# Computational Methods for Stochastic Relations and Markovian Couplings


Lasse Leskelä
Helsinki University of Technology
PO Box 1100, 02015 TKK, Finland
http://www.iki.fi/lsl/
lasse.leskela@iki.fi



## ABSTRACT

Order-preserving couplings are elegant tools for obtaining robust estimates of time-dependent and stationary distributions of Markov processes that are too complex to be analyzed exactly. The starting point of this paper is to study stochastic relations, which may be viewed as natural generalizations of stochastic orders. This generalization is motivated by the observation that for the stochastic ordering of two Markov processes, it suffices that the generators of the processes preserve some, not necessarily reflexive or transitive, subrelation of the order relation. The main contributions of the paper are an algorithmic characterization of stochastic relations between finite spaces, and a truncation approach for comparing infinite-state Markov processes. The methods are illustrated with applications to loss networks and parallel queues.



## Categories and Subject Descriptors

G.3 [**Probability and Statistics**]: Markov processes; C.4 [**Performance of Systems**]: Modeling techniques, Performance attributes

## General Terms

Algorithms, Performance

## Keywords

Markovian coupling, stochastic comparison, stochastic order, stochastic relation


## 1. INTRODUCTION

Comparison techniques based on stochastic orders [13, 14, 15] are key to obtaining upper and lower bounds for complicated random variables and processes in terms of simpler random elements. Consider for example two ergodic discrete-time Markov processes $X$ and $Y$ with stationary distributions $\mu_X$ and $\mu_Y$, taking values in a common ordered state space, and denote by $\le_{\mathrm{st}}$ the corresponding stochastic order. Then the upper bound

$$\mu_X \le_{\mathrm{st}} \mu_Y \tag{1}$$

can be established [7] without explicit knowledge of $\mu_X$ by verifying that the corresponding transition probability kernels $P$ and $P'$ satisfy

$$x \le y \implies P(x,\cdot) \le_{\mathrm{st}} P'(y,\cdot). \tag{2}$$

Analogous conditions for continuous-time Markov processes on countable spaces have been derived by Whitt [17] and Massey [12], and later extended to more general jump processes by Brandt and Last [2].

Less stringent sufficient conditions for obtaining (1) have recently been found using a new theory of stochastic relations [10]. Two random variables are stochastically related, denoted by $X \sim_{\mathrm{st}} Y$, if there exists a coupling $(\hat{X}, \hat{Y})$ of $X$ and $Y$ such that $\hat{X} \sim \hat{Y}$ almost surely, where $\sim$ denotes some relation between the state spaces of $X$ and $Y$. The main motivation for this definition is that (2) is by no means necessary for (1); a less stringent sufficient condition is that

$$x \sim y \implies P(x,\cdot) \sim_{\mathrm{st}} P'(y,\cdot) \tag{3}$$

for some, not necessarily symmetric or transitive, nontrivial subrelation of the underlying order relation. Another advantage of the generalized definition is that $X$ and $Y$ are no longer required to take values in the same state space, leading to greater flexibility in the search for bounding random elements $Y$. For example, to study whether $f(X) \le_{\mathrm{st}} g(Y)$ for some given real functions $f$ and $g$ defined on the state spaces of $X$ and $Y$, we may define a relation $x \sim y$ by the condition $f(x) \le g(y)$ [3].

A well-studied approach for estimating the stationary distribution of a given Markov process with transition probability kernel $P$, is to consider a fixed order on the state space, and construct a probability kernel $P'$, based on the structure of $P$ and the order, so that (2) holds. To obtain good bounds, the bounding kernel $P'$ should be computationally tractable, and in some sense minimal. For totally ordered spaces, there exist several fast algorithms for computing bounding kernels with algebraic structure that allows fast numerical computation of the stationary distribution [1, 5, 16].

The approach presented here differs from the above methodology. Instead of searching for a minimal (in some sense) bounding kernel $P'$ for a given probability kernel $P$ and a fixed order relation, in the following we shall start from a fixed pair $(P, P')$ of kernels, and search for a maximal sub-




*Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7607*
*http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7607*

relation of a given order for which (3) holds. The advantage with this approach is its generality: beside orders, the algorithms presented here can be applied to arbitrary relations, and the state spaces of the processes may be different. A disadvantage is that the candidate bounding kernel $P'$ must be chosen beforehand. However, in many applications such candidates can naturally be found by intuitive reasoning (see Section 5).

The rest of the paper is outlined as follows. After recalling the basic definitions, Section 2 presents a numerical algorithm for verifying stochastic relations between finite spaces, together with an analysis of computational complexity. Section 3 recalls how a recursive subrelation algorithm may be used to find Markovian couplings preserving a given relation. In Section 4, a new truncation approach is presented that allows to precisely compute truncated outcomes of the subrelation algorithm for infinite-state Markov processes with locally bounded jumps. Section 5 discusses applications to loss networks and parallel queues, and Section 6 concludes the paper.

# 2. STOCHASTIC RELATIONS

## 2.1 General definitions

We shall here recall the definitions of stochastic relations between countable spaces. A probability measures $\mu$ on a countable state space $S$ shall be viewed as a probability vector via identifying $\mu(x) = \mu(\{x\})$. For a treatment on more general spaces, see [10].

A *relation* between $S_1$ and $S_2$ is a subset of $S_1 \times S_2$. Given a nontrivial ($R \neq \emptyset$) relation $R$ between $S_1$ and $S_2$, we write

$$x \sim y,$$

if $(x, y) \in R$. The relation $R$ may equivalently be viewed as a matrix so that $R(x, y) = 1$ if $x \sim y$ and $R(x, y) = 0$ otherwise. A *coupling* of probability vectors $\mu$ on $S_1$ and $\nu$ on $S_2$ is a probability vector $\lambda$ on $S_1 \times S_2$ with marginals $\mu$ and $\nu$, that is,

$$\sum_{y \in S_2} \lambda(x, y) = \mu(x) \quad \text{for all } x \in S_1,$$
$$\sum_{x \in S_1} \lambda(x, y) = \nu(y) \quad \text{for all } y \in S_2.$$

For probability vectors $\mu$ on $S_1$ and $\nu$ on $S_2$ we denote

$$\mu \sim_{\text{st}} \nu,$$

and say that $\mu$ is stochastically related to $\nu$, if there exists a coupling $\lambda$ of $\mu$ and $\nu$ such that

$$\sum_{(x,y) \in R} \mu(x, y) = 1.$$

The relation $R_{\text{st}} = \{(\mu, \nu) : \mu \sim_{\text{st}} \nu\}$ is called the *stochastic relation* generated by $R$. Observe that two Dirac masses satisfy $\delta_x \sim_{\text{st}} \delta_y$ if and only if $x \sim y$. In this way the stochastic relation $R_{\text{st}}$ may be regarded as a natural randomization of the underlying relation $R$.

