Exact and Ordinary Lumpability in Finite Markov Chains[†]

Peter Buchholz, *Universität Dortmund*[*]

Abstract

Exact and ordinary lumpability in finite Markov chains is considered. Both concepts naturally define an aggregation of the Markov chain yielding an aggregated chain that allows the exact determination of several stationary and transient results for the original chain. We show which quantities can be determined without an error from the aggregated process and describe methods to calculate bounds on the remaining results. Furthermore, the concept of lumpability is extended to nearly lumpability yielding approximative aggregation.

AGGREGATION, STATIONARY AND TRANSIENT ANALYSIS, PERFORMANCE AND RELIABILITY MODELLING [††]

1 Introduction

The concept of lumpability and weak lumpability of partitions on state spaces of finite Markov chains (Markov chains) has been known for a long time [8, 9, 11, 12]. In those papers lumpability is defined by the fact that the process resulting from the observation of the Markov chain by masking out transitions and states inside a partition group is also a Markov chain. If this is the case for an arbitrary initial vector, then the partition is lumpable, if only some initial vectors yield a Markov chain, the partition is weakly lumpable for these initial vectors. It is known that lumpability can also be characterized on the transition matrix, since the sums of transition probabilities from each state in a partition group to all states of a fixed partition group have to be equal (row sum criterion). Both characterizations of lumpability are equivalent (see [8]). Recently Schweitzer [15] has introduced "exact lumpability" of a partition, which says that a partition is exactly lumpable if the sums of transition probabilities from states of a fixed partition group into each state of another or the same partition group are all equal (column sum criterion). It can be shown that exact lumpability is a special case of weak lumpability. Following [16] we will name lumpability "ordinary lumpability" in this paper and use the notation "lumpable partition" for a partition that is either ordinarily or exactly lumpable.

[†]to appear in Journal of Applied Probability (scheduled: March 1994)

[*]Postal address: Universität Dortmund, Informatik IV, P.O. Box 50 05 00,D-44221 Dortmund, Germany

[††]AMS (MOS) subject classification. Primary: 60J10, Secondary: 68M20

The characterizations of ordinary and exact lumpability on the transition matrix are of practical importance, because they can be used to develop constructive methods for the generation of lumpable partitions for an arbitrary Markov chain. Such methods are extremely useful in performance and reliability modelling, since the knowledge of a lumpable partition allows the generation of an aggregated Markov chain that is smaller than the original one, but can be used to determine several results for the original Markov chain without an error. Methods for the construction of lumpable partitions have been published in [2, 3] for a general class of hierarchical models. In [2] exact or ordinary lumpability is extended to nearly exact or ordinary lumpability which can be characterized straightforwardly on the transition matrix. Other approaches yielding lumpable partitions and aggregated processes are based on symmetries in Petri net or reliability models (see [1, 5, 4, 14]). However, what has not yet been done is to analyse which results can be determined from an aggregated Markov chain resulting from an ordinarily or exactly lumpable partition and which information is lost by aggregating the Markov chain. Such results are of theoretical and practical importance, since often the generation of the huge overall Markov chain for a complex performance or reliability model can be avoided by the direct construction of the aggregated Markov chain using the methods mentioned previously.

In Section 2, the basic notation and definitions are introduced. Although the paper considers discrete-time chains, most of the results hold also for continuous-time Markov chains which are uniformizable [16]. Section 3 describes stationary quantities and shows which of them can be determined from an aggregated chain. In Section 4, the same is done for several transient measures. Nearly exact and ordinary lumpability are introduced in Section 5. A small example is presented in Section 6 and the last section contains the conclusions.

2 Lumpability and aggregation

Let X be a homogeneous, finite and irreducible Markov chain on state space $Z = \{1 \ldots n\}$ with an aperiodic transition matrix $\underline{P}$. Denote by $\underline{\pi}$ the row vector containing the equilibrium distribution of the chain ($\underline{\pi P} = \underline{\pi}$ and $\underline{\pi} e^T = 1.0$) and by $\underline{\pi}^k$ ($= \underline{\pi}^{k-1}\underline{P}$) the distribution after k jumps starting with an initial distribution $\underline{\pi}^0$. Obviously, $\lim_{k\to\infty} \underline{\pi}^k = \underline{\pi}$. Let $\Omega = \{\Omega(1) \ldots \Omega(N)\}$ be a partition of the state space such that

$$\Omega(I) \subseteq Z\ , \qquad \Omega(I) \neq \emptyset\ , \qquad \Omega(I) \cap \Omega(J) = \emptyset \qquad \cup_{I=1}^{N} \Omega(I) = Z \qquad (1)$$

for $I, J \in \{1 \ldots N\}$. n_I is the number of states in partition group $\Omega(I)$. We assume that states belonging to one partition group are grouped together, which yields the following structure on $\underline{P}$:

$$\underline{P} = \begin{pmatrix} \underline{P}_{1,1} & \cdots & \underline{P}_{1,N} \\ \cdots & & \cdots \\ \underline{P}_{N,1} & \cdots & \underline{P}_{N,N} \end{pmatrix} . \tag{2}$$

Submatrix $\underline{P}_{I,J}$ includes all transitions between states from $\Omega(I)$ and $\Omega(J)$. Let $\underline{\pi}_I$ ($\underline{\pi}_I^k$) be a subvector of $\underline{\pi}$ ($\underline{\pi}^k$) including the state probabilities for all states $i \in \Omega(I)$ (i.e. $\underline{\pi} = (\underline{\pi}_1 \ldots \underline{\pi}_N)$). According to Ω the distributions over the partition groups $\underline{\Pi}$ and $\underline{\Pi}^k$ are given by

$$\Pi(I) = \sum_{i \in \Omega(I)} \pi(i) , \qquad \Pi^k(I) = \sum_{i \in \Omega(I)} \pi^k(i) , \qquad \text{for } 1 \leq I \leq N . \tag{3}$$

The conditional stationary distribution inside partition group I is defined as

$$\underline{\bar{\pi}}_I = \frac{\underline{\pi}_I}{\underline{\pi}_I \underline{e}^T} . \tag{4}$$

Partition Ω can be used to construct an aggregated Markov chain $\hat{X}$ on state space $\hat{Z} = \{1 \ldots N\}$ by substituting each subset of states $\Omega(I)$ by a single state. The transition matrix of the aggregated Markov chain $\underline{\hat{P}}$ is generated using a collector matrix $\underline{V}$ and a distributor matrix $\underline{W}$ as shown in the following equation.

$$\underline{\hat{P}} = \underline{WPV} \tag{5}$$

where $\underline{V} \in \mathbb{R}^{n \times N}$, $V(i,I) = 1$ if $i \in \Omega(I)$ and 0 otherwise, $\underline{W} = \underline{D}^{-1}(\underline{V})^T$, $\underline{D} = diag(\underline{\alpha V})$ and $\underline{\alpha} > \underline{0}$, where $diag(\underline{\alpha})$ is a diagonal matrix with $\alpha(i)$ in position i.

The vector $\underline{\alpha}$ assigns a non-zero weight to each state $i \in Z$. The weights determine the transition probabilities in the transition matrix of the aggregated Markov chain.

