Deciding the Value 1 Problem for Reachability in 1-Clock Decision Stochastic Timed Automata

Nathalie Bertrand, Thomas Brihaye, Blaise Genest

Deciding the value 1 problem for reachability in 1-clock decision stochastic timed automata ⋆

Nathalie Bertrand[1], Thomas Brihaye[2], and Blaise Genest[3]

[1] Inria, Team SUMO, UMR IRISA, Rennes, France
[2] Université de Mons, Mons, Belgium
[3] CNRS, Team SUMO, UMR IRISA, Rennes, France

Abstract. We consider reachability objectives on an extension of stochastic timed automata (STA) with nondeterminism. Decision stochastic timed automata (DSTA) are Markov decision processes based on timed automata where delays are chosen randomly and choices between enabled edges are nondeterministic. Given a reachability objective, the value 1 problem asks whether a target can be reached with probability arbitrary close to 1. Simple examples show that the value can be 1 and yet no strategy ensures reaching the target with probability 1. In this paper, we prove that, the value 1 problem is decidable for single clock DSTA by non-trivial reduction to a simple almost-sure reachability problem on a finite Markov decision process. The ε-optimal strategies are involved: the precise probability distributions, even if they do not change the winning nature of a state, impact the timings at which ε-optimal strategies must change their decisions, and more surprisingly these timings cannot be chosen uniformly over the set of regions.

1 Introduction

Stochastic timed automata (STA) were originally defined in [2,3] as a probabilistic semantics for timed automata, with the motivation to rule out 'unrealistic' paths in timed automata, and therefore alleviate some drawbacks of the mathematical model such as infinite precision of the clocks and instantaneous events. Of course, STA also form a new stochastic timed model, interesting on its own. Informally, the semantics of a stochastic timed automaton consists of an infinite-state infinitely-branching Markov chain whose underlying graph is the timed transition system associated with a timed automaton. The transitions between states are governed by the following: first, a delay is sampled randomly among possible delays, and second, an enabled transition is chosen randomly among enabled ones.

[⋆] The research leading to these results has received funding from the ARC project (number AUWB-2010-10/15-UMONS-3), a grant "Mission Scientifique" from the F.R.S.-FNRS, the FRFC project (number 2.4545.11), the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement number 601148 (CASSTING), the project ANR-13-BS02-0011-01 STOCH-MC, and the *professeur invité ISTIC-Rennes 1* programme.

Several models combining dense-time, continuous probabilities, and nondeterminism have been studied [7,8,11] and most result focus on qualitative questions, such as deciding the existence of a strategy ensuring a reachability objective with probability 1 (see the related work section).

A model that extends stochastic timed automata with nondeterminism was defined in [5]: the delays are random, but the choice between enabled transitions is nondeterministic. For this model, optimal strategies always exist for the *time-bounded* reachability problem. Yet, a simple example also shows that optimal strategies do not always exist for the reachability problem: there might be strategies to ensure a probability arbitrary close to 1 to reach a target location, and no strategy achieving probability 1.

More generally, the value 1 problem asks whether for every $\varepsilon > 0$ there exists a strategy ensuring a given objective with probability at least $1 - \varepsilon$. It can be defined in various game-like contexts, ranging from probabilistic finite automata (PFA) to concurrent games. In most models where the agent has full information, the value 1 problem coincides with the almost-sure problem, that is, whether there exists a strategy to ensure a given objective with probability 1. For partial observation models however, the value 1 problem and the almost-sure problem often differ: for concurrent games, both are decidable [12,9], whereas the value 1 problem is undecidable for PFA [14], and decidable only for subclasses [13,10].

In this paper, we consider a probabilistic and nondeterministic version of stochastic timed automata, called decision stochastic timed automata (DSTA), in which delays are random but edges are selected by the player. Contrary to most existing frameworks on stochastic and timed models, we do not assume the distributions over delays to be exponential. We consider (time-unbounded) reachability objectives on DSTA with a single clock. The restriction to 1-clock DSTA derives from the fact that even for purely stochastic models without decisions (*i.e.* STA), the decidability of the almost-sure reachability problem is open, for models with at least two clocks. Using the classical region abstraction we show that the existence of an almost-surely winning strategy is decidable for reachability objectives on 1-clock DSTA. Interestingly, in our context, the value 1 problem does not coincide with the almost-sure problem, although the agent has full information. The main contribution of the paper is then to prove that the value 1 problem is decidable too. To do so, we build an *ad hoc* abstraction based on a refinement of regions, and reduce to an almost-sure reachability question in the derived finite-state Markov decision process (MDP). The correctness proof is complex, and ε-optimal strategies are involved: first they are not uniform within a region as actions they dictate depend on the comparison of the precise clock value with some cutpoint. Second, and more surprisingly, these cutpoints cannot be chosen uniformly over the set of regions.

Related work

In stochastic timed games [7], locations are partitioned into locations owned by three players, a reachability player (who has a time-bounded reachability objective), a safety player (who has the opposite objective), and an environment player (who makes random moves). In a location of the reachability or safety player, the respective player decides both the sojourn time and the edge to fire, whereas in the environment's locations, the delay as well as the edge are chosen randomly. For this model, it was shown that, assuming there is a single player and the underlying timed automaton has only one clock, the existence of a strategy for a reachability goal almost-surely (resp. with positive probability) is PTIME-complete (resp. NLOGSPACE-complete). For two-player games, quantitative questions are undecidable. Simple examples show that even for one player and 1-clock timed automata, the value 1 and probability 1 problems differ. This is due to strict inequalities in guards, that prevent the player to choose an optimal delay. We believe that our proof techniques can be adapted to solve the value 1 problem in 1-player stochastic timed games over 1-clock timed automata.

In stochastic real-time games [8], environment nodes (in which the behaviour is similar to continuous time Markov decision processes (CTMDPs)) and control nodes (where players choose a distribution over actions) induce a probability distribution on runs. The objective for player 0 is to maximise the probability that a run satisfies a specification given by a deterministic timed automaton (DTA). The main result states that if player 0 has an almost-sure winning strategy, then she also has a simple one which can be described by a DTA.

Markovian timed automata (MTA) consist in an extension of timed automata with exponentially distributed sojourn time. Optimal probabilities can be approximated for time-(un)bounded reachability properties in MTA [11].

2 Definitions and problem statement

2.1 Timed automata

Timed automata [1] were introduced in the early nineties. We recall the definition and semantics of one-clock timed automata. Given a clock x, a *guard* is a finite conjunction of expressions of the form $x \sim c$ where $c \in \mathbb{N}$ is an integer, and $\sim$ is one of the symbols $\{<, \leq, =, \geq, >\}$. We denote by $\mathcal{G}(x)$ the set of guards over clock x. Often, for $g \in \mathcal{G}(x)$ a guard and t a clock value, we will write $t \in g$ to express that t satisfies the constraints expressed in g.

Definition 1. *A* one-clock timed automaton *is a tuple $(L, \ell_0, E, \mathcal{I})$ such that: L is a finite set of locations, $\ell_0 \in L$ is the initial location, $E \subseteq L \times \mathcal{G}(x) \times 2^{\{x\}} \times L$ is a finite set of edges, and $\mathcal{I} : L \to \mathcal{G}(x)$ assigns an invariant to each location.*

In the following, we assume all timed automata to be *well-formed*: for every location $\ell \in L$, $\mathcal{I}(\ell) = \bigcup_{(\ell,g,a,r,\ell') \in E} g$, that is, the invariant in a location coincides with the union of the guards on its outgoing edges. This implies in particular that the union of guards outgoing a location is an interval.

The semantics of a one-clock timed automaton $(L, \ell_0, E, \mathcal{I})$ is a timed transition system $\mathcal{T} = (S, s_0, \delta)$ where $S = L \times \mathbb{R}_{\geq 0}$, $s_0 = (\ell_0, 0)$ and the transition function $\delta \subseteq S \times (\mathbb{R}_{\geq 0} \cup E) \times S$ is composed of

- Delay transitions: $\big((\ell, t), \tau, (\ell, t + \tau)\big) \in \delta$ whenever $[t, t + \tau] \subseteq \mathcal{I}(\ell)$

- Discrete transitions: $\big((\ell,t),e,(\ell',t')\big) \in \delta$ as soon as the edge $e=(\ell,g,r,\ell') \in E$ satisfies $t \in g$ and if $r=\{x\}$, $t'=0$ else $t'=t$.

