Available at
www.ComputerScienceWeb.com
POWERED BY SCIENCE @ DIRECT®

The Journal of Logic and
Algebraic Programming 56 (2003) 69–97

THE JOURNAL OF
LOGIC AND
ALGEBRAIC
PROGRAMMING

www.elsevier.com/locate/jlap

Model-checking large structured Markov chains

Peter Buchholz [a,*], Joost-Pieter Katoen [b], Peter Kemper [c], Carsten Tepper [c]

[a] *Fakultät Informatik, TU Dresden, D-01062 Dresden, Germany*
[b] *Formal Methods and Tools Group, Faculty of Computer Science, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands*
[c] *Informatik IV, Universität Dortmund, D-44221 Dortmund, Germany*

Abstract

This paper presents algorithms and experimental results for model-checking continuous-time Markov chains (CTMCs) based on a structured analysis approach. In this approach, a CTMC is represented as a term in Kronecker algebra that reflects the component structure of the system model. Such representations can be obtained in a natural way from various high-level specification formalisms, such as stochastic extensions of Petri nets, process algebras or activity networks. Properties are expressed in continuous stochastic logic (CSL) which includes means to express transient, steady-state and path performance measures. This paper describes novel model-checking algorithms for CSL that fully exploit the compositional description of the CTMC. This yields an effective way to combat the state-space explosion problem and enables the model-checking of fairly large Markov chains. Furthermore, we show how state-space aggregation (modulo bisimulation) and the elimination of vanishing states can be done in a component-wise manner. To demonstrate the applicability of the approach, and to assess the efficiency of our algorithms, we analyze a stochastic Petri net-model of a workstation cluster system and a simple queuing network.
© 2002 Elsevier Science Inc. All rights reserved.

Keywords: Markov chains; Model checking; Kronecker algebra; Stochastic bisimulation; Structured analysis

1. Introduction

Continuous-time Markov chains (CTMCs) are the basic mathematical model underlying many formalisms for the description and analysis of stochastic discrete-event dynamic systems. Examples of significant formalisms which are mapped onto CTMCs are stochastic Petri nets (SPNs) [18], stochastic process algebras (SPAs) [9,31,35], queueing networks (QNs) [14,43] and stochastic automata networks (SANs) [42,43]. All these formalisms

* Corresponding author. Tel.: +49-351-463-8360; fax: +49-351-463-8460.
E-mail addresses: p.buchholz@inf.tu-dresden.de (P. Buchholz), katoen@cs.utwente.nl (J.-P. Katoen), kemper@ls4.cs.uni-dortmund.de (P. Kemper), tepper@ls4.cs.uni-dortmund.de (C. Tepper).

allow a convenient and structured model description which is mapped onto a CTMC and subsequently analyzed using established numerical techniques for the stationary or transient analysis [43]. Measures of interest are usually specified at the model description level and are then transformed into rewards at the CTMC level. This yields a stochastic process, typically a Markov reward process (MRP) [30], which can be analyzed using numerical methods similarly to a CTMC.

However, although MRPs allow an integration of process and measures of interest, the pure specification of rewards is not always the right way to describe the property one wants to evaluate for a system. Often the requirements for a system state that the system has to observe at least a specific level of performance and a model is used to check that the system fulfills this task. For (untimed) concurrent systems, a popular approach for such a verification is to employ automated verification techniques such as *model checking* [23]. Model checking amounts to a systematic check of the validity of a requirement, formulated as a formula in an appropriate temporal logic, in all states of the model. Model checking has been widely used in practice [23] and has recently been extended in various directions including an extension for model-checking CTMCs. For this purpose, the temporal logic CSL was developed [1,2] as a stochastic extension of the branching-time temporal logic computation tree logic (CTL). In several follow-up papers this logic has been extended [6], applications have been described [4,29] and sophisticated model-checking algorithms have been introduced [36].

Like traditional CTMC analysis, the major problem of model-checking realistic examples is the *state-space explosion*. A possible approach to attack this state-space explosion problem is the use of symbolic techniques which extend approaches from standard model checking [23] and represent transition matrices using directed graph structures like multi-terminal binary decision diagrams (MTBDDs) [3,22]. For CSL model checking, this approach has recently been pursued in [36]. However, although MTBDDs may reduce the memory requirements, their operations become very slow and, typically, the model structure is not reflected by the symbolic representation. Besides, reduction using equivalences such as bisimulation does not always lead to a smaller BDD-representation of the model at hand.

Another approach to alleviate state-space explosion for the numerical analysis of CTMCs and also for several functional analysis steps are so-called *structured analysis* approaches. The idea of these approaches is to represent the huge generator or transition matrix as a sum of Kronecker products of small component matrices. The approach has been first developed for networks of stochastic automata [42,43] and has been subsequently extended in various directions [8,14,19,24,38]. Structured analysis approaches allow the efficient analysis of large Markov chains and, due to the compositional description of the transition matrix, several other steps exploiting the model structure like aggregation due to equivalence or local elimination of vanishing states can be performed efficiently. As shown in [16,40] also CTL (or LTL) model checking can be realized very efficiently in the context of structured description of (ordinary) transition systems.

In this paper, we present a structured model-checking approach for CSL, thus combining and extending structured analysis and model checking of CTL-formulas. The presented approach has been implemented and integrated in the abstract Petri net notation (APNN) toolbox, a modular tool environment for the qualitative and quantitative analysis of discrete-event systems [15]. To summarize, this paper:

- presents model-checking algorithms for CSL that exploit structured analysis techniques based on a Kronecker representation of the CTMC,
- shows how bisimulation equivalence (i.e., lumping) can be used to minimize system descriptions component-wise while preserving the validity of CSL-formulas,

- allows the consideration of vanishing states in CSL model checking, and
- provides empirical results substantiating the claim that structured analysis techniques are powerful means to combat the state-space explosion problem in model-checking CSL.

The techniques presented in this paper provide ample means to exploit the compositional nature of the model description—being it a superposed GSPN, a SPA specification, or the like—for both the model-checking procedures as well as for model reduction prior to verification.

1.1. Organization of the paper

The following two sections introduce the basic model class, the underlying structured description of the CTMC, some notations and basic algorithms for stationary and transient CTMC analysis. Section 4 reviews the logic CSL and presents the basic steps of model-checking CSL. Section 5 combines both approaches and presents the model-checking approach which uses the Kronecker representation of the generator matrix of the CTMC. Section 6 shows under which conditions vanishing states can be eliminated locally in the components. Section 7 introduces how bisimulation equivalence, which preserves the validity of CSL-formulas, can be used to reduce components and how the entire components can be substituted by reduced components in the structured description. Section 8 reports on the practical experiments that we conducted. Section 9 concludes the paper. The appendix summarizes the Kronecker operations.

2. Basic definitions and notations

Structured descriptions have been proposed in different forms and notations vary significantly. This paper uses the model class and most of the notation of [14,19]. For the representation of the state space we exploit the extensions given in [16]. The structured model-checking approach we present here can be easily extended to other Kronecker representations of the generator matrix like the hierarchical approach [8,12,17] or matrix diagrams [19,20]. However, to avoid an overloading of the notation, we restrict ourselves to the basic model class of [14].

2.1. Notational conventions

Throughout the paper we use italic letters for scalars, calligraphic letters for sets, lower case boldface Roman or Greek letters for vectors and upper case boldface Roman letters for matrices. Vectors are usually row vectors, unless stated otherwise. $\mathbb{R}$, $\mathbb{R}_{\geqslant 0}$ and $\mathbb{R}_{>0}$ are the sets of real numbers, non-negative real numbers and strictly positive real numbers. Matrix and vector elements are indicated using brackets, e.g., $\mathbf{A}(i, j)$ and $\mathbf{a}(i)$ indicate the i, jth of matrix $\mathbf{A}$ and the ith element of vector $\mathbf{a}$, respectively. $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{a}^{\mathrm{T}}$ denote the transposed matrix and vector, respectively. $\mathbf{I}_n$ indicates the identity matrix of order n. The value n is omitted if it is clear from the context. $\mathbf{1}$ and $\mathbf{0}$ are row vectors with 1 or 0 in all positions. Their length follows from the context. Numbering of elements in vectors and matrices starts with 0 and ends with $n-1$ for an n-dimensional vector or matrix. Square brackets are used to denote sub-matrices or sub-vectors, e.g., if $\mathbf{A} \in \mathbb{R}^{n\times n}$ and $\mathscr{A}, \mathscr{B} \subseteq \{0, \ldots, n-1\}$, then $\mathbf{A}[\mathscr{A}, \mathscr{B}]$ is a matrix including all rows with indices belonging to $\mathscr{A}$ and all columns with indices belonging to $\mathscr{B}$. $\mathbf{A}[\mathscr{A}, \mathscr{B}](i, j)$ with $0 \leqslant i < |\mathscr{A}|$ and $0 \leqslant j < |\mathscr{B}|$ denotes element i, j in the sub-matrix. In the same way sub-vectors are defined, e.g., $\mathbf{a}[\mathscr{A}]$ is the sub-vector of $\mathbf{a}$ including all elements belonging to $\mathscr{A}$.

2.2. Continuous-time Markov chains

As the basic mathematical model we consider CTMCs with state space $\mathscr{T}$ and generator matrix $\mathbf{Q}$. We are interested in stationary or transient results. Both types of results can be derived from the stationary or transient distribution vector of the CTMC. The stationary solution vector $\boldsymbol{\pi}$ is the solution of

$$\boldsymbol{\pi} \cdot \mathbf{Q} = \mathbf{0} \quad \text{subject to } \boldsymbol{\pi} \cdot \mathbf{1}^{\mathrm{T}} = 1. \tag{1}$$

The stationary solution is unique if the CTMC is ergodic. The transient solution $\boldsymbol{\pi}_t$ at time t is the solution of the differential equation $\dot{\boldsymbol{\pi}}_t = \boldsymbol{\pi}_t \cdot \mathbf{Q}$ and is expressed as:

$$\boldsymbol{\pi}_t = \boldsymbol{\pi}_0 \cdot \exp(\mathbf{Q} \cdot t), \tag{2}$$

where $\boldsymbol{\pi}_0$ is the initial distribution at time 0. For stationary and transient analysis of CTMCs a large number of different solution techniques exist, for an excellent overview see [43].

Results related to CTMCs are often expressed in terms of rewards which can be assigned to states (*rate-based rewards*) or transitions (*impulse-based rewards*). Rewards are usually defined with respect to some measures-of-interest of the describing high-level model, e.g., populations in a queue of a QN. We consider rate-based rewards. (For the model-checking algorithm these rewards are between 0 and 1.) The column vector $\boldsymbol{\rho} \in \mathbb{R}_{\geqslant 0}^{|\mathscr{T}|}$ assigns non-negative rewards to the states of a CTMC with state space $\mathscr{T}$. The mean stationary reward is given by $\boldsymbol{\pi} \cdot \boldsymbol{\rho}$, the mean transient reward at time t equals $\boldsymbol{\pi}_t \cdot \boldsymbol{\rho}$, and the accumulated reward in the interval $[t_1, t_2)$ is computed as $\int_{t_1}^{t_2} \boldsymbol{\pi}_t \cdot \boldsymbol{\rho}\, \mathrm{d}t$.

2.3. Component-based descriptions

We consider Markov models consisting of components that interact via synchronized transitions. Such descriptions are common for SPAs [9,31,35], superposed SPNs [24,38], SANs [42,43] or QNs [14,43].

The model consists of K components numbered 1 through K that communicate via synchronization events from a set $\mathscr{E}_S$. Additionally, a set of local events $\mathscr{E}_L$ describes transitions which do not belong to a communication. The sets $\mathscr{E}_{S,k}$ and $\mathscr{E}_{L,k}$ contain the synchronized and local events in which component k participates ($\mathscr{E}_{S,k} \cap \mathscr{E}_{L,k} = \emptyset$). Each $e \in \mathscr{E}_L$ belongs to exactly one set $\mathscr{E}_{L,k}$ whereas each $e \in \mathscr{E}_S$ belongs to at least two sets $\mathscr{E}_{S,k}$ as it describes a communication between components. $\text{rate}(e)$ is the rate of transition $e \in \mathscr{E} = \mathscr{E}_S \cup \mathscr{E}_L$.

Let $\widehat{\mathscr{T}}^k$ be the state space of component k with $n_k = |\widehat{\mathscr{T}}^k|$. States in $\widehat{\mathscr{T}}^k$ are numbered 0 through $n_k - 1$. For the moment we assume that $\widehat{\mathscr{T}}^k$ contains only tangible states and consider the problem of vanishing in Section 6. $\mathbf{W}_{k,e} \in \mathbb{R}_{\geqslant 0}^{n_k \times n_k}$ is a matrix describing the transitions according to event e in component k. Define

$$\mathbf{R}_k = \sum_{e \in \mathscr{E}_{L,k}} \text{rate}(e) \cdot \mathbf{W}_{k,e}$$

the matrix of local transition rates for component k. For notational convenience let $\mathbf{W}_{k,e} = \mathbf{I}_{n_k}$ for $e \notin \mathscr{E}_{L,k} \cup \mathscr{E}_{S,k}$. Boolean matrices $\mathbf{W}_{k,e}$ are sufficient to indicate whether an event can take place or not; numerical entries in $\mathbf{W}_{k,e}$ other than 0 and 1 can be used to encode state dependent scaling factors that speed up or slow down an event. This feature is also used to account for the effect of probabilistic output bags, where an event that takes place in a state can make transitions to several, different states and the selection of a transition takes places according to a probability distribution. These probabilities give the numerical entries of $\mathbf{W}_{k,e}$. Values of $\mathbf{W}_{k,e}$ can also be used to carry path probabilities for sequences of state transitions passing through vanishing states. Vanishing states are states with a zero sojourn time. Since those states are somewhat irrelevant for the timing behavior, it is common practice in performance evaluation to eliminate vanishing states and retain only the probabilities through sequences (paths) of vanishing states. $\mathbf{W}_{k,e}$ contains probabilities of such sequences that immediately follow an initial event e.

Let $n_1^0 = 1$ and for $l \leqslant m$ let $n_l^m = \prod_{k=l}^m n_k$. The *potential state space* of the composed model with components $1, \ldots, K$ equals $\widehat{\mathscr{T}} = \widehat{\mathscr{T}}^1 \times \cdots \times \widehat{\mathscr{T}}^K$ and contains n_1^K states. Each state s from $\widehat{\mathscr{T}}$ is described by a K-dimensional vector $(x^1, \ldots, x^K)$ where $0 \leqslant x^k < n_k$; this can be linearized to $s = \sum_{k=1}^K x^k \cdot n_{k+1}^K$ $(0 < s < n_1^K)$. We use both representations interchangeably in this paper.

