# Bisimulations and Logical Characterizations on Continuous-time Markov Decision Processes

Lei Song[1], Lijun Zhang[2], and Jens Chr. Godskesen[3]

[1] Max-Planck-Institut für Informatik and Saarland University, Saarbrücken, Germany
[2] State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Sciences
[3] Programming, Logic, and Semantics Group, IT University of Copenhagen, Denmark

**Abstract.** In this paper we study strong and weak bisimulation equivalences for continuous-time Markov decision processes (CTMDPs) and the logical characterizations of these relations with respect to the continuous-time stochastic logic (CSL). For strong bisimulation, it is well known that it is strictly finer than CSL equivalence. In this paper we propose strong and weak bisimulations for CTMDPs and show that for a subclass of CTMDPs, strong and weak bisimulations are both sound and complete with respect to the equivalences induced by CSL and the sub-logic of CSL without next operator respectively. We then consider a standard extension of CSL, and show that it and its sub-logic without X can be fully characterized by strong and weak bisimulations respectively over arbitrary CTMDPs.

## 1 Introduction

Recently, continuous-time Markov decision processes (CTMDPs) have received extensive attention in the model checking community, see for example [5, 39, 27, 28, 12, 31]. Analysis techniques for CTMDPs suffer especially from the state space explosion problem. Thus, as for other stochastic models, bisimulation relations have been proposed for CTMDPs. In [27], strong bisimulation was shown to be sound with respect to the continuous-time stochastic logic [2] (CSL). This result guarantees that one can first reduce a CTMDP up to bisimulation equivalence before analysing it. On the other hand, as indicated in [27], strong bisimulation is not complete with respect to CSL, i.e., logically equivalent states might be not bisimilar.

CTMDPs extend Markov decision processes (MDPs) with exponential sojourn time distributions, and subsume models such as labelled transition systems and Markov chains as well. While linear and branching time equivalences have been studied for these sub-models [38, 37, 6, 34], we extend these results to the setting of CTMDPs. In this paper we study strong and weak bisimulation relations for CTMDPs, and the logical characterization problem of these relations with respect to CSL and its sub-logics.

We start with a slightly coarser notion of strong bisimulation than the one in [27], and then propose *weak bisimulation* for CTMDPs. We study the relationship between strong and weak bisimulations and the logical equivalences induced by CSL and CSL$_{\setminus X}$ – the sub-logic of CSL without next operators. Our first contribution is to identify a subclass of CTMDPs under which our strong and weak bisimulations coincide with CSL

and $CSL_{\backslash X}$ equivalences respectively. We discuss then how this class of CTMDPs can be efficiently determined, and moreover, we argue that most models arising in practice are among this class.

As for labelled transition systems and MDPs, we also define an extension of CSL, called $CSL^*$, which is more distinguishable than CSL. Surprisingly, $CSL^*$ is able to fully characterize strong bisimulation over arbitrary CTMDPs, similarly for the sub-logic without next operator and weak bisimulation.

Since CTMDPs can be seen as models combining MDPs and continuous-time Markov Chains (CTMCs), we will discuss the downward compatibility of the relations with those for MDPs [32] and CTMCs in [6]. Summarizing, the paper contains the following contributions:

1. We extend strong probabilistic bisimulation defined in [32] over probabilistic automata to CTMDPs, and then prove that it coincides with CSL equivalence for a subclass of CTMDPs;
2. We propose a scheme to determine the subclass of CTMDPs efficiently, and show that many models in practice are in this subclass;
3. We introduce a new notion of weak bisimulation for CTMDPs, and show its characterization results with respect to $CSL_{\backslash X}$;
4. We present a standard extension of CSL that is shown to be both sound and complete with respect to strong and weak bisimulations for arbitrary CTMDPs.

*Related work.* Logical characterizations of bisimulations have been studied extensively for stochastic models. For CTMCs, CSL characterizes strong bisimulation, while CSL without next operator characterizes weak bisimulation [6]. Our results in this paper are conservative extensions for both strong and weak bisimulations from CTMCs to CTMDPs. In [18], the results are extended to CTMCs with continuous state spaces.

For CTMDPs, the first logical characterization result is presented in [27]. It is shown that strong bisimulation is sound, but not complete with respect to CSL equivalence. In this paper, we introduce strong and weak bisimulation relations for CTMDPs. For a subclass of CTMDPs, i.e., those without *2-step recurrent* states, we show that they are also complete for CSL and $CSL_{\backslash X}$ equivalences respectively.

For probabilistic automata (PAs), Hennessy-Milner logic has been extended to characterize bisimulations in [23, 15, 21]. In [17], Desharnais *et al.* have shown that weak bisimulation agrees with $PCTL^*$ equivalence for alternative PAs. Another related paper for PAs is our previous paper [34], in which we have introduced $i$-depth bisimulations to characterize logical equivalences induced by $PCTL^*$ and its sub-logics.

All proofs are found in the full version of this paper [35].

*Organization of the paper.* Section 2 recalls the definition of CTMDPs and the logic CSL. Variants of bisimulation relations and their corresponding logical characterization results are studied in Section 3. In Section 4 we present the extension of CSL that fully characterizes strong and weak bisimulations. We discuss in Section 5 related work with MDPs and CTMCs. Section 6 concludes the paper.

## 2   Preliminaries

For a finite set $S$, a distribution is a function $\mu : S \to [0,1]$ satisfying $|\mu| := \sum_{s \in S} \mu(s) = 1$. We denote by $Dist(S)$ the set of distributions over $S$. We shall use $s, r, t, \ldots$ and $\mu, \nu \ldots$ to range over $S$ and $Dist(S)$, respectively. The support of $\mu$ is defined by $Supp(\mu) = \{s \in S \mid \mu(s) > 0\}$. Given a finite set of non-negative real numbers $\{p_j\}_{j \in J}$ and distributions $\{\mu_j\}_{j \in J}$ such that $\sum_{j \in J} p_i = 1$ for each $j \in J$, $\sum_{j \in J} p_j \cdot \mu_j$ is the distribution such that $(\sum_{j \in J} p_j \cdot \mu_j)(s) = \sum_{j \in J} p_j \cdot \mu_j(s)$ for each $s \in S$. For an equivalence relation $\mathcal{R}$ over $S$, we write $\mu \; \mathcal{R} \; \nu$ if it holds that $\mu(C) = \nu(C)$ for all equivalence classes $C \in S/\mathcal{R}$ where $\mu(C) = \sum_{s \in C} \mu(s)$, and moreover $[s]_{\mathcal{R}} = \{r \mid s \; \mathcal{R} \; r\}$ is the equivalence class of $S/\mathcal{R}$ containing $s$. The subscript $\mathcal{R}$ will be omitted if it is clear from the context. A distribution $\mu$ is called *Dirac* if $|Supp(\mu)| = 1$, and we let $\mathcal{D}_s$ denote the Dirac distribution such that $\mathcal{D}_s(s) = 1$. We let $R^{\geq 0}$ and $R^{>0}$ denote the set of non-negative and positive real numbers respectively.

### 2.1   Continuous-time Markov Decision Processes

Below follows the definition of CTMDPs, which subsume both MDPs and CTMCs.

**Definition 1 (Continuous-time Markov Decision Processes).** *A tuple* $\mathcal{C} = (S, \to, AP, L, s_0)$ *is a* CTMDP *where* $s_0 \in S$ *is the initial state,* $S$ *is a finite but non-empty set of states,* $AP$ *is a finite set of atomic propositions,* $L : S \mapsto 2^{AP}$ *is a labelling function, and* $\to \subseteq S \times R^{>0} \times Dist(S)$ *is a finite transition relation such that for each* $s \in S$, *there exists* $\lambda$ *and* $\mu$ *with* $(s, \lambda, \mu) \in \to$.

From Definition 1 we can see that there are both non-deterministic and probabilistic transitions in a CTMDP. We write $s \xrightarrow{\lambda} \mu$ if $(s, \lambda, \mu) \in \to$, where $\lambda$ is called exit rate of the transition. Let $Suc(s) = \{r \mid \exists(s \xrightarrow{\lambda} \mu).\mu(r) > 0\}$ denote the successor states of $s$, and let $Suc^*(s)$ be its transitive closure. A state $s$ is said to be *silent* iff for all $s_1, s_2 \in Suc^*(s)$, $L(s_1) = L(s_2)$ and $s_1 \xrightarrow{\lambda} \mu_1$ implies $s_2 \xrightarrow{\lambda} \mu_2$. Intuitively, a state $s$ is silent if all its reachable states have the same labels as $s$. In addition, they have transitions with the same exit rates as transitions of $s$. States like $s$ are called silent, since it is not distinguishable from all its successors, either by labels or sojourn time of states. Therefore a silent state $s$ and all its successors can be represented by a single state which is the same as $s$ but with all its outgoing transitions leading to itself. A CTMC is a deterministic CTMDP satisfying the condition: $s \xrightarrow{\lambda} \mu$ and $s \xrightarrow{\lambda'} \mu'$ imply $\lambda = \lambda'$ and $\mu = \mu'$ for any $s \in S$.

### 2.2   Paths, Uniformization, and Measurable Schedulers

Let $\mathcal{C} = (S, \to, AP, L, s_0)$ be a CTMDP fixed for the remainder of the paper. Let $Paths^n(\mathcal{C}) = S \times (R^{>0} \times S)^n$ denote the set containing paths of $\mathcal{C}$ with length $n$. The set of all finite paths of $\mathcal{C}$ is the union of all finite paths $Paths^*(\mathcal{C}) = \cup_{n \geq 0} Paths^n(\mathcal{C})$. Moreover, $Paths^\infty(\mathcal{C}) = S \times (R^{>0} \times S)^\infty$ contains all infinite paths and $Paths(\mathcal{C}) = Paths^*(\mathcal{C}) \cup Paths^\infty(\mathcal{C})$ is the set of all (finite and infinite) paths of $\mathcal{C}$. Intuitively, a path

is comprised of an alternation of states and their sojourn time. To simplify the discussion we introduce some notations. Given a path $\omega = s_0, t_0, s_1, t_1 \cdots s_n \in Paths^n(\mathcal{C})$, $|\omega| = n$ is the length of $\omega$, $\omega \downarrow = s_n$ is the last state of $\omega$, $\omega|^i = s_0, t_0, \cdots, s_i$ is the prefix of $\omega$ ending at the $(i + 1)$-th state, and $\omega|_i = s_i, t_i, s_{i+1}, \cdots$ is the suffix of $\omega$ starting from the $(i + 1)$-th state, and $\omega \frown (t_n, s_{n+1})$ is the path obtained by extending $\omega$ with $(t_n, s_{n+1})$. Let $\omega[i] = s_i$ denote the $(i + 1)$-th state where $i \leq n$ and $time(\omega, i) = t_i$ the sojourn time in the $(i + 1)$-th state with $i < n$. Let $\omega@t$ be the state at time $t$ in $\omega$, that is, $\omega@t = \omega[j]$ where $j$ is the smallest index such that $\sum_{i=0}^{j} t_i > t$. Moreover, $Steps(s) = \{(\lambda, \mu) \mid (s, \lambda, \mu) \in \rightarrow\}$ is the set of all available choices at state $s$. Let $\{I_i \subseteq [0, \infty)\}_{0 \leq i \leq k}$ denote a set of non-empty closed intervals, then $C(s_0, I_0, \cdots, I_k, s_{k+1})$ is the *cylinder set* of paths $\omega \in Paths^\infty(\mathcal{C})$ such that $\omega[i] = s_i$ for $0 \leq i \leq k + 1$ and $time(\omega, i) \in I_i$ for $0 \leq i \leq k$. Let $\mathfrak{F}_{Paths^\infty(\mathcal{C})}$ be the smallest $\sigma$ algebra on $Paths^\infty(\mathcal{C})$ containing all cylinder sets.

As shown in [4], model checking of CTMCs can be reduced to the problem of computing transient state probabilities, which can be solved efficiently, for instance by uniformization. In a uniformized CTMC, all states will evolve at the same speed, i.e., all transitions have the same exit rates. Similarly, we can also define uniformization of a CTMDP by uniformizing the exit rate of all its transitions. Below we recall the notion of *uniformization* for CTMDPs [12, 28].

**Definition 2 (Uniformization).** *Given a* CTMDP $\mathcal{C} = (S, \rightarrow, AP, L, s_0)$, *the uniformized* CTMDP *is denoted as* $\bar{\mathcal{C}} = (\bar{S}, \rightarrow', AP, \bar{L}, \bar{s}_0)$ *where*

1. $\bar{S} = \{\bar{s} \mid s \in S\}$, $\bar{s}_0 \in \bar{S}$ *is the initial state,*
2. $\bar{L}(\bar{s}) = L(s)$ *for each* $s \in S$, *and*
3. $(\bar{s}, E, \bar{\mu}) \in \rightarrow'$ *iff there exists* $(s, \lambda, \mu) \in \rightarrow$ *and* $\bar{\mu} = \frac{\lambda}{E} \cdot \mu' + (1 - \frac{\lambda}{E}) \cdot \mathcal{D}_{\bar{s}}$ *such that* $\mu'(\bar{r}) = \mu(r)$ *for each* $r \in Supp(\mu)$,

*Here* $E$ *is the uniformization rate for* $\bar{\mathcal{C}}$, *which is a real number equal or greater than all the rates appearing in* $\mathcal{C}$.

By uniformization for each transition $(s, \lambda, \mu)$ we add a self loop to $s$ with rate equal to $E$ minus the original rate $\lambda$. After uniformization every state will have a unique exit rate on all its transitions. As we will show later, this transformation will not change the properties we are interested in under certain classes of schedulers.

