Verification of Open Interactive Markov Chains

Tomáš Brázdil[1], Holger Hermanns[2], Jan Krčál[1], Jan Křetínský[1,3], and Vojtěch Řehák[1]

1 **Faculty of Informatics, Masaryk University, Czech Republic**
{brazdil,krcal,jan.kretinsky,rehak}@fi.muni.cz
2 **Saarland University – Computer Science, Saarbrücken, Germany**
hermanns@cs.uni-saarland.de
3 **Institut für Informatik, Technical University Munich, Germany**

Abstract

Interactive Markov chains (IMC) are compositional behavioral models extending both labeled transition systems and continuous-time Markov chains. IMC pair modeling convenience - owed to compositionality properties - with effective verification algorithms and tools - owed to Markov properties. Thus far however, IMC verification did not consider compositionality properties, but considered closed systems. This paper discusses the evaluation of IMC in an open and thus compositional interpretation. For this we embed the IMC into a game that is played with the environment. We devise algorithms that enable us to derive bounds on reachability probabilities that are assured to hold in any composition context.

1998 ACM Subject Classification D.4.8 Performance

Keywords and phrases IMC, compositional verification, synthesis, time bounded reachability, discretization

Digital Object Identifier 10.4230/LIPIcs.xxx.yyy.p

1 Introduction

With the increasing complexity of systems and software reuse, component based development concepts gain more and more attention. In this setting developers are often facing the need to develop a component with only partial information about the surrounding components at hand, especially when relying on third-party components to be inter-operated with. This motivates verification approaches that ensure the functionality of a component in an environment whose behavior is unknown or only partially known. *Compositional verification* approaches aim at methods to prove guarantees on isolated components in such a way that when put together, the entire system's behavior has the desired properties based on the individual guarantees.

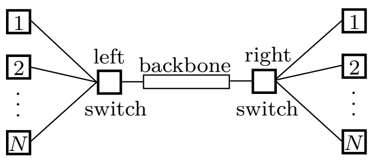
The assurance of reliable functioning of a system relates not only to its correctness, but also to its performance and dependability. This is a major concern especially in embedded system design. A natural instantiation of the general component-based approach in the continuous-time setting are interactive Markov chains [23]. Interactive Markov chains (IMC) are equipped with a sound compositional theory. IMC arise from classical labeled transition systems by incorporating the possibility to change state according to a random delay governed by some negative exponential distribution. This twists the model to one that is running in continuous real time. State transitions may be triggered by delay expirations, or may be triggered by the execution of actions. By dropping the new type of transitions, labeled transition systems are regained in their entirety. By dropping action-labeled transitions instead, one arrives at one of the simplest but also most widespread class of performance and dependability models, *continuous-time Markov chains* (CTMCs). IMC have a well-understood compositional theory, rooted in process algebra [3], and are in use as semantic backbones for dynamic fault trees [6], architectural description languages [5, 8], generalized stochastic Petri nets [24] and Statemate [4] extensions, and are applied in a large spectrum of practical applications, ranging from networked hardware on chips [15] to water treatment facilities [20] and ultra-modern satellite designs [16].

In recent years, various analysis techniques have been proposed [18, 26, 22, 25, 31, 19] for IMC. The pivotal verification problem considered is that of *time-bounded reachability*. It is the problem to calculate or approximate the probability that a given state (set) is reached within a given deadline. However, despite the fact that IMC support compositional model generation minimization very well, the analysis techniques considered thus far are not compositional. They are all bound to the assumption that the analyzed IMC is closed, i.e. does not depend on interaction with the environment. Technically, this is related to the *maximal-progress assumption* governing the interplay of delay and action execution of an IMC component: Internal actions are assumed to happen instantaneously and therefore take precedence over delay transitions while external actions do not. External actions are the process algebraic means for interaction with other components. Remarkably, in all the published IMC verification approaches, all occurring actions are assumed to be internal (respectively internalized by means of a hiding operator prior to analysis).

In this paper, we instead consider *open* IMC, where the control over external actions is in the hands of and possibly delayed by an environment. The environment can be thought of as summarizing the behavior of one or several interacting components. As a consequence, we find ourselves in the setting of a timed game, where the environment has the (timed) control over external actions, while the IMC itself controls choices over internal actions. The resulting game turns out to be remarkably difficult, owed to the interplay of timed moves with external and internal moves of both players.

Concretely, assume we are given an IMC $\mathcal{C}$ which contains some internal non-deterministic transitions and also offers some external actions for synchronization to an unknown environment. Our goal is to synthesize a scheduler controlling the internal transitions which maximizes the probability of reaching a set G of goal states, in time T no matter what and when the environment E decides to synchronize with the external actions. The environment E ranges over all possible IMC able to synchronize with the external actions of $\mathcal{C}$.

To get a principal understanding of the complications faced, we need to consider a restricted setting, where $\mathcal{C}$ does not enable internal and external transitions at the same state. We provide an algorithm which approximates the probability in question up to a given precision $\varepsilon > 0$ and also computes an ε-optimal scheduler. The algorithm consists of two steps. First, we reduce the problem to a game where the environment is not an IMC but can decide to execute external actions at *non-deterministically* chosen time instances. In a second step, we solve the resulting game on $\mathcal{C}$ using discretization. Our discretization is based on the same approach as the algorithm of [31]. However, the algorithm as well as its proof of correctness is considerably more complicated due to presence of non-deterministic choices of the second player. We finally discuss what happens if we allow internal and external transitions to be enabled at the same time.



Example. To illustrate the concepts by an example application, we can consider a variant of the *fault-tolerant workstation cluster* [21] depicted on the right. The overall system consists of two sub-clusters connected via a backbone; each of them contains N workstations.

Any component can fail and then needs to be repaired to become operational again. There is a single repair unit (not depicted) which must take decisions what to repair next when multiple components are failed. The entire system can be modelled using the IMC composition operators [21], but we are now also in the position to study a partial model, where some components, such as one of the switches, are left unspecified. We seek for the optimal repair schedule regardless of how the unknown components are implemented. It can answer questions such as: *"What is the worst case probability to hit a state in which premium service is not guaranteed within T time units?"* with premium service only being guaranteed if there are at least N operational workstations connected to each other via operational switches.

Our contribution. We investigate the problem of compositionally verifying open IMC. In particular, we introduce the problem of synthesizing optimal control for time-bounded reachability in an IMC interacting in an unknown environment, provided no state enables internal and external transition. Thereafter, we solve the problem of finding ε-optimal schedulers using the established method of discretization, give bounds on the size of the game to be solved for a given ε and thus establish upper complexity bound for the problem.

Related work. Model checking of *open* systems has been proposed in [27]. The synthesis problem is often stated as a *game* where the first player controls a component and the second player simulates an environment [30]. There is a large body of literature on games in verification, including recent surveys [1, 13]. *Stochastic* games have been applied to e.g. concurrent program synthesis [12] and for collaboration strategies among compositional stochastic systems [14]. Although most papers deal with discrete time games, lately games with stochastic *continuous-time* have gained attention [7, 29, 9, 10]. Technically, some of the games we consider in the present paper exploit special cases of the games considered in [7, 10]. However, both papers prove decidability only for qualitative reachability problems and do not discuss compositionality issues.

The *time-bounded reachability* problem for closed IMC has been studied in [22, 31] and compositional abstraction techniques to compute it are developed in [25]. In the closed interpretation, IMC have some similarities with continuous-time Markov decision processes, CTMDPs. Algorithms for time-bounded reachability in CTMDPs and corresponding games are developed in [2, 9, 29]. A numerically stable algorithm for time-bounded properties for CTMDPs is developed in [11].

2 Interactive Markov Chains

In this section, we introduce the formalism of interactive Markov chains together with the standard way to compose them. After giving the operational interpretation for closed systems, we define the fundamental problem of our interest, namely we define the value of time-bounded reachability and introduce the studied problems.

We denote by $\mathbb{N}$, $\mathbb{N}_0$, $\mathbb{R}_{>0}$, and $\mathbb{R}_{\geq 0}$ the sets of natural numbers, natural numbers with zero, positive real numbers and non-negative real numbers, respectively.

▶ **Definition 1** (IMC)**.** An interactive Markov chain (IMC) is a tuple $\mathcal{C} = (S, \mathbb{A}\mathrm{ct}^\tau, \hookrightarrow, \rightsquigarrow, s_0)$ where S is a finite set of *states*, $\mathbb{A}\mathrm{ct}^\tau$ is a finite set of *actions* containing a designated *internal action* τ, $s_0 \in S$ is an *initial* state,

- $\hookrightarrow \subseteq S \times \mathbb{A}\mathrm{ct}^\tau \times S$ is an *interactive transition* relation, and
- $\rightsquigarrow \subseteq S \times \mathbb{R}_{>0} \times S$ is a *Markovian transition* relation.

Elements of $\mathbb{A}\mathrm{ct} := \mathbb{A}\mathrm{ct}^\tau \setminus \{\tau\}$ are called *external actions.* We write $s \overset{a}{\hookrightarrow} t$ whenever $(s, a, t) \in \hookrightarrow$, and further $\mathrm{succ}_e(s) = \{t \in S \mid \exists a \in \mathbb{A}\mathrm{ct} : s \overset{a}{\hookrightarrow} t\}$ and $\mathrm{succ}_\tau(s) = \{t \in S \mid$

$s \overset{\tau}{\hookrightarrow} t\}$. Similarly, we write $s \overset{\lambda}{\rightsquigarrow} t$ whenever $(s, \lambda, t) \in \rightsquigarrow$ where λ is called a *rate* of the transition, and $\mathrm{succ}_M(s) = \{t \in S \mid \exists \lambda : s \overset{\lambda}{\rightsquigarrow} t\}$. We assume w.l.o.g. that for each pair of states s and t, there is at most one Markovian transition from s to t. We say that an external, or internal, or Markovian transition is available in s if $\mathrm{succ}_e(s) \neq \emptyset$, or $\mathrm{succ}_\tau(s) \neq \emptyset$, or $\mathrm{succ}_M(s) \neq \emptyset$, respectively.

We also define a *total exit rate* function $\mathbf{E} : S \to \mathbb{R}_{\geq 0}$ which assigns to each state the sum of rates of all outgoing Markovian transitions, i.e. $\mathbf{E}(s) = \sum_{s \overset{\lambda}{\rightsquigarrow} t} \lambda$ where the sum is zero if $\mathrm{succ}_M(s)$ is empty. Furthermore, we define a probability matrix $\mathbf{P}(s,t) = \lambda / \mathbf{E}(s)$ if $\mathbf{E}(s) \neq 0$ and $s \overset{\lambda}{\rightsquigarrow} t$; and $\mathbf{P}(s,t) = 0$, otherwise.

IMC are well suited for compositional modeling, where systems are built out of smaller ones using composition operators. Parallel composition and hiding operators are central to a modeling style, where parallel components synchronize using shared action, and further synchronization can be prohibited by hiding (i.e. internalizing) some actions. IMC employ the *maximal progress assumption*: Internal actions take precedence over the advance of time [23].

▶ **Definition 2** (Parallel composition). For IMC $\mathcal{C}_1 = (S_1, \mathbb{A}\mathrm{ct}_1^\tau, \hookrightarrow_1, \rightsquigarrow_1, s_{01})$ and $\mathcal{C}_2 = (S_2, \mathbb{A}\mathrm{ct}_2^\tau, \hookrightarrow_2, \rightsquigarrow_2, s_{02})$ and a *synchronization alphabet* $A \subseteq \mathbb{A}\mathrm{ct}_1 \cap \mathbb{A}\mathrm{ct}_2$, the parallel composition $\mathcal{C}_1 \parallel_A \mathcal{C}_2$ is the IMC $\mathcal{C}_1 = (S_1 \times S_2, \mathbb{A}\mathrm{ct}_1^\tau \cup \mathbb{A}\mathrm{ct}_2^\tau, \hookrightarrow, \rightsquigarrow, (s_{01}, s_{02}))$ where $\hookrightarrow$ and $\rightsquigarrow$ are defined as the smallest relations satisfying

- $s_1 \overset{a}{\hookrightarrow} s_1'$ and $s_2 \overset{a}{\hookrightarrow} s_2'$ and $a \in A$ implies $(s_1, s_2) \overset{a}{\hookrightarrow} (s_1', s_2')$,
- $s_1 \overset{a}{\hookrightarrow} s_1'$ and $a \notin A$ implies $(s_1, s_2) \overset{a}{\hookrightarrow} (s_1', s_2)$ for each $s_2 \in S_2$,
- $s_2 \overset{a}{\hookrightarrow} s_2'$ and $a \notin A$ implies $(s_1, s_2) \overset{a}{\hookrightarrow} (s_1, s_2')$ for each $s_1 \in S_1$,
- $s_1 \overset{\lambda}{\rightsquigarrow} s_1'$ implies $(s_1, s_2) \overset{\lambda}{\rightsquigarrow} (s_1', s_2)$ for each $s_2 \in S_2$, and
- $s_2 \overset{\lambda}{\rightsquigarrow} s_2'$ implies $(s_1, s_2) \overset{\lambda}{\rightsquigarrow} (s_1, s_2')$ for each $s_1 \in S_1$.

▶ **Definition 3** (Hiding). For an IMC $\mathcal{C} = (S, \mathbb{A}\mathrm{ct}^\tau, \hookrightarrow, \rightsquigarrow, s_0)$ and a *hidden alphabet* $A \subseteq \mathbb{A}\mathrm{ct}$, the hiding $\mathcal{C} \setminus A$ is the IMC $(S, \mathbb{A}\mathrm{ct}^\tau \setminus A, \hookrightarrow', \rightsquigarrow, s_0)$ where $\hookrightarrow'$ is the smallest relation satisfying for each $s \overset{a}{\hookrightarrow} s'$ that $a \in A$ implies $s \overset{\tau}{\hookrightarrow}' s'$, and $a \notin A$ implies $s \overset{a}{\hookrightarrow}' s'$.

The analysis of IMC has thus far been restricted to *closed* IMC [18, 26, 22, 25, 31, 19]. In a closed IMC, external actions do not appear as transition labels (i.e. $\hookrightarrow \subseteq S \times \{\tau\} \times S$). In practice, this is achieved by an outermost hiding operator $\setminus \mathbb{A}\mathrm{ct}$ closing the composed system. Non-determinism among internal τ transitions is resolved using a (history-dependent) scheduler σ [31].

Let us fix a *closed* IMC $\mathcal{C} = (S, \mathbb{A}\mathrm{ct}^\tau, \hookrightarrow, \rightsquigarrow, s_0)$. The IMC $\mathcal{C}$ under a scheduler σ moves from state to state, and in every state may wait for a random time. This produces a *run* which is an infinite sequence of the form $s_0\, t_0\, s_1\, t_1 \cdots$ where s_n is the n-th visited state and t_n is the time spent there. After n steps, the scheduler resolves the non-determinism based on the *history* $\mathfrak{h} = s_0\, t_0 \cdots s_{n-1}\, t_{n-1}\, s_n$ as follows.

▶ **Definition 4** (Scheduler). A scheduler[1] for an IMC $\mathcal{C} = (S, \mathbb{A}\mathrm{ct}^\tau, \hookrightarrow, \rightsquigarrow, s_0)$ is a measurable[2] function $\sigma : (S \times \mathbb{R}_{\geq 0})^* \times S \to S$ such that for each history $\mathfrak{h} = s_0\, t_0\, s_1 \cdots s_n$ with $\mathrm{succ}_\tau(s_n) \neq \emptyset$ we have $\sigma(\mathfrak{h}) \in \mathrm{succ}_\tau(s_n)$. The set of all schedulers for $\mathcal{C}$ is denoted by $\mathfrak{S}(\mathcal{C})$.

The decision of the scheduler $\sigma(\mathfrak{h})$ determines t_n and s_{n+1} as follows. If $\mathrm{succ}_\tau(s_n) \neq \emptyset$, then the run proceeds immediately, i.e. in time $t_n := 0$, to the state $s_{n+1} := \sigma(\mathfrak{h})$. Otherwise,

[1] For the sake of simplicity, we only consider deterministic schedulers in this paper.
[2] More precisely, $\sigma^{-1}(s)$ is measurable in the product topology of the discrete topology on S and the Borel topology on $\mathbb{R}_{\geq 0}$.

if $\text{succ}_\tau(s_n) = \emptyset$, then only Markovian transitions are available in s_n. In such a case, the run moves to a randomly chosen next state s_{n+1} with probability $\mathbf{P}(s_n, s_{n+1})$ after waiting for a random time t_n chosen according to the exponential distribution with the rate $\mathbf{E}(s_n)$.

One of the fundamental problems in verification and performance analysis of continuous-time stochastic systems is the time-bounded reachability. Given a set of goal states $G \subseteq S$ and a time bound $T \in \mathbb{R}_{\geq 0}$, the *value of time-bounded reachability* is defined as $\sup_{\sigma \in \mathfrak{S}(\mathcal{C})} \mathcal{P}^\sigma_\mathcal{C}\left[\Diamond^{\leq T} G\right]$ where $\mathcal{P}^\sigma_\mathcal{C}\left[\Diamond^{\leq T} G\right]$ denotes the probability that a run of $\mathcal{C}$ under the scheduler σ visits a state of G before time T. The pivotal problem in the algorithmic analysis of IMC is to compute this value together with a scheduler which achieves the supremum. As the value is not rational in most cases, the aim is to provide an efficient approximation algorithm and compute an ε-optimal scheduler. The value of time-bounded reachability can be approximated up to a given error tolerance $\varepsilon > 0$ in time $\mathcal{O}(|S|^2 \cdot (\lambda T)^2/\varepsilon)$ [28], where λ is the maximal rate of $\mathcal{C}$, and the procedure also yields an ε-optimal scheduler. We generalize both the notion of the value as well as approximation algorithms to the setting of *open* IMC, i.e. those that are not closed, and motivate this extension in the next section.

3 Compositional Verification

In this section we turn our attention to the central questions studied in this paper. How can we decide how well an IMC component $\mathcal{C}$ performs (w.r.t. time-bounded reachability) when acting in parallel with an unknown environment? And how to control the component to establish a guarantee as high as possible?

Speaking thus far in vague terms, this amounts to finding a scheduler σ for $\mathcal{C}$ which maximizes the probability of reaching a target set G before T no matter what environment E is composed with $\mathcal{C}$. As we are interested in compositional modeling using IMC, the environments are supposed to be IMC with the same external actions as $\mathcal{C}$ (thus resolving the external non-determinism of $\mathcal{C}$). We also need to consider all resolutions of the internal non-determinism of E as well as the non-determinism arising from synchronization of $\mathcal{C}$ and E using another scheduler π. So we are interested in the following value:

$$\sup_{\sigma} \inf_{E,\pi} \mathcal{P}[G \text{ is reached in composition of } \mathcal{C} \text{ and } E \text{ before } T \text{ using } \sigma \text{ and } \pi].$$

Now, let us be more formal and fix an IMC $\mathcal{C} = (S, \mathbb{A}\text{ct}^\tau, \hookrightarrow, \rightsquigarrow, s_0)$. For a given environment IMC E with the same action alphabet $\mathbb{A}\text{ct}^\tau$, we introduce a composition $\mathcal{C}(E) = (\mathcal{C} \parallel_{\mathbb{A}\text{ct}} E) \setminus \mathbb{A}\text{ct}$ where all open actions are hidden, yielding a closed system. Note that the states of $\mathcal{C}(E)$ are pairs (c, e) where c is a state of $\mathcal{C}$ and e is a state of E. We consider a scheduler σ of $\mathcal{C}$ and a scheduler π of $\mathcal{C}(E)$ respecting σ on internal actions of $\mathcal{C}$. We say that π *respects* σ, denoted by $\pi \in \mathfrak{S}(\mathcal{C}(E), \sigma)$, if for every history $\mathfrak{h} = (c_0, e_0)\, t_0 \cdots t_{n-1} (c_n, e_n)$ of $\mathcal{C}(E)$ the scheduler π satisfies one of the following conditions:

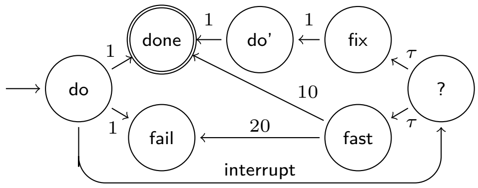
- $\pi(\mathfrak{h}) = (c, e)$ where $c_n \overset{a}{\hookrightarrow} c$ and $e_n \overset{a}{\hookrightarrow} e$ (π resolves synchronization)
- $\pi(\mathfrak{h}) = (c_n, e)$ where $e_n \overset{\tau}{\hookrightarrow} e$ (π chooses a move in the environment)
- $\pi(\mathfrak{h}) = (\sigma(\mathfrak{h}_\mathcal{C}), e_n)$ where $\mathfrak{h}_\mathcal{C} = c_0 t_0 \cdots t_{n-1} c_n$ (π chooses a move in $\mathcal{C}$ according to σ).

