Composition and Equivalence of Markovian and Non-Markovian Models

Peter Buchholz
Department of Computer Science
TU Dortmund
Dortmund, Germany
Email: peter.buchholz@udo.edu

Miklós Telek
Department of Telecommunications
Technical University of Budapest
Budapest, Hungary
Email: telek@hit.bme.hu

***Abstract*—Compositional modeling and the aggregation of components according to equivalence relations based on stochastic bisimulation are often used to handle the problem of state space explosion in Markov models. The paper presents a general class of equivalence relations between Markov models that include stochastic bisimulation or lumpability as specific cases and proves the congruence property of the new equivalence with respect to the composition of components. It is shown that the equivalence relates Markovian and non-Markovian representations but requires some restrictions for the composition which are automatically observed if stochastic bisimulation is used as equivalence relation. Nevertheless, the approach offers the possibility of state space reduction beyond stochastic bisimulation without loosing the possibility of analyzing the resulting stochastic process by means of numerical methods.**

***Keywords*-Compositional Modeling, Equivalence, Aggregation, Markov Models, Rational Processes;**

I. Introduction

In the last two decades compositional modeling and the use of equivalence relations to reduce the state space of components that are afterwards composed to build a complex model are the major steps to handle state space explosion in Markov modeling. Several approaches exist [1], [2], [3], [4], [5], [6] which differ in various details but are based on two central steps, namely the state space reduction of components to a behaviorally equivalent component using stochastic bisimulation [7], [8] which is based on some form of lumpability [9], [10] and a composition of components which preserves the equivalence, i.e. the substitution of a component by an equivalent component results in an equivalent model. The advantage of this approach steams from the combinatorial growth of the state space in terms of the sizes of component state spaces. Consequently, if for a component with $m^{(1)}$ states an equivalent component with $n^{(1)}$ ($< m^{(1)}$) states can be found and the component is composed with another component with $m^{(2)}$ states, then the resulting model has $n^{(1)}m^{(2)}$ rather than $m^{(1)}m^{(2)}$ states. I.e., the state space is reduced by a factor $n^{(1)}/m^{(1)}$. If it is also possible to find for the second component an equivalent representation with $n^{(2)}$ ($< m^{(2)}$) states, then the overall state space reduction is $n^{(1)}n^{(2)}/(m^{(1)}m^{(2)})$.

The theoretical advantages of compositional model can be exploited in practice since efficient methods exist for computing the largest bisimulation relation for a component [11], [12], [13] which can then be used to generate the reduced component by state aggregation. Furthermore, the compositional model structure can be exploited for an efficient analysis by representing the generator matrix of the resulting Markov chain using Kronecker operations [14], [6], [15], graph based data structures like MTBDDs [16] or combinations of both [17].

Available results which use bisimulation for state space reduction are based on the observation that bisimulation is a congruence according to model composition. However, equivalence of component behavior goes beyond bisimulation. In [18] a general equivalence definition has been introduced for components with state spaces of the same sizes. This relation is defined algebraically and relates Markovian and non-Markovian representations. The non-Markovian representations are denoted as Rational Arrival Processes in the literature [19]. In [20], [21] it is shown that these more general process types can be analyzed numerically similarly to Markov chains. Recently, it has been shown in [21], [22] that the equivalence of [18] can be extended to components with different state space sizes and includes bisimulation, lumpability and weak lumpability as specific cases.

In this paper, we consider the general equivalence relations proposed in [21], [22] and analyze their compositionality. Interestingly, it turns out that although bisimulation is a congruence according to the composition operations this does not hold in all cases for the more general equivalence relations. Thus, we either have to restrict composition or the equivalence relations for compositional modeling.

In other words, there are cases when a Markovian component of size $m^{(1)}$ can not be represented with a smaller Markovian component, but there exists a non-Markovian representation of the same component of size $n^{(1)}$ ($n^{(1)} < m^{(1)}$). In these cases the composition of non-Markovian representation of components can avoid state space explosion. To apply this methodology the paper collects composition rules for non-Markovian components.

The paper is structured as follows. In the next section

the basic model class is defined. Section III introduces equivalence relations for our models. Afterwards, in section IV the preservation of equivalence after the composition of components is analyzed. In Section V, several examples are presented to show the possibilities of state space reduction based on equivalences. The paper ends with the conclusion.

II. Compositional Models

We begin with the introduction of Markov models that are afterwards extended to a class of non-Markovian models which is derived from rational arrival processes [19]. Finally, we consider the composition of components which results in a Markovian or non-Markovian model depending on whether the components are Markovian or non-Markovian.

A. Markov Components

A Markov component is defined as $\mathcal{A} = (\mathcal{S}, \pi, \mathbf{E}_e(e \in \mathcal{E}), \Lambda)$, where $\mathcal{S} = \{0, \ldots, m-1\}$ is the finite state space, $\pi \in \mathbb{R}^{1,m}$ is the initial probability distribution, $\mathcal{E}$ is a finite set of events, $\mathbf{E}_e \in \mathbb{R}^{m,m}$ is the transition weight matrix according to event e and $\Lambda = (\lambda_e(e \in \mathcal{E}))$ is a rate vector which contains for every event $e \in \mathcal{E}$ a rate $\lambda_e > 0$. Transition weights are non-negative, i.e., $\mathbf{E}_e(x,y) \geq 0$ for all $e \in \mathcal{E}$ and $x, y \in \mathcal{S}$, where $\mathbf{E}_e(x,y)$ refers to the x,y element of matrix $\mathbf{E}_e$. We number the elements in a vector or matrix from 0 to $m-1$. Furthermore, define

$$\mathbf{D}_e = diag(\mathbf{E}_e \mathbb{1}) \tag{1}$$

where $\mathbb{1}$ is a column vector of ones and $diag(\alpha)$ is a diagonal matrix with element $\alpha(x)$ in position (x,x). We assume that $\mathcal{E}$ contains a specific event ϵ that is not observable and we denote $\mathcal{E}_s = \mathcal{E} \setminus \{\epsilon\}$.

The behavior of the component is as follows. In state $x \in \mathcal{S}$ an exponential distribution with rate $\lambda_e \mathbf{D}_e(x,x)$ is enabled for event $e \in \mathcal{E}$. Consequently, we have possibly up to $|\mathcal{E}|$ exponential distributions running in parallel such that the sojourn time in the state is exponentially distributed with rate $\sum_{e \in \mathcal{E}} \lambda_e \mathbf{D}_e(x,x)$. If the exponential distribution for event e expires first, successor state y is chosen probabilistically with probability $\mathbf{E}_e(x,y)/\mathbf{D}_e(x,x)$. Observe that $x = y$ is included. The probability that $e \in \mathcal{E}$ expires first is given by $\lambda_e \mathbf{D}_e(x,x) / \left(\sum_{f \in \mathcal{E}} \lambda_f \mathbf{D}_f(x,x)\right)$. Since ϵ is not observable we define $\mathbf{Q}_\epsilon = \lambda_\epsilon(\mathbf{E}_\epsilon - \mathbf{D}_\epsilon)$.

Thus, we can define the stationary or transient behavior of a component. Let

$$\mathbf{Q} = \mathbf{Q}_\epsilon + \sum_{e \in \mathcal{E}_s} \lambda_e \left(\mathbf{E}_e - \mathbf{D}_e\right) \tag{2}$$

be the generator matrix of the Markov chain defined by the component. For stationary analysis we assume that $\mathbf{Q}$ is irreducible such that

$$\psi \mathbf{Q} = \mathbf{0} \text{ and } \psi \mathbb{1} = 1 \tag{3}$$

has a unique and non-negative solution which is the stationary vector of the Markov chain. From the stationary vector the stationary throughput of event $e \in \mathcal{E}_s$ can be computed as $\psi \mathbf{D}_e \mathbb{1}$ and the transient throughput of event $e \in \mathcal{E}_s$ at time t is $\pi e^{\mathbf{Q}t} \mathbf{D}_e \mathbb{1}$.

The joint density for a sequence of k observations $(e_1, t_1, e_2, t_2, \ldots, e_k, t_k)$, where $e_i \in \mathcal{E}_s$ is the event that is observed and $t_i \geq 0$ is the interevent time, is given by

$$f_{\mathcal{A}}(e_1, t_1, \ldots, e_k, t_k) = \pi \left(\prod_{i=1}^{k} e^{\mathbf{R}t_i} \lambda_{e_i} \mathbf{E}_{e_i} \right) \mathbb{1}, \tag{4}$$

where

$$\mathbf{R} = \mathbf{Q}_\epsilon - \sum_{e \in \mathcal{E}_s} \lambda_e \mathbf{D}_e. \tag{5}$$

Of course, by choosing $\pi = \psi$ the Markov renewal process starting in steady state is considered. Based on the definition of $\mathbf{R}$, $\mathbf{D}_e$ and the non-negativity of $\mathbf{E}_e$ it is straightforward to show [22] that $f_{(e_1,t_1,\ldots,e_k,t_k)}$ is a valid density, i.e. $f_{(e_1,t_1,\ldots,e_k,t_k)} \geq 0$ and

$$\int_{\tau_1} \cdots \int_{\tau_k} \sum_{e_1} \cdots \sum_{e_k} f_{(e_1,t_1,\ldots,e_k,t_k)} = 1.$$

We did not consider the specification of components yet. However, components can be easily described using stochastic automata [6], [5], stochastic process algebra [4], [8], stochastic Petri nets [2], [14], interactive Markov chains [3] or similar approaches.