Due to the existence of non-deterministic choices in CTMDPs, we need to resolve them to define probability measures. As usual, non-deterministic choices in CTMDPs are resolved by schedulers (or policies or adversaries), which generate a distribution over the available transitions based on the given history information. Different classes of schedulers can be defined depending on the information a scheduler can use in order to choose the next transition. However not all of them are suitable for our purposes, which we will explain later. In this paper, we shall focus on one specific class of schedulers, called *measurable total time positional schedulers* (TTP) [28], which is defined as follows:

**Definition 3 (Schedulers).** *A scheduler* $\pi : S \times R^{\geq 0} \times (R^{>0} \times Dist(S)) \mapsto [0, 1]$ *is measurable if* $\pi(s, t, \cdot) \in Dist(Steps(s))$ *for all* $(s, t) \in S \times R^{\geq 0}$ *and* $\pi(\cdot, tr)$ *are measurable for all* $tr \in 2^{(R^{>0} \times Dist(S))}$, *where*

- $\pi(s,t,\cdot)$ *is a distribution such that* $\pi(s,t,\cdot)(\lambda,\mu) = \pi(s,t,\lambda,\mu)$*, and*
- $\pi(\cdot,tr) : (S \times R^{\geq 0}) \mapsto [0,1]$ *is a function such that for each* $(s,t) \in S \times R^{\geq 0}$*, it holds* $\pi(\cdot,tr)(s,t) = \sum_{(\lambda,\mu) \in tr} \pi(s,t,\lambda,\mu)$*.*

The schedulers defined in Definition 3 are total time positional, since they make decisions only based on the current state and total elapsed time, which are the first and second parameters of $\pi$ respectively. The third parameter and fourth parameter of $\pi$ denote the rate and the resulting distribution of the chosen transition respectively. Given the current state $s$, the total elapsed time $t$, and a transition $(\lambda,\mu)$, $\pi$ will return the probability with which $(\lambda,\mu)$ will be chosen. This is a special case of the general definition of schedulers, which can make decisions based on the full history, for instance visited states and the sojourn time at each state. Given a scheduler $\pi$, a unique probability measure $Pr_{\pi,s}$ can be determined on the $\sigma$-algebra $\mathfrak{F}_{Paths^{\infty}(\mathcal{C})}$ inductively as below: $Pr_{\pi,s}(C(s_0, I_0, \cdots, s_n), tt) =$

$$
\begin{cases}
1 & n = 0 \land s = s_0 \quad \text{(1a)} \\
0 & s \neq s_0 \quad \text{(1b)} \\
\int\limits_{t \in I_0} \sum\limits_{(\lambda,\mu) \in tr} \pi(s_0, tt)(\lambda,\mu) \cdot \mu(s_1) \cdot \lambda e^{-\lambda t} \cdot Pr_{\pi,s_1} dt & \text{otherwise} \quad \text{(1c)}
\end{cases}
$$

where $Pr_{\pi,s_1}$ is an abbreviation of $Pr_{\pi,s_1}(C(s_1, \ldots, s_n), tt + t)$, $tr = Steps(s_0)$ and $tt$ is the parameter denoting the total elapsed time. One nice property of TTP schedulers is that uniformization does not change time-bounded reachability under TTP schedulers [28, 31]. This result can be extended to cover more properties like $CSL_{\backslash X}$ and $CSL_{\backslash X}^*$, which shall be introduced soon.

Besides TTP schedulers, there are other different classes of schedulers for CTMDPs, some of which are insensitive to uniformization, whereas some of which may gain or lose information after uniformization, i.e., properties of a CTMDP may be changed by uniformization. To avoid technical overhead in the presentation, we refer to [28] for an in-depth discussion of these different classes of schedulers and their relation to uniformization.

### 2.3   Continuous Stochastic Logic

Logical formulas are important for verification purpose, since they offer a rigorous and unambiguous way to express properties one may want to check. Probabilistic computation tree logic (PCTL) [19] is often used to express properties of probabilistic systems. In order to deal with probabilistic systems with exponential sojourn time distributions like CTMCs and CTMDPs, the continuous stochastic logic (CSL) was introduced to reason about CTMCs [2, 4], and recently extended to reason about CTMDPs in [27]. CSL contains both state[4] and path formulas whose syntax is defined by the following BNFs:

$$\varphi ::= a \mid \neg\varphi \mid \varphi \land \varphi \mid \mathcal{P}_{\bowtie p}(\psi),$$
$$\psi ::= \mathsf{X}^I \varphi \mid \varphi \, \mathsf{U}^I \varphi,$$

---

[4] The steady-state operator is omitted in this paper for simplicity of presentation.

where $a \in AP$, $p \in [0,1]$, $\bowtie \in \{<, \leq, \geq, >\}$, and $I \subseteq [0, \infty)$ is a non-empty closed interval.

We use $s \models \varphi$ to denote that $s$ satisfies the state formula $\varphi$, while $\omega \models \psi$ denotes that $\omega$ satisfies the path formula $\psi$. The satisfaction relation for atomic proposition and Boolean operators is standard. Below we give the satisfaction relation for the remaining state and path formulas:

$$s_0 \models \mathcal{P}_{\bowtie p}(\psi) \text{ iff } \forall \pi . Pr_{\pi, s_0}(\{\omega \in Paths^\infty(\mathcal{C}) \mid \omega \models \psi\}) \bowtie p,$$

$$\omega \models \mathsf{X}^I \varphi \text{ iff } \omega[1] \models \varphi \wedge time(\omega, 0) \in I,$$

$$\omega \models \varphi_1 \, \mathsf{U}^I \, \varphi_2 \text{ iff } \exists i . (\sum_{0 \leq j < i} time(\omega, j) \in I \wedge \omega[i] \models \varphi_2 \wedge (\forall 0 \leq j < i . \omega[j] \models \varphi_1)).$$

Intuitively, a state $s_0$ satisfies $\mathcal{P}_{\bowtie p}(\psi)$ iff no matter how we schedule the transitions of $s_0$ and its successors, the probability of paths starting from $s_0$ and satisfying $\psi$ is always $\bowtie p$. This operator has the same semantics as in PCTL. Compared to PCTL, the main difference arises in the semantics of the path formulas. Given a path $\omega$, we say $\omega \models \mathsf{X}^I \varphi$, iff the second state in $\omega$ satisfies $\varphi$, moreover the sojourn time in the first state of $\omega$ is within the time interval $I$. We say $\omega \models \varphi_1 \, \mathsf{U}^I \, \varphi_2$, iff along $\omega$, a state satisfying $\varphi_2$ can be reached at some time point in $I$, and all the preceding states if any satisfy $\varphi_1$. If all time bounds are defined to be equal to $[0, \infty)$, i.e., removing time restrictions, CSL will degenerate to PCTL.

Different from [4] where the semantics of CSL is continuous, in this paper we consider pointwise semantics of CSL. This is mainly because the semantics of CSL$^*$ introduced in Section 4 is also pointwise. However, results in Section 3 are also valid if we consider continuous semantics.

*Logic Equivalences.* Let $\mathcal{L}$ denote some logic. We say that $s$ and $r$ are $\mathcal{L}$-equivalent, denoted by $s \sim_{\mathcal{L}} r$, if they satisfy the same set of $\mathcal{L}$ state formulas, that is, $s \models \varphi$ iff $r \models \varphi$ for all state formulas $\varphi$ in $\mathcal{L}$, similarly for $\sim_{\mathcal{L}_{\backslash \mathsf{X}}}$, where $\mathcal{L}_{\backslash \mathsf{X}}$ denotes the sub-logic of $\mathcal{L}$ without the $\mathsf{X}^I$ operator. In this paper, $\mathcal{L}$ will denote either CSL or CSL$^*$, which we shall introduce in Section 4.

## 3   Bisimilarity and CSL Equivalence

In this section, we first introduce the concept of strong bisimulation for CTMDPs, which can be seen as a variant of strong bisimulation for MDPs. Then we define a sub-class of CTMDPs, called non 2-step recurrent CTMDPs, and show that strong bisimulation can be fully characterized by CSL for non 2-step recurrent CTMDPs. We extend the work to the weak setting and show similar results for weak bisimulation. Finally, we propose an efficient scheme to determine non 2-step recurrent CTMDPs and we show that almost all CTMDP models in practice fall into this class.

### 3.1   Strong Bisimulation

The definition of strong bisimulation we shall introduce in this section slightly generalizes the one introduced in [27]. The reason is that we adopt the notion of combined

transitions, used in [32] to define *strong probabilistic bisimulation* for PAs. Combined transitions allow transitions induced by convex combinations of several transitions. We shall lift its definition to the setting of CTMDPs. Let $s \xrightarrow{\lambda}_P \mu$ iff there exists $\{s \xrightarrow{\lambda} \mu_j\}_{j \in J}$ and $\{p_j\}_{j \in J}$ such that $\sum_{j \in J} p_j = 1$, and $\sum_{j \in J} p_j \cdot \mu_j = \mu$. The combined transitions of a CTMDP are almost the same as those for PAs except we need to take care of the rate of each transition. Here we only allow to combine transitions with the same rate, otherwise we may change non-trivial properties of a CTMDP, which we will explain soon. Below follows the definition of strong bisimulation:

**Definition 4 (Strong Bisimulation).** *Let $\mathcal{R} \subseteq S \times S$ be an equivalence relation. $\mathcal{R}$ is a strong bisimulation iff $s \mathrel{\mathcal{R}} r$ implies that $L(s) = L(r)$ and for each $s \xrightarrow{\lambda} \mu$, there exists $r \xrightarrow{\lambda}_P \mu'$ such that $\mu \mathrel{\mathcal{R}} \mu'$.*

*We write $s \sim r$ whenever there exists a strong bisimulation $\mathcal{R}$ such that $s \mathrel{\mathcal{R}} r$. Let strong bisimilarity $\sim$ denote the largest strong bisimulation, which is equal to the union of all strong bisimulation relations.*

For $s$ and $r$ to be strong bisimilar, the same set of atomic propositions should hold at $s$ and $r$. Furthermore, $s$ should be able to mimic $r$ stepwise and vice versa, that is, whenever $s$ has a transition with label $\lambda$ leading to a distribution $\mu$, $r$ should also be able to perform a (combined) transition with the same label to a distribution $\nu$ such that $\mu$ and $\nu$ match with each other, i.e., $\mu$ and $\nu$ assign the same probability to each equivalence class $C \in S/\mathcal{R}$. Strong bisimulation defined in Definition 4 is a conservative extension of strong probabilistic bisimulation for PAs defined in [32], in the sense that it coincides with strong probabilistic bisimulation if we replace $\lambda$ with actions.

The relation defined above is slightly coarser than the one considered in [27], where the combined transition $r \xrightarrow{\lambda}_P \mu'$ is replaced by the normal transition $r \xrightarrow{\lambda} \mu'$. In [27], it was also shown that strong bisimulation is only sound but not complete with respect to CSL equivalence. Even though our definition of strong bisimulation is slightly coarser, it is still too fine for CSL equivalence as shown in the following theorem:

**Theorem 1 ([27]).** $\sim \subsetneq \sim_{CSL}$.

The proof in [27] can be directly adapted to prove the soundness of our slightly more general strong bisimulation. The inclusion in Theorem 1 is strict which is illustrated by the following example:

*Example 1.* Suppose we are given two states $s_0$ and $r_0$ of a CTMDP depicted in Fig. 1 (a) and (b) respectively, where all states have different atomic propositions except $L(s_0) = L(r_0)$. Assume $u_i$ are silent for $i = 1, 2, 3$, our aim is to show that $s_0$ and $r_0$ satisfy the same set of CSL formulas, while they are not strong bisimilar by Definition 4.

We first show that $s_0 \sim_{CSL} r_0$, i.e., $s_0 \models \varphi$ implies $r_0 \models \varphi$ for any $\varphi$ and vice versa. The only non-trivial cases are the time-bounded reachabilities from $s_0$ and $r_0$ to states in $C \subseteq \{u_1, u_2, u_3\}$. For instance the maximal probability from $s_0$ and $r_0$ to $\{u_2, u_3\}$ in time interval $[a, b]$ is equal to $0.7 \cdot (e^{-a} - e^{-b})$, irrelevant of the middle transition of $r_0$. Similarly, we can check that for other $C$, the maximal (or minimal) probabilities from $s_0$ and $r_0$ to $C$ in time interval $I$ are all independent from the middle transition of $r_0$. Therefore we conclude that $s_0 \sim_{CSL} r_0$.

Fig. 1. Counterexample of the completeness of strong bisimulation.

Secondly, we show that it does not hold that $s_0 \sim r_0$ according to Definition 4. We prove by contradiction. Assume that there exists a strong bisimulation $\mathcal{R}$ such that $s_0 \mathcal{R} r_0$. By Definition 4, for the middle transition of $r_0$, i.e., $r_0 \xrightarrow{1} \mu'$ where $\mu'(u_1) = 0.4, \mu'(u_2) = 0.3$, and $\mu'(u_3) = 0.3$, we need to find a transition $s_0 \xrightarrow{1}_P \mu$ of $s_0$ such that $\mu \mathcal{R} \mu'$. Since $u_1, u_2$, and $u_3$ have different atomic propositions, $(u_i, u_j) \notin \mathcal{R}$ for any $1 \leq i \neq j \leq 3$. Therefore the only possibility is that $\mu(u_1) = 0.4, \mu(u_2) = 0.3$, and $\mu(u_3) = 0.3$. However that is impossible, such $\mu$ cannot be the resulting distribution of any (combined) transition of $s_0$. Otherwise there would exist $w_1, w_2 > 0$ such that $w_1 + w_2 = 1, 0.3 \cdot w_1 + 0.5 \cdot w_2 = 0.4$, and $0.3 \cdot w_1 + 0.4 \cdot w_2 = 0.3$ according to the definition of combined transition, which is clearly not possible. Hence we conclude that $s_0 \not\sim r_0$, and $\sim$ is finer than $\sim_{\text{CSL}}$. □

In [31] randomized schedulers allow to combine transitions with different rates, i.e., the combined transition is defined as: $s \xrightarrow{\lambda}_P \mu$ iff there exist $\{s \xrightarrow{\lambda_i} \mu_i\}_{i \in I}$ and $\{p_i\}_{i \in I}$ such that $\sum_{i \in I} p_i \cdot \lambda_i = \lambda$ and $\sum_{i \in I} p_i \cdot \mu_i = \mu$, where $p_i \in [0, 1]$ for each $i \in I$ and $\sum_{i \in I} p_i = 1$. By adopting this definition of combined transition in Definition 4, we will obtain a coarser strong bisimulation. However it turns out that this new definition of strong bisimulation is too coarse for CSL equivalence, since there exist two states

which are strong bisimilar according to the new definition, but they satisfy different CSL formulas. Refer to the following example:

*Example 2.* Suppose that we have two states $s_1$ and $r_1$ such that $s_1$ has two non-deterministic transitions which can evolve into $u_1$ with rates 1 or 4 respectively. The state $r_1$ is the same as $s_1$ except that it can evolve into $u_1$ with an extra transition of rate 2. Also we assume that $L(s_1) = L(r_1)$ and $u_1$ is a silent state with $L(u_1) \nsubseteq L(s_1)$. Suppose that we adopt the new definition of combined transition in Definition 4 by allowing to combine transitions with different rates, we shall show that $s_1$ and $r_1$ are strong bisimilar, but they are not CSL-equivalent.