Given a set of goal states $G \subseteq S$ and a time bound $T \in \mathbb{R}_{\geq 0}$, the *value of compositional time-bounded reachability* is defined as

$$\sup_{\sigma \in \mathfrak{S}(\mathcal{C})} \inf_{\substack{E \in \text{ENV} \\ \pi \in \mathfrak{S}(\mathcal{C}(E),\sigma)}} \mathcal{P}^\pi_{\mathcal{C}(E)}\left[\Diamond^{\leq T} G_E\right] \qquad (*)$$

where ENV denotes the set of all IMC with the action alphabet $\mathbb{A}\mathrm{ct}^\tau$ and $G_E = G \times S_E$ where S_E is the set of states of E. As for the closed IMC, our goal is to efficiently approximate this value together with a maximizing scheduler. Before we present an approximation algorithm based on discretization, we illustrate some of the effects of the open system perspective.

Example. The figure on the right depicts an IMC on which we approximate the value $(*)$ for $T = 2$ and $G = \{\mathsf{done}\}$. From the initial state do, the system may go randomly either to the target done or to state fail. Concurrently, an external action interrupt may switch the run to state ?, where the scheduler σ chooses between two successors (1) state fast allowing fast but risky run to the target and (2) state fix that guarantees reaching the target but takes longer time. The value $(*)$ is approximately 0.47 and the optimal scheduler goes to fix only if there are more than 1.2 minutes left. Note that the probability of reaching the target in time depends on when the external action interrupt is taken. The most adversarial "scheduler" of the environment executes interrupt after 0.8 minutes from the start.



Results. We now formulate our main result concerning efficient approximation of the value of compositional time-bounded reachability. In fact, we provide an approximation algorithm for a restricted subclass of IMC defined by the following two assumptions:

▶ **Assumption 1.** *Each cycle contains a Markovian transition.*

This assumption is standard over all analysis techniques published for IMC [18, 26, 22, 25, 31, 19]. It implies that the probability of taking infinitely many transitions in finite time, i.e. of Zeno behavior, is zero. This is a rather natural assumption and does not restrict the modeling power much, since no real system will be able to take infinitely many transitions in finite time anyway. Furthermore, the assumed property is a compositional one, i.e. it is preserved by parallel composition and hiding.

▶ **Assumption 2.** *Internal and external actions are not enabled at the same time, i.e. for each state s, either* $\mathrm{succ}_e(s) = \emptyset$ *or* $\mathrm{succ}_\tau(s) = \emptyset$.

Note that both assumptions are met by the above mentioned example. However, Assumption 2 is not compositional; specifically, it is not preserved by applications of the hiding operator. A stronger assumption would require the environment not to trigger external actions in zero time after a state change. This is indeed implied by Assumption 2 which basically asks *internal* transitions of the component to be executed before any *external* actions are taken into account.[3] In fact, the reverse precedence cannot be implemented in real systems, if internal actions are assumed to be executed without delay. Any procedure implemented in $\mathcal{C}$ for checking the availability of external actions will involve some non-zero delay (unless one resorts to quantum effects). From a technical point of view, lifting Assumption 2 makes the studied problems considerably more involved; see Section 6 for further discussion.

▶ **Theorem 5.** *Let $\varepsilon > 0$ be an approximation bound and $\mathcal{C} = (S, \mathbb{A}\mathrm{ct}^\tau, \hookrightarrow, \rightsquigarrow, s_0)$ be an IMC satisfying Assumptions 1 and 2. Then one can approximate the value of compositional time-bounded reachability of $\mathcal{C}$ up to ε and compute an ε-optimal scheduler in time $\mathcal{O}(|S|^2 \cdot (\lambda T)^2/\varepsilon)$, where λ is the maximal rate of $\mathcal{C}$ and T is the reachability time-bound.*

[3] To see this one can construct a weak simulation relation between a system violating Assumption 2 and one satisfying it, where any state with both internal and external transitions is split into two: the first one enabling the internal transitions and a new τ to the second one only enabling the external ones.

In the remainder of the paper, we prove this theorem and discuss its restrictions. First, we introduce a new kind of real-time games, called CE games, that are played on open IMC. Then we reduce the compositional time-bounded reachability of $\mathcal{C}$ to time-bounded reachability objective in the CE game played just on the component $\mathcal{C}$ (see Proposition 6). In Section 5, we show how to reduce, using discretization, the time-bounded reachability in CE games to step-bounded reachability in discrete-time stochastic games (see Proposition 8), that in turn can be solved using simple backward propagation. Finally, we show, in Proposition 9, how to transform optimal strategies in the discretized stochastic games to ε-optimal schedulers for $\mathcal{C}$.

4 Game of Controller and Environment

In order to approximate ($*$), the value of compositional time-bounded reachability, we turn the IMC $\mathcal{C}$ into a two-player *controller–environment game* (CE game) $\mathcal{G}$. The CE game naturally combines two approaches to real-time systems, namely the stochastic flow of time as present in CTMCs with the non-deterministic flow of time as present in timed automata. The game $\mathcal{G}$ is played on the graph of an IMC $\mathcal{C}$ played by two players: **con** (controlling the component $\mathcal{C}$) and **env** (controlling/simulating the environment). In essence, **con** chooses in each state with internal transitions one of them, and **env** chooses in each state with external (and hence synchronizing) transitions either which of them should be taken, or a delay $t_e \in \mathbb{R}_{>0}$. Note that, due to Assumption 2, the players control the game in disjoint sets of states, hence $\mathcal{G}$ is a turn-based game. The internal and external transitions take zero time to be execute once chosen. If no zero time transition is chosen, the delay t_e determined by **env** competes with the Markovian transitions, i.e. with a random time sampled from the exponential distribution with the rate $\mathbf{E}(s)$. We consider time-bounded reachability objective, so the goal of **con** is to reach a given subset of states G before a given time T, and **env** opposes it.

Formally, let us fix an IMC $\mathcal{C} = (S, \mathbb{A}\mathrm{ct}^\tau, \hookrightarrow, \rightsquigarrow, s_0)$ and thus a CE game $\mathcal{G}$. A *run* of $\mathcal{G}$ is again an infinite sequence $s_0\, t_0\, s_1\, t_1 \cdots$ where $s_n \in S$ is the n-th visited state and $t_n \in \mathbb{R}_{\geq 0}$ is the time spent there. Based on the *history* $s_0\, t_0 \cdots t_{n-1}\, s_n$ went through so far, the players choose their moves as follows.

- If $\mathrm{succ}_\tau(s_n) \neq \emptyset$, the player **con** chooses a state $s_\tau \in \mathrm{succ}_\tau(s_n)$.
- Otherwise, the player **env** chooses either a state $s_e \in \mathrm{succ}_e(s_n)$, or a delay $t_e \in \mathbb{R}_{>0}$. (If $\mathrm{succ}_e(s_n) = \emptyset$ only a delay can be chosen.)

Subsequently, Markovian transitions (if available) are resolved by randomly choosing a target state s_M according to the distribution $\mathbf{P}(s_n, \cdot)$ and randomly sampling a time t_M according to the exponential distribution with rate $\mathbf{E}(s_n)$. The next waiting time t_n and state s_{n+1} are given by the following rules in the order displayed.

- If $\mathrm{succ}_\tau(s_n) \neq \emptyset$ and s_τ was chosen, then $t_n = 0$ and $s_{n+1} = s_\tau$.
- If $\mathrm{succ}_e(s_n) \neq \emptyset$ and s_e was chosen, then $t_n = 0$ and $s_{n+1} = s_e$.
- If t_e was chosen then:
  - if $\mathrm{succ}_M(s_n) = \emptyset$, then $t_n = t_e$ and $s_{n+1} = s_n$;
  - if $t_e \leq t_M$, then $t_n = t_e$ and $s_{n+1} = s_n$;
  - if $t_M < t_e$, then $t_n = t_M$ and $s_{n+1} = s_M$.

According to the definition of schedulers in IMC, we formalize the choice of **con** as a *strategy* $\sigma : (S \times \mathbb{R}_{\geq 0})^* \times S \to S$ and the choice of **env** as a strategy $\pi : (S \times \mathbb{R}_{\geq 0})^* \times S \to S \cup \mathbb{R}_{>0}$. We denote by Σ and Π the sets of all strategies of the players **con** and **env**, respectively. In order to keep CE games out of Zeno behavior, we consider in Π only those strategies of the player **env** for which the induced Zeno runs have zero measure, i.e. the sum of the chosen delays diverges almost surely no matter what **con** is doing.

Given a set of goal states $G \subseteq S$ and a time bound $T \in \mathbb{R}_{\geq 0}$, the *value of* $\mathcal{G}$ is defined as

$$\sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}\left[\Diamond^{\leq T} G\right] \qquad (**)$$

where $\mathcal{P}_{\mathcal{G}}^{\sigma,\pi}\left[\Diamond^{\leq T} G\right]$ is the probability of all runs of $\mathcal{G}$ induced by σ and π and reaching a state of G before time T. We now show that the value of CE game coincides with the value of compositional time-bounded reachability.

▶ **Proposition 6.** $(*) = (**)$, *i.e.*

$$\sup_{\sigma \in \mathfrak{S}(\mathcal{C})} \inf_{\substack{E \in \mathrm{ENV} \\ \pi \in \mathfrak{S}(\mathcal{C}(E),\sigma)}} \mathcal{P}_{\mathcal{C}(E)}^{\pi}\left[\Diamond^{\leq T} G_E\right] = \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}\left[\Diamond^{\leq T} G\right]$$

Proof Idea. We start with the inequality $(*) \geq (**)$. Let $\sigma \in \Sigma$ $(= \mathfrak{S}(\mathcal{C}))$ and let us fix an environment E together with a scheduler $\pi \in \mathfrak{S}(\mathcal{C}(E), \sigma)$. The crucial observation is that the *only* purpose of the environment E (controlled by π) is to choose delays of external actions (the delay is determined by a sequence of internal and Markovian actions of E executed before the external action), which is in fact similar to the role of the player **env** in the CE game. The only difference is that the environment E "chooses" the delays randomly as opposed to deterministic strategies of **env**. However, using a technically involved argument, we show how to get rid of this randomization and obtain a strategy π' in the CE game satisfying $\mathcal{P}_{\mathcal{G}}^{\sigma,\pi'}\left[\Diamond^{\leq T} G\right] \leq \mathcal{P}_{\mathcal{C}(E)}^{\pi}\left[\Diamond^{\leq T} G_E\right]$.

Concerning the second inequality $(*) \leq (**)$, we show that every strategy of **env** can be (approximately) implemented using a suitable environment together with a scheduler π. The idea is to simulate every deterministic delay, say t, chosen by **env** using a random delay tightly concentrated around t (roughly corresponding to an Erlang distribution) that is implemented as an IMC. We show that the imprecision of delays introduced by this randomization induces only negligible alteration to the value. ◀

5 Discretization

In this section we show how to approximate the value $(**)$ of the CE game up to an arbitrarily small error $\varepsilon > 0$ by reduction to a discrete-time (turn-based) stochastic game Δ.

A stochastic game Δ is played on a graph $(V, \mapsto)$ partitioned into $V_\Box \uplus V_\Diamond \uplus V_\bigcirc$. A play starts in the initial vertex v_0 and forms a run $v_0 v_1 \cdots$ as follows. For a history $v_0 \cdots v_i$, the next vertex v_{i+1} satisfying $v_i \mapsto v_{i+1}$ is determined by a strategy $\sigma \in \Sigma_\Delta$ of player $\Box$ if $v_i \in V_\Box$ and by a strategy $\pi \in \Pi_\Delta$ of player $\Diamond$ if $v_i \in V_\Diamond$. Moreover, v_{i+1} is chosen randomly according to a fixed distribution $Prob(v_i)$ if $v_i \in V_\bigcirc$. For a formal definition, see, e.g., [17].

Let us fix a CE game $\mathcal{G}$ and a discretization step $\delta > 0$ that divides the time bound T into $N \in \mathbb{N}$ intervals of equal length (here $\delta = T/N$). We construct a discrete-time stochastic game Δ by substituting each state of $\mathcal{G}$ by a gadget of one or two vertices (as illustrated in Figure 1).[4] Intuitively, the game Δ models passing of time as follows. Each discrete step "takes" either time δ or time 0. Each step from a vertex of $V_\bigcirc$ takes time δ whereas each step

[4] We assume w.l.o.g. that (1) states with internal transitions have no Markovian transitions available and (2) every state has at least one outgoing transition.This is no restriction since (1) Markovian transitions are never taken in such states and (2) any state without transitions can be endowed with a Markovian self-loop transition without changing the time-bounded reachability.

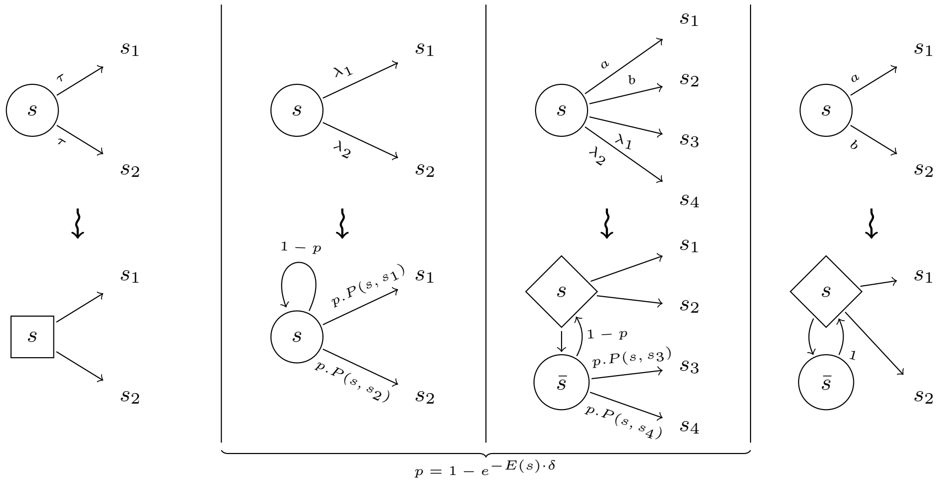


Figure 1 Four gadgets for transforming a CE game into a discrete game. The upper part shows types of states in the original CE game, the lower part shows corresponding gadgets in the transformed discrete game. In the lower part, the square-shaped, diamond-shaped and circle-shaped vertices belong to $V_\Box$, $V_\Diamond$ and $V_\bigcirc$, respectively. Binary branching is displayed only in order to simplify the figure.

from vertex of $V_\Box \cup V_\Diamond$ takes zero time. The first gadget transforms internal transitions into edges of player $\Box$ taking zero time. The second gadget transforms Markovian transitions into edges of player $\bigcirc$ taking time δ where the probability p is the probability that any Markovian transition is taken in $\mathcal{G}$ before time δ. The third gadget deals with states with both external and Markovian transitions available where the player $\Diamond$ decides in vertex s in zero time whether an external transition is taken or whether the Markovian transitions are awaited in $\bar{s}$ for time δ. The fourth gadget is similar, but no Markovian transition can occur and from $\bar{s}$ the play returns into s with probability 1 (see Appendix A.6 for more details).

Similarly to $(*)$ and $(**)$, we define the *value of the discrete-time game* Δ as

$$\sup_{\sigma\in\Sigma_\Delta} \inf_{\pi\in\Pi_\Delta} \mathcal{P}^{\sigma,\pi}_{\Delta}\left[\Diamond^{\#_\circ \le N} G\right] \qquad (***)$$

where $\mathcal{P}^{\sigma,\pi}_{\Delta}\left[\Diamond^{\#_\circ \le N} G\right]$ is the probability of all runs of Δ induced by σ and π that reach G before taking more than N steps from vertices in $V_\bigcirc$. According to the intuition above, such a step bound corresponds to a time bound $N \cdot \delta = T$.

We say that a strategy *is counting* if it only considers the last vertex and the current count $\#_\circ$ of steps taken from vertices in $V_\bigcirc$. We may represent it as a function $V \times \{0, \ldots, N\} \to V$ since it is irrelevant what it does after more than N steps.

▶ **Lemma 7.** *There are counting strategies optimal in* $(***)$*. Moreover, they can be computed together with* $(***)$ *in time* $\mathcal{O}(N|V|^2)$*.*

We now show that the value $(***)$ of the discretized game Δ approximates the value $(**)$ of the CE game $\mathcal{G}$ and give the corresponding error bound.

▶ **Proposition 8** (Error bound). *For every approximation bound* $\varepsilon > 0$ *and discretization step* $\delta \le \varepsilon/(\lambda^2 T)$ *where* $\lambda = \max_{s\in S} \mathbf{E}(s)$*, the value* $(***)$ *induced by* δ *satisfies*

$$(***) \ \le\ (**) \ \le\ (***) + \varepsilon.$$

Sketch of Proof. The proof is inspired by the techniques for closed IMC [28]. Yet, there are several new issues to overcome, caused mainly by the fact that the player **env** in the CE game may choose an arbitrary real delay $t_e > 0$ (so the game is uncountably branching). The discretized game Δ is supposed to simulate the original CE game but restricts possible behaviors as follows: (1) Only one Markovian transition is allowed in any interval of length δ. (2) The delay t_e chosen by player $\Diamond$ (which simulates the player **env** from the CE game) must be divisible by δ. We show that none of these restrictions affects the value.

ad (1) As pointed out in [28], the probability of two or more Markovian transitions occurring in an interval $[0, \delta]$ is bounded by $(\lambda\delta)^2/2$ where $\lambda = \max_{s \in S} \mathbf{E}(s)$. Hence, the probability of multiple Markovian transitions occurring in any of the discrete steps of Δ is $\leq \varepsilon$.

ad (2) Assuming that at most one Markovian transition is taken in $[0, \delta]$ in the CE game, we reduce the decision when to take external transitions to minimization of a linear function on $[0, \delta]$, which in turn is minimized either in 0, or δ. Hence, the optimal choice for the player **env** in the CE game is either to take the transitions immediately at the beginning of the interval (before the potential Markovian transition) or to wait for time δ (after the potential Markovian transition). ◀

Finally, we show how to transform an optimal counting strategy $\sigma : V \times \{0, \ldots, N\} \to V$ in the discretized game Δ into an ε-optimal scheduler $\overline{\sigma}$ in the IMC $\mathcal{C}$. For every $\mathfrak{p} = s_0\, t_0 \cdots s_{n-1}\, t_{n-1}\, s_n$ we put $\overline{\sigma}(\mathfrak{p}) = \sigma(s_n, \lceil (t_0 + \ldots + t_{n-1})/\delta \rceil)$.

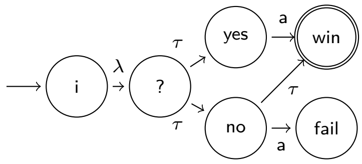
▶ **Proposition 9** (ε-optimal scheduler). *Let $\varepsilon > 0$, Δ be a corresponding discrete game, and $\overline{\sigma}$ be induced by an optimal counting strategy in Δ, then*

$$(*) \quad \leq \inf_{\substack{E \in \mathrm{ENV} \\ \pi \in \mathfrak{S}(\mathcal{C}(E), \overline{\sigma})}} \mathcal{P}^{\pi}_{\mathcal{C}(E)} \left[\Diamond^{\leq T} G_E \right] + \varepsilon$$

This together with the complexity result of Lemma 7 finishes the proof of Theorem 5.

