Expected Reachability-Price Games

Shibashis Guha[1]([✉]) and Ashutosh Trivedi[2]

[1] Université libre de Bruxelles, Brussels, Belgium
shibashis.guha@ulb.ac.be
[2] University of Colorado Boulder, Boulder, USA
ashutosh.trivedi@colorado.edu

Abstract. Probabilistic timed automata(PTA) model real-time systems with non-deterministic and stochastic behavior. They extend Alur-Dill timed automata by allowing probabilistic transitions and a price structure on the locations and transitions. Thus, a PTA can be considered as a Markov decision process (MDP) with uncountably many states and transitions. Expected reachability-price games are turn-based games where two players, player Min and player Max, move a token along the infinite configuration space of PTA. The objective of player Min is to minimize the expected price to reach a target location, while the goal of the Max player is the opposite. The undecidability of computing the value in the expected reachability-price games follows from the undecidability of the corresponding problem on timed automata. A key contribution of this work is a characterization of sufficient conditions under which an expected reachability-price game can be reduced to a stochastic game on a stochastic generalization of corner-point abstraction (a well-known finitary abstraction of timed automata). Exploiting this result, we show that expected reachability-price games for PTA with single clock and price-rates restricted to $\{0, 1\}$ are decidable.

1 Introduction

Two-player zero-sum games on finite automata were introduced by Ramadge and Wonham [27] as a mechanism for supervisory controller synthesis of discrete event systems. In this setting the two players—called Min and Max—represent the *controller* and the *environment*, and controller synthesis corresponds to finding a winning (or optimal) strategy of the controller for some given performance objective. Timed automata [2](TA) extend finite automata by providing a mechanism to model real-time behaviour, while priced timed automata are timed automata with (time-dependent) prices attached to the locations of the automata. If the game structure or objectives are dependent on time or price, e.g. when the objective corresponds to completing a given set of tasks within some deadline or within some cost, then games on timed automata are a well-established approach for controller synthesis, see e.g. [1,3,6,11,15].

We study an extension of the above approach to a setting that is quantitative in terms of both timed and probabilistic behavior. For this purpose, we consider an extension of probabilistic timed automata (PTA) [5,19,24]—a model for

real-time systems exhibiting nondeterministic and probabilistic behavior—with a partition of locations between two players. In our model, priced probabilistic timed game arena (PTGA), a token is placed on a configuration of a PTA and a play of the game corresponds to a player selecting a timed move (i.e. a time delay and an enabled action) and the token is moved according to the probabilistic transition function of the PTA. Players Min and Max choose their moves in order to minimize and maximize, respectively, the objective function. The *upper value* of a game is the minimum expected value that Min can ensure, while the *lower value* of a game is the maximum expected value that Max can ensure. A game is *determined* if the lower and upper values are equal, and in this case the *optimal value* of the game exists and equals the upper and lower values.

We are interested in *reachability-price* objectives, which express the *expected price* to reach a given target set. It is well known that two-player reachability-price games on timed arenas often lead to undecidability [1,11,12], even in non-probabilistic setting [9,14]. We study restrictions of PTGA in order to recover decidability. Our approach makes use of *boundary region abstraction* (BRA) for probabilistic timed automata [17]. Boundary region abstraction is similar to PTGA except that it restricts time delays to region boundaries. Boundary region abstraction has the property that starting from any state, only a finitely many other states can be reached. In particular, the reachable sub-graph of the boundary region abstraction from the initial state is same as the corner-point abstraction [8,10].

We characterize sufficient conditions under which the expected reachability-price games on PTA can be reduced to expected reachability-price games on the corresponding boundary region abstraction. This in particular characterizes conditions under which, for starting states with integral clock valuations, the expected reachability-price games can be reduced to corresponding corner-point abstractions. Using this result, we show the decidability of expected reachability-price problem for one-clock binary-priced PTGAs by reducing it to solving stochastic games on corresponding corner-point abstractions. To our best knowledge, this is the first decidability result for games on PTGA.

To understand the importance of our decidability result, let us review the challenges in solving games on timed automata on various related sub-classes. Brihaye et al. [12] showed the undecidability of deciding the existence of winning strategy for reachability-price games (on non-stochastic timed game arenas) with both positive and negative price-rates and two or more clocks. The next two examples highlight that permitting negative prices, or permitting non-binary prices may require non-positional or non-boundary strategies even in non-stochastic one-clock setting. The first example demonstrates that if negative prices are allowed, positional strategies may not be sufficient even for one-clock PTGA. The second example shows that even when positive prices are allowed, region boundary strategies may not be optimal even for one-clock PTGA.

Example 1 (Negative prices may require non-positional strategies). The Min player locations are represented using circles while the Max player locations are represented using squares. The number in each location is the cost that is incurred
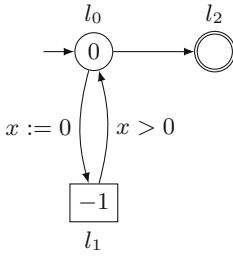
Fig. 1. Timed game arena with negative prices: no player has positional strategies.
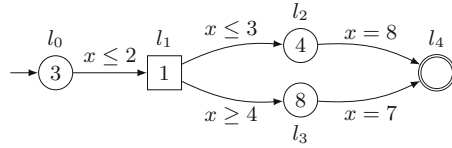
Fig. 2. Timed game arena with non-negative prices other than 0 and 1: boundary strategies may not be optimal; for Min player, the optimal transition from l_0 to l_1 takes place at a time $\frac{4}{3}$.

when the token stays in that location for one time unit. The example in Fig. 1 appears in [12] where location l_1 of player Max has a negative cost that is -1. It has been shown in [12] that player Max needs an infinite memory strategy by staying for a duration $\varepsilon/2^n$ for the n-th visit to location l_1 to ensure a payoff of $-\varepsilon$. Since this is true for every ε, it leads to a value 0 of the game. Besides, player Min also needs a finite memory in order to achieve an arbitrarily small payoff.

Example 2. Consider the timed game arena in Fig. 2. A similar example with two clocks appears in [11]. However, in this example there is only one clock x that is never reset. We show that for the timed automaton in this example, the reachability price problem *cannot* be reduced to the same problem in the corresponding corner-point abstractions/boundary region abstraction. The cost function at location l_1 to reach the goal location l_4 when the token reaches l_4 with a clock value of $x \leq 2$ is $\max(28 - x, 32 - 4x)$. This leads to the optimal reachability price strategy for player Min to reach the goal location l_4 to have an associated cost of $30\frac{2}{3}$ when the transition from location l_0 to location l_1 is taken at time $\frac{4}{3}$.

Related Work. Hoffman and Wong-Toi [18] were the first to define and solve the optimal controller synthesis problem for timed automata. For a detailed introduction to the topic of qualitative games on timed automata, see e.g. [4]. Asarin and Maler [3] initiated the study of quantitative games on timed automata by providing a symbolic algorithm to solve reachability-time objectives. The works of [13] and [22] show that the decision problem for such games over timed automata with at least two clocks is EXPTIME-complete. The tool UPPAAL Tiga [6] is capable of solving reachability and safety objectives for games on timed automata. Jurdziński and Trivedi [23] show the EXPTIME-completeness for average-time games on automata with two or more clocks.

A natural extension of games with reachability-time objectives are games on priced timed automata where the objective concerns the cumulated price of reaching a target. Both [1] and [11] present semi-algorithms for computing the

value of such games for linear prices, while the semi-algorithms always terminate under strongly Non-Zeno assumption on prices. In [12], it has been shown that the problem is decidable for a class of one-clock bivalued timed automata where the location prices can be any two from the set $\{-1, 0, 1\}$. In [14] the problem of checking the existence of optimal strategies is shown to be undecidable, with [9] showing undecidability holds even for three clocks and stopwatch prices.

Regarding one-player games on PTAs, [20] uses simple functions to devise a symbolic algorithm for computing minimum reachability-time. In [7] the problem of deciding whether a target can be reached within a given price and probability bound is shown to be undecidable for PTAs with three clocks and binary prices. Jurdziński et al. [21] show that the optimal expected cost problem is decidable for concavely-priced probabilistic timed automata.

Organization. The structure of the paper is the following. In the next section we recall required definitions to introduce turn-based priced probabilistic timed game arena and boundary region abstraction. We introduce expected reachability-price games in Sect. 3. In Sect. 4 we characterize conditions on the value of expected reachability-price game on the boundary region abstraction such that this value is equal to the value of the expected reachability-price game on the corresponding PTGA. Finally, in Sect. 5 we use these conditions to prove decidability of the expected reachability-price games on one-clock binary-priced PTGA.