B. Non-Markovian Components

The Markov models described above are often used in stochastic modeling and have a clear probabilistic interpretation. However, if one uses an algebraic interpretation of the vectors and matrices, then it is possible to consider non-Markovian representations. For a valid component it is only required that $f_{(e_1,t_1,\ldots,e_k,t_k)}$ as defined in (4) is a valid joint density. This yields the more general interpretation of Rational Arrival Processes as defined in [19]. Since we do not consider specifically arrival processes, we denote the corresponding processes as rational processes (RPs) in the sequel.

In contrast to Markov models matrices $\mathbf{E}_e$ and vector π of RPs may contain negative entries but we still require $\pi \mathbb{1} = 1$. Since matrix $\mathbf{Q}$ is computed as in (2), we also have $\mathbf{Q}\mathbb{1} = \mathbf{0}$. In contrast to the Markov case, RPs have no probabilistic interpretation at the state level but the observation of the occurrence of events still describes a stochastic process.

The more general class of components is more powerful than Markov models, but on the other hand, it is rather hard to check if a set of matrices and an initial vector define a valid joint density. Consequently, it is hard to define specification techniques to specify valid RPs. In this paper we go another way. We start with a Markov model and define equivalence relations that allow us to

relate Markovian and non-Markovian representations. Since equivalence means that the joint densities are identical, a non-Markovian model resulting from an equivalent Markov model is a valid stochastic model. We come back to this point in Section III.

C. Composition of Components

Composition is the major step to build complex models from simpler components. It alleviates to some extent state space explosion if the explicit generation of the composed state space can be avoided. We consider here the composition of two components. However, since the composition is associative, it can be extended to an arbitrary number of components.

Let $\mathcal{A}^{(1)} = (\mathcal{S}^{(1)}, \pi^{(1)}, \mathbf{E}_e^{(1)}(e \in \mathcal{E}), \Lambda)$ and $\mathcal{A}^{(2)} = (\mathcal{S}^{(2)}, \pi^{(2)}, \mathbf{E}_e^{(2)}(e \in \mathcal{E}), \Lambda)$ be two components that should be composed. We assume that the event sets $\mathcal{E}$ and the rate vectors Λ of length $|\mathcal{E}|$ are identical which can always be achieved by scaling the values in the matrices $\mathbf{E}_e^{(i)}$ and adding pseudo events with $\mathbf{E}_e^{(i)} = \mathbf{I}$ ($i = 1, 2$, $e \in \mathcal{E}_s$), where $\mathbf{I}$ is the identity matrix, if necessary. The size of a component is the cardinality of its state space and is denoted as $n^{(i)}$.

Composition is performed over the set of signals $\mathcal{E}$. Signals from $\mathcal{C} \subseteq \mathcal{E}_s$ occur as synchronized signals in both components whereas signals from $\mathcal{N} = \mathcal{E}_s \setminus \mathcal{C}$ and signal ϵ occur independently in the components. This is the usual way of defining composition in different modeling formalisms [4], [2], [14].

We define the composition now using Kronecker operations. This is, however, only a compact way to write composition down and prove some results later. The results concerning the preservation of equivalence after composition are, of course, independent of the mathematical formalism we use for its presentation.

Let $\mathcal{A}^{(0)} = \mathcal{A}^{(1)} \|_{\mathcal{C}} \mathcal{A}^{(2)}$ be the composed model that is considered over the state space $\mathcal{S}^{(0)} = \mathcal{S}^{(1)} \times \mathcal{S}^{(2)}$. For the definition of synchronization, the definition of the rate of synchronized transitions is one important design decision, several possibilities exist [23]. In conjunction with Kronecker representations usually the product of weights is used resulting in a Kronecker product of matrices [6], [15]. For $\mathbf{G} \in \mathbb{R}^{n,n}$, $\mathbf{H} \in \mathbb{R}^{m,m}$ the Kronecker product is defined as

$$\mathbf{F} = \mathbf{G} \otimes \mathbf{H} \in \mathbb{R}^{nm,nm}$$

such that $\mathbf{F}(i_a m + i_b, j_a m + j_b) = \mathbf{G}(i_a, j_a)\mathbf{H}(i_b, j_b)$ ($0 \leq i_a, j_a < n, 0 \leq i_b, j_b < m$). The Kronecker sum of matrices is defined as [15]

$$\mathbf{F} \oplus \mathbf{H} = \mathbf{F} \otimes \mathbf{I}_m + \mathbf{I}_n \otimes \mathbf{H},$$

where $\mathbf{I}_m$ is the identity matrix of size m.

Definition 1: The matrices of the composed model $\mathcal{A}^{(0)}$ are defined as

$$\begin{aligned}
\mathbf{E}_e^{(0)} &= \begin{cases} \mathbf{E}_e^{(1)} \oplus \mathbf{E}_e^{(2)} & \text{if } e \in \mathcal{N} \cup \{\epsilon\}, \\ \mathbf{E}_e^{(1)} \otimes \mathbf{E}_e^{(2)} & \text{if } e \in \mathcal{C}, \end{cases} \\
\mathbf{D}_e^{(0)} &= diag(\mathbf{E}_e^{(0)} \mathbb{1}) \qquad \text{if } e \in \mathcal{E}_s, \\
\mathbf{Q}_\epsilon^{(0)} &= \mathbf{Q}_\epsilon^{(1)} \oplus \mathbf{Q}_\epsilon^{(2)} = \lambda_\epsilon \left(\mathbf{E}_\epsilon^{(1)} \oplus \mathbf{E}_\epsilon^{(2)} - \mathbf{D}_\epsilon^{(1)} \oplus \mathbf{D}_\epsilon^{(2)} \right).
\end{aligned} \tag{6}$$

Based on this definition for $e \in \mathcal{N}$ we have

$$\begin{aligned}
\mathbf{D}_e^{(0)} &= diag(\mathbf{E}_e^{(0)} \mathbb{1}_{nm}) = diag((\mathbf{E}_e^{(1)} \oplus \mathbf{E}_e^{(2)})\ \mathbb{1}_{nm}) = \\
&= diag((\mathbf{E}_e^{(1)} \otimes \mathbf{I}_{n^{(2)}} + \mathbf{I}_{n^{(1)}} \otimes \mathbf{E}_e^{(2)})(\mathbb{1}_{n^{(1)}} \otimes \mathbb{1}_{n^{(2)}})) = \\
&= diag(\mathbf{E}_e^{(1)} \mathbb{1}_{n^{(1)}} \otimes \mathbb{1}_{n^{(2)}}) + diag(\mathbb{1}_{n^{(1)}} \otimes \mathbf{E}_e^{(2)} \mathbb{1}_{n^{(2)}}) \\
&= diag(\mathbf{E}_e^{(1)} \mathbb{1}_{n^{(1)}}) \otimes diag(\mathbb{1}_{n^{(2)}}) \\
&\quad + diag(\mathbb{1}_{n^{(1)}}) \otimes diag(\mathbf{E}_e^{(2)} \mathbb{1}_{n^{(2)}}) \\
&= diag(\mathbf{E}_e^{(1)} \mathbb{1}_{n^{(1)}}) \otimes \mathbf{I}_{n^{(2)}} + \mathbf{I}_{n^{(1)}} \otimes diag(\mathbf{E}_e^{(2)} \mathbb{1}_{n^{(2)}}) \\
&= diag(\mathbf{E}_e^{(1)} \mathbb{1}_{n^{(1)}}) \oplus diag(\mathbf{E}_e^{(2)} \mathbb{1}_{n^{(2)}}) = \mathbf{D}_e^{(1)} \oplus \mathbf{D}_e^{(2)},
\end{aligned}$$

and consequently

$$\begin{aligned}
&\mathbf{Q}_\epsilon^{(0)} - \sum_{e \in \mathcal{N}} \lambda_e \mathbf{D}_e^{(0)} = \\
&\left(\mathbf{Q}_\epsilon^{(1)} - \sum_{e \in \mathcal{N}} \lambda_e \mathbf{D}_e^{(1)} \right) \oplus \left(\mathbf{Q}_\epsilon^{(2)} - \sum_{e \in \mathcal{N}} \lambda_e \mathbf{D}_e^{(2)} \right).
\end{aligned} \tag{7}$$

In a similar way $\mathbf{D}_e^{(0)} = \mathbf{D}_e^{(1)} \otimes \mathbf{D}_e^{(2)}$ for $e \in \mathcal{C}$.

In stochastic automata networks (SANs) [6], [15] the basic rates λ_e are usually inserted in one of the matrices. It is also possible to extend the approach by defining some other form of computing the rates of synchronized transitions as for example done in the stochastic process algebra PEPA [4]. However, we will not consider this case here although it can be incorporated in the approach we present. The initial vector of $\mathcal{A}^{(0)}$ is given by

$$\pi^{(0)} = \pi^{(1)} \otimes \pi^{(2)} \tag{8}$$

such that $\pi^{(1)}\mathbb{1} = \pi^{(2)}\mathbb{1} = 1$ implies $\pi^{(0)}\mathbb{1} = 1$.

If $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$ are Markov models, then $\mathcal{A}^{(0)}$ is also a Markov model since $\pi^{(0)} \geq \mathbf{0}$ and $\mathbf{E}_e^{(0)} \geq \mathbf{0}$ in this case. For the moment we consider only the Markov case, the non-Markovian case will be described later in conjunction with equivalence relations. The generator matrix of the resulting Markov process is computed using (2). It should be remarked that $\mathcal{S}^{(0)}$ may contain unreachable states which implies that $\mathbf{Q}$ is not irreducible. Since we define the initial vector, transient and stationary distribution of the resulting process are well defined but for an efficient numerical computation it might be preferable to choose some other representation which contains only reachable states [14], [24], [17]. We will not consider these representations since the numerical analysis is not the topic of this paper and the necessary methods to handle unreachable states are available and established.