6 Discussion

In this subsection we argue that lifting Assumption 2 makes analysis considerably more involved as the studied game may contain imperfect information and concurrent decisions. Let us illustrate the problems on an example.



Consider an IMC depicted on the right hand side. This IMC violates Assumption 2 in its state no. Let us fix an arbitrary environment E (controlled by π) and a scheduler σ. Since internal transitions of E take zero time, the environment must spend almost all the time in states without internal transitions. Hence, E is almost surely in such a state when ? is entered. Assume E is in a state with the action a being available. The scheduler σ wins if he chooses the internal transition to yes since the synchronizing transition a is then taken immediately, and fails if he chooses to proceed to no, as a (reasonable) scheduler π will now force synchronization on action a. If, otherwise, on entering state ?, E is in a state without the action a being available, the scheduler σ fails if he chooses yes because a (reasonable) environment never synchronizes, and wins if he chooses no since the environment E cannot immediately synchronize and the τ transition is taken. Note that the scheduler σ cannot observe whether a is available in the current state of E. As this is crucial for the further evolution of the game from state ?, the game is intrinsically of imperfect information.

We conjecture that solving even this special case of imperfect information games is PSPACE-hard. Yet, the complexity might only increase in the number of internal transitions that can be taken in a row. For systems, where a bound on the length of internal transition sequences can be assumed, this problem would then still be feasible.

Another possibility of dealing with this issue (that is also interesting on its own) is to limit the knowledge and power of the environment. In our approach, the environment knows the scheduler and the current state of the component and, moreover, can choose whether a synchronizing transition, an internal transition in E, or an internal transition in $\mathcal{C}$ is taken. One could consider giving some of this power either to the scheduler σ or to a third player resolving the synchronization of $\mathcal{C}$ and E who is either random or non-deterministic.

7 Summary

This paper has discussed the computation of maximal timed bounded reachability for IMC operating in an unknown IMC environment to synchronize with. All prior analysis approaches considered closed systems, implicitly assuming that external actions do happen in zero time. Our analysis for open IMC works essentially with the opposite assumption, which is arguably more realistic. We have shown that the resulting stochastic two-player game has the same extremal values as a CE-game, where the player controlling the environment can choose exact times. The latter is approximated up to a given precision by a discretization approach. The resulting control strategy can be translated back to a scheduler of the IMC achieving the bound.

References

1 K. Apt and E. Grädel, editors. *Lectures in Game Theory for Computer Scientists*. Cambridge, 2011.

2 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. Comp. Sci.*, 345(1):2–26, 2005.

3 J.A. Bergstra, A. Ponse, and S.A. Smolka, editors. *Handbook of Process Algebra*. Elsevier, 2001.

4 E. Böde, M. Herbstritt, H. Hermanns, S. Johr, T. Peikenkamp, R. Pulungan, J. Rakow, R. Wimmer, and B. Becker. Compositional dependability evaluation for STATEMATE. *IEEE Trans. on Soft. Eng.*, 35(2):274–292, 2009.

5 H. Boudali, P. Crouzen, B.R. Haverkort, M. Kuntz, and M. I. A. Stoelinga. Architectural dependability evaluation with Arcade. In *Proc. of DSN*, pages 512–521. IEEE, 2008.

6 H. Boudali, P. Crouzen, and M. Stoelinga. A rigorous, compositional, and extensible framework for dynamic fault tree analysis. *IEEE Trans. on Dependable Sec. Comp.*, 7(2):128–143, 2010.

7 P. Bouyer and V. Forejt. Reachability in stochastic timed games. In *Proc. of ICALP*, volume 5556 of *LNCS*, pages 103–114. Springer, 2009.

8 M. Bozzano, A. Cimatti, J.-P. Katoen, V.Y. Nguyen, T. Noll, and M. Roveri. Safety, dependability and performance analysis of extended AADL models. *The Computer Journal*, 54(5):754–775, 2011.

9 T. Brázdil, V. Forejt, J. Krčál, J. Křetínský, and A. Kučera. Continuous-time stochastic games with time-bounded reachability. In *Proc. of FSTTCS*, volume 4 of *LIPIcs*, pages 61–72. Schloss Dagstuhl, 2009.

10 T. Brázdil, J. Krčál, J. Křetínský, A. Kučera, and V. Řehák. Stochastic real-time games with qualitative timed automata objectives. In *Proc. of CONCUR*, volume 6269 of *LNCS*, pages 207–221. Springer, 2010.

11 P. Buchholz and I. Schulz. Numerical Analysis of Continuous Time Markov Decision processes over Finite Horizons. *Computers and Operations Research*, 38:651–659, 2011.

12 P. Černý, K. Chatterjee, T.A. Henzinger, A. Radhakrishna, and R. Singh. Quantitative synthesis for concurrent programs. In *Proc. of CAV*, volume 6806 of *LNCS*, pages 243–259. Springer, 2011.

13 K. Chatterjee and T.A. Henzinger. A survey of stochastic ω-regular games. *J. Comput. Syst. Sci.*, 78(2):394–413, 2012.

14 T. Chen, V. Forejt, M.Z. Kwiatkowska, D. Parker, and A. Simaitis. Automatic verification of competitive stochastic systems. In *Proc. of TACAS*, volume 7214 of *LNCS*, pages 315–330. Springer, 2012.

15 N. Coste, H. Hermanns, E. Lantreibecq, and W. Serwe. Towards performance prediction of compositional models in industrial GALS designs. In *Proc. of CAV*, volume 5643, pages 204–218. Springer, 2009.

16 M.-A. Esteve, J.-P. Katoen, V.Y. Nguyen, B. Postma, and Y. Yushtein. Formal correctness, safety, dependability and performance analysis of a satellite. In *Proc. of ICSE*. ACM and IEEE press, 2012.

17 J. Filar and K. Vrieze. *Competitive Markov Decision Processes.* Springer, 1996.

18 H. Garavel, R. Mateescu, F. Lang, and W. Serwe. CADP 2006: A toolbox for the construction and analysis of distributed processes. In *Proc. of CAV*, volume 4590 of *LNCS*, pages 158–163. Springer, 2007.

19 D. Guck, T. Han, J.-P. Katoenand, and M.R. Neuhäußer. Quantitative timed analysis of interactive markov chains. In *Proc. of NFM*, volume 7226 of *LNCS*, pages 8–23. Springer, 2012.

20 B.R. Haverkort, M. Kuntz, A. Remke, S. Roolvink, and M.I.A. Stoelinga. Evaluating repair strategies for a water-treatment facility using Arcade. In *Proc. of DSN*, pages 419–424, 2010.

21 H. Hermanns and S. Johr. Uniformity by construction in the analysis of nondeterministic stochastic systems. In *DSN*, pages 718–728. IEEE Computer Society, 2007.

22 H. Hermanns and S. Johr. May we reach it? Or must we? In what time? With what probability? In *Proc. of MMB*, pages 125–140. VDE Verlag, 2008.

23 H. Hermanns and J.-P. Katoen. The how and why of interactive Markov chains. In *Proc. of FMCO*, volume 6286 of *LNCS*, pages 311–337. Springer, 2009.

24 H. Hermanns, J.-P. Katoen, M. R. Neuhäußer, and L. Zhang. GSPN model checking despite confusion. Technical report, RWTH Aachen University, 2010.

25 J.-P. Katoen, D. Klink, and M. R. Neuhäußer. Compositional abstraction for stochastic systems. In *Proc. of FORMATS*, volume 5813 of *LNCS*, pages 195–211. Springer, 2009.

26 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. *Performance Evaluation*, 68(2):90–104, 2011.

27 O. Kupferman and M. Vardi. Module checking. In *CAV*, volume 1102 of *LNCS*, pages 75–86. Springer, 1996.

28 M.R. Neuhäußer. *Model checking nondeterministic and randomly timed systems.* PhD thesis, University of Twente, 2010.

29 M.N. Rabe and S. Schewe. Finite optimal control for time-bounded reachability in CTMDPs and continuous-time Markov games. *Acta Informatica*, 48(5-6):291–315, 2011.

30 P.J.G. Ramadge and W.M. Wonham. The control of discrete event systems. *Proceedings of the IEEE*, 77(1), 1989.

31 L. Zhang and M.R. Neuhäußer. Model checking interactive Markov chains. In *Proc. of TACAS*, volume 6015 of *LNCS*, pages 53–68. Springer, 2010.

A Proofs

A.1 Notation and preliminaries

For the whole appendix, we fix a IMC $\mathcal{C} = (S, \mathbb{A}\mathrm{ct}^{\tau}, \hookrightarrow, \rightsquigarrow, s_0)$ satisfying Assumptions 2 and 1, a set of goal states G, and a time bound T. Without loss of generality, we assume that the goal states are absorbing, i.e. $G \subseteq S_M$ and for every $s \in G$ we have $\mathbf{P}(s, s') > 0$ only if $s' \in G$. We further assume, that there are no Markovian self-loop transition. This is w.l.o.g. since to any IMC $\mathcal{C}$ we can construct an equivalent IMC where intuitively each state is duplicated and the Markovian self-loops are converted into a pair of transitions going from a state s into its duplicate s' and back from s' into s. Formally, $\mathcal{C}' = (S', \mathbb{A}\mathrm{ct}^{\tau}, \hookrightarrow', \rightsquigarrow', s_0)$ where $S' = S \cup \{s' \mid s \in S\}$, $\hookrightarrow' = \hookrightarrow \cup \{(s', a, s'') \mid (s, a, s'') \in \hookrightarrow\}$, and $\rightsquigarrow' = \{(s_1, r, s_2), (s_1', r, s_2) \mid (s_1, r, s_2) \in \rightsquigarrow, s_1 \neq s_2\} \cup \{(s, r, s'), (s', r, s) \mid (s, r, s) \in \rightsquigarrow\}$.

We denote by S_τ, S_e, S_M the set of states with only internal, external, and Markovian transitions available, respectively. By S_{e+M} we denote the set of states with both external and Markovian transitions available. From Assumption 2 and from the maximal progress assumption, we have $S = S_\tau \uplus S_e \uplus S_M \uplus S_{e+M}$. Furthermore, by λ we denote the maximal rate of a state in $\mathcal{C}$, i.e. $\lambda = \max_{s \in S} \mathbf{E}(s)$. For the sake of readability, we will consistently use the notions *histories* and *strategies* in the context of CE games and the notions *paths* (instead of histories) and *schedulers* in the context of IMC. The set of all paths of a IMC $\mathcal{C}$ is denoted by $\mathbb{P}\mathrm{aths}(\mathcal{C})$, the set of all histories of a CE game $\mathcal{G}$ is denoted by $\mathbb{H}\mathrm{istories}(\mathcal{G})$. For a history $\mathfrak{h} = s_0 t_0 s_1 t_1 \dots t_{n-1} s_n$, we denote by $\sum \mathfrak{h}$ the total time of the history, i.e. $\sum_{i=0}^{n-1} t_i$.

Let A be a finite or countably infinite set. A *probability distribution* on A is a function $f : A \to \mathbb{R}_{\geq 0}$ such that $\sum_{a \in A} f(a) = 1$. The set of all distributions on A is denoted by $\mathcal{D}(A)$. A σ*-field* over a set Ω is a set $\mathcal{F} \subseteq 2^{\Omega}$ that includes Ω and is closed under complement and countable union. A *measurable space* is a pair $(\Omega, \mathcal{F})$ where Ω is a set called *sample space* and $\mathcal{F}$ is a σ-field over Ω whose elements are called *measurable sets*. Given a measurable space $(\Omega, \mathcal{F})$, we say that a function $f : \Omega \to \mathbb{R}$ is a random variable if the inverse image of any real interval is a measurable set. A *probability measure* over a measurable space $(\Omega, \mathcal{F})$ is a function $\mathcal{P} : \mathcal{F} \to \mathbb{R}_{\geq 0}$ such that, for each countable collection $\{X_i\}_{i \in I}$ of pairwise disjoint elements of $\mathcal{F}$, we have $\mathcal{P}\left[\bigcup_{i \in I} X_i\right] = \sum_{i \in I} \mathcal{P}[X_i]$ and, moreover, $\mathcal{P}[\Omega] = 1$. A *probability space* is a triple $(\Omega, \mathcal{F}, \mathcal{P})$, where $(\Omega, \mathcal{F})$ is a measurable space and $\mathcal{P}$ is a probability measure over $(\Omega, \mathcal{F})$. All integrals in the following text should be understood as Lebesgue integral even when we use Riemann-like notation.

A.2 Semantics of CE games

Let us formally define the semantics of the CE game $\mathcal{G}$ induced by $\mathcal{C}$. We define the σ-field over the set of histories by $\mathcal{F} = \sigma(\bigcup_{n=0}^{\infty} 2_0^S \otimes \mathfrak{B}_0^{\geq 0} \otimes \cdots \otimes \mathfrak{B}_{n-1}^{\geq 0} \otimes 2_n^S)$ where $\sigma(X)$ is the σ-field generated from the set X, the operator $\otimes$ denotes product σ-field, $\mathfrak{B}^{\geq 0}$ denotes the Borel σ-field over the set $\mathbb{R}_{\geq 0}$.

A (randomizing) strategy σ of the controller is a measurable function $\sigma : \mathbb{H}\mathrm{istories}(\mathcal{G}) \to \mathcal{D}(S)$ such that for each history $\mathfrak{h} = s_0 t_0 \cdots t_{n-1} s_n$ with $\mathrm{succ}_\tau(s_n) \neq \emptyset$ we have for each $s \in S$ that $\sigma(\mathfrak{h})(s) > 0$ implies $s \in \mathrm{succ}_\tau(s_n)$, i.e. σ can assign positive probability only to the internal successors of s_n. A (randomizing) strategy π of the environment is a measurable function that assigns to each history a probability measure over the measurable space $(S \cup \mathbb{R}_{>0}, \sigma(2^S \cup \mathfrak{B}^{>0})$. Furthermore, for each history $\mathfrak{h} = s_0 t_0 \cdots t_{n-1} s_n$ with $\mathrm{succ}_e(s_n) \neq \emptyset$ it must hold for each $s \in S$ that $\pi(\mathfrak{h})(\{s\}) > 0$ implies $s \in \mathrm{succ}_e(s_n)$.

We say that a strategy $\sigma \in \Sigma$ is deterministic if for any history $\mathfrak{h}$ it holds $\sigma(\mathfrak{h})(s) = 1$ for some $s \in S$. Similarly, we say that a strategy $\pi \in \Pi$ is deterministic if for any history $\mathfrak{h}$ it either holds $\pi(\mathfrak{h})(s) = 1$ for some state $s \in S$ or it holds $\pi(\mathfrak{h})(w) = 1$ for some $w \in \mathbb{R}_{>0}$.

For a pair of strategies σ and π and a starting history $\mathfrak{h}_0$ we define a Markov chain as follows.

- The state space is the set $\mathbb{H}$istories with its σ-field $\mathcal{F}$;
- the chain starts in the history $\mathfrak{h}_0$, i.e. the initial measure μ_0 satisfies $\mu_0(A) = 1$ if $\mathfrak{h}_0 \in A$ and $\mu_0(A) = 0$, otherwise;
- the chain moves according to the following rules expressed by a transition kernel $P^{\sigma,\pi}_{\mathfrak{h}_0}$. Let $\mathfrak{h} = s_0 t_0 \cdots t_{n-1} s_n$ be a history. If $s_n \in S_\tau$, we have for any $s_{n+1} \in \mathrm{succ}_\tau(s_n)$

$$P^{\sigma,\pi}_{\mathfrak{h}_0}(\mathfrak{h}, \{\mathfrak{h} 0 s_{n+1}\}) = \sigma(\mathfrak{h})(s_{n+1})$$

If $s \in S_e$, only the environment chooses the waiting time, for any $A \subseteq \mathbb{R}_{>0}$ it holds

$$P^{\sigma,\pi}_{\mathfrak{h}_0}(\mathfrak{h}, \{\mathfrak{h} w s_n \mid w \in A\}) = \pi(\mathfrak{h})(A),$$

and for any $s_{n+1} \in \mathrm{succ}_e s_n$ it holds

$$P^{\sigma,\pi}_{\mathfrak{h}_0}(\mathfrak{h}, \{\mathfrak{h}\, 0\, s_{n+1}\}) = \pi(\mathfrak{h})(\{s_{n+1}\}).$$

If $s \in S_M$, we have for each $s_{n+1} \in S$

$$P^{\sigma,\pi}_{\mathfrak{h}_0}(\mathfrak{h}, \{\mathfrak{h}\, w\, s_{n+1} \mid w \in [a,b]\}) = \int_a^b \mathbf{E}(s_n) \cdot e^{-\mathbf{E}(s_n)\cdot x} \cdot \mathbf{P}(s_n, s_{n+1})\, dx.$$

Finally, if $s \in S_{e+M}$, the chain either takes the external transition

$$P^{\sigma,\pi}_{\mathfrak{h}_0}(\mathfrak{h}, \{\mathfrak{h}\, 0\, s_{n+1}\}) = \pi(\mathfrak{h})(\{s_{n+1}\}),$$

or it stays in the same state if the delay picked by the player **env**occurs sooner than any Markovian transition

$$P^{\sigma,\pi}_{\mathfrak{h}_0}(\mathfrak{h}, \{\mathfrak{h}\, w\, s_n \mid w \in [a,b]\}) = \int_{x \in [a,b]} e^{-\mathbf{E}(s_n)\cdot x}\, d\pi(\mathfrak{h}),$$

or it takes a Markovian transition into s_{n+1} with probability that it occurs sooner than the delay of the player **env**

$$P^{\sigma,\pi}_{\mathfrak{h}_0}(\mathfrak{h}, \{\mathfrak{h}\, w\, s_{n+1} \mid w \in [a,b]\}) = \int_a^b \mathbf{E}(s_n) e^{-\mathbf{E}(s_n)\cdot x} \cdot \mathbf{P}(s_n, s_{n+1}) \cdot \pi(\mathfrak{h})((x,\infty))\, dx.$$

By $\mathcal{P}^{\sigma,\pi}_{\mathcal{G},\mathfrak{h}_0}$ we denote the probability measure on the set of runs of the Markov chain defined above. Furthermore, by $\mathcal{P}^{\sigma,\pi}_{\mathcal{G}}$ we denote the probability measure of the Markov chain that starts in the initial state s_0, i.e. $\mathcal{P}^{\sigma,\pi}_{\mathcal{G},s_0}$.

A.3 Value in the CE game

Some parts of proofs in this section are inspired by [28]. Let $\Diamond^{\leq T} G$ denote the set of runs that reach the set of goal states G in time T, i.e. that have a prefix $\mathfrak{h} = s_0 t_0 \cdots t_{n-1} s_n$ such that $s_n \in G$ and $\sum \mathfrak{h} \leq T$. Similarly, let $\Diamond^{\leq T}_{\leq k} G$ denote the set of runs that reach the set of goal states in time T and in at most k *non-self-loop* transitions, i.e. that have a prefix $\mathfrak{h} = s_0 t_0 \cdots t_{n-1} s_n$ such that $s_n \in G$, $\sum \mathfrak{h} \leq T$, and $|\{i \mid 0 \leq i < n, s_i \neq s_{i+1}\}| \leq k$. Let

$$v(s,t) = \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}[\Diamond^{\leq T-t} G]$$
$$v_k(s,t) = \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}[\Diamond^{\leq T-t}_{\leq k} G]$$

denote the respective values when starting in state s with remaining time $T - t$.

