Computing Laboratory

EXPECTED REACHABILITY-TIME GAMES

Vojtěch Forejt[1]
Marta Kwiatkowska[2]
Gethin Norman[3]
Ashutosh Trivedi[2]

[1] Faculty of Informatics, Masaryk University, Botanická 68a, Brno, Czech Republic

[2] Oxford University Computing Laboratory, Parks Road, Oxford, OX1 3QD, UK

[3] Department of Computing Science, University of Glasgow, Glasgow, G12 8RZ, UK

CL-RR-10-07

Oxford University Computing Laboratory
Wolfson Building, Parks Road, Oxford OX1 3QD

Abstract

In an expected reachability-time game (ERTG) two players, Min and Max, move a token along the transitions of a probabilistic timed automaton, so as to minimise and maximise, respectively, the expected time to reach a target. These games are concurrent since at each step of the game both players choose a timed move (a time delay and action under their control), and the transition of the game is determined by the timed move of the player who proposes the shorter delay. A game is turn-based if at any step of the game, all available actions are under the control of precisely one player. We show that while concurrent ERTGs are not always determined, turn-based ERTGs are positionally determined. Using the boundary region graph abstraction, and a generalisation of Asarin and Maler's simple function, we show that the decision problems related to computing the upper/lower values of concurrent ERTGs, and computing the value of turn-based ERTGs are decidable and their complexity is in NEXPTIME $\cap$ co-NEXPTIME.

1 Introduction

Two-player zero-sum games on finite automata, as a mechanism for supervisory controller synthesis of discrete event systems, were introduced by Ramadge and Wonham [1]. In this setting the two players—called Min and Max—represent the *controller* and the *environment*, and control-program synthesis corresponds to finding a winning (or optimal) strategy of the controller for some given performance objective. If the objectives are dependent on time, e.g. when the objective corresponds to completing a given set of tasks within some deadline, then games on timed automata are a well-established approach for controller synthesis, see e.g. [2, 3, 4, 5, 6].

In this paper we extend this approach to objectives that are quantitative both in terms of timed *and* probabilistic behaviour. Probabilistic behaviour is important in modelling, e.g., faulty or unreliable components, the random coin flips of distributed communication and security protocols, and performance characteristics. We consider games on probabilistic timed automata (PTAs) [7, 8, 9], a model for real-time systems exhibiting nondeterministic and probabilistic behaviour. We concentrate on *expected reachability-time games* (ERTGs), which are games on PTAs where the performance objective concerns the minimum expected time the controller can ensure for the system to reach a target, regardless of the uncontrollable (environmental) events that occur. This approach has many practical applications, e.g., in job-shop scheduling, where machines can be faulty or have variable execution time, and both routing and task graph scheduling problems. For real-life examples relevant to our setting, see e.g. [10, 6].

In the games that we study, a token is placed on a configuration of a PTA and a play of the game corresponds to both players proposing a timed move of the PTA, i.e. a time delay and action under their control (we assume each action of the PTA is under the control of precisely one of the players). Once the players have made their choices, the timed move with the shorter delay[1] is performed and the token is moved according to the probabilistic transition function of the PTA. Players Min and Max choose their moves in order to minimise and maximise, respectively, the payoff function (the time till the first visit of a target in the case of ERTGs). It is well known, see,

[1] Min and Max represent two different forms of non-determinism called *angelic* and *demonic*. To prevent the introduction of a third form, we assume the move of Max (the environment) is taken if the delays are equal. The converse can be used without changing the presented results.

e.g. [11], that concurrent timed games are not *determined*, which means the *upper value* of the game (the minimum expected time to reach a target that Min can ensure) is strictly greater that the *lower value* of the game (the maximum expected time to reach a target 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 show that a subclass of ERTGs, called *turn-based* ERTGs, where at each step of the game only one of the players has available actions are *positionally determined*, i.e. both players have ε-optimal (optimal up to a given precision $\varepsilon>0$) positional (history-independent and non-randomised) strategies.

The problem we consider is inspired by Asarin and Maler [2] who studied the *brachystochronic problem* for timed automata. This work focused on reachability-time games, i.e. games on a timed automata where the objective concerns the time to reach a target. The techniques of [2] exploit properties of a special class of functions called *simple functions*. The importance of simple functions is also observed by Courcoubetis and Yannakakis [12] in the context of one-player games. Simple functions have also enabled the computation of a uniform solution for (turn-based) reachability-time games [13] and the proof of correctness of game-reduction for turn-based average-time games [14]. However, we show that the concept of simple functions is not sufficient in the setting of PTAs.

Contribution. We show that the problem of deciding whether the upper (lower, or the optimal when it exists) value of an ERTG is at most a given bound is decidable. An important contribution of the paper is the generalisation of simple functions to *quasi-simple functions*. By using this class of functions and the boundary region abstraction [15, 16], we give a novel proof of positional determinacy of *turn-based* ERTGs. We demonstrate that the problem of finding the upper and lower value of general ERTGs is in NEXPTIME $\cap$ co-NEXPTIME. An EXPTIME-hardness lower bound follows from the EXPTIME-completeness of the corresponding optimisation problem [16]. From [17] it follows that the problem is not NEXPTIME-hard, unless NP equals co-NP. Extending this work we get the similar results for expected discounted-time games.

Related Work. Hoffman and Wong-Toi [18] were the first to define and solve optimal controller synthesis problem for timed automata. For a detailed introduction to the topic of qualitative games on timed automata, see e.g. [19]. Asarin and Maler [2] initiated the study of quantitative games on timed automata by providing a symbolic algorithm to solve reachability-time games. The work of [20] and [13] show that the decision version of the reachability-time game is EXPTIME-complete for timed automata with at least two clocks. The tool UPPAAL Tiga [5] is capable of solving reachability and safety objectives for games on timed automata. Jurdziński and Trivedi [14] show the EXPTIME-completeness for average-time games on automata with two or more clocks.

A natural extension of reachability-time games are games on priced timed automata where the objective concerns the cumulated price of reaching a target. Both [3] and [4] present semi-algorithms for computing the value of such games for linear prices. In [21] the problem of checking the existence of optimal strategies is shown to be undecidable with [22] showing undecidability holds even for three clocks and stopwatch prices.

We are not aware of any previous work studying two-player quantitative games on PTAs. For a significantly different model of stochastic timed games, deciding whether a target is reachable within a given bound is undecidable [23]. Regarding one-player games on PTAs, [16] reduce a number of optimisation problems on concavely-priced PTAs to solving the corresponding problems on the boundary region abstraction and [24] solve expected reachability-price problems for linearly-priced PTAs using digital clocks. In [25] the problem of deciding whether a target can be reached within a given price and probability bound is shown to be undecidable for priced PTAs with three clocks and stopwatch prices. By a simple modification of the proofs in [25] it can be demonstrated that checking the existence of optimal strategies is undecidable for reachability-price turn-based games on priced (probabilistic) timed automata with three clocks and stopwatch prices.

This is a the technical report version of the paper [26].

2 Expected Reachability Games

Expected reachability games (ERGs) are played between two players Min and Max on a state-transition graph, whose transitions are nondeterministic and probabilistic, by jointly resolving the nondeterminism to move a token along the transitions of the graph. The objective for player Min in the game is to reach the final states with the smallest accumulated reward, while Max tries to do the opposite.

Before we give a formal definition, we need to introduce the concept of discrete probability distributions. A *discrete distribution* over a (possibly uncountable) set Q is a function $d : Q{\to}[0,1]$ such that $supp(d){=}\ \{q \in Q \mid d(q){>}0\}$ is at most countable and $\sum_{q \in Q} d(q){=}1$. Let $\mathcal{D}(Q')$ denote the set of all discrete distributions over Q. We say a distribution $d \in \mathcal{D}(Q)$ is a *point distribution* if $d(q){=}1$ for some $q \in Q$.

Definition 1 *An ERG is a tuple* $\mathsf{G}{=}(S, F, A_{\mathrm{Min}}, A_{\mathrm{Max}}, p_{\mathrm{Min}}, p_{\mathrm{Max}}, \pi_{\mathrm{Min}}, \pi_{\mathrm{Max}})$ *where:*

- S *is a (possibly uncountable) set of* states *including a set of* final states F*;*
- A_{Min} *and* A_{Max} *are (possibly uncountable) sets of actions* controlled *by players Min and Max and* $\perp$ *is a distinguished action such that* $A_{\mathrm{Min}} \cap A_{\mathrm{Max}} = \{\perp\}$*;*
- $p_{\mathrm{Min}} : S{\times}A_{\mathrm{Min}} \to \mathcal{D}(S)$ *and* $p_{\mathrm{Max}} : S{\times}A_{\mathrm{Max}} \to \mathcal{D}(S)$ *are the partial probabilistic transition functions for players Min and Max such that* $p_{\mathrm{Min}}(s, \perp)$ *and* $p_{\mathrm{Max}}(s, \perp)$ *are undefined for all* $s \in S$*;*
- $\pi_{\mathrm{Min}} : S{\times}A_{\mathrm{Min}} \to \mathbb{R}_{\geq 0}$ *and* $\pi_{\mathrm{Max}} : S{\times}A_{\mathrm{Max}} \to \mathbb{R}_{\geq 0}$ *are the reward functions for players Min and Max.*

We say that the ERG is *finite* if both S and A are finite. For any state s, we let $A_{\mathrm{Min}}(s)$ denote the set of actions available to player Min in s, i.e., the actions $a \in A_{\mathrm{Min}}$ for which $p_{\mathrm{Min}}(s, a)$ is defined, letting $A_{\mathrm{Min}}(s){=}\perp$ if no such action exists. Similarly, $A_{\mathrm{Max}}(s)$ denotes the actions available to player Max in s and we let $A(s){=}A_{\mathrm{Min}}(s){\times}A_{\mathrm{Max}}(s)$. We say that s is *controlled* by Min (Max) if $A_{\mathrm{Max}}(s){=}\{\perp\}$ ($A_{\mathrm{Min}}(s){=}\{\perp\}$) and the game G is *turn-based* if there is a partition $(S_{\mathrm{Min}}, S_{\mathrm{Max}})$ of S such that all states in S_{Min} (S_{Max}) are controlled by Min (Max).

A game G starts with a token in some initial state and players Min and Max construct an infinite play by repeatedly choosing enabled actions, and then moving the token to a successor state determined by their probabilistic transition functions where the reward of the move is determined by their reward functions. More precisely, if in state s players Min and Max choose actions a and b respectively, then if $\pi_{\mathrm{Min}}(s,a)<\pi_{\mathrm{Max}}(s,b)$ or $b=\bot$ the probabilistic transition function and reward value are determined by Min's choice, i.e. by the transition function $p_{\mathrm{Min}}(s,a)$ and reward value $\pi_{\mathrm{Min}}(s,a)$, and otherwise are determined by Max's choice. Formally we introduce the following auxiliary functions of an ERG which return the transition function and reward value of the game.

Definition 2 *Let* G *be an ERG. The* probabilistic transition function $p : S{\times}A_{\mathrm{Min}}{\times}A_{\mathrm{Max}}{\to}\mathcal{D}(S)$ *and the* reward function $\pi : S{\times}A_{\mathrm{Min}}{\times}A_{\mathrm{Max}}{\to}\mathbb{R}_{\geq 0}$ *of* G *are such that for any* $s \in S$ *and* $(a,b) \in A_{\mathrm{Min}}$*:*

$$
\begin{aligned}
p(s,a,b) &= \begin{cases} \textit{undefined} & \textit{if } a=b=\bot \\ p_{\mathrm{Min}}(s,a) & \textit{if } a\neq\bot \textit{ and either } b=\bot \textit{ or } \pi_{\mathrm{Min}}(s,a)<\pi_{\mathrm{Max}}(s,b) \\ p_{\mathrm{Max}}(s,b) & \textit{otherwise} \end{cases} \\
\pi(s,a,b) &= \begin{cases} \pi_{\mathrm{Min}}(s,a) & \textit{if } b=\bot \textit{ or } \pi_{\mathrm{Min}}(s,a)<\pi_{\mathrm{Max}}(s,b) \\ \pi_{\mathrm{Max}}(s,b) & \textit{otherwise.} \end{cases}
\end{aligned}
$$

From the conditions imposed on the probabilistic transition function, it follows that $(a,b)\in A(s)$ if and only if $p(s,a,b)$ is defined. Using these definitions, if in state s the action pair $(a,b)\in A(s)$ is chosen, then the probability of making a transition to s' equals $p(s'|s,a,b)\overset{\mathrm{def}}{=}p(s,a,b)(s')$ and the reward equals $\pi(s,a,b)$.

A transition of G is a tuple $(s,(a,b),s')$ such that $p(s'|s,a,b)>0$ and a play is an finite or infinite sequence $\langle s_0,(a_1,b_1),s_1,\ldots\rangle$ such that $(s_i,(a_{i+1},b_{i+1}),s_{i+1})$ is a transition for all $i\geq 0$. For a finite play $\rho=\langle s_0,(a_1,b_1),s_1,\ldots,s_k\rangle$, let $\mathit{last}(\rho)$ denote the last state s_k of the play. We write Play ($\mathit{Play}_{\mathit{fin}}$) for the sets of (finite) plays in G and $\mathit{Play}(s)$ ($\mathit{Play}_{\mathit{fin}}(s)$) for the sets of (finite) plays starting from $s\in S$.