Based on this representation, the rate-matrix of the composed model is given by:

$$\widehat{\mathbf{R}} = \sum_{e \in \mathscr{E}} \text{rate}(e) \cdot \bigotimes_{k=1}^{K} \mathbf{W}_{k,e} = \bigoplus_{k=1}^{K} \mathbf{R}_k + \sum_{e \in \mathscr{E}_S} \text{rate}(e) \cdot \bigotimes_{k=1}^{K} \mathbf{W}_{k,e}, \tag{3}$$

where $\otimes$ and $\oplus$ are the Kronecker sum and product, respectively (for details about these operations see [14,42] and the appendix). Thus, $\widehat{\mathbf{R}}$ is defined as the sum of (a Kronecker sum of) transition matrices $\mathbf{R}_k$ that refer to all transitions local to component k, and (a Kronecker product of) the transitions resulting from synchronizations between components.

Define $\widehat{\mathbf{Q}} = \widehat{\mathbf{R}} - \text{diag}(\widehat{\mathbf{R}} \cdot \mathbf{1}^{\mathrm{T}})$ where diag($\mathbf{a}$) is a square matrix having the elements of vector $\mathbf{a}$ on the diagonal. The set of *reachable states* equals $\mathscr{T} \subseteq \widehat{\mathscr{T}}$ and we let $n = |\mathscr{T}|$ denote its cardinality. Let $\mathbf{Q} = \widehat{\mathbf{Q}}[\mathscr{T}, \mathscr{T}]$. As often $\mathscr{T} \subset \widehat{\mathscr{T}}$, we have $\mathbf{Q} \neq \widehat{\mathbf{Q}}$ [14]. The set $\mathscr{T}$ can be efficiently computed from the representation (3), since for $x \in \mathscr{T}$, we have that $\widehat{\mathbf{R}}(x, y) > 0$ implies $y \in \mathscr{T}$. For details about the corresponding algorithms we refer to [16]. We assume in the sequel that for each $y^k \in \widehat{\mathscr{T}}^k$ there exists a state $(x^1, \ldots, x^K) \in \mathscr{T}$ such that $x^k = y^k$; otherwise, we may delete y^k from $\widehat{\mathscr{T}}^k$. Observe that Eq. (3) implicitly also includes a description of the reachability graph of the system, where rates are neglected. If only the reachability graph (and not the generator matrix) is of interest, then it suffices to consider Boolean (instead of real) matrices. The structured representation, however, remains the same.

2.4. Representing sets of states

Different analysis steps work on different subsets of the state space, e.g., numerical analysis methods require knowledge of $\mathscr{T}$ (rather than $\widehat{\mathscr{T}}$) while model-checking algorithms typically consider states which fulfill a sub-formula of the formula to be verified. For $\mathscr{S} \subseteq \widehat{\mathscr{T}}$, let $\Xi_{\mathscr{S}}$ be a function that maps any state $s \in \mathscr{S}$ onto a natural number that is smaller than $|\mathscr{S}|$ such that $\Xi_{\mathscr{S}}$ is monotonic with respect to $<$, i.e., $s < s' \Rightarrow \Xi_{\mathscr{S}}(s) < \Xi_{\mathscr{S}}(s')$ for $s, s' \in \mathscr{S}$. This function is used to map the K-dimensional representation of reachable states with a lexicographical order to an index set $\{0, \ldots, |\mathscr{S}|-1\}$ while preserving this order.

We now describe a data structure that represents $\mathscr{S}$ in a compact way while allowing fast access to individual states. First, notice that $\widehat{\mathscr{T}}$ can be represented by a tree with K levels, where each node at level k has n_k sons. A path in the tree corresponds to a state $(x^1, \ldots, x^K)$. Elimination of all paths to states from $\widehat{\mathscr{T}} \setminus \mathscr{S}$, yields a representation of $\mathscr{S}$. The remaining nodes at level k describe states from $\widehat{\mathscr{T}}^k$ belonging to $\mathscr{S}$. Let $RS_{\mathscr{S}}(x^1, \ldots, x^k) = \{\mathbf{y} \in \mathscr{S} \mid x^1 = y^1 \wedge \cdots \wedge x^k = y^k\}$, the subset of states from $\mathscr{S}$ with fixed states for the components 1 through k. $RS_{\mathscr{S}}(x^1, \ldots, x^k)$ refers to a sub-tree with root at level $k+1$; this sub-tree is denoted $RS_{\mathscr{S}}^{k+1}(x^1, \ldots, x^k)$ (which is the subset of states from $\widehat{\mathscr{T}}^{k+1}$ which belong to states from $\mathscr{S}$ given that the state of components 1 through k equals $(x^1, \ldots, x^k)$). Finally, $RS_{\mathscr{S}}(\ ) = \mathscr{S}$ and $RS_{\mathscr{S}}^1(\ )$ is the root of the tree.

Let us now discuss how we can minimize this tree representation. Sub-trees $RS_{\mathscr{S}} \times (x^1, \ldots, x^k)$ and $RS_{\mathscr{S}}(y^1, \ldots, y^k)$ are equal, if and only if

$$RS_{\mathscr{S}}^k(x^1, \ldots, x^{k-1}) = RS_{\mathscr{S}}^k(y^1, \ldots, y^{k-1})$$

and all pairs of sub-trees $RS_{\mathscr{S}}(x^1, \ldots, x^k, x^{k+1}), RS_{\mathscr{S}}(y^1, \ldots, y^k, y^{k+1})$ with $x^{k+1} = y^{k+1}$ are equal. Using a folding operation in a bottom-up manner, like in ordered BDDs [7], equal sub-trees are represented once. For a given ordering of components, this yields a unique acyclic graph (DAG) that represents $\mathscr{S}$. By introducing appropriate arc inscriptions it is possible to derive for each state $s \in \mathscr{S}$ the number $\Xi_{\mathscr{S}}(s)$. To realize this mapping, one basically has to count the number of leaves to the left of the path of a state x in the tree. The cardinality of leaves in isomorphic sub-trees is equal, and thus remains invariant under folding. By assigning corresponding weights on arcs of the DAG, $\Xi_{\mathscr{S}}$ is evaluated for a path $x^1, \ldots, x^K$ in the DAG by summing arc weights at each node which leave from the left positions of x^k in $(x^1, \ldots, x^{k-1})$ plus the position of x^K in $RS_{\mathscr{S}}(x^1, \ldots, x^{K-1})$. An implementation typically precomputes such weights to avoid the local summation at each node.

Example 2.1. Consider the transition graphs of components M_1, M_2 and M_3 depicted in Fig. 1. Rates are omitted in this example, as we are only interested in a representation of the reachable state space. Labels l_i ($i = 1, 2, 3$) refer to local transitions in component i, all remaining labels describe synchronizations. Labels a, b indicate synchronizations between M_1 and M_2, whereas c, d indicate synchronizations between M_2 and M_3. Initial states are indicated by double circles. $\widehat{\mathscr{T}}$ consists of 18 states of which 12 are reachable. The DAG on the right-hand side of Fig. 1 represents $\mathscr{T}$. Values for $\Xi_{\mathscr{T}}$ are derived by tra-

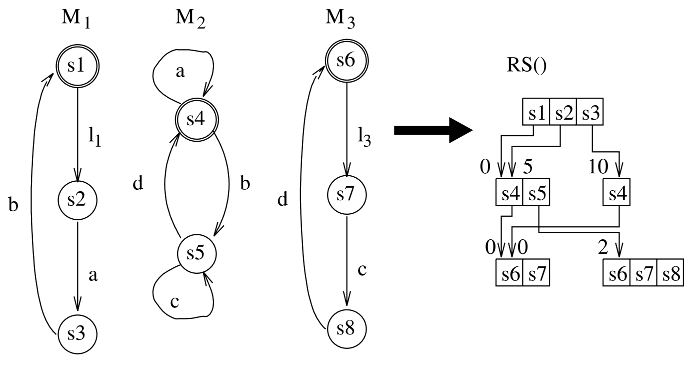

Fig. 1. Example with three components and DAG representation of $\mathscr{T}$.

versing the path through the graph from top to bottom, adding all arc weights and adding at the end the lexicographical position of the element in the leaf node (starting with 0 for the first element), e.g., $\Xi_{\mathscr{T}}(s2, s5, s7) = 5 + 2 + 1 = 8$ and $\Xi_{\mathscr{T}}(s3, s4, s6) = 10 + 0 + 0 = 10$.

3. Structured analysis of Markov chains

In the previous section, compact representations for $\widehat{\mathbf{Q}}$ and $\mathscr{T}$ have been presented. These representations are only useful if they can be exploited in analysis algorithms. Fortunately, this is the case for the Kronecker representation of the generator matrix and the DAG representation of $\mathscr{T}$. In this section, we give a brief overview of structured analysis techniques and refer to the literature for further details. Transient analysis is considered in some more detail, because it is the basic technique to verify time-bounded until operators in CSL (see Section 4).

3.1. Stationary analysis

For the determination of the stationary vector $\boldsymbol{\pi}$ according to (1), iterative numerical solution techniques are applied which compute consecutive approximations $\mathbf{p}^{(i)}$ until the iteration vector is sufficiently near to $\boldsymbol{\pi}$. A large number of iterative methods exist for the stationary analysis of CTMCs [43]. Many of them are applicable in conjunction with the Kronecker representation of the generator matrix. Distinguishing aspects of the iterative methods are whether they perform iterations on $\widehat{\mathscr{T}}$ (states from $\widehat{\mathscr{T}} \setminus \mathscr{T}$ receive zero probability during iteration) or on (preferably) $\mathscr{T}$ and whether iteration is performed row-wise or column-wise. Different algorithms to multiply a vector with a Kronecker representation of $\mathbf{Q}$ are introduced and tested in [14]. That paper also presents basic numerical solution algorithms based on these multiplication algorithms. Advanced structured solution techniques are presented in [13].

3.2. Transient analysis

For transient analysis, i.e., the computation of $\boldsymbol{\pi}_t$ according to (2), *randomization* [27,43] is the most efficient analysis technique for almost all CTMCs. Randomization in conjunction with the Kronecker representation of the generator matrix has been considered in [39]. Since randomization is also applied for model-checking CSL (as discussed in [5]), we present here the main steps. Randomization requires computing Poisson probabilities for a rate λ and a time horizon t according to $\beta_i(\lambda, t) = \mathrm{e}^{-\lambda \cdot t}((\lambda \cdot t)^i / i!)$, i.e., $\beta_i(\lambda, t)$ is the probability that a Poisson process with rate λ generates i events in an interval of length t. A straightforward computation of $\beta_i(\lambda, t)$ is numerically stable only for relatively small values of $\lambda \cdot t$, for arbitrary values of $\lambda \cdot t$, the approach of Fox and Glynn [26] is frequently employed, which gives left and right truncation points $0 \leqslant l < r$, such that $\beta_i(\lambda, t) = 0$ for $i < l$ and $i > r$.

Let $\mathbf{p}^{(0)} = \boldsymbol{\pi}_0$ be the initial vector. Then $\boldsymbol{\pi}_t$ is given by:

$$\boldsymbol{\pi}_t = \sum_{i=0}^{\infty} \mathbf{p}^{(i)} \cdot \beta_i(\lambda, t), \tag{4}$$

where $\mathbf{p}^{(i+1)} = \mathbf{p}^{(i)} + \lambda^{-1} \cdot \mathbf{p}^{(i)} \cdot \mathbf{Q}$ ($i \geqslant 0$) and $\lambda \geqslant \max_{x \in \{0,\ldots,|\mathscr{T}|\}}(|\mathbf{Q}(x,x)|)$. The above sum can be truncated by defining i_ε such that $1 - \sum_{i=0}^{i_\varepsilon} \beta_i(\lambda, t) \leqslant \varepsilon$. In practice i_ε is often significantly smaller than r, the right truncation point. Truncation of the sum in (4) yields a result vector where the error can be bounded by ε.

Often measures-of-interest are computed for a reward vector $\boldsymbol{\rho}$. In this case, (4) becomes

$$\mathbf{p}^{(0)} \cdot \left(\sum_{i=0}^{i_\varepsilon} \beta_i(\lambda, t) \cdot \left(\mathbf{I} + \lambda^{-1} \cdot \mathbf{Q} \right)^i \right) \cdot \boldsymbol{\rho} \tag{5}$$

to compute the expected reward at time t. If $\mathbf{Q}$ is defined as sub-matrix of $\widehat{\mathbf{Q}}$—which has a Kronecker representation—then $\mathbf{I} + \lambda^{-1}\mathbf{Q}$ also has a Kronecker representation, i.e.,

$$\mathbf{I} + \lambda^{-1} \cdot \mathbf{Q} = \underbrace{(\mathbf{I} - \lambda^{-1} \cdot \mathrm{diag}(\widehat{\mathbf{R}} \cdot \mathbf{1}^{\mathrm{T}})[\mathscr{T}, \mathscr{T}])}_{=\mathrm{diag}(\mathbf{d})} + \lambda^{-1} \cdot \widehat{\mathbf{R}}[\mathscr{T}, \mathscr{T}]. \tag{6}$$

Here $\mathbf{d}$ is a vector of length $\mathscr{T}$ that contains diagonal entries. Recall that $\widehat{\mathbf{R}}$ is represented in a structured way, cf. Eq. (3). The scalar value λ^{-1} can be multiplied with matrices of Kronecker sums and one matrix of each Kronecker product such that a similar structure as for $\mathbf{Q}$ is obtained. This facilitates the use of the multiplication algorithms for vectors and sub matrices of $\widehat{\mathbf{R}}$ as presented in [14].

In principle, the iteration (5) can be performed from left-to-right or from right-to-left. In both cases, vector–matrix products have to be computed, either by row or by column. A multiplication from right-to-left is preferable if results have to be computed for different initial vectors $\mathbf{p}^{(0)}$, because for each vector only a vector product has to be computed (provided the product of the reward vector with the matrix has been precomputed). This will appear during CSL model-checking, cf. Section 5. Multiplication from left-to-right is preferable when results are computed for a single initial vector. In this case, the computation for different reward vectors only requires a single vector product (provided the product of the initial distribution with the matrices has been precomputed).

Further optimizations of randomization in combination with a Kronecker representation of the generator matrix are considered in [39].

4. Model-checking Markov chains

The temporal logic CSL, developed originally in [1,2] and extended in [6], provides a powerful means to specify path-based as well as traditional state-based measures on CTMCs in a concise, flexible and unambiguous way. This paper considers model-checking CSL based on a Kronecker representation of the CTMC under consideration. This section presents the important basic concepts of paths and state-labeling, recalls the logic CSL, and summarizes its main model-checking algorithms.

4.1. Context

CSL is based on the well-known branching-time temporal logic CTL [21] and probabilistic CTL (PCTL) [28]; a steady-state operator, a time-bounded until, and a probabilistic (path) operator constitute its main ingredients. It allows one to state, for example, that the probability of reaching a certain set of goal-states within a specified real-valued time bound, provided that all paths to these states obey certain properties, is at least/at most some probability value.