III. Equivalence Relations

We begin with a general definition of equivalent components of the same sizes [18]. Afterwards, two different equivalences relating components with different state space sizes are introduced which are both based on the general equivalence definition and allow state space reduction due to redundancy according to the closing or initial vector [22].

A. An Algebraic Equivalence Relation

We use the following definition of equivalence for two components and assume in the sequel that components are defined over the same event sets $\mathcal{E}$.

Definition 2: Two components $\mathcal{A}^{(1)} = (\mathcal{S}^{(1)}, \pi^{(1)}, \mathbf{E}_e^{(1)}(e \in \mathcal{E}), \Lambda)$ and $\mathcal{A}^{(2)} = (\mathcal{S}^{(2)}, \pi^{(2)}, \mathbf{E}_e^{(2)}(e \in \mathcal{E}), \Lambda)$ are equivalent, if and only if

$$f_{\mathcal{A}^{(1)}}(e_1, t_1, \ldots, e_k, t_k) = f_{\mathcal{A}^{(2)}}(e_1, t_1, \ldots, e_k, t_k)$$

for all $k > 0$, $e_i \in \mathcal{E}_s$ and $t_i > 0$.

Definition 3: Two vectors of size m, $\pi^{(1)}$ and $\pi^{(2)}$, and two ordered sets of matrices of size $m \times m$, $\mathcal{M}^{(1)}$ and $\mathcal{M}^{(2)}$, are in similarity relation C_0 if a non-singular $m \times m$ matrix $\mathbf{B}$ exists such that

- $\mathbf{B}\mathbb{1} = \mathbb{1}$,
- $\pi^{(2)} = \pi^{(1)}\mathbf{B}$ and
- $\mathbf{F}^{(2)} = \mathbf{B}^{-1}\mathbf{F}^{(1)}\mathbf{B}$, for all $\mathbf{F}^{(i)} \in \mathcal{M}^{(i)}$ $(i = 1, 2)$ associated matrices of the two sets.

Definition 4: The following sets of matrices are defined for component $\mathcal{A} = (\mathcal{S}, \pi, \mathbf{E}_e(e \in \mathcal{E}), \Lambda)$.

- $\mathcal{M}_\sim = \{\mathbf{Q}_\epsilon - \sum_{e \in \mathcal{E}_s} \lambda_e \mathbf{D}_e, \mathbf{E}_e(e \in \mathcal{E}_s)\}$,
- $\mathcal{M}_\simeq = \{\mathbf{Q}_\epsilon - \sum_{e \in \mathcal{N}} \lambda_e \mathbf{D}_e, \mathbf{E}_e(e \in \mathcal{E}_s), \mathbf{D}_e(e \in \mathcal{C})\}$,
- $\mathcal{M}_\approx = \{\mathbf{Q}_\epsilon, \mathbf{E}_e(e \in \mathcal{E}_s), \mathbf{D}_e(e \in \mathcal{E}_s)\}$.

$\mathcal{M}_\sim$ and $\mathcal{M}_\approx$ are independent of the set of synchronized signals $\mathcal{C} \subseteq \mathcal{E}_s$, but for $\mathcal{M}_\simeq$ set $\mathcal{C}$ and consequently $\mathcal{N} = \mathcal{E}_s \setminus \mathcal{C}$ have to be defined. For notational convenience this dependence is not indicated explicitly.

The following definition relates two components with the same state space $\mathcal{S}$.

Definition 5: Let $\mathcal{A}^{(1)} = (\mathcal{S}, \pi^{(1)}, \mathbf{E}_e^{(1)}(e \in \mathcal{E}), \Lambda)$ and $\mathcal{A}^{(2)} = (\mathcal{S}, \pi^{(2)}, \mathbf{E}_e^{(2)}(e \in \mathcal{E}), \Lambda)$ with $|\mathcal{S}| = m$ be two components. We define three equivalence relations between $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$ associated with similarity relation C_0.

- $\mathcal{A}^{(1)} \sim_0 \mathcal{A}^{(2)}$ iff C_0 holds for $(\pi^{(1)}, \mathcal{M}_\sim^{(1)})$ and $(\pi^{(2)}, \mathcal{M}_\sim^{(2)})$,
- $\mathcal{A}^{(1)} \simeq_0 \mathcal{A}^{(2)}$ iff C_0 holds for $(\pi^{(1)}, \mathcal{M}_\simeq^{(1)})$ and $(\pi^{(2)}, \mathcal{M}_\simeq^{(2)})$,
- $\mathcal{A}^{(1)} \approx_0 \mathcal{A}^{(2)}$ iff C_0 holds for $(\pi^{(1)}, \mathcal{M}_\approx^{(1)})$ and $(\pi^{(2)}, \mathcal{M}_\approx^{(2)})$.

From these definition it follows that $\approx_0 \Rightarrow \simeq_0 \Rightarrow \sim_0$, for $\mathcal{C} = \emptyset$, $\sim_0 \Leftrightarrow \simeq_0$ and for $\mathcal{C} = \mathcal{E}_s$, $\simeq_0 \Leftrightarrow \approx_0$. The relations may relate Markovian and non-Markovian representations. The following theorem shows, however, that independent of the representation, components in relation $\sim_0$ are equivalent which implies that components in relation $\simeq_0$ and $\approx_0$ are also equivalent since components that are in relation $\simeq_0$ or $\approx_0$ are also in relation $\sim_0$.

Theorem 1: If $\mathcal{A}^{(1)} \sim_0 \mathcal{A}^{(2)}$, then $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$ are equivalent.

Proof: We have to show that $f_{\mathcal{A}^{(1)}}((e_1, t_1, \ldots, e_k, t_k) = f_{\mathcal{A}^{(2)}}(e_1, t_1, \ldots, e_k, t_k)$ for all $k > 0$, $e_i \in \mathcal{E}_s$ and $t_i \geq 0$. Observe that $\mathbf{B}\mathbb{1} = \mathbb{1}$ implies (by multiplying both sides by $\mathbf{B}^{-1}$) $\mathbf{B}^{-1}\mathbb{1} = \mathbb{1}$. Let $\mathbf{R}^{(j)} = \mathbf{Q}_\epsilon^{(j)} - \sum_{e \in \mathcal{E}_s} \lambda_e \mathbf{D}_e^{(j)}$ $(j = 1, 2)$, then $\mathbf{R}^{(2)} = \mathbf{B}^{-1}\mathbf{R}^{(1)}\mathbf{B}$ holds due to $\mathcal{A}^{(1)} \sim_0 \mathcal{A}^{(2)}$. We have (see also [22])

$$
\begin{aligned}
& f_{\mathcal{A}^{(1)}}((e_1, t_1, \ldots, e_k, t_k) = \\
& \pi^{(1)} \left(\prod_{i=1}^{k} exp(\mathbf{R}^{(1)} t_i) \lambda_{e_i} \mathbf{E}_{e_i}^{(1)} \right) \mathbb{1} = \\
& \pi^{(1)} \left(\prod_{i=1}^{k} \sum_{j=0}^{\infty} \frac{(\mathbf{R}^{(1)} t_i)^j}{j!} \lambda_{e_i} \mathbf{E}_{e_i}^{(1)} \right) \mathbb{1} = \\
& \pi^{(1)} \mathbf{B} \left(\prod_{i=1}^{k} \sum_{j=0}^{\infty} \frac{(\mathbf{B}^{-1}\mathbf{R}^{(1)}\mathbf{B} t_i)^j}{j!} \lambda_{e_i} \mathbf{B}^{-1} \mathbf{E}_{e_i}^{(1)} \mathbf{B} \right) \mathbf{B}^{-1}\mathbb{1} = \\
& \pi^{(2)} \left(\prod_{i=1}^{k} \sum_{j=0}^{\infty} \frac{(\mathbf{R}^{(2)} t_i)^j}{j!} \lambda_{e_i} \mathbf{E}_{e_i}^{(2)} \right) \mathbb{1} = \\
& f_{\mathcal{A}^{(2)}}((e_1, t_1, \ldots, e_k, t_k) \ .
\end{aligned}
$$

■

The above relations can be used to prove whether $\mathcal{A}$ describes a valid process even if it is a non-Markovian representation. If it is possible to find some Markovian representation $\mathcal{A}'$ with $\mathcal{A} \sim_0 \mathcal{A}'$, then $\mathcal{A}$ describes a valid process since a Markov representation always defines a process. This approach is used in [18] to transform non-Markovian representations resulting form moment fitting into Markovian representations.

The transform relation of the diagonal matrices that is used in $\simeq_0$ and $\approx_0$ in Definition 5 is a very strict condition as shown in the following theorem.

Theorem 2: If $\mathbf{D}_e^{(2)} = \mathbf{B}^{-1}\mathbf{D}_e^{(1)}\mathbf{B}$ for two diagonal matrices $\mathbf{D}_e^{(i)}$ of size m, then $\mathbf{B}(x, y) = 0$ or $\mathbf{D}_e^{(1)}(x, x) = \mathbf{D}_e^{(2)}(y, y)$.