We define a higher-order operator $\Omega : (S \times \mathbb{R}_{\geq 0} \to [0,1]) \to (S \times \mathbb{R}_{\geq 0} \to [0,1])$ that characterizes the functions v and v_k:

$$\Omega(F)(s,t) = \begin{cases} 0 & \text{if } t > T \text{ or } s \in S_e \setminus G \\ 1 & \text{if } t \leq T \text{ and } s \in G, \\ \max\limits_{s' \in \mathrm{succ}_\tau(s)} F(s',t) & \text{if } t \leq T \text{ and } s \in S_\tau \setminus G, \\ \mathcal{A}(T-t) & \text{if } t \leq T \text{ and } s \in S_M \setminus G, \\ \min\limits_{\substack{w \geq 0, \\ s' \in \mathrm{succ}_e(s)}} \big(e^{-\mu w} \cdot F(s', t+w) + \mathcal{A}(w)\big) & \text{if } t \leq T \text{ and } s \in S_{e+M} \setminus G, \end{cases}$$

where $\mu = \mathbf{E}(s)$ and $\mathcal{A}(u) = \int_0^u \mu e^{-\mu x} \cdot \sum_{s'' \in S} \mathbf{P}(s,s'') \cdot F(s'', t+x) dx$. Furthermore, let us denote by $f_0 : S \times \mathbb{R}_{\geq 0} \to [0,1]$ a constant zero function.

▶ **Lemma 10** (k-step value)**.** *For any $k \in \mathbb{N}_0$ we have $\Omega^{k+1}(f_0) = v_k$ and $v_k(s,\cdot)$ is a measurable function that is continuous on $[0,T]$.*

Proof. By induction on k. For $k = 0$ the zero-step value v_0 is obviously 1 for $s \in G$ and $t \leq T$ and 0 elsewhere; hence, $v_0 = f_0$ and v_0 is measurable and continuous on $[0,T]$.

Let $k > 0$ and let $s \in S$ and $t \in \mathbb{R}_{\geq 0}$. If $t > T$, then $v_k(s,t) = 0$ because for any $\sigma \in \Sigma$ and $\pi \in \Pi$ the measure of the set $\Diamond^{\leq T-t} G$ is zero For the following we assume that $t \leq T$.

- If $s \in S_e \setminus G$, then also $v_k(s,t) = 0$ because for any $\sigma \in \Sigma$ and a strategy π that starts by waiting $T - t + 1$ the measure of the set $\Diamond^{\leq T-t} G$ is again zero.
- If $s \in G$, then obviously $v_k(s,t) = 1$.
- If $s \in S_\tau \setminus G$, we have

$$\begin{aligned} v_k(s,t) &= \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}[\Diamond^{\leq T-t}_{\leq k} G] \\ &= \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \sum_{s' \in \mathrm{succ}_\tau(s)} \sigma(s)(s') \cdot \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s\,0s'}[\Diamond^{\leq T-t}_{\leq k-1} G] \\ &= \sup_{\rho \in \mathcal{D}(\mathrm{succ}_\tau(s))} \sum_{s' \in \mathrm{succ}_\tau(s)} \rho(s') \cdot \sup_{\sigma \in \Sigma_\rho} \inf_{\pi \in \Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s\,0s'}[\Diamond^{\leq T-t}_{\leq k-1} G], \end{aligned}$$

where Σ_ρ denotes the set of strategies that choose ρ for history s. The linear combination is maximized by giving weight 1 to any maximal element, i.e.

$$= \max_{s' \in \mathrm{succ}_\tau(s)} \sup_{\sigma \in \Sigma_{\mathbf{1}_{s'}}} \inf_{\pi \in \Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s\,0s'}[\Diamond^{\leq T-t}_{\leq k-1} G],$$

where $\mathbf{1}_{s'}$ is the Dirac distribution that assigns probability 1 to s'. By Markov property,

$$= \max_{s' \in \mathrm{succ}_\tau(s)} \sup_{\sigma \in \Sigma_{\mathbf{1}_{s'}}} \inf_{\pi \in \Pi} \mathcal{P}^{\sigma[s\,0],\pi[s\,0]}_{\mathcal{G},s'}[\Diamond^{\leq T-t}_{\leq k-1} G],$$

where $\sigma[s\,0](\mathfrak{h}) := \sigma(s\,0\,\mathfrak{h})$ behaves in s' as σ behaves in $s\,0\,s'$ and similarly for π. Finally,

$$= \max_{s' \in \mathrm{succ}_\tau(s)} \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s'}[\Diamond^{\leq T-t}_{\leq k-1} G],$$
$$= \max_{s' \in \mathrm{succ}_\tau(s)} v_{k-1}(s',t) \;=\; \Omega(v_{k-1})(s,t) \;=\; \Omega^{k+1}(f_0)(s,t)$$

The function v_k is obviously measurable and continuous on $[0,T]$.

- If $s \in S_M$, we have from the definition of the Markov chain and by the same arguments as above

$$\begin{aligned}
v_k(s,t) &= \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \int_0^\infty \mu\cdot e^{-\mu x}\cdot \sum_{s'\in \mathrm{succ}_M(s)} \mathbf{P}(s,s')\cdot \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s\,x\,s'}[\Diamond^{\leq T-t-x}_{\leq k-1} G]\,dx,\\
&= \int_0^\infty \mu\cdot e^{-\mu x}\cdot \sum_{s'\in \mathrm{succ}_M(s)} \mathbf{P}(s,s')\cdot \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s'}[\Diamond^{\leq T-t-x}_{\leq k-1} G]\,dx,\\
&= \int_0^\infty \mu\cdot e^{-\mu x}\cdot \sum_{s'\in \mathrm{succ}_M(s)} \mathbf{P}(s,s')\cdot v_{k-1}(s',t+x)\,dx,\\
&= \int_0^{T-t} \mu\cdot e^{-\mu x}\cdot \sum_{s'\in \mathrm{succ}_M(s)} \mathbf{P}(s,s')\cdot v_{k-1}(s',t+x)\,dx\\
&= \Omega(v_{k-1})(s,t) \;=\; \Omega^{k+1}(f_0)(s,t).
\end{aligned}$$

Again, the function v_k is measurable and continuous on $[0,T]$.

- Finally, if $s \in S_{e+M}$, we need to perform a nested induction on the count n of self-loops performed by strategy π before taking a non-self-loop transition. We denote by

$$v_{k,n}(s,t) \;=\; \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi_n} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}[\Diamond^{\leq T-t}_{\leq k} G],$$

where Π_m is the set of strategies that start by assigning positive probability to waiting at most m times before taking an external transition with probability one. We can separate the first decision ρ of π from the rest of the strategy and interchange with σ since σ cannot influence the first step,

$$= \inf_\rho \underbrace{\sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi_\rho\cap\Pi_n} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}[\Diamond^{\leq T-t}_{\leq k} G]}_{=:v^\rho_{k,n}(s,t)}$$

and denote the outcome of ρ w.r.t. next k steps by $v^\rho_{k,n}$. Let $n=0$. We analyze the outcomes of deterministic decisions ρ, i.e. $\rho[\{s'\}]=1$ for some $s' \in \mathrm{succ}_e(s)$. It holds

$$v^\rho_{k,0}(s,t) \;=\; \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi_\rho\cap\Pi_n} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s\,0\,s'}[\Diamond^{\leq T-t}_{\leq k-1} G] \;=\; \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s'}[\Diamond^{\leq T-t}_{\leq k-1} G] \;=\; v_{k-1}(s',t).$$

By this we obtain that $v_{k,0} = \min_{s'\in \mathrm{succ}_e(s)} v_{k-1}(s',t)$ since by randomizing over finite number of successor one cannot achieve less then the minimum. Let $n=1$. We again start with deterministic decisions ρ. If $\rho[\{s'\}]=1$ for some $s' \in \mathrm{succ}_e(s)$, we also get $v^\rho_{k,1}(s,t) = v_{k-1}(s',t)$. If $\rho[\{w\}]=1$ for some $w>0$, it holds

$$v^\rho_{k,1}(s,t) \;=\; \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi_\rho\cap\Pi_1} \Big(e^{-\mu w}\cdot \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s\,w\,s}[\Diamond^{\leq T-t-w}_{\leq k} G] \;+$$

$$
\begin{aligned}
&\int_0^w \mu e^{-\mu x} \sum_{s' \in \mathrm{succ}_M(s)} \mathbf{P}(s,s') \cdot \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s\,x\,s'}[\Diamond^{\leq T-t-x}_{\leq k-1} G]\,dx\Big) \\
&= e^{-\mu w} \cdot \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi_0} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}[\Diamond^{\leq T-t-w}_{\leq k} G] \;+ \\
&\int_0^w \mu e^{-\mu x} \sum_{s' \in \mathrm{succ}_M(s)} \mathbf{P}(s,s') \cdot \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s'}[\Diamond^{\leq T-t-x}_{\leq k-1} G]\,dx \\
&= e^{-\mu w} \cdot \min_{s' \in \mathrm{succ}_e(s)} v_{k-1}(s', t+w) + \\
&\int_0^w \mu e^{-\mu x} \sum_{s' \in \mathrm{succ}_M(s)} \mathbf{P}(s,s') \cdot v_{k-1}(s', t+x)\,dx.
\end{aligned}
$$

We denote by ρ_w the deterministic decision for waiting time $w > 0$. Mixing the deterministic waiting decision cannot yield better outcome than $\inf_{w>0} v_k^{\rho_w}(s,t)$ since $\int_{x>0} f(x)\,d\rho \geq \inf_{x>0} f(x)$ for any probability measure ρ over positive real numbers and for any real measurable function f. Hence,

$$
\begin{aligned}
v_{k,1}(s,t) &= \min\{ \min_{s' \in \mathrm{succ}_e(s)} v_{k-1}(s',t), \inf_{w>0} v_{k,1}^{\rho_w}(s,t)\} \\
&= \inf_{w \geq 0, s' \in \mathrm{succ}_e(s)} \left(e^{-\mu w} \cdot v_{k-1}(s', t+w) + \mathcal{A}(w) \right).
\end{aligned}
$$

From the continuity of v_{k-1} on $[0,T]$ and from the fact that for any $w, w' > T$ we have $v_{k-1}(s',w) = v_{k-1}(s',w')$ and $\mathcal{A}(w) = \mathcal{A}(w')$ the function attains minimum; it holds

$$
\begin{aligned}
&= \min_{w \geq 0, s' \in \mathrm{succ}_e(s)} \left(e^{-\mu w} \cdot v_{k-1}(s', t+w) + \mathcal{A}(w) \right) \\
&= \Omega(v_{k-1})(s,t) \;=\; \Omega^{k+1}(f_0)(s,t).
\end{aligned}
$$

At last, let $n > 1$. For deterministic decision $\rho[\{s'\}] = 1$ for some $s' \in \mathrm{succ}_e(s)$, we again get $v^{\rho}_{k,n}(s,t) = v_{k-1}(s',t)$. If $\rho[\{w\}] = 1$ for some $w > 0$, we get

$$
\begin{aligned}
v^{\rho}_{k,n}(s,t) &= e^{-\mu w} \cdot \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi_{n-1}} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}[\Diamond^{\leq T-t-w}_{\leq k} G] \;+ \\
&\int_0^w \mu e^{-\mu x} \sum_{s' \in \mathrm{succ}_M(s)} \mathbf{P}(s,s') \cdot \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s'}[\Diamond^{\leq T-t-x}_{\leq k-1} G]\,dx \\
&= e^{-\mu w} \cdot v_{k,n-1}(s, t+w) + \int_0^w \mu e^{-\mu x} \sum_{s' \in \mathrm{succ}_M(s)} \mathbf{P}(s,s') \cdot v_{k-1}(s', t+x)\,dx \\
&= e^{-\mu w} \cdot \inf_{w' \geq 0, s' \in \mathrm{succ}_e(s)} \left(e^{-\mu w'} \cdot v_{k-1}(s', t+w+w') + \mathcal{A}(w') \right) + \mathcal{A}(w) \\
&= \inf_{w' \geq 0, s' \in \mathrm{succ}_e(s)} \left(e^{-\mu(w+w')} \cdot v_{k-1}(s', t+w+w') + +\mathcal{A}(w+w') \right).
\end{aligned}
$$

Hence, we again obtain

$$
\begin{aligned}
v_{k,n}(s,t) &= \min\{ \min_{s' \in \mathrm{succ}_e(s)} v_{k-1}(s',t), \inf_{w>0} v_{k,n}^{\rho_w}(s,t)\} \\
&= \inf_{w\geq 0} \inf_{w' \geq 0, s' \in \mathrm{succ}_e(s)} \left(e^{-\mu(w+w')} \cdot v_{k-1}(s', t+w+w') + \mathcal{A}(w+w') \right) \\
&= \inf_{w \geq 0, s' \in \mathrm{succ}_e(s)} \left(e^{-\mu(w)} \cdot v_{k-1}(s', t+w) + \mathcal{A}(w) \right)
\end{aligned}
$$

$$= \min_{w \geq 0, s' \in \text{succ}_e(s)} \Big(e^{-\mu(w)} \cdot v_{k-1}(s', t+w) + \mathcal{A}(w) \Big) \;=\; \Omega^{k+1}(f_0)(s,t)$$

Now we show that it suffices to consider strategies from $\bigcup_{n=0}^{\infty} \Pi_n$. Let π be any non-Zeno strategy, i.e. for any σ we have that $\mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}[Z] = 0$ where Z is the set of Zeno runs. Then also $\mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}[X_\infty] = 0$ where X_∞ is the set of runs with infinite amount of self-loop transitions before any non-self-loop transition. For any $\varepsilon > 0$ there must be an amount of self-loop transitions $k \in \mathbb{N}_0$ such that $\mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}[X_k] \leq \varepsilon$. Hence, a strategy $\pi' \in \Pi_k$ that emulates π in first k self-loop transitions and then takes an arbitrary external transition (and also emulates π in all following transitions) guarantees the same value as π up to ε. Hence, $\lim_{n\to\infty} v_{k,n} = v_k$.

The function v_k is again measurable and continuous on $[0, T]$.

◀

▶ **Lemma 11** (Tail bound on steps). *For each $\varepsilon > 0$ there is $k \in \mathbb{N}$ such that for every pair of strategies $\sigma \in \Sigma$, $\pi \in \Pi$ we have*

$$\mathcal{P}^{\sigma,\pi}_{\mathcal{G}}(\Diamond^{\leq T}(G) \setminus \Diamond^{\leq T}_{\leq k}(G)) \leq \varepsilon$$

Proof. The proof is based on the Assumption 1 and on a tail bound for Poisson distribution. From the Assumption 1, at least one transition on any cycle in Markovian. This provides a bound on number of cycles that can be traversed: for any $\varepsilon > 0$ there is k' such that $Pr[X > k'] < \varepsilon$ for a random variable $X \sim \text{Pois}(\lambda \cdot T)$ distributed according to the Poisson distribution with rate $\lambda \cdot T$.

Let $n \leq |S|$ be the maximal length of a cycle in the state space. We can set $k := k' \cdot n$ and get the desired property. ◀

▶ **Lemma 12** (Value). *The function v is a fixed point of the operator Ω.*

Proof. From Lemma 11 we get that v_k uniformly converges to v for $k \to \infty$. From the measurability of v_k from Lemma 10 we immediately get that v is measurable. The fact that v is a fixed point follows from the fact that for any $s \in S$ and $t \in \mathbb{R}_{\geq 0}$ it holds

$$v(s,t) = \lim_{k\to\infty} v_k(s,t) = \lim_{k\to\infty} v_{k+1}(s,t)$$

which is from Lemma 10 and from continuity of Ω equal to

$$\begin{aligned} &= \lim_{k\to\infty} \Omega(v_k)(s,t) = \Omega(\lim_{k\to\infty} v_k)(s,t) \\ &= \Omega(v)(s,t) \end{aligned}$$

◀

▶ **Lemma 13.** *Let us fix $\delta > 0$ such that $T = \delta n$ for some $n \in \mathbb{N}$. The function v satisfies for any $k \in \mathbb{N}_0$ the following. First, $v(s, k\delta) = 0$ if $k\delta > T$ or $s \in S_e$, and $v(s, k\delta) = 1$ if $s \in G$ and $k\delta \leq T$. Second, for any $s \notin G$ and $k\delta \leq T$ it holds*

$$v(s,k\delta) = \begin{cases} \max\limits_{s' \in \text{succ}_\tau(s)} v(s', k\delta) & \text{if } s \in S_\tau, \\ \mathcal{A}(\delta) + e^{-\mu\delta} \cdot v(s,(k+1)\delta) & \text{if } s \in S_M, \\ \min\Bigg(\mathcal{A}(\delta) + e^{-\mu\delta} \cdot v(s,(k+1)\delta), & \\ \qquad \min\limits_{\substack{s' \in \text{succ}_e(s),\\ 0 \leq w \leq \delta}} \big(e^{-\mu w} \cdot v(s', k\delta + w) + \mathcal{A}(w) \big) \Bigg) & \text{if } s \in S_{e+M}, \end{cases}$$

where $\mu = E(s)$ *and* $\mathcal{A}(u) = \int_0^u \mu e^{-\mu x} \cdot \sum_{s'' \in S} \mathbf{P}(s, s'') \cdot v(s'', k\delta + x) dx$.

Proof. From Lemma 12 we immediately get that $v(s, k\delta) = 0$ if $k\delta > T$ or $s \in S_e$; $v(s, k\delta) = 1$ if $s \in G$ and $k\delta \leq T$; and $v(s, k\delta) = \max_{s' \in \mathrm{succ}_\tau(s)} v(s', k\delta)$ if $s \in S_\tau \setminus G$ and $k\delta \leq T$. Now let $s \in S_M \setminus G$ and $k\delta \leq T$. It holds that

$$v(s, k\delta) = \int_0^{T-t} \mu e^{-\mu x} \cdot \sum_{s'' \in S} \mathbf{P}(s, s'') \cdot v(s'', k\delta + x) dx$$

which can be rewritten by splitting the integral and further by substituting $x + \delta$ for x as

$$\begin{aligned}
&= \mathcal{A}(\delta) + \int_\delta^{T-t} \mu e^{-\mu x} \cdot \sum_{s'' \in S} \mathbf{P}(s, s'') \cdot v(s'', k\delta + x) dx \\
&= \mathcal{A}(\delta) + \int_0^{T-t-\delta} \mu e^{-\mu (x+\delta)} \cdot \sum_{s'' \in S} \mathbf{P}(s, s'') \cdot v(s'', k\delta + x + \delta) dx \\
&= \mathcal{A}(\delta) + e^{-\mu\delta} \cdot \int_0^{T-t-\delta} \mu e^{-\mu x} \cdot \sum_{s'' \in S} \mathbf{P}(s, s'') \cdot v(s'', k\delta + x + \delta) dx \\
&= \mathcal{A}(\delta) + e^{-\mu\delta} \cdot v(s, (k+1) + \delta)
\end{aligned}$$

Finally, for $s \in S_{e+M} \setminus G$ and $k\delta \leq T$, it holds

$$\begin{aligned}
v(s, k\delta) &= \min_{w \geq 0, s' \in \mathrm{succ}_e(s)} \left(e^{-\mu w} \cdot F(s', t + w) + \mathcal{A}(w)\right) \\
&= \min\{ \min_{s' \in \mathrm{succ}_e(s), w \geq \delta} \left(e^{-\mu w} \cdot v(s', k\delta + w) + \mathcal{A}(w)\right), \\
&\qquad \min_{s' \in \mathrm{succ}_e(s), 0 \leq w \leq \delta} \left(e^{-\mu w} \cdot v(s', k\delta + w) + \mathcal{A}(w)\right)\}.
\end{aligned}$$

In the first line we can substitute $w + \delta$ for w and yield similarly to the previous case

$$\begin{aligned}
&= \min\{\mathcal{A}(\delta) + e^{-\mu\delta} \cdot v(s, (k+1)\delta), \\
&\qquad \min_{s' \in \mathrm{succ}_e(s), 0 \leq w \leq \delta} \left(e^{-\mu w} \cdot v(s', k\delta + w) + \mathcal{A}(w)\right)\}.
\end{aligned}$$ ◀

We say that a strategy $\rho \in \Sigma \cup \Pi$ is *total time positional* if for any two histories $\mathfrak{h} = s_0 t_0 \cdots t_{n-1} s_n$ and $\mathfrak{h}' = s'_0 t'_0 \cdots t'_{m-1} s'_m$ with $s_n = s'_m$ and $\sum \mathfrak{h} = \sum \mathfrak{h}'$ it holds $\rho(\mathfrak{h}) = \rho(\mathfrak{h}')$. We denote by $\rho(s, t)$ the decision of a total time positional strategy ρ for any history $\mathfrak{h}$ with $\sum \mathfrak{h} = t$ that ends with state s. Furthermore, we say that a deterministic total time positional strategy π of the player **env**is *consistent* if it satisfies the following implication. Let $s \in S$ and $t \leq T$. If $\pi(s, t) = w$ for $w > 0$, we have for any $y < w$ that $\pi(s, t + y) = w - y$ and either the strategy waits beyond the bound T, i.e. $t + w > T$ or it then chooses a transition, i.e. $\pi(s, t + w) = s'$ for some $s' \in \mathrm{succ}_e(s)$. For any deterministic total time positional strategy σ of the player **con**we say that it is consistent. For the following lemma, let Σ' and Π' denote the set of consistent strategies.