Verification of a given finite-state CTMC against a CSL formula is performed using model checking. The model-checking problem for CSL is decidable for rational time bounds [1,2]. Approximate CSL model-checking algorithms have been studied in [6] where the satisfaction of time-bounded until formulas is shown to be based on solving a (recursive) Volterra equation system. More recently, verifying time-bounded until formulas has been reduced to the problem of computing transient-state probabilities for CTMCs [5]. This significant result employs a formula-dependent transformation of the CTMC and—more importantly—allows one to adopt efficient techniques like *randomization* [27,43] for verifying time-bounded until-formulas. A naive approach to model-check time-bounded until properties requires to perform this procedure for each state. An improvement suggested in [36] accumulates the required probabilities for all states simultaneously, yielding an improvement of $O(|\mathscr{T}|)$ time. Two prototype implementations for CSL model checking do exist: ETMCC [32] that is based on a representation of the CTMC by sparse matrices, and PRISM [36,41], a symbolic CSL model checker using MTBDDs.

4.2. State-labeling and paths

For model-checking purposes, CTMCs are extended with a state-labeling that indicates which elementary propositions hold in a state. In practice, atomic propositions connect a logic with a specific modeling formalism like SPNs or SPAs. Here, we consider state-based propositions, i.e., an atomic proposition is an (in)equality over functions with state variables as their parameters. A concrete example of a set of atomic propositions are (in)equalities of arithmetic expressions that are constructed from state variables, constants (e.g., reals) and arithmetic operations (such as $\{+, -, \cdot\}$). In a Petri net context, state variables are places and the value of a state variable in a given state is defined by the number of tokens at the place. Example atomic propositions are, e.g., "the number of tokens in place P exceeds 4", or "$P1$ and $P2$ together contain at least 2 tokens". More concrete examples are provided in Section 8.

Let AP be a fixed, finite set of atomic propositions. Function L is a *labeling* function which assigns to each state $s \in \mathscr{T}$ the set $L(s)$ of atomic propositions that are valid in s. To allow for a standard interpretation of temporal operators such as the until-operator in the untimed case, we do allow self-loops. We thus allow the system to occupy the same state before and after a transition. This does neither alter the transient nor the stationary behavior of the CTMC since diagonal entries of $\mathbf{Q}$ contain the negative row sums of off-diagonal values, so presence or absence of self-loops does not make a difference and is visible in $\mathbf{R}$, but not in $\mathbf{Q}$. Let $\mathbf{R} : \mathscr{T} \times \mathscr{T} \to \mathbb{R}_{\geqslant 0}$ be the rate matrix of the CTMC, with $\mathbf{R}(s, s) > 0$ for states equipped with a self-loop, and $\mathbf{e}(s) = \sum_{s' \in \mathscr{T}} \mathbf{R}(s, s')$. The probability of taking a transition from s within t time units thus equals $1 - \mathrm{e}^{-\mathbf{e}(s) \cdot t}$. A state s is absorbing if $\mathbf{R}(s, s') = 0$ for all states s'. We have $\mathbf{Q} = \mathbf{R} - \mathrm{diag}(\mathbf{e})$.

A path through a CTMC M is an alternating sequence $\sigma = s_0 t_0 s_1 t_1 s_2 \ldots$ with $\mathbf{R}(s_i, s_{i+1}) > 0$ and $t_i \in \mathbb{R}_{>0}$ for all i. The time stamps t_i denote the amount of time spent in state s_i. Let Path^M denote the set of paths through M. $\sigma[i]$ denotes the $(i+1)$th state of σ, i.e., $\sigma[i] = s_{i+1}$. $\sigma @ t$ denotes the state of σ occupied at time t, i.e., $\sigma @ t = \sigma[i]$ with i the largest index such that $t < \sum_{j=0}^{i} t_j$. Let $\Pr_s$ denote the unique probability measure on sets of paths that start in s [6].

4.3. The temporal logic CSL

Let $a \in AP$, $p \in [0, 1]$, comparison operator $\trianglelefteq \in \{\leqslant, \geqslant\}$ and $I \subseteq \mathbb{R}_{\geqslant 0}$. The syntax of CSL is:

$$\Phi ::= \mathrm{tt} | a | \Phi \wedge \Phi | \neg\Phi | \mathscr{S}_{\trianglelefteq p}(\Phi) | \mathscr{P}_{\trianglelefteq p}(\Phi\, \mathscr{U}^I\, \Phi).$$

The other Boolean connectives are derived in the usual way, i.e., $\mathrm{ff} = \neg\mathrm{tt}$, $\Phi \vee \Psi = \neg(\neg\Phi \wedge \neg\Psi)$, and $\Phi \to \Psi = \neg\Phi \vee \Psi$. $\mathscr{S}_{\trianglelefteq p}(\Phi)$ asserts that the steady-state probability for a Φ-state meets the bound $\trianglelefteq p$. $\mathscr{P}_{\trianglelefteq p}(\Phi\mathscr{U}^I\Psi)$ asserts that the probability measure of the paths satisfying $\Phi\mathscr{U}^I\Psi$ (see below) meets the bound given by $\trianglelefteq p$. (For the sake of simplicity, we do not consider the next state operator in this paper). The semantics of CSL is defined by [5,6]:

$$\begin{array}{ll}
s \models \mathrm{tt} & \text{for all } s \in \mathscr{T}, \\
s \models a & \text{iff } a \in L(s), \\
s \models \neg\Phi & \text{iff } s \not\models \Phi, \\
s \models \mathscr{S}_{\trianglelefteq p}(\Phi) & \text{iff } \lim_{t\to\infty} \Pr_s\{\sigma \in \mathrm{Path}^M \mid \sigma @ t \models \Phi\} \trianglelefteq p, \\
s \models \mathscr{P}_{\trianglelefteq p}(\Phi\, \mathscr{U}^I\, \Psi) & \text{iff } \mathrm{Prob}^M(s, \Phi\, \mathscr{U}^I\, \Psi) \trianglelefteq p.
\end{array}$$

The limit in the fourth equation always exists as M contains finitely many states [43]. $\mathrm{Prob}^M(\cdot)$ is defined by:

$$\mathrm{Prob}^M(s, \Phi\, \mathscr{U}^I\, \Psi) = \Pr_s\{\sigma \in \mathrm{Path}^M \mid \sigma \models \Phi\, \mathscr{U}^I\, \Psi\}.$$

$\Phi\mathscr{U}^I\Psi$ asserts that Ψ will be satisfied at some time instant in the interval I and that at all preceding time instants Φ holds:

$$\sigma \models \Phi\, \mathscr{U}^I\, \Psi \quad \text{iff } \exists\, t \in I \cdot (\sigma @ t \models \Psi \wedge \forall \tau < t, \sigma @ \tau \models \Phi).$$

Note that for $I = \emptyset$ the formula $\Phi\, \mathscr{U}^I\, \Psi$ is not satisfiable The standard until operator is obtained by taking I equal to $[0, \infty)$.

4.4. Model-checking algorithms

Model-checking CSL [5,6] is performed in the same way as for CTL [21] and PCTL [28], by recursively computing the set $\mathrm{Sat}(\Phi)$, i.e., the set of states that satisfy Φ. For the Boolean operators this is exactly as for CTL.

For determining $\mathrm{Sat}(\mathscr{S}_{\trianglelefteq p}(\Phi))$, first $\mathrm{Sat}(\Phi)$ is computed, and a graph analysis is carried out to determine the bottom strongly connected components (BSCCs) of M, i.e. the set of SCCs in M that, once entered, cannot be left any more. The steady-state probability distribution $\boldsymbol{\pi}^B$ inside each BSCC B is determined using standard means [43]: by solving a linear equation system in the size of the BSCC at hand. Then, the probabilities of reaching a BSCC B from a given state s are computed for each B. State s now satisfies $\mathscr{S}_{\trianglelefteq p}(\Phi)$ if:

$$\sum_B \left(\Pr\{\text{reach } B \text{ from } s\} \cdot \sum_{s' \in B \cap \mathrm{Sat}(\Phi)} \pi^B(s') \right) \trianglelefteq p.$$

Checking time-bounded until formulas is based on determining the least solution of a set of integral equations. For instance, for the simplest case, $I = [0, t]$ we have that $\mathrm{Prob}^M(s, \Phi\, \mathscr{U}^{[0,t]}\, \Psi)$ equals 1 if $s \in \mathrm{Sat}(\Psi)$,

$$\mathrm{Prob}^M(s, \Phi\,\mathscr{U}^{[0,t]}\,\Psi) = \int_0^t \sum_{s' \in \mathscr{S}} \frac{\mathbf{R}(s,s')}{\mathbf{e}(s)} \cdot \mathbf{e}(s) \cdot \mathrm{e}^{-\mathbf{e}(s)\cdot\tau} \cdot \mathrm{Prob}^M(s', \Phi\,\mathscr{U}^{[0,t-\tau]}\,\Psi)\,\mathrm{d}\tau$$

if $s \in \mathrm{Sat}(\Phi \wedge \neg\Psi)$, and equals 0 otherwise. Here, $\mathbf{e}(s) \cdot \mathrm{e}^{-\mathbf{e}(s)\cdot\tau}$ denotes the probability density of taking some outgoing transition from s at time τ. For intervals $I = [t, t']$ with $t > 0$ a similar, but more complicated equation is obtained (see [4]).

For CTMC $M = (\mathscr{S}, \mathbf{R}, L)$ and CSL-formula Φ let CTMC $M[\Phi] = (\mathscr{S}, \mathbf{R}', L)$ with $\mathbf{R}'(s,s') = \mathbf{R}(s,s')$ if $s \not\models \Phi$ and 0 otherwise. We have $M[\Phi][\Psi] = M[\Phi \vee \Psi]$.

It has been shown in [5] that for a given CTMC M and state s in M, the measure $\mathrm{Prob}^M(s, \Phi\,\mathscr{U}^I\,\Psi)$ can be calculated by means of a transient analysis of the CTMC M', which can easily be derived from M using the $[\cdot]$ operator. Let $\prod_t^M(s,s')$ denote the probability of being in state s' at time t given that the system started in state s, i.e. $\prod_t^M(s,s') = \mathrm{Pr}_s\{\sigma \in \mathrm{Path}^M \mid \sigma @ t = s'\}$. For intervals $I = [0, t]$ the following result applies [5].

Theorem 4.1. $\mathrm{Prob}^M(s, \Phi\,\mathscr{U}^{[0,t]}\,\Psi) = \sum_{s' \models \Psi} \prod_t^{M[\neg\Phi\vee\Psi]}(s,s')$.

Note that $M[\neg\Phi \vee \Psi] = M[\neg(\Phi \vee \Psi)][\Psi]$, i.e., all $\neg(\Phi \vee \Psi)$-states and all Ψ-states in M are made absorbing. The former is correct since $\Phi\,\mathscr{U}^{[0,t]}\,\Psi$ is violated as soon as some state is visited that neither satisfies Φ nor Ψ. The latter is correct since, once a Ψ-state in M has been reached (along a Φ-path) before time t, then $\Phi\,\mathscr{U}^{[0,t]}\,\Psi$ holds, regardless of which states will be visited later on.

For intervals $I = [t, t']$ with $t > 0$ the following result applies [5].

Theorem 4.2. *For CTMC M and $0 < t \leqslant t'$:*

$$\mathrm{Prob}^M(s, \Phi\,\mathscr{U}^{[t,t']}\,\Psi) = \sum_{s' \models \Phi} \sum_{s'' \models \Psi} \prod_t^{M[\neg\Phi]}(s,s') \cdot \prod_{t'-t}^{M[\neg\Phi\vee\Psi]}(s',s'').$$

Corollary 4.1. *For CTMC M* : $\mathrm{Prob}^M(s, \Phi\,\mathscr{U}^{[t,t]}\,\Psi) = \sum_{s' \models \Phi\wedge\Psi} \prod_t^{M[\neg\Phi]}(s,s')$.

5. Structured model-checking of CSL

This section details how the Kronecker representation of the generator matrix can be exploited for model-checking CSL. It presents the basic algorithms for the several operators (the propositional fragment, the steady-state operator, and the time-bounded until operator). Empirical results for the presented algorithms are given in Section 8.

Since any CSL-formula Φ describes a Boolean function $f_\Phi : \mathscr{S} \to \{\mathrm{ff}, \mathrm{tt}\}$, resp. $\{0,1\}$, model-checking amounts to a transformation of the parse tree of Φ into an explicit representation of f_Φ. This is done by a recursive post-fix traversal of the parse tree, that transforms each node of the tree into its Boolean function. There are different ways to represent Boolean functions by data structures. A bit-vector/array representation is straightforward and will be used in the following. BDDs or MDDs (multi-valued decision diagrams) are more space-efficient, but not employed here for several reasons. As the numerical computations for formulas of $\mathscr{P}_{\trianglelefteq p}(\Phi\,\mathscr{U}^I\,\Psi)$ and $\mathscr{S}_{\trianglelefteq p}(\Phi)$ are the dominating factor in terms of space and time consumption for CSL model-checking, the complexity for Boolean functions is of minor importance. Exchanging Boolean vectors by BDDs or MDDs is possible but is outside the scope of this paper.

5.1. The propositional fragment of CSL

For the propositional part of CSL, model-checking is performed as for CTL [21]. CTL model-checking using Kronecker representations has initially been presented in [40] and we use these results here. Evaluation of atomic proposition $a \in AP$ yields a Boolean function $f_a : \mathscr{T} \to \{\text{ff}, \text{tt}\}$. As atomic propositions are state-based propositions, their evaluation requires the values of state variables. To evaluate an atomic proposition for state $s = (x^1, \dots, x^K)$ in a DAG representation of $\mathscr{T}$ (cf. Section 2.4) the values of state variables related to each x^k for $0 < k \leqslant K$ have to be determined. For an explicit representation of f_a by a bit-vector, each state of $\mathscr{T}$ needs to be considered. A straightforward approach uses $RS_{\mathscr{T}}(\ )$, evaluates $s = (x^1, \dots, x^K)$ for each path $(x^1, \dots, x^K)$ and stores the resulting Boolean value in a bit-vector of length $|\mathscr{T}|$ at position $\Xi_{\mathscr{T}}(s)$.