Proof: Since $\mathbf{D}_e^{(i)}$ are diagonal matrices, we have

$$\mathbf{D}_e^{(1)}\mathbf{B} = \mathbf{B}\mathbf{D}_e^{(2)} \Rightarrow \mathbf{D}_e^{(1)}(x, x)\mathbf{B}(x, y) = \mathbf{B}(x, y)\mathbf{D}_e^{(2)}(y, y)$$

for all $0 \leq x, y < m$. ■

The relations defined in this section all relate representations of the same sizes. In the following two subsections we show how this approach can be extended to relations between representations of different sizes. To define a relation of components with different sizes some redundancy has to be in the larger representation. This redundancy can be related to the so called closing vector which is $\mathbb{1}$ in our case or the initial vector π. The approach we present is related to minimal representations in linear system theory [25] and has been adopted for non-Markov models in [21], [22].

B. Redundancy According to the Closing Vector

We now extend the equivalence relations of Definition 5 to define a relation between $\mathcal{A}^{(1)}$ with m states and $\mathcal{A}^{(2)}$ with n ($< m$) states. To distinguish the sizes of vectors $\mathbb{1}$ we use $\mathbb{1}_m$ and $\mathbb{1}_n$ for the vector of ones of length m and n, respectively. Furthermore, we have to distinguish between zero and arbitrary submatrices and use $\mathbf{0}$ for the former and $*$ for the latter. Again we assume that components are defined over the same alphabet $\mathcal{E}$ with the same vector Λ.

Definition 6: Two vectors, $\pi^{(1)}$ of size m and $\pi^{(2)}$ of size n ($n < m$), and two ordered sets of matrices, $\mathcal{M}^{(1)}$ composed by matrices of size $m \times m$ and $\mathcal{M}^{(2)}$ composed by matrices of size $n \times n$, are in similarity relation C_1 if a non-singular $m \times m$ matrix $\mathbf{B}$ exists such that

- $\mathbf{B}^{-1}\mathbb{1}_m = \begin{pmatrix} \mathbb{1}_n \\ \mathbf{0} \end{pmatrix}$,
- $\pi^{(1)}\mathbf{B} = \left(\pi^{(2)}, *\right)$, and
- $\mathbf{B}^{-1}\mathbf{F}^{(1)}\mathbf{B} = \begin{pmatrix} \mathbf{F}^{(2)} & * \\ \mathbf{0} & * \end{pmatrix}$, for all $\mathbf{F}^{(i)} \in \mathcal{M}^{(i)}$ ($i = 1, 2$) associated matrices of the two sets.

Definition 7: The equivalence relations $\sim_1$, $\simeq_1$ and $\approx_1$ are defined in the same way as $\sim_0$, $\simeq_0$ and $\approx_0$ in Definition 5 using similarity relation C_1.

We use the following partition of the matrices $\mathbf{B}$ and $\mathbf{B}^{-1}$,

$$\mathbf{B} = (\mathbf{V}, *) \text{ and } \mathbf{B}^{-1} = \begin{pmatrix} \mathbf{W} \\ * \end{pmatrix} .$$

$\mathbf{V}$ is a $m \times n$ and $\mathbf{W}$ a $n \times m$ matrix.

Theorem 3: The following relations hold for matrices $\mathbf{V}$ and $\mathbf{W}$ according to Definition 7:

- $\mathbf{W}\mathbb{1}_m = \mathbb{1}_n$, $\mathbf{V}\mathbb{1}_n = \mathbb{1}_m$ and $\mathbf{WV} = \mathbf{I}$,
- $\pi^{(1)}\mathbf{V} = \pi^{(2)}$,
- $\mathbf{W}\mathbf{F}^{(1)}\mathbf{V} = \mathbf{F}^{(2)}$ and $\mathbf{F}^{(1)}\mathbf{V} = \mathbf{V}\mathbf{F}^{(2)}$.

Proof: The proof follows by a simple substitution of the matrices. E.g.,

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} = \mathbf{B}^{-1}\mathbf{B} = \begin{pmatrix} \mathbf{W} \\ * \end{pmatrix} (\mathbf{V}, *) = \begin{pmatrix} \mathbf{WV} & * \\ * & * \end{pmatrix} .$$

■

Theorem 4: If $\mathcal{A}^{(1)} \sim_1 \mathcal{A}^{(2)}$, then $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$ are equivalent.

Proof: We have

$$\begin{aligned}
& f_{\mathcal{A}^{(1)}}((e_1, t_1, \ldots, e_k, t_k) && = \\
& \pi^{(1)} \left(\prod_{i=1}^{k} \sum_{j=0}^{\infty} \frac{(\mathbf{R}^{(1)} t_i)^j}{j!} \lambda_{e_i} \mathbf{E}_{e_i}^{(1)} \right) \mathbb{1}_m && = \\
& \pi^{(1)} \left(\prod_{i=1}^{k} \sum_{j=0}^{\infty} \frac{(\mathbf{R}^{(1)} t_i)^j}{j!} \lambda_{e_i} \mathbf{E}_{e_i}^{(1)} \right) \mathbf{V}\mathbb{1}_n && = \\
& \pi^{(1)}\mathbf{V} \left(\prod_{i=1}^{k} \sum_{j=0}^{\infty} \frac{(\mathbf{R}^{(2)} t_i)^j}{j!} \lambda_{e_i} \mathbf{E}_{e_i}^{(2)} \right) \mathbb{1}_n && = \\
& f_{\mathcal{A}^{(2)}}((e_1, t_1, \ldots, e_k, t_k) \ . &&
\end{aligned}$$

■

The relations define equivalent representations but do not describe a method to decide whether two representations are equivalent or, more important, describe a method to compute for a component $\mathcal{A}^{(1)}$ of size m an equivalent representation $\mathcal{A}^{(2)}$ of minimal size. However, such an approach is available in linear system theory (see e.g. [26]) and has been adopted to Markov and non-Markov models in [27]. For the lack of space we do not present the details of the algorithms here but use them to compute equivalent representations in Section V.

These equivalences are related to lumpability [10], [9] and bisimulation [7], [4] as shown in the following theorem that can be found in [22].

Theorem 5: If $\mathcal{A}^{(2)}$ of size n results from $\mathcal{A}^{(1)}$ of size m ($> n$) by aggregating states according to a stochastic bisimulation relation, then $\mathbf{V} \in \{0, 1\}^{m,n}$ and $\mathbf{V}$ contains one element equal to 1 per row and at least one element equal to 1 per column.

In this case $\mathbf{E}_e^{(1)}\mathbf{V} = \mathbf{V}\mathbf{E}_e^{(2)} \Rightarrow \mathbf{D}_e^{(1)}\mathbf{V} = \mathbf{V}\mathbf{D}_e^{(2)}$ which implies $\sim_1 \Leftrightarrow \simeq_1 \Leftrightarrow \approx_1$.

C. Redundancy According to the Initial Vector

The equivalence relations from the previous paragraph can also be defined in a symmetric form which will be done now.

Definition 8: Two vectors, $\pi^{(1)}$ of size m and $\pi^{(2)}$ of size n ($n < m$), and two ordered sets of matrices, $\mathcal{M}^{(1)}$ composed by matrices of size $m \times m$ and $\mathcal{M}^{(2)}$ composed by matrices of size $n \times n$, are in similarity relation C_2 if a non-singular $m \times m$ matrix $\mathbf{B}$ exists such that

- $\pi^{(1)}\mathbf{B} = \left(\pi^{(2)}, \mathbf{0}\right)$,
- $\mathbf{B}^{-1}\mathbb{1}_m = \begin{pmatrix} \mathbb{1}_n \\ * \end{pmatrix}$, and
- $\mathbf{B}^{-1}\mathbf{F}^{(1)}\mathbf{B} = \begin{pmatrix} \mathbf{F}^{(2)} & \mathbf{0} \\ * & * \end{pmatrix}$, for all $\mathbf{F}^{(i)} \in \mathcal{M}^{(i)}$ ($i = 1, 2$) associated matrices of the two sets.

Definition 9: The equivalence relations $\sim_2$, $\simeq_2$ and $\approx_2$ are defined in the same way as $\sim_0$, $\simeq_0$ and $\approx_0$ in Definition 5 using similarity relation C_2.

Similar to Theorems 3 and 4, the following two theorems can be proved.

Theorem 6: The following relations hold for matrices $\mathbf{V}$ and $\mathbf{W}$ according to Definition 9:

- $\pi^{(1)} = \pi^{(2)}\mathbf{W}$,
- $\mathbf{W}\mathbb{1}_m = \mathbb{1}_n$, $\mathbf{V}\mathbb{1}_n = \mathbb{1}_m$ and $\mathbf{WV} = \mathbf{I}$,
- $\mathbf{W}\mathbf{F}^{(1)}\mathbf{V} = \mathbf{F}^{(2)}$ and $\mathbf{W}\mathbf{F}^{(1)} = \mathbf{F}^{(2)}\mathbf{W}$.

Theorem 7: If $\mathcal{A}^{(1)} \sim_2 \mathcal{A}^{(2)}$, then $\mathcal{A}^{(1)}$ and $\mathcal{A}^{(2)}$ are equivalent.

The proof of the theorem follows the same pattern as the one of Theorem 4.

The equivalence relation defined in Definition 9 is an extension of weak lumpability [10] or exact performance equivalence [7], [28] as shown in [22]. Algorithms to compute minimal representations according to one of the

equivalence relations from Definition 9 can be derived from the approaches used to compute minimal relations according to the equivalence relations of Definition 7 [27].