When convenient, we will use the alternative notations $(\ell,t) \xrightarrow{\tau} (\ell,t+\tau)$ and $(\ell,t) \xrightarrow{e} (\ell',t')$. Edge e is said *enabled* in state $s=(\ell,t)$, whenever there exists $s' \in S$ such that $s \xrightarrow{e} s'$.

2.2 Decision stochastic timed automata

We now introduce the concept of *decision stochastic timed automaton* (DSTA). Roughly speaking, a decision stochastic timed automaton is a one-clock timed automaton equipped with probability distributions over delays. The semantics of DSTA is provided by an infinite-state MDP, in the spirit of [5].

In the following, given $X \subseteq \mathbb{R}_{\geq 0}$, we denote by $\mathsf{Dist}(X)$ the set of probability distributions on X.

Definition 2. *A* decision stochastic timed automaton *is a tuple* $\mathcal{A} = (L,\ell_0,E,\mathcal{I},\mu)$ *where* $(L,\ell_0,E,\mathcal{I})$ *is a one-clock timed automaton and* $\mu=(\mu_{\ell,t})$ *is a family of* distributions, *one for each state* $(\ell,t) \in L \times \mathbb{R}_{\geq 0}$, *and such that* $\mu_{\ell,t} \in \mathsf{Dist}(\mathcal{I}(\ell) \cap [t,+\infty[)$.

Intuitively, for every state $(\ell,t) \in S$, for every interval $I \subseteq \mathbb{R}_{\geq 0}$, $\mu_{\ell,t}(I)$ is the probability that from (ℓ,t) a delay d_0, such that $t+d_0 \in I$, is chosen by μ.

We make some reasonable assumptions on the distributions. For every location ℓ, the function must satisfy the following sanity conditions:

(c_1) for every $t \in \mathcal{I}(\ell)$, and any non-punctual interval $I \subseteq \mathcal{I}(\ell) \cap [t,+\infty[$, $\mu_{\ell,t}(I)>0$; also if $[t,+\infty[\cap\mathcal{I}(\ell) \neq \{t\}$, then for any $a \in \mathbb{R}_{\geq 0}$, $\mu_{\ell,t}(\{a\})=0$;
(c_2) for every $t<t' \in \mathcal{I}(\ell)$, and $I \subseteq [t',+\infty[$, $\mu_{\ell,t}(I) \leq \mu_{\ell,t'}(I)$;
(c_3) if $|\mathcal{I}(\ell)|=\infty$, then for every $t,t' \in \mathcal{I}(\ell)$, for every $a,b \in \mathbb{R}_{\geq 0}$, $\mu_{\ell,t}([t+a,t+b])=\mu_{\ell,t'}(t'+a,t'+b)$;
if $|\mathcal{I}(\ell)|<\infty$, and $m=\sup\{t \mid t \in \mathcal{I}(\ell)\}$, then for every $t,t' \in \mathcal{I}(\ell)$, for every $a,b \in \mathbb{R}_{\geq 0}$, $\mu_{\ell,t}(t+\frac{a}{m-t},t+\frac{b}{m-t})=\mu_{\ell,t'}(t'+\frac{a}{m-t'},t'+\frac{b}{m-t'})$.

Let us comment on these conditions. First, (c_1) states that the distributions are equivalent to the Lebesgue measure: they do not assign 0 measure to interval with non-empty interior, and do not assign positive probability to points. Then, (c_2) is a monotonicity condition: the higher the clock value, the more likely a fixed interval is to be sampled. Last, with (c_3) one assumes that distributions depend only on the location, not on the precise clock value. More precisely, in case the invariant is not bounded, the distributions should be equal in all states; and if the invariant is bounded, they should coincide up to normalisation. It is important to notice the classical exponential and uniform distributions satisfy these three hypotheses.

Notice that stochastic timed automata (STA) [2,3] and DSTA share the same syntax, and only differ in their semantics: STA are interpreted as purely stochastic system whereas DSTA are interpreted as stochastic and nondeterministic systems. Let $\mathcal{A}$ be a decision stochastic timed automaton. Its semantics is given in terms of an infinite state MDP (or equivalently a 1-1/2 player game), based on the timed transition system $\mathcal{T}$ of the underlying timed automaton. The set of states is composed of two copies $S_\Box$ and $S_\Diamond$ of S: stochastic states $S_\Diamond = \{\langle s\rangle \mid s \in S\}$ and player states $S_\Box = \{[s] \mid s \in S\}$. The transitions are of the form:

- *stochastic transition*: $\langle s\rangle \xrightarrow{\tau} [s']$ if $(s,\tau,s') \in \delta$;
- *player transition*: $[s] \xrightarrow{e} \langle s'\rangle$ if $(s,e,s') \in \delta$.

The result of each transition is thus deterministic. However stochastic transitions are not played in an arbitrary way, but follow the family of probability distributions $(\mu_{\ell,t})$. Precisely, for $I \subseteq \mathbb{R}_{\geq 0}$ an interval, the probability from $\langle \ell, t\rangle$ to reach a clock value in I is given by $\mathbb{P}(\langle \ell,t\rangle \xrightarrow{\tau} [\ell,t'] \wedge t' \in I) = \mu_{\ell,t}(I)$.

Decisions of the nondeterministic player are specified through the notion of strategy. A *history* is a finite path in the MDP, ending in a player state: $\langle s_0\rangle \xrightarrow{\tau_0} [s_0'] \xrightarrow{e_0} \langle s_1\rangle \xrightarrow{\tau_1} [s_1'] \cdots \langle s_n\rangle \xrightarrow{\tau_n} [s_n']$. The set of all histories is denoted Hist. A strategy dictates the decision in states of $S_\Box$, given the history so far. Formally, a *strategy* is a function $\sigma : \mathsf{Hist} \to E$ such that $\sigma(\langle s_0\rangle \xrightarrow{\tau_0} [s_0'] \xrightarrow{e_0} \langle s_1\rangle \xrightarrow{\tau_1} [s_1'] \cdots \langle s_n\rangle \xrightarrow{\tau_n} [s_n'])$ is enabled in s_n'.

As pointed out in [16] in the context of continuous-time Markov decision processes, not all strategies are meaningful. The same phenomenon appears for DSTA, and in the following we thus restrict to so-called *measurable strategies* that induce measurable sets of runs for reachability objectives.

For a fixed measurable strategy σ, and an initial state $s_0 \in S_\Diamond \cup S_\Box$, the decision stochastic timed automaton $\mathcal{A}$ gives rise to a stochastic process. For a measurable event $\mathcal{E}$, we write $\mathbb{P}^{\mathcal{A}}_{\sigma}\big(s_0 \models \mathcal{E}\big)$ for the probability of $\mathcal{E}$ starting from s_0 and under strategy σ. Given a target set $F \subseteq S_\Diamond \cup S_\Box$ in the DSTA, the event $\Diamond F$, denotes the set of paths that eventually visit F.

2.3 Problem definition

Let $\mathcal{A}$ be a decision stochastic timed automaton, $\mathsf{Goal} \subseteq L$ and $s \in S_\Box \cup S_\Diamond$. We define $F = \{\langle \ell, t\rangle \mid \ell \in \mathsf{Goal}\}$. The value of s, with respect to the objective Goal, is the supremum, over all strategies, of the probability from s to reach F.