As often state propositions are defined in terms of the state variables of one or a few components, this scheme can be improved by exploiting *locality of expressions*. This works as follows. The decomposition of a model into K components implies that state variables are split among components, so any component k is aware of a certain subset of state variables to define a local state x^k. Proposition a is *determined* by component k if all its state variables belong to component k. Then, $f_a(x^1, \dots, x^k) = f_a^k(x^k)$. Proposition a is *independent* of component k if none of its state variables belongs to component k. If proposition a is *determined* by component k, then its evaluation can be performed in two steps. First, we evaluate a locally on $\mathscr{T}^k$ with $f_a^k : \mathscr{T}^k \to \{\text{ff}, \text{tt}\}$. Then, f_a^k is projected to f_a by assigning the same value $f_a^k(x^k)$ to all elements of subset $RS_{\mathscr{T}}(x^1, \dots, x^k)$. This has the advantage that the second step can proceed without evaluating the atomic proposition (i.e., no consideration of state variables) and that for subsets with $|RS_{\mathscr{T}}(x^1, \dots, x^k)| > 1$, f_a is piece-wise constant.

If atomic propositions are combined by Boolean operators and both operands are local to component k, the second step can be omitted. In this case, the Boolean operator can be performed locally and the projection to the overall state space is postponed until evaluation of a formula requires it.

Evaluating formulas from the propositional fragment of CSL only uses the DAG representation of the reachable state space $\mathscr{T}$, but does not involve the Kronecker representation. This is different for steady-state and path formulas.

5.2. Structured analysis of steady-state formulas

The evaluation of formula $\Omega = \mathscr{S}_{\trianglelefteq p}(\Phi)$ involves numerical computations resulting in a column vector $\mathbf{q}$ (of length $|\mathscr{T}|$), such that for any state $s \in \mathscr{T}$, $s \models \Omega \iff \mathbf{q}(s) \trianglelefteq p$. Since the evaluation of $\mathbf{q}(s) \trianglelefteq p$ is trivial once $\mathbf{q}$ is known, we focus on the derivation of $\mathbf{q}$.

If $\mathscr{T}$ is strongly connected, we can employ steady-state analysis for Kronecker representations [14] to compute the solution of (1). Distribution $\boldsymbol{\pi}$ is used to derive $r = \sum_{s \in \text{Sat}(\Phi)} \boldsymbol{\pi}(s)$, and we have $\mathbf{q}(s) = r$ for all $s \in \mathscr{T}$. If $\mathscr{T}$ consists of m $(m > 1)$ BSCCs $\{B_1, \dots, B_m\}$ and transient states $B^t = \mathscr{T} \setminus \cup_{i=1}^m B_i$, we need a steady-state analysis m times, once for each BSCC, and, finally, a step to compute $\mathbf{q}[B^t]$. The partition of the state space into BSCCs and B^t is obtained by adapting the classical depth-first-search algorithm by Tarjan [44] to Kronecker representations. To that end, a function $\mathrm{succ} : \widehat{\mathscr{T}} \to 2^{\widehat{\mathscr{T}}}$ is used which gives the set of successor states $\mathrm{succ}(s)$ for any state $s = (x^1, \dots, x^K) \in \widehat{\mathscr{T}}$. For a Kronecker representation (3), the successor function of $s = (x^1, \dots, x^K)$ is defined by:

$$\mathrm{succ}(s) = \left\{(y^1, \dots, y^K) \mid \exists\, e \in \mathscr{E},\ \forall 0 < k \leqslant K : \mathbf{W}_{k,e}(x^k, y^k) \neq 0\right\}.$$

This function is used for all kinds of search algorithms on Kronecker representations, e.g., for computing $\mathscr{T}$ in [37] and for model-checking CTL path-formulas in [40]. Worst-case time complexity of Tarjan's algorithm is $\mathrm{O}(|\mathscr{T}| + e)$, where $\mathrm{O}(e)$ is the effort to compute all successors (outgoing edges) of states in $\mathscr{T}$. Roughly, $\mathrm{O}(e) = \mathrm{O}(|\mathscr{T}| \cdot |\mathscr{E}| \cdot K)$. Due to (i) relatively small values of K, (ii) special and more efficient treatment of Kronecker sums, and (iii) extreme sparseness of the involved matrices $\mathbf{W}_{k,e}$, this value is not critical in practice.

For each BSCC B_i a DAG is constructed for set Ξ_{B_i} and a steady-state analysis is carried out:

$$\boldsymbol{\pi}_{B_i} \cdot \mathbf{Q}[B_i, B_i] = 0 \quad \text{subject to } \boldsymbol{\pi}_{B_i} \cdot \mathbf{1}^{\mathrm{T}} = 1, \tag{7}$$

where $\boldsymbol{\pi}_{B_i}$ is the stationary distribution of BSCC B_i. This is done by the Kronecker-based solution approach of [14]. The resulting distribution $\boldsymbol{\pi}_{B_i}$ is then used to achieve a reward $r_i = \sum_{s \in \mathrm{Sat}(\Phi)} \boldsymbol{\pi}_{B_i}(s)$. This yields the vector $\boldsymbol{\rho}$, defined by $\boldsymbol{\rho}(s) = r_i$ if state $s \in B_i$ ($0 < i \leqslant m$) and $\boldsymbol{\rho}(s) = 0$ otherwise. For states of BSCCs, r_i also gives the resulting values for $\mathbf{q}$, i.e., $\mathbf{q}(s) = \boldsymbol{\rho}(s)$ if $s \in B_i$.

It remains to determine $\mathbf{q}(s)$ for transient states $s \in B^t$. Let $\mathbf{D}$ be the transition probability matrix of the embedded DTMC of $\mathbf{Q}$, i.e., $\mathbf{D} = \mathbf{I} + \mathrm{diag}(\mathbf{z}^{-1}) \cdot \mathbf{Q}$ where $\mathbf{z}(s) = -\mathbf{Q}(s, s)$ if $\mathbf{Q}(s, s) \neq 0$ and 1 otherwise. The Kronecker representation of $\mathbf{D}$ is obtained from $\widehat{\mathbf{R}}$ as in (6). For states of B^t we compute $\mathbf{q}$ from path probabilities towards BSCCs as follows:

$$\mathbf{q}[B^t] = \sum_{i=0}^{\infty} \mathbf{D}^i[B^t, B^t] \cdot \boldsymbol{\rho}', \tag{8}$$

where $\boldsymbol{\rho}' = \mathbf{D}[B^t, \mathscr{T} \backslash B^t] \cdot \boldsymbol{\rho}$. As B^t is not strongly connected, the matrix potentials converge towards $\mathbf{0}$ and the infinite sum naturally truncates at a finite point within the numerical precision. Note that Eq. (8) is computed by successive matrix–vector multiplications using the equivalent recursive formulation $\mathbf{x}^{(i+1)} = \mathbf{D}[B^t, B^t] \cdot \mathbf{x}^{(i)}$ and $\mathbf{x}^{(0)} = \boldsymbol{\rho}'$, such that $\mathbf{q}[B^t] = \sum_{i=0}^{\infty} \mathbf{x}^{(i)}$ and $\mathbf{x}^{(i)}$ becomes $\mathbf{0}$ for $i \to \infty$. The use of $\mathbf{D}[B^t, B^t]$ requires a DAG with function Ξ_{B^t} to allow for matrix–vector multiplications with Kronecker representations.

5.3. Structured analysis of time-bounded until-formulas

According to Corollary 4.1 and Theorems 4.1 and 4.2, model-checking time-bounded until-formulas can be done using randomization [27,43] provided the CTMC is appropriately transformed. Three cases for interval $I = [t, t']$ are distinguished: (i) $0 < t = t'$, (ii) $0 = t < t'$, and (iii) $0 < t < t'$. Randomization is based on successive matrix–vector multiplications which makes it amenable to Kronecker representations, i.e., one can use matrix–vector multiplication algorithms of various kind as in [14]. Let $\mathbf{P} = \mathbf{I} + \lambda^{-1} \cdot \mathbf{Q}$ denote the matrix used in randomization with a Kronecker representation as given in (6).

Evaluating $\mathscr{P}_{\trianglelefteq p}(\Phi\,\mathscr{U}^I\,\Psi)$ requires a transformation of the CTMC that makes certain sets of states absorbing. Let $\mathscr{S}_1$ denote the set of non-absorbing states. Since $\mathbf{P}$ is given by a Kronecker representation, we need to avoid matrix transformations in computations, e.g., we select a relevant sub-matrix like $\mathbf{P}[\mathscr{S}_1, \mathscr{S}_1]$ in a matrix–vector multiplication involving some vector $\mathbf{x}$ of dimension $|\mathscr{S}_1|$ using a DAG for $\Xi_{\mathscr{S}_1}$ (as in multiplication algorithms presented in [14]). A straightforward extension uses two DAGs $\Xi_{\mathscr{S}_1}$ and $\Xi_{\mathscr{S}_2}$ to perform multiplications on arbitrary sub-matrices $\mathbf{P}[\mathscr{S}_1, \mathscr{S}_2]$ for sets $\mathscr{S}_1, \mathscr{S}_2 \subseteq \mathscr{T}$.

A key observation in [36] is to perform randomization backwards, i.e., one fixes a reward vector and evaluates equations like (5) from the right side for an unknown, arbitrary initial distribution. In combination with the preceding remarks on absorbing states, it is clear that only absorbing states that have a positive reward can have an impact on the computed values for $\mathscr{S}_1$. Let $\mathscr{S}_2$ denote the set of such absorbing states and $\boldsymbol{\rho}$ a reward vector. The impact of absorbing states on $\mathscr{S}_1$ is then given by $\mathbf{P}[\mathscr{S}_1, \mathscr{S}_2] \cdot \boldsymbol{\rho}[\mathscr{S}_2]$. We will use this observation in the following whenever absorbing states contribute positive rewards.

5.3.1. *Point intervals $I = [t, t']$ with $0 < t = t'$*

This case basically requires computing the transient distribution π_t at a time point t and quantify it with 1-rewards for states in $\mathrm{Sat}(\Phi \wedge \Psi)$, cf. Corollary 4.1. To avoid computing the transient distribution for each state $s \in \mathscr{T}$, we follow the observation of [36] and evaluate an equation like (5) from right-to-left. Let $\mathscr{S}_1 = \mathrm{Sat}(\Phi)$ and $\mathscr{S}_2 = \mathscr{T} \backslash \mathscr{S}_1$. Reward values of states are $\boldsymbol{\rho}(s) = 1$ if $s \in \mathrm{Sat}(\Psi) \cap \mathscr{S}_1$ and 0 otherwise. States of $\mathscr{S}_2$ are made absorbing according to Corollary 4.1, i.e., $\mathbf{q}(s) = 0$ if $s \in \mathscr{S}_2$. It remains to determine $\mathbf{q}[\mathscr{S}_1]$. Since $\boldsymbol{\rho}(s) = 0$ for $s \in \mathscr{S}_2$, states of $\mathscr{S}_2$ have no impact on the results for states in $\mathscr{S}_1$ and can be ignored. We have

$$\mathbf{q}[\mathscr{S}_1] = \sum_{i=0}^{\infty} \beta_i(\lambda, t) \cdot \mathbf{P}^i[\mathscr{S}_1, \mathscr{S}_1] \cdot \boldsymbol{\rho}[\mathscr{S}_1]. \tag{9}$$

The matrix potentials are computed by successive matrix–vector multiplication using the equivalent recursive definition $\mathbf{x}^{(i+1)} = \mathbf{P}[\mathscr{S}_1, \mathscr{S}_1] \cdot \mathbf{x}^{(i)}$ and $\mathbf{x}^{(0)} = \boldsymbol{\rho}[\mathscr{S}_1]$, such that $\mathbf{q}[\mathscr{S}_1] = \sum_{i=0}^{\infty} \beta_i(\lambda, t) \cdot \mathbf{x}^{(i)}$. The infinite summation is truncated as soon as at least one of the following truncation (termination) criteria holds.

First, the derivation of the Poisson distribution by the approach of [26] provides only a finite number of values, so the right truncation point r imposed by [26] limits the summation in a natural way. Since for $i > r$, $\beta_i(\lambda, t) = 0$ no further summation is necessary. Secondly, if the iteration reaches a fix-point at some step j, i.e., $\mathbf{x}^{(j+1)} = \mathbf{x}^{(j)}$, then we need not compute further iteration vectors since they remain the same. So we can directly compute the value of the infinite sum $\mathbf{q}[\mathscr{S}_1] = \sum_{i=0}^{j} \beta_i(\lambda, t) \cdot \mathbf{x}^{(i)} + \left(1 - \sum_{i=0}^{j} \beta_i(\lambda, t)\right) \cdot \mathbf{x}^{(j)}$. Thirdly, if the intermediate results are sufficiently precise to determine whether the constraint $\trianglelefteq p$ holds or not, then further iterations would only improve on the numerical precision of the transient distribution but not change the Boolean result we want to compute for the CSL formula. Note that this phenomenon is possible because the iterative computation of $\mathbf{q}[\mathscr{S}_1]$ is a summation of non-negative vectors which implies that vector entries are non-decreasing. Let $\mathbf{q}^i[\mathscr{S}_1]$ denote the sum of the first i terms. By construction we have $\mathbf{q}^i[\mathscr{S}_1] \leqslant \mathbf{q}^{i+1}[\mathscr{S}_1]$ for all $i > 0$. On the other hand, $\mathbf{P}$ is a probabilistic matrix and $\boldsymbol{\rho}(s) \in \{0, 1\}$ for all $s \in \mathscr{T}$, so $\mathbf{x}^{(j)}(s) \in [0, 1]$ for all $j \geqslant 0$. This allows us to conclude that $\sum_{i=j+1}^{\infty} \beta_i(\lambda, t) \cdot \mathbf{x}^{(i)}(s) \leqslant \left(1 - \sum_{i=0}^{j} \beta_i(\lambda, t)\right) \cdot 1 = \delta(j)$ for any $s \in \mathscr{T}$. We obtain

$$\underbrace{\mathbf{q}^i[\mathscr{S}_1](s)}_{=\mathrm{lb}(s)} \leqslant \mathbf{q}[\mathscr{S}_1](s) \leqslant \underbrace{\mathbf{q}^i[\mathscr{S}_1](s) + \delta(i)}_{=\mathrm{ub}(s)}. \tag{10}$$

For formula $\Omega = \mathscr{P}_{\trianglelefteq p}(\Phi\, \mathscr{U}^I\, \Psi)$ with $p \notin [\mathrm{lb}(s), \mathrm{ub}(s)]$ (for all $s \in \mathscr{S}_1$) we can now terminate, since the evaluation of Ω is fully determined. Note that the computation of bounds is computationally inexpensive, since $\delta(j)$ is simply the remaining probability $1 - \sum_{i=0}^{j} \beta_i(\lambda, t)$. All termination criteria together help to minimize the number of iteration steps.

