y_0, Y_1) are π and the conditional distribution $\mathrm{P}(Y_1 = w \mid Y_0 = v) =: q(v, w)$ for $v, w \in \mathbb{E}$ is obtained from p as follows:

$$\begin{aligned} q(a, d) &= p(a, d) - \frac{\alpha}{\pi(a)}, \qquad & q(c, b) &= p(c, b) - \frac{\alpha}{\pi(c)}, \\ q(a, b) &= p(a, b) + \frac{\alpha}{\pi(a)}, \qquad & q(c, d) &= p(c, d) + \frac{\alpha}{\pi(c)}, \end{aligned} \tag{4.12}$$
$$q(u, v) = p(u, v) \quad \text{otherwise.}$$

Now consider a homogeneous Markov chain $\boldsymbol{Y}$ with the so-defined transition matrix q, and consider $\boldsymbol{X}$ and $\boldsymbol{Y}$ as routeing chains of a network process, where $\boldsymbol{Y}$ is obtained from $\boldsymbol{X}$ by a perturbation through the Robin-Hood transformation. Then, according to Corollary 2.1 and Theorem 3.1 of [4], the routeing governed by $\boldsymbol{Y}$ is more concordant than the routeing governed by $\boldsymbol{X}$. (We will generalize this theorem in Section 5, below, so as to make it appropriate for our network processes.)

From the construction, the transition kernel q obtained by the Robin-Hood transformation is more strongly connected with the ordering of the nodes then the kernel p. This is also expressed in the mentioned cc-ordering of the respective two-dimensional vectors: more weight is given to $(a, b) \prec^2 (a, d)$ and $(b, c) \prec^2 (c, d)$.

It is therefore tempting to conjecture that, for a pair of Jackson network processes $\widetilde{\boldsymbol{X}}$ with routeing q and $\boldsymbol{X}$ with routeing p, correlation inequalities like (4.2) should occur if the ordering of the state space $\mathbb{N}^J$ respects the ordering '$\prec$' of the nodes as well. The latter statement can be given the following precise meaning.

Definition 4.2. Let $(\mathbb{E}, \prec)$ be a countable partially ordered set. The generalized partial sum order '$\prec_*$' on $\mathbb{N}^{\mathbb{E}}$ is defined, for $x = (x_i : i \in \mathbb{E})$ and $y = (y_i : i \in \mathbb{E}) \in \mathbb{N}^{\mathbb{E}}$, by

$$x \prec_* y \quad \Longleftrightarrow \quad \text{for all decreasing } K \subseteq \mathbb{E}, \sum_{k \in K} x_k \leq \sum_{k \in K} y_k \text{ holds.}$$

The order '$\prec_*$' is indeed a partial order because reflexivity and transitivity are immediate, and antisymmetry can be seen as follows. Denote $\{i\}^{\downarrow} = \{j \in \mathbb{E} \colon j \prec i\}$. For $x \in \mathbb{N}^{\mathbb{E}}$, we find from $x \prec_* y \wedge y \prec_* x$ that, for $i \in \mathbb{E}$, $\sum_{k \in \{i\}^{\downarrow}} x_k = \sum_{k \in \{i\}^{\downarrow}} y_k$ holds. Because $\{i\}^{\downarrow} - \{i\}$ is decreasing, we have $\sum_{k \in \{i\}^{\downarrow} - \{i\}} x_k = \sum_{k \in \{i\}^{\downarrow} - \{i\}} y_k$, which yields $x_i = y_i$.

Now consider a Jackson network where the node set $\mathbb{E} = \tilde{J} = \{1, \ldots, J\}$ is a partially ordered set $(\tilde{J}, \prec)$ and the customers flow in line with the directions prescribed by this partial order, i.e. for the routeing matrix $R = [r_{i,j}]_{i,j \in \tilde{J}}$ (see [6])

$$r(i, j) > 0 \quad \Longrightarrow \quad i \prec j \vee j \prec i$$

holds. Then the Jackson network process $\boldsymbol{X} = (X_t \colon t \geq 0)$ has the up–down property on the state space $\mathbb{N}^J$ with respect to '$\prec_*$', which means that, for the generator $Q^{\boldsymbol{X}} = (Q^{\boldsymbol{X}}(x, y) \colon x, y \in \mathbb{N}^{\tilde{J}})$ of $\boldsymbol{X}$,

$$Q^{\boldsymbol{X}}(x, y) > 0 \quad \Longrightarrow \quad x \prec_* y \vee y \prec_* x$$

holds. The proof follows by directly checking the required inequalities.

A case of special interest for this ordering is the partial sum order, which has been studied in the literature. This order on $\mathbb{R}^J$ is defined for a linear order on the node set $\{1, 2, \ldots, J\}$ (index set of respective vectors) by, for $\underline{x} = (x_1, \ldots, x_J)$ and $\underline{y} = (y_1, \ldots, y_J) \in \mathbb{R}^J$,

$$\underline{x} \leq_* \underline{y} \quad \Longleftrightarrow \quad \sum_{i=1}^{k} x_i \leq \sum_{i=1}^{k} y_i \quad \text{for all } k = 1, \ldots, J. \tag{4.13}$$

We begin with a fundamental lemma which is of independent interest. It will readily establish the main theorem, but yields even more advantages because of its relevance to further special networks.

Lemma 4.1. *Consider an ergodic Jackson network with extended routeing matrix*

$$R = [r_{i,j}]_{i,j=0,1,\ldots,J}$$

defined in (3.2) and queue length process $\boldsymbol{X}$. We assume that the node set $\tilde{J} = \{1, \dots, J\}$ is a partially ordered set $(\tilde{J}, \prec)$. For some nodes $a, b, c, d \in \tilde{J}$ (not necessarily distinct), let $a \prec c$ and $b \prec d$, and, for some $\alpha > 0$, let

$$r(a, d) \geq \frac{\alpha}{\pi(a)} \quad \textit{and} \quad r(c, b) \geq \frac{\alpha}{\pi(c)}.$$

Define a new network with queue length process $\widetilde{\boldsymbol{X}}$ as follows. The nodes, the nodes' structure, and the external arrival processes are the same as in the original network. The routeing matrix $\widetilde{R} = [\widetilde{r}_{i,j}]_{i,j=0,1,\dots,J}$ is computed by the Robin-Hood transformation, (4.12), with fixed $a, b, c, d \in \tilde{J}$.

Consider a pair of comonotone functions $f, g \colon \mathbb{N}^{\tilde{J}}$ (either both increasing or both decreasing) such that, for all $n \in \mathbb{N}$, $(f(n + e_c) - f(n + e_a))(g(n + e_d) - g(n + e_b)) \geq 0$ holds.
Then

$$\langle f, Q^{\boldsymbol{X}} g\rangle_\pi \leq \langle f, Q^{\widetilde{\boldsymbol{X}}} g\rangle_\pi. \tag{4.14}$$

Proof. Because the construction of the new routeing matrix in (4.12) leaves the stationary distribution of the associated Markov chains invariant, it follows that the solution of the extended traffic equations of the networks fulfill $\tilde{\eta}_i = \eta_i$, $i = 0, 1, \dots, J$, and, therefore, the stochastic solutions of the extended traffic equations $\tilde{\xi} = (\tilde{\xi}_i, i = 0, 1, \dots, J)$ and $\xi = (\xi_i, i = 0, 1, \dots, J)$ are the same. Therefore, the assumptions of Proposition 4.1 are satisfied. From the proof of Proposition 4.1, respectively, from Corollary 4.1, we see that we have to evaluate

$$\frac{\lambda}{\xi_0}\biggl(\sum_{\underline{n}\in\mathbb{N}^J} \pi(\underline{n}) \sum_{i=0}^{J}\sum_{j=0}^{J} \xi_j f_j(\underline{n}) g_i(\underline{n}) r_{ji} - \sum_{\underline{n}\in\mathbb{N}^J} \pi(\underline{n}) \sum_{i=0}^{J}\sum_{j=0}^{J} \xi_j f_j(\underline{n}) g_i(\underline{n}) \widetilde{r}_{ji}\biggr). \tag{4.15}$$

We consider the case of increasing functions f and g.