Definition 3. *The* value *of state s is* $\mathsf{val}_{\mathcal{A}}(s) = \sup_\sigma \mathbb{P}^{\mathcal{A}}_{\sigma}\big(s \models \Diamond F\big)$.
The value 1 problem *asks, given a decision stochastic timed automaton $\mathcal{A}$, a target set $\mathsf{Goal} \subseteq L$ and an initial state $s \in S_\Box \cup S_\Diamond$, whether* $\mathsf{val}_{\mathcal{A}}(s) = 1$.
F is said limit-surely *reachable from s if* $\mathsf{val}_{\mathcal{A}}(s) = 1$.

Notice that this definition is different from the *almost-sure* reachability problem, which asks whether there exists a strategy σ such that $\mathbb{P}^{\mathcal{A}}_{\sigma}\big(s \models \Diamond F\big) = 1$. From a state with value 1, for every ε, there exists a strategy achieving probability $1-\varepsilon$ to reach F, yet it does not imply that some strategy realises the objective with probability 1. For finite-state Markov decision processes, and in many simple frameworks, value 1 and probability 1 coincide. However, this is untrue for DSTA, as shown in the example below.

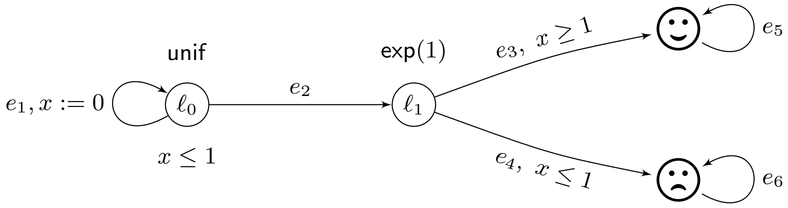

Fig. 1. A simple example of decision stochastic timed automaton.

Example 1. Figure 1 shows a basic example of a DSTA, where the distributions μ are uniform in ℓ_0 and exponential with rate 1 in location ℓ_1. The smiley location is limit-surely reachable from the initial location ℓ_0 with clock value 0. Indeed, the idea, in order to ensure a high probability $1-\varepsilon$ to reach the target is to loop on ℓ_0 until a player state $[\ell_0, 1-\tau]$ is reached (this happens almost surely) for a small τ, and then to move to location ℓ_1. Now, the probability from $\langle \ell_1, 1-\tau \rangle$ to reach the target converges to 1 as τ converges to 0. Yet, no strategy can ensure to reach the target with probability 1. This is thus a simple example where limit-sure reachability and almost-sure reachability differ. Such phenomena are not due to invariants, and already occur in DSTA where only exponential distributions are allowed. Indeed, one can adapt the above example and consider an exponential distribution in ℓ_0, while transferring the invariant $x \leq 1$ to the guard of e_2.

2.4 Limit corner-point MDP

As an extension of timed automata, DSTA have infinitely many states, because of continuous time. The usual technique to deal with this issue for timed automata, is to resort to the region abstraction [1], which we recall here. For one-clock timed automata, the number of regions is linear [15]: they all are intervals, either punctual $\{c\}$, open and bounded (c, d), or open and unbounded $(c, +\infty)$, for $c, d \in \mathbb{N}$.

We write R for the set of such regions, and r denotes a typical element of R.

Beyond these classical regions, we will use *pointed regions*, similar to the notion introduced for the corner-point abstraction by [6]: every bounded open region (c, d) is duplicated into $(\underline{c}, d)$ and $(c, \underline{d})$, with the intuitive meaning of being close to the left limit, or to the right limit of the interval. Other regions (unbounded or punctual) are kept as is. When a timed automaton is fixed, $\mathbf{R}$ denotes the set of pointed regions, with $\mathbf{r}$ a typical element, and it is partitioned into: $\mathbf{R}_{\mathsf{right}}$ (resp. $\mathbf{R}_{\mathsf{left}}$) for the set of pointed regions of the form $(c, \underline{d})$ (resp. $(\underline{c}, d)$), and $\mathbf{R}_{\mathsf{plain}}$ for punctual regions or the unbounded region. Pointed regions are equipped with a natural total order $<$; for example $\{0\} < (\underline{0}, 1) < (0, \underline{1})$. We say that pointed region $\mathbf{r}'$ is a *successor* of $\mathbf{r}$ if $\mathbf{r} < \mathbf{r}'$. The *immediate open* *successor* of $\mathbf{r}$ is the least region $\mathbf{r}'$, for the order $<$, that is different from $\mathbf{r}$ and open. The set of all successors of a region $\mathbf{r}$ is denoted $\overrightarrow{\mathbf{r}}$. A pointed region $\mathbf{r}'$ is said *negligible* with respect to location ℓ and region $\mathbf{r}$, if $\mathbf{r}' \in \overrightarrow{\mathbf{r}}$, $\mathbf{r}'$ is punctual and $\mathcal{I}(\ell) \cap \overrightarrow{\mathbf{r}}$ is not.

We now define the limit corner-point MDP associated with a DSTA.

Definition 4. *Given $\mathcal{A} = (L, \ell_0, E, \mathcal{I}, \mu)$ a DSTA, its* limit corner-point MDP *is $\mathcal{A}_{\mathsf{cp}} = (\mathcal{S}, \mathbf{s}_0, \mathsf{Act}, \Delta)$, where*

- $\mathcal{S} = \mathcal{S}_\Box \cup \mathcal{S}_\Diamond$ *is partitioned into player states and stochastic states:* $\mathcal{S}_\Box = \{[\ell, \mathbf{r}] \mid \ell \in L,\ \mathbf{r} \in \mathbf{R}\}$ *and* $\mathcal{S}_\Diamond = \{\langle \ell, \mathbf{r} \rangle \mid \ell \in L,\ \mathbf{r} \in \mathbf{R}\}$*;*
- $\mathbf{s}_0 = \langle \ell_0, \{0\} \rangle$*;*
- $\mathsf{Act} = E \cup E^{\mathsf{limit}}$*, where* E^{limit} *is a copy of* E*;*
- Δ *consists of the following transitions:*
  - $\langle \ell, \mathbf{r} \rangle \xrightarrow{\tau} [\ell, \mathbf{r}']$ *as soon as* $\mathbf{r}' \geq \mathbf{r}$ *and* $\mathbf{r}'$ *is not negligible w.r.t.* ℓ*, and the probabilities are uniform over all* τ*-successors;*
  - $[\ell, \mathbf{r}] \xrightarrow{e} \langle \ell', \{0\} \rangle$ *as soon as* $e = (\ell, g, \{x\}, \ell') \in E$*, and* $\mathbf{r} \models g$*;*
  - $[\ell, \mathbf{r}] \xrightarrow{e} \langle \ell', \mathbf{r} \rangle$ *as soon as* $e = (\ell, g, \emptyset, \ell') \in E$*, and* $\mathbf{r} \models g$*;*
  - $[\ell, \mathbf{r}] \xrightarrow{e^{\mathsf{limit}}} \langle \ell', \mathbf{r}' \rangle$ *as soon as* $\mathbf{r} \in \mathbf{R}_{\mathsf{right}}$*,* $e = (\ell, g, \emptyset, \ell') \in E$*,* $\mathbf{r} \models g$*,* $\mathbf{r}'$ *is the immediate open successor of* $\mathbf{r}$*, and* $\mathbf{r}' \models \mathcal{I}(\ell')$*.*

With the exception of limit-edges, the definition of the limit corner-point MDP is natural since it mimics the behaviour of the DSTA, at the region level and abstracting precise probabilities. Limit-edges are particular to the value 1 problem. Roughly speaking, they offer the player, from region $[\ell, r]$, the possibility to play as if the clock value was *arbitrarily close* to the right border of r, therefore as if it was in r' the immediate open successor of r. In particular, there cannot be two consecutive transitions starting with a limit edge and staying in the same pointed region $[\ell, \mathbf{r}] \xrightarrow{e^{\mathsf{limit}}} \xrightarrow{\tau} [\ell', \mathbf{r}]$.

Example 2. Let us illustrate Definition 4 on the example of Fig. 2 below. Its limit corner-point MDP is represented in Fig. 3. For readability reasons we only represented states with left-pointed region for (1, 2), since the behaviour is exactly the same from right-pointed regions.