The following result in [10], which is rephrased here for ease of reference, provides an analytical method to check whether a pair of probability measures are stochastically related.

THEOREM 1. *[10] Two probability vectors $\mu$ and $\nu$ are stochastically related with respect to $R$ if and only if*

$$\sum_{x \in U} \mu(x) \leq \sum_{y \in S_2} \left( \max_{x \in U} R(x, y) \right) \nu(y) \tag{4}$$

*for all finite $U \subset S_1$, or equivalently, if and only if*

$$\sum_{y \in V} \nu(y) \leq \sum_{x \in S_1} \left( \max_{y \in V} R(x, y) \right) \mu(x) \tag{5}$$

*for all finite $V \subset S_2$.*

A random variable $X$ is *stochastically related* to a random variable $Y$, denoted by $X \sim_{\text{st}} Y$, if the distribution of $X$ is stochastically related to the distribution of $Y$. Observe that $X$ and $Y$ do not need to be defined on the same probability space. Recall that a coupling of random variables $X$ and $Y$ is a bivariate random variable whose distribution couples the distributions of $X$ and $Y$. Hence $X \sim_{\text{st}} Y$ if and only if there exists a coupling $(\hat{X}, \hat{Y})$ of $X$ and $Y$ such that $\hat{X} \sim \hat{Y}$ almost surely.

*Example 1.* If $\leq$ is an order (reflexive and transitive) relation on a space $S$, then the corresponding stochastic relation $\leq_{\text{st}}$ is called a stochastic order. Using Strassen's classical theorem [13], we see that $X \leq_{\text{st}} Y$ if and only if $\mathrm{E}\, f(X) \leq \mathrm{E}\, f(Y)$ for all positive increasing functions $f$ on $S$.

## 2.2 Stochastic relations between finite spaces

Let $R$ be a relation between finite spaces $S_1$ and $S_2$, and denote by $R_{\text{st}}$ the corresponding stochastic relation. Then Theorem 1 may be used to determine whether $\mu \sim_{\text{st}} \nu$. However, this requires to check the inequality (4) for all subsets of $S_1$, which is computationally infeasible unless the spaces are small. The following result shows that less checks may be sufficient. We shall denote the support of a probability vector $\mu$ by $U_\mu = \{x : \mu(x) > 0\}$. Moreover, we denote by $F(U, \mathbb{Z}_2)$ the set of vectors with components in $\{0, 1\}$ indexed by elements of $U$, which may also be identified as the set of all subsets of $U$.

THEOREM 2. *Two probability vectors $\mu$ and $\nu$ with supports $U_\mu$ and $U_\nu$ are stochastically related with respect to $R$ if and only if*

$$\sum_{x \in U_\mu} f(x)\, \mu(x) \leq \sum_{y \in U_\nu} \max_{x \in U_\mu} [f(x) R(x, y)]\, \nu(y) \tag{6}$$

*for all $f \in F(U_\mu, \mathbb{Z}_2)$, or equivalently, if and only if*

$$\sum_{y \in U_\nu} g(y)\, \nu(y) \leq \sum_{x \in U_\mu} \max_{y \in U_\nu} [R(x, y) g(y)]\, \mu(x) \tag{7}$$

*for all $g \in F(U_\nu, \mathbb{Z}_2)$.*

PROOF. In light of Theorem 1, it suffices to show the equivalence of (4) and (6), and the equivalence of (5) and (7). Observe that (4) directly implies (6), because the members of $F(U_\mu, \mathbb{Z}_2)$ may be identified with the indicator functions of subsets of $U_\mu$. To prove the converse, assume that (6) holds, and let $U$ be an arbitrary subset of $S_1$. Define $f(x) =$

*Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7607*
*http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7607*

$1(x \in U \cap U_\mu)$. Then

$$\begin{aligned}\sum_{x\in U} \mu(x) &= \sum_{x\in U_\mu} f(x)\mu(x) \\ &\le \sum_{y\in U_\nu} \max_{x\in U_\mu} [f(x)R(x,y)]\, \nu(y) \\ &= \sum_{y\in S_2} \left(\max_{x\in U_\mu\cap U} R(x,y)\right) \nu(y) \\ &\le \sum_{y\in S_2} \left(\max_{x\in U} R(x,y)\right) \nu(y).\end{aligned}$$

Hence (4) holds. Proving the equivalence of (5) and (7) is completely analogous. □

**Algorithm 1** Determining whether $\mu \sim_{\rm st} \nu$.

```
U_μ ← {x ∈ S_1 : μ(x) > 0}
U_ν ← {y ∈ S_2 : ν(y) > 0}
if #U_μ > #U_ν then
    flip μ ↔ ν, U_μ ↔ U_ν, S_1 ↔ S_2
end if
b ← true
for i = 1, . . . , 2^{#U_μ} do
    f ← i-th vector in F(U_μ, Z_2)
    v_l ← Σ_{x∈U_μ} f(x) μ(x)
    v_r ← Σ_{y∈U_ν} [max_{x∈U_μ} f(x)R(x, y)] ν(y)
    if v_l > v_r then
        b ← false
        break
    end if
end for
return b
```

Algorithm 1 describes how Theorem 2 can be applied to numerically determine whether $\mu \sim_{\rm st} \nu$. The interchange of the variables in the beginning corresponds to using (6) if the support of $\mu$ is smaller than $\nu$, and (7) otherwise. Inspection of Algorithm 1 shows that the computational complexity of determining whether $\mu \sim_{\rm st} \nu$ is of the order

$$O(\max(n_1', n_2') 2^{\min(n_1', n_2')}),$$

where $n_1'$ and $n_2'$ denote the cardinalities of the supports of $\mu$ and $\nu$. The algorithm is very slow when both state spaces are large and $\mu$ and $\nu$ have positive mass in all states. However, in many applications related to structured Markov chains we may assume that $\mu$ and $\nu$ have small supports (see Section 5).