We first show that $s_1$ and $r_1$ are strong bisimilar. Let $\mathcal{R}$ be an equivalence relation only equating $s_1$ and $r_1$, it suffices to prove that $\mathcal{R}$ is a strong bisimulation. The only non-trivial case is when $r_1 \xrightarrow{2} \mathcal{D}_{u_1}$, we need to find a matching transition of $s_1$. Since we allow to combine transitions of different rates, a combined transition $s_1 \xrightarrow{2}_P \mathcal{D}_{u_1}$ can be obtained by assigning weights $\frac{2}{3}$ and $\frac{1}{3}$ to transitions $s_1 \xrightarrow{1} \mathcal{D}_{u_1}$ and $s_1 \xrightarrow{4} \mathcal{D}_{u_1}$ respectively. Therefore we conclude that $s_1$ and $r_1$ are strong bisimilar.

Secondly, we show that $s_1$ and $r_1$ are not CSL equivalent. It suffices to find a formula $\varphi$ such that $s_1 \models \varphi$ but $r_1 \nvDash \varphi$. Let $\psi = \mathsf{X}^{[a,b]} L(u_1)$ where $0 \le a < b$. The probabilities for paths starting from $s_1$ and satisfying $\psi$ by choosing the transitions with rates 1, 2, and 4 are equal to $e^{-a} - e^{-b}$, $e^{-2a} - e^{-2b}$, and $e^{-4a} - e^{-4b}$ respectively. We need only to find $a$ and $b$ such that $e^{-2a} - e^{-2b} > \max\{e^{-a} - e^{-b}, e^{-4a} - e^{-4b}\}$. Let $a = 0.2$ and $b = 1$, then $e^{-a} - e^{-b} \approx 0.45$, $e^{-2a} - e^{-2b} \approx 0.53$, and $e^{-4a} - e^{-4b} \approx 0.43$. Let $\varphi = \mathcal{P}_{\le 0.46}(\mathsf{X}^{[0.2,1]} L(u_1))$, obviously $s_1 \models \varphi$, but $r_1 \nvDash \varphi$, which means that $s_1$ and $r_1$ are not CSL-equivalent. $\qquad\square$

Example 2 also shows that in order for two states satisfying the same CSL formulas, it is necessary for them to have transitions with the same exit rates, otherwise we can always find CSL formulas distinguishing them, which also justifies that we only allow to combine transitions with the same rate in Definition 4.

We have shown in Example 1 that $\sim$ is not complete with respect to $\sim_{\mathrm{CSL}}$. However in the sequel we shall identify a special class of CTMDPs, in which the completeness holds. We first give two examples for inspiration:

*Example 3.* In this example, we show that, it is impossible to construct similar states as $s_0$ and $r_0$ in Example 1 such that they are not strong bisimilar but only have 2 distinct successors.

Let $s_2$ and $r_2$ denote the two states depicted in Fig. 2, where $x \in [0, 1]$ denotes an arbitrary or unknown probability and all states have different atomic propositions except that $L(s_2) = L(r_2)$. Our aim is to show that states in form of $s_2$ and $r_2$ must be strong bisimilar, provided that $s_2 \sim_{\mathrm{CSL}} r_2$. First we show that $x \in [\frac{1}{4}, \frac{1}{2}]$ in order that $s_2 \sim_{\mathrm{CSL}} r_2$. This is done by contradiction. Assume that $x > \frac{1}{2}$ and let $\psi = \mathsf{X}^{[0,\infty)}(L(u_1))$. Then the maximal probability of paths starting from $s_2$ and satisfying $\psi$ is equal to $\frac{1}{2}$, while the maximal probability of paths starting from $r_2$ and satisfying $\psi$ is equal to $x$. Since $x > \frac{1}{2}$, $s_2 \models \mathcal{P}_{\le \frac{1}{2}}(\psi)$, while $r_2 \nvDash \mathcal{P}_{\le \frac{1}{2}}(\psi)$, therefore $s_2 \nsim_{\mathrm{CSL}} r_2$. Similarly, we can show that it is not possible for $x < \frac{1}{4}$, hence it holds that $x \in [\frac{1}{4}, \frac{1}{2}]$.

**Fig. 2.** $s_2$ can always simulate the middle transition of $r_2$, as long as $\frac{1}{4} \leq x \leq \frac{1}{2}$.

Secondly, we show that $s_2 \sim r_2$ given that $x \in [\frac{1}{4}, \frac{1}{2}]$. Let $\mathcal{R}$ be an equivalence relation only equating $s_2$ and $r_2$, it suffices to show that $\mathcal{R}$ is a strong bisimulation according to Definition 4. Let $\mu_1, \mu_2$, and $\mu_3$ be distributions defined in Fig. 2. The only non-trivial case is when $r_2 \xrightarrow{1} \mu_2$, we need to show that there exists $w_1$ and $w_2$ such that $w_1 + w_2 = 1$, $(w_1 \cdot \mu_1 + w_2 \cdot \mu_3) \, \mathcal{R} \, \mu_2$. Let $w_1 = 2 - 4x$ and $w_2 = 4x - 1$, it is easy to verify that $w_1, w_2 \in [0, 1]$ and $w_1 + w_2 = 1$, since $x \in [\frac{1}{4}, \frac{1}{2}]$. Moreover, $w_1 \cdot \mu_1 + w_2 \cdot \mu_3 = \mu_2$, since $w_1 \cdot \frac{1}{4} + w_2 \cdot \frac{1}{2} = x$ and $w_1 \cdot \frac{3}{4} + w_2 \cdot \frac{1}{2} = 1 - x$. Therefore $s_2 \xrightarrow{1}_P \mu_2$ as desired, and $\mathcal{R}$ is indeed a strong bisimulation.                □

In order for Example 1 being a valid counterexample for $\sim_{\text{CSL}} \subseteq \sim$, we have made another assumption that $u_i$ $(i = 1, 2, 3)$ are silent, i.e., they cannot evolve into other states not equivalent to themselves with positive probability. This assumption is also crucial which can be seen by the following example:

*Example 4.* Consider again the two states $s_0$ and $r_0$ introduced in Example 1, where we prove that $s_0$ and $r_0$ are CSL equivalent. Now suppose that $u_3$ is not silent, but can evolve into some state $u_3'$ with rate 1, where $u_3'$ is a state with different atomic propositions from all the others. We are going to show that $s_0$ and $r_0$ are not CSL equivalent anymore with this slight change. Consider the path formula: $\psi = (L(s_0) \vee L(u_3)) \, \mathsf{U}^{[0,b]} (L(u_2) \vee L(u_3'))$, we can show that the probabilities of paths starting from $r_0$ and satisfying $\psi$ by choosing the left, middle, and right transitions are equal to: $L = 0.3 \cdot w_1 + 0.4 \cdot w_2$, $M = 0.3 \cdot w_1 + 0.3 \cdot w_2$, and $R = 0.4 \cdot w_1 + 0.1 \cdot w_2$ respectively, where $w_1 = 1 - e^{-b}$ and $w_2 = 1 - e^{-b} - b \cdot e^{-b}$. It suffices to find a $b$ such that $M < \min\{L, R\}$, which means that the middle transition of $r_0$ dominates the minimal probability of satisfying $\psi$. Such $b$ exists, for instance, by letting $b = 1$ we obtain: $L \approx 0.295$, $M \approx 0.269$, and $R \approx 0.279$, apparently, $M < \min\{L, R\}$. In other words, let $b = 1$ in $\psi$, we have $s_0 \models \mathcal{P}_{\geq R}(\psi)$, but $r_0 \not\models \mathcal{P}_{\geq R}(\psi)$, since there exists a scheduler of $r_0$, i.e., the one choosing the middle transition of $r_0$ such that the probability of satisfying $\psi$ is equal to $M$, which is strictly less than $R$. Therefore $s_0 \not\sim_{\text{CSL}} r_0$.                □

In Example 1, we have shown that $s_0$ and $r_0$ satisfy the same CSL formulas, but they are not strong bisimilar. However in Examples 3 and 4, we show that without the two assumptions:

 – $s_0$ and $r_0$ should have more than 2 states among their successors;
 – there exists no successor which can evolve into a state not CSL equivalent to other states with positive probability,

we can guarantee that either $s_0$ and $r_0$ are strong bisimilar, or they are not CSL equivalent. These intuitions lead us to the special class of CTMDPs, which we call *non 2-step recurrent* CTMDPs in the sequel.

**Definition 5 (2-step Recurrent).** *Let $\mathcal{R}$ be an equivalence relation on $S$. A state $s$ is said to be* 2*-step recurrent with respect to $\mathcal{R}$ iff $s$ is not silent, $|Suc(s)| > 2$, and*

$$\exists(s \xrightarrow{\lambda} \mu).(\forall s' \in (Supp(\mu) \setminus [s]_{\mathcal{R}}).\forall(s' \xrightarrow{\lambda'} \nu).\nu(C) = 1), \qquad \text{(r1)}$$

*where $C = ([s]_{\mathcal{R}} \cup [s']_{\mathcal{R}})$.*

*We say $\mathcal{C}$ is* 2*-step recurrent with respect to $\mathcal{R}$, iff there exists $s \in S$ such that $s$ is 2-step recurrent with respect to $\mathcal{R}$, otherwise it is non 2-step recurrent with respect to $\mathcal{R}$. Moreover, we say that $s$ (or $\mathcal{C}$) is (non) 2-step recurrent iff it is (non) 2-step recurrent with respect to $\sim_{CSL}$.*

In other words, for a state $s$ to be 2-step recurrent, it must be not silent and have more than 2 successors. Remind that each silent state can be replaced by a single state without changing properties of a CTMDP. After doing so, each silent state will only have one successor which is itself, so the requirement of non silence can be subsumed by $|Suc(s)| > 2$ in this case. Let us explain the more involved condition given in Eq. (r1). Eq. (r1) says that a 2-step recurrent state $s$ must also satisfy: There exists $s \xrightarrow{\lambda} \mu$ such that for all states in $Supp(\mu)$ except those in $[s]_{\mathcal{R}}$, they can only evolve into states equivalent to $s$ or themselves.

*Example 5.* We show some examples of (non) 2-step recurrent states. First of all, states $s_0$ and $r_0$ in Example 1 are 2-step recurrent, since they are not silent and have more than 2 successors. Moreover all successors $u_i$ ($i = 1, 2, 3$) are silent, i.e., can only evolve into states which are CSL equivalent to themselves. However if we add an extra transition to $u_3$ as in Example 4, $s_0$ will be non 2-step recurrent, since $u_3$ can reach the state $u_3'$ with probability 1, where $u_3'$ is not CSL equivalent to either $u_3$ or $s_0$. For similar reasons, $r_0$ is also non 2-step recurrent.

Secondly, States $s_1$ and $r_1$ in Example 2 and $s_2$ and $r_2$ in Example 3 are trivially non 2-step recurrent, since the number of their successors is $\leq 2$. $\qquad \square$

Definition 5 seems tricky, however, we shall show that there exists an efficient scheme to check whether a given CTMDP is 2-step recurrent or not. More importantly, we shall see later in Remark 1 that the class of non 2-step recurrent CTMDPs contains an important part of CTMDP models, in particular those found in practice.

Now we are ready to show the main contribution of this paper. By restricting to the set of non 2-step recurrent CTMDPs, we are able to prove that the classical strong bisimulation defined in Definition 4 is both sound and complete with respect to the CSL equivalence, which is formalized in the following theorem.

**Theorem 2.** *If $\mathcal{C}$ is non 2-step recurrent, $\sim \; = \; \sim_{CSL}$.*

### 3.2   Weak Bisimulation

In this section we will introduce a novel notion of *weak bisimulation* for CTMDPs. Our definition of weak bisimulation is directly motivated by the well-known fact that uniformization does not alter time-bounded reachabilities for CTMDPs [28, 31] when TTP schedulers are considered. Similar as in Section 3.1, we also show that weak bisimulation is both sound and complete for $CSL_{\setminus X}$ over non 2-step recurrent CTMDPs. We shall introduce the definition of weak bisimulation first.

**Definition 6 (Weak bisimulation).** *We say that states $s$ and $r$ in $\mathcal{C}$ are weak bisimilar, denoted by $s \approx r$, whenever $\bar{s} \sim \bar{r}$ in the uniformized CTMDP $\bar{\mathcal{C}}$.*

The way we define weak bisimulation here is different from the definition of weak bisimulation for CTMCs in [6], where a conditional measure is considered, see Definition 7 for the detailed definition. Moreover we will show in Section 5.2 that for CTMCs our weak bisimulation coincides with weak bisimulation defined in [6]. Even though the resulting uniformized CTMDP depends on the chosen rate $E$ as shown in Definition 2, it is worth mentioning that weak bisimulation given in Definition 6 is independent of $E$. Since if two states are strong bisimilar in a uniformized CTMDP, they will be strong bisimilar in any uniformized CTMDP no matter which value we choose for $E$.

The following lemma establishes some properties:

**Lemma 1.**

1. $\sim \ \subseteq \ \approx$,
2. *for uniformized* CTMDP*s,* $\sim \ = \ \approx$.

As we mentioned above, by uniformizing a CTMDP we will not change its satisfiability of $CSL_{\setminus X}$ provided that only TTP schedulers are considered. Therefore we have the following lemma saying that if two states satisfy the same formulas in $CSL_{\setminus X}$, then they will satisfy the same formulas in CSL after uniformization and vice versa.

**Lemma 2.** $s \sim_{CSL_{\setminus X}} r$ in $\mathcal{C}$ iff $\bar{s} \sim_{CSL} \bar{r}$ in $\bar{\mathcal{C}}$.

The following theorem says that our weak bisimulation is sound for $\sim_{CSL_{\setminus X}}$, and particularly when the given CTMDP is non 2-step recurrent, weak bisimulation can be used to fully characterize $CSL_{\setminus X}$ equivalence.

**Theorem 3.** $\approx \ \subsetneq \ \sim_{CSL_{\setminus X}}$. *If $\bar{\mathcal{C}}$ is non 2-step recurrent,* $\approx \ = \ \sim_{CSL_{\setminus X}}$.