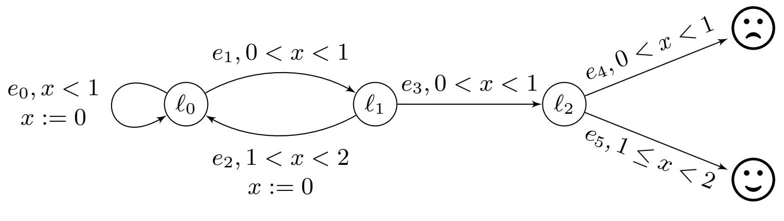

Fig. 2. The first illustrating example.

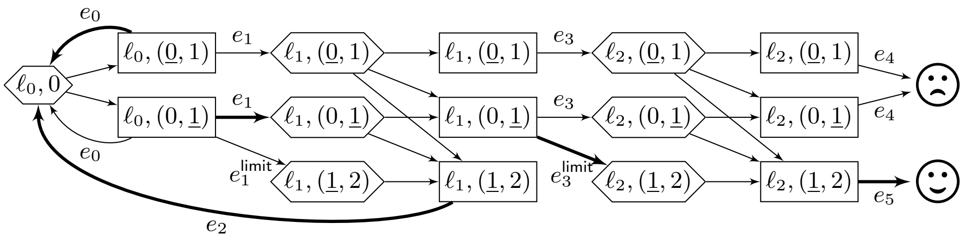
e_0
$\ell_0, (\underline{0}, 1)$
e_1
$\ell_1, (\underline{0}, 1)$
$\ell_1, (\underline{0}, 1)$
e_3
$\ell_2, (\underline{0}, 1)$
$\ell_2, (\underline{0}, 1)$
e_4
$\ell_0, 0$
$\ell_0, (0, \underline{1})$
e_1
$\ell_1, (0, \underline{1})$
$\ell_1, (0, \underline{1})$
e_3
$\ell_2, (0, \underline{1})$
$\ell_2, (0, \underline{1})$
e_4
e_0
e_1^{limit}
$\ell_1, (\underline{1}, 2)$
$\ell_1, (\underline{1}, 2)$
e_3^{limit}
$\ell_2, (\underline{1}, 2)$
$\ell_2, (\underline{1}, 2)$
e_5
e_2

Fig. 3. The limit corner-point MDP for the example from Fig. 2.

We let $\mathcal{F} = \{[\ell, \mathbf{r}] \mid \ell \in \mathsf{Goal}\}$ in the rest of the paper. $\mathcal{A}_{\mathsf{cp}}$ is a finite MDP, and one can define strategies in the usual way. In the following, for $\mathbf{s}$ a state of $\mathcal{A}_{\mathsf{cp}}$, and $\mathcal{F} \subseteq \mathcal{S}$ we write $\mathbb{P}^{\mathcal{A}_{\mathsf{cp}}}_{\max}(\mathbf{s} \models \Diamond\mathcal{F})$ for the maximum probability, over all strategies, to reach $\mathcal{F}$ from $\mathbf{s}$.

Last, we introduce some notations: First, for any region $r \in R$, we define $\bullet r \in \mathbf{R}_{\mathsf{left}} \cup \mathbf{R}_{\mathsf{plain}}$ (resp. $r\bullet \in \mathbf{R}_{\mathsf{right}} \cup \mathbf{R}_{\mathsf{plain}}$) with: $\bullet(c, d) = (\underline{c}, d)$ (resp. $(c, d)\bullet = (c, \underline{d})$), $\bullet\{c\} = \{c\} = \{c\}\bullet$ and $\bullet(c, +\infty) = (c, +\infty) = (c, +\infty)\bullet$. Now, given $t \in \mathbb{R}_{\geq 0}$, $\mathbf{r}_{\mathsf{left}}(t)$ (resp. $\mathbf{r}_{\mathsf{right}}(t)$) represents the left (resp. right) pointed region t belongs to: if $t \in r$, then $\mathbf{r}_{\mathsf{left}}(t) = \bullet r$ (resp. $\mathbf{r}_{\mathsf{right}}(t) = r\bullet$).

3 Main results

We now state the main results of our paper. We start with an expected result:

Proposition 1. *The almost-sure reachability problem is decidable in* PTIME *for DSTA.*

Proposition 1 is not a consequence of the decidability result in [7]. Although our model of DSTA can be encoded into the stochastic timed games of [7], the naive encoding requires an additional clock, in order to prevent players from letting time elapse. Since their decidability result applies only to stochastic timed games with a single clock, this simple reduction is of no help here. We believe their techniques can be adapted though. An alternative, which we take here, is to use the region abstraction, in order to solve the almost sure reachability problem. Details are provided Section 5.

As value 1 and probability 1 do not coincide for DSTA, the following theorem is non trivial, and is the main contribution of this paper.

Theorem 1. *The value 1 problem is decidable in* PTIME *for DSTA.*

To obtain Theorem 1, we reduce the value 1 problem for DSTA to the almost-sure reachability problem in the limit corner-point abstraction.

Proposition 2. *Let $\mathcal{A}$ be a decision stochastic timed automaton, $\mathcal{A}_{\mathsf{cp}}$ its limit corner point abstraction, and $\ell \in L$ a location, and $t \in \mathbb{R}_{\geq 0}$ a clock value. Then*

$$\begin{aligned} \mathsf{val}_{\mathcal{A}}([\ell,t]) = 1 \quad &\iff \quad \mathbb{P}^{\mathcal{A}_{cp}}_{\max}([\ell,\mathbf{r}_{\mathit{left}}(t)]) \models \Diamond\mathcal{F}) = 1 \ \textit{ and} \\ \mathsf{val}_{\mathcal{A}}(\langle\ell,t\rangle) = 1 \quad &\iff \quad \mathbb{P}^{\mathcal{A}_{cp}}_{\max}(\langle\ell,\mathbf{r}_{\mathit{left}}(t)\rangle) \models \Diamond\mathcal{F}) = 1 \ . \end{aligned}$$

Example 3. Let us illustrate the result of Proposition 2 on the example of Fig. 2, whose limit corner-point MDP is represented on Fig. 3. Bold edges give a winning strategy in the MDP for the almost-sure reachability of the smiley state. According to Proposition 2, the set of states with value 1 for the target ☺ in the stochastic timed automaton, is thus $(\ell_0,[0,1)) \cup (\ell_1,[1,2)) \cup (\ell_2,[1,2))$. (Here, we use brackets as a short-cut, rather than square brackets or angle brackets, not to distinguish stochastic and player states.) Intuitively, an ε-optimal strategy from $\langle\ell_0,0\rangle$ to reach ☺ is the following: stay in ℓ_0 until a large clock value is sampled, then move to ℓ_1; if then the sampled clock value is above 1, move back to ℓ_0 and iterate the same process, otherwise, proceed to ℓ_2; finally, reach ☺ or ☹ from ℓ_2 depending on the last sampled clock value.

Theorem 1 is a consequence of Proposition 2. To obtain a polynomial-time algorithm, one exploits that for 1-clock decision stochastic timed automaton, the number of regions, and thus the number of states in the limit corner-point MDP is linear [15], and almost-sure reachability properties can be checked in polynomial-time for finite MDP.

In order to show Proposition 2 we proceed in two steps: first, we prove it for player states $[\ell,t] \in S_\Box$. This suffices to prove it as well for stochastic states $\langle\ell,t\rangle \in S_\Diamond$ thanks to the structure of the limit corner-point MDP. Given $\mathcal{A}_{\mathsf{cp}}$, we write $\mathcal{W} \subseteq \mathcal{S}$ for the set of states from which there exists an almost-sure winning strategy for the reachability objective. We now detail what having left and/or right pointed region winning in the limit corner-point MDP abstraction implies:

Proposition 3. *Let $r \in R$ be a region and $\ell \in L$ a location.*

- *If $[\ell,\bullet r] \in \mathcal{W}$, then $[\ell,r\bullet] \in \mathcal{W}$.*
- *If $[\ell,\bullet r] \in \mathcal{W}$, then for every $t \in r$, $\mathsf{val}_{\mathcal{A}}([\ell,t]) = 1$;*
- *Else, if $[\ell,r\bullet] \in \mathcal{W}$, then for every $t \in r$, $\mathsf{val}_{\mathcal{A}}([\ell,t]) < 1$;*
- *Else, there exists $\varepsilon > 0$ such that for every $t \in r$, $\mathsf{val}_{\mathcal{A}}([\ell,t]) \leq 1-\varepsilon$.*

The first item is a simple observation: for every location ℓ and region $r \in R$, any winning strategy from $[\ell,\bullet r]$ can be mimicked from $[\ell,r\bullet]$, and is also winning from there. Section 4 is devoted to the rest of the proof of Proposition 3.