Theorem 1 *$\underline{P}$ is an irreducible transition matrix $\Rightarrow$ $\underline{\hat{P}}$ is an irreducible transition matrix for each partition Ω and each weight vector $\underline{\alpha} > \underline{0}$.*

Proof:

We have to show that for each pair $I_1, I_J \in \{1 \ldots N\}$ exists a sequence of states $\{I_1, \ldots, I_J\}$ such that $\hat{P}(I_{j-1}, I_j) > 0.0$ $(1 < j \leq J)$. For each pair of states $i_1 \in \Omega(I_1)$ and $i_J \in \Omega(I_J)$ exists a sequence of states $\{i_1 \ldots i_J\}$ with $P(i_{j-1}, i_j) > 0.0$ $(1 < j \leq J)$ since $\underline{P}$ is irreducible . Now let I_j be chosen such that $i_j \in \Omega(I_j)$ $(1 < j < J)$. The transition probability $\hat{P}(I_{j-1}, I_j)$ is given by

$$\sum_{i \in \Omega(I_{j-1})} \frac{\alpha(i)}{\sum\limits_{l \in \Omega(I_{j-1})} \alpha(l)} \sum_{k \in \Omega(I_j)} P(i,k) .$$

Since

$$\frac{\alpha(i)}{\sum\limits_{k \in \Omega(I_{j-1})} \alpha(k)} > 0.0$$

for all i and at least one $P(i,k) > 0.0$ ($P(i_{j-1}, i_j)$), also $\hat{P}(I_{j-1}, I_j) > 0.0$. □

Notice that the opposite direction in theorem 1 is not true. Since we have assumed that $\underline{P}$ is irreducible, also $\underline{\hat{P}}$ is irreducible and the stationary solution of the aggregated Markov chain $\underline{\hat{\pi}}$ exists uniquely as $\underline{\hat{\pi}}\underline{\hat{P}} = \underline{\hat{\pi}}$ and $\underline{\hat{\pi}}\underline{e}^T = 1.0$. In the same way the distribution of the aggregated Markov chain after k jumps starting with a distribution $\underline{\hat{\pi}}^0$ is defined as $\underline{\hat{\pi}}^k$ $(= \underline{\hat{\pi}}^{k-1}\underline{\hat{P}})$.

The usability of the aggregated Markov chain instead of the original Markov chain depends on the relation between the results of the two Markov chains, in particular, on the relation between $\underline{\pi}$, $\underline{\Pi}$ and $\underline{\hat{\pi}}$ for stationary analysis and between $\underline{\pi}^k$, $\underline{\Pi}^k$ and $\underline{\hat{\pi}}^k$ for transient analysis. Sections 3 and 4 are devoted to these relations. If we assume an arbitrary partition Ω and an arbitrary weight vector $\underline{\alpha}$, there is, of course, no relation between both Markov chains. Therefore we consider only special types of aggregated processes.

The first aggregate is the "ideal aggregate" or "pseudo aggregate" [6, 13] resulting from an arbitrary partition Ω and a weight vector $\underline{\alpha} = \underline{\pi}$. The transition matrix of this aggregated Markov chain is denoted by $\underline{\hat{P}}_{ID}$. Ideal aggregates can be generated for all partitions, however, the construction requires the knowledge of the stationary distribution for the original Markov chain and is therefore only of theoretical interest. More important are aggregates resulting from partitions fulfilling special properties, namely lumpability.

Definition 1 *Let $\underline{P}$ be the irreducible transition matrix of a finite Markov chain X on state space Z and $\Omega = \{\Omega(1) \ldots \Omega(N)\}$ a partition of the state space with collector matrix $\underline{V}$.*

- *Ω is ordinarily lumpable, iff for all $I \in \{1 \ldots N\}$ and all $i, j \in \Omega(I)$: $(\underline{e}_i - \underline{e}_j)\underline{PV} = \underline{0}$*
- *Ω is exactly lumpable, iff for all $I \in \{1 \ldots N\}$ and all $i, j \in \Omega(I)$: $(\underline{e}_i - \underline{e}_j)\underline{P}^T\underline{V} = \underline{0}$*
- *Ω is strictly lumpable, iff it is ordinarily and exactly lumpable.*

$\underline{e}_i$ is a row vector with 1.0 in position i and 0 elsewhere.

The above definition of ordinary/exact lumpability defines unique constants $\xi_{I,J} = \underline{e}_i\underline{P}_{I,J}\underline{e}^T$ for ordinarily lumpable partitions and $\eta_{I,J} = \underline{e}\underline{P}_{I,J}\underline{e}_i{}^T$ for exactly lumpable partitions, which are independent from $i \in \Omega(I)$.

Although we consider only discrete-time Markov chains in this paper, the results hold also for the continuous-time case. Let Y be a continuous-time Markov chain with generator matrix $\underline{Q}$ and let the transition rates $Q(i,j)$ be bounded by some positive constant $q < \infty$ (i.e. the Markov chain is uniformizable). Then there exists a discrete time Markov chain X with transition matrix $\underline{P} = \underline{Q}/q + \underline{I}$ which is strongly related to the continuous time Markov chain (see [16]), since

- the stationary solution vectors are identical;

- the distribution vector $\underline{\pi}(t)$ of the continuous time Markov chain at time t starting with $\underline{\pi}^0$ at time 0 is given by $\underline{\pi}(t) = \sum_{k=0}^{\infty} exp(-qt)[(qt)^k/k!]\underline{\pi}^k$, where $\underline{\pi}^k$ is the distribution vector of X after k jumps;
- a partition Ω that is ordinarily/exactly lumpable on $\underline{Q}$ is also ordinarily/exactly lumpable on $\underline{P}$ and vice versa.

3 Stationary measures

Stationary measures are related to the knowledge of the aggregated state probabilities $\underline{\Pi}$ or the overall stationary solution vector $\underline{\pi}$. Which of both vectors is needed depends on the goal of the analysis and the partition used for aggregation. In the remainder of this section we show which information $\underline{\hat{\pi}}$, the stationary solution vector of the aggregated Markov chain, provides about $\underline{\Pi}$ and $\underline{\pi}$ for aggregated Markov chains resulting from ordinarily/exactly lumpable partitions and for the ideal aggregate. The results are borrowed from Schweitzer [15] and Sumita, Rieders [16].

Theorem 2 [[16], §2] *If $\underline{\hat{P}}_{ID}$ is an ideal aggregate, then $\underline{\hat{\pi}} = \underline{\Pi}$.*

The ideal aggregate yields an exact computation of the vector of aggregated state probabilities, the stationary state probabilities $\underline{\pi}$ cannot be determined with the knowledge $\underline{\hat{\pi}}$ and $\underline{\hat{P}}_{ID}$ (of course, $\underline{\pi}$ has to be known to construct an ideal aggregate).

Theorem 3 [[15], Section 5] *If Ω is an exactly lumpable partition on the state space Z of a finite Markov chain with transition matrix $\underline{P}$ and $\underline{\hat{P}} = \underline{WPV}$ is the transition matrix of the aggregated chain according to Ω and resulting from $\underline{W} = diag(\underline{eV})^{-1}\underline{V}^T$, then $\underline{\hat{\pi}} = \underline{\Pi}$ and $\underline{\pi}_I = \hat{\pi}(I)/n_I\underline{e}$.*
The elements of $\underline{\hat{P}}$ are given by $\hat{P}(I,J) = \frac{n_J}{n_I}\eta_{IJ}$.

An aggregated Markov chain resulting from an exactly lumpable partition allows the computation of the stationary solution vector $\underline{\pi}$ without an error. Aggregated Markov chains resulting from ordinarily lumpable partitions do not provide such detailed results, but the aggregated state probabilities can be determined without an error.

Theorem 4 [[16], §3] *If Ω is an ordinarily lumpable partition on the state space Z of a finite Markov chain with transition matrix $\underline{P}$, then $\underline{\hat{P}} = \underline{WPV}$ the transition matrix of the aggregated chain according to Ω is independent from the weight vector $\underline{\alpha}$ and $\underline{\hat{\pi}} = \underline{\Pi}$.*
The elements $\hat{P}(I,J)$ are equal to $\xi_{I,J}$.