*Remark 1.* The verification of $\mu \sim_{\rm st} \nu$ can be carried out faster, if the underlying relation $R$ has some structure that can be employed. For example, for a totally ordered space with $n$ elements, $\mu \le_{\rm st} \nu$ can be verified in $O(n)$ time. However, linear-time algorithms for checking $\mu \sim_{\rm st} \nu$ for relations that are not total orders seem scarce. For example, to verify whether $\mu \le_{\rm st} \nu$ for a general order relation requires to check whether $\mu(U) \le \nu(U)$ for all upper sets $U$, and the number of upper sets usually grows rapidly with the size of the state space.

## 3. MARKOV PROCESSES

### 3.1 Markovian couplings

All state spaces in the following shall be assumed finite or countably infinite. The keep the presentation short, all results shall be formulated for continuous-time Markov processes, which without further mention shall be assumed non-explosive. For a Markov process $X$ with values in $S$ we denote by $X(x,t)$ the value of the process at time $t$ given that was started at state $x$. A Markov process $\hat{X} = (\hat{X}_1, \hat{X}_2)$ taking values in $S_1 \times S_2$ is called a *Markovian coupling* of $X_1$ and $X_2$ if $\hat{X}(x,t)$ couples $X_1(x_1,t)$ and $X_2(x_2,t)$ for all $t$ and all $x = (x_1, x_2)$. A common approach for showing that the time-dependent distributions of two Markov processes $X_1$ and $X_2$ are stochastically ordered, is to find a Markovian coupling $\hat{X} = (\hat{X}_1, \hat{X}_2)$ of $X_1$ and $X_2$ such that

$$x_1 \le x_2 \implies \hat{X}_1(x_1,t) \le \hat{X}_2(x_2,t) \tag{8}$$

almost surely for all $t$ [13]. Observe that if (8) holds, then for all $t$ and all increasing functions $f$,

$$\begin{aligned}\mathrm{E}\, f(X_1(x_1,t)) &= \mathrm{E}\, f(\hat{X}_1(x_1,t)) \\ &\le \mathrm{E}\, f(\hat{X}_2(x_2,t)) = \mathrm{E}\, f(X_2(x_2,t)),\end{aligned}$$

whenever $x_1 \le x_2$. Hence $X_1(x_1,t) \le_{\rm st} X_2(x_2,t)$. If both $X_1$ and $X_2$ have unique stationary distributions, it follows by taking limits that the stationary distributions are stochastically ordered.

Observe that (8) is not necessary for the stochastic ordering of the stationary distributions of $X_1$ and $X_2$. To formulate a less stringent sufficient condition, we shall use the following definitions. Let $R$ be a relation between $S_1$ and $S_2$. A pair of Markov processes $X_1$ in $S_1$ and $X_2$ in $S_2$ is said to *stochastically preserve the relation* $R$, if

$$x_1 \sim x_2 \implies X_1(x_1,t) \sim_{\rm st} X_2(x_2,t) \quad \text{for all } t \ge 0.$$

Moreover, a set $B$ is called *invariant* for a Markov process $X$ if $x \in B$ implies $X(x,t) \in B$ for all $t$ almost surely. The following result was proved in [10], which we rephrase here for convenience.

Theorem 3. *[10] The following are equivalent:*

*(i) $X_1$ and $X_2$ stochastically preserve the relation $R$.*

*(ii) There exists a Markovian coupling of $X_1$ and $X_2$ for which $R$ is invariant.*

Assume now that $X_1$ and $X_2$ are Markov processes with values in an ordered space $S$, and assume $R$ is a subrelation of the order that is stochastically preserved by $X_1$ and $X_2$. Then $X_1(x_1,t) \le_{\rm st} X_2(x_2,t)$ for all $x_1 \sim x_2$. Especially, if $X_1$ and $X_2$ have unique stationary distributions, then a sufficient condition for the stochastic ordering of the stationary distributions is that $X_1$ and $X_2$ stochastically preserve some nontrivial subrelation of the order relation.

### 3.2 Subrelation algorithm

Recall that a matrix $Q$ with entries $Q(x,y)$, $x,y \in S$ is called a *rate matrix*, if $Q(x,y) \ge 0$ for all $x \ne y$ and $Q(x,x) = -\sum_{y\ne x} Q(x,y)$ for all $x$. If $X$ is a Markov process with rate matrix $Q$, then $Q(x,y)$ is the transition rate of $X$ from state $x$ into $y$, and we denote by $q(x) = -Q(x,x)$ the total transition rate of $X$ out of state $x$.

Given Markov processes $X_1$ and $X_2$ with rate matrices $Q_1$ and $Q_2$, define the relation-to-relation mapping $M_{Q_1,Q_2}$ by

$$M_{Q_1,Q_2}(R) = \{(x,y) \in R : (\mu_{x,y}, \nu_{x,y}) \in R_{\rm st}\}, \tag{9}$$

*Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7607*
*http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7607*

where the probability measures $\mu_{x,y}$ and $\nu_{x,y}$ are defined by

$$\mu_{x,y}(u) = q_{x,y}^{-1} Q_1(x,u) + \delta_x(u), \quad (10)$$

$$\nu_{x,y}(v) = q_{x,y}^{-1} Q_2(y,v) + \delta_y(v), \quad (11)$$

and where $q_{x,y} = 1 + q_1(x) + q_2(y)$. When there is no risk of confusion, we denote $M_{Q_1,Q_2} = M$. Moreover, define recursively the sequence $M^k(R)$ by setting $M^0(R) = R$, $M^k(R) = M(M^{k-1}(R))$ for $k \geq 1$, and denote the limit of the sequence by

$$M^*(R) = \cap_{k=0}^{\infty} M^k(R).$$

THEOREM 4. *[10] The relation $M^*(R)$ is the maximal subrelation of $R$ that is stochastically preserved by $X_1$ and $X_2$. Especially:*