2   Preliminaries

A *discrete probability distribution* over a countable set Q is a function $d : Q \to [0, 1]$ such that $\sum_{q \in Q} d(q) = 1$. For a possible uncountable set Q', we define $\mathcal{D}(Q')$ to be the set of functions $d : Q' \to [0, 1]$ such that the support set $supp(d) = \{q \in Q \mid d(q) > 0\}$ is countable and d is a distribution over $supp(d)$. We say that $d \in \mathcal{D}(Q)$ is a *point* if $d(q) = 1$ for some $q \in Q$.

2.1   Markov Decision Processes (MDPs)

We next introduce MDPs as modeling formalism for systems exhibiting nondeterministic and probabilistic behavior.

Definition 1 (Markov Decision Processe (MDP)). *An MDP is a tuple* $\mathsf{M} = (S, F, A, p, \pi)$ *where:*

- *S is the set of* states *including the set F of final states;*
- *A is the set of* actions*;*
- *$p : S \times A \to \mathcal{D}(S)$ is the* probabilistic transition function*;*
- *$\pi : S \times A \to \mathbb{R}_{\geqslant 0}$ is the* cost function*.*

We write $A(s)$ for the set of actions available at s, i.e., the set of actions a for which $p(s, a)$ is defined. For technical convenience we assume that $A(s)$ is nonempty for all $s \in S$. In an MDP $\mathcal{M}$, if the current state is s, then there is a non-deterministic choice between the actions in $A(s)$ and if action a is chosen the probability of reaching the state $s' \in S$ is denoted by $p(s' | s, a) \overset{\text{def}}{=} p(s, a)(s')$.

2.2   Probabilistic Timed Automata

We fix a constant $k \in \mathbb{N}$ and finite set of *clocks* $\mathcal{Y}$. Let $[\![k]\!]_{\mathbb{R}}$ denote the set of reals in $[0, k]$, while $[\![k]\!]_{\mathbb{N}}$ denotes the set of naturals $\{0, 1, \ldots, k\}$. A (k-bounded) *clock valuation* is a function $\nu : \mathcal{Y} \to [\![k]\!]_{\mathbb{R}}$ and we write V for the set of clock valuations. Although clocks are usually allowed to take arbitrary non-negative values, for technical convenience [8,17], we have restricted their values to be bounded by the constant k.

If $\nu \in V$ and $t \in \mathbb{R}_{\geqslant 0}$ then we write $\nu + t$ for the clock valuation defined by $(\nu + t)(c) = \nu(c) + t$, for all $c \in \mathcal{Y}$. For $C \subseteq \mathcal{Y}$, we write $\nu[C := 0]$ for the clock valuation where $\nu[C := 0](c) = 0$ if $c \in C$, and $\nu[C := 0](c) = \nu(c)$ otherwise. For $X \subseteq V$, we write $\overline{X}$ for the smallest closed (topological) set in V containing X. Let $X \subseteq V$ be a convex subset of clock valuations and let $F : X \to \mathbb{R}$ be a continuous function. We write $\overline{F}$ for the unique continuous function $F' : \overline{X} \to \mathbb{R}$, such that $F'(\nu) = F(\nu)$ for all $\nu \in X$.

The set of *clock constraints* over $\mathcal{Y}$ is the set of conjunctions of *simple constraints*, which are constraints of the form $c \bowtie i$ or $c - c' \bowtie i$, where $c, c' \in \mathcal{Y}, i \in [\![k]\!]_{\mathbb{N}}$, and $\bowtie \in \{<, >, =, \leqslant, \geqslant\}$. For every $\nu \in V$, let $\mathrm{CC}(\nu)$ be the set of simple constraints that hold in ν. A *clock region* is a maximal set $\zeta \subseteq V$, such that $\mathrm{CC}(\nu) = \mathrm{CC}(\nu')$ for all $\nu, \nu' \in \zeta$. Every clock region is an equivalence class of the indistinguishability-by-clock-constraints relation, and vice versa. Note that ν and ν' are in the same clock region if and only if the integer parts of the clocks and the partial orders of the clocks, determined by their fractional parts, are the same in ν and ν', and if the fractional part of a clock c be 0 in ν, then it should be 0 in ν', and if it is positive in ν, then so it should be in ν'. We write $[\nu]$ for the clock region of ν and, if $\zeta = [\nu]$, write $\zeta[C := 0]$ for the clock region $[\nu[C := 0]]$.

A *clock zone* is a convex set of clock valuations, that is a union of a set of clock regions. We write $\mathcal{Z}$ for the set of clock zones. For any clock zone W and clock valuation ν, we use the notation $\nu \in W$ to denote that $[\nu] \subseteq W$. A set of clock valuations is a clock zone if and only if it is definable by a clock constraint. Observe that, for every clock zone W, the set $\overline{W}$ is also a clock zone.

Definition 2 (Syntax). *A priced probabilistic timed automaton (PPTA) is a tuple* $\mathsf{T} = (L, L_F, \mathcal{Y}, \mathit{Inv}, \mathit{Act}, E, \delta, r)$ *where:*

- *L is the finite set of* locations *including the set L_F of* final locations*;*
- *$\mathcal{Y}$ is the finite set of* clocks*;*
- *$\mathit{Inv} : L \to \mathcal{Z}$ is the* invariant condition*;*
- *Act is the finite set of* actions*;*
- *$E : L \times \mathit{Act} \to \mathcal{Z}$ is the* action enabledness function*;*
- *$\delta : (L \times \mathit{Act}) \to \mathcal{D}(2^{\mathcal{Y}} \times L)$ is the* transition probability function*;*
- *$r : L \cup L \times \mathit{Act} \to \mathbb{R}_{\geqslant 0}$ is the* price information function*. A PPTA is binary-priced when $r(\ell) \in \{0, 1\}$ for all $\ell \in L$.*

A *probabilistic timed automaton* (PTA) is a PPTA in which the cost on every edge is 0, while the cost on the locations are all 1, i.e., $r((\ell, \ell')) = 0$ for all $(\ell, \ell') \in L \times L$ and $r(\ell) = 1$ for all $\ell \in L$. A *timed automaton* is a PTA with the property that $\delta(\ell, a)$ is a point distribution for all $\ell \in L$ and $a \in \mathit{Act}$.

A *configuration* of a PPTA T is a pair (ℓ, ν), where $\ell \in L$ is a location and $\nu \in V$ is a clock valuation over $\mathcal{Y}$ such that $\nu \in Inv(\ell)$. For any $t \in \mathbb{R}_{\geq 0}$, we let $(\ell, \nu) + t$ equal the configuration $(\ell, \nu + t)$. Informally, the behaviour of a PPTA is as follows: In configuration (ℓ, ν) time passes before an available action from Act is triggered, after which a discrete probabilistic transition occurs. Time passage is available only if the invariant condition $Inv(\ell)$ is satisfied while time elapses, and an action a can be chosen after time t elapses only if it is enabled after time elapse, i.e., if $\nu + t \in E(\ell, a)$. Both the time and the action chosen are nondeterministic. If an action a is chosen, then the probability of moving to a location ℓ' and resetting all of the clocks in $C \subseteq \mathcal{Y}$ to 0 is given by $\delta[\ell, a](C, \ell')$.

Formally, the semantics of a PPTA is given by an MDP which has both an uncountable number of states and an uncountable number of transitions.

Definition 3 (Semantics). *Let* $\mathsf{T} = (L, L_F, \mathcal{Y}, Inv, Act, E, \delta, r)$ *be a PPTA. The semantics of* T *is the MDP* $[\![\mathsf{T}]\!] = (S, F, A, p, \pi)$ *where*

- $S \subseteq L \times V$, *the set of states, is such that* $(\ell, \nu) \in S$ *if and only if* $\nu \in Inv(\ell)$;
- $F = S \cap (L_F \times V)$ *is the set of final states;*
- $A = \mathbb{R}_{\geq 0} \times Act$ *is the set of* timed actions;
- $p : S \times A \to \mathcal{D}(S)$ *is the probabilistic transition function such that for* $(\ell, \nu) \in S$ *and* $(t, a) \in A$, *we have* $p((\ell, \nu), (t, a)) = d$ *if and only if*
  - $\nu + t' \in Inv(\ell)$ *for all* $t' \in [0, t]$;
  - $\nu + t \in E(\ell, a)$;
  - $d((\ell', \nu')) = \sum_{C \subseteq \mathcal{Y} \wedge (\nu + t)[C := 0] = \nu'} \delta[\ell, a](C, \ell')$ *for all* $(\ell', \nu') \in S$.
- $\pi : S \times A \to \mathbb{R}$ *is the price function where* $\pi(s, (t, a)) = r(\ell) \cdot t + r(\ell, a)$ *for* $s = (\ell, \nu) \in S$ *and* $(t, a) \in A$.