5.3.2. *Anchored intervals $I = [t, t']$ for $0 = t < t'$*

Let $\mathscr{S}_1 = \mathrm{Sat}(\Phi) \cap \mathrm{Sat}(\neg\Psi)$ and $\mathscr{S}_2 = \mathrm{Sat}(\Psi)$. States of $\mathscr{T}\backslash\mathscr{S}_1$ are made absorbing (cf. Theorem 4.1). The reward values are $\boldsymbol{\rho}(s) = 1$ if $s \in \mathscr{S}_2$ and 0 otherwise. By construction, $\mathbf{q}(s) = 1$ if $s \in \mathscr{S}_2$ and 0 if $s \in \mathscr{T}\backslash(\mathscr{S}_1 \cup \mathscr{S}_2)$. So it remains to determine $\mathbf{q}[\mathscr{S}_1]$. Since $\boldsymbol{\rho}$ provides positive rewards for states that are not in $\mathscr{S}_1$, we cannot simply use $\boldsymbol{\rho}[\mathscr{S}_1]$ for the iteration as in the case above, but we have to account for the impact of $\boldsymbol{\rho}[\mathscr{S}_2]$. Nevertheless, the iteration can focus on $\mathscr{S}_1$ if we define $\boldsymbol{\rho}'[\mathscr{S}_1] = \mathbf{P}[\mathscr{S}_1, \mathscr{S}_2] \cdot \boldsymbol{\rho}[\mathscr{S}_2]$. Note that $\mathscr{S}_2$ is made absorbing, such that $\mathbf{P}[\mathscr{S}_2, \mathscr{S}_2] = \mathbf{I}$ and the sum of rewards gained so far by $\boldsymbol{\rho}'[\mathscr{S}_1]$ adds up in each step. Using Theorem 4.1 we get

$$\mathbf{q}[\mathscr{S}_1] = \beta_1(\lambda, t') \cdot \boldsymbol{\rho}'[\mathscr{S}_1] + \sum_{i=2}^{\infty} \beta_i(\lambda, t') \cdot \sum_{j=0}^{i-1} \mathbf{P}^j[\mathscr{S}_1, \mathscr{S}_1] \cdot \boldsymbol{\rho}'[\mathscr{S}_1]. \tag{11}$$

Note that for $i = 0$, the reward is $\mathbf{0}$ for $\mathscr{S}_1$, and for $i = 1$ vector $\boldsymbol{\rho}'$ gives the reward earned by reaching $\mathscr{S}_2$ in a single transition. The matrix potentials and their summation are computed by successive matrix–vector multiplications using the equivalent recursive definition $\mathbf{x}^{(i+1)} = \mathbf{P}[\mathscr{S}_1, \mathscr{S}_1] \cdot \mathbf{x}^{(i)} + \boldsymbol{\rho}'[\mathscr{S}_1]$ and $\mathbf{x}^{(1)} = \boldsymbol{\rho}'[\mathscr{S}_1]$, such that $\mathbf{q}[\mathscr{S}_1] = \sum_{i=1}^{\infty} \beta_i(\lambda, t') \cdot \mathbf{x}^{(i)}$. The iteration is truncated by the termination criteria mentioned above. The specific observation in Eq. (11) is that once a Ψ-state is reached along a Φ-path, this path remains in the Ψ-state forever and contributes to the reward accumulation in each subsequent step. If we write out the summation of (11) in a matrix where index i refers to a row, we obtain a lower tridiagonal matrix of infinite dimension (see Table 1).

We can also write up these terms according to columns using the exponent of $\mathbf{P}$ as an index of terms. Since values of β_i are probabilities, $\sum_{i=0}^{\infty} \beta_i(\lambda, t') = 1$ and we define weights $\gamma_i(\lambda, t') = 1 - \sum_{j=0}^{i-1} \beta_j(\lambda, t')$ to obtain $\sum_{j=i}^{\infty} \beta_j(\lambda, t') \cdot \mathbf{P}^{i-1}[\mathscr{S}_1, \mathscr{S}_1] \cdot \boldsymbol{\rho}'[\mathscr{S}_1] = \gamma_i(\lambda, t') \cdot \mathbf{P}^{i-1}[\mathscr{S}_1, \mathscr{S}_1] \cdot \boldsymbol{\rho}'[\mathscr{S}_1]$ for the sum of terms in column $i - 1$. This yields the following alternative formulation of (11):

$$\mathbf{q}[\mathscr{S}_1] = \gamma_1(\lambda, t') \cdot \boldsymbol{\rho}'[\mathscr{S}_1] + \sum_{i=2}^{\infty} \gamma_i(\lambda, t') \cdot \mathbf{P}^{i-1}[\mathscr{S}_1, \mathscr{S}_1] \cdot \boldsymbol{\rho}'[\mathscr{S}_1]. \tag{12}$$

Table 1
Matrix of terms in Eq. (11)

0	$\cdots$			
$\beta_1(\lambda, t') \cdot \boldsymbol{\rho}'[\mathscr{S}_1]$	0	$\cdots$		
$\beta_2(\lambda, t') \cdot \boldsymbol{\rho}'[\mathscr{S}_1]$	$\beta_2(\lambda, t') \cdot \mathbf{P}^1[\mathscr{S}_1, \mathscr{S}_1] \cdot \boldsymbol{\rho}'[\mathscr{S}_1]$	0	$\cdots$	
$\beta_3(\lambda, t') \cdot \boldsymbol{\rho}'[\mathscr{S}_1]$	$\beta_3(\lambda, t') \cdot \mathbf{P}^1[\mathscr{S}_1, \mathscr{S}_1] \cdot \boldsymbol{\rho}'[\mathscr{S}_1]$	$\beta_3(\lambda, t') \cdot \mathbf{P}^2[\mathscr{S}_1, \mathscr{S}_1] \cdot \boldsymbol{\rho}'[\mathscr{S}_1]$	0	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\ddots$

The latter variant is more efficient, since it saves a vector addition per iteration step. In addition, the test for termination uses an adapted $\delta'(i) = \sum_{j=i+1}^{\infty} \gamma_j(\lambda, t')$; note that $\gamma_j = 0$ if j exceeds the right truncation point of the computed Poisson distribution. Since the implementation of (12) iterates different vectors $\mathbf{x}$ than the one of (11), a fix-point may be reached at a different number of steps (if reached at all) and the termination criterion using the upper and lower bound is less effective due to possibly larger values of δ'. So the gain in efficiency per step can be outweighed by an increased number of iteration steps, as discussed in Section 8.

5.3.3. *Non-anchored intervals $I = [t, t']$ with $0 < t < t'$*

This case is the most complex one. It encompasses a randomization for $[0, t'-t]$ and a randomization for $[t, t]$, cf. Theorem 4.2. Let $\mathscr{S}_1 = \mathrm{Sat}(\Phi) \cap \mathrm{Sat}(\neg\Psi)$, $\mathscr{S}_2 = \mathrm{Sat}(\Psi)$, and $\mathscr{S}_3 = \mathrm{Sat}(\Phi)$. It is clear, that $\mathbf{q}(s) = 0$ if $s \notin \mathscr{S}_3$ and that we need to compute $\mathbf{q}[\mathscr{S}_3]$. We need different reward vectors for each randomization. Let $\boldsymbol{\rho}(s) = 1$ if $s \in \mathscr{S}_2$ and 0 otherwise; $\boldsymbol{\rho}(s) = 1$ thus denotes that terminal state s fulfills Ψ. Such states shall be reached within $t'-t$ time units via Φ-states. Hence, we define $\boldsymbol{\rho}'[\mathscr{S}_1] = \mathbf{P}[\mathscr{S}_1, \mathscr{S}_2] \cdot \boldsymbol{\rho}[\mathscr{S}_2] = \mathbf{P}[\mathscr{S}_1, \mathscr{S}_2] \cdot \mathbf{1}^{\mathrm{T}}$, since $\boldsymbol{\rho}[\mathscr{S}_2] = \mathbf{1}$ (as in the case for the anchored interval above). At time point t, the reward for each state in $\mathscr{S}_1$ to reach $\mathscr{S}_2$ within time $t'-t$ follows from (12):

$$\boldsymbol{\rho}''[\mathscr{S}_1] = \gamma_1(\lambda, t'-t) \cdot \boldsymbol{\rho}'[\mathscr{S}_1] + \sum_{i=2}^{\infty} \gamma_i(\lambda, t'-t) \cdot \mathbf{P}^{i-1}[\mathscr{S}_1, \mathscr{S}_1] \cdot \boldsymbol{\rho}'[\mathscr{S}_1].$$

Computing the latter reward requires basically a randomization for interval $[0, t'-t]$. Let $\boldsymbol{\rho}''[\mathscr{S}_3](s) = 0$ for states $s \in \mathscr{S}_3 \backslash \mathscr{S}_1$. Then $\mathbf{q}$ is obtained by using the approach for a point interval (9):

$$\mathbf{q}[\mathscr{S}_3] = \sum_{i=0}^{\infty} \beta_i(\lambda, t) \cdot \mathbf{P}^i[\mathscr{S}_3, \mathscr{S}_3] \cdot \boldsymbol{\rho}''[\mathscr{S}_3]. \tag{13}$$

The computation takes place in two phases using the above solutions for interval $[0, t'-t]$ and $[t, t]$ but with adapted reward vectors of different dimensions.

6. Eliminating vanishing states

In addition to transitions with an exponential delay (*timed transitions*), many modeling formalisms include transitions without delay, i.e., transitions that take place immediately (*immediate transitions*). Examples of such formalisms are GSPNs [18] and SPAs with immediate transitions [31,33]. For such models the state space contains two classes of states: *vanishing* states where immediate transitions are enabled and *tangible* states where no immediate transition is enabled. Typically, immediate transitions are eliminated a priori. This results in a state space of a CTMC with only tangible states. Such elimination is possible, provided the model fulfills some quite natural conditions [18]. However, as we will clarify in this section, vanishing states may cause some problems in structured analysis approaches and in model-checking CSL.

6.1. *Vanishing states in structured analysis*

In structured analysis it is often assumed that all immediate transitions are local, i.e., they do not participate in a synchronization. If this is the case, then vanishing states can be eliminated locally in the component state spaces. This yields state spaces $\widehat{\mathscr{T}}^k$ and matrices $\mathbf{W}_{k,e}$ as introduced above. The problem of synchronization via immediate transitions in structured analysis has been considered in [25]. In this case, stationary analysis is performed on an embedded DTMC with a state space that contains vanishing and tangible states. As this approach cannot be applied for transient analysis, we require that all synchronized transitions are timed and immediate transitions occur only locally in the components.

For CSL model-checking the problem of vanishing states has not been considered yet. We present here a first approach for structured model description which, however, can be easily adopted to a conventional description of a CTMC without any structure (i.e., the degenerated structured case with one component only).

First, we describe how vanishing states can be eliminated from component state spaces, such that the resulting component k is described by state space $\widehat{\mathscr{T}}^k$ and matrices $\mathbf{W}_{k,e}$. After state-space generation, the state space consists of $\widehat{\mathscr{T}}^k$, the set of tangible states, and $\widehat{\mathscr{V}}^k$, the set of vanishing states. We assume w.l.o.g. that states are ordered such that vanishing states precede tangible ones. Furthermore, we assume that all immediate transitions are labeled with $\tau \notin \mathscr{E}$ where τ is not used for synchronization. Consequently, the set $\mathscr{E}'$ of events of the complete model equals $\mathscr{E} \cup \{\tau\}$. Let $n_k = |\widehat{\mathscr{T}}^k|$ and $m_k = |\widehat{\mathscr{V}}^k|$. The dynamics of component k are described by several matrices. Matrix $\mathbf{V}^{tt}_{k,e} \in \mathbb{R}^{n_k \times n_k}$ contains rates of transitions labeled with $e \in \mathscr{E}$ which end in a tangible state. Similarly, matrix $\mathbf{V}^{tv}_{k,e} \in \mathbb{R}^{n_k \times m_k}$ contains rates of transitions labeled with $e \in \mathscr{E}$ which end in a vanishing state. Matrix $\mathbf{U}^{vv}_{k,\tau} \in \mathbb{R}^{m_k \times m_k}$ contains transition weights of immediate transitions which end in vanishing states and $\mathbf{U}^{vt}_{k,\tau} \in \mathbb{R}^{m_k \times n_k}$ is a matrix of transition weights of immediate transitions which end in tangible states. We assume that the weights of immediate transitions are normalized such that:

$$\mathbf{U}^{vv}_{k,\tau} \cdot \mathbf{1}^{\mathrm{T}} + \mathbf{U}^{vt}_{k,\tau} \cdot \mathbf{1}^{\mathrm{T}} = \mathbf{1}^{\mathrm{T}}.$$

This can always be achieved by normalization of the matrix rows. Remember that all weights are non-negative and in each vanishing state at least one immediate transition is enabled.

The required matrices $\mathbf{W}_{k,e}$ can be computed as

$$\mathbf{W}_{k,e} = \mathbf{V}^{tt}_{k,e} + \mathbf{V}^{tv}_{k,e} \cdot \left(\mathbf{I}_{n_k} - \mathbf{U}^{vv}_{k,\tau}\right)^{-1} \cdot \mathbf{U}^{vt}_{k,\tau}. \tag{14}$$

The inverse matrix exists, if the model includes no trap of immediate transitions which is assumed here. Different methods to generate $\mathbf{W}_{k,e}$ exist, often it is not necessary to really perform the matrix inversion.

Matrices $\mathbf{W}_{k,e}$ and state space $\widehat{\mathscr{T}}^k$ can then be used in structured analysis approaches as described above.

6.2. Component-wise elimination of vanishing states

The question is whether we can apply the same approach to CSL model-checking. Propositional formulas are uncritical because they can be evaluated in vanishing states like in tangible states. Slightly more complex is model checking of $\mathscr{S}_{\trianglelefteq p}(\Phi)$ with vanishing states. Assume, like in Section 5, that the states space $\mathscr{T}$ consists of m BSCCs $\{D_1, \ldots, D_m\}$ and transient states D^t. In contrast to Section 5 we now assume that D_i contains tangible and vanishing states, i.e., $D_i = B_i \cup C_i$ (and $D^t = B^t \cup C^t$) where C_i (C^t) is the set of vanishing states in the ith BSCC and B_i (B^t) the set of transient states, respectively. $\mathscr{S}_{\trianglelefteq p}(\Phi)$ provides no problems for vanishing states inside a BSCC, because the stationary probability of each vanishing state is 0. Vanishing states in C_i thus receive the same values $\boldsymbol{\rho}(s)$ as tangible states B_i. All states from D_i (or no state in D_i) fulfill $\mathscr{S}_{\trianglelefteq p}(\Phi)$ and this can be derived from the stationary results by considering only tangible states. For vanishing states in C^t path probabilities have to be computed. For $y^k \in \widehat{\mathscr{V}}^k$ and $x^k \in \widehat{\mathscr{T}}^k$ let $A\,\mathrm{Prob}(y^k, x^k)$ be the probability of reaching x^k as the first tangible state when the component is in vanishing state y^k. The probability $A\,\mathrm{Prob}(y^k, x^k)$ can be computed as:

$$\mathbf{e}_{y^k} \cdot \sum_{i=0}^{\infty} \left(\mathbf{U}_{k,\tau}^{vv}\right)^i \cdot \mathbf{U}_{k,\tau}^{vt} \left(\mathbf{e}_{x^k}\right)^{\mathrm{T}} = \mathbf{e}_{y^k} \cdot \left(\mathbf{I}_{n_k} - \mathbf{U}_{k,\tau}^{vv}\right)^{-1} \cdot \mathbf{U}_{k,\tau}^{vt} \left(\mathbf{e}_{x^k}\right)^{\mathrm{T}},$$

where $\mathbf{e}_x$ is a row vector of length n_k with 1 in position x and 0 elsewhere. Let $s = (y^1, \ldots, y^K) \in C^t$, then

$$\mathbf{q}[C^t](s) = \sum_{s'=(x^1,\ldots,x^K)\in\mathscr{T}} \left(\prod_{k=1}^{K} AP(y^k, x^k)\right) \cdot \mathbf{q}(s').$$

Knowing $\mathbf{q}(s)$ for vanishing states $\mathscr{S}_{\trianglelefteq p}(\Phi)$ can be easily evaluated, i.e., $s \models \mathscr{S}_{\trianglelefteq p}(\Phi)$ if $\mathbf{q}(s) \trianglelefteq p$.