▶ **Lemma 14.** *Consistent strategies suffice for both players, i.e.*

$$\sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\leq T} G] \;=\; \sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi'} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\leq T} G] \;=\; \sup_{\sigma \in \Sigma'} \inf_{\pi \in \Pi} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\leq T} G].$$

Proof. Let us define a pair of consistent strategies σ^* and π^* using v as follows. Let us fix an arbitrary linear order $\preceq$ over S. For any state s and total time $t \in \mathbb{R}_{\geq 0}$

- the strategy σ^* chooses the minimal state $s' \in \mathrm{succ}_\tau(s)$ w.r.t. $\preceq$ out of those that maximize $v(s',t)$, and an arbitrary state if $\mathrm{succ}_\tau(s) = \emptyset$;
- the strategy π^* is defined by the following rules. If $\mathrm{succ}_e(s) = \emptyset$, π^* chooses the delay $T + 1 - t$. Otherwise, let

$$(s', w) = \mathop{\arg\min}_{s' \in \mathrm{succ}_e(s), w \geq 0} e^{-\mu w} \cdot v(s', t + w) + \int_0^w \mu e^{-\mu x} \cdot \sum_{s'' \in S} \mathbf{P}(s, s'') \cdot v(s'', t + x)\, dx. \tag{1}$$

  where the state s' is minimal w.r.t. $\preceq$ if there are more states that minimize the equation above. The strategy π^* then chooses the state s' if $w = 0$; and π^* chooses the delay w if $w > 0$.

Clearly, σ^* and π^* are total time positional and deterministic. By the Markov property of the equation that π^* minimizes, it is easy to show that π^* is consistent.

Now, let $\varepsilon > 0$. We show that σ^* and π^* are ε-optimal. Observe that it suffices for showing that both strategies are optimal. Indeed, it follows immediately since we fix ε arbitrarily. Let $\sigma_{(s,t)}$ and $\pi_{(s,t)}$ denote some strategies that are $(\varepsilon/2)$-optimal w.r.t. the value $v(s,t)$. For any $n > 0$, we define strategies σ_n and π_n as follows. Let $\mathfrak{h} = s_0 t_0 \cdots t_{k-1} s_k$ be a history.

- We set $\sigma_n(\mathfrak{h}) = \sigma^*(\mathfrak{h})$ and $\pi_n(\mathfrak{h}) = \pi^*(\mathfrak{h})$ if there are $\leq n$ non-self-loop transition in $\mathfrak{h}$ (i.e., σ_n and π_n behave as σ^* and π^* in the first n steps, respectively).
- Otherwise, we set $\sigma_n(\mathfrak{h}) = \sigma_{(s_m,t)}(\mathfrak{h}'')$ and $\pi_n(\mathfrak{h}) = \pi_{(s_m,t)}(\mathfrak{h}'')$ where $t = \sum \mathfrak{h}'$, $\mathfrak{h}' = s_0 t_0 \cdots t_{m-1} s_m$ is the shortest prefix of $\mathfrak{h}$ with n non-self-loop transitions, and $\mathfrak{h}'' = s_m t_m \cdots t_{k-1} s_k$ be the remaining part of $\mathfrak{h}$ (i.e., σ_n and π_n then behave as an ε-optimal strategy).

▶ Claim 15. For any $n \in \mathbb{N}_0$, any $s \in S$, and any $t \in \mathbb{R}_{\geq 0}$ it holds

$$\inf_{\pi \in \Pi} \mathcal{P}^{\sigma_n, \pi}_{\mathcal{G}, s}[\Diamond^{\leq T-t} G] \geq v(s,t) - \varepsilon,$$
$$\sup_{\sigma \in \Sigma} \mathcal{P}^{\sigma, \pi_n}_{\mathcal{G}, s}[\Diamond^{\leq T-t} G] \leq v(s,t) + \varepsilon.$$

Proof. By induction on n. The strategies σ_0 and π_0 behave directly as ε-optimal strategies yielding the claim. Let $n > 0$. For $t > T$, $s \in G$ or $s \in S_e$, the claim is straightforward. Otherwise:

- If $s \in S_\tau$, let $s' = \sigma_n(s)$. We get

$$\inf_{\pi \in \Pi} \mathcal{P}^{\sigma_n, \pi}_{\mathcal{G}, s}[\Diamond^{\leq T-t} G] = \inf_{\pi \in \Pi} \mathcal{P}^{\sigma_{n-1}, \pi}_{\mathcal{G}, s'}[\Diamond^{\leq T-t} G] \geq v(s',t) - \varepsilon = v(s,t) - \varepsilon,$$
$$\begin{aligned}\sup_{\sigma \in \Sigma} \mathcal{P}^{\sigma, \pi_n}_{\mathcal{G}, s}[\Diamond^{\leq T-t} G] &= \max_{s' \in \mathrm{succ}_\tau(s)} \sup_{\sigma \in \Sigma} \mathcal{P}^{\sigma, \pi_{n-1}}_{\mathcal{G}, s'}[\Diamond^{\leq T-t} G] \\ &\geq \max_{s' \in \mathrm{succ}_\tau(s)} v(s',t) - \varepsilon = v(s,t) - \varepsilon.\end{aligned}$$

- If $s \in\in S_M$, let $\mu = \mathbf{E}(s)$. We have

$$\begin{aligned}\inf_{\pi \in \Pi} \mathcal{P}^{\sigma_n, \pi}_{\mathcal{G}, s}[\Diamond^{\leq T-t} G] &= \int_0^{T-t} \mu e^{-\mu x} \cdot \sum_{s' \in \mathrm{succ}_M(s)} \mathbf{P}(s,s') \inf_{\pi \in \Pi} \mathcal{P}^{\sigma_{n-1}, \pi}_{\mathcal{G}, s'}[\Diamond^{\leq T-t-x} G]\, dx \\ &\geq \int_0^{T-t} \mu e^{-\mu x} \cdot \sum_{s' \in \mathrm{succ}_M(s)} \mathbf{P}(s,s')(v(s', t+x) - \varepsilon)\, dx \\ &\geq v(s,t) - \varepsilon.\end{aligned}$$

  and analogously for $\sup_{\sigma \in \Sigma} \mathcal{P}^{\sigma, \pi_n}_{\mathcal{G}, s}[\Diamond^{\leq T-t} G]$.

- Finally, if $s \in\in S_{M+e}$, let (s', w) be the successor state and waiting time minimizing (1) and $\mu = \mathbf{E}(s)$. We have

$$\begin{aligned}\sup_{\sigma\in\Sigma} \mathcal{P}^{\sigma,\pi_n}_{\mathcal{G},s}[\Diamond^{\leq T-t}G] &= e^{-\mu w} \sup_{\sigma\in\Sigma} \mathcal{P}^{\sigma,\pi_{n-1}}_{\mathcal{G},s'}[\Diamond^{\leq T-t-w}G] + \\ &\quad \int_0^w \mu e^{-\mu x} \sum_{s''\in S} \mathbf{P}(s,s'') \sup_{\sigma\in\Sigma} \mathcal{P}^{\sigma,\pi_{n-1}}_{\mathcal{G},s''}[\Diamond^{\leq T-t-x}G]\,dx \\ &\leq e^{-\mu w}(v(s', t+w)+\varepsilon) + \\ &\quad \int_0^w \mu e^{-\mu x} \sum_{s''\in S} \mathbf{P}(s,s'')(v(s'', t+x)+\varepsilon)\,dx \\ &\leq v(s,t)+\varepsilon\end{aligned}$$

For $\inf_{\pi\in\Pi} \mathcal{P}^{\sigma_n,\pi}_{\mathcal{G},s}[\Diamond^{\leq T-t}G]$, it does not depend on σ_n in the first step (analogously to $\sup_{\sigma\in\Sigma} \mathcal{P}^{\sigma,\pi_n}_{\mathcal{G},s}[\Diamond^{\leq T-t}G]$ for $s \in S_\tau$). Therefore, it can be easily shown by similar arguments as in Lemma 10. ◀

We conclude the proof of the lemma by a simple observation. Notice that for any $\pi \in \Pi$

$$\mathcal{P}^{\sigma_n,\pi}_{\mathcal{G}}[\Diamond^{\leq T}G] \;=\; \mathcal{P}^{\sigma_n,\pi}_{\mathcal{G}}[\Diamond^{\leq T}_{\leq n}G \uplus X] \;=\; \mathcal{P}^{\sigma_n,\pi}_{\mathcal{G}}[\Diamond^{\leq T}_{\leq n}G] + \mathcal{P}^{\sigma_n,\pi}_{\mathcal{G}}[X] \;=\; \mathcal{P}^{\sigma^*,\pi}_{\mathcal{G}}[\Diamond^{\leq T}_{\leq n}G] + \mathcal{P}^{\sigma_n,\pi}_{\mathcal{G}}[X]$$

where $X = \Diamond^{\leq T}G \setminus \Diamond^{\leq T}_{\leq n}G$. Since from Lemma 11, $\mathcal{P}^{\sigma_n,\pi}_{\mathcal{G}}[X] \to 0$ as $n \to \infty$, we get that $\mathcal{P}^{\sigma_n,\pi}_{\mathcal{G}}[\Diamond^{\leq T}G] \to \mathcal{P}^{\sigma^*,\pi}_{\mathcal{G}}[\Diamond^{\leq T}G]$. Hence, there is n such that

$$\inf_{\pi\in\Pi} \mathcal{P}^{\sigma^*,\pi}_{\mathcal{G},s}[\Diamond^{\leq T}G] \;\geq\; \inf_{\pi\in\Pi} \mathcal{P}^{\sigma_n,\pi}_{\mathcal{G},s}[\Diamond^{\leq T}G] - \varepsilon/2 \;\geq\; v(s_0,t) - \varepsilon.$$

It proves that σ^* is ε-optimal, and analogously that π^* is ε-optimal, concluding the proof. ◀

A.4 Proof of Proposition 6

By Lemma 14 we can slightly abuse the notation and for the rest of the appendix denote by Σ and Π the set of *deterministic* strategies.

Proposition 6. $(*) = (**)$, i.e.

$$\sup_{\sigma\in\mathfrak{S}(\mathcal{C})} \inf_{\substack{E\in\mathrm{ENV}\\ \pi\in\mathfrak{S}(\mathcal{C}(E),\sigma)}} \mathcal{P}^{\pi}_{\mathcal{C}(E)}\left[\Diamond^{\leq T}G_E\right] = \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G}}\left[\Diamond^{\leq T}G\right]$$

Proof. Firstly, we prove the inequality $(*) \geq (**)$. This amounts to showing that an arbitrary environment E can be "simulated" by the player **env**in the CE game. Formally, it is sufficient to prove

$$\forall\sigma\in\Sigma\ \exists\sigma'\in\mathfrak{S}(\mathcal{C})\ \forall E\in\mathrm{ENV}\ \forall\pi\in\mathfrak{S}(\mathcal{C}(E),\sigma')\ \exists\pi_E\in\Pi : \qquad \mathcal{P}^{\pi}_{\mathcal{C}(E)}\left[\Diamond^{\leq T}G_E\right] \geq \mathcal{P}^{\sigma,\pi_E}_{\mathcal{G}}\left[\Diamond^{\leq T}G\right] \quad (\heartsuit)$$

Note that every strategy σ of the player **con**is actually also a scheduler for $\mathcal{C}$. Thus we set $\sigma' := \sigma$ and then for every environment E and its scheduler π, we give a strategy π_E of the player **env**that makes "equivalent" decisions as π in the "equivalent" history. We then prove that π_E guarantees the same value as π does.

The idea of the simulation is the following. Whenever π synchronizes on an external action a, π_E chooses a. Whenever E waits with a rate λ, π_E chooses to wait, too. Here we use randomizing strategies so that we can combine all waiting times $t \in \mathbb{R}_{>0}$ with the exponential distribution with rate λ. In other words, π_E simulates the random waiting of E using randomizing (which we can consider due to Lemma 14). One can view this kind of randomizing strategies as strategies where the co-domain contains not only $\mathbb{R}_{>0}$ but also exponential distributions and we will use this notation.

Let thus σ, E, π be arbitrary but fixed. In the following, we define π_E through a function $\mathrm{WC} : \mathbb{H}\mathrm{istories}(\mathcal{G}) \to \mathbb{P}\mathrm{aths}(\mathcal{C}(E))$ transforming the histories of the CE game into paths of $\mathcal{C}(E)$, which π_E uses to ask what π would do. Since E can have probabilistic branching, we need to pick one of possibly more paths of $\mathcal{C}(E)$ corresponding to the history of the simulating play in $\mathcal{G}$. We will pick one where the future chances are the best for the environment, i.e. worst for the time bounded reachability, hence WC for the "worst case".

The functions π_E and WC are defined inductively and only on the reachable histories; one can define them arbitrarily elsewhere. We start with $\mathrm{WC}(s_0^{\mathcal{C}}) := (s_0^{\mathcal{C}}, s_0^E)$. For history $\mathfrak{h}$ with $\mathrm{WC}(\mathfrak{h})$ ending in (c_1, e_1), we first define what π does after a (possibly empty) sequence of internal steps in E. Formally, let $n \in \mathbb{N}$ be the greatest such that there is a sequence $(c_1, e_1), \ldots, (c_n, e_n)$ where $\pi(\mathrm{WC}(\mathfrak{h})(c_2, e_2) \cdots (c_i, e_i)) = (c_{i+1}, e_{i+1})$ and $c_i = c_1$ and $e_i \overset{\tau}{\hookrightarrow} e_{i+1}$ for $i < n$. If $n = 1$ then $(c_2, e_2) \cdots (c_n, e_n)$ is empty. Depending on the type of the transition between (c_n, e_n) and $(c, e) := \pi\big(\mathrm{WC}(\mathfrak{h})(c_2, e_2) \cdots (c_n, e_n)\big)$, we define $\pi_E(\mathfrak{h})$ as follows:

- If either $c_n \overset{a}{\hookrightarrow} c$ and $e_n \overset{a}{\hookrightarrow} e$ for some $a \in \mathbb{A}\mathrm{ct}$, /* external transition */
  or $c_n \overset{\tau}{\hookrightarrow} c$ and $e_n = e$ /* internal transition in $\mathcal{C}$ */
  then $\pi_E(\mathfrak{h}) := c$.
  The new history is then $\mathfrak{h}' = \mathfrak{h}\, 0\, c$ and we set $\mathrm{WC}(\mathfrak{h}') := \mathrm{WC}(\mathfrak{h})(c_2, e_2) \cdots (c_n, e_n)(c, e)$.
- Else /* only Markovian transition(s) are enabled and $\pi[\sigma]$ is thus ignored*/
  - if $\mathrm{succ}_M(e_n) = \emptyset$ then $\pi_E(\mathfrak{h}) := T + 1$; /* E is blocked */
    the new history is then either longer than T if no Markovian transition from c_n occurs before T, or else a Markovian transition occurs after t still before T and we set $\mathfrak{h}' = \mathfrak{h}\, t\, c_M$ given by the respective $c_M \in \mathrm{succ}_M(c_n)$, and further $\mathrm{WC}(\mathfrak{h}') := \mathrm{WC}(\mathfrak{h})(c_2, e_2), \ldots, (c_n, e_n)(c_M, e_n)$;
  - else $\pi_E(\mathfrak{h}) := \mathrm{Exp}(\mathbf{E}(e))$; /* E waits */
    then either a Markovian transition $c \rightsquigarrow c_M$ happens before t, in which case $\mathfrak{h}'$ and $\mathrm{WC}(\mathfrak{h}')$ are defined as in the previous case; or else pick arbitrary $e_M \in \mathrm{succ}_M(e_n)$ minimizing
    $$\mathcal{P}^{\pi}_{\mathcal{C}(E)}\Big[\Diamond^{\le T} G_E \mid \mathrm{WC}(\mathfrak{h}')\Big]$$
    where $\mathfrak{h}' := \mathfrak{h}\, t\, c_n$ and $\mathrm{WC}(\mathfrak{h}') := \mathrm{WC}(\mathfrak{h})(c_2, e_2), \ldots, (c_n, e_n)(c_n, e_M)$.

▶ **Lemma 17.** *For every $\sigma \in \mathfrak{S}(\mathcal{C}), E \in \mathrm{ENV}, \pi \in \mathfrak{S}(\mathcal{C}(E), \sigma)$, we have*

$$\mathcal{P}^{\pi}_{\mathcal{C}(E)}\big[\Diamond^{\le T} G_E\big] \ge \mathcal{P}^{\sigma,\pi_E}_{\mathcal{G}}\big[\Diamond^{\le T} G\big]$$

Proof. If there are no probabilistic choices in E then the values are the same. Indeed, the only difference of the simulating probabilistic space to the original one is that whenever there is a probabilistic choice, the environment is always "lucky". Since the minimum of elements is never greater than their affine combination, the result follows.

Formally, we proceed as follows.

Firstly, we define a measure $\mathcal{P}^{E,\pi}_{\mathcal{G}}$ on infinite histories of $\mathcal{G}$ directly induced by E and π. As opposed to π_E, the probabilistic choices of the environment are reflected here. Let $\mathrm{RealStep} : \mathbb{P}\mathrm{aths}(\mathcal{C}(E)) \to \mathbb{H}\mathrm{istories}(\mathcal{G})$ project all internal transitions of the environment out, i.e. it maps a run $(c_0, e_0)t_0(c_1, e_1)t_1 \cdots$ to a run $c_0\, t_0 \cdots$ where each $c_i\, t_i$ is omitted whenever $c_i = c_{i-1}$ and $e_{i-1} \xrightarrow{\tau} e_i$. Then we define $\mathcal{P}^{E,\pi}_{\mathcal{G}} := \mathcal{P}^{\pi}_{\mathcal{C}(E)} \circ \mathrm{RealStep}^{-1}$. Clearly, as τ transitions take no time we have[5]

$$\mathcal{P}^{\pi}_{\mathcal{C}(E)}[\Diamond^{\leq T} G_E] = \mathcal{P}^{E,\pi}_{\mathcal{G}}[\Diamond^{\leq T} G]$$

Secondly, for $i \in \mathbb{N}_0$, consider the set $\mathbb{H}\mathrm{istories}_i \subseteq \mathbb{H}\mathrm{istories}(\mathcal{G})$ of histories of length i, i.e. after the ith step is taken. Let $\mathbf{p_i} \in \mathcal{D}(\mathbb{H}\mathrm{istories}_i)$ denote the transient probability measure according to $\mathcal{P}^{E,\pi[\sigma]}_{\mathcal{G}}$ after i steps. Further let $\mathbf{r_i} : \mathbb{H}\mathrm{istories}_i \to [0,1]$ be given by $\mathbf{r}_i(\mathfrak{h}) = \mathcal{P}^{E,\pi}_{\mathcal{G}}[\Diamond^{\leq T} G_E \mid \mathfrak{h}]$. Clearly, as states of G are absorbing we have

$$\mathcal{P}^{E,\pi}_{\mathcal{G}}[\Diamond^{\leq T} G_E] = \int \mathbf{r_i}\, d\mathbf{p_i}$$

Thirdly, let $\mathbf{q_i} \in \mathcal{D}(\mathbb{H}\mathrm{istories}_i)$ be the transient probability measure according to $\mathcal{P}^{\sigma,\pi_E}_{\mathcal{G}}$ after the ith step is taken. A simple induction with case distinction from the definition of π_E reveals that

$$\int \mathbf{r_i}\, d\mathbf{p_i} \geq \int \mathbf{r_i}\, d\mathbf{q_i}$$

Indeed, all but one case preserve equality. The only interesting case is the Markovian transition in E. As the minimum of elements is never greater than their affine combination, we obtain the desired inequality.