We first assume that $a \neq b$ and $c \neq d$. Then (4.15) is, for fixed $\underline{n} = (n_1, \dots, n_J) \in \mathbb{N}^J$,

$$\begin{aligned}
&\sum_{\underline{n}\in\mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) \biggl(\sum_{j=1}^{J} \lambda_j g(\underline{n} + e_j) + \sum_{j=1}^{J} \mu_j(n_j) r(j, 0) g(\underline{n} - e_j) \\
&\qquad + \sum_{j=1}^{J} \mu_j(n_j) \sum_{i=1}^{J} r(j, i) g(\underline{n} - e_j + e_i)\biggr) \\
&- \sum_{\underline{n}\in\mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) \biggl(\sum_{j=1}^{J} \lambda_j g(\underline{n} + e_j) + \sum_{j=1}^{J} \mu_j(n_j) r(j, 0) g(\underline{n} - e_j) \\
&\qquad + \sum_{\substack{j=1\\ j\neq a,c}}^{J} \mu_j(n_j) \sum_{i=1}^{J} r(j, i) g(\underline{n} - e_j + e_i) \\
&\qquad + \mu_a(n_a) \biggl(\sum_{\substack{i=1\\ i\neq a,b,d}}^{J} r(a, i) g(\underline{n} - e_a + e_i) + \biggl(r(a, d) - \frac{\alpha}{\xi_a}\biggr) g(\underline{n} - e_a + e_d) \\
&\qquad\qquad + \biggl(r(a, b) + \frac{\alpha}{\xi_a}\biggr) g(\underline{n} - e_a + e_b)\biggr)
\end{aligned}$$

$$
\begin{aligned}
&\quad + \mu_c(n_c)\Bigg(\sum_{\substack{i=1\\ i\neq c,b,d}}^{J} r(c,i) g(\underline{n} - e_c + e_i) \\
&\qquad + \left(r(c,d) + \frac{\alpha}{\xi_c} \right) g(\underline{n} - e_c + e_d) \\
&\qquad + \left(r(c,b) - \frac{\alpha}{\xi_c} \right) g(\underline{n} - e_c + e_b) \Bigg)\Bigg) \\
&= \alpha \sum_{\underline{n}\in\mathbb{N}^J} \pi(\underline{n}) f(\underline{n}) \Bigg(\mu_a(n_a) \frac{1}{\xi_a} (g(\underline{n} - e_a + e_d) - g(\underline{n} - e_a + e_b)) \\
&\qquad + \mu_c(n_c) \frac{1}{\xi_c} (g(\underline{n} - e_c + e_b) - g(\underline{n} - e_c + e_d)) \Bigg) \\
&= \alpha G^{-1} \Bigg(\sum_{(n_j\in\mathbb{N}:\ j\in\{1,\dots,J\}-\{a\})} \prod_{\substack{j=1\\ j\neq a}}^{J} \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)} \sum_{n_a=1}^{\infty} \prod_{k=1}^{n_a} \frac{\eta_a}{\mu_a(k)} \\
&\qquad \times \mu_a(n_a) (f(\underline{n}) g(\underline{n} - e_a + e_d) - f(\underline{n}) g(\underline{n} - e_a + e_b)) \frac{1}{\xi_a} \\
&\qquad + \sum_{(n_j\in\mathbb{N}:\ j\in\{1,\dots,J\}-\{c\})} \prod_{\substack{j=1\\ j\neq c}}^{J} \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)} \sum_{n_c=1}^{\infty} \prod_{k=1}^{n_c} \frac{\eta_c}{\mu_c(k)} \\
&\qquad \times \mu_c(n_c) (f(\underline{n}) g(\underline{n} - e_c + e_b) - f(\underline{n}) g(\underline{n} - e_c + e_d)) \frac{1}{\xi_c} \Bigg) \\
&= \alpha G^{-1} \Bigg(\sum_{(n_j\in\mathbb{N}:\ j\in\{1,\dots,J\}-\{a\})} \prod_{\substack{j=1\\ j\neq a}}^{J} \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)} \sum_{n_a=0}^{\infty} \prod_{k=1}^{n_a} \frac{\eta_a}{\mu_a(k)} \\
&\qquad \times \frac{\eta_a}{\xi_a} (f(\underline{n} + e_a) g(\underline{n} + e_d) - f(\underline{n} + e_a) g(\underline{n} + e_b)) \\
&\qquad + \sum_{(n_j\in\mathbb{N}:\ j\in\{1,\dots,J\}-\{c\})} \prod_{\substack{j=1\\ j\neq c}}^{J} \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)} \sum_{n_c=0}^{\infty} \prod_{k=1}^{n_c} \frac{\eta_c}{\mu_c(k)} \\
&\qquad \times \frac{\eta_c}{\xi_c} (f(\underline{n} + e_c) g(\underline{n} + e_b) - f(\underline{n} + e_c) g(\underline{n} + e_d)) \Bigg) \\
&= \alpha \frac{\lambda}{\xi_0} G^{-1} \sum_{\underline{n}\in\mathbb{N}^J} \prod_{j=1}^{J} \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)} ([f(\underline{n} + e_a) g(\underline{n} + e_d) - f(\underline{n} + e_a) g(\underline{n} + e_b)] \\
&\qquad + [f(\underline{n} + e_c) g(\underline{n} + e_b) - f(\underline{n} + e_c) g(\underline{n} + e_d)]) \\
&= \alpha \frac{\lambda}{\xi_0} G^{-1} \sum_{\underline{n}\in\mathbb{N}^J} \prod_{j=1}^{J} \prod_{k=1}^{n_j} \frac{\eta_j}{\mu_j(k)} ([g(\underline{n} + e_d) - g(\underline{n} + e_b)][f(\underline{n} + e_a) - f(\underline{n} + e_c)]) \\
&= \alpha \frac{\lambda}{\xi_0} \sum_{\underline{n}\in\mathbb{N}^J} \pi(\underline{n}) ([g(\underline{n} + e_b) - g(\underline{n} + e_d)][f(\underline{n} + e_a) - f(\underline{n} + e_c)]) \\
&\leq 0.
\end{aligned}
$$

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

For $a = b$ and $c = d$, we find, with slight modifications to the proof, that (4.15) is equal to

$$-\alpha\frac{\lambda}{\xi_0}\sum_{\underline{n}\in\mathbb{N}^J}\pi(\underline{n})([g(\underline{n}+e_a)-g(\underline{n}+e_c)][f(\underline{n}+e_a)-f(\underline{n}+e_c)])\leq 0.$$

Similarly, for $a = b$ and $c \neq d$, (4.15) is equal to

$$-\alpha\frac{\lambda}{\xi_0}\sum_{\underline{n}\in\mathbb{N}^J}\pi(\underline{n})([g(\underline{n}+e_a)-g(\underline{n}+e_d)][f(\underline{n}+e_a)-f(\underline{n}+e_c)])\leq 0,$$

and, for $a \neq b$ and $c = d$, (4.15) is equal to

$$-\alpha\frac{\lambda}{\xi_0}\sum_{\underline{n}\in\mathbb{N}^J}\pi(\underline{n})([g(\underline{n}+e_b)-g(\underline{n}+e_c)][f(\underline{n}+e_a)-f(\underline{n}+e_c)])\leq 0.$$

This completes the proof.

The lemma and its proof show that, for the partially ordered node set $(\widetilde{J}, \prec)$ and nodes $a, b, c, d \in \widetilde{J}$ (not necessarily distinct) such that $a \prec c$ and $b \prec d$ hold, and for functions $f, g\colon \mathbb{N}^{\tilde{J}} \to \mathbb{R}$ which fulfill either

$$f(n+e_c)-f(n+e_a)\geq 0 \quad\text{and}\quad f(n+e_d)-f(n+e_b)\geq 0 \quad\text{for all } n\in\mathbb{N}^{\tilde{J}}$$

or

$$f(n+e_c)-f(n+e_a)\leq 0 \quad\text{and}\quad f(n+e_d)-f(n+e_b)\leq 0 \quad\text{for all } n\in\mathbb{N}^{\tilde{J}},$$

we immediately obtain the correlation inequality, (4.14). For example, consider, for $a \prec c$, the projection $\Upsilon_c^{\tilde{J}}\colon \mathbb{N}^{\tilde{J}} \to \mathbb{R}$, $(n_i : i \in \widetilde{J}) \to n_c$. It is readily seen that $\Upsilon_c^{\tilde{J}}$ fulfills the first pair of conditions, while $\Upsilon_a^{\tilde{J}}$ fulfills the second pair of conditions.