Theorem 3 works if we restrict to only TTP schedulers. However, this is not a restriction. Since it has been proved in [31, 11] that there always exists an optimal scheduler in TTP for any path property in $CSL_{\setminus X}$.

### 3.3   Determining 2-step Recurrent CTMDPs

In Theorem 2 and 3, the completeness holds only for CTMDPs which are non 2-step recurrent. Hence it is important that 2-step recurrent CTMDPs can be checked efficiently. This section discusses a simple procedure for determining (non) 2-step recurrent CTMDPs. Before presenting the decision scheme, we shall introduce the following lemma, which holds by applying the definition of 2-step recurrent CTMDPs directly:

**Lemma 3.** *Given two equivalence relations $\mathcal{R}$ and $\mathcal{R}'$ over $S$ such that $\mathcal{R} \subseteq \mathcal{R}'$, if $\mathcal{C}$ is 2-step recurrent with respect to $\mathcal{R}$, then it is 2-step recurrent with respect to $\mathcal{R}'$, or equivalently if $\mathcal{C}$ is non 2-step recurrent with respect to $\mathcal{R}'$, then it is non 2-step recurrent with respect to $\mathcal{R}$.*

Lemma 3 suggests a simple way to check whether a given CTMDP $\mathcal{C}$ is 2-step recurrent. Given an arbitrary equivalence relation $\mathcal{R}$ such that $\sim \ \subseteq \ \sim_{\mathrm{CSL}} \ \subseteq \mathcal{R}$, by Lemma 3, we can first check whether $\mathcal{C}$ is 2-step recurrent with respect to $\mathcal{R}$. Proper candidates for $\mathcal{R}$ should be as fine as possible, but also can be determined efficiently. For instance, we can let $\mathcal{R} = \{(s, r) \mid L(s) = L(r)\}$, or a finer equivalence relation defined as follows: $s \ \mathcal{R} \ r$ iff for each $C \in S/\mathcal{R}$ and $s \xrightarrow{\lambda} \mu$, there exists $r \xrightarrow{\lambda} \mu'$ such that $\mu'(C) \geq \mu(C)$. Such $\mathcal{R}$ is coarser than $\sim_{\mathrm{CSL}}$, and can be computed efficiently in polynomial time.

If $\mathcal{C}$ is not 2-step recurrent with respect to $\mathcal{R}$, we know that $\mathcal{C}$ is non 2-step recurrent with respect to $\sim_{\mathrm{CSL}}$ either. Otherwise we continue to check whether $\mathcal{C}$ is 2-step recurrent with respect to $\sim$, if the answer is yes, then $\mathcal{C}$ is 2-step recurrent with respect to $\sim_{\mathrm{CSL}}$ too. Note that $\sim$ can also be computed in polynomial time, see [40] for details. In the remaining cases, namely when $\mathcal{C}$ is 2-step recurrent with respect to $\mathcal{R}$, but not for $\sim$, we cannot conclude anything, instead the relation $\sim_{\mathrm{CSL}}$ shall be computed first for a definite answer.

As we discussed above, sometimes we need to use $\sim_{\mathrm{CSL}}$ to decide whether a given CTMDP is 2-step recurrent or not. But it turns out that $\sim_{\mathrm{CSL}}$ is hard to compute in general. Actually, we can prove the following lemma showing that the decision of $\sim_{\mathrm{CSL}}$ and $\sim_{\mathrm{CSL}\setminus \mathsf{x}}$ is NP-hard.

**Lemma 4.** *It is NP-hard to decide whether $s \ \sim_{CSL} \ r$ and $s \ \sim_{CSL\setminus \mathsf{x}} \ r$.*

*Remark 1.* We have implemented the above described scheme to check whether some models in practice are 2-step recurrent or not. Even though the implemented classification scheme is not complete since we do not compute CSL equivalence, it has been shown quite useful in practice. Our initial experiments show that the non 2-step recurrent CTMDPs consist of most models in practice. For instance the models of "Erlang Stages" [41], "Stochastic Job Scheduling" [10], "Fault-Tolerant Work Station Cluster" [20, 24], and "European Train Control System" [7] are all non 2-step recurrent, which means that strong bisimulation coincides with $\sim_{\mathrm{CSL}}$ on these models. To be more confident, we also checked MDP models from the PRISM [26] benchmark interpreted as CTMDP models by interpreting all probabilities as rates. We found that all of them are non 2-step recurrent. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

## 4   Bisimilarity and CSL$^*$ Equivalence

In this section we study the relation between bisimilarity and CSL$^*$ equivalence. We first introduce CSL$^*$, then show that strong bisimulation can be fully characterized by CSL$^*$ for arbitrary CTMDPs. Then we extend the work to weak bisimulation.

### 4.1   CSL*

As CTL* and PCTL* can be seen as extensions of CTL and PCTL respectively, CSL* can also be seen as an extension of CSL, where the path formula is defined by the Metric Temporal Logic (MTL) [25]. MTL extends linear temporal logic [30] by associating each temporal operator with a time interval. It is a popular logic used to specify properties of real-time systems and has been extensively studied in the literature [1, 29, 8, 22]. The logic MTL was also extended to CTMCs in [13], where the authors studied the problem of model checking CTMCs against MTL specifications. Formally, the syntax of CSL* is defined by the following BNFs:

$$\varphi ::= a \mid \neg\varphi \mid \varphi \wedge \varphi \mid \mathcal{P}_{\bowtie p}(\psi),$$
$$\psi ::= \varphi \mid \neg\psi \mid \psi \wedge \psi \mid \mathsf{X}^I\,\psi \mid \psi\,\mathsf{U}^I\,\psi.$$

The semantics of state formulas is the same as CSL, while the semantics of path formulas is more involved, since we may have different and embedded time bounds. As for MTL, there are two different semantics for the path formulas: continuous semantics and pointwise semantics. These two semantics make non-trivial differences in real-time systems, see [29] for details. We shall focus on the pointwise semantics as for CSL in this paper. Given a path $\omega$ and a path formula $\psi$ of CSL*, the satisfiability $\omega \models \psi$ is defined inductively as follows: $\omega \models a$ iff $a \in L(\omega[0])$, $\omega \models \neg\psi$ iff $\omega \not\models \psi$, $\omega \models \psi_1 \wedge \psi_2$ iff $\omega \models \psi_1 \wedge \omega \models \psi_2$, $\omega \models \mathsf{X}^I\,\psi$ iff $\omega|_1 \models \psi \wedge time(\omega,0) \in I$, and

$$\omega \models \psi_1\,\mathsf{U}^I\,\psi_2 \text{ iff } \exists i.(\omega|_i \models \psi_2 \wedge \sum_{0 \le j < i} time(\omega,j) \in I \wedge (\forall 0 \le j < i.\omega|_j \models \psi_1)).$$

### 4.2   Strong Bisimulation

In this section we prove the soundness and completeness of strong bisimulation with respect to CSL* equivalence. Different from CTL and its extension CTL*, whose equivalences coincide on labelled transition systems [9], the extension from CSL to CSL* is non-trivial, as we shall show in this section that CSL* can fully characterize strong bisimulation for arbitrary CTMDPs. We reconsider Example 1 for inspiration:

*Example 6.* Let $s_0$ and $r_0$ be the states introduced in Example 1, where we have shown that $s_0$ and $r_0$ are not bisimilar, but satisfy the same CSL formula. However if we consider CSL*, $s_0$ and $r_0$ are not CSL* equivalent. It suffices to find a formula $\varphi$ in CSL* such that $s_0 \models \varphi$, but $r_0 \not\models \varphi$. Let $\psi := (L(s_0)\,\mathsf{U}^{[0.6,\infty)}\,L(u_1)) \vee (L(s_0)\,\mathsf{U}^{[1,\infty)}\,L(u_3))$, then the maximal probability of paths starting from $s_0$ and satisfying $\psi$ is equal to $\max\{0.3 \cdot e^{-0.6} + 0.4 \cdot e^{-1}, 0.5 \cdot e^{-0.6} + 0.1 \cdot e^{-1}\} < 0.312$, while the probability for $r_0$ is equal to $\max\{0.3 \cdot e^{-0.6} + 0.4 \cdot e^{-1}, 0.4 \cdot e^{-0.6} + 0.3 \cdot e^{-1}, 0.5 \cdot e^{-0.6} + 0.1 \cdot e^{-1}\} > 0.312$, thus $s_0 \models \mathcal{P}_{\le 0.312}(\psi)$, while $r_0 \not\models \mathcal{P}_{\le 0.312}(\psi)$, which indicates $s_0 \not\sim_{\text{CSL}^*} r_0$. Note $\psi$ is not a valid formula in CSL, since it is the disjunction of two until operators.   $\square$

In the remainder of this section, we shall focus on the proof of $\sim = \sim_{\text{CSL}^*}$. First, we introduce the following lemma in [33]:

**Lemma 5 (Theorem 5 [33]).** *Given a path formula $\psi$ of $CSL^*$ and a state $s$, there exists a set of cylinder sets $Cyls$ such that $Sat(\psi) = \cup_{C \in Cyls} C$.*

As a direct result of Lemma 5, $Sat(\psi)$ is measurable for any path formula $\psi$ of $CSL^*$, as $Sat(\psi)$ can be represented by a countable set of measurable cylinders.

Now we are ready to present the main result of this section, i.e., strong bisimulation coincides with $CSL^*$ equivalence for arbitrary CTMDPs:

**Theorem 4.** *For any $CTMDP$, $\sim \; = \; \sim_{CSL^*}$.*

### 4.3   Weak Bisimulation

In this section we shall discuss the relation between weak bisimulation and the equivalence induced by $CSL^*_{\backslash X}$. Similar as in Section 4.2 for strong bisimulation, weak bisimulation can be fully characterized by $CSL^*_{\backslash X}$.

Since our weak bisimulation is defined as strong bisimulation on the uniformized CTMDPs, foremost we shall make sure that $CSL^*_{\backslash X}$ is preserved by uniformization under TTP schedulers, that is, we shall prove the following lemma:

**Lemma 6.** $s \; \sim_{CSL^*_{\backslash X}} \; r$ in $C$ iff $\bar{s} \; \sim_{CSL^*} \; \bar{r}$ in $\bar{C}$.

As a side contribution, we extend the result in [28, 31] and show that uniformization also does not change properties specified by $CSL^*_{\backslash X}$, provided TTP schedulers are considered. Given Lemma 6, the soundness and completeness of $\approx$ with respect to $\sim_{CSL^*_{\backslash X}}$ are then straightforward from Definition 6 and the fact that $\sim$ is both sound and complete with respect to $CSL^*$.

**Theorem 5.** *For any $CTMDP$, $\approx \; = \; \sim_{CSL^*_{\backslash X}}$.*

Currently, we only prove Theorem 5 with respect to TTP schedulers. However, the optimal scheduler for a $CSL^*$ formula may be not a TTP scheduler. Refer to the following example:

*Example 7.* Let $C$ be a CTMDP as in Fig. 3, where the letter on above of each state denotes its label. Moreover states $s_8$ and $s_9$ only have self-loop transitions which are omitted. Let $\psi = ((a \vee b) \, \mathsf{U}^I \, d) \vee ((a \vee c) \, \mathsf{U}^I \, e)$ be a path formula of $CSL^*$. We show that there exists a non-TTP scheduler $\pi$ such that

$$Pr_{\pi,s_4}(\{\omega \in Paths^\infty(C) \mid \omega \models \psi\}) > Pr_{\pi',s_4}(\{\omega \in Paths^\infty(C) \mid \omega \models \psi\})$$

for any TTP scheduler $\pi'$. Let $I = [0, \infty]$. Since $\pi'$ is a TTP scheduler, it can only make decision based on the elapsed time and the current state. When at $s_7$, $\pi'$ will choose either the transition to $s_8$ or the transition to $s_9$ at each time point. Therefore the maximal probability of satisfying $\psi$ is 0.5. However for a general scheduler $\pi$, it can make decision based on the full history. For instance when at $s_7$, we can let $\pi$ choose the transition to $s_8$, if the previous state is $s_5$, otherwise $s_9$. Under this scheduler, the maximal probability of satisfying $\psi$ is equal to 1, which cannot be obtained by any TTP scheduler. From this example, we can see that an optimal scheduler for a $CSL^*$ formula may make it decision based on the elapsed time as well as the states visited.

**Fig. 3.** TTP schedulers are not enough to obtain optimal values for CSL$^*$ properties.

Example 7 shows that it is not enough to consider TTP schedulers in the setting of CSL$^*$. In [28] another class of schedulers called *Total Time History dependent schedulers* (TTH) is introduced. We conjecture that for TTH schedulers: i) they preserve CSL$^*_{\backslash \mathsf{X}}$ properties after uniformization, and ii) they are powerful enough to obtain optimal values for CSL$^*_{\backslash \mathsf{X}}$ properties. Condition i) guarantees that Theorem 5 is valid, while condition ii) makes Theorem 5 general enough. We leave the proof of the conjecture as our future work.

*Remark 2.* The expressiveness of CSL$^*$ may be considered too powerful in certain cases. For instance, path formulas like $\Box(a \, \mathsf{U}^{[2,10]} \, b)$ [5] will be satisfied with probability 0 for any CTMDP. In general, if $\psi$ can only be satisfied with probability strictly less than 1, the probability of satisfying $\psi$ forever will be 0 for any CTMDP.

In the other hand, a small fragment of CSL$^*$ is enough to characterize strong bisimulation. Let CSL$^\vee$ denote the fragment of CSL$^*$ whose path formulas are defined by the following syntax: $\psi ::= \mathsf{X}^I \, \varphi \mid \psi \vee \psi$. We have shown in [35] that $\sim \; = \; \sim_{\text{CSL}^\vee}$ for any CTMDP. Therefore any subset of CSL$^*$ which subsumes CSL$^\vee$ will be strong enough to fully characterize strong bisimulation.

## 5   Relation to MDPs and CTMCs

In this section, we compare related work on other stochastic models: MDPs and CTMCs.

### 5.1   Relation to (Weak) Bisimulation for MDPs

For MDPs, it is known that strong (probabilistic) bisimulation is only sound but not complete with respect to PCTL [32]–the counterpart of CSL in discrete setting. Differently, the completeness does not hold either even if we restrict to non 2-step recurrent MDPs, which can be defined in a straightforward way given Definition 5. Refer to the following example:

---

[5] $\Box\psi \equiv \neg((a \wedge \neg a) \, \mathsf{U}^{[0,\infty)} \, \neg\psi)$ for some $a$, i.e., $\psi$ holds forever.