For the sake of notational convenience, we often write $d(\ell, \nu)$ for $d((\ell, \nu))$.

2.3   Priced Probabilistic Timed Game Arena

Definition 4. *A priced probabilistic timed game arena (PTGA)* $\mathcal{T}$ *is a triplet* $(\mathsf{T}, L_{Min}, L_{Max})$ *where* $\mathsf{T} = (L, L_F, \mathcal{Y}, Inv, Act, E, \delta, r)$ *is a priced probabilistic timed automaton and* (L_{Min}, L_{Max}) *is a partition of* L.

The semantics of a PTGA $\mathcal{T}$ is the stochastic game arena $[\![\mathcal{T}]\!] = ([\![\mathsf{T}]\!], S_{\text{Min}}, S_{\text{Max}})$ where $[\![\mathsf{T}]\!] = (S, F, A, p, \pi)$ is the semantics of T, and $S_{\text{Min}} = S \cap (L_{\text{Min}} \times V)$ and $S_{\text{Max}} = S \setminus S_{\text{Min}}$. Intuitively S_{Min} is the set of states controlled by player Min, and S_{Max} is the set of states controlled by player Max.

In a turn-based game on $\mathcal{T}$, players Min and Max move a token along the states of the PPTA in the following manner. If the current state is s, then the player controlling the state chooses an action $(t, a) \in A(s)$ after which state $s' \in S$ is reached with probability $p(s'|s, a)$. In the next turn, the player controlling the state s' chooses an action in $A(s')$ and a probabilistic transition is made accordingly. We say that $(s, (t, a), s')$ is a transition in $\mathcal{T}$ if $p(s'|s, (t, a)) > 0$ and a *play* of $\mathcal{T}$ is a sequence $\langle s_0, (t_1, a_1), s_1, \ldots \rangle \in S \times (A \times S)^*$ such that $(s_i, (t_{i+1}, a_{i+1}), s_{i+1})$ is a transition for all $i \geq 0$. We write Runs (FRuns) for the

set of infinite (finite) plays and Runs_s (FRuns_s) for the sets of infinite (finite) plays starting from state s. For a finite play η let $\text{Last}(\eta)$ denote the last state of the play. Let X_i and Y_i denote the random variables corresponding to i^{th} state and action of a play.

A *strategy* of player Min in $\mathcal{T}$ is a partial function $\mu : \text{FRuns} \to \mathcal{D}(A)$, defined for $\eta \in \text{FRuns}$ if and only if $\text{Last}(\eta) \in S_{\text{Min}}$, such that $supp(\mu(\eta)) \subseteq A(\text{Last}(\eta))$. Strategies of player Max are defined analogously. We write Σ_{Min} and Σ_{Max} for the set of strategies of players Min and Max, respectively. Let $\text{Runs}_s^{\mu,\chi}$ denote the subset of Runs_s which corresponds to the set of plays in which the players play according to $\mu \in \Sigma_{\text{Min}}$ and $\chi \in \Sigma_{\text{Max}}$, respectively. A strategy σ is *pure* if $\sigma(\eta)$ is a Dirac distribution for all $\eta \in \text{FRuns}$ for which it is defined, while it is *positional* if $\text{Last}(\eta){=}\text{Last}(\eta')$ implies $\sigma(\eta){=}\sigma(\eta')$ for all $\eta, \eta' \in \text{FRuns}$. We write Π_{Min} and Π_{Max} for the set of positional strategies of player Min and player Max.

To analyse a stochastic game on $\mathcal{T}$ under a strategy pair (μ, χ), for every state s of $\mathcal{T}$, we define a probability space $(\text{Runs}_s^{\mu,\chi}, \mathcal{F}_{\text{Runs}_s^{\mu,\chi}}, Prob_s^{\mu,\chi})$ over the set of infinite plays under strategies μ and χ with s as the initial state. Given a *real-valued random variable* $f : \text{Runs} \to \mathbb{R}$, we can then define the expectation of this variable $\mathbb{E}_s^{\mu,\chi}\{f\}$ with respect to strategy pair (μ, χ) when starting in s. For technical reasons we make the following assumption [26] (a similar assumption is required for finite MDP [16]):

Assumption 1 (Stopping Game assumption): *For every strategy pair* $(\mu, \chi) \in \Sigma_{Min} \times \Sigma_{Max}$, *and state* $s \in S$ *we have that* $\lim_{i \to \infty} Prob_s^{\mu,\chi}(X_i \in F) = 1$.

Given any PPTA without Assumption 1, a PPTA can be constructed for which Assumption 1 holds using standard attractor computation in a two-player game. This can be done by constructing the region graph of the PPTA.

2.4   Boundary Region Abstraction

A *region* is a pair (ℓ, ζ), where ℓ is a location and ζ is a clock region such that $\zeta \subseteq Inv(\ell)$. For every $s{=}(\ell, \nu)$, we write $[s]$ for the region $(\ell, [\nu])$ and, we denote by $\mathcal{R}$ the set of regions. A set $Z \subseteq L \times V$ is a *zone* if, for every $\ell \in L$, there is a clock zone W_ℓ (possibly empty), such that $Z = \{(\ell, \nu) \mid \ell \in L \wedge \nu \in W_\ell\}$. For a region $R{=}(\ell, \zeta) \in \mathcal{R}$, we write $\overline{R}$ for the zone $\{(\ell, \nu) \mid \nu \in \overline{\zeta}\}$, recall $\overline{\zeta}$ is the smallest closed set in V containing ζ.

For $R, R' \in \mathcal{R}$, we say that R' is in the future of R, or that R is in the past of R', if there is $s \in R$, $s' \in R'$ and $t \in \mathbb{R}_{\geqslant 0}$ such that $s' = s + t$; we then write $R \to_* R'$. We say that R' is the *time successor* of R if $R \to_* R'$, $R{\neq}R'$, and $R \to_* R'' \to_* R'$ implies $R''{=}R$ or $R''{=}R'$ and denote it by both $R \to_{+1} R'$ and $R' \leftarrow_{+1} R$. We say that a region $R \in \mathcal{R}$ is *thin* if $[s] \neq [s+\varepsilon]$ for every $s \in R$ and $\varepsilon{>}0$; other regions are called *thick*. We write $\mathcal{R}_{\text{Thin}}$ and $\mathcal{R}_{\text{Thick}}$ for the sets of thin and thick regions, respectively. Note that if $R \in \mathcal{R}_{\text{Thick}}$ then, for every $s \in R$, there is an $\varepsilon > 0$, such that $[s] = [s+\varepsilon]$. Observe that the time successor of a thin region is thick, and vice versa.

We say $(\ell, \nu) \in L \times V$ is in the *closure of the region* (ℓ, ζ), and we write $(\ell, \nu) \in \overline{(\ell, \zeta)}$, if $\nu \in \overline{\zeta}$. For any $\nu \in V$, $b \in [\![k]\!]_{\mathbb{N}}$ and $c \in \mathcal{Y}$ such that $\nu(c){\leqslant}b$,

we let $time(\nu, (b, c)) \stackrel{\text{def}}{=} b - \nu(c)$. Intuitively, $time(\nu, (b, c))$ returns the amount of time that must elapse from ν before clock c reaches the integer value b. Note that, for any $(\ell, \nu) \in L \times V$ and $a \in Act$, if $t = time(\nu, (b, c))$ is defined, then $(\ell, [\nu + t]) \in \mathcal{R}_{\text{Thin}}$ and $supp(p(\cdot \mid (\ell, \nu), (t, a))) \subseteq \mathcal{R}_{\text{Thin}}$. Observe that, for every $R' \in \mathcal{R}_{\text{Thin}}$, there is a number $b \in [\![k]\!]_{\mathbb{N}}$ and a clock $c \in \mathcal{Y}$, such that, for every $R \in \mathcal{R}$ in the past of R', we have $s \in R$ implies $s + (b - s(c)) \in R'$; and we write $R \to_{b,c} R'$.

Intuition. The boundary region abstraction is motivated by the following. Consider an $a \in Act$, $s = (\ell, \nu)$ and $R = (\ell, \zeta) \to_* R' = (\ell, \zeta')$ such that $s \in R$ and $R' \in E(\ell, a)$.