The situation is more complex for $\mathscr{P}_{\trianglelefteq p}(\Phi\, \mathscr{U}^I\, \Psi)$, e.g., consider $\mathscr{P}_{\trianglelefteq p}(\Phi\, \mathscr{U}^I\, \Psi)$ and assume that Φ or Ψ hold in all tangible states, but in vanishing states neither Φ nor Ψ do hold. Consequently, each path which passes through vanishing states does not fulfill $\mathscr{P}_{\trianglelefteq p}(\Phi\, \mathscr{U}^I\, \Psi)$ because Φ and Ψ are not observed for a zero amount of time. This, unfortunately, is not visible after the elimination of vanishing states.

A straightforward solution to the above problem would be to say that the system cannot be observed during a time interval of length zero and therefore it does not matter which CSL formulas hold or do not hold in vanishing states because results are only interesting if they are observed for some time. This solution is easy and allows us to eliminate vanishing states locally without taking care of the CSL formulas to be checked with the model. However, the solution is against the definition of CSL path-formulas where results are defined according to the states and transitions which are observed independently of the time. Thus, we consider in the sequel the case where immediate states have to be considered for the verification of CSL path-formulas. Let $\mathrm{Path}_{im}(s, s')$ be the set of all paths σ that start in state s, end in state s' and pass only through vanishing states.

Definition 6.1. $\mathscr{P}_{\trianglelefteq p}(\Phi\, \mathscr{U}^I\, \Psi)$ is invariant under immediate transitions, iff
- for all $s, s' \in \mathscr{T}$ with $s, s' \models \Phi \wedge s, s' \models \neg\Psi$, and all $\sigma \in \mathrm{Path}_{im}(s, s')$: $\sigma[i] \models \Phi \wedge \sigma[i] \models \neg\Psi$ for all $0 < i \leqslant |\sigma|$, and
- for all $s, s' \in \mathscr{T}$ with $s \models \Phi \wedge s \models \neg\Psi$, $s' \models \Psi$ and all $\sigma \in \mathrm{Path}_{im}(s, s')$: exists $i \leqslant |\sigma|$ such that $\sigma[j] \models \Phi \wedge \sigma[i] \models \neg\Psi$ for $j < i$ and $\sigma[k] \models \Psi$ for $k \geqslant i$,

where $|\sigma|$ is the length of path σ.

In general it is hard to check at state-space level whether a CSL-formula is invariant under immediate transitions or not. However, such a check can be done locally for some component k if Φ and Ψ are logical combinations of atomic propositions which are determined by k. In this case, invariance under immediate transitions can be checked using the matrices for component k only, since the functions f^k_Φ and f^k_Ψ can be evaluated for each state of k and the properties can be proved locally.

Let $\models_M$ denote the satisfaction relation $\models$ (on CSL) for M, and let $\widehat{M}$ denote the result of removing vanishing states in M. Then:

Corollary 6.1. *If $\Omega = \mathscr{P}_{\trianglelefteq p}(\Phi\,\mathscr{U}^I\,\Psi)$ is invariant under immediate transitions:*

$$\forall s \in \widehat{M} : s \models_{\widehat{M}} \Omega \quad \textit{iff}\; s \models_M \Omega.$$

6.3. Exploiting locality

Alternatively, for formulas whose validity is determined by one component only, vanishing states can be eliminated. We consider two extreme cases for the dependency of a CSL-formula on the state of a component, viz. either it is independent or (fully) determined by such state.

Definition 6.2. Φ is independent of component k iff

$$(x^1, \ldots, x^{k-1}, x^k, x^{k+1}, \ldots, x^K) \models \Phi \Longrightarrow (x^1, \ldots, x^{k-1}, y^k, x^{k+1}, \ldots, x^K) \models \Phi$$

for all $y^k \in \widehat{\mathscr{T}}^k$.
Φ is *determined* by component k iff

$$(x^1, \ldots, x^{k-1}, x^k, x^{k+1}, \ldots, x^K) \models \Phi \Longrightarrow (y^1, \ldots, y^{k-1}, x^k, y^{k+1}, \ldots, y^K) \models \Phi$$

for all $y^l \in \widehat{\mathscr{T}}^l$ and $l \in \{1, \ldots, K\} \setminus \{k\}$.

As before it is in general hard to decide whether Φ is *independent* or *determined* by a component. However, we can define two simple sufficient conditions, namely if Φ contains only atomic propositions which are *determined* by component k, then Φ is *determined* by component k and if Φ contains only atomic propositions which are *independent* of component k, then Φ is *independent* of component k. Observe that the conditions always hold in cases where model checking is used to analyze the behavior of some component embedded in an environment. If $\mathscr{P}_{\trianglelefteq p}(\Phi\,\mathscr{U}^I\,\Psi)$ is independent of component k, then its validity is preserved under the elimination of vanishing states locally in k.

More interesting, though, is the case where the formula is determined by k. Elimination of vanishing states can be made independently of the interval I. We first add two new absorbing tangible states x^k_s and x^k_f with $x^k_s \models \Psi$ and $x^k_f \models \neg(\Phi \vee \Psi)$. State x^k_s is used to model a successful transition into a state where Ψ holds and state x^k_f models the situation that a state is entered where Φ and Ψ both do not hold. Both states are used to preserve the effects which occur in vanishing states after elimination of vanishing states. The set of vanishing states $\widehat{\mathscr{V}}^k$ is decomposed into three subsets, $\mathscr{S}_1$ contains all states where Φ but not Ψ holds, $\mathscr{S}_2$ contains all states where Ψ holds and $\mathscr{S}_3$ contains the remaining states. All transitions from $x \in \mathscr{S}_1$ to $y \in \mathscr{S}_2$ are substituted by a transition from x to tangible state x^k_s. The weights remain the same and are possibly added, if x has more than one successor $\mathscr{S}_2$. Similarly, all transitions from $x \in \mathscr{S}_1$ to $y \in \mathscr{S}_3$ are substituted by a transition from x to tangible state x^k_f. Furthermore, transitions from $x \in \mathscr{T}^k$ into $\mathscr{S}_2$ and $\mathscr{S}_3$ are substituted by transitions into x^k_s and x^k_f, respectively. Let $\mathbf{V}^{tt}_{k,e}$, $\mathbf{V}^{tv}_{k,e}$, $\mathbf{U}^{vv}_{k,\tau}$ and $\mathbf{U}^{vt}_{k,\tau}$

be the matrices prior to the transformations. We add states x_s^k and x_f^k as the last and second last state to $\widehat{\mathscr{T}}^k$ and obtain matrices $\mathbf{V}_{k,e}^{*tt}$ and $\mathbf{U}_{k,\tau}^{*vt}[\mathscr{S}_1, \widehat{\mathscr{T}}^k]$ on the new state space with

$$\mathbf{V}_{k,e}^{*tt} = \begin{pmatrix} \mathbf{V}_{k,e}^{tt} & \mathbf{V}_{k,e}^{tv}[\widehat{\mathscr{T}}^k, \mathscr{S}_2] \cdot \mathbf{1}^{\mathrm{T}} & \mathbf{V}_{k,e}^{tv}[\widehat{\mathscr{T}}^k, \mathscr{S}_3] \cdot \mathbf{1}^{\mathrm{T}} \\ \mathbf{0} & 0 & 0 \\ \mathbf{0} & 0 & 0 \end{pmatrix},$$

$$\mathbf{U}_{k,\tau}^{*vt}[\mathscr{S}_1, \widehat{\mathscr{T}}^k] = \left(\mathbf{U}_{k,\tau}^{vt}[\mathscr{S}_1, \widehat{\mathscr{T}}^k] \quad \mathbf{U}_{k,\tau}^{vv}[\mathscr{S}_1, \mathscr{S}_2] \cdot \mathbf{1}^{\mathrm{T}} \quad \mathbf{U}_{k,\tau}^{vv}[\mathscr{S}_1, \mathscr{S}_3] \cdot \mathbf{1}^{\mathrm{T}} \right).$$

The remaining matrices need not be modified. Matrices $\mathbf{W}_{k,e}$ can then be computed by:

$$\mathbf{W}_{k,e}^{*} = \mathbf{V}_{k,e}^{*tt} + \mathbf{V}_{k,e}^{tv}[\mathscr{T}, \mathscr{S}_1] \cdot \left(\mathbf{I}_{n_k} - \mathbf{U}_{k,\tau}^{vv}[\mathscr{S}_1, \mathscr{S}_1]\right)^{-1} \cdot \mathbf{U}_{k,\tau}^{*vt}[\mathscr{S}_1, \widehat{\mathscr{T}}^k]. \qquad (15)$$

Observe that $\mathbf{W}_{k,e}$ has dimension $n_k + 2$ instead of n_k because the two additional states have been added to make immediate events visible for CTMC analysis. Since the two additional states make changes according to formulas Φ and Ψ visible in the tangible state space, the following corollary holds.

Let M^k denote the CTMC obtained from M by representing its generator matrix in a structured form using matrices $\mathbf{W}_{l,e}$ for components $l \neq k$ and matrices $\mathbf{W}_{k,e}^{*}$ for component k. Then:

Corollary 6.2. *If $\Omega = \mathscr{P}_{\trianglelefteq p}(\Phi\, \mathscr{U}^I\, \Psi)$ is determined by component k, then*

$$\forall s \in M^k : s \models_M \Omega \quad \textit{iff } s \models_{M^k} \Omega.$$

7. On the role of bisimulations

As shown in [5], the validity of CSL-formulas is preserved under stochastic bisimulation equivalence (i.e., lumping). Thus it is possible to first build a minimal representation of a CTMC up to bisimulation and then model-check the reduced instead of the entire CTMC. Moreover, stochastic bisimulation is a congruence with respect to the composition of components by synchronization [9,31]. Bisimulation minimization can thus be applied in a compositional way. This approach can be naturally combined with structured analysis techniques where the generator matrix is represented as the sum of Kronecker products of component matrices. Components are first reduced modulo bisimulation, and a reduced representation of the generator matrix of the entire model is then given in a structured form by replacing the entire component matrices by the matrices for the respective reduced component. This section adapts this approach to model-checking CSL.

We first recall some notions and results from [5] where we confine ourselves to bisimulations over sets of atomic propositions. As we consider structured CTMC descriptions, we start with bisimulation at a component level as our basic notion.

Definition 7.1. For component k with state space $\widehat{\mathscr{T}}^k$ as part of a composed model consisting of components 1 through K and set of atomic propositions AP, let $\simeq_{AP}$ be defined as the equivalence relation on $\widehat{\mathscr{T}}^k$ such that for all $x^k, y^k \in \widehat{\mathscr{T}}^k$ and all $a \in AP$: $x^k \simeq_{AP} y^k$ iff

$$(x^1, \ldots, x^k, \ldots, x^K) \models a \Leftrightarrow (x^1, \ldots, y^k, \ldots, x^K) \models a \quad \text{for all } x^l \in \widehat{\mathscr{T}}^l,\ \ l \neq k.$$

Note that $\simeq_{AP}$ depends on the embedding realized by the components 1 through K. In many cases, it is possible to define $\simeq$ independent from the environment, e.g., if all propositions from AP are either determined by k or are independent of k. We now define Φ-bisimulation, a somewhat simplified version of F-bisimulation [4]. For CSL-formula Φ, let AP_Φ denote the set of atomic propositions in Φ.

Definition 7.2. A Φ-bisimulation is an equivalence relation Rel on $\widehat{\mathscr{T}}^k$ such that whenever $(x, y) \in$ Rel, then:

$$x \simeq_{AP_\Phi} y \text{ and } \sum_{z\in\mathscr{C}} \mathbf{W}_{k,e}(x, z) = \sum_{z\in\mathscr{C}} \mathbf{W}_{k,e}(y, z)$$

for all $\mathscr{C} \in \widehat{\mathscr{T}}^k/\text{Rel}$ and $e \in \mathscr{E}_{L,k} \cup \mathscr{E}_{S,k}$. States x and y are Φ-bisimilar iff there exists a Φ-bisimulation Rel that contains (x, y).

The largest Φ-bisimulation, denoted by $\sim_\Phi$, can be defined in the standard way [9]. Algorithms for computing this equivalence are given in [11,34]; the algorithm in [11] has been implemented in the APNN toolbox [15].

Let $\tilde{n}_k$ ($\leqslant n_k$) be the number of equivalence classes of $\sim_\Phi$. The Φ-bisimilar quotient of component k has state space $\widehat{\mathscr{T}}^k/\sim_\Phi = \{0, \ldots, \tilde{n}_k-1\}$, and is characterized by matrices $\widetilde{\mathbf{W}}_{k,e} \in \mathbb{R}^{\tilde{n}_k\times\tilde{n}_k}$ defined element-wise by:

$$\widetilde{\mathbf{W}}_{k,e}([x]_{\sim_\Phi}, [y]_{\sim_\Phi}) = \sum_{z\in[y]_{\sim_\Phi}} \mathbf{W}_{k,e}(x, z). \tag{16}$$

Reduced components may be computed for all components. This yields a CTMC described by (3) where matrices $\mathbf{W}_{k,e}$ are replaced by $\widetilde{\mathbf{W}}_{k,e}$. The resulting state space has $\prod_{k=1}^{K} \tilde{n}_k$ instead of $\prod_{k=1}^{K} n_k$ potential reachable states.

The following result follows from the facts that bisimulation preserves the validity of CSL-formulas [5] and is a congruence for the composition operators used here [9]. Let $M/\sim_\Phi$ denote the CTMC obtained from M by replacing each component in M by its Φ-bisimilar equivalent. Then:

Theorem 7.1. *For all $x \in M : x \models_M \Phi$ iff $[x]_{\sim_\Phi} \models_{M/\sim_\Phi} \Phi$.*

Thus, model-checking CSL-formulas can be carried out on the modularly obtained reduced representation of the structured CTMC under consideration.