Proposition 3 suffices to prove Proposition 2, in the case of player states. Indeed, the second item (whose proof is in Section 4.2) shows the right-to-left implication of Proposition 2, and the third and fourth items show the other implication by contraposition (the proof is in Section 4.1). Remark that we can be more specific in the third case and state: when $[\ell,\bullet r] \notin \mathcal{W}$ and $[\ell,r\bullet] \in \mathcal{W}$, then $\sup_{t\in r} \mathsf{val}_{\mathcal{A}}([\ell,t]) = 1$. This explains the difference between the third and fourth items.

Before moving to the proofs, we show an example where ε-optimal strategies need to be conceptually complex: the cutpoint inside regions where the strategy changes decision cannot be chosen independently of the location.

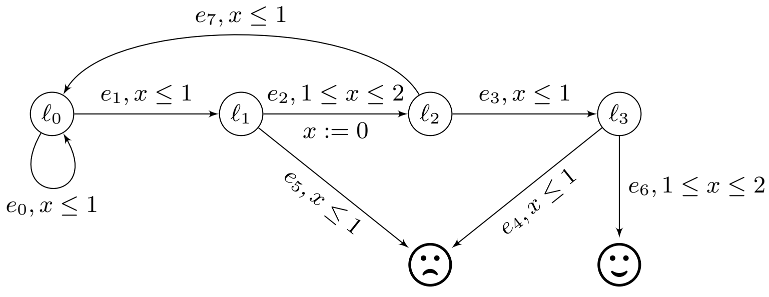

Fig. 4. An example where non-uniform strategies are needed.

Consider the example from Fig. 4, where the implicit probability distributions over delays are all uniformly distributed. Decisions can only be taken in locations ℓ_0 and ℓ_2, where transitions with overlapping guards are possible. Intuitively, from ℓ_0 to reach ☺, transition e_1 needs to be taken, with a risk that once ℓ_1 is reached, transition e_5 is triggered. Let t_0 be the cutpoint in ℓ_0 such that if $t_0 < t < 1$, the player decides to take e_1 from (ℓ_0, t). In the same way, let $t_2 \in (0, 1)$ be the cutpoint in ℓ_2 such that if $t_2 < t < 1$, the player decides to take e_3 from (ℓ_2, t). To reach a contradiction, we assume $t_0 = t_2$, and write τ for this value. From $[\ell_2, t]$ with $1 - \tau < t < 1$, a simple calculation shows that the probability to lose is $p^{T}_{\mathsf{lose}}(\ell_2, t) = (1-t)/(2-t)$. Also, from $[\ell_0, t]$, the losing probability is lower bounded by the probability to lose in two steps, directly from ℓ_1, hence $p^{T}_{\mathsf{lose}}(\ell_0, t) \geq (1-t)/(2-t)$. Moreover, $p^{T}_{\mathsf{win}}(\ell_0, t) \leq p^{T}_{\mathsf{win}}(\ell_1, t) \leq (1-t)p^{T}_{\mathsf{win}}(\ell_3, t) \leq (1-t)$. Hence, $p^{T}_{\mathsf{lose}}(\ell_0, t) > p^{T}_{\mathsf{win}}(\ell_0, t)/2$ for all $t > 1 - \tau$, that is $\mathbb{P}_T(\langle \ell_0, t\rangle \models \Diamond F) < 2/3$. This shows that F is not limit-surely reachable under simple strategies, defined by constant mappings. Yet, the ☺-state is limit-surely reachable from $[\ell_0, 0]$. However, to achieve this, t_0 needs to be set to a much lower value than t_2, *e.g.* $t_2 = \tau$ and $t_0 = \tau^2$.

4 Deciding the value 1 problem

The goal of this section is to provide a proof of Proposition 3 (and thus of Proposition 2 and of Theorem 1). For the sake of completeness, we recall the algorithm to compute the set of states of an MDP from which there exists a strategy to reach a target set $\mathcal{F}$ almost-surely (see *e.g.* [4]). We denote by $\mathcal{W}$ the winning states for this objective, and $\mathcal{W}_\Box$ the subset of winning player states. The algorithm that computes $\mathcal{W}$ computes at the same time for every player state $\mathbf{w} \in \mathcal{W}_\Box$, the largest set of safe actions $\mathsf{Safe}(\mathbf{w})$, *i.e.* the set of all actions that ensure staying in $\mathcal{W}$.

- Initially: $\mathcal{L} = \emptyset$, and for every $\mathbf{s} \in \mathcal{S}_\Box$, $\mathsf{Safe}(\mathbf{s}) = \{e \in E \cup E^{\mathsf{limit}} \mid \exists \mathbf{s} \xrightarrow{e} \mathbf{s}'\}$.
- Perform steps 1 and 2 until convergence.
  - Step 1: Move to $\mathcal{L}$ every $[\ell, \mathbf{r}]$ from which there is no path to $\mathcal{F}$ via states in $\mathcal{S} \setminus \mathcal{L}$ only.
  - Step 2: Remove from $\mathsf{Safe}([\ell, \mathbf{r}])$ any e such that if $[\ell, \mathbf{r}] \xrightarrow{e} \langle \ell', \mathbf{r}' \rangle$, there exists $[\ell', \mathbf{r}''] \in \mathcal{L}$ with $\langle \ell', \mathbf{r}' \rangle \to [\ell', \mathbf{r}'']$.
    Move to $\mathcal{L}$ every $[\ell, \mathbf{r}]$ such that $\mathsf{Safe}([\ell, \mathbf{r}]) = \emptyset$.
- Return $(\mathcal{W} = \mathcal{S} \setminus \mathcal{L}, \mathsf{Safe})$.

The rest of the section is organised as follows. Subsection 4.1 establishes the third and fourth items of Proposition 3: states whose left region is not winning in the limit corner-point MDP, do not have value 1 in the decision stochastic timed automaton. Then, Subsection 4.2 shows its second item: states whose left region is winning in the limit corner-point MDP do have value 1 in the decision stochastic timed automaton.

4.1 Non limit-surely winning states

We first aim at showing the right-to-left implication in Proposition 2, by contraposition: $\mathbb{P}^{\mathcal{A}_{\mathsf{cp}}}_{\max}((\ell, \mathbf{r}_{\mathsf{left}}(t)) \models \Diamond \mathcal{F}) < 1$ implies $\mathsf{val}_{\mathcal{A}}([\ell, t]) < 1$. This corresponds to proving the third and fourth items of Proposition 3.

Lemma 1. *Let $\ell \in L$ be a location and $r \in R$ a region.*

- *If $[\ell, r\bullet] \in \mathcal{L}$, then there exists $\varepsilon_{\ell,r} > 0$ such that for every $t \in r$, $\mathsf{val}_{\mathcal{A}}([\ell, t]) \leq 1 - \varepsilon_{\ell,r}$;*
- *If $[\ell, \bullet r] \in \mathcal{L}$, then for every $t \in r$, there exists $\varepsilon_{\ell,t} > 0$ such that $\mathsf{val}_{\mathcal{A}}([\ell, t]) \leq 1 - \varepsilon_{\ell,t}$; Moreover, one can pick non-increasing values for the $\varepsilon_{\ell,t}$'s, that is, $\varepsilon_{\ell,t} \leq \varepsilon_{\ell,t'}$ as soon as $t \geq t'$.*

Proof (Sketch). The proof is by induction on the moment in the MDP algorithm at which $[\ell, r]$ has been moved to $\mathcal{L}$. We thus define $\emptyset = \mathcal{L}_0 \subseteq \mathcal{L}_1 \subseteq \cdots \mathcal{L}_n = \mathcal{L}$ to describe the evolution of $\mathcal{L}$ during time with $|\mathcal{L}_i| = i$. Notice that this decomposition is finer than steps (this is important for step 2 of the MDP algorithm).