Finally, it remains to prove that

$$\lim_{i \to \infty} \int \mathbf{r_i}\, d\mathbf{q_i} = \mathcal{P}^{\sigma,\pi_E}_{\mathcal{G}}[\Diamond^{\leq T} G]$$

i.e. that the gains of the gradual replacements of the strategy converge to the gain of the limiting strategy. This follows from $\mathbf{r_i}(\mathfrak{h})$ being zero or one for each path $\mathfrak{h}$ longer than T only depending on the state at time T, and from the fact that the set of runs that never exceed T is of zero measure due to Assumption 1. ◀

The previous lemma proves ($\heartsuit$) by which the proof of $(*) \geq (**)$ is concluded.

Secondly, we prove the inequality $(**) \geq (*)$. We can divide the proof in three steps: (a) we show that in the CE game *grid* strategies are sufficient for both players; (b) this result is further employed in showing that *exponential* strategies are sufficient for the player **env**; furthermore, (c) any exponential strategy of the player **env**can be straightforwardly simulated by a specific environment and scheduler in the IMC. Let us first define the necessary notions.

We say that a strategy is a *grid* strategy on a grid of size $\delta > 0$ if

- it decides only according to the current state and the integer k such that the total time of the history belongs to the interval $[k\delta, (k+1)\delta)$, i.e. for any two histories $\mathfrak{h} = s_0 t_0 \ldots t_{n-1} s_n$ and $\mathfrak{h}' = s'_0 t'_0 \ldots t'_{m-1} s'_m$ with $s_n = s'_m$ and $\sum \mathfrak{h}, \sum \mathfrak{h}' \in [k\delta, (k+1)\delta)$ for some $k \in \mathbb{N}_0$ we have $\sigma(\mathfrak{h}) = \sigma(\mathfrak{h}')$; and
- it is either a strategy of the player **con**or it chooses waiting times only on the δ-grid, i.e. for any history $\mathfrak{h} = s_0 t_0 \ldots t_n s_n$ the strategy σ either chooses an action or a time step t_{n+1} such that $\sum \mathfrak{h} + t_{n+1} = k\delta$ for some $k \in \mathbb{N}$.

[5] Note that E and $\pi[\sigma]$ do *not* induce any strategy that would copy the IMC behavior completely. For this, one would need the notion of a strategy with a *stochastic update*, i.e. a strategy that can change its "state" randomly and thus model where in E the original path currently is.

Furthermore, for $\lambda \in \mathbb{R}$ we say that a strategy π of the player **env**is *exponential* with rate λ if

- it chooses to wait solely with the exponential distribution with rate λ;
- for any history $\mathfrak{h} = s_0 t_0 \dots t_n s_n$ with $n > 0$, $t_n > 0$, and $s_n \neq s_{n-1}$ it chooses to wait;
- for other histories it behaves as a grid strategy for some $\delta > 0$, i.e. we have $\sigma(\mathfrak{h}) = \sigma(\mathfrak{h}')$ for any two histories $\mathfrak{h} = s_0 t_0 \dots t_{n-1} s_n$ and $\mathfrak{h}' = s'_0 t'_0 \dots t'_{m-1} s'_m$ with
  - either $n = m = 0$ or $t_{n-1} = t'_{m-1} = 0$ or $s_n = s_{n-1} = s'_m = s'_{m-1}$, and /* the conditions above negated */
  - $s_n = s'_m$ and $\sum \mathfrak{h}, \sum \mathfrak{h}' \in [k\delta, (k+1)\delta)$ for some $k \in \mathbb{N}_0$. /* grid strategy */

Intuitively, a λ-exponential strategy cannot take an action right after a Markovian transition (resulting in $s_n \neq s_{n-1}$ and $t_{n-1} > 0$). The set of all grid strategies is denoted by $\Sigma_\#$ and $\Pi_\#$, the set of all λ-exponential strategies is denoted by Π_λ.

▶ **Lemma 18.** *Grid strategies are sufficient for both players, i.e.*

$$\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\le T} G] \;=\; \sup_{\sigma\in\Sigma_\#}\inf_{\pi\in\Pi} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\le T} G] \;=\; \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi_\#} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\le T} G]$$

Proof. The grid strategies closely correspond to the counting strategies in the discretized game Δ. Therefore, we do the technical work on this lemma in Section A.8. Denoting the three values by (1), (2), and (3), observe that (2) $\le$ (1) and (1) $\le$ (3) hold trivially. As regards (1) $\le$ (2), we denote by $\Pi_{\#,\delta}$ the set of grid strategies on a grid of size δ. Lemma 25 implies for any $\varepsilon > 0$ and any $\delta \le 2\varepsilon/(\lambda^2 T)$

$$v_\delta(s_0, 0) \;\le\; \sup_{\sigma\in\Sigma_{\#,\delta}}\inf_{\pi\in\Pi} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\le T} G],$$

which combined with the first part of Lemma 27 results in

$$\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\le T} G] \;\le\; \sup_{\sigma\in\Sigma_{\#,\delta}}\inf_{\pi\in\Pi} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\le T} G] + \varepsilon,$$

which proves the inequality by taking $\varepsilon \to 0$. Similarly for (3) $\le$ (1) because we have by combining the second part of Lemma 27 with the first part of Lemma 25

$$\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi_\#} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\le T} G] \;\le\; \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\le T} G] + \varepsilon.$$ ◀

▶ **Lemma 19.** *Exponential strategies for the player* **env***are sufficient against grid strategies, i.e. for any grid strategy σ we have*

$$\inf_{\pi\in\Pi_\#} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\le T} G] \;=\; \inf_{\lambda\in\mathbb{R}_{>0},\pi_\lambda\in\Pi_\lambda} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi_\lambda}[\Diamond^{\le T} G]$$

Proof. We fix arbitrary strategies $\sigma \in \Sigma_\#$ and $\pi \in \Pi_\#$ of the same grid size. We need to find a sequence of strategies π_λ for any λ such that

$$\mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\le T} G] \;\ge\; \lim_{\lambda\to\infty} \mathcal{P}_{\mathcal{G}}^{\sigma,\pi_\lambda}[\Diamond^{\le T} G].$$

For any $\lambda > 0$, we define $\pi_\lambda(\mathfrak{h})$ for $\mathfrak{h} = s_0 t_0 \cdots t_{n-1} s_n$ using π as follows. Intuitively, if π chooses to wait for time t and then makes action a, the simulating strategy π_λ repeatedly waits for random time with exponential distribution until the sum of the random waiting times exceeds t and then makes action a; the larger the rate λ, the more precise is this simulation. Notice that the history of the play with strategy π_λ contains a lot of waiting steps that are not in the history of the play with strategy π. Therefore, we need a mapping *destutter* that removes these superfluous waiting steps. We define it inductively by $destutter(s_0) = s_0$ and

$$destutter(\mathfrak{h}'\, ts\, t's') = \begin{cases} destutter(\mathfrak{h}'\, ts)\; t's' & \text{if } s \neq s'\text{; and} \\ destutter(\mathfrak{h}'\, ts) & \text{if } s = s'. \end{cases}$$

Notice that the second case corresponds to the situation when the last step is the waiting step of the strategy π_λ or any self-loop transition.

Furthermore, let a' be the first action taken by π at total time t' for history $destutter(\mathfrak{h})$ if no Markovian transition occurs (notice that strategy π may decide to wait subsequently for several times before it chooses an action; a' is the first action taken by π if none of the waiting is interrupted by a Markovian transition). We finally set

$$\pi_\lambda(\mathfrak{h}) = \begin{cases} \mathrm{Exp}(\lambda) & \text{if either } \sum \mathfrak{h} < t' \text{ or both } t_{n-1} > 0 \text{ and } s_n \neq s_{n-1}; \\ a' & \text{if } \sum \mathfrak{h} \geq t' \text{ and either } t_{n-1} = 0 \text{ or } s_n = s_{n-1}. \end{cases}$$

where $\mathrm{Exp}(\lambda)$ denotes the exponential distribution with rate λ. Notice that the strategy π_λ is from definition λ-exponential.

We now define a set of runs X_λ in the game with π_λ where the imprecision in the simulation does not cause any difference with respect to the time bounded reachability. Let $\delta > 0$ be the grid size of σ and π. A run in the CE game with strategies σ, π_λ belongs to X_λ if for all $k \in \{0, 1, \ldots, T/\delta\}$ we have that

- no non-self-loop transition occurs at the total time neither in the interval $[k\delta, k\delta + \delta/\sqrt{\lambda}]$ nor in the interval $[(k+1)\delta - \delta/\sqrt{\lambda}, (k+1)\delta]$.
- the first transition after total time $k\delta$ is a self-loop transition and occurs in the interval $[k\delta, k\delta + \delta/\sqrt{\lambda}]$;

The proof of the lemma is concluded by the following claim. ◀

▶ Claim 20. For $\lambda \to \infty$ we have

$$\mathcal{P}_{\mathcal{G}}^{\sigma,\pi_\lambda}[X_\lambda] \;\to\; 1 \tag{2}$$

$$\mathcal{P}_{\mathcal{G}}^{\sigma,\pi_\lambda}[\Diamond^{\leq T} G \mid X_\lambda] \;\to\; \mathcal{P}_{\mathcal{G}}^{\sigma,\pi}[\Diamond^{\leq T} G] \tag{3}$$

Proof. As regards (2), we deal with the conditions on runs in X_λ one by one. First, notice that the Lebesgue measure of all the forbidden intervals tend to 0 as λ goes to infinity; hence, the probability of a Markovian transition occurring in any such interval tends to 0. Second, we can underestimate the probability of X_λ by considering only the waiting transitions of π_λ as self-loops. The probability that the waiting transition occurs in each such interval can be bounded by

$$\left(1 - e^{\lambda\cdot\delta/\sqrt{\lambda}}\right)^{T/\delta} \;=\; \left(1 - e^{\sqrt{\lambda}\delta}\right)^{T/\delta} \;\to\; 1$$

since T/δ is constant and $e^{\sqrt{\lambda}\delta} \to 0$ as $\lambda \to \infty$.

As regards (3), notice that the delay caused by the exponential simulation does not qualitatively change the behavior. Namely, under the condition of X_λ,

- any transition made by π is simulated by π_λ at most $\delta/\sqrt{\lambda}$ later; the player **con**cannot interfere meanwhile because if there is an external transition enabled, there cannot be any internal transitions enabled by the Assumption 2;
- also no Markovian transition occurs meanwhile;

- the decision of the players after the delayed transition are the same as in the original play, since both players have grid strategies.

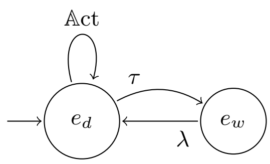
The change is only quantitative because we limit the Markovian transitions, but this change tends to zero as the probability of the set we condition by goes to one. ◀

▶ **Lemma 21.** *An exponential strategy in $\mathcal{G}$ can be simulated by an IMC environment of $\mathcal{C}$, i.e. for any scheduler σ we have*

$$\inf_{\lambda\in\mathbb{R}_{>0},\pi'\in\Pi_\lambda} \mathcal{P}^{\sigma',\pi'}_{\mathcal{G}'}[\Diamond^{\leq T} G] \quad\geq\quad \inf_{E\in\text{ENV},\pi\in\mathfrak{S}(\mathcal{C}(E))} \mathcal{P}^{\pi[\sigma]}_{\mathcal{C}(E)}[\Diamond^{\leq T} G_E]$$

Proof. We fix an arbitrary scheduler σ and use the same strategy σ (observe that a scheduler has the same type as a strategy of the player **con**). Furthermore, we fix an arbitrary $\lambda \in \mathbb{R}_{>0}$ and a λ-exponential strategy π'. We choose E to be a two-state environment with rate λ depicted below and set π to be strategy that in a state (c, e_d) chooses the transition to (c, e_w) if π' chooses exponential waiting, and that chooses the transition to (c', e_d) if π' chooses c' (as a result of a synchronization over some external action a, taking the self-loop in E).

Formally, for a path $\mathfrak{p} = (c_0, e_0)\, t_0\, (c_1, e_1)\, t_1 \cdots t_{n-1}\, (c_n, e_n)$ where each c_i is the state of the IMC component and each $e_i \in \{e_d, e_w\}$ is the state of E, we set



- $\pi(\mathfrak{p}) = (c_n, e_w)$ if $\pi'(proj_1(\mathfrak{p}))$ chooses exponential waiting,
- $\pi(\mathfrak{p}) = (c_{n+1}, e_d)$ if $\pi'(proj_1(\mathfrak{p}))$ chooses $c_{n+1} \in \text{succ}_e(c_n)$

where $proj_1 : \mathbb{P}\text{aths}(C(E_\lambda)) \to \mathbb{H}\text{istories}(\mathcal{G}')$ is the first projection of the path (leaving out the states of the environment).

Again, it remains to show that

$$\mathcal{P}^{\sigma',\pi'}_{\mathcal{G}'}[\Diamond^{\leq T} G] \;=\; \mathcal{P}^{\pi[\sigma]}_{\mathcal{C}(E)}[\Diamond^{\leq T} G_E]$$

The key observation is that any path ending with a state of the form (c, e_w) where the scheduler π cannot do anything is mapped by $proj_1$ on a history where the λ-exponential strategy must wait. Furthermore, in all other situations the decisions of the schedulers and strategies coincide w.r.t. $proj_1$. Again, it is easy to see that for any measurable set of runs X in $\mathcal{G}$ we have $\mathcal{P}^{\sigma',\pi'}_{\mathcal{G}}[X] = \mathcal{P}^{\sigma,\pi}_{C(E_\lambda)}[proj_1^{-1}(X)]$. ◀

Finally, the proof of $(**) \geq (*)$ follows easily from Lemmata 18, 19, and 21 since we have

$$\begin{aligned}\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G}}[\Diamond^{\leq T} G] \;&=\; \sup_{\sigma\in\Sigma_\#}\inf_{\pi\in\Pi_\#} \mathcal{P}^{\sigma,\pi}_{\mathcal{G}}[\Diamond^{\leq T} G] \;=\; \sup_{\sigma\in\Sigma_\#}\inf_{\lambda\in\mathbb{R}_{>0},\pi_\lambda\in\Pi_\lambda} \mathcal{P}^{\sigma,\pi_\lambda}_{\mathcal{G}}[\Diamond^{\leq T} G] \\ &\geq \sup_{\sigma\in\Sigma}\inf_{E\in\text{ENV},\pi\in\mathfrak{S}(\mathcal{C}(E))} \mathcal{P}^{\pi[\sigma]}_{\mathcal{C}(E)}[\Diamond^{\leq T} G_E]\end{aligned}$$

◀

A.5 Value preservation under uniformization

We say that a game $\mathcal{G}$ is *uniform* if all the states with Markovian transitions have the maximal rate λ, i.e. every $s \in S_M \cup S_{e+M}$ satisfies $\mathbf{E}(s) = \lambda$. After showing that the value of $\mathcal{C}(E)$ equals the value of $\mathcal{G}$ we define a uniformized game $\mathcal{G}'$ with value equal to $\mathcal{G}$.

Recall that we assume that there are no Markovian self-loops in $\mathcal{G}$. The game $\mathcal{G}'$ is obtained by adding to each state $s \in S_M \cup S_{e+M}$ with $\mathbf{E}(s) < \lambda$ a Markovian self-loop with rate $\lambda - \mathbf{E}(s)$ yielding an uniform game. The game now contains Markovian self-loops which is not an issue for the following sections.

▶ **Lemma 22.**

$$\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}^{\sigma,\pi}_{\mathcal{G}}[\Diamond^{\leq T}G] \;=\; \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}^{\sigma,\pi}_{\mathcal{G}'}[\Diamond^{\leq T}G]$$

Proof. Recall that thanks to Lemma 14 it is sufficient to consider consistent strategies. For the rest of the proof, let Σ and Π denote consistent strategies. Intuitively, the values equal since the decision of same strategies in both games are completely the same (thanks to the consistency) and the resulting Markov chains have the same probabilities to reach the target by standard arguments.

More formally, we denote by $\mathbf{E}$ and $\mathbf{P}$ the total rate function and the probability matrix of $\mathcal{G}'$, by $\mathbf{E}^-$ and $\mathbf{P}^-$ the total rate function and the probability matrix of $\mathcal{G}$. It is sufficient to show that the k-step values coincide in the two games.

For $k = 0$ it is the same as in Lemma 10. Let $k > 0$ and $s \in S_M$. We proceed by nested induction on the count n of Markovian self-loops that occur before any non-self-loop transition. We denote by $v'_{k,n}(s,t)$ the k-step value of $\mathcal{G}'$ conditioned by the event that *at most* n Markovian self-loops precede the first non-self-loop transition; for $n = 0$, we have directly from the definition of the Markov chain .

$$\begin{aligned}
v'_{k,0}(s,t) &= \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\int_0^\infty \mathbf{E}^-(s)\cdot e^{-\mathbf{E}^-(s)x}\cdot\sum_{s'\in\mathrm{succ}_M(s)}\mathbf{P}^-(s,s')\cdot\mathcal{P}^{\sigma,\pi}_{\mathcal{G}',s\,x\,s'}[\Diamond^{\leq T-t-x}_{\leq k-1}G]\,dx\\
&= \int_0^\infty \mathbf{E}^-(s)\cdot e^{-\mathbf{E}^-(s)x}\cdot\sum_{s'\in\mathrm{succ}_M(s)}\mathbf{P}^-(s,s')\cdot\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\mathcal{P}^{\sigma,\pi}_{\mathcal{G}',s\,x\,s'}[\Diamond^{\leq T-t-x}_{\leq k-1}G]\,dx\\
&= \int_0^\infty \mathbf{E}^-(s)\cdot e^{-\mathbf{E}^-(s)x}\cdot\sum_{s'\in\mathrm{succ}_M(s)}\mathbf{P}^-(s,s')\cdot v_{k-1}(s',t+x)\,dx\\
&= \int_0^{T-t} \mathbf{E}^-(s)\cdot e^{-\mathbf{E}^-(s)x}\cdot\sum_{s'\in\mathrm{succ}_M(s)}\mathbf{P}^-(s,s')\cdot v_{k-1}(s',t+x)\,dx\\
&= v_k(s,t).
\end{aligned}$$

For $n > 0$, we have

$$\begin{aligned}
v'_{k,n}(s,t) &= \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi}\int_0^\infty \mathbf{E}(s)\cdot e^{-\mathbf{E}(s)x}\cdot\Bigg(\sum_{s'\in\mathrm{succ}_M(s),s'\neq s}\mathbf{P}(s,s')\cdot\mathcal{P}^{\sigma,\pi}_{\mathcal{G}',s\,x\,s'}[\Diamond^{\leq T-t-x}_{\leq k-1}G]\\
&\qquad\qquad + \mathbf{P}(s,s)\cdot\mathcal{P}^{\sigma,\pi}_{\mathcal{G}',s\,x\,s}[\Diamond^{\leq T-t-x}_{\leq k}G]\Bigg)\,dx,\\
&= \int_0^\infty \mathbf{E}(s)\cdot e^{-\mathbf{E}(s)x}\cdot\sum_{s'\in\mathrm{succ}_M(s),s'\neq s}\mathbf{P}(s,s')\cdot v_{k-1}(s',t+x)\,dx \;+\\
&\quad \int_0^\infty \mathbf{E}(s)\cdot e^{-\mathbf{E}(s)x}\cdot\mathbf{P}(s,s)\cdot v'_{k,n-1}(s,t+x)\,dx.
\end{aligned}$$

By the induction hypothesis, $v'_{k,n-1}(s,t+x) = v_k(s,t+x)$; hence,

$$\begin{aligned}
&= \int_0^\infty \mathbf{E}(s)\cdot e^{-\mathbf{E}(s)x}\cdot\sum_{s'\in\mathrm{succ}_M(s)}\mathbf{P}^-(s,s')\cdot\frac{\mathbf{E}^-(s)}{\mathbf{E}(s)}\cdot v_{k-1}(s',t+x)\,dx \;+\\
&\quad \int_0^\infty \mathbf{E}(s)\cdot e^{-\mathbf{E}(s)x}\cdot\mathbf{P}(s,s)\cdot\\
&\qquad \left(\int_0^\infty \mathbf{E}^-(s)\cdot e^{-\mathbf{E}^-(s)y}\cdot\sum_{s'\in\mathrm{succ}_M(s)}\mathbf{P}^-(s,s')\cdot v_{k-1}(s',t+x+y)\,dy\right)dx.
\end{aligned}$$

Since the density function of the convolution of two exponential distributions with rates $\mathbf{E}(s)$ and $\mathbf{E}^-(s)$ equals $\mathbf{E}(s)\cdot\mathbf{E}^-(s)\cdot(e^{-\mathbf{E}^-(s)}/\xi - e^{-\mathbf{E}(s)}/\xi)$ where $\xi = \mathbf{E}(s) - \mathbf{E}^-(s)$ is the rate of the self-loop transition, and since $\mathbf{P}(s,s)\cdot\mathbf{E}(s) = \xi$, we have

$$\begin{aligned}
&= \int_0^\infty \mathbf{E}^-(s)\cdot e^{-\mathbf{E}(s)x}\cdot \sum_{s'\in \mathrm{succ}_M(s)} \mathbf{P}^-(s,s')\cdot v_{k-1}(s',t+x)dx \ + \\
&\quad \int_0^\infty \mathbf{E}^-(s)\cdot \left(e^{-\mathbf{E}^-(s)x} - e^{-\mathbf{E}(s)x}\right)\cdot \sum_{s'\in \mathrm{succ}_M(s)} \mathbf{P}^-(s,s')\cdot v_{k-1}(s',t+x)dx \\
&= \int_0^\infty \mathbf{E}^-(s)\cdot e^{-\mathbf{E}^-(s)x}\cdot \sum_{s'\in \mathrm{succ}_M(s)} \mathbf{P}^-(s,s')\cdot v_{k-1}(s',t+x)dx \\
&= v_k(s,t).
\end{aligned}$$

This proves that $v_k(s,t) = v'_k(s,t)$ since $v'_{k,n}(s,t) \to v'_k(s,t)$ as $n\to\infty$.