If several formulas are to be verified, several reduced representations have to be computed. Fortunately, this can often be made more efficient using the following observation.

Corollary 7.1. *For CSL-formulas $\Phi, \Psi : AP_\Phi \subseteq AP_\Psi \Rightarrow \sim_\Phi \subseteq \sim_\Psi$.*

Proof. Since $AP_\Phi \subseteq AP_\Psi$ and the transition weights/rates are the same, every Ψ-bisimulation is also a Φ-bisimulation. $\square$

Assume that formulas $\Phi_1, \ldots, \Phi_M$ have to be checked. The above result suggests to compute $\sim_{\Phi_i}$ before $\sim_{\Phi_j}$ if $AP_{\Phi_i} \subset AP_{\Phi_j}$. $\sim_{\Phi_j}$ can then be computed by refining the equivalence classes of $\sim_{\Phi_i}$. This often makes the bisimulation computations and the generation of reduced components more efficient.

To summarize the results of this section, we present the following algorithm:

```
(1) Input: Structured description with components $1, \dots, K$ and $\mathscr{T}$
(2)       Formulas $\Phi_1, \dots, \Phi_M$
(3)     for ($i = 1$ to $M$) do
(4)       for ($k = 1$ to $K$) do
(5)         compute $\sim_{\Phi_i}$ for component $k$
(6)         generate matrices $\widetilde{\mathbf{W}}_{k,e}$ according to $\sim_{\Phi_i}$ using (16)
(7)       check formula $\Phi_i$ on the reduced CTMC to construct Sat($\Phi_i$)
```

To model-check formulas in the reduced model, the structured algorithms of Section 5 are used that exploit the reduced Kronecker representation. The reduction factor depends on both the structure of the components and the formulas.

A different strategy would be to exploit bisimulations among different components rather than on individual components (as we did here). For instance, it is well known that for identical components which are symmetrically connected, bisimulations can be generated by abstracting from the order of the components. Such approach can be integrated with the structured analysis approach (see [10]).

8. An application example

We illustrate our approach by model-checking a cluster of workstations consisting of two sub-clusters connected via a backbone connection (taken from [29]), and model-checking a simple tandem QN (taken from [32]). All reported experiments are carried out on a Sun Enterprise 250 with 400 MHz CPU and 2 Gb main memory running Solaris 5.7. The reported execution times are all wall clock times.

8.1. Workstation cluster

Each sub-cluster consists of N workstations, connected in a star topology with a central switch that provides the interface to the backbone. Each system of the component (workstations, switches, and backbone) can break down. There is single repair unit that takes care of repairing failed components. The computing power of the cluster is over-dimensioned, in order to be able to accommodate varying levels of traffic volume, as well as to cope with component failures. The system operation is subject to the following informal constraints [29]:

- In order to provide *minimum* quality of service (QoS), at least k ($k < N$) workstations have to be operational, and these workstations have to be connected to each other via operational switches. If in each sub-cluster the number of operational workstations drops below k then an operational backbone is required to ensure that in total at least k operational workstations are still connected to provide minimum service.
- *Premium* quality of service requires at least N operational workstations, with the same connectivity constraints as mentioned above.

The CTMC of the system model is obtained from the generalized SPN (GSPN) depicted in Fig. 2 using similar assumptions as in [29], so the immediate transitions are modeled by timed events with short delays in order to obtain CTMCs of same size. For simplicity, we assume that transitions rates are not state-dependent. The first two "rows" in the GSPN represent the two groups of workstations. In each group, individual workstations can fail (transition WorkstationFail). Once failed, the workstation is Down, and the

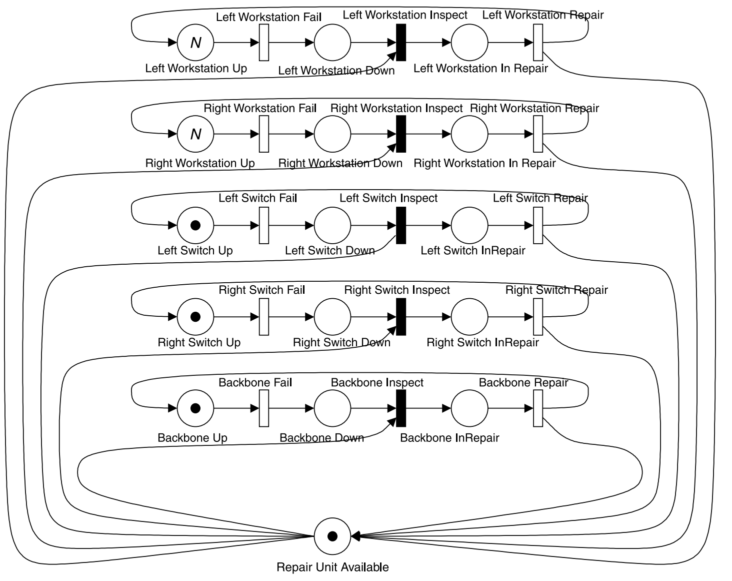

Fig. 2 GSPN model of the workstation cluster (taken from [29]).

Table 2
Abbreviations of basic formulas for the workstation cluster

LeftOperational$_i$	(#LeftWorkstationUp $\geqslant i$) $\wedge$ (#LeftSwitchUp > 0)
RightOperational$_i$	(#RightWorkstationUp $\geqslant i$) $\wedge$ (#RightSwitchUp > 0)
Conn	(#LeftSwitchUp > 0) $\wedge$ (#BackboneUp > 0) $\wedge$(#RightSwitchUp > 0)
Operational$_i$	(#LeftWorkstationUp $+$ #RightWorkstationUp $\geqslant i$)$\wedge$ *Conn*
Minimum	*LeftOperational$_k$* $\vee$ *RightOperational$_k$* $\vee$ *Operational$_k$*
Premium	*LeftOperational$_N$* $\vee$ *RightOperational$_N$* $\vee$ *Operational$_N$*

repair unit is needed to repair the workstation. The repair unit is depicted at the bottom of the figure. If a repair unit (token) is Available, repair starts almost immediately (transition WorkstationInspect), the workstation is InRepair. Once repaired successfully (transition WorkstationRepair), the workstation is Up again. The evolution of the other components of the system is very similar. The next two “rows” in Fig. 2 represent the behavior of the two switches, and the last row represents the backbone.

The CSL-formulas for the workstation cluster example are constructed from atomic propositions that range over the individual markings of the place of the GSPN. For place P and natural number n, proposition #P$\trianglelefteq n$ is valid if and only if place P contains $\trianglelefteq n$ tokens. Appropriate abbreviations for the case study at hand are listed in Table 2.

8.2. *Experimental results for the workstation cluster*

We decompose the model into three components: the left workstations, the right workstations, and the remaining parts (switches, backbone and repair unit). We consider the

Table 3
Experimental data for the workstation cluster

N	$\lvert\mathscr{T}\rvert$	$\otimes$	Ω_1		Ω_2		Ω_3	
			$RS_{\mathscr{S}_1}(\ )$	Time	$RS_{\mathscr{S}_1}(\ )$	Time	$RS_{\mathscr{S}_1}(\ )$	Time
32	38,676	67.0	10.5	0.02	3.1	0.03	30.3	0.02
256	2,373,652	104.5	78.6	2.61	20.1	3.57	1385.1	2.43
512	9,465,876	190.5	156.4	10.63	39.6	14.7	5390.9	13.61
1024	37,806,100	362.5	312.1	43.55	78.5	59.90	21181.2	58.42

model for increasing numbers of N (up to 1024) workstations. Note that only the size of the state space of the first two components increases in dimension while the rest remains constant. For analysis, the Kronecker representation and the set of reachable states $\mathscr{T}$ are generated in a pre-processing step. We apply the technique of [16]: first, state spaces of components are generated in isolation to achieve a Kronecker representation of $\mathbf{Q}$ (as given by Eq. (3)). Then, for each component, the reachability graph is minimized using an inverse bisimulation [12]. The set of reachable states is generated on the minimized components. For this model this procedure results in an aggregated overall state space with three states for all considered values of N, i.e., there are three equivalence classes reachable in the aggregated model. Inverse bisimulations are equivalence relations that preserve reachability. This means that from the set of reachable equivalence classes in the minimized model, the set of reachable states can be easily recreated by a disaggregation step where a reachable aggregated state (i.e., a reachable equivalence class) is substituted by all detailed states belonging to the equivalence class. After a dis-aggregation step of its DAG representation, we achieve a DAG of $RS_{\mathscr{T}}(\ )$. The generation of a Kronecker representation and $RS_{\mathscr{T}}(\ )$ takes less than 10 s for all considered values of N. The first three columns of Table 3 list the values considered for N, the size of the state space, and the size of the Kronecker representation (in Kb), respectively. We model-check $\Omega_s = \mathscr{S}_{\geqslant 0.7}(\mathit{Premium})$ and three different formulas of type $\mathscr{P}_{\trianglelefteq p}(\Phi\,\mathscr{U}^I\,\Psi)$. We focus on the latter since these formulas are most complex. Let $k = 3$.

As the underlying CTMC of the workstation cluster is strongly connected, model-checking Ω_s reduces to the computation of a stationary distribution for a single BSCC $B = \mathscr{T}$. This is performed by a Jacobi iteration (with a relaxation factor of 0.9) for a model with $N = 1024$. The iteration vector requires 302.4 Mb, i.e., space consumption for iteration vectors clearly dominates memory requirements for the CTMC. A single iteration step takes 113.2 s, resulting in a total time of 100,985 s to converge to a maximum residual of 1.13×10^{-5} in 890 iteration steps.

We consider $\Omega_1 = \mathscr{P}_{<0.1}(\mathrm{tt}\,\mathscr{U}^I \mathit{Minimum})$ with $I = [0, 85]$ to allow for a direct comparison with results presented in [29] and a corresponding case study with PRISM.[1] It asserts that the chance that QoS drops below minimum quality within 85 time units is less than 10%. To model-check this formula, one randomization is required (case $I = [0, t]$) with an iterative procedure that considers states in $\mathscr{S}_1 = \mathscr{T} \setminus \mathrm{Sat}(\neg \mathit{Minimum})$ since $\Phi = \mathrm{tt}$ for all states $s \in \mathscr{T}$. The fourth and fifth column of Table 3 list the space requirements (to represent set $\mathscr{S}_1$ for Ω_1; in Kb) and the time (in s) per iteration for verifying Ω_1.

[1] http://www.cs.bham.ac.uk/~dxp/prism/cluster.html/.

Formula $\Omega_2 = \neg \mathit{Minimum} \Rightarrow \mathscr{P}_{<0.3}(\mathrm{tt}\,\mathscr{U}^I\,\neg \mathit{Minimum})$ for $I = [2, 2]$ is build upon a logical implication that involves the solution of a formula $\mathscr{P}_{\trianglelefteq p}(\Phi\,\mathscr{U}^I\,\Psi)$ as an intermediate result. It is an interval of kind $I = [t, t]$ with $t = 2$, where the iteration considers a set $\mathscr{S}_1 = \mathscr{T}$ as $\Phi = \mathrm{tt}$ for all states $s \in \mathscr{T}$. Since $\mathscr{S}_1$ is slightly larger for Ω_2 than for Ω_1, computation times increase but remain within the same order of magnitude (cf. Table 3, sixth and seventh column).

Finally, we consider formula $\Omega_3 = \mathscr{P}_{>0.8}(\mathit{Minimum}\,\mathscr{U}^I\,\mathit{Premium})$, where $I = [t, t'] = [20, 40]$. (These formulas can neither be checked with the current implementation of ETMCC [32] nor with PRISM [41].) This formula requires a double randomization: one for the interval $[0, t'-t] = [0, 20]$ and another for the interval $[0, t] = [0, 20]$. Randomization in both phases is carried out with different sets $\mathscr{S}_1$ and $\mathscr{S}_3$, e.g., for $N = 1024$ we need 1029 steps for the first phase and 895 steps for the second phase. The entire computation requires 112,757 s. In the second phase, the iteration does not reach a fix-point but stops since the resulting Boolean values are determined for all states.

8.3. A simple tandem queue

This example is taken from [32]. It is a simple tandem queuing network of a $M/\mathrm{Cox}_2/1$ queue and a $M/M/1$ queue which both have finite capacity c, such that a blocking phenomenon can be observed. We use this model to illustrate the effect of the different termination criteria we employ. Table 4 gives the number of iteration steps to evaluate $\mathscr{P}_{>p}(\mathrm{tt}\,\mathscr{U}^I\,\mathit{full})$ with $I = [0, t]$ for different values of p and t but fixed capacity $c = 255$ which implies $|\mathscr{T}| = 130, 816$. The atomic proposition *full* identifies situations in which both queues reach their capacities. The iterative computation of vectors $x^{(i)}$ reaches a fix-point

Table 4
Tandem QN with blocking

p	t	# steps	
		(11)	(12)
0.1	5	173	228
0.1	10	315	394
0.1	15	375	552
0.1	20	492	706
0.3	5	165	228
0.3	10	317	394
0.3	15	389	552
0.3	20	492	706
0.7	5	169	228
0.7	10	318	394
0.7	15	413	552
0.7	20	526	706
0.9	5	174	228
0.9	10	332	394
0.9	15	481	552
0.9	20	526	706

Table 5
Tandem QN with blocking

t	# steps	Time per step	Total time
5	168	41.62	7217
10	313	44.80	14,236
15	462	41.85	19,552
20	607	42.64	26,100
25	753	41.86	31,737
30	881	41.96	37,205

after 526 steps independently from values of p and t if we use the implementation of (11). We exercised the tandem model for different values of p and different time horizons t and both implementations of (11) and (12). Results are given in Table 4, where columns # steps either refer to the implementation of (11) or (12) respectively. In this model, the termination criterion appears to be rather insensitive to the selection of p but very sensitive to the selection of t in case of (11) and neither sensitive to p nor t for (12). If $t > 20$ the mode of the Poisson distribution exceeds 526 and the fix-point termination criterion turns out to be the most effective in case of (11), for small values of t the decision criterion is effective and causes termination much ahead of the right truncation point of Fox and Glynn [26]. It is interesting to see, that for this example the number of steps for the iteration according to (11) is significantly less than for (12) which is more efficient per step. In summary, we see that the variety of different termination criteria is useful and it is not clear in general whether (11) or (12) is more efficient in terms of computation time, as far as space is concerned, (12) is clearly superior to (11) since it uses one vector less.