We show one important subcase here. Assume $[\ell, (\underline{a}, b)] \in \mathcal{L}$ because of step 2. Every transition e in the DSTA are associated with a transition e in the MDP which leads from $[\ell, (\underline{a}, b)]$ to $\langle \ell', \mathbf{r}' \rangle$, and there exists $\langle \ell', \mathbf{r}' \rangle \xrightarrow{\tau} [\ell', \mathbf{r}'']$ with $[\ell', \mathbf{r}''] \in \mathcal{L}_i$. The hardest case is when $\mathbf{r}'' = \mathbf{r}' = \mathbf{r} = (\underline{a}, b)$. Other cases are actually easier to treat and lead to a uniform bound ε over t. Let $\nu_{\ell,t} = \mu_{\ell,t}(t, (t+b)/2)$ for every $t \in (a, b)$. Observe that $\nu_{\ell,t} > 0$ and is non increasing with $t \in (a, b)$, by assumption (c_3) on the measure functions μ. We then set $\varepsilon^{\ell,t}_e = \nu_{\ell,t} \cdot \varepsilon_{\ell',(t+b)/2} > 0$ for all $t \in (a, b)$. Note that $\varepsilon^{\ell,t}_e$ depends upon t. Last, we define $\varepsilon_{\ell,t} = \min_e \varepsilon^{\ell,t}_e$, the minimum over all transitions e outgoing from $[\ell, t]$. So defined, $\varepsilon_{\ell,t}$ is positive and non increasing because $\nu_{\ell,t}$ and $\varepsilon'_{\ell',(t+b)/2}$ are positive and non increasing. □

4.2 Limit-surely winning states

We now prove that $\mathbb{P}^{\mathcal{A}_{\mathsf{cp}}}_{\max}((\ell, \mathbf{r}_{\mathsf{left}}(t)) \models \Diamond\mathcal{F}) = 1$ implies $\mathsf{val}_{\mathcal{A}}([\ell, t]) = 1$. This amounts to show the second item of Proposition 3.

Covering forest and golden paths. Let $\mathcal{A}$ be a decision stochastic timed automaton, and $\mathcal{A}_{\mathsf{cp}}$ the associated limit corner-point MDP. From the almost sure winning set of states and actions $(\mathcal{W}, \mathsf{Safe})$ of $\mathcal{A}_{\mathsf{cp}}$, we extract a *covering forest* whose roots are elements of $\mathcal{F}$. Each edge of the forest from a player state $\mathbf{w} \in \mathcal{W}_{\Box}$ to its unique parent is a transition of $\mathcal{A}_{\mathsf{cp}}$, labelled with action $\mathtt{sel}(\mathbf{w}) \in \mathsf{Safe}(\mathbf{w})$. Globally, for every $\mathbf{w} \in \mathcal{W}_{\Box}$, the unique path to a root of the forest $\mathbf{w}'_n \in \mathcal{F}$ is a path in $\mathcal{A}_{\mathsf{cp}}$: there are $\mathbf{w}_1 \cdots \mathbf{w}_n \in \mathcal{W}_{\Box}$ and $\mathbf{w}'_1 \cdots \mathbf{w}'_n \in \mathcal{W}_{\Diamond}$ with

$$\mathbf{w} \xrightarrow{\mathtt{sel}(\mathbf{w})} \mathbf{w}' \xrightarrow{\tau} \mathbf{w}_1 \xrightarrow{\mathtt{sel}(\mathbf{w}_1)} \mathbf{w}'_1 \cdots \mathbf{w}_{n+1} \xrightarrow{\mathtt{sel}(\mathbf{w}_n)} \mathbf{w}'_n \ ,$$

and such a path is called a *golden path* in the following. Notice that many edges emanating from stochastic states do not appear in this forest. They may lead to states that are further from $\mathcal{F}$. The intuition is that this forest represents the ideal situation. Even if it is not guaranteed to take these ideal edges from stochastic states, there is a chance to follow the forest towards $\mathcal{F}$, and we will show it is sufficient.

Example 4. Fig. 5 represents the covering forest, here a tree, on our running example from Fig. 2. Remark here that e_0 is selected in $[\ell_0, (\underline{0}, 1)]$ and e_1 in $[\ell_0, (0, \underline{1})]$. The fact that they differ reflects that the decision for an optimal strategy should not be uniform within region $(0, 1)$ in location ℓ_0.

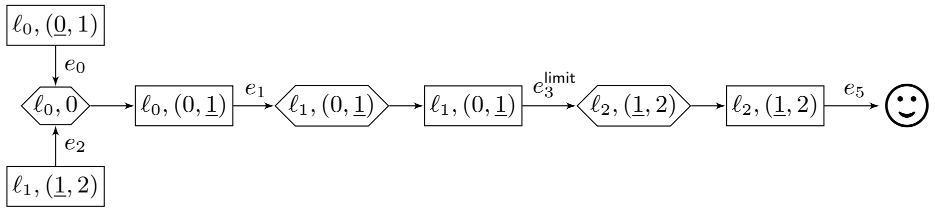
$\ell_0, (\underline{0}, 1)$
e_0
$\ell_0, 0$
$\ell_0, (0, \underline{1})$
e_1
$\ell_1, (0, \underline{1})$
$\ell_1, (0, \underline{1})$
e_3^{limit}
$\ell_2, (\underline{1}, 2)$
$\ell_2, (\underline{1}, 2)$
e_5
☺
e_2
$\ell_1, (\underline{1}, 2)$

Fig. 5. The covering forest on our running example.

Assume now that $\mathcal{F}$ is almost-surely reachable from $(\ell, \mathbf{r}_{\mathsf{left}}(t))$ in the limit corner-point $\mathcal{A}_{\mathsf{cp}}$, and let us exhibit a family of ε-optimal strategies, showing that F is limit-surely reachable from $\langle \ell, t \rangle$ in $\mathcal{A}$. More precisely, these ε-optimal strategies are positional but not region uniform: for every location ℓ and bounded open region $r = (c, d)$, there can be a cut-point $c + \tau \in (c, d)$ such that the decisions in the left part of r, $(c, c{+}\tau)$ and in its right part $(c{+}\tau, d)$ differ, but are uniform over each of these sub-intervals. These cut-points can be defined through a mapping $T : L \times R \to \mathbb{R}_{\geq 0}$. When T is fixed, we write $\mathsf{left}_T(\ell, r)$ for the subregion of (ℓ, r), to the left of its cutpoint $T(\ell, r)$. Similarly, $\mathsf{right}_T(\ell, r)$ denotes the states to the right of the cutpoint. For $t \in r$, we write $\mathbf{r}_T([\ell, t]) = [\ell, \mathbf{r}_{\mathsf{left}}(t)]$ if $t \in \mathsf{left}_T(\ell, r)$, and $\mathbf{r}_T([\ell, t]) = [\ell, \mathbf{r}_{\mathsf{right}}(t)]$ if $t \in \mathsf{right}_T(\ell, r)$. We will abuse the notation and call a path $s_1 \cdots s_n$ in the DSTA $\mathcal{A}$ a golden path if for $(\ell_i, \mathbf{r}_i)$ the golden path associated with $\mathbf{r}_T(s_1)$, we have $\mathbf{r}_T(s_{2i+1}) = [\ell_{2i+1}, \mathbf{r}_{2i+1}]$ for all i (if a limit edge is played, it is not possible to agree on the stochastic state).