*(i) $X_1$ and $X_2$ stochastically preserve $R$ if and only if $M(R) = R$.*

*(ii) $X_1$ and $X_2$ stochastically preserve a nontrivial subrelation of $R$ if and only if $M^*(R) \neq \emptyset$.*

**Algorithm 2** Computation of $R' = M(R)$.

```
R' ← n1-by-n2 zero matrix
for (x, y) ∈ R do
  q ← 1 + q^(1)(x) + q^(2)(y)
  μ ← q^{-1} Q^(1)(x, ·) + δ_x(·)
  ν ← q^{-1} Q^(2)(y, ·) + δ_y(·)
  Check whether μ ~st ν (use Algorithm 1)
  if μ ~st ν then
    R'(x, y) ← 1
  end if
end for
```

When the state spaces $S_1$ and $S_2$ are finite, Algorithm 2 describes how to numerically compute $M(R)$, and Algorithm 3 describes the computation of $M^*(R)$. Observe that for finite state spaces, Algorithm 3 computes the apparently infinite intersection $\cap_{k=0}^{\infty} M^k(R)$ in finite time, because as long as $M^k(R)$ and $M^{k-1}(R)$ are not equal, they differ by at least one element, and the sequence $M^k(R)$ is decreasing.

**Algorithm 3** Computation of $R^* = \cap_{n=0}^{\infty} M^k(R)$.

```
R' ← M(R) (use Algorithm 2)
while R' ≠ R do
  R ← R'
  R' ← M(R) (use Algorithm 2)
end while
R* ← R'
```

# 4. TRUNCATION APPROACH

## 4.1 Truncation of Markov processes

If $Q$ is a rate matrix of a Markov process on a countably infinite space $S$, and $S_n$ is a finite subset of $S$, we define the *truncation* of $Q$ into $S_n$ by

$$Q_n(x,y) = \begin{cases} Q(x,y), & x \neq y, \quad x,y \in S_n, \\ -\sum_{y \in S_n, y \neq x} Q(x,y), & x = y, \quad x \in S_n. \end{cases} \quad (12)$$

We shall later approximate $Q$ by $Q_n$ and use the finite subrelation algorithm applied to $Q_n$. To understand the approximation error, we need to study how the untruncated process may escape the set $S_n$. Given a rate matrix $Q$ on a countable space $S$, we say that an increasing sequence of finite sets $S_n \subset S$ is a *truncation sequence* for $Q$, if $\cup_{n=0}^{\infty} S_n = S$, and

$$\{y : Q(x,y) > 0\} \subset S_{n+1} \quad (13)$$

for all $x \in S_n$.

*Example 2.* Let $Q$ be the rate matrix of a Markov process on $\mathbb{Z}_+$ that is skip-free to the right, so that $Q(i,j) = 0$ for all $j > i+1$. Then $S_n = \mathbb{Z}_+ \cap [0,n]$ is a truncation sequence for $Q$.

The next result shows that truncation sequences can be constructed for most Markov processes encountered in applications. We say that a Markov process $X$ with rate matrix $Q$ has *locally bounded jumps*, if the set $\{y : Q(x,y) > 0\}$ is finite for all $x$.

LEMMA 1. *Any rate matrix $Q$ of a Markov process with locally bounded jumps possesses a truncating sequence.*

PROOF. Because $S$ is countable, we may choose an increasing sequence of finite sets $K_n$ such that $\cup_{n=0}^{\infty} K_n = S$. Using this sequence we may recursively define the sets $S_n$ by setting $S_0 = K_0$, and

$$S_{n+1} = S_n \cup K_n \cup J(S_n), \quad n \geq 0,$$

where

$$J(S_n) = \{y : Q(x,y) > 0 \text{ for some } x \in S_n\}$$

denotes the set of states that are reachable from $S_n$ by one jump. Then $\cup_{n=0}^{\infty} S_n = S$, because $K_n \subset S_{n+1}$ for all $n \geq 0$. Moreover, $J(S_n) \subset S_{n+1}$ implies that (13) holds for all $n$, and induction shows that the sets $S_n$ are finite, because $Q$ has locally bounded jumps. □

## 4.2 Truncation of stochastic relations

Let $R$ be a relation between countably infinite state spaces $S_1$ and $S_2$. If $S_i'$ is a finite subset of $S_i$, $i = 1, 2$, we define the truncation of $R$ by

$$R' = R \cap (S_1' \times S_2') \quad (14)$$

The corresponding stochastic relation $R'_{\mathrm{st}}$, a relation between probability measures on the finite spaces $S_1'$ and $S_2'$, is called the *truncation of* $R_{\mathrm{st}}$ into $S_1' \times S_2'$.

If $\mu_i$ is probability measure on $S_i$ having all its mass on $S_i'$, we may regard $\mu_i$ as a probability measure $\mu_i'$ on $S_i'$ by identifying subsets of $S_i'$ as subsets of $S_i$.

LEMMA 2. *Let $\mu_i$ be a probability measure on $S_i$ such that $\mu_i(S_i') = 1$, $i = 1, 2$. Then $(\mu_1, \mu_2) \in R_{\mathrm{st}}$ if and only if $(\mu_1', \mu_2') \in R'_{\mathrm{st}}$.*

PROOF. The claim follows using [10, Lemma 5.2], after observing that $R'$ is the relation induced from $R$ by the pair $(\phi_1, \phi_2)$, where $\phi_i$ is the natural embedding of $S_i'$ into $S_i$, and $\mu_i' = \mu_i \circ \phi_i$. □

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7607
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7607

## 4.3 Truncated subrelation algorithm

Given a pair of Markov processes $X_1$ and $X_2$ with rate matrices $Q_1$ and $Q_2$, and a relation $R$ between $S_1$ and $S_2$, Algorithm 3 describes how to recursively calculate $R^* = M^*_{Q_1,Q_2}(R)$ as the limit of the sequence $R^k = M^k_{Q_1,Q_2}(R)$. If the state spaces are infinite, $R^k$ cannot be computed using finite time and memory. Nevertheless, when the processes have locally bounded jumps, the truncations $T_N(R^k)$ of $R^k$ into suitable truncation sets $S_{1,N} \times S_{2,N}$ can be computed precisely, as shall be shown next.