Finally, we consider the tandem network for a large configuration with $c = 4095$ to achieve results comparable to [36]. The time to generate the Kronecker representation and $RS_{\mathscr{T}}(\ )$ is 242 s; the bisimulation approach is not applied due to the size of component state spaces, i.e. the first queue results in a state space with 8191 states, the second has 4096 states. The whole model has $|\mathscr{T}| = 33{,}550{,}336$ states. We compute a solution of formula $\Omega = \mathscr{P}_{>0.5}(\mathrm{tt}\,\mathscr{U}^{I}\ \mathit{full})$ for interval $I = [0, t]$ with different values of t. Table 5 give results obtained over a range of parameter values t (column 1). Column 2 gives the number of iteration steps, column 3 the time per step in seconds, and the last column gives the total time in seconds using the implementation of (11).

The DAG $RS_{\mathscr{S}_1}(\ )$ encodes 33,542,144 states, consists of 2 nodes and uses 163,844 bytes. The randomization procedure requires several vectors, where each uses 268,337,152 bytes (double precision), which creates the bottleneck for this approach in terms of space. For the considered values of t, termination happens in all cases due to the fact that intermediate results are sufficient to determine the formula. This kind of termination is effective in case of relatively small time horizons or a slow rate of convergence.

9. Conclusions

This paper considers a new approach for model-checking CSL formulas based on a combination of CSL model-checking algorithms and structured analysis approaches for large Markov chains. It is shown that with this combination CSL model checking can be applied for fairly large models which are structured as the parallel composition of interacting components. We presented in detail the evaluation of $\mathscr{P}_{\trianglelefteq p}(\Phi\,\mathscr{U}^I\,\Psi)$ in structured models which is the most complex formula. Furthermore, we show that due to the compositional model structured state space reduction by minimization of components using bisimulation and the elimination of vanishing states can be performed very efficiently.

Since the presented algorithms have been implemented and integrated in the APNN toolbox, they are ready to be used for realistic examples. First results are encouraging, but there is still the need to perform more experiments to compare the different realizations of CSL model-checking algorithms that have been proposed.

Appendix A. Kronecker operations for matrices

Kronecker products and sums are widely used in different branches of mathematics and system analysis. Their application for the analysis of stochastic automata dates back to Plateau [42]. This appendix gives a brief overview of the operations.

Definition 9.1. The Kronecker product of two matrices $\mathbf{A} \in \mathbb{R}^{r_1 \times c_1}$ and $\mathbf{B} \in \mathbb{R}^{r_2 \times c_2}$ is defined as $\mathbf{C} = \mathbf{A} \otimes \mathbf{B}$, $\mathbf{C} \in \mathbb{R}^{r_1 r_2 \times c_1 c_2}$, where

$$\mathbf{C}(i_1 r_2 + i_2, j_1 c_2 + j_2) = \mathbf{A}(i_1, j_1)\mathbf{B}(i_2, j_2)$$

$(0 \leqslant i_x < r_x,\ 0 \leqslant j_x < c_x,\ x \in \{1, 2\})$.

The Kronecker sum of two matrices $\mathbf{A} \in \mathbb{R}^{n_1 \times n_1}$ and $\mathbf{B} \in \mathbb{R}^{n_2 \times n_2}$ is defined as

$$\mathbf{A} \oplus \mathbf{B} = \mathbf{A} \otimes \mathbf{I}_{n_2} + \mathbf{I}_{n_1} \otimes \mathbf{B}.$$

For Kronecker products and sums the following properties hold if the ordinary matrix products are defined (see [42,43]).

1. $(\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C} = \mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C})$ and $(\mathbf{A} \oplus \mathbf{B}) \oplus \mathbf{C} = \mathbf{A} \oplus (\mathbf{B} \oplus \mathbf{C})$ associativity,
2. $(\mathbf{AB}) \otimes (\mathbf{CD}) = (\mathbf{A} \otimes \mathbf{C})(\mathbf{B} \otimes \mathbf{D})$ compatibility with multiplication,
3. $(\mathbf{A} + \mathbf{B}) \otimes (\mathbf{C} + \mathbf{D}) = \mathbf{A} \otimes \mathbf{C} + \mathbf{A} \otimes \mathbf{D} + \mathbf{B} \otimes \mathbf{C} + \mathbf{B} \otimes \mathbf{D}$ distributivity over addition,
4. $(\mathbf{A} \otimes \mathbf{B})^{\mathrm{T}} = \mathbf{A}^{\mathrm{T}} \otimes \mathbf{B}^{\mathrm{T}}$ compatibility with transposition,
5. $(\mathbf{A} \otimes \mathbf{B})^{-1} = \mathbf{A}^{-1} \otimes \mathbf{B}^{-1}$ compatibility with inversion.

Since Kronecker products and sums are associative, the following generalisations are well-defined.

$$\bigotimes_{j=1}^{J} \mathbf{A}_j = \mathbf{A}_1 \otimes \bigotimes_{j=2}^{J} \mathbf{A}_j \quad \text{and} \quad \bigotimes_{j=1}^{J} \mathbf{A}_j = \mathbf{A}_1 \oplus \bigoplus_{j=2}^{J} \mathbf{A}_j.$$

Generalised Kronecker products and sums can be transformed into some standardized form applying the above properties several times.

$$\bigotimes_{j=1}^{J} \mathbf{A}_j = \prod_{j=1}^{J} \mathbf{I}_{l_j} \otimes \mathbf{A}_j \otimes \mathbf{I}_{u_j} \quad \text{and} \quad \bigoplus_{j=1}^{J} \mathbf{A}_j = \sum_{j=1}^{J} \mathbf{I}_{l_j} \otimes \mathbf{A}_j \otimes \mathbf{I}_{u_j},$$

where $\mathbf{A}_j \in \mathbb{R}^{r_j, c_j}$, $l_j = \prod_{i=1}^{j-1} c_i$ and $u_j = \prod_{i=j+1}^{J} r_i$.

References

[1] A. Aziz, K. Sanwal, V. Singhal, R. Brayton, Verifying continuous time Markov chains, in: R. Alur, T.A. Henzinger (Eds.), Computer-Aided Verification, LNCS 1102, Springer-Verlag, Berlin, 1996, pp. 269–276.

[2] A. Aziz, K. Sanwal, V. Singhal, R. Brayton, Model checking continuous time Markov chains, ACM Trans. Comput. Logic 1 (1) (2000) 162–170.

[3] I. Bahar, E. Frohm, C. Gaona, G. Hachtel, E. Macii, A. Pardo, F. Somenzi, Algebraic decision diagrams and their applications, Formal Meth. Syst. Des. 10 (2/3) (1997) 171–206.

[4] C. Baier, B.R. Haverkort, H. Hermanns, J.-P. Katoen, On the logical characterisation of performability properties, in: U. Montanari, J.D.P. Rolim, E. Welzl (Eds.), Automata, Languages, and Programming (ICALP) LNCS 1853, Springer-Verlag, Berlin, 2000, pp. 780–792.

[5] C. Baier, B.R. Haverkort, H. Hermanns, J.-P. Katoen, Model checking continuous-time Markov chains by transient analysis, in: E.A. Emerson, A.P. Sistla (Eds.), Computer-Aided Verification, LNCS 1855, Springer-Verlag, Berlin, 2000, pp. 358–372.

[6] C. Baier, J.-P. Katoen, H. Hermanns, Approximate symbolic model checking of continuous-time Markov chains, in: J.C.M. Baeten, S. Mauw (Eds.), Concurrency Theory, LNCS 1664, Springer-Verlag, Berlin, 1999, pp. 146–162.

[7] R. Bryant, Graph-based algorithms for Boolean function manipulation, IEEE Trans. Comput. 35 (8) (1986) 677–691.

[8] P. Buchholz, Die Strukturierte Analyse Markovscher Modelle, IFB 282, Springer, Berlin, 1991.

[9] P. Buchholz, Markovian process algebra: composition and equivalence, in: U. Herzog, M. Rettelbach (Eds.), Proc. of the 2nd Work. on Process Algebras and Perf. Modelling, vol. 27, Arbeitsberichte des IMMD, University of Erlangen, 1994, pp. 11–30.

[10] P. Buchholz, Hierarchical Markovian models: symmetries and reduction, Perform. Eval. 22 (1995) 93–110.

[11] P. Buchholz, Efficient computation of equivalent and reduced representations for stochastic automata, Int. J. Comput. Syst. Sci. Eng. 15 (2) (2000) 93–103.

[12] P. Buchholz, Hierarchical structuring of superposed GSPNs, IEEE Trans. Softw. Eng. 25 (2) (1999) 166–181.

[13] P. Buchholz, Structured analysis approaches for large Markov chains, Appl. Numer. Math. 31 (4) (1999) 375–404.

[14] P. Buchholz, G. Ciardo, S. Donatelli, P. Kemper, Complexity of Kronecker operations and sparse matrices with applications to the solution of Markov models, INFORMS J. Comput. 12 (3) (2000) 203–222.

[15] P. Buchholz, M. Fischer, P. Kemper, C. Tepper, New features in the APNN toolbox, in: P. Kemper (Ed.), Tools of Aachen 2001 Int. Multiconference on Measurement, Modeling and Evaluation of Computer-Communication Systems, Universität Dortmund, Fachbereich Informatik, Forschungsbericht Nr. 760, 2001, pp. 62–68.

[16] P. Buchholz, P. Kemper, Efficient computation and representation of large reachability sets for composed automata, Discrete Event Dyn. Syst.—Theory Applic. 12 (3) (2002) 265–286.

[17] J. Campos, M. Silva, S. Donatelli, Structured solution of asynchronously communicating stochastic modules, IEEE Trans. Softw. Eng. 25 (2) (1999) 147–165.

[18] G. Chiola, M. Ajmone-Marsan, G. Balbo, G. Conte, Generalized stochastic Petri nets: a definition at the net level and its implications, IEEE Trans. Softw. Eng. 19 (2) (1993) 89–107.

[19] G. Ciardo, Distributed and structured analysis approaches to study large and complex systems, in: E. Brinksma, H. Hermanns, J.-P. Katoen (Eds.), Lectures on Formal Methods and Performance Analysis, LNCS 2090, Springer-Verlag, Berlin, 2001, pp. 344–374.

[20] G. Ciardo, A.S. Miner, A data structure for the efficient Kronecker solution of GSPNs, in: R. Buchholz, M. Silva (Eds.), Proc. 8th Int. Workshop on Petri Nets and Perf. Models, IEEE CS-Press, 1999, pp. 22–31.

[21] E. Clarke, E. Emerson, A. Sistla, Automatic verification of finite-state concurrent systems using temporal logic specifications, ACM Trans. Program. Languages Syst. 8 (1986) 244–263.

[22] E. Clarke, M. Fujita, P.C. McGeer, J.C.-Y. Yang, Multi-terminal binary decision diagrams: an efficient data structure for matrix representation, Formal Meth. Syst. Des. 10 (2/3) (1997) 149–169.

[23] E.M. Clarke, O. Grumberg, D.A. Peled, Model Checking, MIT Press, 1999.

[24] S. Donatelli, Superposed stochastic automata: a class of stochastic Petri nets amenable to parallel solution, Perform. Eval. 18 (1994) 21–36.

[25] S. Donatelli, P. Kemper, Integrating synchronization with priority into a Kronecker representation, Perform. Eval. 44 (1–4) (2001) 73–96.

[26] B.L. Fox, P.W. Glynn, Computing Poisson probabilities, Commun. ACM 31 (1988) 440–445.

[27] D. Gross, D. Miller, The randomization technique as a modeling tool and solution procedure for transient Markov processes, Oper. Res. 32 (2) (1984) 926–944.

[28] H. Hansson, B. Jonsson, A logic for reasoning about time and reliability, Formal Aspects Comput. 6 (1994) 512–535.

[29] B.R. Haverkort, H. Hermanns, J.-P. Katoen, The use of model checking techniques for quantitative dependability evaluation, IEEE Symp. on Reliable Distr. Sys., IEEE CS Press, 2000, pp. 228–238.

[30] B.R. Haverkort, K.S. Trivedi, Specification techniques for Markov reward models, Discrete Event Syst.: Theory Applic. 3 (1993) 219–247.

[31] H. Hermanns, U. Herzog, J.-P. Katoen, Process algebra for performance evaluation, Theor. Comput. Sci. 274 (2002) 43–87.

[32] H. Hermanns, J.-P. Katoen, J. Meyer-Kayser, M. Siegle, A Markov chain model checker, in: S. Graf, M. Schwartzbach (Eds.), Tools and Algorithms for the Construction and Analysis of Systems, LNCS 1785, Springer-Verlag, Berlin, 2000, pp. 347–362.

[33] H. Hermanns, M. Rettelbach, T. Weiss, Formal characterisation of immediate actions in SPA with nondeterministic branching, The Comput. J. 38 (7) (1995) 530–542.

[34] H. Hermanns, M. Siegle, Bisimulation algorithms for stochastic process algebras and their BDD-based implementation, in: J.-P. Katoen (Ed.), Formal Methods for Real-Time and Probabilistic Systems, LNCS 1601, Springer-Verlag, Berlin, 1999, pp. 244–265.

[35] J. Hillston, A Compositional Approach to Performance Modelling, Cambridge University Press, 1996.

[36] J.-P. Katoen, M. Kwiatkowska, G. Norman, D. Parker, Faster and symbolic CTMC model checking, in: L. de Alfaro, S. Gilmore (Eds.), Process Algebra and Probabilistic Methods, LNCS 2165, Springer-Verlag, Berlin, 2001, pp. 23–38.

[37] P. Kemper, Reachability analysis based on structured representations, in: J. Billington, W. Reisig (Eds.), Application and Theory of Petri Nets 1996, LNCS 1091, Springer-Verlag, Berlin, 1996, pp. 269–288.

[38] P. Kemper, Numerical analysis of superposed GSPNs, IEEE Trans. Softw. Eng. 22 (9) (1996) 615–628.

[39] P. Kemper, Transient analysis of superposed GSPNs, IEEE Trans. Softw. Eng. 25 (2) (1999) 182–193.

[40] P. Kemper, R. Lübeck, Model checking based on Kronecker algebra, Technical Report 669, Fachbereich Informatik, Universität Dortmund (Germany), 1998.

[41] M. Kwiatkowska, G. Norman, D. Parker, Probabilistic symbolic model checking with PRISM: a hybrid approach, in: J.-P. Katoen, P. Stevens (Eds.), Tools and Algorithms for the Construction and Analysis of Systems, LNCS 2280, Springer-Verlag, Berlin, 2002, pp. 52–67.

[42] B. Plateau, On the stochastic structure of parallelism and synchronisation models for distributed algorithms, Perform. Eval. Rev. 13 (1985) 142–154.

[43] W.J. Stewart, Introduction to the Numerical Solution of Markov Chains, Princeton University Press, 1994.

[44] R. Tarjan, Depth-first search and linear graph algorithms, SIAM. J. Comput. 1 (1972) 146–160.