Based on the covering forest of selected actions, we build a family of strategies $(\sigma_T)_{T:L\times R\to\mathbb{R}_{\geq 0}}$, parametrised by a cutpoint function $T : L \times R \to \mathbb{R}_{\geq 0}$: if $s \in \mathsf{left}_T(\ell, (c, d))$, then $\sigma_T(s) = \mathtt{sel}((\ell, (\underline{c}, d)))$, else $s \in \mathsf{right}_T(\ell, (c, d))$ and $\sigma_T(s) = \mathtt{sel}((\ell, (c, \underline{d})))$. In short, when T is fixed, $\sigma_T(s) = \mathtt{sel}(\mathbf{r}_T(s))$. Since limit edges do not exist in the DSTA, by $\sigma_T(s) = e^{\mathsf{limit}}$, we implicitly mean $\sigma_T(s) = e$, where $e \in E$ is the unique edge associated with $e^{\mathsf{limit}} \in E^{\mathsf{limit}}$.

We define the following subset of states in the DSTA $\mathcal{A}$: $S_{\mathsf{win}} = \{[\ell, t] \in S_\Box \mid [\ell, \mathbf{r}_{\mathsf{left}}(t)] \in \mathcal{W}\}$. We show that all states in S_{win} have value 1 for the reachability objective in $\mathcal{A}$. Further, $S_T^{\mathsf{right}} = \{[\ell, t] \in S_\Box \mid \mathbf{r}_T([\ell, t]) = [\ell, \mathbf{r}_{\mathsf{right}}(t)] \in \mathcal{W}\}$ is the set of states belonging to the right-part (as specified by the cut-point function T) of limit winning bounded open regions; Note that S_T^{right} and S_{win} are not necessarily disjoint. Our objective is to show that for every $\varepsilon > 0$, there exists a mapping T such that, from any state $s \in S_{\mathsf{win}}$, the probability to reach F under σ_T is at least $1 - \varepsilon$.

Lemma 2. *For every $s \in S_{\mathsf{win}}$, $\mathbb{P}^{\mathcal{A}}_{\sigma_T}(s \models \Diamond(F \cup S_T^{\mathsf{right}})) = 1$.*

When the mapping T is fixed, we let $[\![\mathcal{W}]\!]_T = \{s \mid \mathbf{r}_T(s) \in \mathcal{W}\}$, that is, states whose pointed-region (relatively to T) is winning in the limit corner point. Notice that $[\![\mathcal{W}]\!]_T = S_T^{\mathsf{right}} \cup S_{\mathsf{win}}$. We show that the probability to leave $[\![\mathcal{W}]\!]_T$ can be made arbitrarily small.

Lemma 3. *For every $\varepsilon > 0$, there exists a function $T : L \times R \to \mathbb{R}_{\geq 0}$ such that, writing $\bigcirc$ for the next-step operator, for every $s \in [\![\mathcal{W}]\!]_T$*

- *if $\sigma_T(s) \in E^{\mathsf{limit}}$ then $\mathbb{P}^s_{\sigma_T}(\bigcirc[\![\mathcal{W}]\!]_T) \geq 1 - \varepsilon$,*
- *else $\mathbb{P}^s_{\sigma_T}(\bigcirc[\![\mathcal{W}]\!]_T) = 1$.*

Taking a limit edge in the limit corner-point MDP is the only case where a losing state can be reached in the concrete DSTA. However, staying in the current region when the decision in the abstraction was a limit-edge may not necessarily lead to a losing state. Actually, limit-edges are not always the best choice: the only way to reach F might be to stay in the current region (and therefore avoid limit-edges) even if the probability to stay there is very small. This is illustrated on the example from Fig. 3 in which from $[\ell_0, (0, \underline{1})]$, in order to eventually reach $\mathcal{F}$, one should pick e_1 rather than e_1^{limit}.

We now explain that the probability to follow the covering forest towards $\mathcal{F}$, although small, will be arbitrarily bigger than the probability to follow the forest until reaching a losing state.

Given T a mapping assigning cut-points to each bounded open region, and $[\ell, t]$ a player state in $[\![\mathcal{W}]\!]_T$, we write $p^T_{\mathsf{win}}(\ell, t)$ for the probability, from $[\ell, t]$ and under σ_T, to execute a golden path (which therefore reach F). Also, $p^T_{\mathsf{lose}}(\ell, t)$ is the probability to execute a golden path until a losing state in $S \setminus [\![\mathcal{W}]\!]_T$ is reached. If from a stochastic state, a golden path is not executed, and yet $\mathcal{L}$ is not immediately reached either (this corresponds to behaviours not "counted" in p^T_{lose} and p^T_{win}), then a winning region will be reached, possibly further away from $\mathcal{F}$.

Lemma 4. *For every $\varepsilon > 0$, there exists a function $T : L \times R \to \mathbb{R}_{\geq 0}$ such that for every $[\ell, t] \in [\![\mathcal{W}]\!]_T$*

$$p^T_{\mathit{win}}(\ell, t) \cdot \varepsilon \geq p^T_{\mathit{lose}}(\ell, t) \enspace .$$

Given a tolerance ε, the proof of Lemma 4 details how to define a mapping, denoted $T(\varepsilon)$ to make the dependency explicit, under which the probability to reach $\mathcal{F}$ by progressing in the covering forest is arbitrarily bigger than the probability to reach a losing state. The definition of $T(\varepsilon)$ is non trivial and is done by induction on the distance to $\mathcal{F}$ in the covering forest. Recall that once a cut-point mapping T is fixed, the strategy σ_T is perfectly defined. It now remains to justify that, the strategies $(\sigma_{T(\varepsilon)})$ form a family of limit-sure strategies.

Lemma 5. *For every $s \in S_{\mathit{win}}$, $\mathbb{P}^{\mathcal{A}}_{\sigma_{T(\varepsilon)}}(s \models \Diamond F) \geq 1 - \varepsilon$.*

Proof. Let $\varepsilon > 0$, $T(\varepsilon) : L \times R \to (0, 1)$ the mapping as defined in Lemma 4, and $\sigma_{T(\varepsilon)}$ the corresponding strategy. To establish Lemma 5 we provide a lower bound on the probability, under $\sigma_{T(\varepsilon)}$ to reach F from winning states.

To do so, we consider the set X of runs under $\sigma_{T(\varepsilon)}$ that stay forever in $[\![\mathcal{W}]\!]_T \setminus F$. Such runs never reach the target, and also stay away from the losing states. We will show that $\mathbb{P}^{\mathcal{A}}_{\sigma_{T(\varepsilon)}}(s \models X) = 0$. To do so, we again partition X into three categories: X_1 gathers runs with infinitely many resets; X_2 consists of runs with finitely many resets and ending in the unbounded clock region (M, ∞); and X_3 is the set of runs with finitely many resets eventually staying in a bounded region (c, d).

Let us first consider X_1. Runs in X_1 necessarily visit some some state $(\ell_0, 0)$ infinitely often. Since, at each visit of $(\ell_0, 0)$, there is a strictly positive probability to execute a golden path and thus reach F, we reach a contradiction. Thus $\mathbb{P}^{\mathcal{A}}_{\sigma_{T(\varepsilon)}}(s \models X_1) = 0$. We now consider runs in X_2, that ultimately stay in the unbounded region. As explained earlier, almost surely F will be reached, a contradiction: $\mathbb{P}^{\mathcal{A}}_{\sigma_{T(\varepsilon)}}(s \models X_2) = 0$.

Last, for runs in X_3, that ultimately stay in a bounded region (c, d), the reasoning is exactly the same as for runs of X_2. Thus, $\mathbb{P}^{\mathcal{A}}_{\sigma_{T(\varepsilon)}}(s \models X_3) = 0$.