*Example 8.* Let $s_0$ and $r_0$ be two states as in Example 4, which will be viewed as two states in an MDP. Moreover we assume that $u_3'$ only has a self loop. Since $u_3'$ has atomic propositions different from $s_0$ ($r_0$) and $u_3$, therefore $s_0$ and $r_0$ are not 2-step recurrent. However $s_0$ and $r_0$ satisfy the same PCTL formulas, since the maximal and minimal probabilities from $s_0$ and $r_0$ to any subset of $\{u_1, u_2, u_3, u_3'\}$ are the same. As mentioned before, the middle transition of $r_0$ cannot be simulated by any combined transition of $s_0$, hence they are not strong probabilistic bisimilar. This indicates that strong (probabilistic) bisimulation is not complete with respect to PCTL equivalence even that the given MDP is non 2-step recurrent.                                          □

The counterpart of CSL$^*$ in discrete setting is PCTL$^*$ [3]. Similar as in the continuous case, the equivalence induced by PCTL$^*$ is strictly finer than $\sim_{\text{PCTL}}$ [34]. However, different from the continuous case, $\sim_{\text{PCTL}^*}$ is still coarser than strong (probabilistic) bisimulation for MDPs, that is, strong (probabilistic) bisimulation is not complete with respect to PCTL$^*$:

*Example 9.* Let $s_0$ and $r_0$ be two states as in Example 1, where we have shown that $s_0$ and $r_0$ are neither strong bisimilar nor CSL$^*$ equivalent. However in [34] $s_0$ and $r_0$ are shown to be PCTL$^*$ equivalent by viewing them as two states in an MDP. Therefore CSL$^*$ gains more expressiveness by adding time bounds to the logic.                                          □

The case for weak bisimulation is similar and omitted here.

## 5.2   Relation to (Weak) Bisimulation for CTMCs

In this section we show that our bisimulations are downward compatible to those for CTMCs. Different from CTMDPs, there is no non-deterministic transitions in CTMCs, i.e., each state has only one transition, which will be denoted by $s \xrightarrow{\lambda_s} \mu_s$. The notion of weak bisimulation can be found in [6] for CTMCs, which is repeated as follows:

**Definition 7  (Weak Bisimulation of CTMCs).** *For* CTMC*s, an equivalence relation* $\mathcal{R}$ *is a weak bisimulation iff for all* $s \mathcal{R} r$ *it holds: i)* $L(s) = L(r)$*, and ii)* $\lambda_s \cdot \mu_s(C) = \lambda_r \cdot \mu_r(C)$ *for all equivalence classes* $C \neq [s]_{\mathcal{R}}$*.*
    *States* $s, r$ *are weak bisimilar, denoted by* $s \approx_{CTMC} r$*, iff there exists a weak bisimulation* $\mathcal{R}$ *such that* $s \mathcal{R} r$*.*

Strong bisimulation for CTMCs is defined if in addition $\lambda_s \cdot \mu_s(C) = \lambda_r \cdot \mu_r(C)$ holds for $C = [s]_{\mathcal{R}} = [r]_{\mathcal{R}}$ as well. States $s, r$ are strong bisimilar, denoted by $s \sim_{CTMC} r$, iff there exists a strong bisimulation $\mathcal{R}$ such that $s \mathcal{R} r$.
    Below we prove that, restricted to CTMCs, our strong and weak bisimulations agree with strong and weak bisimulations for CTMCs, respectively:

**Lemma 7.** *For CTMCs, it holds that* $\sim \; = \; \sim_{CTMC}$ *and* $\approx \; = \; \approx_{CTMC}$*.*

The lemma above shows that $\sim$ and $\approx$ are conservative extensions of strong and weak bisimulations for CTMCs in [6], and so are their logical characterization results except that they only work on a subset of CTMDPs free of 2-step recurrent states.
    Since CTMCs are sub-models of CTMDPs, Theorem 4 and 5 also hold for CTMCs. Together with Lemma 7, we have the following result:

**Corollary 1.**    *1.* $\sim_{CSL^*}$  =  $\sim$  =  $\sim_{CTMC}$  =  $\sim_{CSL}$,
*2.* $\sim_{CSL^*_{\backslash X}}$  =  $\approx$  =  $\approx_{CTMC}$  =  $\sim_{CSL_{\backslash X}}$.

Corollary 1 shows that $CSL^*$ gains no more distinguishing power than CSL on CTMCs without non-determinism, similarly for their sub-logics without the next operator.

## 6   Conclusion and Future Work

In this paper, we have proposed both strong and weak bisimulations for CTMDPs, which are shown to be able to fully characterize CSL and $CSL_{\backslash X}$ equivalences respectively, but over non 2-step recurrent CTMDPs. For a standard extension of CSL – $CSL^*$, we show that strong and weak bisimulations are both sound and complete with respect to $CSL^*$ and $CSL^*_{\backslash X}$ respectively for arbitrary CTMDPs. Moreover, we give a simple scheme to determine non 2-step recurrent CTMDPs, and show almost all CTMDPs found in practice are non 2-step recurrent CTMDPs. We note that the work in this paper can be extended to the simulation setting in a straightforward way.

For future work we would like to consider the approximation of bisimulations and simulations on CTMDPs as well as their logic characterization, along [16]. Moreover, the model checking of $CSL^*$ against CTMCs and CTMDPs will be also worthwhile to exploit. Another interesting direction is to consider the continuous semantics of $CSL^*$.

## Acknowledgements

## References

1. R. Alur and T. A. Henzinger. A really temporal logic. *J. ACM*, 41(1):181–203, 1994.
2. A. Aziz, K. Sanwal, V. Singhal, and R. K. Brayton. Verifying continuous time Markov chains. In *CAV*, volume 1102 of *LNCS*, pages 269–276. Springer-Verlag, 1996.
3. A. Aziz, V. Singhal, and F. Balarin. It usually works: The temporal logic of stochastic systems. In *CAV*, volume 939 of *LNCS*, pages 155–165. Springer-Verlag, 1995.
4. C. Baier, B. Haverkort, H. Hermanns, and J.-P. Katoen. Model-checking algorithms for continuous-time Markov chains. *IEEE Trans. Softw. Eng.*, 29(6):524–541, 2003.
5. C. Baier, H. Hermanns, J.-P. Katoen, and B. R. Haverkort. Efficient computation of time-bounded reachability probabilities in uniform continuous-time Markov decision processes. *Theor. Comput. Sci.*, 345(1):2–26, 2005.
6. C. Baier, J.-P. Katoen, H. Hermanns, and V. Wolf. Comparative branching-time semantics for Markov chains. *Inf. Comput.*, 200(2):149–214, 2005.

7. E. Böde, M. Herbstritt, H. Hermanns, S. Johr, T. Peikenkamp, R. Pulungan, R. Wimmer, and B. Becker. Compositional performability evaluation for STATEMATE. In *QEST*, pages 167–178. IEEE, 2006.

8. P. Bouyer, N. Markey, J. Ouaknine, and J. Worrell. The cost of punctuality. In *LICS*, pages 109–120. IEEE, 2007.

9. M. C. Browne, E. M. Clarke, and O. Grümberg. Characterizing finite Kripke structures in propositional temporal logic. *Theor. Comput. Sci.*, 59(1-2):115–131, July 1988.

10. J. Bruno, P. Downey, and G. N. Frederickson. Sequencing tasks with exponential service times to minimize the expected flow time or makespan. *J. ACM*, 28(1):100–113, 1981.

11. P. Buchholz, E. M. Hahn, H. Hermanns, and L. Zhang. Model checking algorithms for CT-MDPs. In *Proceedings of the 23rd international conference on Computer aided verification*, LNCS, pages 225–242. Springer-Verlag, 2011.

12. P. Buchholz and I. Schulz. Numerical analysis of continuous time Markov decision processes over finite horizons. *Computers & Operations Research*, 38(3):651–659, 2011.

13. T. Chen, M. Diciolla, M. Kwiatkowska, and A. Mereacre. Time-bounded verification of CTMCs against real-time specifications. In *FORMATS*, volume 6919 of *LNCS*, pages 26–42. Springer-Verlag, 2011.

14. T. H. Cormen, C. Stein, R. L. Rivest, and C. E. Leiserson. *Introduction to Algorithms*. McGraw-Hill Higher Education, 2nd edition, 2001.

15. P. R. D'Argenio, N. Wolovick, P. S. Terraf, and P. Celayes. Nondeterministic labeled Markov processes: Bisimulations and logical characterization. In *QEST*, pages 11–20. IEEE, 2009.

16. J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Metrics for labelled Markov processes. *Theor. Comput. Sci.*, 318(3):323–354, 2004.

17. J. Desharnais, V. Gupta, R. Jagadeesan, and P. Panangaden. Weak bisimulation is sound and complete for pCTL$^*$. *Inf. Comput.*, 208(2):203–219, 2010.

18. J. Desharnais and P. Panangaden. Continuous stochastic logic characterizes bisimulation of continuous-time Markov processes. *J. Log. Algebr. Program.*, 56(1-2):99–115, 2003.

19. H. Hansson and B. Jonsson. A logic for reasoning about time and reliability. *Formal aspects of computing*, 6(5):512–535, 1994.

20. B. R. Haverkort, H. Hermanns, and J.-P. Katoen. On the use of model checking techniques for dependability evaluation. In *SRDS*, pages 228–237, 2000.

21. H. Hermanns, A. Parma, R. Segala, B. Wachter, and L. Zhang. Probabilistic logical characterization. *Inf. Comput.*, 209(2):154–172, 2011.

22. M. Jenkins, J. Ouaknine, A. Rabinovich, and J. Worrell. Alternating timed automata over bounded time. In *LICS*, pages 60–69. IEEE, 2010.

23. B. Jonsson, K. Larsen, and Y. Wang. Probabilistic extensions of process algebras. In J. Bergstra, A. Ponse, and S. Smolka, editors, *Handbook of Process Algebra*, pages 685–710. Elsevier, 2001.

24. J.-P. Katoen, I. S. Zapreev, E. M. Hahn, H. Hermanns, and D. N. Jansen. The ins and outs of the probabilistic model checker MRMC. In *QEST*, pages 167–176, 2009.

25. R. Koymans. Specifying real-time properties with metric temporal logic. *Real-Time Syst.*, 2(4):255–299, Oct. 1990.

26. M. Kwiatkowska, G. Norman, and D. Parker. PRISM 4.0: verification of probabilistic real-time systems. In *CAV*, volume 6806 of *LNCS*, pages 585–591. Springer-Verlag, 2011.

27. M. R. Neuhäußer and J.-P. Katoen. Bisimulation and logical preservation for continuous-time Markov decision processes. In *CONCUR*, volume 4703 of *LNCS*, pages 412–427. Springer-Verlag, 2007.

28. M. R. Neuhäußer, M. Stoelinga, and J.-P. Katoen. Delayed nondeterminism in continuous-time Markov decision processes. In *FOSSACS*, volume 5504 of *LNCS*, pages 364–379, 2009.

29. J. Ouaknine and J. Worrell. On the decidability of metric temporal logic. In *LICS*, pages 188–197. IEEE, 2005.
30. A. Pnueli. The temporal logic of programs. In *FOCS*, pages 46–57. IEEE, 1977.
31. M. N. Rabe and S. Schewe. Finite optimal control for time-bounded reachability in CTMDPs and continuous-time Markov games. *Acta Inf.*, 48(5-6):291–315, 2011.
32. R. Segala and N. A. Lynch. Probabilistic simulations for probabilistic processes. *Nord. J. Comput.*, 2(2):250–273, 1995.
33. A. Sharma and J.-P. Katoen. Weighted lumpability on Markov chains. In *Perspectives of Systems Informatics*, volume 7162 of *LNCS*, pages 322–339. Springer-Verlag, 2012.
34. L. Song, L. Zhang, and J. Godskesen. Bisimulations meet PCTL equivalences for probabilistic automata. In *CONCUR*, volume 6901 of *LNCS*, pages 108–123, 2011.
35. L. Song, L. Zhang, and J. C. Godskesen. The branching time spectrum for continuous-time mdps. *CoRR*, abs/1204.1848, 2012.
36. M. Tracol, J. Desharnais, and A. Zhioua. Computing distances between probabilistic automata. In *QAPL*, pages 148–162, 2011.
37. R. J. van Glabbeek. The linear time - branching time spectrum ii. In *CONCUR*, volume 715 of *LNCS*, pages 66–81, 1993.
38. R. J. van Glabbeek. The linear time - branching time spectrum i. In J. Bergstra, A. Ponse, and S. Smolka, editors, *Handbook of Process Algebra*, pages 3–99. Elsevier, 2001.
39. N. Wolovick and S. Johr. A characterization of meaningful schedulers for continuous-time Markov decision processes. In *FORMATS*, volume 4202 of *LNCS*. Springer-Verlag, 2006.
40. L. Zhang, H. Hermanns, F. Eisenbrand, and D. N. Jansen. Flow faster: Efficient decision algorithms for probabilistic simulations. *Logical Methods in Computer Science*, 4(4), 2008.
41. L. Zhang and M. R. Neuhäußer. Model Checking Interactive Markov Chains. In *TACAS*, volume 6015 of *LNCS*, pages 53–68. Springer-Verlag, 2010.

# A    Proofs of Section 3

## A.1    Proof of Theorem 2

*Proof.* Theorem 1, states the direction $\sim \quad \subseteq \quad \sim_{\text{CSL}}$, so here we only prove that $\sim_{\text{CSL}} \subseteq \sim$.

Let $\mathcal{R} = \sim_{\text{CSL}}$ and $s \, \mathcal{R} \, r$, where $\mathcal{R}$ is obviously an equivalence relation. We need to prove that $L(s) = L(r)$ and for each $s \xrightarrow{\lambda} \mu$, there exists $r \xrightarrow{\lambda}_{\text{P}} \mu'$ such that $\mu \, \mathcal{R} \, \mu'$. The proof is along the same line as the proof of Theorem 4: we only need to consider the $|Suc(s)| > 2$, as the formula constructed there contains disjunctions.