A *strategy* of Min is a function $\mu : \mathit{Play}_{\mathit{fin}}{\to}\mathcal{D}(A_{\mathrm{Min}})$ such that $\mathit{supp}(\mu(\rho))\subseteq A_{\mathrm{Min}}(\mathit{last}(\rho))$ for all finite plays $\rho\in\mathit{Play}_{\mathit{fin}}$, i.e. for any finite play, a strategy returns a distribution over actions available to Min in the last state of the play. A strategy χ of Max is defined analogously and we let Σ_{Min} and Σ_{Max} denote the sets of strategies of Min and Max, respectively. A strategy σ is *pure* if $\sigma(\rho)$ is a point distribution for all $\rho\in\mathit{Play}_{\mathit{fin}}$, while it is *stationary* if $\mathit{last}(\rho){=}\mathit{last}(\rho')$ implies $\sigma(\rho){=}\sigma(\rho')$ for all $\rho,\rho'\in\mathit{Play}_{\mathit{fin}}$. A strategy is *positional* if it is pure and stationary and let Π_{Min} and Π_{Max} denote the set of positional strategies of Min and Max, respectively.

For any state s and strategy pair $(\mu,\chi)\in\Sigma_{\mathrm{Min}}{\times}\Sigma_{\mathrm{Max}}$, let $\mathit{Play}^{\mu,\chi}(s)$ denote the infinite plays in which Min and Max play according to μ and χ, respectively. Using standard results from probability theory, see e.g., [27], we can construct a probability measure $\mathit{Prob}^{\mu,\chi}_s$ over the set $\mathit{Play}^{\mu,\chi}(s)$. Let X_i and Y_i denote the random variables corresponding to i^{th} state and action of a play (i.e., for play $\langle s_0,(a_1,b_1),s_1,\ldots\rangle$ we have $X_i{=}s_i$ and $Y_{i+1}{=}(a_{i+1},b_{i+1})$), and given a *real-valued random variable* $f : \mathit{Play}\to\mathbb{R}$, let $\mathbb{E}^{\mu,\chi}_s\{f\}$ denote the expected value of f with respect to the probability measure $\mathit{Prob}^{\mu,\chi}_s$. To keep the presentation simple, for the rest of the paper we only consider *transient stochastic games* [28, Chapter 4] (games where every play is finite with probability 1) and for this reason we make the following assumption[2].

[2]Techniques (see, e.g., *positive stochastic games* [28, Chapter 4]) for lifting such an assumption are orthogonal to the main idea presented in this paper.

Assumption 1 *For any strategy pair* $(\mu, \chi) \in \Sigma_{\mathrm{Min}} \times \Sigma_{\mathrm{Max}}$*, and state* $s \in S$ *there is* $q > 0$ *and* $n \in \mathbb{N}$ *such that* $Prob_s^{\mu,\chi}(X_n \in F) \geq q$*.*

Recall that the objective for Min is to reach a final state with the smallest accumulated reward, while for Max it is the opposite. Starting from s, if Min uses the strategy μ and Max χ, then the expected reward accumulated before reaching a final state is given by:

$$\mathrm{EReach}^{\mu,\chi}(s) \stackrel{\mathrm{def}}{=} \mathbb{E}_s^{\mu,\chi}\left\{\sum_{i=0}^{\min\{k-1 \mid X_k \in F\}} \pi(X_i, Y_{i+1})\right\}.$$

Observe when starting at state s, Max can choose actions such that the expected reward is *at least* a value arbitrarily close to $\sup_{\chi \in \Sigma_{\mathrm{Max}}} \inf_{\mu \in \Sigma_{\mathrm{Min}}} \mathrm{EReach}^{\mu,\chi}(s)$. This is called the *lower value* $\mathrm{Val}_*(s)$ of the game when starting at state s. For $\chi \in \Sigma_{\mathrm{Max}}$ let $\mathrm{Val}_\chi(s) = \inf_{\mu \in \Sigma_{\mathrm{Min}}} \mathrm{EReach}^{\mu,\chi}(s)$. We say χ is *optimal* (ε*-optimal*), if $\mathrm{Val}_\chi(s) = \mathrm{Val}_*(s)$ ($\mathrm{Val}_\chi(s) \geq \mathrm{Val}_*(s) - \varepsilon$) for all $s \in S$. Similarly, Min can make choices such that the expected reward is *at most* a value arbitrarily close to the *upper value* $\mathrm{Val}^*(s) = \inf_{\mu \in \Sigma_{\mathrm{Min}}} \sup_{\chi \in \Sigma_{\mathrm{Max}}} \mathrm{EReach}^{\mu,\chi}(s)$. In addition, for $\mu \in \Sigma_{\mathrm{Min}}$, we can define $\mathrm{Val}^\mu(s)$ and say when μ is *optimal* or ε*-optimal*.

A game G is *determined* if $\mathrm{Val}_*(s) = \mathrm{Val}^*(s)$ for all $s \in S$ and then we say that the value of the game exists and equals $\mathrm{Val}(s) = \mathrm{Val}_*(s) = \mathrm{Val}^*(s)$. If G is determined, then each player has an ε-optimal strategy for all $\varepsilon > 0$. A game is *positionally determined* if

$$\mathrm{Val}(s) = \inf_{\mu \in \Pi_{\mathrm{Min}}} \sup_{\chi \in \Sigma_{\mathrm{Max}}} \mathrm{EReach}^{\mu,\chi}(s) = \sup_{\chi \in \Pi_{\mathrm{Max}}} \inf_{\mu \in \Sigma_{\mathrm{Min}}} \mathrm{EReach}^{\mu,\chi}(s)$$

for all $s \in S$. It is straightforward to see that if a game is positionally determined, then both players have *positional* ε-optimal strategies for all $\varepsilon > 0$.

Optimality Equations. We complete this section by introducing optimality equations for ERGs. For a game G and function $P : S \to \mathbb{R}_{\geq 0}$, we say that P is a solution of the optimality equations $\mathrm{Opt}^*(\mathsf{G})$, and write $P \models \mathrm{Opt}^*(\mathsf{G})$, if for any $s \in S$:

$$P(s) = \begin{cases} 0 & \text{if } s \in F \\ \inf\limits_{\alpha \in A_{\mathrm{Min}}(s)} \left\{ \sup\limits_{\beta \in A_{\mathrm{Max}}(s)} \left\{ \pi(s, \alpha, \beta) + \sum\limits_{s' \in S} p(s'|s, \alpha, \beta) \cdot P(s') \right\} \right\} & \text{if } s \notin F. \end{cases}$$

and P is a solution of the optimality equations $\mathrm{Opt}_*(\mathsf{G})$, and write $P \models \mathrm{Opt}_*(\mathsf{G})$, if for any $s \in S$:

$$P(s) = \begin{cases} 0 & \text{if } s \in F \\ \sup\limits_{\beta \in A_{\mathrm{Max}}(s)} \left\{ \inf\limits_{\alpha \in A_{\mathrm{Min}}(s)} \left\{ \pi(s, \alpha, \beta) + \sum\limits_{s' \in S} p(s'|s, \alpha, \beta) \cdot P(s') \right\} \right\} & \text{if } s \notin F. \end{cases}$$

The following result demonstrate the correspondence between these equations and the lower and upper values of the expected reachability game.

Proposition 2.1 *For any ERG* G *and bounded function* $P : S \to \mathbb{R}_{\geq 0}$:

- *if* $P \models \mathrm{Opt}^*(\mathsf{G})$*, then* $\mathrm{Val}^*(s) = P(s)$ *for all* $s \in S$ *and for any* $\varepsilon > 0$ *player Min has a* positional *strategy* μ_ε *such that* $\mathrm{Val}^{\mu_\varepsilon}(s) \leq P(s) + \varepsilon$ *for all* $s \in S$*;*
- *if* $P \models \mathrm{Opt}_*(\mathsf{G})$*, then* $\mathrm{Val}_*(s) = P(s)$ *for all* $s \in S$ *and for any* $\varepsilon > 0$ *player Max has a* positional *strategy* χ_ε *such that* $\mathrm{Val}_{\chi_\varepsilon}(s) \geq P(s) - \varepsilon$ *for all* $s \in S$*.*

If G is turn-based, then the equations $\mathrm{Opt}^*(\mathsf{G})$ and $\mathrm{Opt}_*(\mathsf{G})$ are the same and we write $\mathrm{Opt}(\mathsf{G})$ for these equations. The following is a direct consequence of Proposition 2.1.

Proposition 2.2 *If G is a turn-based, $P : S{\to}\mathbb{R}_{\geq 0}$ is a bounded and $P \models \mathrm{Opt}(\mathsf{G})$, then $\mathrm{Val}(s){=}P(s)$ for all $s \in S$ and for any $\varepsilon{>}0$ both players have ε-optimal strategies.*

3 Expected Reachability-Time Games

Expected reachability-time games (ERTGs) are played on the infinite graph of a probabilistic timed automaton where Min and Max choose their moves so that the expected time to reach a final state is minimised or maximised, respectively. Before defining ERTGs, we introduce the concept of clocks, constraints, regions, and zones.

Clocks. Let $\mathcal{C}$ be a finite set of *clocks*. A *clock valuation* on $\mathcal{C}$ is a function $\nu : \mathcal{C}{\to}\mathbb{R}_{\geq 0}$ and we write V for the set of clock valuations. Abusing notation, we also treat a valuation ν as a point in $\mathbb{R}^{|\mathcal{C}|}$. If $\nu \in V$ and $t \in \mathbb{R}_{\geq 0}$ then we write $\nu{+}t$ for the clock valuation defined by $(\nu{+}t)(c) = \nu(c){+}t$ for all $c \in \mathcal{C}$. For $C \subseteq \mathcal{C}$, we write $\nu[C{:=}0]$ for the valuation where $\nu[C{:=}0](c)$ equals 0 if $c \in C$ and $\nu(c)$ otherwise. For $X \subseteq V$, we write $\overline{X}$ for the smallest closed set in V containing X. Although clocks are usually allowed to take arbitrary non-negative values, w.l.o.g [29] we assume that there is an upper bound K such that for every clock $c \in \mathcal{C}$ we have that $\nu(c) \leq K$.

Clock constraints. A *clock constraint* over $\mathcal{C}$ is a conjunction of *simple constraints* of the form $c \bowtie i$ or $c{-}c' \bowtie i$, where $c, c' \in \mathcal{C}$, $i \in \mathbb{N}$, $i{\leq}K$, and $\bowtie \in \{<, >, =, \leq, \geq\}$. For $\nu \in V$, let $\mathrm{SCC}(\nu)$ be the finite set of simple constraints which hold in ν.

Clock regions. A *clock region* is a maximal set $\zeta{\subseteq}V$ such that $\mathrm{SCC}(\nu){=}\mathrm{SCC}(\nu')$ for all $\nu, \nu' \in \zeta$. We write $\mathcal{R}$ for the finite set of clock regions. Every clock region is an equivalence class of the indistinguishability-by-clock-constraints relation, and vice versa. We write $[\nu]$ for the clock region of ν and, if $\zeta{=}[\nu]$, write $\zeta[C{:=}0]$ for $[\nu[C{:=}0]]$.

Clock zones. A *clock zone* is a convex set of clock valuations, which is a union of a set of clock regions. We write $\mathcal{Z}$ for the set of clock zones. 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.

We now introduce ERTGs which extend classical timed automata [30] with discrete distributions and a partition of the actions between two players Min and Max.

Definition 3 (ERTG Syntax) *A (concurrent) expected reachability-time game (ERTG) is a tuple $\mathcal{T}{=}(L, L_F, \mathcal{C}, \mathit{Inv}, \mathit{Act}, \mathit{Act}_{\mathrm{Min}}, \mathit{Act}_{\mathrm{Max}}, E, \delta)$ where*

- *L is a finite set of* locations *including a set of* final locations *L_F;*
- *$\mathcal{C}$ is a finite set of* clocks*;*
- *$\mathit{Inv} : L \to \mathcal{Z}$ is an* invariant condition*;*
- *Act is a finite set of* actions *and $\{\mathit{Act}_{\mathrm{Min}}, \mathit{Act}_{\mathrm{Max}}\}$ is a partition of Act;*

- $E : L{\times}Act \to \mathcal{Z}$ *is an* action enabledness function*;*
- $\delta : L{\times}Act \to \mathcal{D}(2^{\mathcal{C}}{\times}L)$ *is a* probabilistic transition function*.*

When we consider an ERTG as an input of an algorithm, its size is understood as the sum of the sizes of encodings of L, $\mathcal{C}$, Inv, Act, E, and δ. As usual [31], we assume that probabilities are expressed as ratios of two natural numbers, each written in binary.

An ERTG is *turn-based* if for each location ℓ, only one player has enabled actions, i.e. $E(\ell, a){=}\emptyset$ for all $a \in Act_{\mathrm{Min}}$ or $a \in Act_{\mathrm{Max}}$. In this case, we write L_{Min} and L_{Max} for the set of locations where players Min and Max, respectively, have an enabled action. A *one-player ERTG* is a turn-based ERTG where one of the player does not control any location, i.e., either $L_{\mathrm{Min}}{=}\emptyset$ or $L_{\mathrm{Max}}{=}\emptyset$. A (non-probabilistic) *reachability-timed game* is an ERTG such that $\delta(\ell, a)$ is a point distribution for all $\ell \in L$ and $a \in Act$.