IV. Compositionality and Equivalence

To be really useful, equivalence relations should be preserved by composition which means that $\mathcal{A}^{(1)} \bowtie \mathcal{A}^{(2)}$ implies $\mathcal{A}^{(1)} \| _{\mathcal{C}} \mathcal{A}^{(3)} \bowtie \mathcal{A}^{(2)} \| _{\mathcal{C}} \mathcal{A}^{(3)}$ for every $\mathcal{C} \subseteq \mathcal{E}_s$, component $\mathcal{A}^{(3)}$ and $\bowtie \in \{\sim_0, \simeq_0, \approx_0, \sim_1, \simeq_1, \approx_1, \sim_2, \simeq_2, \approx_2\}$. It is known that this relation holds for Markovian representations and bisimulation or inverse bisimulation [4], [7]. But it is important to note that the equivalence relations do not hold in general for Markovian representations and $\sim_j$ if $\mathbf{V}$ and $\mathbf{W}$ are not restricted as much as in bisimulation and inverse bisimulation, i.e., if $\mathbf{V} \notin \{0,1\}^{m,n}$. We now extend the results to non-Markovian representations and the more general equivalence relations.

Theorem 8: If $\mathcal{A}^{(1)} \simeq_j \mathcal{A}^{(2)}$, then $\mathcal{A}^{(1)} \| _{\mathcal{C}} \mathcal{A}^{(3)} \simeq_j \mathcal{A}^{(2)} \| _{\mathcal{C}} \mathcal{A}^{(3)}$ and $\mathcal{A}^{(3)} \| _{\mathcal{C}} \mathcal{A}^{(1)} \simeq_j \mathcal{A}^{(3)} \| _{\mathcal{C}} \mathcal{A}^{(2)}$ $(j = 0, 1, 2)$ if the set $\mathcal{C}$ used to compute $\simeq_j$ equals the set $\mathcal{C}$ used for synchronization.

Proof: We show the proof for $\simeq_1$, the other two proofs are similar. Let $\mathcal{A}^{(13)} = \mathcal{A}^{(1)} \| _{\mathcal{C}} \mathcal{A}^{(3)}$ and $\mathcal{A}^{(23)} = \mathcal{A}^{(2)} \| _{\mathcal{C}} \mathcal{A}^{(3)}$. Let $n^{(i)}$ be the size of $\mathcal{A}^{(i)}$ and let $\mathbf{V}^{(1,2)}$ be a $n^{(1)} \times n^{(2)}$ matrix such that $\mathbf{V}^{(1,2)} \mathbb{1}_{n^{(2)}} = \mathbb{1}_{n^{(1)}}$, $\pi^{(1)} \mathbf{V}^{(1,2)} = \pi^{(2)}$ and $\mathbf{F}^{(1)} \mathbf{V}^{(1,2)} = \mathbf{V}^{(1,2)} \mathbf{F}^{(2)}$ for all $\mathbf{F}^{(i)} \in \mathcal{M}^{(i)}_{\simeq}$ $(i = 1, 2)$. Matrix $\mathbf{V}^{(1,2)}$ exists since $\mathcal{A}^{(1)} \simeq_1 \mathcal{A}^{(2)}$ holds. Observe that $n^{(13)} = n^{(1)} n^{(3)}$ and $n^{(23)} = n^{(2)} n^{(3)}$. To prove $\mathcal{A}^{(13)} \simeq_1 \mathcal{A}^{(23)}$, it has to be shown that a $n^{(13)} \times n^{(23)}$ matrix $\mathbf{V}^{(13,23)}$ exists for which $\mathbf{V}^{(13,23)} \mathbb{1}_{n^{(23)}} = \mathbb{1}_{n^{(13)}}$, $\pi^{(13)} \mathbf{V}^{(13,23)} = \pi^{(23)}$ and $\mathbf{F}^{(13)} \mathbf{V}^{(13,23)} = \mathbf{V}^{(13,23)} \mathbf{F}^{(23)}$ for all $\mathbf{F}^{(i)} \in \mathcal{M}^{(i)}_{\simeq}$ $(i = 13, 23)$ holds. Define $\mathbf{V}^{(13,23)} = \mathbf{V}^{(1,2)} \otimes \mathbf{I}_{n^{(3)}}$. Then

$$\mathbf{V}^{(13,23)} \mathbb{1}_{n^{(13)}} = \left(\mathbf{V}^{(1,2)} \otimes \mathbf{I}_{n^{(3)}} \right) \left(\mathbb{1}_{n^{(1)}} \otimes \mathbb{1}_{n^{(3)}} \right) = \mathbb{1}_{n^{(23)}},$$

and

$$\pi^{(13)} \mathbf{V}^{(13,23)} = \left(\pi^{(1)} \otimes \pi^{(3)} \right) \left(\mathbf{V}^{(1,2)} \otimes \mathbf{I}_{n^{(3)}} \right) = \pi^{(23)}.$$

To prove the relation between the matrices we first prove that the required relation holds for matrices composed with Kronecker products and sums. For the Kronecker product we obtain

$$\begin{aligned} \mathbf{F}^{(13)} \mathbf{V}^{(13,23)} &= \left(\mathbf{F}^{(1)} \otimes \mathbf{F}^{(3)} \right) \left(\mathbf{V}^{(1,2)} \otimes \mathbf{I}_{n^{(3)}} \right) \\ &= \mathbf{F}^{(1)} \mathbf{V}^{(1,2)} \otimes \mathbf{F}^{(3)} \mathbf{I}_{n^{(3)}} \\ &= \mathbf{V}^{(1,2)} \mathbf{F}^{(2)} \otimes \mathbf{I}_{n^{(3)}} \mathbf{F}^{(3)} = \mathbf{V}^{(13,23)} \mathbf{F}^{(23)} \end{aligned}$$

and for the Kronecker sum

$$\begin{aligned} \mathbf{F}^{(13)} \mathbf{V}^{(13,23)} &= \left(\mathbf{F}^{(1)} \oplus \mathbf{F}^{(3)} \right) \left(\mathbf{V}^{(1,2)} \otimes \mathbf{I}_{n^{(3)}} \right) \\ &= \left(\mathbf{F}^{(1)} \otimes \mathbf{I}_{n^{(3)}} \right) \left(\mathbf{V}^{(1,2)} \otimes \mathbf{I}_{n^{(3)}} \right) \\ &\quad + \left(\mathbf{I}_{n^{(1)}} \otimes \mathbf{F}^{(3)} \right) \left(\mathbf{V}^{(1,2)} \otimes \mathbf{I}_{n^{(3)}} \right) \\ &= \left(\mathbf{F}^{(1)} \mathbf{V}^{(1,2)} \otimes \mathbf{I}_{n^{(3)}} \right) + \left(\mathbf{V}^{(1,2)} \otimes \mathbf{F}^{(3)} \right) \\ &= \left(\mathbf{V}^{(1,2)} \mathbf{F}^{(2)} \otimes \mathbf{I}_{n^{(3)}} \right) + \left(\mathbf{V}^{(1,2)} \otimes \mathbf{F}^{(3)} \right) \\ &= \left(\mathbf{V}^{(1,2)} \otimes \mathbf{I}_{n^{(3)}} \right) \left(\mathbf{F}^{(2)} \oplus \mathbf{F}^{(3)} \right) = \mathbf{V}^{(13,23)} \mathbf{F}^{(23)} \end{aligned}$$

These relations prove the theorem for all matrices of $\mathcal{M}_{\simeq}$, since according to (6) and (7) one of the two Kronecker operations relates the associated matrices.

To prove $\mathcal{A}^{(3)} \| _{\mathcal{C}} \mathcal{A}^{(1)} \simeq_1 \mathcal{A}^{(3)} \| _{\mathcal{C}} \mathcal{A}^{(2)}$ matrix $\mathbf{V}^{(31,32)} = \mathbf{I}_{n^{(3)}} \otimes \mathbf{V}^{(1,2)}$ is used. Apart form this the proof is analogous. ■

Corollary 1: If $\mathcal{A}^{(1)} \approx_j \mathcal{A}^{(2)}$, then $\mathcal{A}^{(1)} \| _{\mathcal{C}} \mathcal{A}^{(3)} \approx_j \mathcal{A}^{(2)} \| _{\mathcal{C}} \mathcal{A}^{(3)}$ $(j = 0, 1, 2)$ for all $\mathcal{C} \in \mathcal{E}_s$.

Proof: The proof is identical with the one of Theorem 8 assuming $\mathcal{C} \subseteq \mathcal{E}_s$. ■

Corollary 1 shows that $\approx_j$ is a congruence according to parallel composition but since $\approx_j$ requires conditions for all diagonal matrices, Theorem 2 applies and the possible relations are close to the known bisimulations which also hold for non-Markovian representations. Theorem 8 shows that the strict conditions on the matrices $\mathbf{D}^{(i)}_e$ are only required for events that are used in a synchronization. This implies that the equivalence relation has to be chosen according to the composition. The following theorem shows that $\sim_j$ is also a congruence, if the composition is adequately chosen.

Theorem 9: If $\mathcal{A}^{(1)} \sim_j \mathcal{A}^{(2)}$ and for all $e \in \mathcal{C}$, either $\mathbf{E}^{(1)}_e \mathbb{1}_{n^{(1)}} = \mathbb{1}_{n^{(1)}}$, $\mathbf{E}^{(2)}_e \mathbb{1}_{n^{(2)}} = \mathbb{1}_{n^{(2)}}$ or $\mathbf{E}^{(3)}_e \mathbb{1}_{n^{(3)}} = \mathbb{1}_{n^{(3)}}$, then $\mathcal{A}^{(1)} \| _{\mathcal{C}} \mathcal{A}^{(3)} \sim_j \mathcal{A}^{(2)} \| _{\mathcal{C}} \mathcal{A}^{(3)}$ and $\mathcal{A}^{(3)} \| _{\mathcal{C}} \mathcal{A}^{(1)} \sim_j \mathcal{A}^{(3)} \| _{\mathcal{C}} \mathcal{A}^{(2)}$ $(j = 0, 1, 2)$.