Recall that in this theorem $\mathcal{C}$ is non 2-step recurrent. Let $s_k \in Supp(\mu)$ be a state such that there exists $t \notin [s] \cup [s_k]$, $s_k \xrightarrow{\lambda_2} \nu$ and $\nu(t) > 0$ for some $\lambda_2$ and $\nu$. Since $\mathcal{C}$ is non 2-step recurrent, such $s_k$ always exists. Then the formula for case when $a_k \in (b_k, c_k)$ and $a_j \in (c_j, b_j)$ is given by:

$$\psi = (s \vee s_k) \, \mathsf{U}^{[a,b]} (s_j \vee t)$$

We also distinguish the following three sub-cases:

$\lambda_1 = \lambda_2$: Let

$$\rho_1 = \rho \cdot (e^{-\lambda_1 a} - e^{-\lambda_1 b} + a\lambda_1 e^{-\lambda_1 a} - b\lambda_1 e^{-\lambda_1 b})$$
$$\rho_2 = (e^{-\lambda_1 a} - e^{-\lambda_1 b})$$

then
- the probability of paths starting from $s$ satisfying $\psi$ by choosing transitions $s \xrightarrow{\lambda_1} \mu$ and $s_k \xrightarrow{\lambda_2`} \nu$ is equal to $p(s, \mu) := a_j \cdot \rho_2 + a_k \cdot \rho_1$, and
- the probabilities of paths starting from $r$ satisfying $\psi$ by choosing transitions $r \xrightarrow{\lambda_1} \mu'_1$ and $r \xrightarrow{\lambda_1} \mu'_2$ and then $s_k \xrightarrow{\lambda_2} \nu$ are equal to $p(r, \nu_1) := b_j \cdot \rho_2 + b_k \cdot \rho_1$ and $p(r, \nu_2) = c_j \cdot \rho_2 + c_k \cdot \rho_1$ respectively.

As in Theorem 4, it is sufficient to prove that $\frac{\rho_1}{\rho_2} \in (0, \infty)$, which can be seen as follows:
- Let $b = \infty$, then $\frac{\rho_1}{\rho_2} = \rho \cdot (a\lambda_1 + 1)$ and it is easy to see that there exists $a, b$ such that $\frac{\rho_1}{\rho_2} \in [\rho, \infty)$.
- On the other hand let $a = 0$, then $\rho_1 = \rho(1 - e^{-\lambda_1 b} - b\lambda_1 e^{-\lambda_1 b})$ and $\rho_2 = 1 - e^{-\lambda_1 b}$, so $\frac{\rho_1}{\rho_2} = \rho \cdot (1 - \frac{b\lambda_1 e^{-\lambda_1 b}}{1 - e^{-\lambda_1 b}})$, note here that $\frac{b\lambda_1 e^{-\lambda_1 b}}{1 - e^{-\lambda_1 b}} \in (0, 1)$ since $\frac{b\lambda_1 e^{-\lambda_1 b}}{1 - e^{-\lambda_1 b}}$ can be arbitrary close to 1 when $b$ is close to 0, while $\frac{b\lambda_1 e^{-\lambda_1 b}}{1 - e^{-\lambda_1 b}}$ is arbitrary close to 0 as $b$ increases. As a result $\frac{\rho_1}{\rho_2} \in (0, \rho)$.

$\lambda_1 > \lambda_2$: Then $\rho_2$ will be the same as in the case when $\lambda_1 = \lambda_2$ and

$$\rho_1 = \rho\left(\frac{\lambda_1}{\lambda_1 - \lambda_2}(e^{-\lambda_2 a} - e^{\lambda_2 b}) - \frac{\lambda_2}{\lambda_1 - \lambda_2}(e^{-\lambda_1 a} - e^{\lambda_1 b})\right).$$

Therefore

$$\frac{\rho_1}{\rho_2} = \rho\left(\frac{\lambda_1}{\lambda_1 - \lambda_2}\left(\frac{e^{-\lambda_2 a} - e^{-\lambda_2 b}}{e^{-\lambda_1 a} - e^{-\lambda_1 b}}\right) - \frac{\lambda_2}{\lambda_1 - \lambda_2}\right).$$

When $\lambda_1 > \lambda_2$, $\frac{e^{-\lambda_2 a} - e^{-\lambda_2 b}}{e^{-\lambda_1 a} - e^{-\lambda_1 b}} \in (\frac{\lambda_2}{\lambda_1}, \infty)$, thus $\frac{\rho_1}{\rho_2} \in (0, \infty)$. The remaining arguments are the same as in the case when $\lambda_1 = \lambda_2$.

$\lambda_1 < \lambda_2$: This case is similar as the above case and is omitted.

Thus there always exists $0 \leq a \leq b$ such that paths starting from $r$ will satisfy $\psi$ with higher probability than $s$ for some $a, b$, therefore $s \not\sim_{\text{CSL}} r$, which contradict the assumption.                                                                                $\square$

### A.2   Proof of Lemma 1

*Proof.*   1. Let $\mathcal{R} = \sim$ and $s \, \mathcal{R} \, r$. To show that $\sim$ implies $\approx$, it is enough to prove that $\mathcal{R}$ is a weak bisimulation. Let $\bar{s} \xrightarrow{E} \mu$, according to Definition 6 we need to prove that there exists $\bar{r} \xrightarrow{E} \nu$ such that $\mu \, \mathcal{R} \, \nu$. By Definition 2, $\bar{s} \xrightarrow{E} \mu$ iff there exists $s \xrightarrow{\lambda} \mu'$ such that $\mu = \frac{\lambda}{E} \cdot \mu' + \frac{E-\lambda}{E} \cdot \mathcal{D}_s$. Since $s \sim r$, there exists $r \xrightarrow{\lambda}_{\text{P}} \nu'$ such that $\mu' \, \mathcal{R} \, \nu'$. Note that $r \xrightarrow{\lambda}_{\text{P}} \nu'$ implies $r \xrightarrow{E}_{\text{P}} \nu \equiv \frac{\lambda}{E} \cdot \nu' + \frac{E-\lambda}{E} \cdot \mathcal{D}_r$, apparently $\mu \, \mathcal{R} \, \nu$ as required.
   2. The proof of Clause 2 is straightforward from Definition 6.

$\square$

### A.3   Proof of Lemma 2

*Proof.*   We first prove that if $\mathcal{C}$ is a CTMC, then $s \sim_{\text{CSL}_{\backslash\mathsf{X}}} r$ in $\mathcal{C}$ iff $\bar{s} \sim_{\text{CSL}} \bar{r}$ in $\bar{\mathcal{C}}$. Since uniformization preserves the satisfiability of $\text{CSL}_{\backslash\mathsf{X}}$, we have $\bar{s} \sim_{\text{CSL}_{\backslash\mathsf{X}}} \bar{r}$.

Let $\mathcal{R} = \sim_{\text{CSL}_{\backslash\mathsf{X}}}$ and $\bar{s} \, \mathcal{R} \, \bar{r}$. According to [6], CSL equivalence coincides with strong bisimulation on CTMCs, therefore it suffices to prove that $\mathcal{R}$ is a strong bisimulation. Let $\lambda$ denote the exit rate of $\bar{s}$ and $\bar{r}$, and $\lambda_{\bar{s}}$ denote the rate from $\bar{s}$ to states in $[\bar{s}]_{\mathcal{R}}$ i.e. $\lambda_{\bar{s}} = \lambda \cdot \mu([\bar{s}]_{\mathcal{R}})$ where $\bar{s} \xrightarrow{\lambda} \mu$. We need to prove that there exists $\bar{r} \xrightarrow{\lambda} \nu$ such that $\mu \, R \, \nu$.

The case when $\lambda_{\bar{s}} = \lambda$ is trivial, we assume that $\lambda > \lambda_{\bar{s}}$.

In the following proof, we let $\varphi_C$ be a formula such that $Sat(\varphi_C) = C$ where $C$ is a $\mathcal{R}$ closed set. Now we are going to prove that $\lambda_{\bar{s}} = \lambda_{\bar{r}}$ i.e. the rates for $s$ and $r$ leaving to states in equivalence classes different from $[\bar{s}]_{\mathcal{R}}$ are equal. Let $C = \bar{\mathcal{S}} \setminus [\bar{s}]_{\mathcal{R}}$, then $s \models \mathcal{P}_{\geq p}(\varphi_{[\bar{s}]_{\mathcal{R}}} \, \mathsf{U}^{[a,b]} \, \varphi_C)$ where $p = e^{-\lambda' a} - e^{-\lambda' b}$ and $\lambda' = \lambda - \lambda_{\bar{s}}$. Since $\bar{s} \sim_{\text{CSL}_{\backslash\mathsf{X}}} \bar{r}$, we have $r \models \mathcal{P}_{\geq p}(\varphi_{[\bar{s}]_{\mathcal{R}}} \, \mathsf{U}^{[a,b]} \, \varphi_C)$ for any $0 \leq a < b$. Therefore $\lambda - \lambda_{\bar{s}} = \lambda - \lambda_{\bar{r}}$ which implies $\lambda_{\bar{s}} = \lambda_{\bar{r}}$.

Let $C \in \bar{S}/\mathcal{R}$ be an equivalence relation such that $\bar{s} \notin C$, we know that $\bar{s} \models \varphi := \mathcal{P}_{\geq p}(\varphi_{[\bar{s}]_{\mathcal{R}}} \, \mathsf{U}^{[a,b]} \, \varphi_C)$ where

$$p = \frac{\lambda \cdot \mu(C)}{\lambda - \lambda_{\bar{s}}} \cdot (e^{-\lambda_C \cdot a} - e^{-\lambda_C \cdot b}).$$

Since $\bar{s} \sim_{\text{CSL}_{\backslash\mathsf{X}}} \bar{r}$, we have $\bar{r} \models \varphi$. We show that it must be the case that $\mu(C) = \nu(C)$. We prove by contradiction and distinguish the following cases:

   1. $\mu(C) < \nu(C)$. Let $a = 0$ and $b = \infty$, then $p = \frac{\lambda \cdot \mu(C)}{\lambda - \lambda_{\bar{s}}}$. The probability of the paths starting from $r$ satisfying $(\varphi_{[\bar{s}]_{\mathcal{R}}} \, \mathsf{U}^{[a,b]} \, \varphi_C)$ is $\frac{\lambda \cdot \nu(C)}{\lambda - \lambda_{\bar{r}}}$ which is apparently greater than $p$, given that we have proved that $\lambda_{\bar{s}} = \lambda_{\bar{r}}$. Therefore $r \models$

$\mathcal{P}_{\geq p'}(\varphi_{[\bar{s}]_{\mathcal{R}}} \mathsf{U}^{[a,b]} \varphi_C)$, but $s \not\models \mathcal{P}_{\geq p'}(\varphi_{[\bar{s}]_{\mathcal{R}}} \mathsf{U}^{[a,b]} \varphi_C)$ where $p' = \frac{\lambda \cdot \nu(C)}{\lambda - \lambda_{\bar{r}}}$, this contradicts with our assumption.

2. $\mu(C) > \nu(C)$. This case is similar as the first case by letting $a = 0$ and $b = \infty$, thus is omitted here.

Consequently, we have that $\mu(C) = \nu(C)$ for each $C \in \bar{S}/\mathcal{R}$ except for $[\bar{s}]_{\mathcal{R}}$, moreover $\lambda_{\bar{s}} = \lambda_{\bar{r}}$, hence $\mu \mathcal{R} \nu$ and $\mathcal{R}$ is a strong bisimulation. According to [6] where it is was shown that $\sim$ is both sound and complete for $\sim_{\mathrm{CSL}}$ on CTMC, thus $\bar{s} \sim_{\mathrm{CSL}} \bar{r}$.

We now generalize the result to CTMDPs. If $s \sim_{\mathrm{CSL}\backslash\mathsf{x}} r$, then $\bar{s} \sim_{\mathrm{CSL}\backslash\mathsf{x}} \bar{r}$. Since in a uniformized CTMDP, every execution of $\mathcal{C}$ guided by a given scheduler can be seen as a CTMC, thus $\bar{s} \sim_{\mathrm{CSL}} \bar{r}$ based on the above result. □

### A.4   Proof of Theorem 3

*Proof.* Since in Theorem 2, we have shown that $\sim\ =\ \sim_{\mathrm{CSL}}$ provided that $\mathcal{C}$ is non 2-step recurrent. The proof is straightforward since:

$$(s \approx r) \overset{Def.\ 6}{\Longleftrightarrow} (\bar{s} \sim \bar{r}) \overset{Thm.\ 2}{\Longleftrightarrow} (\bar{s} \sim_{\mathrm{CSL}} \bar{r}) \overset{Lem.\ 2}{\Longleftrightarrow} (s \sim_{\mathrm{CSL}\backslash\mathsf{x}} r).$$
□

### A.5   Proof of Lemma 3

*Proof.* Straightforward from Definition 5. The first two cases are simple since they do not depend on the given relation. We only need to check the third condition. Since $\mathcal{R} \subseteq \mathcal{R}'$ implies $[s]_{\mathcal{R}} \subseteq [s]_{\mathcal{R}'}$ for any $s$. Therefore if there exists $s \xrightarrow{\lambda} \mu$ such that for all $s' \in Supp(\mu)$ and $s' \xrightarrow{\lambda'} \nu$, we always have $\nu(C) = 1$ where $C = [s]_{\mathcal{R}} \cup [s']_{\mathcal{R}}$, it must be the case that $\nu(C') = 1$ where $C' = [s]_{\mathcal{R}'} \cup [s']_{\mathcal{R}'}$, since $C \subseteq C'$. □

### A.6   Proof of Lemma 4

*Proof.* Our proof is inspired by the reduction used in the long version of [36]. We sketch the proof here.

Consider the subset sum problem which is known to be NP-hard [14]: Given a set of $n$ integers $\{k_1, \dots, k_n\}$, is there a non-empty subset whose sum is equal to 0. Note any subset sum problem can be reduced to the following problem by dividing each $k_i$ by $\frac{1}{4n} \cdot \max\{|k_i|\}$ where $1 \leq i \leq n$: Given $n$ decimal numbers $w_1, \dots, w_n$ such that $w_i \in [-\frac{1}{4n}, \frac{1}{4n}]$ for each $i \in [1, n]$, can we find a set $I \subseteq [1, n]$ such that $\sum_{i \in I} w_i = 0$. We show that this problem can also be transformed to a problem of deciding the negation of $\sim_{\mathrm{CSL}}$ by constructing a CTMDP as follows: Suppose we have states $s_0, s_0', r$, and $\{s_i\}_{1 \leq i \leq n}$, all of which have distinct atomic propositions except $L(s_0) = L(s_0')$, and moreover they only have a self loop transition with rate 1 except: $s_0 \xrightarrow{1} \mu$, $s_0' \xrightarrow{1} \nu_1$, and $s_0' \xrightarrow{1} \nu_2$, where for each $1 \leq i \leq n$

- $\mu(s_i) = |w_i| + \epsilon$ with $\epsilon = 10^{-2n}$,
- $\nu_1(s_i) = w_i + |w_i|$,
- $\nu_2(s_i) = -w_i + |w_i|$.