Finally, we mention that the properties relevant in the present discussion can be formulated as being properties of functions with increasing marginal differences, i.e. for $i, j \in \widetilde{J}$ such that $i \prec j$ holds, we have $f(n+e_j) - f(n) \geq f(n+e_i) - f(n)$.

We are now ready to prove a result on the generalized partial sum order for network processes which relies on the order structure of the node set.

Theorem 4.3. *Consider an ergodic Jackson network with extended routeing matrix* $R = [r_{i,j}]_{i,j=0,1,\ldots,J}$ *according to (3.2) and queue length process* $\boldsymbol{X}$. *We assume that the node set* $\tilde{J} = \{1, \ldots, J\}$ *is a partially ordered set* $(\tilde{J}, \prec)$.

Define a new network with queue length process $\widetilde{\boldsymbol{X}}$ *as follows. The nodes, the nodes' structure, and the external arrival processes are the same as in the original network. The routeing matrix* $\widetilde{R} = [\widetilde{r}_{i,j}]_{i,j=0,1,\ldots,J}$ *is computed by a sequence of* $n \geq 1$ *feasible Robin-Hood transformations according to (4.12) for nodes* $i, j \in \{1, \ldots, J\}$.

Then, for any pair of comonotone functions $f, g\colon \mathbb{N}^{\tilde{J}} \to \mathbb{R}_+$ *with respect to the generalized partial sum order '*$\prec_*$*' (either both increasing or both decreasing),*

$$\langle f, Q^{\boldsymbol{X}} g\rangle_\pi \leq \langle f, Q^{\widetilde{\boldsymbol{X}}} g\rangle_\pi$$

holds.

Proof. The proof follows by an n-fold successive application of Lemma 4.1 with successively fixed $a_m, b_m, c_m, d_m \in \tilde{J}$, $1 \le m \le n$, which fulfill the required order relations.

From the assumptions of Theorem 4.3, it follows that the nodes which occur in the mass shifting are linearly ordered. Either $a \prec c \prec b \prec d$ or $b \prec d \prec a \prec c$ holds.

This means that the mass shifting is only between edges from the underlying graph and, therefore, completely in line with our order structure: no new edges for the transition probability graph are generated, but possibly some edges after applying the transformation have zero probability for customers to move across.

5. Dependence orderings for processes and monotonicity properties

In the general theory of dependence ordering of stochastic processes [4], relation (4.2) plays a prominent role. It turns out that in order to obtain similar inequalities for network processes, we have to adapt the general theory to the present context.

In this section we briefly collect necessary definitions and previous results on monotonicity and dependence orderings. We generalize these concepts to prepare a framework which will be used in our investigation of the structure of stochastic networks, especially their internal dependencies. Because the respective proofs are in line with those of the theorems in the previous settings, we omit them here. The definitions in this section and the relevant theorems are valid in a more general state space setting than we will use here; see [4].

Let $\mathbb{E}$ be a partially ordered countable space with discrete topology, σ-algebra $\mathcal{E} = 2^{\mathbb{E}}$, and closed partial order '$\prec$'. Selecting different classes $\mathcal{F}$ of real functions on $\mathbb{E}$, we will construct integral stochastic orders. Equalities and inequalities between integrals are always assumed to hold provided that the respective integrals are well defined.

Definition 5.1. (a) Random elements $\boldsymbol{X}$ and $\boldsymbol{Y}$ of $\mathbb{E}^n$ are called concordant stochastically ordered with respect to $\mathcal{F}$ (written as $\boldsymbol{X} \prec^n_{\mathcal{F}-\mathrm{cc}} \boldsymbol{Y}$ or $\boldsymbol{Y} \succ^n_{\mathcal{F}-\mathrm{cc}} \boldsymbol{X}$, but often abbreviating the notation to $\boldsymbol{X} \prec_{\mathcal{F}-\mathrm{cc}} \boldsymbol{Y}$ and, respectively, $\boldsymbol{Y} \succ_{\mathcal{F}-\mathrm{cc}} \boldsymbol{X}$) if

$$\mathrm{E}\biggl[\prod_{i=1}^{n} f_i(X_i)\biggr] \le \mathrm{E}\biggl[\prod_{i=1}^{n} f_i(Y_i)\biggr]$$

for all $f_i \in \mathcal{I}^*_+(\mathbb{E}) \cap \mathcal{F}$ and all $f_i \in \mathcal{D}^*_+(\mathbb{E}) \cap \mathcal{F}$, $i = 1, \dots, n$.

(b) Let $T \subseteq \mathbb{R}$ be an index set for stochastic processes $\boldsymbol{X} = (X_t : t \in T)$ and $\boldsymbol{Y} = (Y_t : t \in T)$ with $X_t, Y_t : (\Omega, \mathcal{F}, \mathrm{P}) \to (\mathbb{E}, \mathcal{E}, \prec)$, $t \in T$. We say that $\boldsymbol{X}$ and $\boldsymbol{Y}$ are concordant stochastically ordered with respect to a class $\mathcal{F}$ of functions on $(\mathbb{E}, \mathcal{E}, \prec)$ (and write $\boldsymbol{X} \prec_{\mathcal{F}-\mathrm{cc}} \boldsymbol{Y}$) if, for all $n \ge 2$ and all $t_1 < t_2 < \cdots < t_n$, we have, on $\mathbb{E}^n$,

$$(X_{t_1}, \dots, X_{t_n}) \prec_{\mathcal{F}-\mathrm{cc}} (Y_{t_1}, \dots, Y_{t_n}).$$

The setting of (b) will be applied to Markovian models.

In (a), taking for $\mathcal{F}$ the space of all measurable functions $\mathcal{M}$ on $\mathbb{E}$ we obtain the usual concordance ordering, as in [4]. It is easy to see that the two-dimensional marginals of the Markov chains related by the Robin-Hood construction in (4.12) fulfill

$$(X_0, X_1) \le_{\mathcal{M}-\mathrm{cc}} (Y_0, Y_1);$$

see Corollary 2.1 of [4].

In this situation, $\boldsymbol{X} \prec_{\mathcal{F}-\mathrm{cc}} \boldsymbol{Y}$ implies that X_j and Y_j have the same distribution for all $j = 1, \ldots, J$, so, if such $\boldsymbol{X}$ and $\boldsymbol{Y}$ are equilibria of product form networks then $\boldsymbol{X}$ and $\boldsymbol{Y}$ have the same distribution. Furthermore, $\boldsymbol{X} \prec_{\mathcal{M}-\mathrm{cc}} \boldsymbol{Y}$ implies that $\mathrm{cov}(f(X_j), g(X_j)) \leq \mathrm{cov}(f(Y_j), g(Y_j))$ for each $f \in \mathcal{I}_+^*(\mathbb{E})$ and $g \in \mathcal{I}_+^*(\mathbb{E})$. If the class $\mathcal{F}$ is sufficiently rich, these properties will be maintained. For example, if $\mathcal{F}$ contains the indicator functions of point-generated increasing and decreasing sets, $\{i\}^{\uparrow} = \{j \in E : i \prec j\}$ and $\{i\}^{\downarrow} = \{j \in E : j \prec i\}$, for concordant stochastically ordered processes $\boldsymbol{X}$ and $\boldsymbol{Y}$ (with respect to $\mathcal{F}$), we can compare the probability of extreme events like

$$\begin{aligned} &\mathrm{P}(\inf (X_{t_1}, \ldots, X_{t_n}) \succ t) \leq \mathrm{P}(\inf (Y_{t_1}, \ldots, Y_{t_n}) \succ t) \\ \text{and} \quad &\mathrm{P}(\sup (X_{t_1}, \ldots, X_{t_n}) \prec s) \leq \mathrm{P}(\sup (Y_{t_1}, \ldots, Y_{t_n}) \prec s) \end{aligned}$$

for large t and small s. Similar remarks apply throughout to almost all of our subsequent results.

We mention that in most cases $\mathcal{F}$ will be a convex cone of functions which is often additionally closed under pointwise convergence.