– If $R' \in \mathcal{R}_{\text{Thick}}$, then there are infinitely many $t \in \mathbb{R}_{\geqslant 0}$ such that $s + t \in R'$. However, amongst all such t's, for one of the boundaries of ζ', the closer $\nu + t$ is to this boundary, the 'better' the timed action (t, a) becomes for a player's objective. However, since R' is a thick region, the set $\{t \in \mathbb{R}_{\geqslant 0} \mid s + t \in R'\}$ is an open interval, and hence does not contain its boundary values. Observe that the infimum equals $b_- - \nu(c_-)$ where $R \to_{b_-, c_-} R_- \to_{+1} R'$ and the supremum equals $b_+ - \nu(c_+)$ where $R \to_{b_+, c_+} R_+ \leftarrow_{+1} R'$. In our abstraction we include these 'best' timed actions through the actions $((b_-, c_-, a), R')$ and $((b_+, c_+, a), R')$. Stated otherwise, b_- and b_+ respectively denote the lower and the upper boundary of the thick region R' and the clocks c_- and c_+ correspond to the best timed actions.
– If $R' \in \mathcal{R}_{\text{Thin}}$, then there exists a unique $t \in \mathbb{R}_{\geqslant 0}$ such that $(\ell, \nu + t) \in R'$. Moreover since R' is a thin region, there exists a clock $c \in C$ and a number $b \in \mathbb{N}$ such that $R \to_{b,c} R'$ and $t = b - \nu(c)$. In the boundary region abstraction we summarise this 'best' timed action from region R to region R' through the action $((b, c, a), R')$.

Based on this intuition the abstraction is formalized below.

Definition 5. *Let* $\mathsf{T} = (L, L_F, \mathcal{Y}, Inv, Act, E, \delta)$ *be a PPTA. The* boundary region abstraction *of* T *is defined as the MDP* $\widehat{\mathsf{T}} = (\widehat{S}, \widehat{F}, \widehat{A}, \widehat{p}, \widehat{\pi})$ *where*

– $\widehat{S} = \{((\ell, \nu), (\ell, \zeta)) \mid (\ell, \zeta) \in \mathcal{R} \wedge \nu \in \overline{\zeta}\}$ *and* $\widehat{F} = \{((\ell, \nu), (\ell, \zeta)) \in \widehat{S} \mid \ell \in L_F\}$;
– $\widehat{A} \subseteq ([\![k]\!]_{\mathbb{N}} \times \mathcal{Y} \times Act) \times \mathcal{R}$ *is the finite set of boundary actions and for* $R \in \mathcal{R}$ *we let* $\widehat{A}(R) = \{\alpha \in \widehat{A}((\ell, \nu), R) \mid ((\ell, \nu), R) \in \widehat{S}\}$;
– *for* $((\ell, \nu), (\ell, \zeta)) \in \widehat{S}$ *and boundary action* $((b, c, a), (\ell, \zeta_a)) \in \widehat{A}$ *we have* $\widehat{p}((\ell, \nu), (\ell, \zeta), ((b, c, a), (\ell, \zeta_a))) = d$ *if and only if*

$$d((\ell', \nu'), (\ell', \zeta')) = \sum_{\substack{C \subseteq \mathcal{Y} \wedge \nu_a[C := 0] = \nu' \\ \wedge \zeta_a[C := 0] = \zeta'}} \delta[\ell, a](C, \ell')$$

for all $((\ell', \nu'), (\ell', \zeta')) \in \widehat{S}$ *where* $\nu_a = \nu + time(\nu, (b, c))$ *and one of the following conditions holds:*
  • $(\ell, \zeta) \to_{b,c} (\ell, \zeta_a)$ *and* $\zeta_a \in E(\ell, a)$;

- $(\ell, \zeta) \to_{b,c} (\ell, \zeta_-) \to_{+1} (\ell, \zeta_a)$ *for some* (ℓ, ζ_-) *and* $\zeta_a \in E(\ell, a)$; *and*
- $(\ell, \zeta) \to_{b,c} (\ell, \zeta_+) \leftarrow_{+1} (\ell, \zeta_a)$ *for some* (ℓ, ζ_+) *and* $\zeta_a \in E(\ell, a)$.

- $\widehat{\pi} : \widehat{S} \times \widehat{A} \to \mathbb{R}$ *is such that for* $((\ell, \nu), (\ell, \zeta)) \in \widehat{S}$ *and* $((b, c, a), R) \in \widehat{A}(((\ell, \nu), (\ell, \zeta)))$ *we have*

$$\widehat{\pi}(((\ell, \nu), (\ell, \zeta)), ((b, c, a), R)) = r(\ell, a) + r(\ell) \cdot (b - \nu(c)).$$

Although the boundary region abstraction is uncountably infinite, for a fixed initial state we can restrict attention to a finite state subgraph, thanks to the following observation [17].

Lemma 1. *For every state of a boundary region abstraction, its reachable sub-graph is finite. Moreover, the reachable sub-graph from the initial valuation corresponds to the standard corner-point abstraction [8].*

3 Expected Reachability-Price Games

In an expected reachability-price game (ERPG) on $\mathcal{T} = (\mathsf{T}, L_{\mathrm{Min}}, L_{\mathrm{Max}})$ player Min attempts to reach the final states with cost as low as possible, while the objective of player Max is the opposite. In fact, the cost is infinity if player Max has a strategy such that a configuration in $F \times V$, that is one corresponding to a goal location is never reached. More precisely, Min is interested in minimising her losses, while player Max is interested in maximising his winnings where, if player Min uses the strategy $\mu \in \Sigma_{\mathrm{Min}}$ and player Max uses the strategy $\chi \in \Sigma_{\mathrm{Max}}$, player Min loses the following to player Max:

$$\mathrm{EReach}(s, \mu, \chi) \overset{\mathrm{def}}{=} \mathbb{E}_s^{\mu, \chi} \left\{ \sum_{i=1}^{\min\{i \,|\, X_i \in F\}} \pi(X_{i-1}, Y_i) \right\}.$$

Observe that player Max can choose his actions to win at least an amount arbitrarily close to $\sup_{\chi \in \Sigma_{\mathrm{Max}}} \inf_{\mu \in \Sigma_{\mathrm{Min}}} \mathrm{EReach}(s, \mu, \chi)$. This is called the *lower value* $\underline{\mathrm{Val}}(s)$ of the expected reachability-price game starting at s:

$$\underline{\mathrm{Val}}(s) \overset{\mathrm{def}}{=} \sup_{\chi \in \Sigma_{\mathrm{Max}}} \inf_{\mu \in \Sigma_{\mathrm{Min}}} \mathrm{EReach}(s, \mu, \chi).$$

Similarly, player Min can choose to lose at most an amount arbitrarily close to $\inf_{\mu \in \Sigma_{\mathrm{Min}}} \sup_{\chi \in \Sigma_{\mathrm{Max}}} \mathrm{EReach}(s, \mu, \chi)$. This is the *upper value* $\overline{\mathrm{Val}}(s)$ of the game:

$$\overline{\mathrm{Val}}(s) \overset{\mathrm{def}}{=} \inf_{\mu \in \Sigma_{\mathrm{Min}}} \sup_{\chi \in \Sigma_{\mathrm{Max}}} \mathrm{EReach}(s, \mu, \chi).$$

It is easy to verify that $\underline{\mathrm{Val}}(s) \leqslant \overline{\mathrm{Val}}(s)$ for all $s \in S$. We say that the expected reachability-price game is determined if $\underline{\mathrm{Val}}(s) = \overline{\mathrm{Val}}(s)$ for all $s \in S$. In this case we also say that the value of the game exists and denote it by $\mathrm{Val}(s) = \underline{\mathrm{Val}}(s) = \overline{\mathrm{Val}}(s)$ for all $s \in S$. The determinacy of expected reachability-price games follow from Martin's determinacy theorem [25].

Proposition 1. *Every Expected reachability-price game is determined.*

For $\mu \in \Sigma_{\mathrm{Min}}$ and $\chi \in \Sigma_{\mathrm{Max}}$ we define $\mathrm{Val}^{\mu}(s) = \sup_{\chi \in \Sigma_{\mathrm{Max}}} \mathrm{EReach}(s, \mu, \chi)$ and $\mathrm{Val}_{\chi}(s) = \inf_{\mu \in \Sigma_{\mathrm{Min}}} \mathrm{EReach}(s, \mu, \chi)$. For an $\varepsilon > 0$, we say that $\mu \in \Sigma_{\mathrm{Min}}$ or $\chi \in \Sigma_{\mathrm{Max}}$ is ε-*optimal* if $\mathrm{Val}^{\mu}(s) \leqslant \mathrm{Val}(s) + \varepsilon$ or $\mathrm{Val}_{\chi}(s) \geqslant \mathrm{Val}(s) - \varepsilon$, respectively, for all $s \in S$. Since an expected reachability-price game is determined, for every $\varepsilon > 0$, both players have ε-optimal strategies. We say that a game is *positionally-determined* if for every $\varepsilon > 0$ we have strategies $\mu_{\varepsilon} \in \Pi_{\mathrm{Min}}$ and $\chi_{\varepsilon} \in \Pi_{\mathrm{Max}}$ such that for every initial state $s \in S$, we have that

$$\underline{\mathrm{Val}}(s) - \varepsilon \leqslant \mathrm{Val}_{\chi_{\varepsilon}}(s) \text{ and } \overline{\mathrm{Val}}(s) + \varepsilon \geqslant \mathrm{Val}^{\mu_{\varepsilon}}(s).$$

Given an expected reachability-price game $\mathcal{T}$, and initial state $s \in S$, and a bound $B \in \mathbb{R}$, the *expected reachability-price game problem* is to decide whether $\mathrm{Val}(s) \leqslant B$.