If one is interested in the whole stationary solution vector or parts of this vector, aggregated Markov chains resulting form ordinarily lumpable partitions provide no exact results, only bounds can be determined. A trivial and usually unsatisfactory bound for the elements of $\underline{\pi}$ is given by

$0.0 < \pi_I(i) < \hat{\pi}(I)$. Tighter bounds for $\underline{\pi}_I$ are gained using the submatrices $\underline{P}_{I,J}$ and $\underline{P}_{J,I}$ as shown in [7], which requires, of course, the knowledge of the transition matrix for the overall Markov chain, or at least of parts of this matrix. However, ordinary lumpability provides no special improvements concerning the bounds on $\underline{\bar{\pi}}_I$, which lie in the same polyhedron as in the general case. The bounds on $\underline{\pi}$ are, of course, improved by the exact knowledge of $\underline{\Pi}$. We come back to the determination of bounds in section 5.

4 Transient measures

We consider two different kinds of transient measures. The first are related to the behaviour of the ergodic Markov chain with transition matrix $\underline{P}$ as given in (2). We are interested in $\underline{\pi}^k$, the distribution of the Markov chain after k jumps starting with $\underline{\pi}^0$ and its relation to the aggregated state vector $\underline{\hat{\pi}}^k$. From an application viewpoint the convergence of $\underline{\pi}^k$ to $\underline{\pi}$ is often important, since it describes the behaviour of the Markov chain after a disturbance, caused for instance by an overload or underload situation in the system being modelled.

The second kind of transient measures is related to a modified Markov chain with a set of absorbing states. We assume that the states from $\Omega(N)$ are absorbing, yielding the transition matrix

$$\underline{P}^* = \begin{pmatrix} \underline{P}_{1,1} & \cdots & \cdots & \underline{P}_{1,N} \\ \cdots & & & \cdots \\ \underline{P}_{N-1,1} & \cdots & \cdots & \underline{P}_{N-1,N} \\ \underline{0} & \cdots & \underline{0} & \underline{I} \end{pmatrix} . \tag{6}$$

The modified process can be used to determine the distribution of absorption time and describes measures like sojourn times, probability flows, up or down times, availability and similar quantities of the modelled system.

In all cases considered here, no information about the transient behaviour can be gained from the ideal aggregate. However, aggregates resulting from ordinary/exact lumpability provide results in the transient case also. We start with the first kind of measures on the ergodic Markov chain.

Theorem 5 *If Ω is an ordinarily lumpable partition on the state space Z of a finite Markov chain with transition matrix $\underline{P}$ and $\underline{\hat{P}}$ is the transition matrix of the aggregated Markov chain according to Ω as defined in Theorem 4, then $\underline{\hat{\pi}}^k = \underline{\Pi}^k$ for $\underline{\hat{\pi}}^0 = \underline{\pi}^0\underline{V} = \underline{\Pi}^0$.*

Proof:
We prove the result by induction over k and notice that it holds for $k = 0$. Assume that it has been proved for $k-1$, then $\underline{\hat{\pi}}^k = \underline{\hat{\pi}}^{k-1}\underline{\hat{P}} = \underline{\pi}^{k-1}\underline{VWPV}$. Following [8], $\underline{VWPV} = \underline{PV}$ is

equivalent to ordinary lumpability, therefore the last term equals $\underline{\pi}^{k-1}\underline{PV} = \underline{\Pi}^k$. □

The aggregated state vector depends only on the aggregated initial vector. When bounds on the individual state probabilities are computed we obtain, as for the stationary analysis, $0.0 \leq \pi_I^k(i) \leq \hat{\pi}^k(I)$. With the knowledge of the transition matrix of the overall chain the upper bound can be improved to $\pi_I^k(i) \leq \sum_{J=1}^{N} \hat{\pi}^{k-1}(J) \max_{j \in \Omega(J)} P_{J,I}(j,i)$. The upper bound is computed from $\underline{\hat{\pi}}^{k-1}$ without any knowledge of bounds for $\underline{\pi}^{k-1}$. This is especially useful if the dimension of the original Markov chain is very large, and bounds have to be computed only for a few individual state probabilities and a specific value of k. The computation of better bounds is outlined in [15]; however, these bounds require an approximation based on the whole vector $\underline{\pi}^{k-1}$ and are therefore behind the scope of this paper, where results are computed using the aggregated and not the original Markov chain.

For exactly lumpable partitions exact results for individual probabilities of the original Markov chain can be computed for a special class of initial vectors, namely for all vectors for which the partition is weakly lumpable.

Theorem 6 *If Ω is an exactly lumpable partition on the state space Z of a finite Markov chain with transition matrix $\underline{P}$, and $\underline{\hat{P}}$ is the transition matrix of the aggregated Markov chain according to Ω as defined in Theorem 3, then $\underline{\hat{\pi}}^k = \underline{\Pi}^k$ and $\underline{\pi}_I^k = \hat{\pi}^k(I)/n_I \; \underline{e}$ for all initial distributions $\underline{\hat{\pi}}^0 = \underline{\pi}^0\underline{V}$ with $\underline{\pi}_I^0 = \gamma_I \underline{e}$ and $\gamma_I \geq 0.0$, $\sum_{I=1}^{N} \gamma_I = 1.0$.*

Proof:
Assume that the result holds for $k-1$. By induction we show that it also holds for k:

$$\begin{aligned} \underline{\pi}_I^k &= \sum_{J=1}^{N} \underline{\pi}_J^{k-1}\underline{P}_{J,I} = \sum_{J=1}^{N} \hat{\pi}^{k-1}(J)/n_J \; \underline{eP}_{J,I} \\ &= \sum_{J=1}^{N} \hat{\pi}^{k-1}(J)\eta_{J,I}/n_J \; \underline{e} = \sum_{J=1}^{N} \hat{\pi}^{k-1}(J)/n_I \hat{P}(J,I) \, \underline{e} = \hat{\pi}^k(I)/n_I \underline{e} \; . \end{aligned}$$

Since by assumption the result holds for $k = 0$, the proof is complete. □

The class of allowed initial vectors is of practical importance and includes distributions that describe at system level situations where identical components are in an identical state (e.g. all components in a reliability model are up).

The results for the absorbing process are, of course, very similar. Before we give the related theorems, some new quantities have to be defined. The transition matrix of the aggregated Markov chain related to $\underline{P}^*$ is given by

$$\underline{\hat{P}}^* = \begin{pmatrix} \hat{P}(1,1) & \dots & \dots & \hat{P}(1,N) \\ \dots & & & \dots \\ \hat{P}(N-1,1) & \dots & \dots & \hat{P}(N-1,N) \\ 0 & \dots & 0 & 1.0 \end{pmatrix} . \tag{7}$$

Since we are interested in the absorption time distribution of the process, we define $\tau^k = (\underline{\pi}_N^k - \underline{\pi}_N^{k-1})\underline{e}^T$ and $\hat{\tau}^k = \hat{\pi}^k(N) - \hat{\pi}^{k-1}(N)$ (for $k > 0$), the probability flow in the absorbing state(s) in the k-th jump. The following theorems show the relation between τ^k and $\hat{\tau}^k$ for ordinarily and exactly lumpable partitions.