5.1. Discrete-time Markov processes

Let $\boldsymbol{X} = (X_t : t \in \mathbb{Z})$ and $\boldsymbol{Y} = (Y_t : t \in \mathbb{Z})$ with $X_t, Y_t : (\Omega, \mathcal{F}, \mathrm{P}) \to (\mathbb{E}, \mathcal{E}, \prec)$ be discrete time, stationary, homogeneous Markov processes. Assume that π is an invariant (stationary) one-dimensional marginal distribution, the same for both $\boldsymbol{X}$ and $\boldsymbol{Y}$, and denote the one-step transition kernels for $\boldsymbol{X}$ and $\boldsymbol{Y}$ by $K^X : \mathbb{E} \times \mathcal{E} \to [0, 1]$ and $K^Y : \mathbb{E} \times \mathcal{E} \to [0, 1]$, respectively. Denote the respective transition kernels for the time-reversed processes $\overleftarrow{\boldsymbol{X}}$ and $\overleftarrow{\boldsymbol{Y}}$ by $\overleftarrow{K}^X$ and $\overleftarrow{K}^Y$. Note that $\overleftarrow{K}^X$ and $\overleftarrow{K}^Y$ can be seen as adjoint operators on $L_2(\mathbb{E}, \pi)$ of K^X and K^Y, respectively. Indeed, for $f \in L_2(\mathbb{E}, \pi)$, $K^X f$ is defined by $K^X f(x) = \int f(y) K^X(x, \mathrm{d}y)$ and $\langle f, g \rangle_\pi = \int f(x) g(x) \pi(\mathrm{d}x)$. Then $\langle \overleftarrow{K}^X f, g \rangle_\pi = \langle f, K^X g \rangle_\pi$ for all $f, g \in L_2(\mathbb{E}, \pi)$, that is, $\overleftarrow{K}^X = (K^X)^*$, where A^* denotes the adjoint of the operator A. Similarly, $\overleftarrow{K}^Y = (K^Y)^*$. We say that a stochastic kernel $K : \mathbb{E} \times \mathcal{E} \to [0, 1]$ is $\mathcal{F}$-monotone if $\int f(x) K(s, \mathrm{d}x) \in \mathcal{I}_+^*(\mathbb{E}) \cap \mathcal{F}$ for each $f \in \mathcal{I}_+^*(\mathbb{E}) \cap \mathcal{F}$.

A pair $\boldsymbol{X}$ and $\boldsymbol{Y}$ of discrete-time Markov processes having the same invariant probability measure fulfills Property 5.1, below, which recently proved to be useful in comparing second-order properties of Markov processes; see [2], [3], [4], and [7]. It will be convenient to impose this property here as well.

Property 5.1. ($\mathcal{F}$-symmetric monotonicity.) *Either K^Y and $\overleftarrow{K}^X$ are $\mathcal{F}$-monotone or K^X and $\overleftarrow{K}^Y$ are $\mathcal{F}$- monotone.*

The following theorem is new but an analog of Theorem 3.1 of [4], therefore, we omit the proof as it follows the same arguments.

Theorem 5.1. (Concordance ordering under $\mathcal{F}$-symmetric monotonicity.) *For the stationary Markov processes $\boldsymbol{X}$ and $\boldsymbol{Y}$ defined above with a common invariant distribution π, under $\mathcal{F}$-symmetric monotonicity, the following relations are equivalent.*

(i) $\boldsymbol{X} \prec_{\mathcal{F}-\mathrm{cc}} \boldsymbol{Y}$.

(ii) $(X_0, X_1) \prec^2_{\mathcal{F}-\mathrm{cc}} (Y_0, Y_1)$.

(iii) $\langle f, K^X g \rangle_\pi \leq \langle f, K^Y g \rangle_\pi$ *for all* $f, g \in \mathcal{I}_+^*(\mathbb{E}) \cap \mathcal{F}$ *and all* $f, g \in \mathcal{D}_+^*(\mathbb{E}) \cap \mathcal{F}$.

(iv) $\langle f, \overleftarrow{K}^X g \rangle_\pi \leq \langle f, \overleftarrow{K}^Y g \rangle_\pi$ *for all* $f, g \in \mathcal{I}_+^*(\mathbb{E}) \cap \mathcal{F}$ *and all* $f, g \in \mathcal{D}_+^*(\mathbb{E}) \cap \mathcal{F}$.

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

If $\mathcal{F}$ is the set $\mathcal{M}$ of all measurable real functions on $\mathbb{E}$ then we have the statement of Theorem 3.1 of [4].

5.2. Continuous-time Markov processes

Let $\boldsymbol{X} = (X_t : t \in \mathbb{R})$ and $\boldsymbol{Y} = (Y_t : t \in \mathbb{R})$ with $X_t, Y_t : (\Omega, \mathcal{F}, \mathrm{P}) \to (\mathbb{E}, \mathcal{E}, \prec)$ be stationary, homogeneous Markov processes with countable state spaces. Denote the corresponding families of transition kernels of $\boldsymbol{X}$ and $\boldsymbol{Y}$ by $\mathbb{K}^X = (K_t^X : \mathbb{E} \times \mathcal{E} \to [0, 1] : t \geq 0)$ and $\mathbb{K}^Y = (K_t^Y : \mathbb{E} \times \mathcal{E} \to [0, 1] : t \geq 0)$, respectively, and the respective transition kernels for the stationary time-reversed processes $\overleftarrow{X}$ and $\overleftarrow{Y}$ by $\overleftarrow{K}^X = (\overleftarrow{K}_t^X : \mathbb{E} \times \mathcal{E} \to [0, 1] : t \geq 0)$ and $\overleftarrow{K}^Y = (\overleftarrow{K}_t^Y : \mathbb{E} \times \mathcal{E} \to [0, 1] : t \geq 0)$. Assume that π is an invariant distribution common to both $\mathbb{K}^X$ and $\mathbb{K}^Y$; that is, $\int K_t^X(x, \mathrm{d}y)\pi(\mathrm{d}x) = \int K_t^Y(x, \mathrm{d}y)\pi(\mathrm{d}x) = \pi(\mathrm{d}y)$ for all $t > 0$.

We restrict our considerations to Markov processes with bounded generators $Q^X = [Q^X(x, y)]$ and $Q^Y = [Q^Y(x, y)]$ of $\boldsymbol{X}$ and $\boldsymbol{Y}$, respectively. For the time-reversed processes, we use the corresponding notation $\overleftarrow{Q}^X$ and $\overleftarrow{Q}^Y$. We say that $\mathbb{K}^X = (K_t^X : \mathbb{E} \times \mathcal{E} \to [0, 1] : t \geq 0)$ is $\mathcal{F}$-time monotone if, for each $t \geq 0$, K_t^X is $\mathcal{F}$-monotone.

Analogously to the above, a pair $\boldsymbol{X}$ and $\boldsymbol{Y}$ of continuous-time Markov processes having the same invariant probability measure fulfills the following property.

Property 5.2. ($\mathcal{F}$-time symmetric monotonicity.) *Either* $\mathbb{K}^Y$ *and* $\overleftarrow{K}^X$ *are* $\mathcal{F}$*-time monotone or* $\mathbb{K}^X$ *and* $\overleftarrow{K}^Y$ *are* $\mathcal{F}$*-time monotone.*

Using similar arguments as in the proof of Theorem 3.3 of [4] (which is covered by the present theorem if $\mathcal{F}$ is the set $\mathcal{M}$), we have the following theorem.

Theorem 5.2. (Concordance ordering under $\mathcal{F}$-time symmetric monotonicity.) *Suppose that* $(\mathbb{E}, \mathcal{E}, \prec)$ *is countable and that the above-defined stationary chains* $\boldsymbol{X}$ *and* $\boldsymbol{Y}$ *have bounded intensity matrices* Q^X *and* Q^Y*, respectively. Then, under* $\mathcal{F}$*-time symmetric monotonicity, the following properties are equivalent.*

(i) $\boldsymbol{X} \prec_{\mathcal{F}-\mathrm{cc}} \boldsymbol{Y}$.

(ii) $(X_0, X_t) \prec^2_{\mathcal{F}-\mathrm{cc}} (Y_0, Y_t)$ *for all* $t > 0$.