Moreover let $\mu(r) = 1 - \sum_{1 \le i \le n}(|w_i| + \epsilon)$, $\nu_1(r) = 1 - \sum_{1 \le i \le n}(w_i + |w_i|)$, and $\nu_2(r) = 1 - \sum_{1 \le i \le n}(-w_i + |w_i|)$. Clearly $\mu$, $\nu_1$, and $\nu_2$ are full distributions. In order to check whether $s_0 \sim_{\text{CSL}} s_0'$, the only non-trivial cases are formulas like $\mathcal{P}_{\ge p}(\psi)$, where $\psi = \top \mathsf{U}^{[a,b]}(\vee_{s \in C} s)$ for some $C \subseteq \{s_i\}_{1 \le i \le n} \cup \{r\}$. Since the probabilities of paths starting from $s_0$ and $s_0'$ satisfying $\psi$ by choosing transitions to $\mu$, $\nu_1$, and $\nu_2$ are equal to: $\mu(E) \cdot (e^{-a} - e^{-b})$, $\nu_1(E) \cdot (e^{-a} - e^{-b})$, and $\nu_2(E) \cdot (e^{-a} - e^{-b})$ respectively, $s_0 \nsim_{\text{CSL}} s_0'$ iff they exists $E$ such that $\mu(E) > \nu_1(E)$ and $\mu(E) > \nu_2(E)$. We distinguish the following two cases:

1. $r \notin E$ i.e. there exists $I \subseteq [1, n]$ such that $E = \{s_i \mid i \in I\}$.
   In this case we will have

   $$\sum_{i \in I} \mu(s_i) > \sum_{i \in I} \nu_1(s_i), \ \sum_{i \in I} \mu(s_i) > \sum_{i \in I} \nu_2(s_i),$$

   which implies

   $$\sum_{i \in I}(\epsilon + |w_i|) > \sum_{i \in I}(w_i + |w_i|), \ \sum_{i \in I}(\epsilon + |w_i|) > \sum_{i \in I}(-w_i + |w_i|),$$

   which implies

   $$\sum_{i \in I} w_i < \epsilon \cdot |E|, \ -\sum_{i \in I} w_i < \epsilon \cdot |E|.$$

   Since $\epsilon \cdot |E| < 10^{-2n} \cdot n < \frac{1}{4n}$, the only possibility for both $\sum_{i \in I} w_i < \epsilon \cdot |E|$ and $-\sum_{i \in I} w_i < \epsilon \cdot |E|$ hold is that $\sum_{i \in I} w_i = 0$.
2. $r \in E$ i.e. there exists $I \subseteq [1, n]$ such that $E = \{s_i \mid i \in I\} \cup \{r\}$.
   In this case we will have

   $$\mu(r) + \sum_{i \in I} \mu(s_i) > \nu_1(r) + \sum_{i \in I} \nu_1(s_i),$$

   $$\mu(r) + \sum_{i \in I} \mu(s_i) > \nu_2(r) + \sum_{i \in I} \nu_2(s_i),$$

   which implies

   $$1 - \sum_{1 \le i \le n}(\epsilon + |w_i|) + \sum_{i \in I}(\epsilon + |w_i|) > 1 - \sum_{1 \le i \le n}(w_i + |w_i|) + \sum_{i \in I}(w_i + |w_i|),$$

   $$1 - \sum_{1 \le i \le n}(\epsilon + |w_i|) + \sum_{i \in I}(\epsilon + |w_i|) > 1 - \sum_{1 \le i \le n}(-w_i + |w_i|) + \sum_{i \in I}(-w_i + |w_i|),$$

   which implies

   $$-\epsilon \cdot |\bar{I}| > \sum_{i \in \bar{I}} w_i,$$

   $$-\epsilon \cdot |\bar{I}| > -\sum_{i \in \bar{I}} w_i,$$

   where $\bar{I} = [1, n] \setminus I$, which holds iff $\bar{I} = \emptyset$, but this contradicts that $\mu(E) = \nu_1(E) = \nu_2(E) = 1$.

In conclusion, $s_0 \not\sim_{\mathrm{CSL}} s_0'$ iff there exist $I \subseteq [1, n]$ such that $\sum_{i \in I} w_i = 0$. Since the reduction is polynomial, we can say that it is NP-hard to decide $\not\sim_{\mathrm{CSL}}$, which implies that the decision of $\sim_{\mathrm{CSL}}$ is also NP-hard.

The above proof can also be applied to prove that deciding $\sim \mathrm{CSL}_{\backslash \mathsf{X}}$ is NP-hard.

□

## B  Proofs of Section 4

### B.1  Proof of Theorem 4

The proof of Theorem 4 is divided into the following lemmas:

**Lemma 8.** $s \sim r$ *implies* $s \sim_{CSL^*} r$ *for any* $s$ *and* $r$ *i.e.* $\sim \ \subseteq \ \sim_{CSL^*}$.

*Proof.* We shall show that $s \sim r$ implies $s \sim_{\mathrm{CSL}^*} r$ for any $s$ and $r$, that is, $s \sim r$ and $s \models \varphi$ implies that $r \models \varphi$ for any $\varphi$. Given two cylinders $C_1$ and $C_2$, we say that $C_1$ and $C_2$ are strong bisimilar, written as $C_1 \sim C_2$, iff $|C_1| = |C_2|$, $C_1[i] \sim C_2[i]$ for each $0 \le i \le |C_1|$, and $time(C_1, i) = time(C_2, i)$ for each $0 \le i < |C_1|$. Similarly, we can define strong bisimulation of paths.

As usual we prove the following two things simultaneously:

1. $s \models \varphi$ iff $r \models \varphi$ for any $\varphi$, provided that $s \sim r$;
2. $\omega_1 \models \psi$ iff $\omega_2 \models \psi$ for any $\psi$, provided that $\omega_1 \sim \omega_2$.

We only show the proof for case when $\varphi = \mathcal{P}_{\ge q}(\psi)$ and $\psi = \psi_1 \, \mathsf{U}^I \, \psi_2$, since all the other cases are either trivial or similar. Suppose that $s \models \varphi$ i.e. for all schedulers $\pi$, $Pr_{\pi,s}(Sat(\psi)) \ge q$, we shall prove that $Pr_{\pi,r}(Sat(\psi)) \ge q$ for any scheduler $\pi$ of $r$. According to Lemma 5, the set of paths starting from $s$ and satisfying $\psi$ can be represented by a set of cylinders $Cyls$. By induction hypothesis, $Sat(\psi)$ is $\sim$ closed, thus for any $C \in Cyls$, $[C]_\sim \subseteq Sat(\psi)$. Since for any $C_1, C_2 \in Cyls$ such that $C_1 \cap C_2 \ne \emptyset$, there exists a set of disjoint cylinders $\{C_i'\}$ such that $\cup\{C_i'\} = C_1 \cup C_2$, so any $Cyls$ can be transformed to an equivalent set of disjoint cylinders. In the sequel we assume that $Cyls$ contains only disjoint cylinders, therefore

$$Pr_{\pi,s}(\{\omega \in Paths^\infty \mid \omega \models \psi\}) = \sum_{C \in Cyls} Pr_{\pi,s}(C),$$

for any scheduler $\pi$. As a result, it suffices to prove that for each scheduler $\pi_1$ of $s$, there exists a scheduler $\pi_2$ of $r$ such that $Pr_{\pi_1,s}([C]_\sim) = Pr_{\pi_2,r}([C]_\sim)$ for each $C \in Cyls$. Let $C = C(s_0, I_0, \ldots, I_{n-1}, s_n)$ where $s_0 = s$, we prove by induction on $n$. The base case when $n = 0$ is trivial. Assume that $n > 0$, then according to Eq. (1c), $Pr_{\pi_1,s}([C]_\sim) = Pr_{\pi_1,s}([C]_\sim, 0) =$

$$\int_{t \in I_0} \sum_{(\lambda,\mu) \in tr} \pi(s,0)(\lambda,\mu) \cdot \sum_{s' \in [s_1]_\sim} \mu(s') \cdot \lambda e^{-\lambda t} \cdot Pr_{\pi,s'}([C']_\sim, t) dt,$$

where $tr = Steps(s)$ and $C' = C(s_1, I_1, \ldots, s_n)$. Since $s \sim r$, for each $(\lambda, \mu) \in tr$ there exists $r \xrightarrow{\lambda}_{\mathsf{P}} \nu$ such that $\mu \sim \nu$. Let $\pi_2$ mimic exactly what $\pi_1$ does when at

state $r$. Moreover $Pr_{\pi_1,s'}([C']_\sim, t) = Pr_{\pi_2,r'}([C']_\sim, t)$ for each $C' \in Cyls$ such that $|C'| < n$, provided $s' \sim r'$. By induction hypothesis, such $\pi_2$ always exists, and $Pr_{\pi_1,s}([C]_\sim) = Pr_{\pi_2,r}([C]_\sim)$ for each $C$. Consequently, we have $r \models \varphi$.    □

**Lemma 9.** $s \sim_{CSL^*} r$ implies $s \sim r$ for any $s$ and $r$ i.e. $\sim_{CSL^*} \subseteq \sim$.

*Proof.* First we define a sub-logic of CSL$^*$, called CSL$^\vee$, whose state formulas are the same as CSL$^*$, while its path formulas are defined by the following BNFs:

$$\psi ::= \mathsf{X}^I \varphi \mid \psi \vee \psi,$$

that is, the only path formula of CSL$^\vee$ is the disjunction of several next operators.

Secondly, we prove that $\sim_{CSL^\vee} \subseteq \sim$. Let $\mathcal{R} = \{(s,r) \mid s \sim_{CSL^\vee} r\}$ and $s \mathcal{R} r$, where $\mathcal{R}$ is obviously an equivalence relation. The proof of $L(s) = L(r)$ is trivial and omitted here. It suffices now to prove that for each $s \xrightarrow{\lambda} \mu$, there exists $r \xrightarrow{\lambda}_P \mu'$ such that $\mu \mathcal{R} \mu'$.

*Claim.* Fix a $s \xrightarrow{\lambda} \mu$, there exists $r \xrightarrow{\lambda} \mu'$ such that $\mu(C) = \mu'(C) = 1$ for some $\mathcal{R}$-closed set $C$.

To prove the claim we let $\{\lambda_i \mid r \xrightarrow{\lambda_i} \mu'_i \wedge \mu'_i(C) = 1\}_{1 \le i \le n}$. We proceed by contradiction and assume that there does not exist $i$ such that $\lambda_i = \lambda$. Without loss of generality, we assume that $n = 2$. There are three cases we should consider here:

1.  $\lambda_1 < \lambda_2 < \lambda$. Let $\varphi_C$ be a formula such that $Sat(\varphi_C) = C$, since $C$ is $\mathcal{R}$ closed, $\varphi_C$ always exists. Let $\psi = \mathsf{X}^{[0,b]} \varphi_C$, then the maximal probability of paths starting from $s$ satisfying $\psi$ is equal to $1 - e^{-\lambda \cdot b}$, while the probability for $r$ is $1 - e^{-\lambda \cdot b}$ which is obviously less than $1 - e^{-\lambda_2 \cdot b}$. Therefore there exists $p = 1 - e^{-\lambda_2 \cdot b}$, such that $s \models \mathcal{P}_{\le p}(\psi)$, but $r \not\models \mathcal{P}_{\le p}(\psi)$, which contradicts the assumption that $s \sim_{CSL^\vee} r$.
2.  $\lambda < \lambda_1 < \lambda_2$. This case is similar with the above case and omitted here.
3.  $\lambda_1 < \lambda < \lambda_2$. Let $f(x) = e^{-ax} - e^{-bx}$, then $df/dx = b \cdot e^{-bx} - a \cdot e^{-ax}$. We solve the inequation $df/dx > 0$, and get $x < \ln(b/a)/(b-a)$, which means that if $x_1 < x_2 \le \ln(b/a)/(b-a)$ or $x_1 > x_2 \ge \ln(b/a)/(b-a)$, we have

$$e^{-ax_2} - e^{-bx_2} > e^{-ax_1} - e^{-bx_1}.$$

    Let $a, b$ be two real numbers such that $\lambda = \frac{ln(b/a)}{b-a}$, thus it holds that

$$e^{-\lambda \cdot a} - e^{-\lambda \cdot b} > \max\{e^{-\lambda_1 \cdot a} - e^{-\lambda_1 \cdot b}, e^{-\lambda_2 \cdot a} - e^{-\lambda_2 \cdot b}\}.$$

    Therefore there also exists $p$ such that $s \models \mathcal{P}_{\le p}(\psi)$, but $r \not\models \mathcal{P}_{\le p}(\psi)$, which contradicts the assumption. Thus, we have the claim.

To proceed with the proof of the main theorem, we show that for each $s \xrightarrow{\lambda_1} \mu$, there exists $r \xrightarrow{\lambda_1}_P \mu'$ such that $\mu \mathcal{R} \mu'$. Due to the above proven claim, it is enough to focus on transitions with same rates. We proceed by contradiction, and assume there exists a set of transitions $\{\mu'_i \mid r \xrightarrow{\lambda_1} \mu'_i\}$ with $1 \le i \le n$, but there does not exist $\{w_i \in [0,1]\}$ such that $\mu \mathcal{R} \mu'$ where $\mu' = \sum_{1 \le i \le n} w_i \cdot \mu'_i$. In order to get a contradiction, we need to find a formula $\varphi$ which is satisfied by $s$ but not $r$, or the other way around. We consider the following cases:

1. $s_\perp$ i.e. $s$ is a silent state. This case is impossible since all the derivations of $s$ will stay in the same equivalence class $[s]$, as well as $r$, thus there exists $r \xrightarrow{\lambda_1}_P \nu$ such that $\mu([s]) = \nu([s]) = 1$.