Theorem 7 *If Ω is an ordinarily lumpable partition on the state space Z of a finite Markov chain with the modified transition matrix $\underline{P}^*$ and $\underline{\hat{P}}^*$ is the modified transition matrix of the aggregated Markov chain according to Ω, then $\hat{\tau}^k = \tau^k$ for $\underline{\hat{\pi}}^0 = \underline{\pi}^0\underline{V} = \underline{\Pi}^0$.*

Proof:
The proof follows immediately from the proof of Theorem 5. □

Theorem 8 *If Ω is an exactly lumpable partition on the state space Z of a finite Markov chain with the modified transition matrix $\underline{P}^*$, and $\underline{\hat{P}}^*$ is the modified transition matrix of the aggregated Markov chain according to Ω, then $\hat{\tau}^k = \tau^k$ for all initial distributions $\underline{\hat{\pi}}^0 = \underline{\pi}^0\underline{V}$ with $\underline{\pi}_I^0 = \gamma_I\underline{e}$ and $\gamma_I \geq 0.0$, $\sum_{I=1}^{N} \gamma_I = 1.0$.*

Proof:
The proof follows immediately from the proof of theorem 6. □

Typical initial distributions assign zero probabilities to states in the absorbing set $\Omega(N)$. For example, the first entrance into a state from $\Omega(N)$ starting from the stationary state under the condition that the process is not in a state from $\Omega(N)$ or the time between exit to and entry from $\Omega(N)$, are important. The former initial distribution is given by

$$\underline{\pi}_I^0 = \underline{\pi}_I/(1.0 - \underline{\pi}_N\underline{e}^T) \text{ for } I < N \text{ and } \underline{0} \text{ otherwise,}$$

the latter initial distribution can be computed as

$$\underline{\pi}_I^0 = \underline{\pi}_N\underline{P}_{NI}/(\underline{\pi}_N \sum_{J=1}^{N-1} \underline{P}_{NJ}\underline{e}^T) \text{ for } I < N \text{ and } \underline{0} \text{ otherwise.}$$

For exactly lumpable partitions, both initial distributions observe the conditions of Theorem 8.

5 Near lumpability and approximative aggregation

A natural way to extend the concepts of exact and ordinary lumpability is to allow a small difference for elements of one partition according to the column/row sum criterion, yielding nearly exactly/ordinarily lumpable partitions.

Definition 2 *Let $\underline{P}$ be the irreducible transition matrix of a finite Markov chain X on state space Z and $\Omega = \{\Omega(1)\ldots\Omega(N)\}$ a partition of the state space with collector matrix $\underline{V}$.*

- *Ω is nearly ordinarily lumpable, iff for all $I \in \{1\ldots N\}$ and all $i,j \in \Omega(I)$: $(\underline{e}_i - \underline{e}_j)\underline{PV} < \epsilon\underline{e}$;*
- *Ω is nearly exactly lumpable, iff for all $I \in \{1\ldots N\}$ and all $i,j \in \Omega(I)$: $(\underline{e}_i - \underline{e}_j)\underline{P}^T\underline{V} < \epsilon\underline{e}$;*
- *Ω is nearly strictly lumpable, iff it is nearly exactly and nearly ordinarily lumpable*

for $0 < \epsilon << 1.0$.

A concept similar to nearly exact lumpability has been published by Schweitzer [15]. It differs from nearly exactly lumpability slightly since the value ϵ is defined as $\max_{1\le I\le N} d(I)$, where $\underline{d} = (\underline{V}^T(\underline{P}^T\underline{V} - \underline{V}\hat{\underline{P}}^T)\underline{e}^T)$ for some $\hat{\underline{P}} = \underline{WPV}$. The advantage of our definition is that it is independent from a particular aggregated matrix $\hat{\underline{P}}$, and ϵ corresponds directly to the maximum difference of elements in the bounding matrices defined below. Nevertheless, both concepts are based on the same underlying ideas.

If Ω is a nearly exactly/ordinarily lumpable partition, $\underline{P}$ can be represented as $\underline{A} + \epsilon\underline{B}$, where $\underline{A}$ is a matrix which is exactly/ordinarily lumpable according to Ω. Following theorem 3 (or Theorem 4, respectively) an aggregated Markov chain with transition matrix $\hat{\underline{A}}$ is generated that can be used to approximate the behaviour of the original chain. The aggregated Markov chain is used to determine the behaviour of a chain with transition matrix $\underline{A}$ and the results are exact, in the sense defined in the previous section, for this Markov chain. If ϵ is sufficiently small, these results approximate the behaviour of the Markov chain with transition matrix $\underline{P}$ very well, as known from perturbation theory. The emphasis of the remainder of this section is on the determination of bounds on the error made when calculating results for the overall chain from the aggregated chain, as introduced previously for the case of exactly/ordinarily lumpable partitions.

If Ω is a nearly ordinarily lumpable partition, the transition matrix $\hat{\underline{P}}$ depends on the weight vector $\underline{\alpha}$ used for aggregate construction. Possible transition probabilities between the states of the aggregated Markov chain are defined as

$$\Xi_{I,J} = \underline{P}_{I,J}\underline{e}^T \ , \qquad \xi_{I,J}^{min} = \min_{i\in\Omega(I)} \Xi_{I,J}(i) \ , \qquad \xi_{I,J}^{max} = \max_{i\in\Omega(I)} \Xi_{I,J}(i) \ . \tag{8}$$

Two matrices $\underline{L}$ and $\underline{U}$ are defined as $L(I,J) = \xi_{I,J}^{min}$ and $U(I,J) = \xi_{I,J}^{max}$. Obviously, $\underline{L} \leq \underline{\hat{P}} \leq \underline{U}$ and $L(I,J) < \hat{P}(I,J) < U(I,J)$, if $L(I,J) < U(I,J)$. The elements of $\underline{L}$ and $\underline{U}$ differ at most by ϵ. Varying the weight vector $\underline{\alpha}$, each matrix inside the polyhedra defined by $\underline{L}$ and $\underline{U}$ can be generated. $\underline{L}$ and $\underline{U}$ can often be computed very efficiently from the specification of the model (e.g., a queueing network or stochastic Petri net) without generation of the large matrix $\underline{P}$. Bounds for the transient quantities $\underline{\Pi}^k$ are calculated straightforwardly from the matrices $\underline{L}$ and $\underline{U}$, as shown in the following theorem.

Theorem 9 *If Ω is a nearly ordinarily lumpable partition on the state space Z of a finite Markov chain with transition matrix $\underline{P}$, and $\underline{L}/\underline{U}$ are lower/upper bound matrices for the transitions of the aggregated Markov chain according to Ω, then $\underline{\hat{\pi}}_{min}^k \leq \underline{\Pi}^k \leq \underline{\hat{\pi}}_{max}^k$ for $\underline{\hat{\pi}}_{min}^0 = \underline{\hat{\pi}}_{max}^0 = \underline{\pi}^0\underline{V} = \underline{\Pi}^0$, where $\underline{\hat{\pi}}_{min}^k = \underline{\hat{\pi}}_{min}^{k-1}\underline{L}$ and $\underline{\hat{\pi}}_{max}^k = \underline{\hat{\pi}}_{max}^{k-1}\underline{U}$ $(k \geq 1)$.*

Proof:
We prove the result for $\hat{\pi}_{min}^k(I)$ by induction over k and notice that it holds for $k = 0$. Now assume that it has been proofed for $k-1$, then

$$\hat{\pi}_{min}^k(I) = \sum_{J=1}^{N} \hat{\pi}_{min}^{k-1}(J)L(J,I) \leq \sum_{J=1}^{N} \sum_{j \in \Omega(J)} \pi_J^{k-1}(j)L(J,I) \leq \sum_{J=1}^{N} \underline{\pi}_J^{k-1}\underline{P}_{JI}\underline{e}^T = \Pi^k(I).$$

The proof for $\underline{\Pi}^k \leq \underline{\hat{\pi}}_{max}^k$ follows immediately. □

In the same way bounds $\hat{\tau}_{min}^k \leq \tau^k \leq \hat{\tau}_{max}^k$ can be determined from lower and upper bound matrices of the absorbing Markov chain. The difference $\hat{\tau}_{max}^k - \hat{\tau}_{min}^k$ and $\hat{\pi}_{max}^k(I) - \hat{\pi}_{min}^k(I)$ is obviously less or equal $k\epsilon$.