Proof: Note that in the extreme case when $\mathcal{A}^{(1)} \sim_j \mathcal{A}^{(2)}$ and for all $e \in \mathcal{C}$, $\mathbf{E}^{(1)}_e \mathbb{1}_{n^{(1)}} = \mathbb{1}_{n^{(1)}}$, $\mathbf{E}^{(2)}_e \mathbb{1}_{n^{(2)}} = \mathbb{1}_{n^{(2)}}$ then $\mathbf{D}^{(i)}_e = \mathbf{I}_{n^{(i)}}$ and $\mathcal{A}^{(1)} \simeq_j \mathcal{A}^{(2)}$ holds. It means that in this case Theorem 8 applies.

In the general case for the $\mathcal{J}^{(1,2)} \subseteq \mathcal{C}$ subset of synchronized events $\mathbf{E}^{(1)}_e \mathbb{1}_{n^{(1)}} = \mathbb{1}_{n^{(1)}}$, $\mathbf{E}^{(2)}_e \mathbb{1}_{n^{(2)}} = \mathbb{1}_{n^{(2)}}$ and for the $\mathcal{J}^{(3)} = \mathcal{C} \setminus \mathcal{J}^{(1,2)}$ subset of synchronized events $\mathbf{E}^{(3)}_e \mathbb{1}_{n^{(3)}} = \mathbb{1}_{n^{(3)}}$ hold. We prove the theorem for $\sim_1$, the other two proofs are similar. As in Theorem 8 we have to show that $\mathbf{F}^{(13)} \mathbf{V}^{(13,23)} = \mathbf{V}^{(13,23)} \mathbf{F}^{(23)}$ follows from $\mathbf{F}^{(1)} \mathbf{V}^{(1,2)} = \mathbf{V}^{(1,2)} \mathbf{F}^{(2)}$. Again we define $\mathbf{V}^{(13,23)} = \mathbf{V}^{(1,2)} \otimes \mathbf{I}_{n^{(3)}}$. Then it follows from the proof of Theorem 8 that $\mathbf{V}^{(13,23)}$ observes the conditions according to the closing vector $\mathbb{1}_{n^{(i3)}}$ and the initial vector. Furthermore, the relation holds for all matrices $\mathbf{E}^{(i3)}_e$ $(i = 1, 2,\ e \in \mathcal{E}_s)$. It remains to show that the similarity relation also holds for the matrices $\mathbf{R}^{(i3)} = \left(\mathbf{Q}^{(i3)}_\epsilon - \sum_{e \in \mathcal{E}_s} \lambda_e \mathbf{D}^{(i3)}_e \right)$. Using that

$\mathbf{E}_e^{(i)}\mathbb{I} = \mathbb{I}$ implies $\mathbf{D}_e^{(i)} = \mathbf{I}_{n^{(i)}}$, we have

$$\mathbf{R}^{(13)} =$$
$$\Bigg(\left(\mathbf{Q}_\epsilon^{(1)} \oplus \mathbf{Q}_\epsilon^{(3)}\right) - \sum_{e\in\mathcal{N}} \lambda_e \left(\mathbf{D}_e^{(1)} \oplus \mathbf{D}_e^{(3)}\right)$$
$$- \sum_{e\in\mathcal{J}^{(1,2)}} \lambda_e \left(\mathbf{I}_{n^{(1)}} \otimes \mathbf{D}_e^{(3)}\right) - \sum_{e\in\mathcal{J}^{(3)}} \lambda_e \left(\mathbf{D}_e^{(1)} \otimes \mathbf{I}_{n^{(3)}}\right) \Bigg) =$$
$$\Bigg(\left(\mathbf{Q}_\epsilon^{(1)} \otimes \mathbf{I}_{n^{(3)}}\right) + \left(\mathbf{I}_{n^{(1)}} \otimes \mathbf{Q}_\epsilon^{(3)}\right)$$
$$- \sum_{e\in\mathcal{N}} \lambda_e \left(\mathbf{D}_e^{(1)} \otimes \mathbf{I}_{n^{(3)}}\right) - \sum_{e\in\mathcal{N}} \lambda_e \left(\mathbf{I}_{n^{(1)}} \otimes \mathbf{D}_e^{(3)}\right)$$
$$- \sum_{e\in\mathcal{J}^{(1,2)}} \lambda_e \left(\mathbf{I}_{n^{(1)}} \otimes \mathbf{D}_e^{(3)}\right) - \sum_{e\in\mathcal{J}^{(3)}} \lambda_e \left(\mathbf{D}_e^{(1)} \otimes \mathbf{I}_{n^{(3)}}\right) \Bigg) =$$
$$\left(\left(\mathbf{Q}_\epsilon^{(1)} - \sum_{e\in\mathcal{J}^{(3)}\cup\mathcal{N}} \lambda_e \mathbf{D}_e^{(1)}\right) \oplus \left(\mathbf{Q}_\epsilon^{(3)} - \sum_{e\in\mathcal{J}^{(1,2)}\cup\mathcal{N}} \lambda_e \mathbf{D}_e^{(3)}\right)\right)$$

Thus

$$\mathbf{R}^{(13)}\mathbf{V}^{(13,23)} =$$
$$\left(\left(\mathbf{Q}_\epsilon^{(1)} - \sum_{e\in\mathcal{J}^{(3)}\cup\mathcal{N}} \lambda_e \mathbf{D}_e^{(1)}\right) \oplus \left(\mathbf{Q}_\epsilon^{(3)} - \sum_{e\in\mathcal{J}^{(1,2)}\cup\mathcal{N}} \lambda_e \mathbf{D}_e^{(3)}\right)\right)$$
$$\cdot \left(\mathbf{V}^{(1,2)} \otimes \mathbf{I}_{n^{(3)}}\right) =$$
$$\left(\left(\mathbf{Q}_\epsilon^{(1)} - \sum_{e\in\mathcal{J}^{(3)}\cup\mathcal{N}} \lambda_e \mathbf{D}_e^{(1)}\right) \mathbf{V}^{(1,2)} \right) \oplus \left(\mathbf{Q}_\epsilon^{(3)} - \sum_{e\in\mathcal{J}^{(1,2)}\cup\mathcal{N}} \lambda_e \mathbf{D}_e^{(3)}\right). \tag{9}$$

It remains to show that

$$\left(\mathbf{Q}_\epsilon^{(1)} - \sum_{e\in\mathcal{J}^{(3)}\cup\mathcal{N}} \lambda_e \mathbf{D}_e^{(1)}\right) \mathbf{V}^{(1,2)} = \mathbf{V}^{(1,2)} \left(\mathbf{Q}_\epsilon^{(2)} - \sum_{e\in\mathcal{J}^{(3)}\cup\mathcal{N}} \lambda_e \mathbf{D}_e^{(2)}\right),$$

because having that we obtain $\mathbf{R}^{(13)}\mathbf{V}^{(13,23)} = \mathbf{V}^{(13,23)}\mathbf{R}^{(23)}$ using the same steps as in (9) in reverse order. According to $\mathcal{A}^{(1)} \sim_j \mathcal{A}^{(2)}$ the similarity relation holds when we sum over all $e \in \mathcal{E}_s$. Using also that the similarity relation holds for unity matrices we have

$$\left(\mathbf{Q}_\epsilon^{(1)} - \sum_{e\in\mathcal{J}^{(3)}\cup\mathcal{N}} \lambda_e \mathbf{D}_e^{(1)}\right) \mathbf{V}^{(1,2)} =$$
$$\left(\left(\mathbf{Q}_\epsilon^{(1)} - \sum_{e\in\mathcal{E}_s} \lambda_e \mathbf{D}_e^{(1)}\right) + \left(\sum_{e\in\mathcal{J}^{(1,2)}\cup\mathcal{N}} \lambda_e \mathbf{I}_{n^{(1)}}\right)\right) \mathbf{V}^{(1,2)} =$$
$$\mathbf{V}^{(1,2)} \left(\left(\mathbf{Q}_\epsilon^{(2)} - \sum_{e\in\mathcal{E}_s} \lambda_e \mathbf{D}_e^{(2)}\right) + \left(\sum_{e\in\mathcal{J}^{(1,2)}\cup\mathcal{N}} \lambda_e \mathbf{I}_{n^{(2)}}\right)\right) =$$
$$\mathbf{V}^{(1,2)} \left(\mathbf{Q}_\epsilon^{(2)} - \sum_{e\in\mathcal{J}^{(3)}\cup\mathcal{N}} \lambda_e \mathbf{D}_e^{(2)}\right),$$

which completes the proof for $\mathcal{A}^{(1)} \| _{\mathcal{C}} \mathcal{A}^{(3)} \sim_j \mathcal{A}^{(2)} \| _{\mathcal{C}} \mathcal{A}^{(3)}$. Again the proof for $\mathcal{A}^{(3)} \| _{\mathcal{C}} \mathcal{A}^{(1)} \sim_1 \mathcal{A}^{(3)} \| _{\mathcal{C}} \mathcal{A}^{(2)}$ is similar using matrix $\mathbf{V}^{(3)} = \mathbf{I}_{n^{(3)}} \otimes \mathbf{V}^{(1,2)}$ ■

The theorem shows that the weakest relation is a congruence if the composition is *asynchronous* which means that for each event one component is active and the other component accepts the event and does not block it or affect the rate of the synchronized event.