A *configuration* of an ERTG is a pair (ℓ, ν), where ℓ is a location and ν a clock valuation such that $\nu \in Inv(\ell)$. For any $t \in \mathbb{R}$, we let $(\ell, \nu){+}t$ equal the configuration $(\ell, \nu{+}t)$. In a configuration (ℓ, ν), a timed action (time-action pair) (t, a) is available if and only if the invariant condition $Inv(\ell)$ is continuously satisfied while t time units elapse, and a is enabled (i.e. the enabling condition $E(\ell, a)$ is satisfied) after t time units have elapsed. Furthermore, if the timed action (t, a) is performed, then the next configuration is determined by the probabilistic transition relation δ, i.e. with probability $\delta[\ell, a](C, \ell')$ the clocks in C are reset and we move to the location ℓ'.

An ERTG starts at some initial configuration and Min and Max construct an infinite play by repeatedly choosing available timed actions $(t_a, a) \in \mathbb{R}_{\geq 0}{\times}Act_{\mathrm{Min}}$ and $(t_b, b) \in \mathbb{R}_{\geq 0}{\times}Act_{\mathrm{Max}}$ proposing $\perp$ if no timed action is available. The player responsible for the move is Min if the time delay of Min's choice is less than that of Max's choice or Max chooses $\perp$, and otherwise Max is responsible. We assume the players cannot simultaneously choose $\perp$. We now present the formal semantics which is an ERG with potentially infinite number of states and actions. It is straightforward to show the semantics of a turn-based ERTG is a turn-based ERG.

Definition 4 (ERTG Semantics) *Let $\mathcal{T}$ be an ERTG. The semantics of $\mathcal{T}$ is given the ERG $[\![\mathcal{T}]\!]{=}(S, F, A_{\mathrm{Min}}, A_{\mathrm{Max}}, p_{\mathrm{Min}}, p_{\mathrm{Max}}, \pi_{\mathrm{Min}}, \pi_{\mathrm{Max}})$ where*

- *$S \subseteq L{\times}V$ is the (possibly uncountable) set of states such that $(\ell, \nu) \in S$ if and only if $\nu \in Inv(\ell)$ and $F = \{(\ell, \nu) \in S \mid \ell \in L_F\}$ is the set of final states;*
- *$A_{\mathrm{Min}} = (\mathbb{R}_{\geq 0}{\times}Act_{\mathrm{Min}}) \cup \{\perp\}$ and $A_{\mathrm{Max}} = (\mathbb{R}_{\geq 0}{\times}Act_{\mathrm{Max}}) \cup \{\perp\}$ are the sets of timed actions of players Min and Max;*
- *for $\star \in \{\mathrm{Min}, \mathrm{Max}\}$, $(\ell, \nu) \in S$ and $(t, a) \in A_\star$ the probabilistic transition function $p_\star$ is defined when $\nu{+}t' \in Inv(\ell)$ for all $t'{\leq}t$ and $\nu{+}t \in E(\ell, a)$ and for any (ℓ, ν'):*

$$p_\star((\ell, \nu), (t, a))((\ell', \nu')) = \textstyle\sum_{C \subseteq \mathcal{C} \wedge (\nu+t)[C:=0]=\nu'} \delta[\ell, a](C, \ell');$$

- *for $\star \in \{\mathrm{Min}, \mathrm{Max}\}$, $s \in S$ and $(t, a) \in A_{\mathrm{Min}}$ the reward function $\pi_\star$ is given by $\pi_\star(s, (t, a)){=}t$.*

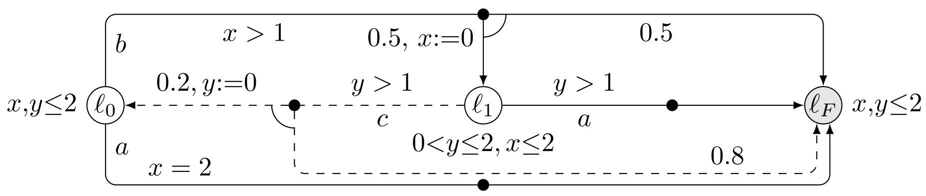

Figure 1: An expected reachability-time game.

The sum in the definitions of p_{Min} and p_{Max} is used to capture the fact that resetting different subsets of $\mathcal{C}$ may result in the same clock valuation (e.g. if all clocks are initially zero, then we end up with the same valuation, no matter which clocks we reset). Also, notice that the reward function of the ERG corresponds to the elapsed time of each move.

For any ERTG $\mathcal{T}$, to ensure Assumption 1 holds on the ERG $[\![\mathcal{T}]\!]$, we require only that the following weaker assumption holds on $[\![\mathcal{T}]\!]$.

Assumption 2 *For any strategy pair* $(\mu, \chi) \in \Sigma_{\mathrm{Min}}{\times}\Sigma_{\mathrm{Max}}$*, and state* $s \in S$ *we have that* $\lim_{n\to\infty} Prob_s^{\mu,\chi}(X_n \in F) = 1$.

Example 3.1 *Consider the ERTG in Figure 1; we use solid and dashed lines to indicate actions controlled by Min and Max respectively. The shaded circle denotes the final location. Considering location* ℓ_1*, the invariant condition is* $0{<}y{\leq}2{\wedge}x{\leq}2$*, actions* a *and* c *are enabled when* $y{>}1$ *and, if* a *is taken, we move to* ℓ_F*, while if* c *is taken, with probability* 0.2 *we move to* ℓ_0 *and reset* y*, and with probability* 0.8 *move to* ℓ_F.

Starting in the configuration[3] $(\ell_0, (0,0))$ *and supposing Min's strategy is to choose* $(1.1, b)$ *(i.e., wait* 1.1 *time units before performing action* b*) in location* ℓ_0 *and then choose* $(0.5, a)$ *in location* ℓ_1*, while Max's strategy in location* ℓ_1 *is to choose* $(0.2, c)$*. One possible play under this strategy pair is*

$$\langle(\ell_0,(0,0)),\ ((1.1,b),\bot),\ (\ell_1,(0,1.1)),\ ((0.5,a),\ (0.2,c)),\ (\ell_0,(0.2,0)),\ ((1.1,b),\bot),\ (\ell_F,(1.3,1.1))\rangle$$

which has probability $0.5{\cdot}0.2{\cdot}0.5 = 0.05$ *and time* $1.1{+}0.2{+}1.1 = 2.4$ *of reaching the final location. Using the optimality equations* $\mathrm{Opt}^*(\mathsf{G})$ *and* $\mathrm{Opt}_*(\mathsf{G})$*, we obtain upper and lower value in state* $(\ell_0, (0,0))$ *of* $\frac{10}{9}$ *and* 1*, respectively. For details of the equations see the appendix.*

Example 3.1 above demonstrates that in general expected reachability-time games are not determined. However, our results yield a novel proof of the following fundamental result for turn-based expected reachability-time games.

Theorem 3.2 *Turn-based ERTGs are positionally determined.*

Since the general ERTG are not determined, we study the following decision problem related to computing the upper-value of a configuration. All presented results also apply to the corresponding lower value problem, and the value problem, if the value exists.

[3]We suppose the first (second) coordinate in a clock valuation correspond to the clock x (y).

Definition 5 (ERTG Decision Problem) *The decision problem for an ERTG $\mathcal{T}$, a state s of $[\![\mathcal{T}]\!]$, and a bound $T \in \mathbb{Q}$ is to decide whether* $\mathrm{Val}^*(s) \leq T$.

We now present the second fundamental result of the paper.

Theorem 3.3 *The ERTG decision problem is in NEXPTIME $\cap$ co-NEXPTIME.*

From [16] we know that the ERTG problem is EXPTIME-complete even for one player ERTGs with two or more clocks. Hence the ERTG problem for general (two-player, concurrent) ERTG is at least EXPTIME-hard. Moreover, from the results of [17] and [32] it follows that ERTG problem is not NEXPTIME-hard, unless NP = co-NP.

4 Proofs of Theorems 3.2 and 3.3

This section is dedicated to the correctness of Theorems 3.2 and 3.3. We begin by defining *boundary region abstraction* (BRA) (an instance of an ERG) of an ERTG. In Section 4.2 we show that the solution of the optimality equations for a BRA always exists and is unique. While Section 4.3 demonstrates (Theorem 4.4) that the solution of the optimality equations of the BRA can be used to construct a solution of the optimality equations of the ERTG. Using these results we can then prove our main results.

Proof outline of Theorem 3.2. Using Theorem 4.4, a bounded solution of the equations for the upper and lower values of a ERTG always exists, and hence Proposition 2.1 implies both players have positional ε-optimal strategies. Since for turn-based ERTGs both equations are equivalent, from Proposition 2.2 positional determinacy of turn-based ERTGs follows.

Proof outline of Theorem 3.3. From Theorem 4.4 the upper value of a state of a ERTG can be derived from that of the boundary region abstraction. Since in the BRA the sub-graph of reachable states from any state is finite (Lemma 4.1) and its size is at most exponential in size of its ERTG, the upper value of a state in BRA can be computed by analysing an ERG of exponential size. The membership of the ERTG problem in NEXPTIME $\cap$ co-NEXPTIME then follows from the fact that a non-deterministic Turing machine needs to guess a (rational) solution of optimality equations only for exponentially many states, and it can verify in exponential time whether it is indeed a solution.

4.1 Boundary region abstraction

The region graph [30] is useful for solving time-abstract optimisation problems on timed automata. The region graph, however, is not suitable for solving timed optimisation problems and games on timed automata as it abstracts away the timing information. The corner-point abstraction [33] is an abstraction of timed automata which retains some timing information, but it is not convenient for the dynamic programming based proof techniques used in this paper. The boundary region abstraction (BRA) [13], a generalisation of the corner-point abstraction, is more suitable for such proof techniques. More precisely, we need to prove certain properties of values in ERTG, which we can do only when reasoning about all states of the ERTG. In the corner point abstraction we cannot do this since it represents only states corresponding to corner points of regions. Here, we generalise the BRA of [13] to handle ERTG.

Timed Successor Regions. Recall that $\mathcal{R}$ is the set of clock regions. For $\zeta, \zeta' \in \mathcal{R}$, we say that ζ' is in the future of ζ, denoted $\zeta \xrightarrow{*} \zeta'$, if there exist $\nu \in \zeta$, $\nu' \in \zeta'$ and $t \in \mathbb{R}_{\geq 0}$ such that $\nu' = \nu{+}t$ and say ζ' is the *time successor* of ζ if $\nu{+}t' \in \zeta \cup \zeta'$ for all $t' \leq t$ and write $\zeta \rightarrow \zeta'$, or equivalently $\zeta' \leftarrow \zeta$, to denote this fact. For regions $\zeta, \zeta' \in \mathcal{R}$ such that $\zeta \xrightarrow{*} \zeta'$ we write $[\zeta, \zeta']$ for the zone $\cup\{\zeta'' \,|\, \zeta \xrightarrow{*} \zeta'' \wedge \zeta'' \xrightarrow{*} \zeta'\}$.

Thin and Thick Regions. We say that a region ζ is *thin* if $[\nu]{\neq}[\nu{+}\varepsilon]$ for every $\nu \in \zeta$ and $\varepsilon{>}0$ and *thick* otherwise. We write $\mathcal{R}_{\mathrm{Thin}}$ and $\mathcal{R}_{\mathrm{Thick}}$ for the sets of thin and thick regions, respectively. Observe that if $\zeta \in \mathcal{R}_{\mathrm{Thick}}$ then, for any $\nu \in \zeta$, there exists $\varepsilon{>}0$, such that $[\nu]{=}[\nu{+}\varepsilon]$ and the time successor of a thin region is thick, and vice versa.

Intuition for the Boundary Region Graph. Recall K is an upper bound on clock values and let $[\![K]\!]_{\mathbb{N}} = \{0, 1, \ldots, K\}$. For any $\nu \in V$, $b \in [\![K]\!]_{\mathbb{N}}$ and $c \in \mathcal{C}$ we define $\mathsf{time}(\nu, (b, c)) \stackrel{\text{def}}{=} b{-}\nu(c)$ if $\nu(c){\leq}b$, and $\mathsf{time}(\nu, (b, c)) \stackrel{\text{def}}{=} 0$ if $\nu(c){>}b$. Intuitively, $\mathsf{time}(\nu, (b, c))$ returns the amount of time that must elapse in ν before the clock c reaches the integer value b. Observe that, for any $\zeta' \in \mathcal{R}_{\mathrm{Thin}}$, there exists $b \in [\![K]\!]_{\mathbb{N}}$ and $c \in \mathcal{C}$, such that $\nu \in \zeta$ implies $(\nu{+}(b{-}\nu(c)) \in \zeta'$ for all $\zeta \in \mathcal{R}$ in the past of ζ' and write $\zeta \rightarrow_{b,c} \zeta'$. The boundary region abstraction is motivated by the following. Consider $a \in Act$, (ℓ, ν) and $\zeta \xrightarrow{*} \zeta'$ such that $\nu \in \zeta$, $[\zeta, \zeta'] \subseteq Inv(\ell)$ and $\nu' \in E(\ell, a)$.

- If $\zeta' \in \mathcal{R}_{\mathrm{Thick}}$, then there are infinitely many $t \in \mathbb{R}_{\geq 0}$ such that $\nu{+}t \in \zeta'$. 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 ζ' is a thick region, the set $\{t \in \mathbb{R}_{\geq 0} \,|\, \nu{+}t \in \zeta'\}$ is an open interval, and hence does not contain its boundary values. Observe that the infimum equals $b_{\mathsf{inf}} - \nu(c_{\mathsf{inf}})$ where $\zeta \rightarrow_{b_{\mathsf{inf}}, c_{\mathsf{inf}}} \zeta_{\mathsf{inf}} \rightarrow \zeta'$ and the supremum equals $b_{\mathsf{sup}}{-}\nu(c_{\mathsf{sup}})$ where $\zeta \rightarrow_{b_{\mathsf{sup}}, c_{\mathsf{sup}}} \zeta_{\mathsf{sup}} \leftarrow \zeta'$. In the boundary region abstraction we include these 'best' timed actions through the actions $(b_{\mathsf{inf}}, c_{\mathsf{inf}}, a, \zeta')$ and $(b_{\mathsf{sup}}, c_{\mathsf{sup}}, a, \zeta')$.