Optimality Equations. We now review optimality equations for characterising the value in an expected reachability-price game. Let $\mathcal{T}$ be a priced probabilistic timed game arena and let $P : S \to \mathbb{R}_{\geqslant 0}$. We say that P is a solution of optimality equations $\mathrm{Opt}(\mathcal{T})$, and we write $P \models \mathrm{Opt}(\mathcal{T})$ if, for all $s \in S$:

$$P(s) = \begin{cases} 0 & \text{if } s \in F \\ \inf_{\tau \in A(s)} \{\pi(s, \tau) + \sum_{s' \in S} p(s'|s, \tau) \cdot P(s')\} & \text{if } s \in S_{\mathrm{Min}} \backslash F \\ \sup_{\tau \in A(s)} \{\pi(s, \tau) + \sum_{s' \in S} p(s'|s, \tau) \cdot P(s')\} & \text{if } s \in S_{\mathrm{Max}} \backslash F. \end{cases}$$

Under Assumption 1, we have the following proposition.

Proposition 2. *If $P \models \mathrm{Opt}(\mathcal{T})$, then $\mathrm{Val}(s) = P(s)$ for all $s \in S$ and, for every $\varepsilon > 0$, both players have* pure ε-optimal strategies.

Proof. We show that for every $\varepsilon > 0$, there exists a pure strategy $\mu_{\varepsilon} : \mathrm{FRuns} \to A$ for player Min, such that for every strategy χ for player Max, we have $\mathrm{EReach}(s, \mu_{\varepsilon}, \chi) \leqslant P(s) + \varepsilon$. The proof, that for every $\varepsilon > 0$, there exists a pure strategy $\chi_{\varepsilon} : \mathrm{FRuns} \to A$ for player Max, such that for every strategy μ for player Min, we have $\mathrm{EReach}(s, \mu, \chi_{\varepsilon})) \geqslant P(s) - \varepsilon$, follows similarly. Together, these claims imply that P is equal to the value function of the expected reachability-time game, and the pure strategies μ_{ε} and χ_{ε}, defined in the proof below for all $\varepsilon > 0$, are ε-optimal.

Let us fix $\varepsilon > 0$ and μ_{ε} be a pure strategy where for every $n \in \mathbb{N}$ and finite play $r \in \mathrm{FRuns}$ of length n, we have $\mu_{\varepsilon}(r) = (t, a)$ such that

$$\pi(t, a) + \sum_{s' \in S} p(s'|\mathrm{Last}(r), (t, a)) \cdot P(s') \leqslant P(\mathrm{Last}(r)) + \frac{\varepsilon}{2^{n+1}}.$$

Observe that for every state $s \in S_{\mathrm{Min}}$ and for every $\varepsilon' > 0$, there is a ε'-optimal timed action because $P \models \mathrm{Opt}(\mathcal{T})$.

Again using the fact that $P \models \mathrm{Opt}(\mathcal{T})$, it follows that, that for every $s \in S_{\mathrm{Max}} \setminus F$ and $(t, a) \in A(s)$, we have

$$P(s) \geqslant \pi(t, a) + \sum_{s' \in S} p(s' | s, a) \cdot P(s'). \tag{1}$$

Now for an arbitrary strategy χ for player Max, it follows by induction that for every $n \geqslant 1$:

$$P(s) \geqslant \mathbb{E}_s^{\mu_\varepsilon, \chi} \left\{ \sum_{i=1}^{\min\{i \,|\, X_i \in F\}} \pi(X_{i-1}, Y_i) \right\} \\ + \sum_{s' \in S \setminus F} Prob_s^{\mu_\varepsilon, \chi}(X_n = s') \cdot P(s') - (1 - \tfrac{1}{2^n}) \cdot \varepsilon. \tag{2}$$

Using Assumption 1, we have $\lim_{n \to \infty} \sum_{s' \in S \setminus F} Prob_s^{\mu_\varepsilon, \chi}(X_n = s') = 0$, and therefore taking the limit in (2) we get the inequality:

$$P(s) \geqslant \mathbb{E}_s^{\mu, \chi} \Big\{ \sum_{i=1}^{\min\{i \,|\, X_i \in F\}} \pi(X_{i-1}, Y_i) \Big\} - \varepsilon = \mathrm{EReach}(s, \mu_\varepsilon, \chi) - \varepsilon,$$

which completes the proof.   □

Using Proposition 2, it follows that the problem of solving an expected reachability-price game on $\mathcal{T}$ can be reduced to solving the optimality equations $\mathrm{Opt}(\mathcal{T})$. Forejt et al. [17] showed that solving optimality equations for a reachability-time game on a probabilistic timed automata $\mathcal{T}$ can be reduced to solving a reachability-time game on an abstraction, called the boundary region abstraction. In the following section we study expected reachability-price games on boundary region abstraction.

4   ERPG on Boundary Region Abstractions

For the rest of the paper, we assume that $\mathcal{T}$ is a binary-priced PTGA. The partition of the locations of a PTGA $\mathcal{T} = (\mathsf{T}, L_{\mathrm{Min}}, L_{\mathrm{Max}})$ gives rise to a partition $(\widehat{S}_{\mathrm{Min}}, \widehat{S}_{\mathrm{Max}})$ of the set of states $\widehat{S}$ of its boundary region graph and let $\widehat{\mathcal{T}} = (\widehat{\mathsf{T}}, \widehat{S}_{\mathrm{Min}}, \widehat{S}_{\mathrm{Max}})$. Before we present our main theorem, we study the properties of non-expansive and monotone value functions. For brevity, we call such functions *nice functions*.

4.1   Nice Functions over Clock Valuations

Let $X \subseteq V$ be a subset of valuations. A function $F : X \to \mathbb{R}$ is *non-expansive* if $|F(\nu) - F(\nu')| \leqslant \|\nu - \nu'\|$ for all $\nu, \nu' \in X$.

Lemma 2 (Properties of Nice Functions). *A function $F : X \to \mathbb{R}$ is nice if it is non-expansive and monotonically decreasing. We say a function $F : \widehat{S} \to \mathbb{R}_{\geqslant 0}$ is regionally nice if for every region $(\ell, \zeta) \in \mathcal{R}$ the function $F((\ell, \cdot), (\ell, \zeta))$ is nice. The nice functions satisfy the following properties.*

1. **Continuous Closure.** *If $F : X \to \mathbb{R}$ is nice, then its unique continuous closure $\overline{F} : \overline{X} \to \mathbb{R}$ is also a nice function.*
2. **Minimum and Maximum.** *If the functions $F, F' : \widehat{S} \to \mathbb{R}$ are regionally nice functions, then $\max(F, F')$ and $\min(F, F')$ are also regionally nice.*
3. **Convex Combination.** *The $\langle f_i \rangle_{i=1}^{n}$ are nice functions then for $\langle p_i \in [0,1] \rangle_{i=1}^{n}$ with $\sum_{i=1}^{n} p_i = 1$, the function $\sum_{i=1}^{n} p_i \cdot f_i$ is nice.*
4. **Limit.** *The limit of a sequence of nice functions is nice.*

The following property of nice functions is useful in proving the correctness of the reduction to the boundary region abstraction.

Lemma 3. *Consider a binary-priced PTGA $\mathcal{T}$. Let $s = (\ell, \nu) \in S$ and $(\ell, \zeta) \in \mathcal{R}$ such that $(\ell, [\nu]) \to_* (\ell, \zeta)$. If $F : \widehat{S} \to \mathbb{R}$ is regionally nice, then the function $F_{s,\zeta,a}^{\oplus} : I \to \mathbb{R}$ defined as*

$$t \mapsto \pi(s, (t, a)) + \sum_{(C,\ell') \in 2^{\mathcal{Y}} \times L} \delta[\ell, a](C, \ell') \cdot F((\ell', \nu_C^t), (\ell', \zeta^C))$$

is continuous and monotone, where $I = \{t \in \mathbb{R}_{\geqslant 0} \mid \nu + t \in \zeta\}$, $\nu_C^t = \nu + t[C := 0]$ and $\zeta^C = \zeta[C := 0]$.