For $s \in S_{e+M}$ it is similar (only more technical). Let w we the delay chosen by a strategy π in state s. Until leaving s or until time w, the system behaves again as a CTMC (thanks to the consistency of π), yielding again $v_k(s,t) = v'_k(s,t)$ for each $n \in \mathbb{N}_0$. ◀

Thanks to this value preservation, we will assume in the following sections that $\mathcal{G}$ is uniform (and may contain Markovian self-loops). We will denote all rates by λ.

A.6 Definition of the discrete game

Let us fix a discretization step $\delta > 0$ that divides the time bound T into $k \in \mathbb{N}$ pieces of equal length, i.e. $T = k\cdot\delta$. We construct a discrete game $\Delta = (V, \mapsto, (V_\Box, V_\Diamond, V_\bigcirc), \mathit{Prob}, v_0)$ for two players $\Box$ and $\Diamond$, who have antagonistic objectives, and a random player $\bigcirc$. The game is played on a finite graph with vertices V and edges $\mapsto$. The vertices are assigned to players according to the partitioning $V_\Box \cup V_\Diamond \cup V_\bigcirc = V$, i.e. s is assigned to $\odot$ if $s \in V_\odot$, where $\odot \in \{\Box, \Diamond, \bigcirc\}$. A play starts in the initial vertex v_0 and moves step by step from vertex to vertex forming an infinite sequence $v_0v_1v_2\cdots$, called a *run*. In each vertex v_i the assigned player chooses a successor v_{i+1} such that $v_i \mapsto v_{i+1}$. The player $\bigcirc$ chooses the successor randomly according to the fixed distribution $\mathit{Prob}(v_i)$, the players $\Box$ and $\Diamond$ choose according to their strategies. In our setting, *strategy* is a function that assigns to each finite path $v_0\cdots v_i$ a successor vertex v_{i+1}. A pair of such strategies $\sigma \in \Sigma_\Delta$ and $\pi \in \Pi_\Delta$ of respective players $\Box$ and $\Diamond$ and a vertex v determine a probability measure $\mathcal{P}^{\sigma,\pi}_{\Delta,v}$ over the measurable space of all runs such that the play is started in vertex v. For a formal definition, see, e.g., [17].

Formally, the set of vertices is $V = S \cup \{\bar{s} \mid s \in S \wedge \mathrm{succ}_e(s) \neq \emptyset\}$. The initial vertex is $v_0 = s_0$. The partitioning $(V_\Box, V_\Diamond, V_\bigcirc)$ and the transition function $\mapsto$ are constructed as follows.

(i) For each $s \in S_\tau$, we set $s \in V_\Box$ and $s \mapsto s'$ for $s' \in \mathrm{succ}_\tau(s)$.
(ii) For each $s \in S_M$, we set $s \in V_\bigcirc$ and $s \mapsto s'$ for $s' \in \mathrm{succ}_M(s) \cup \{s\}$.
(iii) For each $s \in S_{e+M}$, we set $s \in V_\Diamond$, $s \mapsto s'$ for $s' \in \mathrm{succ}_e(s) \cup \{\hat{s}\}$, $\hat{s} \in V_\bigcirc$, $\hat{s} \mapsto s'$ for $s' \in \mathrm{succ}_M(s) \cup \{s\}$.
(iv) For each $s \in S_e$, we set $s \in V_\Diamond$, $s \mapsto s'$ for $s' \in \mathrm{succ}_e(s) \cup \{\hat{s}\}$, $\hat{s} \in V_\bigcirc$, $\hat{s} \mapsto s$.

The distribution Prob choosing the successors in the states of $V_\bigcirc$ is set as follows.

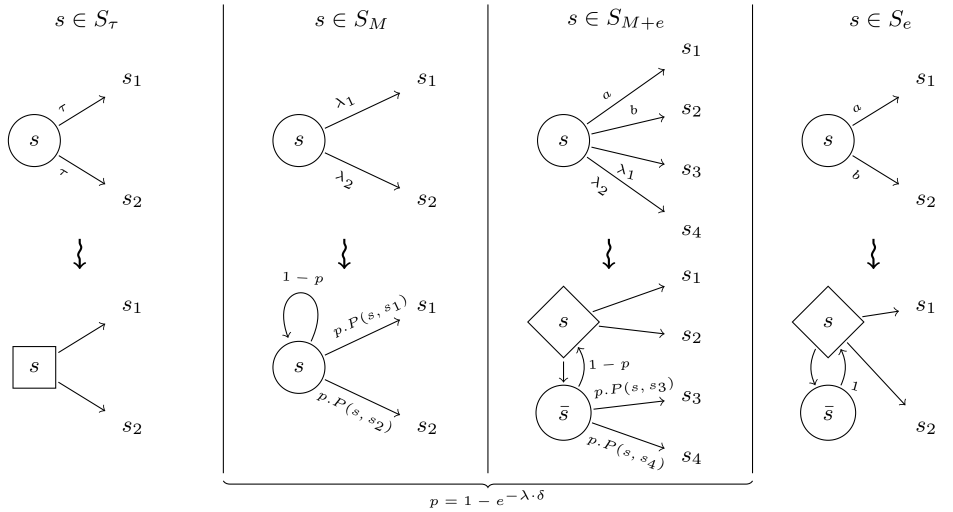


Figure 2 Four gadgets for transforming a CE game into a discrete game. The upper part shows types of states in the CE game, the lower part shows corresponding gadgets in the discrete game. In the lower part, the square-shaped vertices belong to player $\Box$, the diamond-shaped vertices to player $\Diamond$ and the circle-shaped vertices to player $\bigcirc$.

- For each t equal to s of item (ii) or $\hat{s}$ of item (iii), and t' such that $t \mapsto t'$, we set

$$Prob(t,t') = \begin{cases} (1-e^{-\lambda\cdot\delta})\cdot\mathbf{P}(t,t') & \text{if } t \neq t' \\ (1-e^{-\lambda\cdot\delta})\cdot\mathbf{P}(t,t) + e^{-\lambda)\cdot\delta} & \text{if } t = t'. \end{cases}$$

- For each $\bar{s}$ of item (iv), we set $Prob(\bar{s}, s) = 1$.

A.7 Value in the discrete game

Recall that each discrete step from a vertex in $V_{\bigcirc}$ intuitively "takes" time $\delta(= T/N)$ whereas each step from vertex in $V_\Box \cup V_\Diamond$ "takes" zero time. The winning condition is the *step-bounded reachability* of the set of vertices G in up to N steps where only the steps from the vertices $V_{\bigcirc}$ are counted for this limit. Formally,

$$\Diamond^{\#_\circ \leq N} G = \{v_0 v_1 \cdots \mid \exists n \ (v_n \in G \text{ and } \#_\circ(v_0 \cdots v_{n-1}) \leq N)\}$$

where $\#_\circ(v_0 \cdots v_{n-1}) = |\{\, i \mid 0 \leq i < n, v_i \in V_{\bigcirc} \}|$. Further, let us recall that we are interested in the lower value of this game

$$\sup_{\sigma\in\Sigma_\Delta} \inf_{\pi\in\Pi_\Delta} \mathcal{P}^{\sigma,\pi}_{\Delta}\left[\Diamond^{\#_\circ \leq N} G\right] \qquad (***)$$

Finally, recall that we say that a strategy *is counting* if it only considers the last vertex and the current count $\#_\circ$. We may view it as a function $V \times \{0,\ldots,N\} \to V$ since it is irrelevant what it does after more than N steps.

Observe that the value in game Δ is equivalent to the value in the following game Δ' with standard step-bounded reachability objective. In Δ' we enhance the state space with a counter $0 \leq k \leq N$ representing the count $\Diamond^{\#_\circ \leq N}$ of steps from random vertices, and by introducing a new sink (non-goal) state $\underline{s}$ which represents counting above N. Indeed, it is sufficient to consider reachability up to $|S| \cdot N$ steps since from Assumption 1 there cannot be more than $|S|$ steps without a Markovian transition, i.e. without increasing the counter k. Notice that in discrete reachability games it is sufficient to consider memoryless deterministic strategies [17]. Furthermore, notice that memoryless deterministic strategies in Δ' correspond bijectively to counting deterministic strategies in Δ (observe that we actually defined strategies to be deterministic in Δ which is now justified by this bijection). Finally, notice that the game Δ' is acyclic (except for the sink state $\underline{s}$ where the value is trivially 0) which allows to compute the value and the optimal memoryless deterministic by straightforward $(|S| \cdot N)$-fold value iteration in polynomial time as summarized by the following lemma.

Lemma 7. There are counting strategies optimal in $(***)$. Moreover, they can be computed together with $(***)$ in time $\mathcal{O}(N|V|^2)$.

Similarly to v, we now need to characterize the value in the discrete game Δ constructed for a given δ. Let

$$v_\delta(s,k) = \sup_{\sigma \in \Sigma_\Delta} \inf_{\pi \in \Pi_\Delta} \mathcal{P}^{\sigma,\pi}_{\Delta,s}\left[\Diamond^{\#_{\circ} \leq N-k} G\right]$$

denote the value when starting in vertex s with remaining $N-k$ random steps (or equivalently, the value in Δ' from the vertex (s,k)). The following characterization is straightforward from the definition of Δ and Δ' and from the discussion above.

▶ **Lemma 24.** *The function v_δ satisfies for any $k \in \mathbb{N}_0$ and $s \in S$ the following. First, $v_\delta(s,k) = 0$ if $k > N$ or $s \in S_e$, and $v(s,k) = 1$ if $s \in G$ and $k \leq N$. Second, for any $s \notin G$ and $k < N$ it holds*

$$v_\delta(s,k) = \begin{cases} \max\limits_{s' \in \mathrm{succ}_\tau(s)} v_\delta(s',k) & \text{if } s \in S_\tau, \\ \mathcal{B} + e^{-\lambda\delta} \cdot v_\delta(s,k+1) & \text{if } s \in S_M, \\ \min\left(\mathcal{B} + e^{-\lambda\delta} \cdot v_\delta(s,k+1), \min\limits_{s' \in \mathrm{succ}_e(s)} v_\delta(s',k\delta)\right) & \text{if } s \in S_{e+M}, \end{cases}$$

where $\mathcal{B} = (1 - e^{-\lambda\delta}) \cdot \sum_{s'' \in S} \mathbf{P}(s,s'') \cdot v_\delta(s'',k+1)$.

A.8 Relating $(**)$ and $(***)$

Some parts of proofs in this section are also inspired by [28]. Recall that we assume that $\mathcal{G}$ is uniform thanks to Lemma 22. Further, let us recall how a scheduler $\overline{\sigma}$ in the IMC $\mathcal{C}$ is defined using a counting optimal strategy σ in the discrete game Δ constructed for some error $\varepsilon > 0$.

For $\sigma : V \times \{0,\dots,N\} \to V$, we have for any $\mathfrak{p} = s_0\, t_0 \cdots s_{n-1}\, t_{n-1}\, s_n$

$$\overline{\sigma}(\mathfrak{p}) = \sigma(s_n, \lceil (t_0 + \ldots + t_{n-1})/\delta \rceil).$$

Since a scheduler in $\mathcal{C}$ is also a strategy of the player **con**in the CE game $\mathcal{G}$, we use the same symbol $\overline{\sigma}$ also in the CE game setting.

▶ **Lemma 25.** *For any $\delta > 0$, $s \in S$ and any $0 \leq k \leq N$ we have*

$$v_\delta(s,k) \;\leq\; v(s,k\delta) \tag{4}$$

The strategy $\overline{\sigma}$ guarantees in $\mathcal{G}$ at least as much as the value in Δ, i.e.

$$v_\delta(s,k) \;\leq\; \inf_{\pi \in \Pi} \mathcal{P}^{\overline{\sigma},\pi}_{\mathcal{G},s}\left[\Diamond^{\leq T-k\delta} G\right] \tag{5}$$

Proof. We prove it by backwards induction on k relying heavily on the value characterizations in Lemmata 13 and 24. For $k = N$, (4) is immediate from Lemmata 13 and 24. Furthermore, any strategy $\overline{\sigma}$ satisfies (5). Let us fix $0 \leq k < N$.

For $s \in G \cup S_e$, (4) is immediate from Lemmata 13 and 24 and again, any strategy $\overline{\sigma}$ satisfies (5). For the following, we assume $s \notin G$.

For $s \in S_M$, first observe that the function $\mathcal{P}^{\overline{\sigma},\pi}_{\mathcal{G},s}\left[\Diamond^{\leq T-x} G\right]$ is monotonous w.r.t. x for any strategies $\sigma \in \Sigma$, $\pi \in \Pi$ and any starting state $s \in S$; and thus $v(s,\cdot)$ is also monotonous by definition. We get (4) by replacing $v(s'',(k+1)\delta)$ for $v(s'',k\delta+x)$ and applying the induction hypothesis

$$
\begin{aligned}
v(s,k\delta) &= \int_0^\delta \lambda e^{-\lambda x} \cdot \sum_{s''\in S} \mathbf{P}(s,s'') \cdot v(s'',k\delta+x)dx + e^{-\lambda\delta} \cdot v(s,(k+1)\delta) \\
&\geq (1-e^{-\lambda\delta}) \cdot \sum_{s''\in S} \mathbf{P}(s,s'') \cdot v(s'',(k+1)\delta) + e^{-\lambda\delta} \cdot v(s,(k+1)\delta) \\
&\geq (1-e^{-\lambda\delta}) \cdot \sum_{s''\in S} \mathbf{P}(s,s'') \cdot v_\delta(s'',k+1) + e^{-\lambda\delta} \cdot v_\delta(s,k+1) \\
&= v_\delta(s,k).
\end{aligned}
$$

As regards (5), $\overline{\sigma}$ does not guarantee $v(s'',k\delta+x)$ in state s'' and elapsed time $k\delta+x$ because its possible decision $\sigma(s'',\lceil k\delta+x\rceil)$ may not be optimal for the elapsed time $k\delta+x$. Nevertheless, again from monotonicity of $\mathcal{P}^{\overline{\sigma},\pi}_{\mathcal{G},s}\left[\Diamond^{\leq T-x} G\right]$ and from the induction hypothesis, $\overline{\sigma}$ guarantees there at least $v_\delta(s'',(k+1))$. From the definition of $\mathcal{P}^{\overline{\sigma},\pi}_{\mathcal{G},s}$, $\overline{\sigma}$ guarantees in state s and elapsed time $k\delta$ at least

$$
\int_0^\delta \lambda e^{-\lambda x} \cdot \sum_{s''\in S} \mathbf{P}(s,s'') \cdot v_\delta(s'',(k+1)\delta)dx + e^{-\lambda\delta} \cdot v_\delta(s,(k+1)\delta) \ = v_\delta(s,k).
$$

Now we deal with the states where potentially no time is spent, i.e. $s \in S_{e+M} \cup S_\tau$. For these states, we need deal with the value for all elapsed times not only for multiples of δ. We define

$$
v_\downarrow(s,k\delta+w) = \begin{cases} \max_{s'\in \mathrm{succ}_\tau(s)} v_\downarrow(s',k\delta+w) & \text{if } s \in S_\tau \\ \mathcal{C}(s,w) & \text{if } s \in S_M \\ \min\{\mathcal{C}(s,w), \min_{s'\in \mathrm{succ}_e(s)} v_\downarrow(s',k\delta+w)\} & \text{if } s \in S_{e+M} \end{cases}
$$

where $\mathcal{C}(s,w) = e^{-\lambda(\delta-w)} v_\delta(s,k+1) + (1-e^{-\lambda(\delta-w)}) \sum_{s''\in S} \mathbf{P}(s,s'') v_\delta(s'',k+1)$. We need the following claim in order to prove (4) and (5).

▶ Claim 26. For any $0 \leq w \leq \delta$ it holds $v(s,k\delta+w) \geq v_\downarrow(s,k\delta+w)$.

We prove the claim as well as (4) and (5) by a nested induction on the length n of the longest path via internal and external transitions to a state in $S_M \cup S_e \cup G$. Such length is bounded for any state by Assumption 1. For $n = 0$, i.e. for $s \in S_M$, we already discussed (4) and (5), the claim can be obtained by similar arguments as above and as in Lemma 13. Now let $n > 0$.