2. $Suc(s) \leq 2$ i.e. there exists at most two equivalence classes $C_1, C_2 \subseteq C$ such that $\mu(C_1 \cup C_2) = 1$, in other words, $\mu(C_1) = 1 - \mu(C_2)$. In case of $Suc(s)$ is a singleton set, we simply set $C_2 = \emptyset$. We consider the following cases:

   (a) $\mu_1'(C_1) \leq \mu_2'(C_1) < \mu(C_1)$. Let $\psi = \mathsf{X}^{[0,\infty)} \varphi_{C_1}$, the maximal probability of paths starting from $r$ satisfying $\psi$ is $\mu(C_1)$, while the maximal probability for $s$ is $\mu_2'(C_1)$ less than $\mu(C_1)$, thus there exists $p$ such that $s \models \mathcal{P}_{\leq p}(\psi)$, but $r \not\models \mathcal{P}_{\leq p}(\psi)$, which contradict the assumption.

   (b) $\mu_2'(C_1) \geq \mu_1'(C_1) > \mu(C_1)$. This case is similar with the case above, and is omitted here.

   (c) $\mu_1'(C_1) \leq \mu(C_1) \leq \mu_2'(C_1)$. In this case we can make sure that there exists $w_1, w_2$ such that $w_1 + w_2 = 1$ and $w_1 \cdot \mu_1'(C_1) + w_2 \cdot \mu_2'(C_1) = \mu(C_1)$, therefore

$$w_1 \cdot \mu_1'(C_2) + w_2 \cdot \mu_2'(C_2) = w_1 \cdot (1 - \mu_1'(C_1)) + w_2 \cdot (1 - \mu_2'(C_1))$$
$$= w_1 + w_2 - (w_1 \cdot \mu_1'(C_1) + w_2 \cdot \mu_2'(C_1))$$
$$= 1 - \mu'(C_1) = \mu'(C_2)$$

   thus $(w_1 \cdot \mu_1' + w_2 \cdot \mu_2') = \mu'$ such that $\mu \mathrel{\mathcal{R}} \mu'$ as we expect. Note this cannot be generalized to the case when $Suc(s) > 2$.

3. We consider the – most involved – remaining case: $Suc(s) > 2$. Note that every combined transition of $r$ can be seen as a combined transition of two other (combined) transitions of $r$. We fix two arbitrary (combined) transitions of $r$: $r \xrightarrow{\lambda_1}_P \mu_1'$ and $r \xrightarrow{\lambda_1}_P \mu_2'$, thus

$$\forall 0 \leq w_1, w_2 \leq 1. \quad w_1 + w_2 = 1 \tag{2}$$
$$\wedge \mu \mathrel{\mathcal{R}} (w_1 \cdot \mu_1' + w_2 \cdot \mu_2').$$

   Let $Supp(\mu) = \{s_1, s_2, \ldots, s_n\}$. For simplicity we assume that $s_1, \ldots, s_n$ belong to different equivalence classes. For $1 \leq i \leq n$, define: $\mu(s_i) = a_i, \mu_1'(s_i) = b_i$, and $\mu_2'(s_i) = c_i$. According to Eq. (2), for each $k$ there must exist $1 \leq j \neq k \leq n$ such that there does not exist $0 \leq w_1, w_2 \leq 1$ with $w_1 + w_2 = 1$ such that $w_1 \cdot b_k + w_2 \cdot c_k = a_k$ and $w_1 \cdot b_j + w_2 \cdot c_j = a_j$, otherwise $\mu \mathrel{\mathcal{R}} (w_1\mu_1' + w_2\mu_2')$ which contradicts Eq. (2). The idea now is then to construct a formula $\varphi$ which is satisfied by $s$ but not $r$. There are several cases to be considered depending on whether $a_k \in [b_k, c_k]$ and/or $a_j \in [b_j, c_j]$. Most of the cases are trivial except when $a_k \in (b_k, c_k)$ and $a_j \in (c_j, b_j)$ with $c_k \geq b_k$ and $b_j \geq c_j$. For instance if $a_k > b_k, c_k$, $s$ will evolve into $s_k$ with higher probability than $r$, so $\varphi$ is easy to give.

   Let $\psi := (\mathsf{X}^{[a,b]} s_j) \vee (\mathsf{X}^{[a',b']} s_k)$, where the names of states are used as abbreviations of the state formulas characterizing the equivalence classes where they are located. Then the probability of paths starting from $s$ satisfying $\psi$ by choosing transition $s \xrightarrow{\lambda_1} \mu$ is equal to $p(s, \mu) := a_j \cdot \rho_2 + a_k \cdot \rho_1$, where $\rho_1 = (e^{-\lambda_1 a} - e^{-\lambda_1 b})$

and $\rho_2 = (e^{-\lambda_1 a'} - e^{-\lambda_1 b'})$. Similarly, the probabilities of paths starting from $r$ satisfying $\psi$ by choosing transitions $r \xrightarrow{\lambda_1} \mu_1'$ and $r \xrightarrow{\lambda_1} \mu_2'$ are equal to $p(r, \mu_1') := b_j \cdot \rho_2 + b_k \cdot \rho_1$ and $p(r, \mu_2') = c_j \cdot \rho_2 + c_k \cdot \rho_1$ respectively.

Now it is sufficient to prove that we can always find $0 \le a \le b$ and $0 \le a' \le b'$ such that $p(s, \mu) > \max\{p(r, \mu_1'), p(r, \mu_2')\}$.

(a) $\frac{b_j - a_j}{a_k - b_k} < \frac{a_j - c_j}{c_k - a_k}$: Let $\frac{\rho_1}{\rho_2} \in (\frac{b_j - a_j}{a_k - b_k}, \frac{a_j - c_j}{c_k - a_k})$, then we have $a_k \cdot \rho_1 + a_j \cdot \rho_2 > \max\{b_k \cdot \rho_1 + b_j \cdot \rho_2, c_k \cdot \rho_1 + c_j \cdot \rho_2\}$ i.e. $p(s, \mu) > \max\{p(r, \mu_1'), p(r, \mu_2')\}$ as we shall prove. Note that $\frac{\rho_1}{\rho_2} = \frac{e^{-\lambda_1 a} - e^{-\lambda_1 b}}{e^{-\lambda_1 a'} - e^{-\lambda_1 b'}}$ ranges over $[0, \infty)$ by choosing different values for $a, b, a'$, and $b'$, therefore the discriminating formula always exists, we get contradiction. The case when $\frac{b_j - a_j}{a_k - b_k} > \frac{a_j - c_j}{c_k - a_k}$ can be proved in a similar way, and is omitted here.

(b) $\frac{b_j - a_j}{a_k - b_k} = \frac{a_j - c_j}{c_k - a_k}$: This case is impossible, otherwise there exists $0 \le w_1, w_2 \le 1$ such that $w_1 \cdot b_k + w_2 \cdot c_k = a_k$ and $w_1 \cdot b_j + w_2 \cdot c_j = a_j$ with $w_1 + w_2 = 1$, simply let $w_1 = \frac{1}{k+1}$ and $w_2 = \frac{k}{k+1}$ where $k = \frac{a_k - b_k}{c_k - a_k}$.

Since CSL$^\vee$ is a sub-logic of CSL$^*$, trivially $\sim_{\text{CSL}^*} \subseteq \sim_{\text{CSL}^\vee}$, therefore $\sim_{\text{CSL}^*} \subseteq \sim$, which completes the proof. □

## B.2  Proof of Lemma 6

*Proof.* Since in Lemma 5 we have shown that for any $s$ and $\psi$, the paths starting from $s$ and satisfying $\psi$ can be represented by a set of disjoint cylinders. It suffices to prove that for each $\pi$ of $s$, $Pr_{\pi,s}(C) = Pr_{\bar{\pi},\bar{s}}(\bar{C})$ for each cylinder $C$, where $\bar{C}$ is a cylinder same as $C$ except that $\bar{C}[i] = C[i]$ for each $0 \le i \le |C|$, and $\bar{\pi}$ is the scheduler mimicking $\pi$ stepwise. Let $C = s_0, I_0, s_1, \ldots, s_n$, we shall prove by induction on $n$. The case when $n = 0$ is trivial, since $Pr_{\pi,s}(C)$ is either 1 or 0 depending on whether $s_0 = s$. Suppose that $n > 0$, $s_0 = s$, and $I_0 = [a, b]$, Since it has been proved in [31, Sec. 6] that uniformization does not change time-bounded reachability, that is, the probability from $s_0$ to $s_1$ in time interval $I$ is equal to the probability from $\bar{s}_0$ to $\bar{s}_1$ in time interval $I$ for any $I$. Let $F(t)$ denote the probability from $s_0$ to $s_1$ in time interval $[0, t]$ given scheduler $\pi$, and $f(t) = \frac{dF(t)}{dt}$, that is, $f(t)$ is the corresponding probability density function, similarly we can define $\bar{F}(t)$ and $\bar{f}(t)$. According to Eq. (1c), $Pr_{\pi,s}(C) = Pr_{\pi,s}(C, 0) = \int_{t \in I_0} f(t) \cdot Pr_{\pi,s}(C, t) dt$ and $Pr_{\bar{\pi},\bar{s}}(C) = Pr_{\bar{\pi},\bar{s}}(C, 0) = \int_{t \in I_0} \bar{f}(t) \cdot Pr_{\bar{\pi},\bar{s}}(C, t) dt$. Since $F(t) = \bar{F}(t)$ for any $t$, we have $f(t) = \bar{f}(t)$ for any $t$. By induction hypothesis, $Pr_{\pi,s}(C, t) = Pr_{\bar{\pi},\bar{s}}(C, t)$ for any $t$, thus

$$f(t) \cdot Pr_{\pi,s}(C, t) = \bar{f}(t) \cdot Pr_{\bar{\pi},\bar{s}}(C, t)$$

for any $t$, which indicates that $Pr_{\pi,s}(C) = Pr_{\bar{\pi},\bar{s}}(C)$. □

## B.3  Proof of Section 5

*Proof.* The proof can be presented as the following chain:

$$s \approx r \overset{Def.\ 6}{\Longleftrightarrow} \bar{s} \sim \bar{r} \overset{Thm.\ 4}{\Longleftrightarrow} \bar{s} \sim_{\text{CSL}^*} \bar{r} \overset{Lem.\ 6}{\Longleftrightarrow} s \sim_{\text{CSL}^*_{\setminus \mathcal{X}}} r.$$

# C   Proofs of Section 5.2

## C.1   Proof of Lemma 7

*Proof.* The proof of $\sim$   $=$   $\sim_{\text{CTMC}}$ is trivial, since in a CTMC there is only one transition for each state, thus we can simply replace $\to_{\text{P}}$ with $\to$. The condition $\lambda_s \cdot \mu_s(C) = \lambda_r \cdot \mu_r(C)$ for each $C$ coincides with the condition: i) $\lambda_s = \lambda_r$, and ii) $\mu_s \, \mathcal{R} \, \mu_r$.

We first prove that $\approx$ implies $\approx_{\text{CTMC}}$. Let $\mathcal{R} = \approx$ and $s \, \mathcal{R} \, r$. We shall prove that $\mathcal{R}$ is a weak bisimulation as defined in Definition 7. Suppose that $s \xrightarrow{\lambda_s} \mu_s$, we need to prove that $r \xrightarrow{\lambda_r} \mu_r$ such that $\lambda_s \cdot \mu_s(C) = \lambda_r \cdot \mu_r(C)$ for all $C \in S/\mathcal{R}$ with $C \neq [s]_{\mathcal{R}} = [r]_{\mathcal{R}}$. According to Definition 6, $s \approx r$ if $\bar{s} \sim \bar{r}$. By Definition 2, if $s \xrightarrow{\lambda_s} \mu_s$, then $\bar{s} \xrightarrow{E} \mu$ such that $\mu = \frac{E - \lambda_s}{E} \cdot \mathcal{D}_{\bar{s}} + \frac{\lambda_s}{E} \cdot \bar{\mu}_s$ where $\bar{\mu}_s$ is defined as expected. Therefore there exists $\bar{r} \xrightarrow{E} \nu$ such that $\mu \sim \nu$ where $\nu = \frac{E - \lambda_r}{E} \cdot \mathcal{D}_{\bar{r}} + \frac{\lambda_r}{E} \cdot \bar{\mu}_r$. Obviously if there exists $C \in S/\mathcal{R}$ with $C \neq [s]_{\mathcal{R}} = [r]_{\mathcal{R}}$ such that $\lambda_s \cdot \mu_s(C) \neq \lambda_r \cdot \mu_r(C)$, then $\mu(\bar{C}) \neq \nu(\bar{C})$ since $\mu(\bar{C}) = \frac{\lambda_s}{E} \cdot \mu_s(C)$ and $\nu(\bar{C}) = \frac{\lambda_r}{E} \cdot \mu_r(C)$, thus it is impossible for $\mu \sim \nu$.

To show that $\approx_{\text{CTMC}}$ implies $\approx$, it is enough to show that $\mathcal{R} = \approx_{\text{CTMC}}$ is a weak bisimulation according to Definition 6, that is, we need show that $\mathcal{R} = \{(\bar{s}, \bar{r}) \mid s \approx_{\text{CTMC}} r\}$ is a strong bisimulation by Definition 4. Suppose that $\bar{s} \xrightarrow{E} \mu$, then there exists $s \xrightarrow{\lambda_s} \mu_s$ such that $\mu = \frac{E - \lambda_s}{E} \cdot \mathcal{D}_{\bar{s}} + \frac{\lambda_s}{E} \cdot \bar{\mu}_s$. Since $s \approx_{\text{CTMC}} r$, there exists $r \xrightarrow{\lambda_r} \mu_r$ such that $\lambda_s \cdot \mu_s(C) = \lambda_r \cdot \mu_r(C)$ for all equivalence class $C \neq [s]_{\approx_{\text{CTMC}}} = [r]_{\approx_{\text{CTMC}}}$. Therefore there exists $\bar{r} \xrightarrow{E} \nu$ such that $\nu = \frac{E - \lambda_r}{E} \cdot \mathcal{D}_{\bar{r}} + \frac{\lambda_r}{E} \cdot \bar{\mu}_r$ and $\mu(\bar{C}) = \nu(\bar{C})$ for all equivalence class $\bar{C} \neq [\bar{s}]_{\mathcal{R}} = [\bar{r}]_{\mathcal{R}}$, since $\mu(\bar{C}) = \frac{\lambda_s}{E} \cdot \mu_s(C)$ and $\nu(\bar{C}) = \frac{\lambda_r}{E} \cdot \mu_r(C)$ i.e. $\mu \, \mathcal{R} \, \nu$.    □