Proof. To prove this lemma we consider a $t_1 \in I$, and for all $t_2 \in I$ and $t_2 \geq t_1$, we show that $F_{s,\zeta,a}^{\oplus}(t_2) - F_{s,\zeta,a}^{\oplus}(t_1)$ is either greater than or equal to 0, or less than or equal to 0. Since $\mathcal{T}$ is binary-priced, there are two cases: (i) $r(\ell) = 1$ and (ii) $r(\ell) = 0$.

For $r(\ell) = 1$, by definition we have $F_{s,\zeta,a}^{\oplus}(t_2) - F_{s,\zeta,a}^{\oplus}(t_1)$ equals:

$$t_2 - t_1 + \sum_{(C,\ell') \in 2^{\mathcal{Y}} \times L} \delta[\ell, a](C, \ell') \cdot \big(F((\ell', \nu_C^{t_2}), (\ell', \zeta^C)) - F((\ell', \nu_C^{t_1}), (\ell', \zeta^C)) \big)$$

$$= t_2 - t_1 - \sum_{(C,\ell') \in 2^{\mathcal{Y}} \times L} \delta[\ell, a](C, \ell') \cdot \big(F((\ell', \nu_C^{t_1}), (\ell', \zeta^C)) - F((\ell', \nu_C^{t_2}), (\ell', \zeta^C)) \big)$$

$$\geqslant t_2 - t_1 - \sum_{(C,\ell') \in 2^{\mathcal{Y}} \times L} \delta[\ell,a](C,\ell') \cdot (t_2 - t_1) \geqslant 0$$

where the inequality is due to the fact the F is nice.

For the case $r(\ell) = 0$, we have that

$$F_{s,\zeta,a}^{\oplus}(t_2) - F_{s,\zeta,a}^{\oplus}(t_1) \geqslant - \sum_{(C,\ell') \in 2^{\mathcal{Y}} \times L} \delta[\ell,a](C,\ell') \cdot (t_2 - t_1)$$

Hence for the case for the case when $r(\ell) = 0$, given a $t_1 \in I$, for all $t_2 \geqslant t_1$, we have that $F_{s,\zeta,a}^{\oplus}(t_2) - F_{s,\zeta,a}^{\oplus}(t_1) \leqslant 0$, and we are done. □

4.2   Optimality Equations

Consider the optimality equations for an expected reachability-price game on a
boundary region graph $\widehat{\mathcal{T}}$. Let $P : \widehat{S} \to \mathbb{R}_{\geqslant 0}$. We say that P is a solution of
optimality equations $\text{Opt}(\widehat{\mathcal{T}})$, and we write $P \models \text{Opt}(\widehat{\mathcal{T}})$, if for every $s \in \widehat{S}$:

$$
P(s) = \begin{cases}
0 & \text{if } s \in \widehat{F} \\
\min\limits_{\alpha \in \widehat{A}(s)} \left\{ \pi(s, \alpha) + \sum\limits_{s' \in S} p(s'|s, \alpha) \cdot P(s') \right\} & \text{if } s \in \widehat{S}_{\text{Min}} \backslash \widehat{F} \\
\max\limits_{\alpha \in \widehat{A}(s)} \left\{ \pi(s, \alpha) + \sum\limits_{s' \in S} p(s'|s, \alpha) \cdot P(s') \right\} & \text{if } s \in \widehat{S}_{\text{Max}} \backslash \widehat{F}.
\end{cases}
$$

For a given function $f : \widehat{S} \to \mathbb{R}$ over boundary region abstraction, we define
a transfer function $\widetilde{f} : S \to \mathbb{R}$ over PTGA by $\widetilde{f}(\ell, \nu) = f((\ell, \nu), (\ell, [\nu]))$. The fol-
lowing theorem characterizes the conditions under which expected reachability-
price games on PTGAs can be reduced to expected reachability-price games over
the boundary region abstraction.

Theorem 1. *Let $\mathcal{T}$ be a binary-priced priced probabilistic timed game. If $P \models$
$Opt(\widehat{\mathcal{T}})$ and P is regionally nice then $\widetilde{P} \models Opt(\mathcal{T})$.*

Proof. Suppose $P \models \text{Opt}(\widehat{\mathcal{T}})$. To prove this theorem it is sufficient to show that
for every $s = (\ell, \nu) \in S_{\text{Min}}$ we have:

$$
\widetilde{P}(s) = \inf_{(t,a) \in A(s)} \left\{ \pi(s, (t,a)) + \textstyle\sum_{(C,\ell') \in 2^{\mathcal{Y}} \times L} \delta[\ell, a](C, \ell') \cdot \widetilde{P}(\ell', (\nu + t)[C := 0]) \right\}
\tag{3}
$$

and for every $s = (\ell, \nu) \in S_{\text{Max}}$ we have:

$$
\widetilde{P}(s) = \sup_{(t,a) \in A(s)} \left\{ \pi(s, (t,a)) + \textstyle\sum_{(C,\ell') \in 2^{\mathcal{Y}} \times L} \delta[\ell, a](C, \ell') \cdot \widetilde{P}(\ell', (\nu + t)[C := 0]) \right\} .
\tag{4}
$$

In the remainder of the proof we restrict attention to Min states as the
case for Max states follows similarly. Therefore we fix $s = (\ell, \nu) \in S_{\text{Min}}$ for the
remainder of the proof. For $a \in Act$, let $\mathcal{R}_{\text{Thin}}^a$ and $\mathcal{R}_{\text{Thick}}^a$ denote the set of
thin and thick regions respectively that are successors of $[\nu]$ and are subsets of
$E(\ell, a)$. Considering the right hand side (RHS) of (3) we have:

$$
\text{RHS of (3)} = \min_{a \in Act} \{ R_{\text{Thin}}(s, a), R_{\text{Thick}}(s, a) \},
\tag{5}
$$

where $R_{\text{Thin}}(s, a)$ ($R_{\text{Thick}}(s, a)$) is the RHS of (3) over all actions (t, a) such that
$[\nu + t] \in \mathcal{R}_{\text{Thin}}^a$ ($[\nu + t] \in \mathcal{R}_{\text{Thick}}^a$). For the first term of (5) we have that $R_{\text{Thin}}(s, a)$
equals:

$$\min_{\substack{(\ell,\zeta)\in\mathcal{R}^a_{\mathrm{Thin}}}} \inf_{\substack{t\in\mathbb{R}\wedge\\ \nu+t\in\zeta}} \left\{ \pi(s,(t,a)) + \sum_{(C,\ell')\in 2^{\mathcal{Y}}\times L} \delta[\ell,a](C,\ell')\cdot\widetilde{P}(\ell',\nu^t_C) \right\}$$

$$= \min_{\substack{(\ell,\zeta)\in\mathcal{R}^a_{\mathrm{Thin}}}} \inf_{\substack{t\in\mathbb{R}\wedge\\ \nu+t\in\zeta}} \left\{ \pi(s,(t,a)) + \sum_{(C,\ell')\in 2^{\mathcal{Y}}\times L} \delta[\ell,a](C,\ell')\cdot P((\ell',\nu^t_C),(\ell',\zeta_C)) \right\}$$

$$= \min_{\substack{(\ell,\zeta)\in\mathcal{R}^a_{\mathrm{Thin}}}} \left\{ \pi(s,(t^{(\ell,\zeta)},a)) + \sum_{(C,\ell')\in 2^{\mathcal{Y}}\times L} \delta[\ell,a](C,\ell')\cdot P((\ell',\nu^{t^{(\ell,\zeta)}}_C),(\ell',\zeta^C)) \right\}$$

where ν^t_C denotes the clock valuation $(\nu+t)[C{:=}0]$, $t^{(\ell,\zeta)}$ the time to reach the region R from s and ζ^C the region $\zeta[C{:=}0]$. Considering the second term of (5) we have that $R_{\mathrm{Thick}}(s,a)$ equals

$$\min_{\substack{(\ell,\zeta)\in\mathcal{R}^a_{\mathrm{Thick}}}} \inf_{\substack{t\in\mathbb{R}\wedge\\ \nu+t\in\zeta}} \left\{ \pi(s,(t,a)) + \sum_{(C,\ell')\in 2^{\mathcal{Y}}\times L} \delta[\ell,a](C,\ell')\cdot\widetilde{P}(\ell',\nu^t_C) \right\}$$

$$= \min_{\substack{(\ell,\zeta)\in\mathcal{R}^a_{\mathrm{Thick}}}} \inf_{\substack{t\in\mathbb{R}\wedge\\ \nu+t\in\zeta}} \left\{ \pi(s,(t,a)) + \sum_{(C,\ell')\in 2^{\mathcal{Y}}\times L} \delta[\ell,a](C,\ell')\cdot P((\ell',\nu^t_C),(\ell',\zeta^C)) \right\}$$

$$= \min_{\substack{(\ell,\zeta)\in\mathcal{R}^a_{\mathrm{Thick}}}} \inf_{\substack{t^s_{R_-}<t<t^s_{R_+}\\ R\leftarrow_{+1}R_-\\ R\rightarrow_{+1}R_+}} \left\{ \pi(s,(t,a)) + \sum_{(C,\ell')\in 2^{\mathcal{Y}}\times L} \delta[\ell,a](C,\ell')\cdot P((\ell',\nu^t_C),(\ell',\zeta^C)) \right\}$$