- For $s \in S_\tau$, (4) is easy from Lemmata 13 and 24. Furthermore, the strategy $\overline{\sigma}$ which takes any transition that maximizes the value $v_\delta(s,k\delta)$ satisfies (5). Furthermore, the claim can be obtained by similar arguments as in Lemma 13.
- Finally, for $s \in S_{e+M}$ it suffices to prove the claim. Indeed, (5) is trivial for $s \in S_{e+M}$ from the induction hypothesis as $\overline{\sigma}$ is not involved in the first step. Furthermore, from

the claim we obtain (4) by setting $w = 0$ and by observing that $v_\downarrow(s, k\delta) = v_\delta(s, k)$ by definition,

$$v(s, k\delta) \geq \min\{\mathcal{C}(s, 0), \min_{s' \in \mathrm{succ}_e(s)} v_\downarrow(s', k\delta)\} = \min\{\mathcal{C}(s, 0), \min_{s' \in \mathrm{succ}_e(s)} v_\delta(s', k)\} = v_\delta(s, k).$$

As regards the claim, observe that we can obtain by similar arguments as in Lemma 13

$$\begin{aligned} v(s, k\delta + w) = \min\{&\mathcal{A}(\delta - w) + e^{-\lambda(\delta - w)} \cdot v(s, (k + 1)\delta), &(6)\\ &\min_{\substack{s' \in \mathrm{succ}_e(s),\\ 0 \leq w' \leq \delta - w}} \Big(e^{-\lambda w'} \cdot v(s', k\delta + w + w') + \mathcal{A}(w')\Big)\} &(7) \end{aligned}$$

We prove the claim by showing that (6) $\geq v_\downarrow(s, k\delta + w)$ and (7) $\geq v_\downarrow(s, k\delta + w)$. First,

$$\begin{aligned} (6) &= \int_0^{\delta - w} \lambda e^{-\lambda x} \cdot \sum_{s'' \in S} \mathbf{P}(s, s'') \cdot v(s'', k\delta + x)\, dx + e^{-\lambda(\delta - w)} \cdot v(s, (k + 1)\delta)\\ &\geq (1 - e^{-\lambda(\delta - w)}) \cdot \sum_{s'' \in S} \mathbf{P}(s, s'') \cdot v(s'', (k + 1)\delta) + e^{-\lambda(\delta - w)} \cdot v(s, (k + 1)\delta)\\ &\geq (1 - e^{-\lambda(\delta - w)}) \cdot \sum_{s'' \in S} \mathbf{P}(s, s'') \cdot v_\delta(s'', k + 1) + e^{-\lambda(\delta - w)} \cdot v_\delta(s, k + 1)\\ &= \mathcal{C}(s, w) \geq v_\downarrow(s, k\delta + w). \end{aligned}$$

Let us focus on (7). We fix $s' \in \mathrm{succ}_e(s)$ and search for the minimal w. First, we deal with Markovian successor $s' \in S_M$. It holds

$$\begin{aligned} &\min_{0 \leq w' \leq \delta - w} \Big(e^{-\lambda w'} \cdot v(s', k\delta + w + w') + \mathcal{A}(w')\Big) \qquad (\star)\\ \geq &\min_{0 \leq w' \leq \delta - w} \Big(e^{-\lambda w'} \cdot v_\downarrow(s', k\delta + w + w') + \mathcal{A}(w')\Big)\\ = &\min_{0 \leq w' \leq \delta - w} \Big(e^{-\lambda w'} \cdot \mathcal{C}(s', w + w') + \mathcal{A}(w')\Big)\\ \geq &\min_{0 \leq w' \leq \delta - w} \Bigg(e^{-\lambda w'} \cdot \Bigg(e^{-\lambda(\delta - w - w')} v_\delta(s', k + 1) + (1 - e^{-\lambda(\delta - w - w')}) \sum_{s'' \in S} \mathbf{P}(s', s'') v_\delta(s'', k + 1)\Bigg)\\ &\quad + (1 - e^{-\lambda w'}) \sum_{s'' \in S} \mathbf{P}(s, s'') v_\delta(s'', k + 1)\Bigg)\\ = \ & e^{-\lambda(\delta - w)} \Bigg(v_\delta(s', k + 1) - \sum_{s'' \in S} \mathbf{P}(s', s'') v_\delta(s'', k + 1)\Bigg) +\\ &\min_{0 \leq w' \leq \delta - w} \Bigg(e^{-\lambda w'} \sum_{s'' \in S} \mathbf{P}(s', s'') v_\delta(s'', k + 1) + (1 - e^{-\lambda w'}) \sum_{s'' \in S} \mathbf{P}(s, s'') v_\delta(s'', k + 1)\Bigg). \end{aligned}$$

Now comes the crucial observation of the proof: the formula above is minimized either for $w' = 0$ or for $w' = \delta - w$ because of its linearity. By setting $w' = 0$, we get

$$e^{-\lambda(\delta - w)} v_\delta(s', k + 1) + (1 - e^{-\lambda(\delta - w)}) \sum_{s'' \in S} \mathbf{P}(s', s'') v_\delta(s'', k + 1) = \mathcal{C}(s', w).$$

By setting $w' = \delta - w$ and from the definition of v_δ, we get

$$e^{-\lambda(\delta - w)} v_\delta(s', k + 1) + (1 - e^{-\lambda(\delta - w)}) \sum_{s'' \in S} \mathbf{P}(s, s'') v_\delta(s'', k + 1)$$

$$\geq\ e^{-\lambda(\delta-w)} v_\delta(s,k+1) + (1-e^{-\lambda(\delta-w)}) \sum_{s''\in S} \mathbf{P}(s,s'') v_\delta(s'',k+1) = \mathcal{C}(s,w).$$

Now, we consider successor $s' \in S_{e+M}$. It holds

$$(\star) \geq \min_{0\leq w'\leq \delta-w} \left(e^{-\lambda w'} \cdot v_{\downarrow}(s', k\delta+w+w') + \mathcal{A}(w') \right)$$

$$= \min_{0\leq w'\leq \delta-w} \left(e^{-\lambda w'} \cdot \min\left\{ \mathcal{C}(s', w+w'), \min_{s''\in \mathrm{succ}_e(s')} v_{\downarrow}(s'', k\delta+w+w') \right\} + \mathcal{A}(w') \right).$$

Observe, that similarly for $s' \in S_\tau$, we can obtain

$$(\star) \geq \min_{0\leq w'\leq \delta-w} \left(e^{-\lambda w'} \cdot \max_{s''\in \mathrm{succ}_\tau(s')} v_{\downarrow}(s'', k\delta+w+w') + \mathcal{A}(w') \right).$$

For both the cases $s' \in S_\tau \cup S_{e+M}$, we get by unraveling the internal and external transitions of $v_\downarrow$ according to definition

$$= \min_{0\leq w'\leq \delta-w} \left(e^{-\lambda w'} \cdot \operatorname*{minimax_{s'}}_{s''\in S_M\cup S_{e+M}} \mathcal{C}(s'', w+w') + \mathcal{A}(w') \right)$$

where $\mathrm{minimax}_{s'}$ is an operator that chooses a state with Markovian transitions that is reachable from s' (where s' is also reachable from s') that is the most suitable for both players. Technically, it is a sequence of min and max operators that are combined according to the transition structure of $\mathcal{G}$. Similarly to the previous case, it is easy to show that for any fixed $s'' \in S_M \cup S_{e+M}$ it holds $\min_{0\leq w'\leq\delta-w} \left(e^{-\lambda w'} \cdot \mathcal{C}(s'', w+w') + \mathcal{A}(w')\right) \geq \min\{\mathcal{C}(s'',w), \mathcal{C}(s,w)\}$. To sum up all the cases,

$$(7) \geq \min_{s'\in \mathrm{succ}_e(s)} \min \left\{ \mathcal{C}(s,w), \operatorname*{minimax_{s'}}_{s''\in S_M\cup S_{e+M}} \mathcal{C}(s'',w) \right\} = \min \left\{ \mathcal{C}(s,w), \operatorname*{minimax_{s}}_{s''\in S_M\cup S_{e+M}} \mathcal{C}(s'',w) \right\},$$

which is by considering only the first "decision" in $\mathrm{minimax}_s$ again

$$= \min\{\mathcal{C}(s,w), \min_{s'\in \mathrm{succ}_e(s)} v_\downarrow(s', k\delta+w)\} = v_\downarrow(s, k\delta+w). \qquad ◀$$

In the following lemma we denote by $\Pi_{\#,\delta}$ the set of grid strategies on a grid of size δ.

▶ **Lemma 27.** *For any $\varepsilon > 0$, $\delta \leq 2\varepsilon/(\lambda^2 T)$, $s \in S$ and any $k \in \mathbb{N}_0$ we have*

$$v(s,k\delta)\ \leq v_\delta(s,k) + \varepsilon \cdot \frac{T-k\delta}{T} \tag{8}$$

$$\sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi_{\#,\delta}} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}\left[\Diamond^{\leq T-k\delta} G\right]\ \leq v_\delta(s,k) + \varepsilon \cdot \frac{T-k\delta}{T}. \tag{9}$$

Proof. Again, we prove it by backwards induction on k. For $k = N$, (8) is immediate from Lemmata 13 and 24. Let us fix $0 \leq k < N$. For $s \in G \cup S_e$, (8) is immediate from Lemmata 13 and 24. Observe that in these two situations, it is easy to derive (9) by very similar arguments as in Lemma 13. For the following, we assume $s \notin G$.

For $s \in S_M$ we need to analyze first, what is the probability that two or more Markovian transitions are taken in the interval $[0,\delta]$. Let us denote by $\mathrm{Runs}_{\#_{[0,\delta]}>1}$ the set of runs where more than one Markovian transition occurs in the first δ time units. Furthermore, we denote by $\Diamond^{\leq T}_{\#_{[0,\delta]}\leq 1} G = \Diamond^{\leq T} G \setminus \mathrm{Runs}_{\#_{[0,\delta]}>1}$ and $\Diamond^{\leq T}_{\#_{[0,\delta]}>1} G = \Diamond^{\leq T} G \cap \mathrm{Runs}_{\#_{[0,\delta]}>1}$. Next, we have

$$\begin{aligned}
v(s,k\delta) &= \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}\left[\Diamond^{\leq T-k\delta} G\right] \\
&= \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}\left[\Diamond^{\leq T-k\delta}_{\#_{[0,\delta]}\leq 1} G \uplus \Diamond^{\leq T-k\delta}_{\#_{[0,\delta]}>1} G\right] \\
&= \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi} \left(\mathcal{P}^{\sigma,\pi}_{\mathcal{G},(s,k\delta)}\left[\Diamond^{\leq T-k\delta}_{\#_{[0,\delta]}\leq 1} G\right] + \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}\left[\Diamond^{\leq T-k\delta}_{\#_{[0,\delta]}>1} G\right]\right) \\
&\leq \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi} \left(\mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}\left[\Diamond^{\leq T-k\delta}_{\#_{[0,\delta]}\leq 1} G\right] + \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}\left[\mathrm{Runs}_{\#_{[0,\delta]}>1}\right]\right) \\
&= \sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}\left[\Diamond^{\leq T-k\delta}_{\#_{[0,\delta]}\leq 1} G\right] + \mathcal{P}^{\sigma^*,\pi^*}_{\mathcal{G},s}\left[\mathrm{Runs}_{\#_{[0,\delta]}>1}\right]
\end{aligned}$$

where σ^* and π^* are arbitrary strategies. Indeed, notice that strategies in $\mathcal{G}$ have no influence on the frequency of occurrence of Markovian transitions.

$$\leq \underbrace{\sup_{\sigma\in\Sigma}\inf_{\pi\in\Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}\left[\Diamond^{\leq T}_{\#_{[0,\delta]}\leq 1} G\right]}_{v_-(s,k\delta)} + \frac{(\lambda\delta)^2}{2}$$

which follows from the properties of the Poisson distribution with parameter $\lambda\delta$ using the very same arguments as in [28, Lemma 6.2]. Finally, by $\delta \leq 2\varepsilon/(\lambda^2 T)$,

$$\leq v_-(s,k\delta) + \varepsilon \cdot \frac{\delta}{T} \tag{10}$$

This observation allows us to focus on paths where at most one Markovian transition occurs.

$$v_-(s,k\delta) = \int_0^\delta \lambda e^{-\lambda x} \cdot \sum_{s''\in S} \mathbf{P}(s,s'') \cdot \kappa(s'',\delta-x) v_{/\delta-x}(s'',k\delta+x)dx + e^{-\lambda\delta} \cdot v(s,(k+1)\delta)$$

where $\kappa(s'',\delta-x)$ is the probability that (after the first Markovian transition occurs at time x), no other Markovian transition occurs in the remaining time $\delta - x$; and $v_{/\delta-x}(s'',k\delta+x)$ denotes the value in configuration $(s'',k\delta+x)$ with no Markovian transition occurring in the next $\delta - x$ time units. Now we show

$$v_{/\delta-x}(s'',k\delta+x) \leq v(s'',(k+1)\delta) \qquad \text{for any } 0 \leq x \leq \delta. \tag{11}$$

We proceed by nested induction on the length j of the longest path via internal and external transitions to a state in $S_M \cup S_e \cup G$. Such length is bounded for any state by the Assumption 1.

Let $j = 0$. If $s'' \in S_M$, it is clear since the game waits there for sure until time $(k+1)\delta$. If $s'' \in G$, both values equal to 1, and if $s'' \in S_e$, the strategy that waits until time $(k+1)\delta$ gains value $v(s'',(k+1)\delta) = 0$ which is optimal.

Assume $j > 0$,

- if $s'' \in S_\tau$, the player chooses successor state s''' with maximal $v_{/\delta-x}(s''',k\delta+x)$, respectively. Thus, for such s''', $v_{/\delta-x}(s'',k\delta+x)$ equals $v_{/\delta-x}(s''',k\delta+x)$ for which we can use the induction hypothesis;

- if $s'' \in S_{e+M}$, the value equals

$$\min\left(v(s'',(k+1)\delta), \min_{\substack{s'''\in \mathrm{succ}_e(s''),\\ w\in[0,\delta-x]}} v_{/\delta-x-w}(s''',k\delta+x+w)\right) \le v(s'',(k+1)\delta).$$

By (10), (11) and since $0 \le \kappa \le 1$, we get

$$v_{-}(s,k\delta) \le \int_0^\delta \lambda e^{-\lambda x} \cdot \sum_{s''\in S} \mathbf{P}(s,s'') \cdot v(s'',(k+1)\delta)\,dx + e^{-\lambda\delta} \cdot v(s,(k+1)\delta)$$

and further by integrating the density and by applying the induction hypothesis

$$\begin{aligned}
&\le (1-e^{-\lambda\delta}) \cdot \sum_{s''\in S} \mathbf{P}(s,s'') \cdot v_\delta(s'',k+1) + e^{-\lambda\delta} \cdot v_\delta(s,k+1) + \varepsilon \cdot \frac{T-(k+1)\delta}{T}\\
&\le v_\delta(s,k) + \varepsilon \cdot \frac{T-(k+1)\delta}{T}
\end{aligned}$$

which in total yields the sought form $v(s,k\delta) \le v_\delta(s,k) + \varepsilon \cdot \frac{T-k\delta}{T}$. Similarly, as regards (9), we can denote $v_{\#}(s,k\delta) := \sup_{\sigma\in\Sigma} \inf_{\pi\in\Pi_{\#,\delta}} \mathcal{P}^{\sigma,\pi}_{\mathcal{G},s}\left[\Diamond^{\le T-k\delta} G\right]$, and analogously define $v_{\#-}$ and $v_{\#/\delta-x}$. By similar arguments it can be shown that $v_{\#}(s,k\delta) \le v_{\#-}(s,k\delta) + \varepsilon\delta/T$, $v_{\#/\delta-x}(s'',k\delta+x) \le v_{\#}(s'',(k+1)\delta)$ for any $0 \le x \le \delta$ (observe that a grid strategy that for any x chooses to wait until the end of the interval guarantees the value $v_{\#}(s'',(k+1)\delta)$). From these inequalities it is again easy to derive that $v_{\#}(s,k\delta) \le v_\delta(s,k) + \varepsilon \cdot \frac{T-k\delta}{T}$.

Now, let us proceed to the states where potentially no time is spent, i.e. $s \in S_{e+M} \cup S_\tau$. Again, we need to use a nested induction on the length of the longest path to a state in $S_M \cup S_e \cup G$. For length 0, we have already proven it. Assume length > 0,

- for $s \in S_\tau$ we have

$$\begin{aligned}
v(s,k\delta) &= \max_{s'\in \mathrm{succ}_\tau(s)} v(s',k\delta)\\
&\le \max_{s'\in \mathrm{succ}_\tau(s)} \left(v_\delta(s',k\delta) + \varepsilon \cdot \frac{T-k\delta}{T}\right)\\
&= v_\delta(s,k\delta) + \varepsilon \cdot \frac{T-k\delta}{T};
\end{aligned}$$

  By similar arguments as in Lemma 13 it is easy to see that $v_{\#}(s,k\delta) = \max_{s'\in\mathrm{succ}_\tau(s)} v_{\#}(s',k\delta)$ and thus also $v_{\#}(s,k\delta) \le v_\delta(s,k\delta) + \varepsilon(T-k\delta)/T$.

- for $s \in S_{e+M}$ we get

$$v(s,k\delta) = \min\left(\mathcal{A}(\delta) + e^{-\lambda\delta} \cdot v(s,(k+1)\delta), \min_{\substack{s'\in\mathrm{succ}_e(s),\\ 0\le w<\delta}} \left(e^{-\lambda w} \cdot v(s',k\delta+w) + \mathcal{A}(w)\right)\right)$$

  which is surely increased by restricting the choice only to $w \in \{0,\delta\}$ so that we obtain $v(s,k\delta) \le v_{\#}(s,k\delta)$; which is further increased by restricting the choice only to $w \in \{0\}$,

$$v_{\#}(s,k\delta) \le \min\left(\mathcal{A}(\delta) + e^{-\lambda\delta} \cdot v(s,(k+1)\delta), \min_{s'\in\mathrm{succ}_e(s)} v(s',k\delta)\right).$$

  By the same arguments as for $s \in S_M$ we obtain

$$\le \min\left(\mathcal{B} + e^{-\lambda\delta} \cdot v_\delta(s,k+1), \min_{s'\in\mathrm{succ}_e(s)} v_\delta(s',k)\right) + \varepsilon \cdot \frac{T-k\delta}{T}. \qquad ◀$$

Theorem 8. For every approximation bound $\varepsilon > 0$ and discretization step $\delta \leq \varepsilon/(\lambda^2 T)$ where $\lambda = \max_{s \in S} \mathbf{E}(s)$, the value $(***)$ induced by δ satisfies

$$(***) \;\leq\; (**) \;\leq\; (***) + \varepsilon.$$

Proof. Straightforward, by putting $s := s_0$ and $k := 0$ into Lemmata 25 and 27. ◀

Theorem 9. Let $\varepsilon > 0$, Δ be a corresponding discrete game, and $\overline{\sigma}$ be induced by an optimal counting strategy in Δ, then

$$(*) \;\leq\; \inf_{\substack{E \in \mathrm{ENV} \\ \pi \in \mathfrak{S}(\mathcal{C}(E), \overline{\sigma})}} \mathcal{P}^{\pi}_{\mathcal{C}(E)} \left[\Diamond^{\leq T} G_E \right] + \varepsilon$$

Proof. By combining Theorem 8 and Lemma 25 we get that

$$\sup_{\sigma \in \Sigma} \inf_{\pi \in \Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G}} \left[\Diamond^{\leq T} G \right] \;\leq\; \inf_{\pi \in \Pi} \mathcal{P}^{\overline{\sigma},\pi}_{\mathcal{G}} \left[\Diamond^{\leq T} G \right] + \varepsilon$$

and further by Proposition 6

$$\sup_{\sigma \in \mathfrak{S}(\mathcal{C})} \inf_{\substack{E \in \mathrm{IMC} \\ \pi \in \mathfrak{S}(\mathcal{C}(E), \sigma)}} \mathcal{P}^{\pi[\sigma]}_{\mathcal{C}(E)} \left[\Diamond^{\leq T} G_E \right] \;\leq\; \inf_{\pi \in \Pi} \mathcal{P}^{\overline{\sigma},\pi}_{\mathcal{G}} \left[\Diamond^{\leq T} G \right] + \varepsilon.$$

Since the proof of Proposition 6 shows that for each strategy σ it holds

$$\inf_{\pi \in \Pi} \mathcal{P}^{\sigma,\pi}_{\mathcal{G}} \left[\Diamond^{\leq T} G \right] \;\leq\; \inf_{\substack{E \in \mathrm{IMC} \\ \pi \in \mathfrak{S}(\mathcal{C}(E), \sigma)}} \mathcal{P}^{\pi}_{\mathcal{C}(E)} \left[\Diamond^{\leq T} G_E \right],$$

we can conclude the proof by

$$\sup_{\sigma \in \mathfrak{S}(\mathcal{C})} \inf_{\substack{E \in \mathrm{IMC} \\ \pi \in \mathfrak{S}(\mathcal{C}(E), \sigma)}} \mathcal{P}^{\pi}_{\mathcal{C}(E)} \left[\Diamond^{\leq T} G_E \right] \;\leq\; \inf_{\substack{E \in \mathrm{IMC} \\ \pi \in \mathfrak{S}(\mathcal{C}(E), \overline{\sigma})}} \mathcal{P}^{\pi}_{\mathcal{C}(E)} \left[\Diamond^{\leq T} G_E \right] + \varepsilon.$$ ◀