We now exploit Lemma 4 to conclude. Since almost all runs leave $[\![\mathcal{W}]\!]_T \setminus F$, it must be that either a losing states or F is reached, and it suffices to compare the probabilities in each case. Thanks to Lemma 4, for all states in $[\![\mathcal{W}]\!]_T$, it is much more likely to reach F than to reach a losing state. More precisely,

$$\frac{p^T_{\mathsf{win}}(\ell, t)}{p^T_{\mathsf{win}}(\ell, t) + p^T_{\mathsf{lose}}(\ell, t)} \geq \frac{p^T_{\mathsf{win}}(\ell, t)}{p^T_{\mathsf{win}}(\ell, t) + p^T_{\mathsf{win}}(\ell, t) \cdot \varepsilon} \geq \frac{1}{1 + \varepsilon} \geq 1 - \varepsilon \enspace .$$

As a consequence $\mathbb{P}^{\mathcal{A}}_{\sigma_{T(\varepsilon)}}(s \models \Diamond F) \geq (1-\varepsilon)$. □

This ends the proof of the left-to-right implication in Proposition 2.

5 Deciding the probability 1 problem

If one wants to solve the probability 1 problem, rather than the more difficult value 1 problem, it suffices to consider the region MDP $\mathcal{A}_R$. This MDP $\mathcal{A}_R$ is equivalent to the fragment of $\mathcal{A}_{\mathsf{cp}}$ restricted to left and plain regions, and hence without limit edges. As for the value 1 problem, the decidability of the probability 1 problem is given thanks to the following reduction:

Lemma 6. *Let $\mathcal{A}$ be a decision stochastic timed automaton, $\mathcal{A}_R$ its region MDP, $r \in R$ a region, and $\ell \in L$ a location and $t \in \mathbb{R}_{\geq 0}$ a clock value with $t \in r$. Then*

$$\exists \sigma,\ \mathbb{P}^{\mathcal{A}}_{\sigma}([\ell,t] \models \Diamond F) = 1 \quad \Longleftrightarrow \quad \mathbb{P}^{\mathcal{A}_R}_{\max}([\ell,r] \models \Diamond \mathcal{F}) = 1 \ \textit{and}$$
$$\exists \sigma,\ \mathbb{P}^{\mathcal{A}}_{\sigma}(\langle \ell,t \rangle \models \Diamond F) = 1 \quad \Longleftrightarrow \quad \mathbb{P}^{\mathcal{A}_R}_{\max}(\langle \ell,r \rangle \models \Diamond \mathcal{F}) = 1\ .$$

Proof. The proof is not different for player and stochastic states, so we treat them indistinctly, and use brackets in place of square or angle brackets.

Let (ℓ, t) with $t \in R$, be a state of $\mathcal{A}$ such that $\mathbb{P}^{\mathcal{A}_R}_{\max}((\ell,r) \models \Diamond \mathcal{F}) < 1$. One can easily adapt the inductive proof of Lemma 1, showing that if (ℓ, r) is losing in the MDP $\mathcal{A}_R$, then for every $t \in r$, there exists $\varepsilon_{\ell,t}$ with $\mathbb{P}^{\mathcal{A}}_{\sigma}((\ell,t) \models \Diamond F) < 1 - \varepsilon_{\ell,t}$ whatever the strategy σ.
For the other implication, it suffices to mimic faithfully in $\mathcal{A}$ the positional winning strategy $\sigma_{\mathcal{A}_R}$ from $\mathcal{A}_R$. For every (ℓ, r) a winning state in the region MDP $\mathcal{A}_R$, for every $t \in r$, we let $\sigma(\ell,t) = \sigma_{\mathcal{A}_R}(\ell, r)$. Now that the strategy is fixed, we recover the purely probabilistic framework of Stochastic Timed Automata, and can apply the results of [3] to conclude that $\mathbb{P}^{\mathcal{A}}_{\sigma}((\ell,t) \models \Diamond F) = 1$. Alternatively, partitioning runs into three categories, as we did for the proofs of Lemmas 2 and 5, allows one to conclude that σ is almost-surely winning in $\mathcal{A}$. □

Lemma 6, and precisely its right-to-left implication, does not hold in DSTA with at least two clocks. We emphasise here again that the decomposition of runs we use is only valid for 1-clock timed automata.

As an immediate consequence, we obtain the decidability in PTIME of the probability 1 problem for reachability objectives in DSTA (see Proposition 1).

6 Conclusion

This paper shows the decidability in PTIME of the probability 1 and value 1 problems for reachability objectives on an extension of 1-clock timed automata with random delays, and in which edges are chosen according to a strategy. It would be natural to allow for more general objectives (*e.g.* Büchi or parity). We could also investigate the extension of our framework to 2 players, taking decisions in turn or concurrently. Moving to more quantitative questions, such as computing the value would probably require a finer abstraction than the limit corner-point MDP. Last, the class of 1-clock DSTA can seem restrictive, and it is definitely a challenge to tackle already stochastic timed automata without decisions, for which the almost-sure model checking of reachability properties is still open.

References

1. R. Alur and D. L. Dill. A Theory of Timed Automata. *Theoretical Computer Science*, 126(2):183–235, 1994.
2. C. Baier, N. Bertrand, P. Bouyer, Th. Brihaye, and M. Größer. Probabilistic and Topological Semantics for Timed Automata. In *Proceedings of FSTTCS'07*, volume 4855 of *LNCS*, pages 179–191. Springer, 2007.
3. C. Baier, N. Bertrand, P. Bouyer, Th. Brihaye, and M. Größer. Almost-Sure Model Checking of Infinite Paths in One-Clock Timed Automata. In *Proceedings of LICS'08*, pages 217–226. IEEE Comp. Soc. Press, 2008.
4. C. Baier and J.-P. Katoen. *Principles of model checking.* MIT Press, 2008.
5. N. Bertrand and S. Schewe. Playing optimally on timed automata with random delays. In *Proceedings of FORMATS'12*, volume 7595 of *LNCS*, pages 43–58. Springer, 2012.
6. P. Bouyer, E. Brinksma, and K. G. Larsen. Optimal infinite scheduling for multi-priced timed automata. *Formal Methods in System Design*, 32(1):3–23, 2008.
7. P. Bouyer and V. Forejt. Reachability in Stochastic Timed Games. In *Proceedings of ICALP'09*, volume 5556 of *LNCS*, pages 103–114. Springer, 2009.
8. T. Brázdil, J. Krcál, J. Kretínský, A. Kucera, and V. Rehák. Stochastic Real-Time Games with Qualitative Timed Automata Objectives. In *Proceedings of CONCUR'10*, volume 6269 of *LNCS*, pages 207–221. Springer, 2010.
9. K. Chatterjee, L. de Alfaro, and T. A. Henzinger. Qualitative concurrent parity games. *ACM Transactions on Computation Logic*, 12(4):28, 2011.
10. K. Chatterjee and M. Tracol. Decidable problems for probabilistic automata on infinite words. In *Proceedings of LICS'12*, pages 185–194. IEEE, 2012.
11. T. Chen, T. Han, J.-P. Katoen, and A. Mereacre. Reachability Probabilities in Markovian Timed Automata. In *Proceedings of CDC-ECC'11*, pages 7075–7080. IEEE, 2011.
12. L. de Alfaro, T. A. Henzinger, and O. Kupferman. Concurrent reachability games. *Theoretical Computer Science*, 386(3):188–217, 2007.
13. N. Fijalkow, H. Gimbert, and Y. Oualhadj. Deciding the value 1 problem for probabilistic leaktight automata. In *Proceedings of LICS'12*, pages 295–304. IEEE, 2012.
14. H. Gimbert and Y. Oualhadj. Probabilistic automata on finite words: Decidable and undecidable problems. In *Proceedings of ICALP'10*, volume 6199 of *LNCS*, pages 527–538. Springer, 2010.
15. F. Laroussinie, N. Markey, and Ph. Schnoebelen. Model checking timed automata with one or two clocks. In *Proceedings of CONCUR'04*, volume 3170 of *LNCS*, pages 387–401. Springer, 2004.
16. N. Wolovick and S. Johr. A Characterization of Meaningful Schedulers for Continuous-Time Markov Decision Processes. In *Proceedings of FORMATS'06*, volume 4202 of *LNCS*, pages 352–367. Springer, 2006.