Bounds for the stationary case can also be computed from the matrices $\underline{L}$ and $\underline{U}$. The best known bounds are due to Courtois and Semal [7] and are based on the following matrices:

$$\underline{Z}_L = {\underline{\Sigma}_L}^{-1}(\underline{I} - \underline{L})^{-1} \qquad \underline{Z}_U = {\underline{\Sigma}_U}^{-1}(\underline{I} - \underline{U})^{-1} \ , \tag{9}$$

where $\underline{\Sigma}_L = diag((\underline{I} - \underline{L})^{-1}\underline{e}^T)$ and $\underline{\Sigma}_U = diag((\underline{I} - \underline{U})^{-1}\underline{e}^T)$.

Define $\mathcal{J} \subseteq \{1 \ldots N\}$ as a set of column indices including all columns J with at least one element (I,J) with $L(I,J) < U(I,J)$. The inverse of the matrix $\underline{I} - \underline{L}$ always exists and is positive (see [7]). Using $\underline{Z}_L$ and $\mathcal{J}$, the following bounds for $\underline{\Pi}$ can be determined.

$$\begin{aligned} \hat{\pi}_{min}^L(I) &= \max(\min_{J \in \mathcal{J}} Z_L(J,I),\ 1.0 - \sum_{K \neq I} \max_{J \in \mathcal{J}} Z_L(J,K)) \ , \\ \hat{\pi}_{max}^L(I) &= \min(\max_{J \in \mathcal{J}} Z_L(J,I),\ 1.0 - \sum_{K \neq I} \min_{J \in \mathcal{J}} Z_L(J,K)) \ , \\ \underline{\hat{\pi}}_{min}^L &\leq \underline{\Pi} \leq \underline{\hat{\pi}}_{max}^L \ . \end{aligned} \tag{10}$$

If the inverse of the matrix $\underline{I}-\underline{U}$ exists and all row sums and column sums are strictly negative, then another set of bounds $\hat{\pi}_{min}^U(I)$ and $\hat{\pi}_{max}^U(I)$ can be determined by using $\underline{Z}_U$ instead of $\underline{Z}_L$ in the above equation. The maximum of both lower bounds and the minimum of the upper bounds for each component yields improved bound vectors $\underline{\hat{\pi}}_{min}$ and $\underline{\hat{\pi}}_{max}$.

If the matrices $\underline{L}$ and $\underline{U}$ are both known, we can often improve the bounds of [7] as shown below. For the elements of the upper and lower bound matrices the following trivial relations have to hold: $L(I,J) \geq 1.0 - \sum_{K \neq J} U(I,K)$ and $U(I,J) \leq 1.0 - \sum_{K \neq J} L(I,K)$. If the conditions are not observed, the elements of the bounding matrices are improved to meet the relations.

Following [7] the I-th row of $\underline{Z}_L$ is the stationary vector of a stochastic matrix which results from $\underline{L}$ by increasing the elements of column I until the resulting matrix is stochastic. Thus, the stationary vector of any stochastic matrix elementwise larger or equal to $\underline{L}$ can be expressed as a linear combination of the rows of $\underline{Z}_L$. It has been shown by Courtois that the bounds given in (10) are the best, if only $\underline{L}$ is known (it should be noticed that in this case $\mathcal{J}$ equals $\{1 \ldots N\}$). However, the additional knowledge of $\underline{U}$ increases the available information and this is only partially considered by defining the set $\mathcal{J}$ as the set of columns where $\underline{L}$ and $\underline{U}$ include at least one different element. Obviously the set of stochastic matrices which has to be considered includes those matrices which are elementwise not smaller than $\underline{L}$ and, additionally, not larger than $\underline{U}$. This difference is not reflected in the isolated analysis of $\underline{Z}_L$ and $\underline{Z}_U$ since it is implicitly assumed that each element of the matrices can be increased/decreased to yield a stochastic matrix. The definition of $\mathcal{J}$ only captures those columns where both bounding matrices are equal. The question is whether we can get better bounds with an acceptable effort using both matrices $\underline{L}$ and $\underline{U}$. Of course, this depends on the concrete matrices but in many realistic cases the difference $U(I,J) - L(I,J)$ is rather small such that $1.0 - \sum_{K=1}^{N} L(I,K)$ has to be distributed over several elements. In what follows we assume that $\underline{Z}_U$ exists (i.e., the inverse of $(\underline{I} - \underline{U})$ exists and fulfills the above conditions). Following [7] the unknown vector $\underline{\Pi}$ can be expressed as

$$\underline{\Pi} = \underline{\beta}_L \underline{Z}_L = \underline{\beta}_U \underline{Z}_U \qquad \text{where } \underline{\beta}_L,\ \underline{\beta}_U \geq \underline{0} \text{ and } \underline{\beta}_U \underline{e}^T = \underline{\beta}_L \underline{e}^T = 1.0 \tag{11}$$

Obviously the following relation has to hold:

$$\underline{\beta}_L \underline{X}_{LU} = \underline{\beta}_U \qquad \text{where } \underline{X}_{LU} = {\underline{\Sigma}_L}^{-1} (\underline{I} - \underline{L})^{-1} (\underline{I} - \underline{U}) \underline{\Sigma}_U \tag{12}$$

From the previous equation follows that $\underline{\beta}_L \underline{X}_{LU} \geq \underline{0}$ has to hold to yield a valid vector $\underline{\beta}_U$. Thus, the unknown vector $\underline{\Pi}$ can only be expressed by those vectors $\underline{\beta}_L$ which also observe $\underline{\beta}_L \underline{X}_{LU} \geq \underline{0}$. Before we describe how to use this observation to improve the bounds, the following theorem gives some results about the matrix $\underline{X}_{LU}$.

Theorem 10 *For the matrix $\underline{X}_{LU}$ the following relations hold:*

1. $\underline{X}_{LU} \underline{e}^T = \underline{e}^T$.
2. *If* $U(I,J) = 1.0 - \sum_{K=1, K \neq J}^{N} L(I,K)$ *for all* $I \in \{1, \ldots, N\}$, *then* $X_{LU}(I,J) = 0$ *for* $I = J$ *and* $X_{LU}(I,J) > 0$ *for* $I \neq J$.
3. $X_{LU}(I,J) \geq 0$ *for* $I \neq J$.

4. $X_{LU}(J,J) < 0$ *if a* $I \in \{1,\ldots,N\}$ *with* $U(I,J) < 1.0 - \sum_{K=1,K\neq J}^{N} L(I,K)$ *exists.*

5. $X_{LU}(I,J) = 0$ *for* $I \neq J$ *and* $X_{LU}(I,J) < 0$ *for* $I = J$*, if* $L(I,J) = U(I,J)$ *for all* $J \in \{1,\ldots,N\}$*.*

Proof:
Define $\underline{Y} = (\underline{I} - \underline{L})^{-1}(\underline{I} - \underline{U})$. If $Y(I,J) < 0$, then $X_{LU}(I,J) > 0$, since by assumption the diagonal elements of $\underline{\Sigma}_L$ are positive and of $\underline{\Sigma}_U$ are negative. Let $\underline{N} = (\underline{I} - \underline{L})^{-1}$.

1. $\underline{X}_{LU}\underline{e}^T = \underline{Z}_L(\underline{I} - \underline{U})(\underline{I} - \underline{U})^{-1}\underline{e}^T = \underline{e}^T$.