V. Examples

We present numerical results in this section. We begin with some small matrices to clarify the equivalence relations. Afterwards two more complex examples indicate that the introduced equivalences go beyond bisimulation.

A. Small Matrices

The following example matrices are very small and are used to present in detail the equivalence relations. We begin with a component $\mathcal{A}^{(1)} = (\mathcal{S}^{(1)}, \pi^{(1)}, \mathbf{E}_e^{(1)}(e \in \mathcal{E}), \Lambda)$. The component has 6 states (i.e., $\mathcal{S}^{(1)} = \{0, \ldots, 5\}$), $\mathcal{E} = \{\epsilon, a, b\}$ and $\Lambda = (1, 1, 1)$. The following matrices characterize the behavior of the component.

$$\mathbf{E}_\epsilon^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 5 & 2 \\ 3 & 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1.25 & 0 & 0 \\ 3.75 & 1.25 & 0 & 0 & 0 & 0 \\ 1.5 & 0.5 & 4.8 & 1.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$
$$\mathbf{E}_a^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1.25 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix},$$
$$\mathbf{E}_b^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0.5 & 0 & 0 \\ 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 3.2 & 0.8 & 2 & 0 \end{pmatrix}.$$

$\pi^{(1)} = (0.111, 0.037, 0.430, 0.107, 0.176, 0.139)$ is the stationary vector of Component 1 whose generator is

$$\mathbf{Q}^{(1)} = \lambda_\epsilon(\mathbf{E}_\epsilon^{(1)} - \mathbf{D}_\epsilon^{(1)}) + \lambda_a(\mathbf{E}_a^{(1)} - \mathbf{D}_a^{(1)}) + \lambda_b(\mathbf{E}_b^{(1)} - \mathbf{D}_b^{(1)}).$$

Now consider the component $\mathcal{A}^{(2)} = (\mathcal{S}^{(2)}, \pi^{(2)}, \mathbf{E}_e^{(2)}(e \in \mathcal{E}), \Lambda)$ defined for same set $\mathcal{E}$ and the same vector Λ with the matrices

$$\mathbf{E}_\epsilon^{(2)} = \begin{pmatrix} 0 & 0 & 4 & 2 \\ 1 & 0 & 0 & 0 \\ 2 & 6 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$
$$\mathbf{E}_a^{(2)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}, \mathbf{E}_b^{(2)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 4 & 2 & 0 \end{pmatrix}.$$

$\mathcal{A}^{(1)} \sim_2 \mathcal{A}^{(2)}$ since $\mathbf{W}\left(\mathbf{E}_\epsilon^{(1)} - \sum_{e\in\mathcal{E}} \mathbf{D}_e^{(1)}\right) = \left(\mathbf{E}_\epsilon^{(2)} - \sum_{e\in\mathcal{E}} \mathbf{D}_e^{(2)}\right)\mathbf{W}$ and $\mathbf{W}\mathbf{E}_e^{(1)} = \mathbf{E}_e^{(2)}\mathbf{W}$ $(e = a, b)$ with matrix

$$\mathbf{W} = \begin{pmatrix} 0.75 & 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.8 & 0.2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

The non-Markovian component $\mathcal{A}^{(3)} = (\mathcal{S}^{(3)}, \pi^{(3)}, \mathbf{E}_e^{(3)}(e \in \mathcal{E}), \Lambda)$ defined for the same set $\mathcal{E}$ and the same vector Λ with the matrices

$$\mathbf{E}_\epsilon^{(3)} = \begin{pmatrix} 0 & 7 & 1 \\ 1 & 0 & 0 \\ -1 & -3 & 0 \end{pmatrix},$$

$$\mathbf{E}_a^{(3)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0.5 & 0.5 \\ 0 & 1.5 & 1.5 \end{pmatrix}, \mathbf{E}_b^{(3)} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 3.5 & -0.5 \\ 0 & 10.5 & -1.5 \end{pmatrix}$$

is also equivalent to the other two components since $\mathcal{A}^{(3)} \sim_1 \mathcal{A}^{(2)}$. We have $\left(\mathbf{E}_\epsilon^{(2)} - \sum_{e\in\mathcal{E}} \mathbf{D}_e^{(2)}\right)\mathbf{V} = \mathbf{V}\left(\mathbf{E}_\epsilon^{(3)} - \sum_{e\in\mathcal{E}} \mathbf{D}_e^{(3)}\right)$ and $\mathbf{E}_e^{(2)}\mathbf{V} = \mathbf{V}\mathbf{E}_e^{(3)}$ $(e = a, b)$ with matrix

$$\mathbf{V} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1.5 & -0.5 \\ 0 & 0.5 & 0.5 \end{pmatrix}.$$

Observe that only the weakest equivalence $\sim$ holds between the components. However, the simple example shows that equivalence between Markov components goes beyond bisimulation and equivalence between Markovian and non-Markovian representations exists.

We can try to transform this representation in a Markovian representation using relation $\sim_0$ and the approach presented in [18] to transform matrices to Markov representations by some elementary similarity transformations. For our example we obtain

$$\mathbf{E}_\epsilon^{(4)} = \begin{pmatrix} 0 & 0.00219253 & 4.46261 \\ 2.1795 & 0 & 1.44534 \\ 0.00697531 & 0.00120621 & 0 \end{pmatrix},$$

$$\mathbf{E}_a^{(4)} = \begin{pmatrix} 0.00573625 & -0.000338443 & 1.21588 \\ 5.37744 * 10^{-6} & -3.17273 * 10^{-7} & 0.00113983 \\ 0.00940847 & -0.000555107 & 1.99426 \end{pmatrix},$$

$$\mathbf{E}_b^{(4)} = \begin{pmatrix} 0.212434 & 2.3629 & 1.08851 \\ 0.000199146 & 0.0022151 & 0.00102042 \\ 0.34843 & 3.87557 & 1.78535 \end{pmatrix},$$

which is almost a Markovian representation of order 3.

B. A Disk System

The following example is a GSPN model of an IO system which has been taken from [29]. A specification of the model as a colored GSPN is shown in Figure 1. The system consists of d disks which can be accessed via c channels. A request arrives by firing the transition labeled with t_a and chooses with equal probability one of the disks (color $d = \{d_1, \ldots, d_k\}$). To access the disk, the request needs the disk (the token with color d on the upper place) and one channel (one of the c tokens on the place in the middle). We assume that every channel can be used to access every disk. Sending the request to the disk requires a negligible amount of time realized by an immediate transition which releases the channel immediately. Then the disk operation is performed which requires an exponentially distributed time with rate 1. Then a channel is required to transfer the data. We assume that the transfer operation requires an exponentially distributed time with rate 3. Finally, the disk and channel are returned and the request terminates by firing transition labeled with t_b. Thus, the component communicates with its environment by receiving requests (firing transition label t_a) and returning the finished request (transition with label t_b). The state space for the component is generated for up to n simultaneous requests. This implies that the component can be used in an environment with up to n request that can be pending at the component or the component has to block additional request. In the former case relation $\sim_i$ can be used to find a representation with a smaller state space, in the latter case relation $\approx_i$ is required.

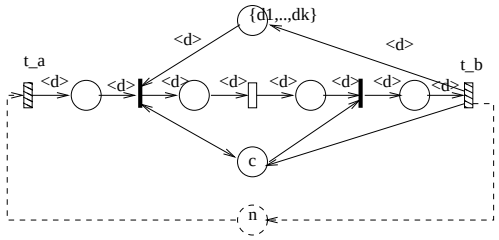

Figure 1. IO system example model.

The example is symmetric according to the usage of the disks which implies that the state space can be reduced using bisimulation and lumpability for state space reduction. We check whether the relations $\sim_i$ and $\approx_i$ go beyond bisimulation for this example.

Results for different configurations of the example are shown in Table I. The table includes in the first three columns the system parameters (n = maximum number of concurrent requests, k = number of disks, c = number of channels) and in the fourth column the sizes of the tangible state spaces of the model without any state space reduction. The fifth column contains the sizes of the reduced state spaces if ordinary lumpability or stochastic bisimulation

Parameters			State space size				
n	k	c	original	ordinary	exact	$\sim_i$	$\approx_i$
4	2	1	59	27	31	27	27
4	2	2	41	23	23	23	23
4	4	1	842	47	61	43	46
4	4	2	444	45	45	43	43
8	2	1	229	101	117	101	101
8	2	2	145	77	77	77	77
8	4	1	15143	541	836	508	524
8	4	2	7779	494	494	433	433
8	6	1	326115	853	1501	738	752
8	6	2	205239	968	971	890	898
8	8	4	444496	530	528	482	482

Table I
STATE SPACE SIZES OF EQUIVALENT REPRESENTATIONS OF THE IO SYSTEM.