(iii) $\langle f, T_t^X g\rangle_\pi \leq \langle f, T_t^Y g\rangle_\pi$ *for all* $f, g \in \mathcal{I}_+^*(\mathbb{E}) \cap \mathcal{F}$ *and all* $f, g \in \mathcal{D}_+^*(\mathbb{E}) \cap \mathcal{F}$*, for all* $t > 0$.

(iv) $\langle f, Q^X g\rangle_\pi \leq \langle f, Q^Y g\rangle_\pi$ *for all* $f, g \in \mathcal{I}_+^*(\mathbb{E}) \cap \mathcal{F}$ *and all* $f, g \in \mathcal{D}_+^*(\mathbb{E}) \cap \mathcal{F}$.

(v) $\langle f, \overleftarrow{Q}^X g\rangle_\pi \leq \langle f, \overleftarrow{Q}^Y g\rangle_\pi$ *for all* $f, g \in \mathcal{I}_+^*(\mathbb{E}) \cap \mathcal{F}$ *and all* $f, g \in \mathcal{D}_+^*(\mathbb{E}) \cap \mathcal{F}$.

Reducing the class of functions from $\mathcal{M}$ to some smaller class $\mathcal{F}$ makes the theorem much more versatile for applications, as we will demonstrate below.

From Theorem 5.2 we conclude that comparing correlations for stochastic network processes in continuous time is an interplay of the following two tasks.

- Proving monotonicity, the form of which we identified as $\mathcal{F}$-time symmetric monotonicity.
- Proving generator inequalities.

(A similar remark applies for discrete time, where we use $\mathcal{F}$-symmetric monotonicity and prove one-step transition inequalities.)

Generator inequalities have been studied in the previous sections. We will continue with studying the concept of time-symmetric monotonicity for network processes.

From the recent literature on dependence structures of Markovian processes with one-dimensional (linearly ordered) discrete state spaces, we conclude that $\mathcal{F}$*-time symmetric monotonicity* (in continuous time) and $\mathcal{F}$*-symmetric monotonicity* (in discrete time) play a central role; see, e.g. [7]. This property occurred independently in the literature several times; see, e.g. [2] and [3, Lemma 3.2].

So, in general, we cannot hope to dispense these assumptions when proving dependence properties in the more complex network setting. Nevertheless, the necessity of these assumptions is still an unsolved problem; some counterexamples, where dependence structures of Markovian processes over a finite-time horizon are proved without $\mathcal{F}$-symmetric monotonicity are provided in [4, Section 3.3].

On the other hand, the need for some monotonicity is emphasized further by the related theory of association in time for Markov processes, which strongly relies on monotonicity of the processes; for a review, see [3] and [14, Chapter II].

Obviously, any pair $\boldsymbol{X}$ and $\boldsymbol{Y}$ of reversible (stationary), strongly stochastic, monotone Markov chains in discrete time or Markov processes in continuous time with the same stationary distribution constitute a pair of $\mathcal{M}$-symmetric monotone chains, where $\mathcal{M}$ is the class of measurable functions. It follows that especially discrete-time and continuous birth–death processes with the same equilibrium constitute such a pair.

For stochastic networks, which are in general not reversible, on the other hand, the property of time-symmetric monotonicity is a natural property. Every Jackson network process $\boldsymbol{X}$ with service rates that are at all nodes nondecreasing functions of the local queue length [3, Corollary 4.1] is stochastically monotone with respect to strong stochastic ordering on the set of all probability measures on $(\mathbb{N}^J, \leq)$. Because the time-reversed process of a Jackson network process is the state process of a suitably defined Jackson network with the same properties for the service rates, any pair of Jackson network processes with the same steady-state distribution fulfills $\mathcal{F}$-time symmetric monotonicity, where $\mathcal{F} = \mathcal{I}^*(\mathbb{N}^J, \leq)$. (Note that this allows us to compare networks with different service rates, a situation which is not covered here.)

We only mention that, by a similar observation, $\mathcal{F}$-time symmetric monotonicity holds for Gordon–Newell networks with respect to strong stochastic ordering.

In the investigations found in the literature, $\mathcal{F}$ is always the class of all (bounded) increasing functions with respect to the natural linear order. The weaker concept of $\mathcal{F}$-(time) symmetric monotonicity for smaller sets of functions is suggested by the concept of integral orders with respect to subclasses of the class of increasing functions; see [16] or [18]. The problems arising with this concept are that we need the *closure property*, and we need the $\mathcal{F}$-functions to be transformed into $\mathcal{F}$-functions, or at least into the maximal generator of the respective order (see [16, Definition 3.2] or [18, Definition 2.3.3]).

The balance between having a small class of $\mathcal{F}$-functions and the necessity of obtaining the closure property is demonstrated next. The first example is in the spirit of the classical Gordon–Newell networks but with a smaller set $\mathcal{F}$. Recall that $\mathcal{L}$ is the set of nonnegative affine-linear functions on $S(I, J)$.

Proposition 5.1. (Linear service rates.) *Consider two Gordon–Newell network processes* $\boldsymbol{X}$ *and* $\widetilde{\boldsymbol{X}}$ *on the state space* $S(I, J) \subseteq \mathbb{N}^J$ *equipped with the coordinatewise order '*$\leq$*', both with stationary distribution* $\pi^{I,J}$*. Assume that the service rates in both networks at all nodes are linear functions of the local queue length,* $\mu_j(n_j) = \mu_j n_j$*,* $n_j \geq 0$ *for all* $j = 1, \ldots, J$*. Then the pair* $\boldsymbol{X}$ *and* $\widetilde{\boldsymbol{X}}$ *of Gordon–Newell network processes is* $\mathcal{L}$*-time symmetric monotone.*

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

Proof. Since Gordon–Newell networks are monotone and their time reversals are Gordon–Newell networks again, it suffices to show that any transition kernel $K_t^X(s, \cdot)$ maps $\mathcal{L}$ into $\mathcal{L}$. Because Q^X is bounded, we have, for $\underline{n}, \underline{m} \in S(I, J)$ and sufficiently small $h \geq 0$,

$$K_t^X(\underline{n}, \underline{m}) = \exp(t(I + hQ^X))(\underline{n}, \underline{m}) = \sum_{k=0}^{\infty} \frac{\mathrm{e}^{-ht}(ht)^k}{k!}(I + hQ^X)^k(\underline{n}, \underline{m}), \qquad t \geq 0.$$

For $f\colon S(I, J) \to \mathbb{R}_+$ in $\mathcal{L}$ with $f(n_1, \dots, n_J) = a + \sum_{i=1}^J \alpha_i n_i$, we find that

$$\begin{aligned}((I + hQ^X)f)(\underline{n}) &= f(\underline{n}) + h \sum_{\underline{m} \in S(I,J)} Q^X(\underline{n}, \underline{m}) f(\underline{m}) \\ &= a + \sum_{j=1}^J n_j \bigg(\alpha_j(1 - h\mu_j) + h\mu_j \sum_{i=1}^J r_{j,i}\alpha_i\bigg), \end{aligned} \tag{5.1}$$

so $((I + hQ^X)f)$ is linear affine and, for sufficiently small h, $((I + hQ^X)f)$ is nonnegative. Iterating and taking limits yields the result.

Proposition 5.2. (Generalized tandem network.) *Consider a generalized tandem network process* $\boldsymbol{X}$ *on the state space* $\mathbb{N}^J$ *equipped with the partial sum order '*$\leq_*$*' with stationary distribution* π. *The routeing for* $\boldsymbol{X}$ *is* linear *as follows:*

- *customers enter the network only through node* 1*:* $\lambda_1 > 0$ *and* $\lambda_j = 0,\ j = 2, \dots, J$*;*
- *customers depart from the network only from node* J*:* $r_{J0} > 0$ *and* $r_{j0} = 0,\ j = 1, \dots, J-1$*;*
- *customers move only stepwise:* $r_{j,j+1} > 0,\ j = 1, \dots, J-1, r_{j,j-1} \geq 0,\ j = 2, \dots, J,$ *and* $r_{j,j} \geq 0,\ j = 1, \dots, J$, *and* $r_{j,i} = 0$ *in any other case.*

Let $\widetilde{\boldsymbol{X}}$ *be another generalized tandem network process with stationary distribution* π, *and with routeing subject to the same restriction as described for* $\boldsymbol{X}$.