2. $Y(I,J) = N(I,J) - \sum_{H=1}^{N} N(I,H)U(H,J) = \sum_{K=1,K\neq J}^{N} (\sum_{H=1}^{N} N(I,H)L(H,K) - N(I,K))$. Each term of the outer sum is 0 for $K \neq I$ and -1 for $K = I$. Therefore the whole sum is 0 for $I = J$ and -1 for $I \neq J$.

3. Since $L(I,J) \leq U(I,J)$
$Y(I,J) = N(I,J) - \sum_{H=1}^{N} N(I,H)U(H,J) \leq N(I,J) - \sum_{H=1}^{N} N(I,H)L(H,J) = 0.$

4. $Y(J,J) = N(J,J) - \sum_{H=1}^{N} N(J,H)U(H,J) > \sum_{K=1,K\neq J}^{N} (\sum_{H=1}^{N} N(J,H)L(H,K) - N(J,K)) = 0$

5. $Y(I,J) = N(I,J) - \sum_{H=1}^{N} N(I,H)U(H,J) = N(I,J) - \sum_{H=1}^{N} N(I,H)L(H,J) = 0$ for $I \neq J$ and 1 for $I = J$. □

The previous theorem shows that $\beta_L(J)$ has to be 0, if $L(I,J) = U(I,J)$ for all $J \in \{1,\ldots,N\}$ (see point 5). Define $\mathcal{I} \subseteq \mathcal{J}$ as the set of column indices J with $U(I,J) = 1 - \sum_{K=1,K\neq J}^{N} L(I,K)$ for all $I \in \{1,\ldots,N\}$. $\beta_L(J) = 1.0$, which describes the situation that the elements of the J-th column of $\underline{L}$ are increased to yield a stochastic matrix, is only allowed for $J \in \mathcal{I}$. If $J \in \mathcal{I}$, then $X_{LU}(J,J) = 0$ (see point 2) and $X_{LU}(I,J) \geq 0$ (see point 3), therefore $\underline{\beta}_L\underline{X}_{LU} = \underline{e}_J\underline{X}_{LU} \geq \underline{0}$. Otherwise $X_{LU}(J,J) < 0$ (see point 4) and $\underline{e}_J\underline{X}_{LU}$ includes one negative element, namely $X_{LU}(J,J)$.

Thus, the determination of bounds for $\Pi(J)$ becomes a linear programming problem:

- Minimize/maximize $\sum_{I\in\mathcal{J}} \beta_L(I)Z_L(I,J)$
- with respect to $\sum_{I\in\mathcal{J}} \beta_L(I) = 1.0$ and $\sum_{I\in\mathcal{J}} \beta_L(I)X_{LU}(I,K) \geq 0.0$ for $K = 1,\ldots,N$.

The optimization can be avoided, if for some $I \in \mathcal{I}$ $Z_L(I,J) \leq (\geq) Z_L(K,J)$ for all $K \in \mathcal{J}$ such that the minimum (maximum) is given by setting $\underline{\beta}_L = e_I$.

The outlined optimization procedure requires additional effort which equals the solution of a linear programming problem using an appropriate algorithm (e.g., the simplex algorithm) for each

component of the bounding vector to be optimized. However, the approach allows to capture upper and lower bounds for the elements of the matrix and results in tighter bounds when the difference between the elements of the lower and upper bounds matrix is smaller than the difference of the row sums from 1.0.

The quality of the bounds depends on several factors (see [7] for additional explanations), the width $|\underline{\hat{\pi}}_{max} - \underline{\hat{\pi}}_{min}|$ becomes smaller, if the difference $\underline{U} - \underline{L}$ decreases (or $\underline{L}$, $\underline{U}$ become "more stochastic"). Better bounding matrices can be found by first determining bounds for the conditional distributions $\underline{\bar{\pi}}_I$ using the same approach as outlined for bounding $\underline{\hat{\pi}}$. The approach is omitted here since it is described in detail in [7]. If bounds for $\underline{\bar{\pi}}_I$ are computed, these bounds can be combined with the bounds for $\underline{\hat{\pi}}$ yielding bounds for the complete vector $\underline{\pi}$. Nevertheless, bounding $\underline{\bar{\pi}}_I$ requires the generation of the matrix $\underline{P}$ and the computation of an inverse matrices of order n_I for each $I \in \{1, \ldots, N\}$. If we assume that near lumpability results from a large number of nearly identical components, then the dimension of $\underline{P}$ is much larger than the dimension of $\underline{\hat{P}}$ and the bounding matrices $\underline{L}$ and $\underline{U}$ often can be generated directly from the model specification. Therefore it is desirable to use only the matrices of the aggregated Markov chain.

For nearly exactly lumpable partitions the matrices $\underline{L}$ and $\underline{U}$ can be used for bounding the aggregation error, but the difference between the elements $U(I,J) - L(I,J)$ needs not to be in any relation to ϵ. Therefore the bounds might become poor. For bounding transient quantities matrices $\underline{\mathcal{L}}$ and $\underline{\mathcal{U}}$ which are determined from the minimum and maximum column sum are often better suited:

$$\begin{aligned} \underline{\Psi}_{I,J} = \underline{e}\underline{P}_{I,J} \qquad \eta^{min}_{I,J} = \min_{j \in \Omega(J)} \Psi_{I,J}(j) \qquad \eta^{max}_{I,J} = \max_{j \in \Omega(J)} \Psi_{I,J}(j)\ , \\ \mathcal{L}(I,J) = \tfrac{n_J}{n_I} \eta^{min}_{I,J} \qquad \mathcal{U}(I,J) = \tfrac{n_J}{n_I} \eta^{max}_{I,J}\ . \end{aligned} \tag{13}$$

With the above matrices bounds on $\underline{\pi}^k$ can be determined for a special class of initial distributions as shown in the following theorem.

Theorem 11 *If Ω is a nearly exactly lumpable partition on the state space Z of a finite Markov chain with transition matrix $\underline{P}$ and $\underline{\mathcal{L}}/\underline{\mathcal{U}}$ are lower/upper bound matrices as defined in (13), then $\underline{\hat{\pi}}^k_{min}(I)/n_I\ \underline{e} \le \underline{\pi}^k_I \le \underline{\hat{\pi}}^k_{max}(I)/n_I\ \underline{e}$ for $\underline{\hat{\pi}}^0_{min} = \underline{\hat{\pi}}^0_{max} = \underline{\pi}^0\underline{V} = \underline{\Pi}^0$ and $\underline{\pi}^0_I = \gamma_I \underline{e}$, $\gamma_I \ge 0.0$, $\sum_{I=1}^{N} \gamma_I = 1.0$, where $\underline{\hat{\pi}}^k_{min} = \underline{\hat{\pi}}^{k-1}_{min}\underline{\mathcal{L}}$ and $\underline{\hat{\pi}}^k_{max} = \underline{\hat{\pi}}^{k-1}_{max}\underline{\mathcal{U}}$ ($k \ge 1$).*

Proof:
We prove $\pi^k_I(i) \ge \hat{\pi}^k_{min}(I)/n_I$ by induction over k and notice that it holds for $k = 0$. Now assume that it has been proved for $k-1$, then

$$\begin{aligned} \pi^k(i) &= \sum_{J=1}^{N} \sum_{j \in \Omega(J)} \pi^{k-1}(j) P_{J,I}(j,i) \ge \sum_{J=1}^{N} \hat{\pi}^{k-1}_{min}(J)/n_J \sum_{j \in \Omega(J)} P_{J,I}(j,i) \\ &\ge \sum_{J=1}^{N} \hat{\pi}^{k-1}_{min}(J) \eta^{min}_{J,I}/n_J = \sum_{J=1}^{N} \hat{\pi}^{k-1}_{min}(J) \mathcal{L}(J,I)/n_I = \hat{\pi}^k_{min}(I)/n_I\ . \end{aligned}$$

The proof for $\pi_I^k(i) \leq \hat{\pi}_{max}^k(I)$ follows immediately. □

Bounds on τ^k can be found in the same way. However, although bounds for transient quantities of aggregated Markov chains resulting from nearly exactly lumpable partitions can be determined as for the nearly ordinarily lumpable case, bounds for $\underline{\Pi}$ cannot be calculated from the matrices $\underline{\mathcal{L}}$ and $\underline{\mathcal{U}}$. The problem is that the transition probabilities of the aggregated Markov chain depend on $\underline{\bar{\pi}}_I$ the distribution inside the partition groups, which depends on the structure of the whole matrix $\underline{P}$.