Now P is regionally nice and, from Lemma 3 it follows that

$$\pi(s,(t,a)) + \sum_{(C,\ell')\in 2^{\mathcal{Y}}\times L} \delta[\ell,a](C,\ell')\cdot P((\ell',\nu^t_C),(\ell',\zeta^C))$$

is continuous and monotone over $\{t\,|\,\nu+t\in\zeta\}$, and hence minimized at one of the boundaries, that is, either at $t^s_{R_-}$ or at $t^s_{R_+}$. Therefore it follows that $R_{\mathrm{Thick}}(s,a)$ equals

$$\min_{\substack{(\ell,\zeta)\in\mathcal{R}^a_{\mathrm{Thick}}}} \min_{\substack{t=t^s_{R_-},t^s_{R_+}\\ (\ell,\zeta)\leftarrow_{+1}R_-\\ (\ell,\zeta)\rightarrow_{+1}R_+}} \left\{ \pi(s,(t,a)) + \sum_{(C,\ell')\in 2^{\mathcal{Y}}\times L} \delta[\ell,a](C,\ell')\cdot P((\ell',\nu^t_C),(\ell',\zeta^C)) \right\}$$

Substituting the values of $R_{\mathrm{Thin}}(s,a)$ and $R_{\mathrm{Thick}}(s,a)$ into (5) and observing that, for every thin region $(\ell,\zeta)\in\mathcal{R}^a_{\mathrm{Thin}}$, there exist $b\in\mathbb{Z}$ and $c\in C$ such that $\nu+(b-\nu(c))\in\zeta$, it follows from Definition 5 that RHS of (3) equals:

$$\min_{\alpha\in\widehat{A}(s,[s])} \left\{ \widehat{\pi}((s,[s]),\alpha) + \sum_{(s',R')\in\widehat{S}} \widehat{p}((s',R')|(s,[s]),\alpha)\cdot P(s',R') \right\}$$

which by definition equals $\widetilde{P}(s)$ as required. $\square$

5   One-Clock Binary-Priced PTGA

Theorem 1 characterizes conditions on the solution of optimality equations of the boundary region abstraction for binary-priced PTGA under which its solution also gives the solution for the corresponding PTGA. In this section we study binary-priced one-clock PTGA and show that the solution of optimality equations for such games remain regionally nice. This result together with Theorem 1 proves the correctness of reduction of the expected reachability-price games for 1-clock binary-priced PTGA to the similar problem on the corresponding boundary region abstraction. Let $\mathcal{T}$ be a one-clock binary-priced PTGA (1BPTA) and we denote by x the only clock of the 1BPTA.

We prove the following property of nice functions in the context of one-clock binary-priced PTGAs.

Lemma 4. *Let $\mathcal{T}$ be a one-clock binary-priced PTGA and $\widehat{\mathcal{T}}$ be its boundary region abstraction. If F is regionally nice, then, for every $R = (\ell, \zeta)$ and $\alpha \in \widehat{A}(R)$, the function $F^{\boxplus}_{(\ell,R,\alpha)} : V \to \mathbb{R}$ defined as*

$$\nu \mapsto \pi(((\ell, \nu), R), \alpha) + \sum_{s' \in S} p(s'|((\ell, \nu), R), \alpha) \cdot F(s')$$

is nice.

Proof. Let us rewrite the function $F^{\boxplus}_{(\ell,R,\alpha)} : V \to \mathbb{R}$ as

$$\nu \mapsto \sum_{s' \in S} p(s'|((\ell, \nu), R), \alpha) \cdot (\pi(((\ell, \nu), R), \alpha) + F(s')).$$

From Lemma 2 (3), it suffices to show that $(\pi(((\ell, \nu), R), \alpha) + F(s'))$ is a nice function. Let $\nu(x) - \nu'(x) = d$. There are several cases to consider.

1. Price-rate of ℓ is 0. There are two cases to consider.
   (a) The action α resets the clock. In this case, the function $(\pi(((\ell, \nu), R), \alpha) + F(s'))$ is constant, and hence nice.
   (b) The action α does not reset the clock. There are two cases to consider. The first case is when α suggests 0 time delay. In this case, $(\pi(((\ell, \nu), R), \alpha) + F(s')) = F(\ell', \nu)$ and that is a nice function in ν as F is a regionally nice function. The second case is when α suggests $d - \nu(x)$ time delay. In this case, $(\pi(((\ell, \nu), R), \alpha) + F(s')) = 0 * (d - \nu(x)) + F(\ell', d)$ is constant, and hence a nice function.
2. Price-rate of ℓ is 1. In this case there are two cases to consider.
   (a) The action α resets the clock. In this case, the function $(\pi(((\ell, \nu), R), \alpha) + F(s'))$ is a simple function, and hence nice.
   (b) The action α does not reset the clock. There are two cases to consider. The first case is when α suggests 0 time delay. In this case, $(\pi(((\ell, \nu), R), \alpha) + F(s')) = F(\ell', \nu)$ and that is a nice function in ν as F is a regionally nice function. The second case is when α suggests $d - \nu(x)$ time delay.

In this case, $(\pi(((\ell, \nu), R), \alpha) + F(s')) = d - \nu(x) + F(\ell', d) = c' - \nu(x)$. It is easy to see that this function is also non-expansive and monotonically decreasing.

The proof is now complete. □

We are now ready to state the key result of this section.

Proposition 3. *Let $\mathcal{T}$ be a one-clock binary-priced PTGA. If $P \models Opt(\widehat{\mathcal{T}})$, then P is regionally nice.*

Proof. Based on the optimality equations, we define the value improvement function $\mathbf{\Psi} : [\widehat{S} \to \mathbb{R}_{\geqslant 0}] \to [\widehat{S} \to \mathbb{R}_{\geqslant 0}]$ such that for any $f : \widehat{S} \to \mathbb{R}_{\geqslant 0}$ and $s \in \widehat{S}$:

$$\mathbf{\Psi}(f)(s) = \begin{cases} 0 & \text{if } s \in \widehat{F} \\ \min\limits_{\alpha \in \widehat{A}(s)} \left\{ \pi(s, \alpha) + \sum\limits_{s' \in S} p(s'|s, \alpha) \cdot f(s') \right\} & \text{if } s \in \widehat{S}_{\text{Min}} \backslash \widehat{F} \\ \max\limits_{\alpha \in \widehat{A}(s)} \left\{ \pi(s, \alpha) + \sum\limits_{s' \in S} p(s'|s, \alpha) \cdot f(s') \right\} & \text{if } s \in \widehat{S}_{\text{Max}} \backslash \widehat{F}. \end{cases} \quad (6)$$

Since we consider stopping assumption on PTGA, (Assumption 1) and in $\widehat{\mathcal{T}}$ every state can only reach a finite sub-graph (Lemma 1), it follows that $\mathbf{\Psi}^N$ is a p-contractive mapping where p is the smallest probability appearing on the transitions of the PTGA and N is a finite upper bound (Lemma 1) on the number of states reachable from any state in the corresponding BRA. Therefore, from Banach's fixed point theorem $\mathbf{\Psi}$ can be used in an iterative scheme to converge to the solution of optimality equations $Opt(\widehat{\mathcal{T}})$.

Starting the iterative scheme with a regionally nice function, from Lemma 4, along with Lemmas 2(1–3), it follows that the intermediate iterates of $\mathbf{\Psi}$ in (6) remain regionally nice. Now Lemma 2 (4) implies that the limit of these sequences in also regionally nice, it follows that the fixpoint is also regionally nice. Hence, for every one-clock binary-priced PTGA, the value of the optimality equations $Opt(\widehat{\mathcal{T}})$ is regionally nice. □

The following theorem follows from Theorem 1 and Proposition 3.