Assume that the arrival rates and the (nondecreasing) service rates in both networks are the same and bounded.

Then the pair $\boldsymbol{X}$ *and* $\widetilde{\boldsymbol{X}}$ *is* $\mathcal{I}^*(\mathbb{R}^J, \leq_*) \cup \mathcal{D}^*(\mathbb{R}^J, \leq_*)$*-time symmetric monotone.*

Proof. It is immediate from the construction that the time-reversed processes of $\boldsymbol{X}$ and $\widetilde{\boldsymbol{X}}$ exhibit the same tandem network structure as $\boldsymbol{X}$ and $\widetilde{\boldsymbol{X}}$. Therefore, [3, Corollary 4.2], which guarantees the $\mathcal{I}^*(\mathbb{R}^J, \leq_*)$ monotonicity for all K_t^X and $\overleftarrow{K}_t^X$, yields the desired result. (Note that in [3] it was assumed for technical reasons that $\tilde{r}_{j,j} = r_{j,j} = 0,\ j = 1, \dots, J$, holds. The statements of Corollary 4.2 and Proposition 4.4 there obviously also hold in the present situation.)

Proposition 5.3. (Functions of the total population size.) *Consider two Jackson networks with linear service rates, i.e.* $\mu_j(n_j) = \mu_j n_j$ *for all* $h \in \mathbb{N}$ *and all* $j = 1, \dots, J$, *which have the same stationary distributions. Furthermore, assume that inside both networks the effective departure rates from all nodes are the same, i.e.* $\mu_j r_{j0}$ *is invariant for all* $j = 1, \dots, J$ *(and, therefore, greater than* 0*). Let*

$$\mathcal{F} = \{f\colon \mathbb{N}^J \to \mathbb{R}_+\colon f(n_1, \dots, n_J) = \hat{f}(n_1 + \cdots + n_J) \text{ for some } \hat{f}\colon \mathbb{R} \to \mathbb{R}_+\}$$

be the set of real-valued functions on $\mathbb{N}^J$, *which depend on the sum of the arguments only. Then these networks constitute an* $\mathcal{F}$*-time symmetric monotone pair.*

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

Proof. By direct computation we show that, for $f \in \mathcal{F}$, it holds that, for the generator Q^X of the associated network process $\boldsymbol{X}$,

$$(I + hQ^X)(f) \in \mathcal{F}.$$

Then the principles of the proof of Proposition 5.1 are applicable.

6. The interplay of time-symmetric monotonicity and correlation inequalities via generators

Jackson's result (3.3) says that the joint queue lengths vector process $\boldsymbol{X} = (X_t : t \geq 0)$ of the Jackson network in equilibrium has one-dimensional marginals with independent coordinates. This does not imply that there are no internal dependencies in $(X_t = (X_1(t), \dots, X_J(t)) : t \in \mathbb{R}_+)$. We have, for example, nonvanishing $\operatorname{cov}(X_i(t), X_i(t+s))$ and $\operatorname{cov}(X_i(t), X_j(t+s))$. For Gordon–Newell networks in equilibrium, even the one-dimensional marginals have negatively correlated coordinates.

Our conjecture is that the internal dependence structure of the Markovian joint queue length network processes is to some extent determined through the internal dependence structure of the customers' routeing behavior. To be more specific, given a prescribed network in equilibrium, we expect that if we make the routeing process more dependent in a specified way then the joint queue length process will be more dependent in some (possibly differently) specified way. Dependence will be characterized by generalized correlation functions with suitable function classes, as described in Section 5.

Let $\rho = (\rho_1, \dots, \rho_J)$ be an ordered sequence of the numbers $\{1, 2, \dots, J\}$ (without repetition), which will serve as a rank vector for the linear factors of functions in

$$\begin{aligned} \mathcal{L}(\rho) = \{ f : S(I, J) &\to \mathbb{R}_+ : f(n_1, \dots, n_J) \\ &= a + \sum_{i=1}^{J} \alpha_i n_i,\ \alpha_i \in \mathbb{R},\ i = 1, \dots, J,\ a \in \mathbb{R}_+,\ \mathcal{R}(\alpha_1, \dots, \alpha_J) = \rho \} \\ \subseteq \mathcal{L}. \end{aligned}$$

Theorem 6.1. *Consider two ergodic Gordon–Newell network processes with common stationary distribution π: $\boldsymbol{X}$ with a doubly stochastic routeing matrix $R = [r_{ij}]$ and $\widetilde{\boldsymbol{X}}$ with the routeing matrix $\widetilde{R} = [\widetilde{r}_{ij}] = RT$ for a doubly stochastic matrix $T = [t_{ij} : i, j = 1, \dots, J]$. The service rates $\mu_j(n_j) = \mu_j n_j$ are the same in both networks.*

Let $A\mathcal{R}(\mu) = \rho = (\rho_1, \dots, \rho_J)$ denote the antirank vector of the unit service intensity vector $\mu = (\mu_1, \dots, \mu_J)$. Then

$$\boldsymbol{X} \geq_{\mathcal{L}(\rho)-\mathrm{cc}} \widetilde{\boldsymbol{X}}.$$

Proof. From Proposition 5.1 we know that the transition kernels of $\boldsymbol{X}$ and of the time reversal of $\widetilde{\boldsymbol{X}}$ map $\mathcal{L}(\rho)$ into the set $\mathcal{L}$ of affine-linear functions on $S(I, J)$. For proving $\mathcal{L}(\rho)$-time symmetric monotonicity, we have to show that the ranks of the transformed coefficient vectors remain invariant.

For simplicity, assume that, for all $j = 1, \dots, J-1$, $\alpha_j \geq \alpha_{j+1}$ holds. Then we have to show that (see (5.1))

$$\alpha_j(1 - h\mu_j) + h\mu_j \sum_{k=1}^{J} r_{j,k}\alpha_k \geq \alpha_{j+1}(1 - h\mu_{j+1}) + h\mu_{j+1} \sum_{k=1}^{J} r_{j+1,k}\alpha_k. \tag{6.1}$$

https://doi.org/10.1239/jap/1222441833 Published online by Cambridge University Press

But this follows because by majorization (recall that R is doubly stochastic) we have

$$\sum_{k=1}^{J} r_{j,k}\alpha_k \geq \sum_{k=1}^{J} r_{j+1,k}\alpha_k,$$

and because of our assumption $\mu_j \lesseqqgtr \mu_{j+1}$, $j = 1, \dots, J-1$. Because a similar property holds for the time reversal of $\boldsymbol{X}$ and $\widetilde{\boldsymbol{X}}$, we have $\mathcal{L}(\rho)$-time symmetric monotonicity of the pair $\boldsymbol{X}$ and $\widetilde{\boldsymbol{X}}$.

If we could show that, for all pairs of functions $f, g \in \mathcal{I}_+^*(\mathbb{N}^J) \cap \mathcal{L}(\rho)$ and all pairs of functions $f, g \in \mathcal{D}_+^*(\mathbb{N}^J) \cap \mathcal{L}(\rho)$, it holds that

$$\langle f, Q^{\tilde{X}} g\rangle_\pi \leq \langle f, Q^{X} g\rangle_\pi, \tag{6.2}$$

we can apply Theorem 5.2 to prove the theorem. But (6.2) follows from Theorem 4.1.

Example 6.1. In many applications the functions in $\mathcal{F}$ serve as cost or reward functions connected with the network's performance. A typical cost function is as follows. Per customer at node j and per time unit, a cost of amount α_j occurs, so $f_j(X_j(t)) = \alpha_j X_j(t)$ is the cost at node j. Incorporate a fixed constant cost a. Then in state $(n_1, \dots, n_J)$ the total cost per time unit is $f(n_1, \dots, n_J) = a + \sum_{i=1}^{J} \alpha_i n_i$. When we put the natural assumption that the costs increase when the service speed decreases, this situation is covered by the preceding theorem.

Our next theorem is in the realm of generalized tandem networks, as described in Proposition 5.2. Robin-Hood transforms under this graph structure are of the following form. Shift (probability) mass $\alpha > 0$ from arcs $r_{j,j+1}$ and $r_{j+1,j}$ to arcs $r_{j,j}$ and $r_{j+1,j+1}$. This has the following consequences.