6 An application

The example to be presented describes two synchronized processes and is an extended version of the model published in [10]. Each process has a local clock and a copy of the value of the clock of the other process. Let C_i be the actual value of the clock of process $i = 1, 2$ and $\bar{C}_i$ be the value of the copy of clock i in process $j \neq i$. At the end of a time interval the clock of process i is incremented by o1 with probability $1 - f_i$. Additionally a message with the actual clock value is sent to the other process. With probability $1 - l_i$ the message from process i reaches j immediately. However, with probability f_i the clock is not incremented and, of course, no message is sent and with probability l_i a message is lost. To synchronize the processes, a process is blocked if the difference between his own clock and the image of the clock of his counterpart becomes too large (i.e., $C_i - \bar{C}_j < \Delta$ has to hold in a non-blocked state, where Δ is the upper bound by which the value of the local clock is allowed to exceed the copy of the clock value of the other process). A blocked process does not increment its clock but still transmits messages at the end of a time interval. For the definition of possible transitions we have to distinguish the cases $C_i - \bar{C}_j < \Delta$ (the process is not blocked) and $C_i - \bar{C}_j = \Delta$ (the process is blocked). In the former case the following transitions and transition probabilities are possible for process $i = 1, 2$.

1. The clock is incremented and the message is received by the other process with probability $p_1^i = (1 - f_i)(1 - l_i)$.
2. The clock is incremented but no message is received by the other process with probability $p_2^i = (1 - f_i)l_i$.
3. The clock is not incremented and no message is received by the other process with probability $p_3^i = 1 - p_1^i - p_2^i = f_i$.

In a blocked state the following two transitions are possible.

1. A message is received by the other process with probability $p_4^i = (1 - f_i)(1 - l_i)$.
2. No message is received by the other process with probability $p_5^i = 1 - p_4^i$.

The above specification describes a Markov chain with an infinite state space; however, a simple modification in the state description yields a finite state space without loosing any significant information. Define the following variables $S_1 = C_1 - \bar{C}_2$, $S_2 = C_2 - \bar{C}_1$ and $S_3 = \bar{C}_1 - \bar{C}_2$. Obviously the variables include all necessary information and since $|S_1|, |S_2| \leq \Delta$ and $-S_2 \leq S_3 \leq S_1$ the state space is finite. For $\Delta = 1$ the state space includes the following 10 states. States are described in the form (S_1, S_2, S_3).

1) $(0,0,0)$ 2) $(1,0,1)$ 3) $(1,-1,1)$ 4) $(0,1,-1)$ 5) $(1,1,0)$

6) $(1,0,0)$ 7) $(-1,1,-1)$ 8) $(0,1,0)$ 9) $(1,1,1)$ 10) $(1,1,-1)$

In what follows states are identified by their number. The transition matrix of the chain is shown below, diagonal elements are denoted by Σ since they are uniquely determined as the difference of 1.0 and the sum of non-diagonal elements.

$$\begin{pmatrix}
\Sigma & p_1^1 p_2^2 & p_1^1 p_3^2 & p_2^1 p_1^2 & p_2^1 p_2^2 & p_2^1 p_3^2 & p_3^1 p_1^2 & p_3^1 p_2^2 & - & - \\
- & \Sigma & - & - & - & - & p_1^2 & - & p_2^2 & - \\
p_1^2 & p_2^2 & \Sigma & - & - & - & - & - & - & - \\
- & - & p_1^1 & \Sigma & - & - & - & - & - & p_2^1 \\
p_4^1 p_4^2 & p_4^1 p_5^2 & - & p_5^1 p_4^2 & \Sigma & - & - & - & - & - \\
p_4^1 p_1^2 & p_4^1 p_2^2 & p_4^1 p_3^2 & p_5^1 p_1^2 & p_5^1 p_2^2 & \Sigma & - & - & - & - \\
p_1^1 & - & - & p_2^1 & - & - & \Sigma & - & - & - \\
p_1^1 p_4^2 & p_1^1 p_5^2 & - & p_2^1 p_4^2 & p_2^1 p_5^2 & - & p_3^1 p_4^2 & \Sigma & - & - \\
- & - & - & - & - & - & p_4^2 & - & \Sigma & - \\
- & - & p_4^1 & - & - & - & - & - & - & \Sigma
\end{pmatrix}.$$

If both processes are identical (i.e., $p_j^1 = p_j^2 = p_j$ $(j = 1 \ldots 5)$), then a strictly lumpable partition $\{(1),(2,4),(3,7),(5),(6,8),(9,10)\}$ exists. This partition results from the symmetry in the model specification. An aggregated Markov chain is generated by substituting each partition group through a single state, states are numbered according to the above ordering of partition groups. The transition matrix of the aggregate chain is given by

$$\begin{pmatrix}
\Sigma & 2p_1p_2 & 2p_1p_3 & {p_2}^2 & 2p_2p_3 & - \\
- & \Sigma & p_1 & - & - & p_2 \\
p_1 & p_2 & \Sigma & - & - & - \\
{p_4}^2 & 2p_4p_5 & - & \Sigma & - & - \\
p_1p_4 & p_1p_5 + p_2p_4 & p_3p_4 & p_2p_5 & \Sigma & - \\
- & - & p_4 & - & - & \Sigma
\end{pmatrix}.$$

The results of the aggregated chain determine stationary results of the original chain completely. Transient results can be computed for the aggregated state probabilities and for specific states of the original chain under the condition that the initial distribution assigns equal probabilities to states in a partition group.

ϵ	Courtois lower bound	optimized lower bound	exact value	optimized upper bound	Courtois upper bound
1.0e-5	6.0635e-1	6.0635e-1	6.0636e-1	6.0636e-1	6.0636e-1
1.0e-4	6.0629e-1	6.0630e-1	6.0632e-1	6.0634e-1	6.0635e-1
1.0e-3	6.0566e-1	6.0581e-1	6.0596e-1	6.0612e-1	6.0627e-1
1.0e-2	5.9934e-1	6.0080e-1	6.0238e-1	6.0400e-1	6.0543e-1
5.0e-2	5.7057e-1	5.7818e-1	5.8598e-1	5.9476e-1	6.0165e-1

Table 1: Bounds and exact value for $\pi(1)$ ($\Delta = 1$).

Now we assume that the clocks are reliable and only message transmission is unreliable (i.e. $f_1 = f_2 = 0, l_1 = l_2 = l > 0.0$). Thus, $p_1 = p_4 = (1-l)$, $p_2 = p_5 = l$ and $p_3 = 0.0$. In this case state 5 of the aggregated chain is no longer reachable and need not to be considered. For the remaining 5 states the partition $\{(1,4),(2,6),(3)\}$ is ordinarily but not exactly lumpable. The aggregated chain can be used to determine exact results for stationary and transient aggregated state probabilities. The matrix of the aggregated chain resulting from the partition is given by

$$\begin{pmatrix} \Sigma & 2l(1-l) & - \\ - & \Sigma & 1-l \\ 1-l & l & \Sigma \end{pmatrix} .$$

For $l_1 = l_2 = 0.5$ and $f_1 = f_2 = 0$, the partition $\{1,2,(3,6),4\}$ is exactly but not ordinarily lumpable. This partition can be used to generate an aggregated Markov chain for the determination of all stationary results and transient results starting with an initial vector $\underline{\pi}^0$ with $\hat{\pi}^0(3) = \hat{\pi}^0(6)$.