- If $\zeta' \in \mathcal{R}_{\mathrm{Thin}}$, then there exists a unique $t \in \mathbb{R}_{\geq 0}$ such that $\nu{+}t \in \zeta'$. Moreover since ζ' is a thin region, there exists a clock $c \in C$ and a number $b \in \mathbb{N}$ such that $\zeta \rightarrow_{b,c} \zeta'$ and $t = b{-}\nu(c)$. In the boundary region abstraction we summarise this 'best' timed action from region ζ via region ζ' through the action (b, c, a, ζ').

Based on this intuition above the boundary region abstraction (BRA) is defined as follows.

Definition 6 *For an ERTG $\mathcal{T} = (L, L_F, \mathcal{C}, Inv, Act, Act_{\mathrm{Min}}, Act_{\mathrm{Max}}, E, \delta)$ the BRA of $\mathcal{T}$ is given by the ERG $\widehat{\mathcal{T}} = (\widehat{S}, \widehat{F}, \widehat{A}_{\mathrm{Min}}, \widehat{A}_{\mathrm{Max}}, \widehat{p}_{\mathrm{Min}}, \widehat{p}_{\mathrm{Max}}, \widehat{\pi}_{\mathrm{Min}}, \widehat{\pi}_{\mathrm{Max}})$ where*

- *$\widehat{S} \subseteq L \times V \times \mathcal{R}$ is the (possibly uncountable) set of states such that $(\ell, \nu, \zeta) \in \widehat{S}$ if and only if $\zeta \in \mathcal{R}$, $\zeta \subseteq Inv(\ell)$, and $\nu \in \overline{\zeta}$;*

- *$\widehat{F} = \{(\ell, \nu, \zeta) \in \widehat{S} \,|\, \ell \in L_F\}$ is the set of final states;*

- *$\widehat{A}_{\mathrm{Min}} \subseteq ([\![K]\!]_{\mathbb{N}} \times \mathcal{C} \times Act_{\mathrm{Min}} \times \mathcal{R}) \cup \{\bot\}$ is the set of actions of player* Min*;*

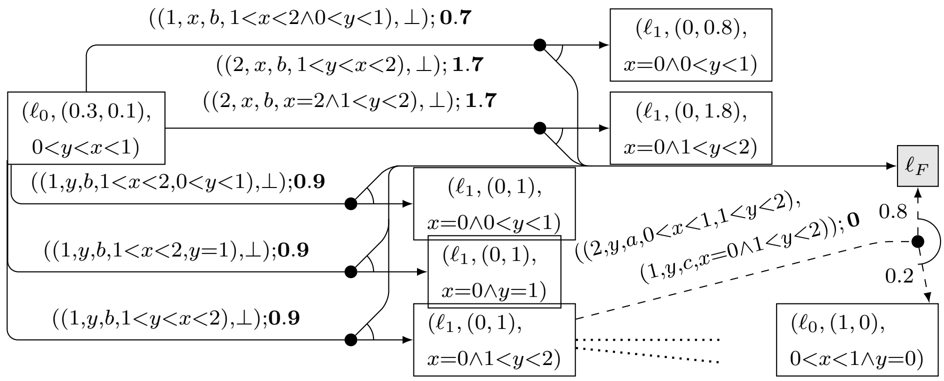

Figure 2: Sub-graph of the boundary region abstraction for the ERTG of Figure 1.

- $\widehat{A}_{\mathrm{Max}} \subseteq ([\![K]\!]_{\mathbb{N}} \times \mathcal{C} \times Act_{\mathrm{Max}} \times \mathcal{R}) \cup \{\perp\}$ *is the set of actions of player* Max*;*
- *for* $\star \in \{\mathrm{Min}, \mathrm{Max}\}$*,* $s = (\ell, \nu, \zeta) \in \widehat{S}$ *and* $\alpha = (b_\alpha, c_\alpha, a_\alpha, \zeta_\alpha) \in \widehat{A}_\star$ *the probabilistic transition function* $p_\star$ *is defined if* $[\zeta, \zeta_\alpha] \subseteq Inv(\ell)$ *and* $\zeta_\alpha \subseteq E(\ell, a_\alpha)$ *and for any* $(\ell', \nu', \zeta') \in \widehat{S}$*:*

$$\widehat{p_\star}(s, \alpha)((\ell', \nu', \zeta')) = \textstyle\sum_{C \subseteq \mathcal{C} \wedge \nu_\alpha[C:=0]=\nu' \wedge \zeta_\alpha[C:=0]=\zeta'} \delta[\ell, a_\alpha](C, \ell')$$

where $\nu_\alpha = \nu + \mathsf{time}(\nu, (b_\alpha, c_\alpha))$ *and one of the following conditions holds:*

  - $\zeta \rightarrow_{b_\alpha, c_\alpha} \zeta_\alpha$*,*
  - $\zeta \rightarrow_{b_\alpha, c_\alpha} \zeta_{\mathsf{inf}} \rightarrow \zeta_\alpha$ *for some* $\zeta_{\mathsf{inf}} \in \mathcal{R}$
  - $\zeta \rightarrow_{b_\alpha, c_\alpha} \zeta_{\mathsf{sup}} \leftarrow \zeta_\alpha$ *for some* $\zeta_{\mathsf{sup}} \in \mathcal{R}$*;*

- *for* $\star \in \{\mathrm{Min}, \mathrm{Max}\}$*,* $(\ell, \nu, \zeta) \in \widehat{S}$ *and* $(b_\alpha, c_\alpha, a_\alpha, \zeta_\alpha) \in \widehat{A}_\star$ *the reward function* $\widehat{\pi}_\star$ *is given by* $\widehat{\pi}_\star((\ell, \nu, \zeta), (b_\alpha, c_\alpha, a_\alpha, \zeta_\alpha)) = b_\alpha - \nu(c_\alpha)$*.*

Although the boundary region abstraction is not a finite ERG, for a fixed initial state we can restrict attention to a finite ERG, thanks to the following result of [15, 16].

Lemma 4.1 *For any state of a boundary region abstraction, its reachable sub-graph is finite and is constructible in time exponential in the size of corresponding ERTG.*

Example 4.2 *Sub-graph of BRA reachable from* $(\ell_0, (0.3, 0.1), 0<y<x<1)$ *for the ERTG of Figure 1 is shown in Figure 2. Edges are labelled* (b, c, a, ζ) *whose intuitive meaning is to wait until clock* c *attains the value* b *and then fire action* a*. The rewards of edges (indicated in bold) correspond to the time delay before the action is fired. Figure 2 includes the actions available in the initial state and one of action pairs available in* $(\ell_1, (0, 1), x=0\wedge 1<y<2)$*. To simplify, the states with location* ℓ_F *are merged together into a single state labelled* ℓ_F *and probabilities that are equal to* 0.5 *are omitted.*

4.2 Solving optimality equations of a boundary region abstraction

Based on the optimality equations $\mathrm{Opt}^*(\widehat{\mathcal{T}})$ (see Section 2), we define the value improvement function $\mathbf{\Psi} : [\widehat{S} \to \mathbb{R}_{\geq 0}] \to [\widehat{S} \to \mathbb{R}_{\geq 0}]$ where for $f : \widehat{S} \to \mathbb{R}_{\geq 0}$ and $s \in \widehat{S}$:

$$\mathbf{\Psi}(f)(s) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } s \in \widehat{F} \\ \min_{\alpha \in \widehat{A}_{\mathrm{Min}}(s)} \left\{ \max_{\beta \in \widehat{A}_{\mathrm{Max}}(s)} \left\{ \widehat{\pi}(s, \alpha, \beta) + \sum_{s' \in \widehat{S}} \widehat{p}(s' | s, \alpha, \beta) \cdot f(s') \right\} \right\} & \text{if } s \notin \widehat{F} \end{cases}$$

By construction, a fixpoint of $\mathbf{\Psi}$ is a solution of $\mathrm{Opt}^*(\widehat{\mathcal{T}})$. The following demonstrates the existence and uniqueness of a fixpoint of $\mathbf{\Psi}$, and thus also the solution of $\mathrm{Opt}^*(\widehat{\mathcal{T}})$.

Proposition 4.3 *For any ERTG $\mathcal{T}$, the value improvement function $\mathbf{\Psi}$ on the BRA $\widehat{\mathcal{T}}$ has a unique fixed point and equals $\lim_{i \to \infty} \mathbf{\Psi}^i(f)$ for an arbitrary $f \in [\widehat{S} \to \mathbb{R}_{\geq 0}]$.*

Proof. From Assumption 2 and Lemma 4.1 follows that every $|L \times \mathcal{R}|$-th iterate of $\mathbf{\Psi}$ is a contraction. Hence using Banach fixed point theorem the result is immediate. □

4.3 Correctness of the boundary region abstraction reduction

In this section we show how the optimality equations for the boundary region abstraction can be used to solve optimality equations for its ERTG. Given an ERTG $\mathcal{T}$ and real-valued function $f : \widehat{S} \to \mathbb{R}$ on the states of the BRA $\widehat{\mathcal{T}}$, we define $\widetilde{f} : S \to \mathbb{R}$ by $\widetilde{f}(\ell, \nu) = f(\ell, \nu, [\nu])$ which gives a real-valued function on the states of $\mathcal{T}$. The following theorem states that, by applying this mapping, the solution of optimality equations for an ERTG is given by that of the optimality equations for its BRA.

Theorem 4.4 *Let $\mathcal{T}$ be an ERTG. If $P \models \mathrm{Opt}^*(\widehat{\mathcal{T}})$, then $\widetilde{P} \models \mathrm{Opt}^*(\mathcal{T})$.*

To prove Theorem 4.4 we first introduce quasi-simple functions and state some of their key properties. Next, we show that for any BRA $\widehat{\mathcal{T}}$ the solution of $\mathrm{Opt}^*(\widehat{\mathcal{T}})$ is regionally quasi-simple (a quasi-simple function for every region). Finally, we sketch how Theorem 4.4 follows from this fact (Proposition 4.7 and Theorem 4.9).

Quasi-simple functions. Asarin and Maler [2] introduced simple functions, a finitely representable class of functions, with the property that every decreasing sequence is finite. Given $X \subseteq V$, a function $F : X \to \mathbb{R}$ is *simple* if there exists $e \in \mathbb{N}$ and either $F(\nu) = e$ for all $\nu \in X$, or there exists $c \in C$ and $F(\nu) = e - \nu(c)$ for all $\nu \in X$. A function $F : \widehat{S} \to \mathbb{R}_{\geq 0}$ is regionally simple if $F(\ell, \cdot, \zeta)$ is simple for all $\ell \in L$ and $\zeta \in \mathcal{R}$.

For timed games, Asarin and Maler showed that if $f : \widehat{S} \to \mathbb{R}_{\geq 0}$ is regionally simple, then $\mathbf{\Psi}(f)$ is regionally simple. Therefore, since $\mathbf{\Psi}$ is a decreasing function, it follows that starting from a regionally simple function in finitely many iterations of $\mathbf{\Psi}$ a fixed point is reached and the upper value in reachability-time games is regionally simple. Also, using the properties of simple functions, [13] shows that for a non-probabilistic reachability-time game, the optimal strategies are *regionally positional*, i.e., in every state of a region the strategy chooses the same action. Unfortunately, in the case of ERTGs, $\mathbf{\Psi}(f)$ is not necessarily regionally simple for any given regionally simple function f. Moreover, as the example below demonstrates, neither is the value of the game necessarily regionally-simple nor optimal strategies regionally positional.

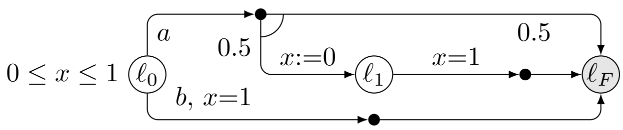

Figure 3: Example demonstrating optimal strategies are not regionally positional.

Example 4.5 *Consider the ERTG shown in Figure 3. Observe that for every state* (ℓ_0, ν) *in the region* $(\ell_0, 0{<}x{<}1)$*, the optimal expected time to reach* ℓ_F *equals*

$$\min\{\inf_{t \geq 0} \{t + 0.5{\cdot}1 + 0.5{\cdot}0\}, 1{-}\nu(x)\} = \min\{0.5, 1{-}\nu(x)\}.$$

Hence optimal expected reachability-time is not regionally simple. Moreover, the optimal strategy is not regionally positional, since if $\nu(x){\leq}0.5$*, then the optimal strategy is to fire* a *immediately, while otherwise the optimal strategy is to wait until* $\nu(x){=}1$ *and fire* b*.*

Due to these results it is not possible to work with simple function, and we define quasi-simple functions. Let $\trianglelefteq \subseteq V{\times}V$ be the partial order on clock valuations, where $\nu \trianglelefteq \nu'$ if and only if there exists a $t \in \mathbb{R}_{\geq 0}$ such that for each clock $c \in C$ either $\nu'(c){-}\nu(c) = t$ or $\nu(c){=}\nu'(c)$. For $x{=}(x_1, \ldots, x_n) \in \mathbb{R}^n$, we let $\|x\|_\infty = \max\{|x_i| \mid 1 \leq i \leq n\}$.