Theorem 2. *The value of the expected reachability-price game on a one-clock binary-priced PTGA $\mathcal{T}$ is equal to the value of the game on the corresponding corner-point abstraction.*

Corollary 1. *The expected reachability-price game problem is decidable for one-clock binary-priced PTGA.*

6   Conclusion

In this work, we consider two-player games with expected reachability-price objective over probabilistic timed automata. We show the decidability for one-clock binary-priced PTGA where the expected reachability-price problem can be

reduced to the same problem over a boundary region graph abstraction. For this purpose, we use the notion of *nice functions* that is a generalization of simple functions introduced by Asarin and Maler in [3].

One clock timed automata have been widely studied for two-player games. They can be used to model the time difference between two actions or time needed to finish an action and so on. For two-player games, reachability time objectives have been studied in [3] and [17]. Having both 0 and 1 costs is certainly more expressive than having only cost 1 and the latter is well studied. On the other hand, in [13] it has been shown that the reachability-price problem becomes undecidable even for timed automata with more than two costs. This along with the different examples shown in the paper indicate that the decidability result is strong enough and may lead to undecidability by generalizing the class of one-clock binary-priced PTGAs.

We note that the definition of boundary region abstraction has been developed for general PTGA and decidability can be recovered for cases where solutions of $Opt(\widehat{\mathcal{T}})$ is regionally nice and the cost functions are minimized or maximized at some boundary of a region. In particular, this has been shown to be the case for one-clock binary-priced PTGA using Lemma 4 and Lemma 3. However, it may be possible to extend the technique to broader classes of PTGA.

References

1. Alur, R., Bernadsky, M., Madhusudan, P.: Optimal reachability for weighted timed games. In: Díaz, J., Karhumäki, J., Lepistö, A., Sannella, D. (eds.) ICALP 2004. LNCS, vol. 3142, pp. 122–133. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-27836-8_13
2. Alur, R., Dill, D.: A theory of timed automata. Theor. Comput. Sci. **126**(2), 183–235 (1994). https://doi.org/10.1016/0304-3975(94)90010-8
3. Asarin, E., Maler, O.: As soon as possible: time optimal control for timed automata. In: Vaandrager, F.W., van Schuppen, J.H. (eds.) Proceedings of HSCC, pp. 19–30 (1999). https://doi.org/10.1007/3-540-48983-5_6
4. Asarin, E., Maler, O., Pnueli, A.: Symbolic controller synthesis for discrete and timed systems. In: Antsaklis, P., Kohn, W., Nerode, A., Sastry, S. (eds.) HS 1994. LNCS, vol. 999, pp. 1–20. Springer, Heidelberg (1995). https://doi.org/10.1007/3-540-60472-3_1
5. Beauquier, D.: On probabilistic timed automata. Theor. Comput. Sci. **292**(1), 65–84 (2003). https://doi.org/10.1016/S0304-3975(01)00215-8
6. Behrmann, G., Cougnard, A., David, A., Fleury, E., Larsen, K.G., Lime, D.: UPPAAL-tiga: time for playing games!. In: Damm, W., Hermanns, H. (eds.) CAV 2007. LNCS, vol. 4590, pp. 121–125. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73368-3_14
7. Berendsen, J., Chen, T., Jansen, D.N.: Undecidability of cost-bounded reachability in priced probabilistic timed automata. In: Chen, J., Cooper, S.B. (eds.) TAMC 2009. LNCS, vol. 5532, pp. 128–137. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-02017-9_16
8. Bouyer, P., Brihaye, T., Bruyère, V., Raskin, J.F.: On the optimal reachability problem on weighted timed automata. Formal Meth. Syst. Des. **31**(2), 135–175 (2007). https://doi.org/10.1007/s10703-007-0035-4

9. Bouyer, P., Brihaye, T., Markey, N.: Improved undecidability results on weighted timed automata. Inf. Process. Lett. **98**, 188–194 (2006). https://doi.org/10.1016/j.ipl.2006.01.012

10. Bouyer, P., Brinksma, E., Larsen, K.G.: Staying alive as cheaply as possible. In: Alur, R., Pappas, G.J. (eds.) HSCC 2004. LNCS, vol. 2993, pp. 203–218. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-24743-2_14

11. Bouyer, P., Cassez, F., Fleury, E., Larsen, K.G.: Optimal strategies in priced timed game automata. In: Lodaya, K., Mahajan, M. (eds.) FSTTCS 2004. LNCS, vol. 3328, pp. 148–160. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30538-5_13

12. Brihaye, T., Geeraerts, G., Krishna, S.N., Manasa, L., Monmege, B., Trivedi, A.: Adding negative prices to priced timed games. In: Proceedings of CONCUR, pp. 560–575 (2014). https://doi.org/10.1007/978-3-662-44584-6_38

13. Brihaye, T., Henzinger, T.A., Prabhu, V.S., Raskin, J.-F.: Minimum-time reachability in timed games. In: Arge, L., Cachin, C., Jurdziński, T., Tarlecki, A. (eds.) ICALP 2007. LNCS, vol. 4596, pp. 825–837. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73420-8_71

14. Brihaye, T., Bruyère, V., Raskin, J.-F.: On optimal timed strategies. In: Pettersson, P., Yi, W. (eds.) FORMATS 2005. LNCS, vol. 3829, pp. 49–64. Springer, Heidelberg (2005). https://doi.org/10.1007/11603009_5

15. Cassez, F., Jessen, J.J., Larsen, K.G., Raskin, J.-F., Reynier, P.-A.: Automatic synthesis of robust and optimal controllers – an industrial case study. In: Majumdar, R., Tabuada, P. (eds.) HSCC 2009. LNCS, vol. 5469, pp. 90–104. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-00602-9_7

16. Alfaro, L.: Computing minimum and maximum reachability times in probabilistic systems. In: Baeten, J.C.M., Mauw, S. (eds.) CONCUR 1999. LNCS, vol. 1664, pp. 66–81. Springer, Heidelberg (1999). https://doi.org/10.1007/3-540-48320-9_7

17. Forejt, V., Kwiatkowska, M., Norman, G., Trivedi, A.: Expected reachability-time games. Theor. Comput. Sci. **631**, 139–160 (2016). https://doi.org/10.1016/j.tcs.2016.04.021

18. Hoffmann, G., Wong-Toi, H.: The input-output control of real-time discrete event systems. In: IEEE Real-Time Systems Symposium (RTSS). pp. 256–265 (1992). https://doi.org/10.1109/REAL.1992.242655

19. Jensen, H.: Model checking probabilistic real time systems. In: Bjerner, B., Larsson, M., Nordström, B. (eds.) Proceedings 7th Nordic Workshop Programming Theory, pp. 247–261. Report 86:247–261, Chalmers University of Technology (1996). https://doi.org/10.1.1.23.2754

20. Jovanović, A., Kwiatkowska, M., Norman, G.: Symbolic minimum expected time controller synthesis for probabilistic timed automata. In: Sankaranarayanan, S., Vicario, E. (eds.) FORMATS 2015. LNCS, vol. 9268, pp. 140–155. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-22975-1_10

21. Jurdziński, M., Kwiatkowska, M., Norman, G., Trivedi, A.: Concavely-priced probabilistic timed automata. In: Bravetti, M., Zavattaro, G. (eds.) CONCUR 2009. LNCS, vol. 5710, pp. 415–430. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04081-8_28

22. Jurdziński, M., Trivedi, A.: Reachability-time games on timed automata. In: Arge, L., Cachin, C., Jurdziński, T., Tarlecki, A. (eds.) ICALP 2007. LNCS, vol. 4596, pp. 838–849. Springer, Heidelberg (2007). https://doi.org/10.1007/978-3-540-73420-8_72

23. Jurdziński, M., Trivedi, A.: Average-time games. In: Hariharan, R., Mukund, M., Vinay, V. (eds.) Proceedings of FSTTCS. Dagstuhl Seminar Proceedings (2008). https://doi.org/10.4230/LIPIcs.FSTTCS.2008.1765
24. Kwiatkowska, M., Norman, G., Segala, R., Sproston, J.: Automatic verification of real-time systems with discrete probability distributions. Theor. Comput. Sci. **282**, 101–150 (2002). https://doi.org/10.1016/S0304-3975(01)00046-9
25. Martos, B.: The direct power of adjacent vertex programming methods. Manage. Sci. **12**(3), 241–252 (1965). https://doi.org/10.1287/mnsc.12.3.241. http://www.jstor.org/stable/2627581
26. Neyman, A., Sorin, S. (eds.): Stochastic Games and Applications. NATO Science Series C, vol. 570. Kluwer Academic Publishers, Dordrecht (2004). https://doi.org/10.1007/978-94-010-0189-2
27. Ramadge, P.J., Wonham, W.M.: The control of discrete event systems. IEEE **77**, 81–98 (1989). https://doi.org/10.1109/5.21072