To show the effects of nearly lumpability we assume the two processes to be similar but not completely identical. Let $l_1 = l_2 = 0.1$, $f_1 = 0.1$ and $f_2 = 0.1 + \epsilon$. With these parameters the first partition is no longer strictly lumpable, but still nearly strictly lumpable for sufficiently small ϵ. In table 1 lower and upper bounds for the value of $\pi(1)$ ($S_1 = S_2 = S_3 = 0$) are given. The Courtois bounds are calculated as shown in [7] resp. equation (10) using the combination bounds from the matrices $\underline{Z}_L$ and $\underline{Z}_U$. Therefore these are the best bounds we can compute with the method. The optimized bounds are the bounds calculated from the optimization using $\underline{X}_{LU}$ and $\underline{Z}_L$.

The bounds are improved by the optimization, however, we cannot expect too much in this small example, since most of the elements of $\underline{L}$ and $\underline{U}$ are equal and the dimension of the matrices is very small. In Table 2 the values $\|\underline{\hat{\pi}}_{\max} - \underline{\hat{\pi}}_{\min}\|_\infty$ and $\|\underline{\hat{\pi}}_{\max} - \underline{\hat{\pi}}_{\min}\|_2$ for both sets of bounds are given. The results confirm the observation for a single component of the solution vector and show that bounds can be improved using the optimization procedure.

In Table 3 bounds for $\pi(1)$ in a system with $\Delta = 5$ are shown. In this case the original Markov chain includes 286 states, the reduced chain resulting from a strictly lumpable partition includes

	l_∞		l_2	
ϵ	Courtois	optimized	Courtois	optimized
1.0e-5	6.0600e-6	3.1662e-6	1.2051e-5	7.8638e-6
1.0e-4	6.0603e-5	3.1665e-5	1.2052e-4	7.8647e-5
1.0e-3	6.0630e-4	3.1692e-4	1.2062e-3	7.8731e-4
1.0e-2	6.0907e-3	3.1961e-3	1.2165e-2	7.9578e-3
5.0e-2	3.1089e-2	1.6578e-2	6.3171e-2	4.1691e-2

Table 2: Difference between lower and upper bounds vectors ($\Delta = 1$).

ϵ	Courtois lower bound	optimized lower bound	exact value	optimized upper bound	Courtois upper bound
1.0e-5	1.0609e-1	1.0610e-1	1.0612e-1	1.0614e-1	1.0618e-1
1.0e-4	1.0584e-1	1.0592e-1	1.0612e-1	1.0631e-1	1.0639e-1
1.0e-3	1.0335e-1	1.0409e-1	1.0603e-1	1.0803e-1	1.0883e-1
1.0e-2	8.0343e-2	8.3319e-2	1.0269e-1	1.1163e-1	1.1846e-1
5.0e-2	2.9429e-2	3.2683e-2	6.3927e-2	1.4020e-1	2.5302e-1

Table 3: Bounds and exact value for $\pi(1)$ ($\Delta = 5$).

only 160 states. The results are similar to the previous example with $\Delta = 1$, however, the absolute difference between lower and upper bound becomes larger. It is worthwhile to mention that the optimization procedure might be used only for specific states (i.e. in the previous example only for the first state).

The simple example is sufficient to show the idea of aggregation based on lumpability. The gains of the method increase, of course, with the number of identical or nearly identical components in a system. Realistic models often consist of very huge state spaces and several identical components allowing a significant reduction of the state space based on lumpability.

7 Conclusions

In this paper we investigate the aggregation of finite Markov chains based on exact and ordinary lumpability. It has been shown that various measures on the original Markov chain can be determined on the aggregated Markov chain without an error. The concepts of nearly ordinary and nearly exactly lumpability are introduced and easily to computable bounds on the aggregation error have been presented.

In practical applications from performance or reliability modelling aggregated Markov chains can often be generated straightforward from the "high-level" model specification using recently developed techniques (see [1, 2, 3, 5, 4, 14]). The aggregation is often based on strictly lumpable

partitions combining the advantages of exact and ordinary lumpability. Nevertheless, the usability of the aggregated Markov chain to determine results for the overall Markov chain is limited. Stationary quantities can be completely determined, transient measures are only computable as aggregated state probabilities or for special classes of initial distributions.

References

[1] AUPPERLE, B.E. AND MEYER, J.F. (1991) State space generation for degredable multiprocessor systems. *Proc. FCTS 21* Montreal, Canada.

[2] BUCHHOLZ, P. (1990) The aggregation of Markovian submodels in isolation. Universität Dortmund, Fachbereich Informatik, Forschungsbericht Nr. 369.

[3] BUCHHOLZ, P. (1992) Hierarchical Markovian models -symmetries and aggregation-. *Proc. 6th Internat. Conf. on Modelling Tools and Techniques for Performance Evaluation.* Edinburgh University Press.

[4] CHIOLA, G. AND FRANCESCHINIS, G. (1989) Colored GSPN models and automatic symmetry detection. *Proc. 3rd Internat. Workshop on Petri Nets and Performance Models* IEEE Press, New York.

[5] CHIOLA, G., DUTHEILLET, C., FRANCESCHINIS, G. AND HADDAD, S. (1991) Stochastic well-formed colored nets and multiprocessor modelling application. *In High Level Petri Nets*, ed. K. Jensen and G. Rosenberg. Springer-Verlag, Berlin

[6] COURTOIS, P. J. (1982) Error minimization in decomposable stochastic models. *In Applied Probability - Computer Science The Interface*, Vol. 1. Birkhäuser, Boston.

[7] COURTOIS, P. J. AND SEMAL, P. (1984) Bounds for the positive eigenvectors of nonnegative matrices and their approximation by decomposition. *J. Assoc. Comput. Mach.* **31**, 804-825.

[8] KEMENY, J. G. AND SNELL, J. L. (1976) *Finite Markov Chains.* Springer-Verlag, New York.

[9] NICOLA, V. F. (1989) Lumping in Markov reward processes. IBM Research Report RC 14719.

[10] PLATEAU, B. D. AND TRIPATHI, S. K. (1988) Performance analysis of synchronization for two communicating processes. *Performance Evaluation* **8**, 305-320.

[11] ROGERS, L. C. G. AND PITMAN, J. W. (1981) Markov functions; *Ann. Prob.* **9**, 573-582.

[12] RUBINO, G. AND SERICOLA, B. (1989) On weak lumpability in Markov chains; *J. Appl. Prob.* **26**, 446-457.

[13] RUBINO, G. AND SERICOLA, B. (1989) Sojourn times in finite Markov processes; *J. Appl. Prob.* **27**, 744-756.

[14] SANDERS, W. H. AND MEYER, J. F. (1991) Reduced base Models construction methods for stochastic activity networks; *IEEE J. Sel. Areas Comm.* **9**, 25-36.

[15] SCHWEITZER, P. (1984) Aggregation methods for large Markov chains; *In Mathematical Computer Performance and Reliability* ed. G. Iazeolla et al. Elsevier-North-Holland, Amsterdam.

[16] SUMITA, U. AND RIDERS, M. (1989) Lumpability and time reversibility in the aggregation-disaggregation method for large Markov chains; *Commun. Statist. - Stoch. Models* **5**, 63-81.