For any relation $R$ between $S_{1,n}$ and $S_{2,n}$, let $M_n(R) = M_{Q_{1,n},Q_{2,n}}(R)$ be the relation given by Algorithm 2 applied to $R$ using the truncations $Q_{1,n}$ and $Q_{2,n}$ of $Q_1$ and $Q_2$ as defined in (13). Moreover, denote by $M_\infty = M_{Q_1,Q_2}$ the corresponding untruncated mapping.

LEMMA 3. *For any pair of Markov processes with locally bounded jumps, the truncation of $M_\infty(R)$ to $S_{1,n} \times S_{2,n}$ satisfies*

$$T_n(M_\infty(R)) = T_n(M_{n+1}(T_{n+1}(R))) \quad \text{for all } n.$$

PROOF. Observe that $T_n(M_\infty(R))$ equals the set of points $(x,y) \in T_n(R)$ such that the measures $\mu_{x,y}$ and $\nu_{x,y}$ defined in (10) and (11) are stochastically related with respect to $R$. Because $S_{1,n}$ and $S_{2,n}$ are truncating sequences for $Q_1$ and $Q_2$, we know that $\mu_{x,y}$ and $\nu_{x,y}$ have supports in $S_{1,n+1}$ and $S_{2,n+1}$, respectively. Hence by Lemma 2, $T_n(M_\infty(R))$ equals the set of points $(x,y) \in T_n(R)$ such that $\mu_{x,y}$ and $\nu_{x,y}$ are stochastically related with respect to $T_{n+1}(R)$. Because for $(x,y) \in T_n(R)$, the measures $\mu_{x,y}$ and $\nu_{x,y}$ remain the same, if we replace $Q_1$ and $Q_2$ by $Q_{1,n+1}$ and $Q_{2,n+1}$ in (10) and (11), the claim follows. □

THEOREM 5. *For any pair of Markov processes with locally bounded jumps, the truncation of $R^k = M^k_\infty(R)$ to $S_{1,n} \times S_{2,n}$ satisfies*

$$T_n(R^k) = T_n(M_{n+1}T_{n+1}) \cdots (M_{n+k}T_{n+k})(R).$$

PROOF. Apply Lemma 3 and induction. □

Algorithm 4 describes how Theorem 5 can be used to compute $T_n(R^k)$ in finite time and memory.

**Algorithm 4** Computation of $R' = T_N(M^K(R))$.

```
R' ← T_{N+K}(R)
for k = 1, ..., K do
    n ← N + K + 1 − k
    Q_{1,n} ← truncation of Q_1 into S_{1,n}
    Q_{2,n} ← truncation of Q_2 into S_{2,n}
    R' ← T_n(R')
    R' ← Algorithm 2 applied to (Q_{1,n}, Q_{2,n}, R')
end for
R' ← T_N(R')
```

The following result is a necessary condition for finding a subrelation of $R$ that is stochastically preserved by a pair of Markov processes.

THEOREM 6. *Let $X_1$ and $X_2$ be Markov processes with locally bounded jumps. If $T_n(M^k(R)) = \emptyset$ for some $k$, then $T_n(M^*(R)) = \emptyset$. Especially, if for all $n$, $T_n(M^k(R)) = \emptyset$ for some $k$, then $X_1$ and $X_2$ do not stochastically preserve any nontrivial subrelation of $R$.*

PROOF. Because the sequence $M^k(R)$ is decreasing, and because the truncation map is monotone with respect to set inclusion, the first claim follows. For the second claim, observe that if $T_n(R^*) = \emptyset$ for all $n$, then because $\cup_n(S_{1,n} \times S_{2,n}) = S$, $R^* = \emptyset$. □

# 5. APPLICATIONS

## 5.1 Multilayer loss networks

### 5.1.1 Overflow routing

Consider a loss system with $K$ customer classes and two layers of servers, where layer 1 contains $m_k$ servers dedicated to class $k$, and layer 2 consists of $n$ servers capable of serving all customer classes. Arriving class-$k$ customers are routed to vacant servers in one of the layers, with preference given to layer 1; or rejected otherwise. For analytical tractability, we assume that the interarrival times and the service requirements of class-$k$ customers are exponentially distributed with parameters $\lambda_k$ and $\mu_k$, respectively, and that all these random variables across all customer classes are independent. For brevity, we denote $m = (m_1, \ldots, m_K)$, $\lambda = (\lambda_1, \ldots, \lambda_K)$, and $\mu = (\mu_1, \ldots, \mu_K)$.

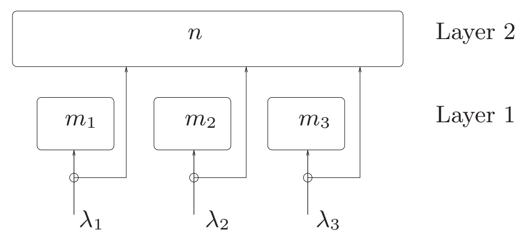


**Figure 1: Two-layer loss network with three customer classes ($K = 3$).**

Denote by $X_{i,k}(t)$ the number of class-$k$ customers being served at layer $i$ at time $t$. The system is described by the Markov process $X = (X_{i,k})$ taking values in

$$S_1 = \left\{ x \in \mathbb{Z}_+^K \times \mathbb{Z}_+^K : \ x_{1,k} \le m_k \ \forall k, \ \textstyle\sum_{k=1}^K x_{2,k} \le n \right\}, \tag{15}$$

and having the transitions

$$x \mapsto \begin{cases} x + e_{1,k}, & \text{at rate } \lambda_k 1(x_{1,k} < m_k), \\ x + e_{2,k}, & \text{at rate } \lambda_k 1(x_{1,k} = m_k, \sum_{k=1}^K x_{2,k} < n), \\ x - e_{1,k}, & \text{at rate } \mu_k x_{1,k}, \\ x - e_{2,k}, & \text{at rate } \mu_k x_{2,k}, \end{cases}$$

where $e_{i,k}$ denotes the unit vector in $\mathbb{Z}_+^K \times \mathbb{Z}_+^K$ corresponding to the coordinate direction $(i,k)$.