Parameters		State space size				
TWS	*TBS*	original	ordinary	exact	$\sim_i$	$\approx_i$
1	1	16	16	14	14	14
2	1	69	69	61	57	57
3	1	217	217	199	194	196
4	1	546	546	504	498	500
1	2	28	28	24	21	23
2	2	120	119	102	96	98
3	2	384	378	339	329	334
4	2	993	972	882	860	863

Table II
STATE SPACE SIZES OF EQUIVALENT REPRESENTATIONS OF THE COURIER PROTOCOL.

is applied for state space reduction [7], [4]. Column six contains the state space sizes if state space reduction is performed according to exact lumpability or exact performance bisimulation [9], [28], column seven contains the reduced state space sizes according to the relation $\sim_1$ and $\sim_2$ which are applied one after the other and column eight contains the state space sizes for relation $\approx_1$ and $\approx_2$. Exploitation of symmetries results in a reduction which is based on an ordinary and exactly lumpable partition. Consequently, whenever the state space sizes according to ordinary and exact lumpability differ, the reduction goes beyond symmetry exploitation.

For several configurations of the example the state space sizes according to ordinary and exact lumpability differ which shows that state space reduction is more than symmetry exploitation. Furthermore, for some configurations it is also possible to reduce the state space further if relation $\sim_i$ and even if relation $\approx_i$ is applied. Although the use of $\sim_i$ sometimes allows an additional reduction, possibly resulting in a non-Markovian representation, the major reduction is , of course, due to exploitation of the inherent symmetry of the model.

It should be noted that the computation of lumpability only requires the comparison of sums of transition rates whereas the computation of reduced representations according to $\sim_i$ requires a repeated singular value decomposition of matrices which means that the ϵ-rank of a matrix is computed (see e.g., [30], and note that ϵ refers to numerical accuracy in this terminology). Even if singular value decomposition is stable, singular values and matrix elements do not completely vanish using floating point arithmetic even if they would be zero under exact arithmetic. Thus, model reduction using the approach from [27] requires, like model reduction in linear systems theory [25], the definition of some threshold ϵ to define which values are interpreted as 0. In all examples we present the absolute value of the largest element in the lower left submatrix (cf. Definition 7) or the upper right submatrix (cf. Definition 9) is below 10^{-10} whereas non-zero matrix elements in the matrices of the reduced component are in $O(1)$ which indicates that the aggregation is exact up to numerical inaccuracies.

C. A Communication Protocol

The second example has been taken from [31]. It describes a unidirectional communication protocol. We consider a compositional description that decomposes the model into four components and has been proposed in [32]. For aggregation the transport layer of the sender, which is described by one component, is used. The component can be parameterized by the communication window size (*TWS*) and the number of messages concurrently accepted from the upper layer protocol (*TBS*).

Table II contains the sizes of the different component state spaces. Although the component is not symmetric, exact and in some cases also ordinary lumpability allow one to reduce the state space. An additional reduction of the state space is possible using relation $\sim_i$ and $\approx_i$. However, as in the previous example this reduction reduces the state space only slightly more than exact lumpability which is for the example more effective than oridnary lumpability.

VI. CONCLUSION

We have presented a framework for compositional modeling. This framework is composed by similarity relations and compositional rules. The key feature which is investigated in the paper is the set of similarity relations and compositional rules which ensures congruence. It turns out that the similarity relations of stochastic bisimulation are strict enough to ensure congruence also with synchronized composition, but it is not the case when more general similarity relations are considered. These more general similarity relations might relate Markovian models with also non-Markovian ones, which are algebraic constructions without stochastic interpretation. Necessary conditions are presented to ensure congruence in case of the set of general similarity relations.

The investigation of efficient numerical analysis of possibly non-Markovian compositional models and the state space reduction due to the general set of similarity relations are future research plans. The paper presents numerical example which demonstrate that general similarity relations might result in smaller models than stochastic bisimulation.

REFERENCES

[1] P. Buchholz, *Die strukturierte Analyse Markovscher Modelle (in German)*, ser. IFB 282. Springer, 1991.

[2] S. Donatelli, "Superposed stochastic automata: a class of stochastic Petri nets amenable to parallel solution," *Performance Evaluation*, vol. 18, pp. 21–36, 1994.

[3] H. Hermanns, *Interactive Markov Chains: The Quest for Quantified Quality*, ser. Lecture Notes in Computer Science. Springer, 2002, vol. 2428.

[4] J. Hillston, "A compositional approach for performance modelling," University of Edinburgh, Dep. of Comp. Sc., PhD thesis, 1994.

[5] B. Plateau and J. M. Fourneau, "A methodology for solving Markov models of parallel systems," *J. Parallel Distrib. Comput.*, vol. 12, no. 4, pp. 370–387, 1991.

[6] B. Plateau, "On the stochastic structure of parallelism and synchronisation models for distributed algorithms," *ACM Performance Evaluation Review*, vol. 13, pp. 142–154, 1985.

[7] P. Buchholz, "Equivalence relations for stochastic automata networks," in *Computations with Markov Chains*, W. J. Stewart, Ed. Kluwer Academic Publishers, 1995, pp. 197–216.

[8] J. Hillston, "Compositional Markovian modelling using a process algebra," in *Computations with Markov Chains*, W. J. Stewart, Ed. Kluwer Academic Publishers, 1995, pp. 177–196.

[9] P. Buchholz, "Exact and ordinary lumpability in finite Markov chains," *Journal of Applied Probability*, vol. 31, pp. 59–75, 1994.

[10] J. G. Kemeny and J. L. Snell, *Finite Markov Chains*. Springer, 1976.

[11] P. Buchholz, "Efficient computation of equivalent and reduced representations for stochastic automata," *Int. Journ. Computer Systems Science and Engineering*, vol. 15, no. 2, pp. 93–103, 2000.

[12] S. Derisavi, H. Hermanns, and W. H. Sanders, "Optimal state-space lumping in markov chains," *Inf. Process. Lett.*, vol. 87, no. 6, pp. 309–315, 2003.

[13] A. Valmari and G. Franceschinis, "Simple $O(m \log n)$ time Markov chain lumping," in *TACAS*, ser. LNCS, J. Esparza and R. Majumdar, Eds., vol. 6015. Springer, 2010, pp. 38–52.

[14] P. Buchholz, G. Ciardo, S. Donatelli, and P. Kemper, "Complexity of Kronecker operations and sparse matrices with applications to the solution of Markov models," *INFORMS Journal on Computing*, vol. 12, no. 3, pp. 203–222, 2000.

[15] W. J. Stewart, *Introduction to the numerical solution of Markov chains*. Princeton University Press, 1994.

[16] H. Hermanns, M. Kwiatkowska, G. Norman, D. Parker, and M. Siegle, "On the use of MTBDDs for performability analysis and verification of stochastic systems," *Journal of Logic and Algebraic Programming*, vol. 56, pp. 23–67, 2003.

[17] G. Ciardo and A. S. Miner, "Storage alternatives for large structured state spaces," in *Proc. 9th Int. Conf. on Modelling Techniques and Tools for Computer Performance Evaluation*, R. Marie and B. Plateau, Eds. Springer LNCS 1245, 1997, pp. 44–57.

[18] M. Telek and G. Horváth, "A minimal representation of Markov arrival processes and a moments matching method," *Performance Evaluation*, vol. 64, no. 9-12, pp. 1153–1168, 2007.

[19] S. Asmussen and M. Bladt, "Point processes with finite-dimensional conditional probabilities," *Stochastic Processes and their Application*, vol. 82, pp. 127–142, 1999.

[20] N. G. Bean and B. F. Nielsen, "Quasi-birth-and-death processes with rational arrival process components," *Stochastic Models*, vol. 26, no. 3, pp. 309–334, 2010.

[21] P. Buchholz and M. Telek, "Stochastic Petri nets with matrix exponentially distributed firing times," *Performance Evaluation*, vol. 67, no. 12, pp. 1373–1385, 2010.

[22] ——, "Rational processes related to communicating Markov processes," *submitted for publication*, 2010.

[23] J. Hillston, "The nature of synchronization," in *Proc. 2nd Workshop on Process Algebras and Performance Modeling*, U. Herzog and M. Rettelbach, Eds., 1994, pp. 51–70.

[24] P. Buchholz and P. Kemper, "Kronecker based matrix representations for large Markov chains," in *Validation of Stochastic Systems*, ser. LNCS, B. Haverkort, H. Hermanns, and M. Siegle, Eds., vol. 2925, 2004, pp. 256–295.

[25] B. de Schutter, "Minimal state-space realization in linear system theory: An overview," *Journal of Computational and Applied Mathematics*, vol. 121, no. 1-2, pp. 331–354, 2000.

[26] P. M. V. Dooren, "The generalized eigenstructure on linear system theory," *IEEE Trans. on Automatic Control*, vol. 26, no. 1, pp. 111–129, 1981.

[27] P. Buchholz and M. Telek, "On minimal representations of rational arrival processes," Madrid Conference on Queueing Theory, MCQT, 2010.

[28] P. Buchholz, "Exact performance equivalence - an equivalence relation for stochastic automata," *Theoretical Computer Science*, vol. 215, no. 1/2, pp. 263–287, 1999.

[29] G. Balbo, C. Bruell, and S. Ghanta, "Combining queueing networks and generalized stochastic Petri nets for the solution of complex models of system behavior," *IEEE Trans. on Comp.*, vol. 37, no. 10, 1988.

[30] G. H. Golub and C. F. van Loan, *Matrix Computations*. John Hopkins University Press, 1996.

[31] C. M. Woodside and Y. Li, "Performance Petri net analysis of communications protocol software by delay equivalent aggregation," in *Proc. 4th Int. Workshop on Petri Nets and Performance Models*. IEEE CS Press, 1991, pp. 64–73.

[32] P. Buchholz, "Hierarchical structuring of superposed GSPNs," *IEEE Trans. Software Eng.*, vol. 25, no. 2, pp. 166–181, 1999.