Definition 7 *Let* $X \subseteq V$*. A function* $F : X \to \mathbb{R}$ *is* quasi-simple *if for all* $\nu, \nu' \in X$*:*

- (Lipschitz Continuous) *there exists* $k{\geq}0$ *such that* $|F(\nu){-}F(\nu')| \leq k \cdot \|\nu{-}\nu'\|_\infty$*;*
- (Monotonically decreasing and nonexpansive w.r.t. $\trianglelefteq$) $\nu \trianglelefteq \nu'$ *implies* $F(\nu){\geq}F(\nu')$ *and* $F(\nu){-}F(\nu') \leq \|\nu{-}\nu'\|_\infty$.

For a convex set $X \subseteq V$ and continuous function $F : X \to \mathbb{R}$, we let $\overline{F} : \overline{X} \to \mathbb{R}$ denote the unique continuous function satisfying $\overline{F}(\nu){=}F(\nu)$ for all $\nu \in X$.

Theorem 4.6 (Properties of Quasi-simple Functions) *Let* $X \subseteq V$*.*

1. *Every simple function is also quasi-simple.*
2. *If* $F : X{\to}\mathbb{R}$ *is quasi-simple, then* $\overline{F} : \overline{X}{\to}\mathbb{R}$ *is quasi-simple.*
3. *If* $F, F' : X{\to}\mathbb{R}$ *are quasi-simple functions, then both the pointwise minimum and maximum of* F *and* F' *are quasi-simple.*
4. *The limit of a sequence of quasi-simple functions is quasi-simple.*

We say that $f : \widehat{S} \to \mathbb{R}_{\geq 0}$ is regionally quasi-simple if $f(\ell, \cdot, \zeta)$ is quasi-simple for all $\ell \in L$ and $\zeta \in \mathcal{R}$. Using Theorem 4.6 and Definition 6 we get the following result.

Proposition 4.7 *If* f *is regionally quasi-simple, then* $\mathbf{\Psi}(f)$ *is regionally quasi-simple.*

From Proposition 4.3 it follows that for an arbitrary function $f : \widehat{S} \to \mathbb{R}_{\geq 0}$ the limit of the sequence $\langle f, \mathbf{\Psi}(f), \mathbf{\Psi}^2(f), \ldots \rangle$ is the solution of $\mathrm{Opt}^*(\widehat{\mathcal{T}})$. From Proposition 4.7 it follows that, if we start from a regionally quasi-simple function f, then all the functions in the sequence $\langle f, \mathbf{\Psi}(f), \mathbf{\Psi}^2(f), \ldots \rangle$ are regionally quasi-simple. Since the limit of quasi-simple functions is quasi-simple, the following proposition is immediate.

Proposition 4.8 *For any ERTG $\mathcal{T}$, if $P \models \mathrm{Opt}^*(\widehat{\mathcal{T}})$, then P is regionally quasi-simple.*

The following result states that, from a regionally quasi-simple solution of the optimality equations for the boundary region abstraction, one can derive the solution of the optimality equations for the expected reachability time-game.

Theorem 4.9 *For any ERTG $\mathcal{T}$, if $P \models \mathrm{Opt}^*(\widehat{\mathcal{T}})$ and P is regionally quasi-simple, then $\widetilde{P} \models \mathrm{Opt}^*(\mathcal{T})$.*

The following observation is crucial for the proof of Theorem 4.9.

Lemma 4.10 *Let $s = (\ell, \nu) \in S$ and $\zeta \in \mathcal{R}$ such that $[\nu] \xrightarrow{*} \zeta$. If $P : \widehat{S} \to \mathbb{R}$ is regionally quasi-simple, then the functions:*

$$t \mapsto t + \textstyle\sum_{s' \in S} p(s'|s, (t, a), \bot) \cdot \widetilde{P}(s') \quad \textit{and} \quad t \mapsto t + \textstyle\sum_{s' \in S} p(s'|s, \bot, (t, b)) \cdot \widetilde{P}(s')$$

are continuous and nondecreasing on the interval $\{t \in \mathbb{R}_{\geq 0} \mid \nu + t \in \zeta\}$.

5 Conclusions

We introduced expected reachability-time games and showed that the natural decision problem is decidable and in NEXPTIME $\cap$ co-NEXPTIME. Furthermore, we proved that the turn-based subclass of these games is positionally determined. We believe that the main contribution of this paper is the concept of quasi-simple function that generalise simple functions to the context of probabilistic timed games. In fact, the techniques introduced in this paper extend to expected discounted-time games (EDTGs)[4] in a straightforward manner, since every expected discounted-time game can be reduced to an expected reachability-time game. Hence all the result presented for ERTGs are valid for EDTGs as well. Regarding other games on probabilistic timed automata, we conjecture that it is possible to reduce expected average-time games to mean payoff games on the boundary region abstraction. However, the techniques presented in this paper are insufficient to demonstrate such a reduction.

Although the computational complexity of solving games on timed automata is high, UPPAAL Tiga [5] is able to solve practical [6, 10] reachability and safety properties for timed games by using efficient symbolic zone-based algorithms. A natural future work is to investigate the possibility of extending similar algorithms for probabilistic timed games.

Acknowledgements. We would like to thank anonymous reviewers for their helpful comments and for finding an error in the submitted version. The authors are supported in part by EPSRC grants EP/D076625, EP/F001096, EP/D07956X/2, and by research center ITI, project No. 1M0545.

[4]In EDTG there is a fixed discount factor $\lambda \in [0,1)$, and when players follow strategies $\mu \in \Sigma_{\mathrm{Min}}$ and $\chi \in \Sigma_{\mathrm{Max}}$ the reward for state s is equal to $\mathbb{E}_s^{\mu,\chi} \left\{ \sum_{i=0}^{\infty} \lambda^i \cdot \pi(X_i, Y_{i+1}) \right\}$.

References

[1] Ramadge, P., Wonham, W.: The control of discrete event systems. In: Proc. IEEE. Volume 77(1). (1989)

[2] Asarin, E., Maler, O.: As soon as possible: Time optimal control for timed automata. In: Proc. HSCC'99. Volume 1569 of LNCS., Springer (1999)

[3] Alur, R., Bernadsky, M., Madhusudan, P.: Optimal reachability for weighted timed games. In: Proc. ICALP'04. Volume 3142 of LNCS., Springer (2004)

[4] Bouyer, P., Cassez, F., Fleury, E., Larsen, K.G.: Optimal strategies in priced timed game automata. In: FSTTCS'04. Volume 3328 of LNCS., Springer (2004)

[5] Behrmann, G., Cougnard, A., David, A., Fleury, E., Larsen, K., Lime, D.: UPPAAL-Tiga: Time for playing games! In: Proc. CAV'07. Volume 4590 of LNCS., Springer (2007)

[6] Cassez, F., Jessen, J., Larsen, K., Raskin, J., Reynier, P.: Automatic synthesis of robust and optimal controllers: an industrial case study. In: HSCC'09. Volume 5469 of LNCS. (2009)

[7] Kwiatkowska, M., Norman, G., Segala, R., Sproston, J.: Automatic verification of real-time systems with discrete probability distributions. Theoretical Computer Science **282** (2002)

[8] Jensen, H.: Model checking probabilistic real time systems. In: Proc. 7th Nordic Workshop on Programming Theory. Report 86:247–261, Chalmers University of Technology (1996)

[9] Beauquier, D.: Probabilistic timed automata. Theoretical Computer Science **292**(1) (2003)

[10] AlAttili, I., Houben, F., Igna, G., Michels, S., Zhu, F., Vaandrager, F.: Adaptive scheduling of data paths using Uppaal Tiga. CoRR **abs/0912.1897** (2009)

[11] de Alfaro, L., Faella, M., Henzinger, T.A., Majumdar, R., Stoelinga, M.: The element of surprise in timed games. In: Proc. CONCUR'03. Volume 2761 of LNCS., Springer (2003)

[12] Courcoubetis, C., Yannakakis, M.: Minimum and maximum delay problems in real-time systems. In: FMSD 1992. Volume 1., Dordrecht, Kluwer (1992)

[13] Jurdziński, M., Trivedi, A.: Reachability-time games on timed automata. In: Proc. ICALP'07. Volume 4596 of LNCS., Springer (2007)

[14] Jurdziński, M., Trivedi, A.: Average-time games. In: Proc. FSTTCS'08. Volume 2 of Leibniz International Proceedings in Informatics., Schloss Dagstuhl (2008)

[15] Jurdziński, M., Trivedi, A.: Concavely-priced timed automata. In: Proc. FORMATS'08. Volume 5215 of LNCS., Springer (2008)

[16] Jurdziński, M., Kwiatkowska, M., Norman, G., Trivedi, A.: Concavely-priced probabilistic timed automata. In: Proc. CONCUR'09. Volume 5710 of LNCS., Springer (2009)

[17] Chadha, R., Legay, A., Prabhakar, P., Viswanathan, M.: Complexity bounds for the verification of real-time software. In: Proc. VMCAI'10. Volume 5944 of LNCS. (2010)

[18] Hoffmann, G., Wong-Toi, H.: The input-output control of real-time discrete event systems. In: Proc. RTS'92, IEEE Press (1992)

[19] Asarin, E., Maler, O., Pnueli, A.: Symbolic controller synthesis for discrete and timed systems. In: Hybrid Systems II. Volume 999 of LNCS., Springer (1995)

[20] Brihaye, T., Henzinger, T.A., Prabhu, V.S., Raskin, J.: Minimum-time reachability in timed games. In: Proc. ICALP'07. Volume 4596 of LNCS., Springer (2007)

[21] Brihaye, T., Bruyère, V., Raskin, J.: On optimal timed strategies. In: Proc. FORMATS'05. Volume 3829 of LNCS., Springer (2005)

[22] Bouyer, P., Brihaye, T., Markey, N.: Improved undecidability results on weighted timed automata. Information Processing Letters **98** (2006)

[23] Bouyer, P., Forejt, V.: Reachability in stochastic timed games. In: Proc. ICALP'09. Volume 5556 of LNCS., Springer (2009)

[24] Kwiatkowska, M., Norman, G., Parker, D., Sproston, J.: Performance analysis of probabilistic timed automata using digital clocks. FMSD **29** (2006)

[25] Berendsen, J., Chen, T., Jansen, D.: Undecidability of cost-bounded reachability in priced probabilistic timed automata. In: Proc. TAMC'09. Volume 5532 of LNCS., Springer (2009)

[26] Forejt, V., Kwiatkowska, M., Norman, G., Trivedi, A.: Expected reachability-time games. In Chatterjee, K., Henzinger, T., eds.: Proc. 8th International Conference on Formal Modelling and Analysis of Timed Systems (FORMATS'10). LNCS, Springer (2010) To appear.

[27] Neyman, A., Sorin, S., eds.: Stochastic Games and Applications. Volume 570 of NATO Science Series C. Kluwer Academic Publishers (2004)

[28] Filar, J., Vrieze, K.: Competitive Markov Decision Processes. Springer (1997)

[29] Bouyer, P., Brihaye, T., Bruyère, V., Raskin, J.: On the optimal reachability problem on weighted timed automata. FMSD **31**(2) (2007)

[30] Alur, R., Dill, D.: A theory of timed automata. Theoretical Computer Science **126** (1994)

[31] Jurdziński, M., Sproston, J., Laroussinie, F.: Model checking probabilistic timed automata with one or two clocks. Logical Methods in Computer Science **4**(3) (2008)

[32] Condon, A.: The complexity of stochastic games. Information and Computation **96** (1992)

[33] Bouyer, P., Brinksma, E., Larsen, K.G.: Staying alive as cheaply as possible. In: Proc. HSCC'04. Volume 2993 of LNCS., Springer (2004)

A Proof of Proposition 2.1

For the remainder of this section we fix any ERG G and we slightly abuse notation and write a positional strategy of player Min as a function $\mu : S \to A_{\mathrm{Min}}$. In this section we show that a bounded solution to $\mathrm{Opt}^*(\mathsf{G})$ equals the upper value of the ERG G and that the existence of a solution of $\mathrm{Opt}^*(\mathsf{G})$ implies existence of a positional optimal strategy for player Min. These results are characterised by the following two lemmas.