### 5.1.2 Maximum packing

To approximate the original two-layer loss system, we consider a modification of the system, where customers are redirected from layer 2 to layer 1 as soon as servers become vacant. This corresponds to the so-called maximum packing policy introduced by Everitt and Macfadyen [4]. Denote by $Y_k(t)$ the total number of customers in the system with maximum packing. Then $t \mapsto Y(t) = (Y_1(t), \ldots, Y_K(t))$ is a Markov process (see [6]) with values in

$$S_2 = \left\{ y \in \mathbb{Z}_+^K : \textstyle\sum_{k=1}^K (y_k - m_k)_+ \le n \right\},$$

Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7607
http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7607

and having the transitions

$$y \mapsto \begin{cases} y + e_k, & \text{at rate } \lambda_k 1(y + e_k \in S_2), \\ y - e_k, & \text{at rate } \mu_k y_k. \end{cases}$$

The structure of the above transition rates implies that the stationary distribution $\pi_Y$ of $Y$ is a product of Poisson distributions truncated to $S_2$ [8], so that

$$\pi_Y(y) = c \prod_{k=1}^{K} \frac{(\lambda_k/\mu_k)^{y_k}}{y_k!}, \quad y \in S_2,$$

where the constant $c$ can be solved from $\sum_{y \in S_2} \pi_Y(y) = 1$. This product form structure allows for fast computation of stationary performance characteristics of the maximum packing system.

### 5.1.3 Stochastic comparison

Table 1 illustrates the outcomes of the subrelation algorithm (computed using [9]) applied to various initial relations, where

$$R^{\text{sum}} = \left\{(x,y) \in S_1 \times S_2 : \textstyle\sum_k (x_{1,k} + x_{2,k}) \le \sum_k y_k \right\},$$
$$R_1^{\text{sum}} = \left\{(x,y) \in S_1 \times S_2 : \textstyle\sum_k x_{1,k} \le \sum_k (y_k \wedge m_k) \right\},$$
$$R_1^{\text{coord}} = \left\{(x,y) \in S_1 \times S_2 : x_{1,k} \le y_k \wedge m_k \text{ for all } k \right\}.$$

Note that $R_1^{\text{sum}}$ relates the state pairs $(x,y)$ where the total number of layer-1 customers corresponding to $x$ is less than that corresponding to $y$. Moreover, $R_1^{\text{coord}}$ relates the state pairs where the populations in layer 1 are coordinatewise ordered. The entry "several" in Table 1 refers to running Algorithm 2 separately for several pseudorandom parameter combinations, and taking the intersection of the produced relations as a final result.

| $R$ | $(\lambda_1, \lambda_2)$ | $(\mu_1, \mu_2)$ | $M^*(R)$ |
|---|---|---|---|
| $R^{\text{sum}}$ | $(1,1)$ | $(1,1)$ | $R^{\text{sum}} \cap R_1^{\text{sum}} \cap R_1^{\text{min}}$ |
| $R^{\text{sum}}$ | several | $(1,1)$ | $R^{\text{sum}} \cap R_1^{\text{coord}}$ |
| $R^{\text{sum}}$ | several | several | $\emptyset$ |
| $R_1^{\text{sum}}$ | $(1,1)$ | $(1,1)$ | $R_1^{\text{sum}}$ |
| $R_1^{\text{sum}}$ | several | $(1,1)$ | $R_1^{\text{coord}}$ |
| $R_1^{\text{sum}}$ | several | several | $R_1^{\text{coord}}$ |
| $R_1^{\text{coord}}$ | several | several | $R_1^{\text{coord}}$ |

**Table 1: Outcomes of the subrelation algorithm for a network with $m_1 = 1, m_2 = 1, n = 2$.**

From Table 1, we can make several conclusions on the behavior of the subrelation algorithm:

- The pair $(X_1, X_2)$ appears to stochastically preserve $R_1^{\text{sum}}$ for all parameter combinations (same parameters in both systems). This fact is in fact not hard to verify analytically.
- When $\mu_1 = \mu_2$, the pair $(X_1, X_2)$ stochastically preserves a nontrivial subrelation of $R^{\text{sum}}$. The maximal subrelation of $R^{\text{sum}}$ preserved may depend on the model parameters.
- $R^{\text{sum}} \cap R_1^{\text{coord}}$ appears to be a subrelation of $R^{\text{sum}}$ that is stochastically preserved for all parameter choices, as long as $\mu_1 = \mu_2$ (Figure 3). This fact is proved in [6].
- When the system is symmetric ($\lambda_1 = \lambda_2$, $\mu_1 = \mu_2$ and $m_1 = m_2$), a larger relation $R^{\text{sum}} \cap R_1^{\text{sum}} \cap R_1^{\text{min}} \supset R^{\text{sum}} \cap R_1^{\text{coord}}$ is stochastically preserved (Figure 2).
- When $\mu_1 \neq \mu_2$, $R^{\text{sum}}$ in general does not have a nontrivial subrelation preserved by the pair. This fact is also reflected in [6, Example 5.2.1], where it was found that the stationary distributions are not in general stochastically ordered with respect to $R^{\text{sum}}$.

As an illustration, the limiting relations $R^{\text{sum}} \cap R_1^{\text{sum}} \cap R_1^{\text{min}}$ and $R^{\text{sum}} \cap R_1^{\text{coord}}$ (filled circles), together with the initial relation $R^{\text{sum}}$ (filled + unfilled circles) are plotted in Figures 2 and 3, respectively.