Theorem 6.2. (General tandem.) *Consider Jackson network processes* $\boldsymbol{X}$ *and* $\widetilde{\boldsymbol{X}}$ *on the state space* $\mathbb{N}^J$ *(equipped with the partial sum order '*$\leq_*$*'; see (4.13)) which have common stationary distribution* π.

Furthermore, assume that, for some fixed $j \in \{1, \dots, J-1\}$ *and* $\alpha > 0$, $r_{j,j+1} > \alpha$ *and* $r_{j+1,j} \geq \alpha$*, and that the routeing for* $\widetilde{\boldsymbol{X}}$ *is obtained by a Robin-Hood transformation according to (4.12), where* $a = b = j$ *and* $c = d = j+1$.

Then, with $\mathcal{PS} := \mathcal{I}^*(\mathbb{R}^J, \leq_*) \cup \mathcal{D}^*(\mathbb{R}^J, \leq_*)$*, we have*

$$\boldsymbol{X} \leq_{\mathcal{PS}-\mathrm{cc}} \widetilde{\boldsymbol{X}}.$$

Proof. We can apply Theorem 5.2(iv), because from Proposition 5.2 we know that $\mathcal{I}^*(\mathbb{N}^J, \leq_*) \cup \mathcal{D}^*(\mathbb{N}^J, \leq_*)$-time symmetric monotonicity holds. Computing $\langle f, Q^X g\rangle_\pi - \langle f, Q^{\widetilde{X}} Y g\rangle_\pi \leq 0$ for all $f, g \in \mathcal{I}_+^*(\mathbb{N}^J, \leq_*)$ and all $f, g \in \mathcal{D}_+^*(\mathbb{N}^J, \leq_*)$ follows exactly the same approach used in the proof of Theorem 4.1 for the case in which $a = b$ and $c = d$.

It is worth mentioning that the Robin-Hood transformation applied to the tandem routeing yields Peskun ordering between the routeing matrices (see Definition 4.1), but we do not need reversibility in the theorem. These are substituted by time-symmetric monotonicity.

7. Further applications and complements

In Proposition 4.3 we have shown that, for the discrete-time Markov chains with transition matrices $I + \varepsilon Q^X$ and $I + \varepsilon Q^{\widetilde{X}}$, the asymptotic variance can be ordered whenever the routeing matrices are ordered by Peskun ordering and both are reversible.

We want to establish here similar properties for networks without reversibility requirements. Our examples are networks which fulfill correlation inequalities and time-symmetric monotonicity. As a prototype example, we take the setting of Theorem 6.1.

Proposition 7.1. *Consider ergodic Gordon–Newell networks with state processes $\boldsymbol{X}$ and $\widetilde{\boldsymbol{X}}$. The networks have the same service intensity vector $\mu = (\mu_1, \dots, \mu_J)$, which means that $\mu_j(n_j) = \mu_j n_j$ for all $n_j \in \mathbb{N}$, and we denote by $A\mathcal{R}(\mu) = \rho = (\rho_1, \dots, \rho_J)$ the antirank vector of μ.*

Assume that the routeing matrices $R = [r_{ij}]$ and $\widetilde{R} = [\widetilde{r}_{ij}]$ are doubly stochastic with $\widetilde{R} = RT$ for some doubly stochastic matrix T.

Then, for any function $f \in \mathcal{L}(\rho)$ (nonnegative linear affine functions on $S(I, J)$ with $\mathcal{R}(\alpha_1, \dots, \alpha_J) = \rho$) which is either increasing or decreasing, we have

$$v(f, I + \varepsilon Q^{\tilde{X}}) \geq v(f, I + \varepsilon Q^{X}).$$

Proof. From (6.2), it follows that

$$\langle f, (I + Q^{\tilde{X}})g\rangle_\pi \leq \langle f, (I + Q^{X})g\rangle_\pi$$

for all pairs of functions $f, g \in \mathcal{I}^*_+(\mathbb{N}^J) \cap \mathcal{L}(\rho)$ and all pairs of functions $f, g \in \mathcal{D}^*_+(\mathbb{N}^J) \cap \mathcal{L}(\rho)$. This set of inequalities serves as (iii) of Theorem 5.1 and, therefore,

$$\langle f, (I + Q^{\tilde{X}})^k g\rangle_\pi \leq \langle f, (I + Q^{X})^k g\rangle_\pi$$

holds for all $k \in \mathbb{N}$ if we can show that

$$I + \varepsilon Q^{X} \quad \text{and} \quad I + \varepsilon Q^{\widetilde{X}}$$

fulfill $\mathcal{L}(\rho)$-symmetric monotonicity. We only show that $I + \varepsilon Q^{X}$ is $\mathcal{L}(\rho)$-monotone, as the other case is similar.

In the proof of Proposition 5.1 we have shown that, for $f \in \mathcal{L}(\rho)$, $(I + \varepsilon Q^{\widetilde{X}})f$ is affine linear; see (5.1). That the rank vector of the linear factors of $(I + \varepsilon Q^{\widetilde{X}})f$ is ρ was shown by (6.1) in the proof of Theorem 6.1. This completes the proof.

Proposition 7.1 raises the question whether there is a similar connection between the general theorem of Peskun [20, Theorem 2.1.1] and a similar consequence from our Theorem 5.1. The answer is positive. We immediately have from Theorem 5.1 the following corollary.

Corollary 7.1. *For stationary Markov chains $\boldsymbol{X}$ and $\boldsymbol{Y}$ with common equilibrium distribution π on $\mathbb{E}$, which fulfill $\mathcal{F}$-symmetric monotonicity for some class $\mathcal{F}$ of real functions on $\mathbb{E}$, we have*

$$\boldsymbol{X} \prec_{\mathcal{F}-\mathrm{cc}} \boldsymbol{Y} \quad \Longrightarrow \quad v(f, K^{X}) \leq v(f, K^{Y})$$

for all functions in $\mathcal{I}^*_+(\mathbb{E}) \cap \mathcal{F}$ and in $\mathcal{D}^*_+(\mathbb{E}) \cap \mathcal{F}$.

For the case in which $\mathcal{F} = \mathcal{M}$, the class of all measurable functions, this is Corollary 4.1 of [4]. We show that the following theorem can be considered as a special case of Corollary 7.1.

Theorem 7.1. ([20].) *Consider reversible Markov chains $\boldsymbol{X}$ and $\boldsymbol{Y}$ with transition kernels K^{X} and K^{Y} with common equilibrium distribution π on $\mathbb{E}$. Then, for all real functions on $\mathbb{E}$,*

$$K^{X} \prec_{\mathrm{P}} K^{Y} \quad \Longrightarrow \quad v(f, K^{Y}) \leq v(f, K^{X}).$$

For the proof that Peskun's theorem is a consequence of Corollary 7.1, we need some preparation. The first necessary observation was made by Neal [19] and entitled 'Looking at one pair of states is enough'. Because Neal only provided an illuminating example, we give a short proof here.

Lemma 7.1. *Consider irreducible transition kernels $K^{\boldsymbol{X}}$ and $K^{\boldsymbol{Y}}$ with common equilibrium distribution π on a finite space $\mathbb{E}$, both of which are reversible with respect to π, such that*

$$K^{\boldsymbol{X}} \prec_{\mathrm{P}} K^{\boldsymbol{Y}}$$

holds. Then $K^{\boldsymbol{X}}$ can be transformed into $K^{\boldsymbol{Y}}$ by a sequence of intermediate kernels $K^{\boldsymbol{X}} = K_0, K_1, K_2, \ldots, K_n, K_{n+1} = K^{\boldsymbol{Y}}$ on $\mathbb{E}$ such that

- *each K_i is reversible with respect to π;*
- *in any step $K_i \to K_{i+1}$ there are only two states $x, y \in \mathbb{E}$, $x \neq y$, where matrix entries are changed such that $K_i(x, y) \leq K_{i+1}(x, y)$ and $K_i(x, x) \geq K_{i+1}(x, x)$, and $K_i(y, x) \leq K_{i+1}(y, x)$ and $K_i(y, y) \geq K_{i+1}(y, y)$, where at least one of the inequalities is strict;*
- *all other entries of K_i and K_{i+1} are identical;*
- *nondiagonal entries of K_i that are already identical to those of $K^{\boldsymbol{Y}}$ are not changed during the step from K_i to K_{i+1}.*