Lemma A.1 *If $P : S \to \mathbb{R}_{\geq 0}$ is a bounded function such that $P \models \mathrm{Opt}^*(\mathsf{G})$, then for every $\varepsilon{>}0$, there exists a positional strategy $\mu_\varepsilon : S \to A_{\mathrm{Min}}$ for player Min, such that $P(s) \geq \sup_{\chi \in \Sigma_{\mathrm{Max}}} \mathrm{EReach}^{\mu_\varepsilon,\chi}(s) - \varepsilon$ for all $s \in S$.*

Lemma A.2 *If $P : S \to \mathbb{R}_{\geq 0}$ is a bounded function such that $P \models \mathrm{Opt}^*(\mathsf{G})$, then $P(s) \leq \sup_{\chi \in \Sigma_{\mathrm{Max}}} \mathrm{EReach}^{\mu,\chi}(s)$ for all strategies μ of player Min and $s \in S$.*

The proof corresponding to lower value optimality equations $\mathrm{Opt}_*(\mathsf{G})$ is analogous and hence omitted.

of Lemma A.1. Let $P : S \to \mathbb{R}_{\geq 0}$ is a bounded function such that $P \models \mathrm{Opt}^*(\mathsf{G})$. We first show that for every $\varepsilon{>}0$, there exists a positional strategy $\mu_\varepsilon : S \to A_{\mathrm{Min}}$ for player Min, such that $\mathrm{EReach}^{\mu_\varepsilon,\chi}(s) \leq P(s){+}\varepsilon$ for all strategies χ of player Max. Consider any $\varepsilon'{>}0$, and let μ_ε be a positional strategy such that for any state $s \in S$:

$$\sup_{b \in A_{\mathrm{Max}}(s)} \left\{ \pi(s, \mu_\varepsilon(s), b) + \sum_{s' \in S} p(s'|s, \mu_\varepsilon(s), b) \cdot P(s') \right\} \leq P(s) + \varepsilon' \,.$$

Observe that, since $P \models \mathrm{Opt}^*(\mathsf{G})$, such an action $\mu_\varepsilon(s)$ exists for all $s \in S$ and $\varepsilon'{>}0$. Now for any strategy χ of player Max, it follows by induction that for any $n \geq 1$:

$$\begin{aligned} P(s) \;\; \geq \;\; & \mathbb{E}_s^{\mu_\varepsilon,\chi} \left\{ \textstyle\sum_{i=0}^{\min\{\{n-1\} \cup \{k-1 \,|\, X_k \in F\}\}} \pi(X_i, Y_{i+1}) \right\} \\ & + \textstyle\sum_{s' \in S \setminus F} \mathit{Prob}_s^{\mu_\varepsilon,\chi}(X_n{=}s') \cdot P(s') - \sum_{i=1}^{n} \mathit{Prob}_s^{\mu_\varepsilon,\chi}(X_i \notin F) {\cdot} \varepsilon' . \end{aligned} \tag{1}$$

Next, using Assumption 1 and standard results of probability theory, we have:

- there exists $r \in \mathbb{R}_{\geq 0}$ such that $\lim_{n\to\infty} \sum_{i=1}^{n} \mathit{Prob}_s^{\mu_\varepsilon,\chi}(X_i \notin F) \leq r$;
- $\lim_{n\to\infty} \sum_{s' \in S \setminus F} \mathit{Prob}_s^{\mu_\varepsilon,\chi}(X_n{=}s') = 0$.

Therefore, taking the limit of (1) as n tends to infinity:

$$P(s) \geq \mathbb{E}_s^{\mu_\varepsilon,\chi} \left\{ \textstyle\sum_{i=0}^{\min\{k-1 \,|\, X_k \in F\}} \pi(X_i, Y_{i+1}) \right\} - \varepsilon' {\cdot} r \,.$$

Finally, if we consider any $\varepsilon{>}0$, then setting $\varepsilon'{=}\varepsilon/r$ it follows that $P(s) \geq \mathrm{EReach}^{\mu_\varepsilon,\chi}(s) - \varepsilon$ and, since the strategy $\chi \in \Sigma_{\mathrm{Max}}$ was arbitrary, we have:

$$P(s) \geq \sup\nolimits_{\chi \in \Sigma_{\mathrm{Max}}} \mathrm{EReach}^{\mu_\varepsilon,\chi}(s) - \varepsilon$$

as required. □

of **Lemma A.2.** Let $P : S \to \mathbb{R}_{\geq 0}$ is a bounded function such that $P \models \mathrm{Opt}^*(\mathsf{G})$. For simplicity let us first consider the case when G is a finite ERG. For each state $s \in S$ and $a \in A_{\mathrm{Min}}(s)$ we let $b_{s,a} \in A_{\mathrm{Max}}(s)$ be a player Max action such that

$$P(s) \leq \pi(s, a, b_{s,a}) + \sum_{s' \in S} p(s'|s, a, b_{s,a}) \cdot P(s') \,.$$

The existence of such an action follows from the fact that $P \models \mathrm{Opt}^*(\mathsf{G})$ and G is a finite ERG. Let $\chi^* : S \to A_{\mathrm{Max}}$ be a positional strategy of player Max such that the following conditions are satisfied for any $s \in S$.

- If $\pi_{\mathrm{Max}}(s, b_{s,a}) > \pi_{\mathrm{Min}}(s, a)$ for all $a \in A_{\mathrm{Min}}(s)$, then $\chi^*(s) = b_{s,a^*}$ for some $a^* \in A_{\mathrm{Min}}(s)$ such that:

$$\pi_{\mathrm{Max}}(s, b_{s,a^*}) = \max_{a \in A_{\mathrm{Min}}(s)} \pi_{\mathrm{Max}}(s, b_{s,a}) \,.$$

- If there exists $a \in A_{\mathrm{Min}}(s)$ such that $\pi_{\mathrm{Max}}(s, b_{s,a}) \leq \pi_{\mathrm{Min}}(s, a)$, then $\chi^*(s) = b_{s,a^*}$ for some $a^* \in A_{\mathrm{Min}}(s)$ such that:

$$\pi_{\mathrm{Max}}(s, b_{s,a^*}) = \min_{a \in A_{\mathrm{Min}}(s) \wedge \pi_{\mathrm{Max}}(s, b_{s,a}) \leq \pi_{\mathrm{Min}}(s,a)} \pi_{\mathrm{Max}}(s, b_{s,a}) \,.$$

Using this strategy χ^*, we will next show that for any state $s \in S$ and $a \in A_{\mathrm{Min}}(s)$:

$$P(s) \leq \pi(s, a, \chi^*(s)) + \sum_{s' \in S} p(s'|s, a, \chi^*(s)) \cdot P(s') \,. \tag{2}$$

Therefore, consider any $s \in S$ and $a \in A_{\mathrm{Min}}$ and suppose that $a^* \in A_{\mathrm{Min}}(s)$ is such that $\chi^*(s) = b_{s,a^*}$, we have the following two cases to consider.

- If $\pi_{\mathrm{Max}}(s, \chi^*(s)) \leq \pi_{\mathrm{Min}}(s, a)$, then we show via a contradiction that $\pi_{\mathrm{Max}}(s, \chi^*(s)) \leq \pi_{\mathrm{Min}}(s, a^*)$. Suppose $\pi_{\mathrm{Max}}(s, \chi^*(s)) > \pi_{\mathrm{Min}}(s, a^*)$, then by the conditions satisfied by χ^*, we have that $\pi_{\mathrm{Max}}(s, b_{s,a'}) > \pi_{\mathrm{Min}}(s, a')$ and $\pi_{\mathrm{Max}}(s, \chi^*(s)) > \pi_{\mathrm{Max}}(s, b_{s,a'})$ for all $a' \in A_{\mathrm{Min}}(s)$. Combining these facts it follows that $\pi_{\mathrm{Max}}(s, \chi^*(s)) > \pi_{\mathrm{Min}}(s, a)$ which contradicts the hypothesis, and hence $\pi_{\mathrm{Max}}(s, \chi^*(s)) \leq \pi_{\mathrm{Min}}(s, a^*)$. Now, since $b_{s,a^*} = \chi^*(s)$, by definition we have:

$$\begin{aligned} P(s) \leq\ & \pi(s, a^*, \chi^*(s)) + \sum_{s' \in S} p(s'|s, a^*, \chi^*(s)) \cdot P(s') \\ =\ & \pi_{\mathrm{Max}}(s, \chi^*(s)) + \sum_{s' \in S} p_{\mathrm{Max}}(s'|s, \chi^*(s)) \cdot P(s') \\ =\ & \pi(s, a, \chi^*(s)) + \sum_{s' \in S} p(s'|s, a, \chi^*(s)) \cdot P(s') \end{aligned}$$

where the first equality follows from the fact that $\pi_{\mathrm{Max}}(s, \chi^*(s)) \leq \pi_{\mathrm{Min}}(s, a^*)$ and the second from the hypothesis that $\pi_{\mathrm{Max}}(s, \chi^*(s)) \leq \pi_{\mathrm{Min}}(s, a)$.

- If $\pi_{\mathrm{Max}}(s,\chi^*(s)){>}\pi_{\mathrm{Min}}(s,a)$, then $\pi_{\mathrm{Max}}(s,b_{s,a}){>}\pi_{\mathrm{Min}}(s,a)$ since otherwise, by conditions satisfied by χ^*, it follows that $\pi_{\mathrm{Max}}(s,\chi^*(s)){\leq}\pi_{\mathrm{Max}}(s,b_{s,a})$ which leads to a contradiction. Now, by definition of $b_{s,a}$:

$$\begin{aligned} P(s) \leq\ & \pi(s,a,b_{s,a}) + \sum_{s'\in S} p(s'|s,a,b_{s,a}) \cdot P(s') \\ =\ & \pi_{\mathrm{Min}}(s,a) + \sum_{s'\in S} p_{\mathrm{Min}}(s'|s,a) \cdot P(s') \\ =\ & \pi(s,a,\chi^*(s)) + \sum_{s'\in S} p(s'|s,a,\chi^*(s)) \cdot P(s') \end{aligned}$$

  where the two equalities follow from the fact that $\pi_{\mathrm{Max}}(s,b_{s,a}){>}\pi_{\mathrm{Min}}(s,a)$ and the hypothesis that $\pi_{\mathrm{Max}}(s,\chi^*(s)){>}\pi_{\mathrm{Max}}(s,b_{s,a})$.

Since these are all the possible cases to consider, it follows that (2) holds for all $s \in S$ and $a \in A_{\mathrm{Min}}(s)$. Now let μ be an arbitrary (nonstationary and mixed) strategy. Consider a finite play $\rho \in \mathit{Play}_{\mathit{fin}}$ where player Min plays with strategy μ and player Max plays with strategy χ^*, since $\mu(\rho)$ is a distribution over actions we have:

$$\begin{aligned} P(s) &= \sum_{a\in A_{\mathrm{Min}}(s)} \mu(\rho)(a) \cdot P(s) \\ &\leq \sum_{a\in A_{\mathrm{Min}}(s)} \mu(\rho)(a) \cdot \Big(\pi(s,a,\chi^*(s)) + \sum_{s'\in S} p(s'|s,a,\chi^*(s)) \cdot P(s')\Big) \quad \text{by (2).} \end{aligned}$$

Thus, by induction it follows that for any $n \in \mathbb{N}$:

$$\begin{aligned} P(s) \ \leq\ & \mathbb{E}_s^{\mu,\chi^*}\left\{\textstyle\sum_{i=0}^{\min\{\{n-1\}\cup\{k-1\,|\,X_k\in F\}\}} \pi(X_i,Y_{i+1})\right\} \\ & + \textstyle\sum_{s'\in S\setminus F} \mathit{Prob}_s^{\mu,\chi^*}(X_n{=}s') \cdot P(s') \end{aligned}$$

and letting n tend to infinity we have $P(s) \leq \mathrm{EReach}^{\mu,\chi^*}(s)$ for all $s \in S$. Hence, for any $s \in S$ we have:

$$P(s) \leq \sup_{\chi\in\Sigma_{\mathrm{Max}}} \mathrm{EReach}^{\mu,\chi}(s),$$

and thus Lemma A.2 holds for finite ERGs.

Now let us consider the general case, i.e. for infinite ERGs and let $P_{max} = \sup_{s\in S} P(s)$ which, since P is bounded, is finite. Let us fix $\varepsilon{>}0$ and an arbitrary strategy μ of player Min. For each $s \in S$ and $a \in A_{\mathrm{Min}}(s)$, let $b^{\varepsilon}_{s,a} \in A_{\mathrm{Max}}(s)$ be a player Max action such that:

$$P(s) - \varepsilon \leq \pi(s,a,b^{\varepsilon}_{s,a}) + \sum_{s'\in S} p(s'|s,a,b^{\varepsilon}_{s,a}) \cdot P(s')\,.$$

The existence of such an action follows from the fact that $P \models \mathrm{Opt}^*(\mathsf{G})$ which implies that for any $s \in S$ and $a \in A_{\mathrm{Min}}(s)$:

$$P(s) \leq \sup_{b\in A_{\mathrm{Max}}(s)} \left\{\pi(s,a,b) + \sum_{s'\in S} p(s'|s,a,b) \cdot P(s')\right\}$$

We now construct a pure (nonstationary) strategy $\chi^{\mu}_{\varepsilon} : \mathit{Play}_{\mathit{fin}} \to A_{\mathrm{Max}}$ such that for any finite play $\rho \in \mathit{Play}_{\mathit{fin}}$ where $s{=}\mathit{last}(\rho)$ the following conditions are satisfied:

- If $\pi_{\mathrm{Min}}(s,a){<}\pi_{\mathrm{Max}}(s,b^{\varepsilon}_{s,a})$ for all $a \in A_{\mathrm{Min}}$, then $\chi^{\mu}_{\varepsilon}(\rho){=}b^{\varepsilon}_{s,a^*}$ for some $a^* \in A_{\mathrm{Min}}(s)$ such that:
$$\sum_{\substack{a \in A_{\mathrm{Min}}(s) \wedge \\ \pi_{\mathrm{Min}}(s,a) < \pi_{\mathrm{Max}}(s,b^{\varepsilon}_{s,a^*})}} \mu(\rho)(a) \geq 1-\varepsilon\,.$$
In addition, when this condition holds we write $A^{\chi^{\mu}_{\varepsilon}}(\rho)$ for the set of actions
$$\{a \in A_{\mathrm{Min}}(s) \mid \pi_{\mathrm{Min}}(s,a){\geq}\pi_{\mathrm{Max}}(s,\chi^{\mu}_{\varepsilon}(\rho))\}.$$
- If there exists $a \in A_{\mathrm{Min}}$ such that $\pi_{\mathrm{Min}}(s,a){\geq}\pi_{\mathrm{Max}}(s,b^{\varepsilon}_{s,a})$, then $\chi^{\mu}_{\varepsilon}(r){=}b^{\varepsilon}_{s,a^*}$ for some $a^* \in A_{\mathrm{Min}}$ such that $\pi_{\mathrm{Min}}(s,a^*){\geq}\pi_{\mathrm{Max}}(s,b^{\varepsilon}_{s,a^*})$ and
$$\sum_{\substack{a \in A_{\mathrm{Min}}(s) \wedge \\ (\pi_{\mathrm{Min}}(s,a) \geq \pi_{\mathrm{Min}}(s,a^*) \vee \pi_{\mathrm{Min}}(s,a) < \pi_{\mathrm{Max}}(s,b^{\varepsilon}_{s,a}))}} \mu(\rho)(a) \geq 1-\varepsilon\,.$$
If this condition holds, then let $A^{\chi^{\mu}_{\varepsilon}}(\rho)$ denote the set of player Min actions:
$$\{a \in A_{\mathrm{Min}}(s) \mid \pi_{\mathrm{Min}}(s,a){<}\pi_{\mathrm{Max}}(s,\chi^{\mu}_{\varepsilon}(\rho)) \wedge \pi_{\mathrm{Min}}(s,a){\geq}\pi_{\mathrm{Max}}(s,b^{\varepsilon}_{s,a})\}\,.$$

Observe that by construction in both cases we have:
$$\sum_{a \in A^{\chi^{\mu}_{\varepsilon}}(\rho)} \mu(\rho)(a) \leq \varepsilon. \tag{3}$$

Similarly, to the finite case, we will now show that for any finite path ρ where $s{=}\mathit{last}(\rho)$ and $a \in A_{\mathrm{Min}}(s) \backslash A^{\chi^{\mu}_{\varepsilon}}(\rho)$:
$$P(s) - \varepsilon \leq \pi(s,a,\chi^{\mu}_{\varepsilon}(\rho)) + \sum_{s' \in S} p(s'|s,a,\chi^{\mu}_{\varepsilon}(\rho)) \cdot P(s')\,. \tag{4}$$

Now supposing that $a^* \in A_{\mathrm{Min}}(s)$ is such that $\chi^{\mu}_{\varepsilon}(\rho){=}b^{\varepsilon}_{s,a^*}$, we have the following two cases to consider.

- If $\pi_{\mathrm{Max}}(s,\chi^{\mu}_{\varepsilon}(s)){\leq}\pi_{\mathrm{Min}}(s,a)$, then we show via contradiction that $\pi_{\mathrm{Max}}(s,\chi^{\mu}_{\varepsilon}(s)) \leq \pi_{\mathrm{Min}}(s,a^*)$. Suppose that $\pi_{\mathrm{Max}}(s,\chi^{\mu}_{\varepsilon}(s)){>}\pi_{\mathrm{Min}}(s,a^*)$, then the conditions required of χ^{μ}_{ε} imply that $\pi_{\mathrm{Max}}(s,b^{\varepsilon}_{s,a'}){>}\pi_{\mathrm{Min}}(s,a')$ for all $a' \in A_{\mathrm{Min}}(s)$ and since $a \not\in A^{\chi^{\mu}_{\varepsilon}}(\rho)$ if follows that $\pi_{\mathrm{Max}}(s,\chi^{\mu}_{\varepsilon}(s)){>}\pi_{\mathrm{Min}}(s,a)$ which contradicts the hypothesis. Now, since $b^{\varepsilon}_{s,a^*}{=}\chi^{\mu}_{\varepsilon}(s)$, we have:
$$\begin{aligned} P(s) - \varepsilon \leq\ & \pi(s,a^*,\chi^{\mu}_{\varepsilon}(s)) + \sum_{s' \in S} p(s'|s,a^*,\chi^{\mu}_{\varepsilon}(s)) \cdot P(s') \\ =\ & \pi_{\mathrm{Max}}(s,\chi^{\mu}_{\varepsilon}(s)) + \sum_{s' \in S} p_{\mathrm{Max}}(s'|s,\chi^{\mu}_{\varepsilon}(s)) \cdot P(s') \\ =\ & \pi(s,a,\chi^{\mu}_{\varepsilon}(s)) + \sum_{s' \in S} p(s'|s,a,\chi^{\mu}_{\varepsilon}(s)) \cdot P(s') \end{aligned}$$
where the first equality follows from the fact that $\pi_{\mathrm{Max}}(s,\chi^{\mu}_{\varepsilon}(s)){\leq}\pi_{\mathrm{Min}}(s,a^*)$ and the second from the hypothesis that $\pi_{\mathrm{Max}}(s,\chi^{\mu}_{\varepsilon}(s)){\leq}\pi_{\mathrm{Min}}(s,a)$.

- If $\pi_{\rm Max}(s,\chi^\mu_\varepsilon(s)){>}\pi_{\rm Min}(s,a)$, then $\pi_{\rm Max}(s,b^\varepsilon_{s,a}){>}\pi_{\rm Min}(s,a)$ since otherwise, by construction of χ^μ_ε, it follows that $\pi_{\rm Max}(s,\chi^\mu_\varepsilon(s)){\leq}\pi_{\rm Max}(s,b^\varepsilon_{s,a})$ which leads to a contradiction. Now, by definition of $b^\varepsilon_{s,a}$:

$$\begin{aligned}
P(s)-\varepsilon \leq\ & \pi(s,a,b^\varepsilon_{s,a}) + \sum_{s'\in S} p(s'|s,a,b^\varepsilon_{s,a})\cdot P(s') \\
=\ & \pi_{\rm Min}(s,a) + \sum_{s'\in S} p_{\rm Min}(s'|s,a)\cdot P(s') \\
=\ & \pi(s,a,\chi^\mu_\varepsilon(s)) + \sum_{s'\in S} p(s'|s,a,\chi^\mu_\varepsilon(s))\cdot P(s')
\end{aligned}$$

  where the two equalities follow from the fact that $\pi_{\rm Max}(s,b^\varepsilon_{s,a}){>}\pi_{\rm Min}(s,a)$ and the hypothesis that $\pi_{\rm Max}(s,\chi^\mu_\varepsilon(s)){>}\pi_{\rm Max}(s,b^\varepsilon_{s,a})$.

Since these are all the possible case, (4) holds for all finite paths ρ and $a\in A_{\rm Min}(\mathit{last}(\rho))\setminus A^{\chi^\mu_\varepsilon}(\rho)$. Now, for any finite play $\rho\in \mathit{Play}_{\mathit{fin}}$ where $s{=}\mathit{last}(\rho)$, since ρ is a probability distribution:

$$\begin{aligned}
P(s) &= \sum_{a\in A_{\rm Min}(s)} \mu(\rho)(a){\cdot}P(s) \\
&= \sum_{a\in A_{\rm Min}(s)} \mu(\rho)(a){\cdot}(P(s){-}\varepsilon)) + \varepsilon && \text{rearranging} \\
&= \sum_{a\in A_{\rm Min}(s)\setminus A^{\chi^\mu_\varepsilon}(\rho)} \mu(\rho)(a){\cdot}(P(s){-}\varepsilon)) \quad + \sum_{a\in A^{\chi^\mu_\varepsilon}(\rho)} \mu(\rho)(a){\cdot}(P(s){-}\varepsilon) + \varepsilon && \text{rearranging} \\
&\leq \sum_{a\in A_{\rm Min}(s)\setminus A^{\chi^\mu_\varepsilon}(\rho)} \mu(\rho)(a){\cdot}(P(s){-}\varepsilon)) + P_{max}{\cdot}\varepsilon + \varepsilon && \text{by (3)} \\
&= \sum_{a\in A_{\rm Min}(s)\setminus A^{\chi^\mu_\varepsilon}(\rho)} \mu(\rho)(a){\cdot}(P(s){-}\varepsilon)) + (P_{max}{+}1){\cdot}\varepsilon && \text{rearranging} \\
&\leq \sum_{a\in A_{\rm Min}(s)\setminus A^{\chi^\mu_\varepsilon}(\rho)} \mu(\rho)(a){\cdot}\left(\pi(s,a,\chi^\mu_\varepsilon(\rho)) + \sum_{s'\in S} p(s'|s,a,\chi^\mu_\varepsilon(\rho)){\cdot}P(s')\right) + (P_{max}{+}1){\cdot}\varepsilon \\
& && \text{by (4)} \\
&\leq \sum_{a\in A_{\rm Min}(s)} \mu(\rho)(a){\cdot}\left(\pi(s,a,\chi^\mu_\varepsilon(\rho)) + \sum_{s'\in S} p(s'|s,a,\chi^\mu_\varepsilon(\rho)){\cdot}P(s')\right) + (P_{max}{+}1){\cdot}\varepsilon
\end{aligned}$$

Using this result and induction, it follows that for any $n\in\mathbb{N}$:

$$\begin{aligned}
P(s) \quad \leq \quad & \mathbb{E}^{\mu,\chi^\mu_\varepsilon}_s\left\{\textstyle\sum_{i=0}^{\min\{\{n-1\}\cup\{k-1\,|\,X_k\in F\}\}}\pi(X_i,Y_{i+1})\right\} \\
& +\textstyle\sum_{s'\in S\setminus F}\mathit{Prob}^{\mu,\chi^\mu_\varepsilon}_s(X_n{=}s')\cdot P(s') \\
& +\textstyle\sum_{i=1}^n \mathit{Prob}^{\mu,\chi^\mu_\varepsilon}_s(X_i\not\in F){\cdot}(P_{max}{+}1){\cdot}\varepsilon
\end{aligned} \tag{5}$$

Next, using Assumption 1 and standard results of probability theory, it follows that:

- there exists $r\in\mathbb{R}_{\geq 0}$ such that $\lim_{n\rightarrow\infty}\sum_{i=1}^n \mathit{Prob}^{\mu,\chi^\mu_\varepsilon}_s(X_i\not\in F)\leq r$;
- $\lim_{n\rightarrow\infty}\sum_{s'\in S\setminus F}\mathit{Prob}^{\mu,\chi^\mu_\varepsilon}_s(X_n{=}s') = 0$.

Hence, taking the limit as n tends to infinity of (5) we obtain:

$$P(s) \leq \mathbb{E}_s^{\mu,\chi_\varepsilon^\mu}\left\{\sum_{i=0}^{\min\{k-1\,|\,X_k\in F\}}\pi(X_i,Y_{i+1})\right\} + r{\cdot}(P_{max}{+}1){\cdot}\varepsilon\,.$$

Therefore, for any $\varepsilon'{>}0$, it follows that $P(s) \leq \mathrm{EReach}^{\mu,\chi_\varepsilon^\mu}(s){+}\varepsilon'$ where $\varepsilon = \varepsilon'/((P_{max}{+}1){\cdot}r)$ which completes the proof in the general case. □

B Optimality equations for Example 3.1

Considering the optimality equations $\mathrm{Opt}^*(\mathsf{G})$, for all ν we have $P(\ell_F,\nu) = 0$ and after some simplifications we have:

$$\begin{aligned}
P(\ell_0,\nu) \;&=\; \min\Big\{(2{-}\nu(x)){+}P(\ell_F,\nu+(2{-}\nu(x))), \\
&\qquad \inf_{\substack{t\in\mathbb{R}_{\geq 0},\\ 1<\nu(x)+t\leq 2,\\ \nu(y)+t\leq 2}}\big\{t{+}0.5{\cdot}P(\ell_1,(0,\nu(y){+}t)){+}0.5{\cdot}P(\ell_F,\nu{+}t)\big\}\Big\} \\
P(\ell_1,\nu) \;&=\; \inf_{\substack{t\in\mathbb{R}_{\geq 0},\\ \nu(x)+t\leq 2,\\ 1<\nu(y)+t\leq 2}}\;\sup_{\substack{t'\in\mathbb{R}_{\geq 0},\\ \nu(x)+t'\leq 2,\\ 1<\nu(y)+t'\leq 2}}\Big(\big\{t{+}P(\ell_F,\nu{+}t)\mid t<t'\big\}\cup \\
&\qquad \big\{t'{+}0.2{\cdot}P(\ell_0,(\nu(x){+}t',0)){+}0.8{\cdot}P(\ell_F,\nu{+}t')\mid t\geq t'\big\}\Big)\,.
\end{aligned}$$

For $\mathrm{Opt}_*(\mathsf{G})$, the only change is the reversed order of $\sup$ and $\inf$ in the last equation.

C Proof of Lemma 4.1

For a real number $r \in \mathbb{R}$ we write $\lfloor r\rfloor$ for the floor of r, i.e., largest integer $n \in \mathbb{N}$ such that $n \leq r$; and we write $\wr r\int$ for the fractional part of r, i.e. $r - \lfloor r\rfloor$. For a clock valuation ν we define its fractional signature $\wr\nu\int$ to be the sequence $(f_0, f_1, \ldots, f_m)$, such that $f_0 = 0$, $f_i < f_j$ if $i < j$, for all $i, j \leq m$, and $f_1, f_2, \ldots, f_m$ are all the non-zero fractional parts of clock values in the clock valuation ν. In other words, for every $i \geq 1$, there is a clock c, such that $\wr\nu(c)\int = f_i$, and for every clock $c \in \mathcal{C}$, there is $i \leq m$, such that $\wr\nu(c)\int = f_i$. Let $(f_0, f_1, \ldots, f_m)$ be the fractional signature $\wr\nu\int$.