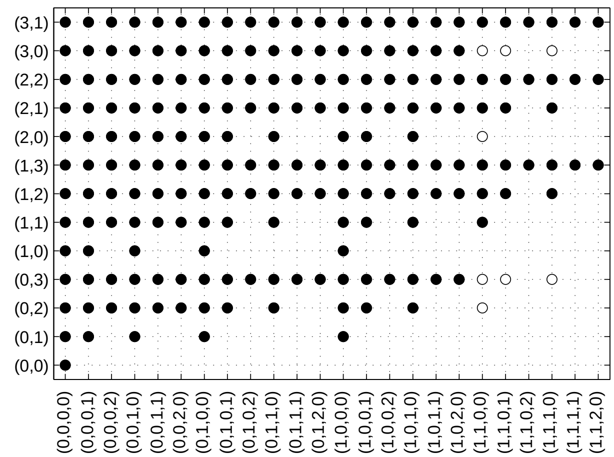


**Figure 2: Relations $R^{\text{sum}}$ and $M^*(R^{\text{sum}})$ for a system with $m_1 = 1, m_2 = 1, n = 2$, for $\lambda = (1,1)$ and $\mu = (1,1)$.**

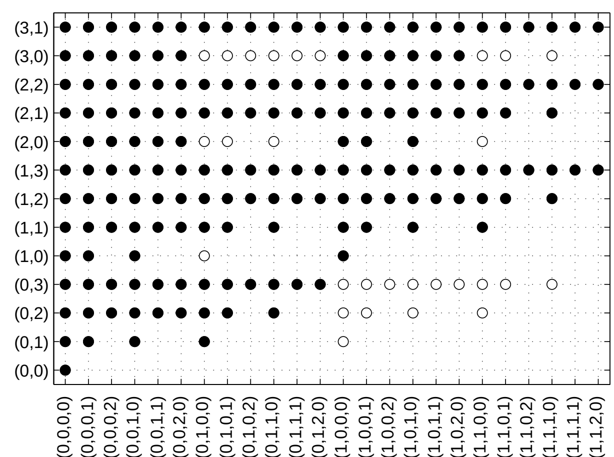


**Figure 3: Relations $R^{\text{sum}}$ and $M^*(R^{\text{sum}})$ for a system with $m_1 = 1, m_2 = 1, n = 2$, for various $\lambda$ and fixed $\mu = (1,1)$.**

## 5.2 Parallel queues

Consider a system of two queues in parallel, where customers of queue $k$ have arrival rate $\lambda_k$ and service rate $\mu_k$. Assuming that all interarrival and service times are independent and exponential, the queue length process $X =$

*Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7607*
*http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7607*

$(X_1, X_2)$ is a Markov process in $\mathbb{Z}_+^2$ with transitions

$$x \mapsto \begin{cases} x + e_k, & \text{at rate } \lambda_k, \\ x - e_k, & \text{at rate } \mu_k 1(x_k > 0). \end{cases}$$

We shall also consider a modification of the system, where load is balanced by routing incoming traffic to the shortest queue, modeled as a Markov process $X^{\text{LB}} = (X_1^{\text{LB}}, X_2^{\text{LB}})$ in $\mathbb{Z}_+^2$ with transitions

$$x \mapsto \begin{cases} x + e_1, & \text{at rate } (\lambda_1 + \lambda_2) 1(x_1 < x_2) + \lambda_1 1(x_1 = x_2), \\ x + e_2, & \text{at rate } (\lambda_1 + \lambda_2) 1(x_1 > x_2) + \lambda_2 1(x_1 = x_2), \\ x - e_1, & \text{at rate } \mu_1 1(x_1 > 0), \\ x - e_2, & \text{at rate } \mu_2 1(x_2 > 0). \end{cases}$$

Common sense suggests that load balancing decreases the number of customers in the system in some sense. The validity of this comparison property can be numerically studied using Algorithm 4. Denote the rate matrix of $X^{\text{LB}}$ by $Q_1$ and the rate matrix of $X$ by $Q_2$, and let $S_n = \mathbb{Z}_+^2 \cap [0, n-1]^2$, $n \geq 1$. Then $S_n$ is a truncation sequence for both $Q_1$ and $Q_2$.

Figure 4 illustrates five iterations of the subrelation algorithm (computed using [9]) truncated to $S_3$ applied to the coordinatewise order

$$R^{\text{coord}} = \left\{ (x, y) \in \mathbb{Z}_+^2 \times \mathbb{Z}_+^2 : x_1 \leq y_1, \ x_2 \leq y_2 \right\},$$

with the parameters $\lambda_1, \lambda_2, \mu_1, \mu_2$ all equal to one. Because $T_3(R^4) = T_3(R^5) = \emptyset$, we conjecture that there exists no nontrivial subrelation of $R^{\text{coord}}$ stochastically preserved by $(X^{\text{LB}}, X)$ (see Theorem 6).

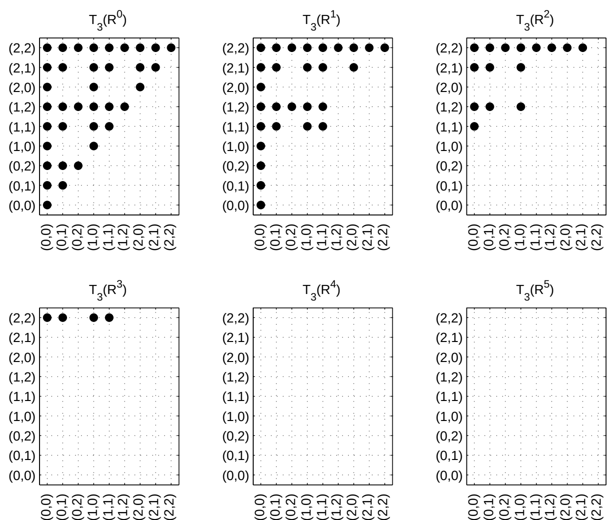


**Figure 4: Truncated subrelation algorithm applied to $R^{\text{coord}}$.**

Let us next study another order on $\mathbb{Z}_+^2$, defined by

$$R^{\text{sum}} = \left\{ (x, y) \in \mathbb{Z}_+^2 \times \mathbb{Z}_+^2 : x_1 + y_1, \ x_2 + y_2 \right\}.$$

Figure 5 illustrates five iterations of the subrelation algorithm (computed using [9]) truncated to $S_3$ applied to $R^0 = R^{\text{sum}}$ with the parameters $\lambda_1, \lambda_2, \mu_1, \mu_2$ all equal to one. The observation that $T_3(R^k)$ remains unchanged from $k = 1$ onwards suggests that some nontrivial subrelation of $R^{\text{sum}}$