Proof. If $K^{\boldsymbol{X}} = K^{\boldsymbol{Y}}$, there is nothing to prove. So we can assume that there exists $x, y \in \mathbb{E}$, $x \neq y$, such that $K^{\boldsymbol{X}}(x, y) = K_0(x, y) < K^{\boldsymbol{Y}}(x, y)$ and $K^{\boldsymbol{X}}(x, x) = K_0(x, x) > K^{\boldsymbol{Y}}(x, x)$ hold. We provide the construction of K_1, and write, for all $v, w \in \mathbb{E}$, $v \neq w$,

$$K_0(v, w) + \underbrace{a(v, w)}_{\geq 0} = K^{\boldsymbol{Y}}(v, w) \quad \text{and obtain} \quad K_0(v, v) - \sum_{w \neq v} a(v, w) = K^{\boldsymbol{Y}}(v, v).$$

Because $K^{\boldsymbol{X}}$ and $K^{\boldsymbol{Y}}$ are reversible with respect to π, we have

$$\pi(x) K_0(x, y) = \pi(y) K_0(y, x),$$
$$\pi(x)(K_0(x, y) + a(x, y)) = \pi(y)(K_0(y, x) + a(y, x)),$$

and $a(x, y) > 0$ implies that $a(y, x) > 0$. Now K_1 is defined by

$$K_1(x, y) = K^{\boldsymbol{Y}}(x, y) \quad \text{and} \quad K_1(x, x) = K_0(x, x) - a(x, y),$$
$$K_1(y, x) = K^{\boldsymbol{Y}}(y, x) \quad \text{and} \quad K_1(y, y) = K_0(y, y) - a(y, x),$$
$$K_1(v, w) = K_0(v, w) \quad \text{otherwise.}$$

We conclude that (i) $K_0 \prec_{\mathrm{P}} K_1$, (ii) K_1 has strictly more entries coincident with $K^{\boldsymbol{Y}}$ than K_0, and all entries other than (x, y), (y, x), (x, x), and (y, y) are maintained, and (iii) K_1 is reversible with respect to π.

We only have to check (iii). For all pairs (v, w) other than (x, y), (y, x), (x, x), and (y, y), local balance equations maintain to hold via $K^{\boldsymbol{X}} = K_0$, and we have, from local balance for $K^{\boldsymbol{Y}}$, $\pi(x) K_1(x, y) = \pi(y) K_1(y, x)$ because of $K^{\boldsymbol{Y}}(x, y) = K_1(x, y)$ and $K^{\boldsymbol{Y}}(y, x) = K_1(y, x)$.

The rest of the proof follows by induction, which will obviously terminate.

Lemma 7.2. *Each of the successive steps in Lemma 7.1 is performed as a (reversed) Robin-Hood transform according to (4.12).*

Proof. We compare the two-dimensional distributions $\pi \otimes K_0$ and $\pi \otimes K_1$, and see immediately that all entries other than (x, y), (y, x), (x, x), and (y, y) are identical. The proposal is that we obtain $\pi \otimes K_1$ from $\pi \otimes K_0$ by shifting equal masses from (x, x) to (x, y) and from (y, y) to (y, x). Because K_1 and K_0 are reversible with respect to π, we have

$$\pi(x)K_1(x, y) = \pi(x)(K_0(x, y) + a(x, y)) = \pi(y)(K_0(y, x) + a(y, x)) = \pi(y)K_1(x, y)$$
$$\text{and} \quad \pi(x)K_0(x, y) = \pi(y)K_0(y, x).$$

Subtracting these equations shows that the masses that have been shifted are

$$\pi(x)a(x, y) = \pi(y)a(y, x) =: \alpha.$$

According to (4.12), this is turned into subtracting

$$\frac{\alpha}{\pi(x)} = a(x, y) \quad \text{and adding} \quad \frac{\alpha}{\pi(y)} = a(y, x)$$

to obtain the respective kernels, which we have done in Lemma 7.1.

Proof of Theorem 7.1. From reversibility of both kernels we conclude that the pair K^X and K^Y fulfills $\mathcal{M}$-symmetric monotonicity with respect to the trivial order '=' on $\mathbb{E}$. Furthermore, the stepwise construction going from K^X to K^Y shows that in any step we find a (reversed) Robin-Hood transformation (with respect to the trivial order) which yields (iii) in Theorem 5.1 and, therefore, iteratively the required cc-ordering.

Acknowledgements

We thank an anonymous referee for his/her careful reading of the first version of the paper and constructive suggestions to improve its presentation.

The work of the first author was supported by Deutscher Akademischer Austauschdienst. The work of the second author was supported by KBN, grant number N N2014079 33 (2007–2009) and by the Marie Curie Fellowship for the Transfer of Knowledge Harmonic Analysis, Nonlinear Analysis and Probability, MTKD-CT-2004-013389.

References

[1] Alberti, P. M. and Uhlmann, A. (1982). *Stochasticity and Partial Order*. VEB Deutscher Verlag der Wissenschaften, Berlin.

[2] Bäuerle, N. and Rolski, T. (1998). A monotonicity result for the workload in Markov-modulated queues. *J. Appl. Prob.* **35,** 741–747.

[3] Daduna, H. and Szekli, R. (1995). Dependencies in Markovian networks. *Adv. Appl. Prob.* **27,** 226–254.

[4] Daduna, H. and Szekli, R. (2006). Dependence ordering for Markov processes on partially ordered spaces. *J. Appl. Prob.* **43,** 793–814.

[5] Gordon, W. J. and G.F. Newell (1967). Closed queueing networks with exponential servers. *Operat. Res.* **15,** 254–265.

[6] Harris, T. G. (1977). A correlation inequality for Markov processes in partially ordered spaces. *Ann. Prob.* **5,** 451–454.

[7] Hu, T. and Pan, X. (2000). Comparisons of dependence for stationary Markov processes. *Prob. Eng. Inf. Sci.* **14,** 299–315.

[8] Iscoe, I. and McDonald, D. (1994). Asymptotics of exit times for Markov jump processes I. *Ann. Appl. Prob.* **22,** 372–397.

[9] Iscoe, I. and McDonald, D. (1994). Asymptotics of exit times for Markov jump processes II. Applications to Jackson networks. *Ann. Appl. Prob.* **22,** 2168–2182.
[10] Jackson, J. R. (1957). Networks of waiting lines. *Operat. Res.* **5,** 518–521.
[11] Jackson, J. R. (1963). Jobshop-like queueing systems. *Manag. Sci.* **10,** 131–142.
[12] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall, London.
[13] Marshall, A. W. and Olkin, I. (1979). *Inequalities: Theory of Majorisation and Its Applications*. Academic Press, New York.
[14] Liggett, T. M. (1985). *Interacting Particle Systems*, (Fundamental Principles Math. Sci. **276**). Springer, Berlin.
[15] Lorek, P. (2007). Speed of convergence to stationarity for stochastically monotone Markov chains. Doctoral thesis, Mathematical Institute, University of Wrocław.
[16] Li, H. and Shaked, M. (1994). Stochastic convexity and concavity of Markov processes. *Math. Operat. Res.* **19,** 477–493.
[17] Li, H. and Xu, S. H. (2000). Stochastic bounds and dependence properties of survival times in a multicomponent shock model. *J. Appl. Prob.* **37,** 1020–1043.
[18] Müller, A. and Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risks*. John Wiley, Chichester.
[19] Neal, R. M. (2004). Improving asymptotic variance of MCMC estimators: Non-reversible chains are better. Tech. Rep. 0406 Department of Statistics, University of Toronto.
[20] Peskun, P. H. (1973). Optimum Monte-Carlo sampling using Markov chains. *Biometrika* **60,** 607–612.
[21] Szekli, R. (1995). *Stochastic Ordering and Dependence in Applied Probability* (Lecture Notes Statist. **97**). Springer, New York.
[22] Szekli, R., Disney, R. L. and Hur, S. (1994). MR/GI/1 queues with positively correlated arrival streams. *J. Appl. Prob.* **31,** 497–514.
[23] Tierney, L. (1998). A note on Metropolis–Hastings kernels for general state spaces. *Ann. Appl. Prob.* **8,** 1–9.