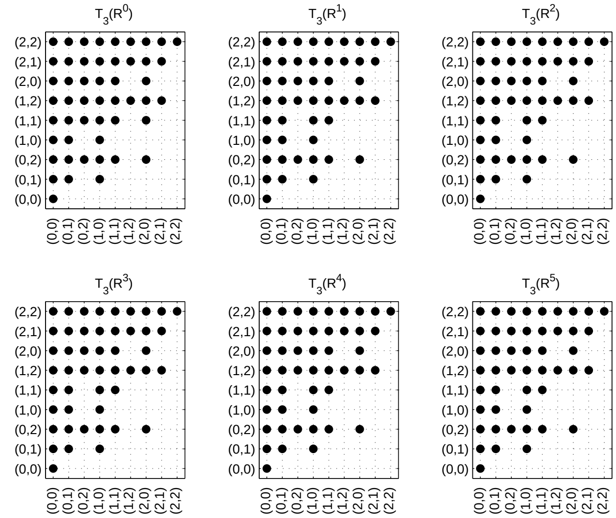


**Figure 5: Truncated subrelation algorithm applied to $R^{\text{sum}}$.**

might be stochastically preserved by $(X^{\text{LB}}, X)$. Indeed, it has been analytically shown [10] that whenever $\mu_1 = \mu_2$, the untruncated subrelation algorithm converges to the relation

$$R^* = \{ (x, y) : x_1 + x_2 \leq y_1 + y_2 \text{ and } x_1 \vee x_2 \leq y_1 \vee y_2 \}.$$

As a consequence of Theorem 4, the pair $(X^{\text{LB}}, X)$ stochastically preserves the relation $R^*$, which may be identified as the weak majorization order on $\mathbb{Z}_+^2$ [11]. Especially,

$$\begin{aligned} X_1^{\text{LB}}(t) + X_2^{\text{LB}}(t) &\leq_{\text{st}} X_1(t) + X_2(t), \\ X_1^{\text{LB}}(t) \vee X_2^{\text{LB}}(t) &\leq_{\text{st}} X_1(t) \vee X_2(t), \end{aligned}$$

for all $t$, whenever the initial states $X^{\text{LB}}(0)$ and $X(0)$ satisfy the same inequalities.

## 6. CONCLUSIONS

This paper presented computational methods for verifying stochastic relations and finding relation-invariant couplings of continuous-time Markov processes on finite and countably infinite state spaces. A key point of the paper is that the stochastic relationship between two probability measures can be quickly numerically checked, if one of the measures has small support (Theorem 2). This result allows the development of a truncation approach for finding relations stochastically preserved by pairs of Markov processes with locally bounded jumps. The truncated subrelation algorithm (Algorithm 4) allows to numerically find candidates for a subrelation of a given relation that is stochastically preserved by a pair of Markov processes. It remains an interesting open problem for future research to study how the truncated subrelation algorithm behaves for structured Markov processes with for example shift-invariant transition rate matrices.

## 7. ACKNOWLEDGMENTS

I thank two anonymous referees for helpful comments and pointers to the literature. This work has been supported by the Academy of Finland.

*Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7607*
*http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7607*

## 8. REFERENCES


[1] M. Ben Mamoun, A. Bušić, and N. Pekergin. Generalized class $\mathcal{C}$ Markov chains and computation of closed-form bounding distributions. *Probab. Engrg. Inform. Sci.*, 21(2):235–260, 2007.

[2] A. Brandt and G. Last. On the pathwise comparison of jump processes driven by stochastic intensities. *Math. Nachr.*, 167:21–42, 1994.

[3] M. Doisy. A coupling technique for stochastic comparison of functions of Markov processes. *J. Appl. Math. Decis. Sci.*, 4(1):39–64, 2000.

[4] D. E. Everitt and N. W. Macfadyen. Analysis of multicellular mobile radiotelephone systems with loss. *Brit. Telecom Technol. J.*, 1(2):37–45, 1983.

[5] J. M. Fourneau, M. Lecoz, and F. Quessette. Algorithms for an irreducible and lumpable strong stochastic bound. *Linear Algebra Appl.*, 386:167–185, 2004.

[6] M. Jonckheere and L. Leskelä. Stochastic bounds for two-layer loss systems. *Stoch. Models*, 24(4):583–603, 2008.

[7] T. Kamae, U. Krengel, and G. L. O'Brien. Stochastic inequalities on partially ordered spaces. *Ann. Probab.*, 5(6):899–912, 1977.

[8] F. P. Kelly. Loss networks. *Ann. Appl. Probab.*, 1:319–378, 1991.

[9] L. Leskelä. STOCHREL v1.0 – A Matlab stochastic relations package. http://www.iki.fi/lsl/software/stochrel/, 2008.

[10] L. Leskelä. Stochastic relations of random variables and processes. *J. Theor. Probab.*, To appear. http://arxiv.org/abs/0806.3562.

[11] A. W. Marshall and I. Olkin. *Inequalities: Theory of Majorization and Its Applications*. Academic Press, 1979.

[12] W. A. Massey. Stochastic orderings for Markov processes on partially ordered spaces. *Math. Oper. Res.*, 12(2):350–367, 1987.

[13] A. Müller and D. Stoyan. *Comparison Methods for Stochastic Models and Risks*. Wiley, 2002.

[14] M. Shaked and J. G. Shanthikumar. *Stochastic Orders*. Springer Series in Statistics. Springer, New York, 2007.

[15] R. Szekli. *Stochastic Ordering and Dependence in Applied Probability*. Springer, New York, 1995.

[16] L. Truffet. Reduction techniques for discrete-time Markov chains on totally ordered state space using stochastic comparisons. *J. Appl. Probab.*, 37(3):795–806, 2000.

[17] W. Whitt. Stochastic comparisons for non-Markov processes. *Math. Oper. Res.*, 11(4):608–618, 1986.



*Digital Object Identifier: 10.4108/ICST.VALUETOOLS2009.7607*
*http://dx.doi.org/10.4108/ICST.VALUETOOLS2009.7607*