For a nonnegative integer $k \leq m$ we define the k-shift of a fractional signature $(f_0, f_1, \ldots, f_m)$ as the fractional signature $(f'_k, f'_{k+1}, \ldots, f'_m, f'_0, \ldots, f'_{k-1})$ such that for all non-negative integers $i \leq m$ we have $f'_i = \wr f_i + 1 - f_k\int$. We say that a fractional signature $(f'_0, f'_1, \ldots, f'_n)$ is a subsequence of another fractional signature $(f_0, f_1, \ldots, f_m)$ if $n \leq m$ and for all nonnegative integers $i < n$ we have $f'_i \leq f'_{i+1}$; and for every nonnegative integer $i \leq n$ there exists a nonnegative integer $j \leq m$ such that $f'_i = f_j$. Since taking boundary timed actions either result in a subsequence of fractional signature (potentially in the case of clock resets) or in a k-shift of the fractional signature, the following proposition is immediate.

Proposition C.1 *Let $\mathcal{T}$ be a ERTG with boundary region abstraction $\widehat{\mathcal{T}}$. For every BRA state $s = (\ell, \nu, \zeta) \in \widehat{S}$ and boundary timed actions $\alpha \in \widehat{A}_{\mathrm{Min}}(s)$ and $\beta \in \widehat{A}_{\mathrm{Max}}(s)$, we have that $\widehat{p}((\ell', \nu', \zeta')|s, \alpha, \beta) > 0$ implies that the fractional signature of ν' is k-shift of a subsequence of the fractional signature of ν.*

The proof of Lemma 4.1 follows from this proposition since the set of valuations having fractional signatures as k-shifts of subsequences of the fractional signatures of a fixed valuation is finite.

D Proof of Theorem 4.6

Recall the partial order $\trianglelefteq \subseteq V \times V$ on clock valuations, where $\nu \trianglelefteq \nu'$ if and only if there exists a $t \in \mathbb{R}_{\geq 0}$ such that for each clock $c \in C$ either $\nu'(c) - \nu(c) = t$ or $\nu(c) = \nu'(c)$. In this case we also write $(\nu' - \nu) = t$. Note that in this case $(\nu' - \nu) = \|\nu - \nu'\|_\infty$.

The proof of Theorem 4.6 follows from Lemmas D.1-D.4 below.

Lemma D.1 *Every simple function is also quasi-simple.*

Proof. Let $X \subseteq V$ be a subset of valuations and $f : X \to \mathbb{R}$ a simple function. If f is constant then the proposition trivially follows. Otherwise, there exists $b \in \mathbb{Z}$ and $c \in C$ such that $f(\nu) = b - \nu(c)$ for all $\nu \in X$. We need to show that f is Lipschitz continuous, and monotonically decreasing and nonexpansive w.r.t $\trianglelefteq$.

1. To prove that f is Lipschitz continuous, notice that $|f(\nu) - f(\nu')| = |b - \nu(c) - b + \nu'(c)| = |\nu'(c) - \nu(c)| \leq \|\nu - \nu'\|_\infty$.
2. For $\nu, \nu' \in X$ such that $\nu \trianglelefteq \nu'$, we have $f(\nu) = b - \nu(c) \geq b - \nu'(c) = f(\nu')$. From the first part of this proof, it trivially follows that $f(\nu) - f(\nu') \leq \nu - \nu'$. □

Lemma D.2 *If $f : X \to \mathbb{R}$ is quasi-simple, then $\overline{f} : \overline{X} \to \mathbb{R}$ is quasi-simple.*

Proof. Note that since every quasi-simple function $f : X \to \mathbb{R}$ is Lipschitz continuous, and hence Cauchy continuous, it can be uniquely extended to closure of its domain X. The properties of quasi-simple function are trivially met by such extensions. □

Lemma D.3 *If $f, f' : X \subseteq V \to \mathbb{R}$ are quasi-simple functions, then $\max(f, f')$ and $\min(f, f')$ are also quasi-simple.*

Proof. Let $f, f' : X \subseteq V \to \mathbb{R}$ be quasi-simple. We need to show that $\max(f, f')$ and $\min(f, f')$ are quasi-simple. Notice that $\max(f, f')$ and $\min(f, f')$ are Lipschitz continuous, as pointwise minimum and maximum of a finite set of Lipschitz continuous functions is Lipschitz continuous. It therefore remains to show that $\max(f, f')$ and $\min(f, f')$ are monotonically decreasing and nonexpansive w.r.t $\trianglelefteq$.

Consider any $\nu_1, \nu_2 \in X$ such that $\nu_1 \trianglelefteq \nu_2$. Since f and f' are quasi-simple, by definition f and f' are monotonically decreasing, and hence $f(\nu_1) \geq f(\nu_2)$ and $f'(\nu_1) \geq f'(\nu_2)$. Now since

$$\max(f, f')(\nu_1) = \max\left\{f(\nu_1), f'(\nu_1)\right\} \geq \max\left\{f(\nu_2), f'(\nu_2)\right\} = \max(f, f')(\nu_2),$$

it follows that $\max(f, f')$ is monotonically decreasing w.r.t $\trianglelefteq$. In an analogous manner we show that $\min(f, f')$ is monotonically decreasing w.r.t $\trianglelefteq$.

Again since f and f' are quasi-simple, we have that they are nonexpansive, i.e., $f(\nu_1) - f(\nu_2) \leq \nu_2 - \nu_1$ and $f'(\nu_1) - f'(\nu_2) \leq \nu_2 - \nu_1$. To show $\max(f, f')$ is nonexpansive, there are the following four cases to consider.

1. If $f(\nu_1) \geq f'(\nu_1)$ and $f(\nu_2) \geq f'(\nu_2)$, then $\max(f, f')(\nu_1) - \max(f, f')(\nu_2) = f(\nu_1) - f(\nu_2) \leq \nu_2 - \nu_1$.
2. If $f'(\nu_1) \geq f(\nu_1)$ and $f'(\nu_2) \geq f(\nu_2)$, then $\max(f, f')(\nu_1) - \max(f, f')(\nu_2) = f'(\nu_1) - f'(\nu_2) \leq \nu_2 - \nu_1$.
3. If $f(\nu_1) \geq f'(\nu_1)$ and $f'(\nu_2) \geq f(\nu_2)$, then $\max(f, f')(\nu_1) - \max(f, f')(\nu_2) = f(\nu_1) - f'(\nu_2) \leq f(\nu_1) - f(\nu_2) \leq \nu_2 - \nu_1$.
4. If $f'(\nu_1) \geq f(\nu_1)$ and $f(\nu_2) \geq f'(\nu_2)$, then $\max(f, f')(\nu_1) - \max(f, f')(\nu_2) = f'(\nu_1) - f(\nu_2) \leq f'(\nu_1) - f'(\nu_2) \leq \nu_2 - \nu_1$.

Since these are all the possible cases to consider, $\max(f, f')$ is nonexpansive w.r.t $\trianglelefteq$. Similarly show $\min(f, f')$ is nonexpansive completing the proof. $\square$

The following lemma follows directly from the fact that the limit of Lipschitz continuous functions is Lipschitz continuous, and the limit of monotonically decreasing and nonexpansive functions is monotonically decreasing and nonexpansive.

Lemma D.4 *The limit of a sequence of quasi-simple functions is quasi-simple.*

E Proof of Proposition 4.7

In this section we fix a ETRG $\mathcal{T}$ with boundary region abstraction $\widehat{\mathcal{T}}$. The following Lemma is central for the proof of Proposition 4.7.

Lemma E.1 *If $f : \widehat{S} \to \mathbb{R}_{\geq 0}$ is regionally quasi-simple, then for every location $\ell \in L$ and region $\zeta \in \mathcal{R}$ we have that the function*

$$\nu \mapsto \widehat{\pi}((\ell, \nu, \zeta), \alpha, \beta) + \textstyle\sum_{s' \in \widehat{S}} \widehat{p}(s'|(\ell, \nu, \zeta), \alpha, \beta) \cdot f(s')$$

is quasi-simple on its domain $\{\nu \in V \,|\, \nu \in \overline{\zeta}\}$ where (α, β) is in the set of boundary actions available to players in location ℓ and clock region ζ, i.e., $(\alpha, \beta) \in \widehat{A}(\ell, \zeta) \stackrel{\text{def}}{=} \{(a, b) \in \widehat{A}(\ell, \nu, \zeta) \,|\, \nu \in \zeta\}$.

Proof. Let f be regionally quasi-simple and consider any location $\ell \in L$, clock region $\zeta \in \mathcal{R}$ and pair of actions $(\alpha, \beta) \in \widehat{A}(\ell, \zeta)$. Moreover, we assume that $\mathsf{Winner}(\nu, \alpha, \beta)$=Min for all $\nu \in \zeta$ (the alternative cases are similar, and hence omitted). if $\alpha = (b, c, a, \zeta_a)$, then $\widehat{\pi}((\ell, \nu, \zeta), \alpha, \beta) = b - \nu(c)$ and we need to show that the function:

$$\nu \mapsto (b - \nu(c)) + \textstyle\sum_{s' \in \widehat{S}} \widehat{p}(s' \mid (\ell, \nu, \zeta), \alpha, \beta) \cdot f(s')$$

is quasi-simple on its domain. Denoting this function by $f^{\oplus}$, by definition of $\widehat{\mathcal{T}}$ we have:

$$\begin{aligned} f^{\oplus}(\nu) &= (b - \nu(c)) + \textstyle\sum_{(C,\ell') \in 2^{\mathcal{C}} \times L} \delta[\ell,a](C,\ell') \cdot f(\ell', \nu_{\alpha,C}, \zeta'[C{:=}0]) \\ &= (b - \nu(c)) + \textstyle\sum_{(C,\ell') \in 2^{\mathcal{C}} \times L} \delta[\ell,a](C,\ell') \cdot f(s_{\ell',\nu,\alpha,C}) \end{aligned} \tag{6}$$

where $\nu_{\alpha,C} = (\nu{+}(b{-}\nu(c)))[C{:=}0]$ and $s_{\ell',\nu,\alpha,C} = (\ell', \nu_{\alpha,C}, \zeta'[C{:=}0])$. We now demonstrate that $f^{\oplus}$ is Lipschitz continuous. If f is Lipschitz continuous with constant κ, then using (6) we have that: $|f^{\oplus}(\nu){-}f^{\oplus}(\nu')|$ equals

$$\begin{array}{rl} |\nu'(c){-}\nu(c)| + \sum_{(C,\ell')\in 2^{\mathcal{C}}\times L}\delta[\ell,a](C,\ell')\cdot \left|f(s_{\ell',\nu',\alpha,C}){-}f(s_{\ell',\nu,\alpha,C})\right| & \\ \leq & |\nu'(c){-}\nu(c)| + \sum_{(C,\ell')\in 2^{\mathcal{C}}\times L}\delta[\ell,a](C,\ell')\cdot\kappa\cdot\|\nu{-}\nu'\|_{\infty} \\ = & |\nu'(c){-}\nu(c)| + \kappa\cdot\|\nu{-}\nu'\|_{\infty} \leq (1+\kappa)\cdot\|\nu{-}\nu'\|_{\infty}. \end{array}$$

and hence it follows that $f^{\oplus}$ is Lipschitz constant with constant $(1+\kappa)$.

It therefore remains to show that $f^{\oplus}$ is monotonically decreasing and nonexpansive w.r.t $\trianglelefteq$. For any $\nu, \nu' \in V$ such that $\nu \trianglelefteq \nu'$ and $\nu'{-}\nu = d$, we have the following two cases to consider.

- If $\nu(c) = \nu'(c)$, then for any set $(C,\ell') \in 2^{\mathcal{C}} \times L$ we have that $(\nu{+}b{-}\nu(c))[C{:=}0] \trianglelefteq (\nu{+}b{-}\nu'(c))[C{:=}\,0]$, and hence $f(s_{\ell',\nu,\alpha,C}){-}f(s_{\ell',\nu',\alpha,C})$ is nonnegative for all $(C,\ell') \in 2^{\mathcal{C}} \times L$. Moreover, since f is nonexpansive, we have that $f(s_{\ell',\nu,\alpha,C}) - f(s_{\ell',\nu',\alpha,C}) \leq d$. It follows that $f^{\oplus}$ is monotonically decreasing and non-expansive as using (6) it follows that:

  $$f^{\oplus}(\nu){-}f^{\oplus}(\nu') = \textstyle\sum_{(C,\ell')\in 2^{\mathcal{C}}\times L}\delta[\ell,a](C,\ell')\cdot(f(s_{\ell',\nu,\alpha,C}) - f(s_{\ell',\nu',\alpha,C}))$$

  and $\nu'{-}\nu = d$.
- If $\nu'(c){-}\nu(c) = d$, then for any $(C,\ell') \subseteq 2^{\mathcal{C}} \times L$ we have that

  $$(\nu'{+}b{-}\nu'(c))[C{:=}\,0] \trianglelefteq (\nu{+}b{-}\nu(c))[C{:=}0]$$

  which implies that $f(s_{\ell',\nu,\alpha,C}){-}f(s_{\ell',\nu',\alpha,C})$ is nonpositive for all $(C,\ell') \in 2^{\mathcal{C}} \times L$. Moreover since f is nonexpansive, we have that $f(s_{\ell',\nu,\alpha,C}){-}f(s_{\ell',\nu',\alpha,C}) \leq d$. Similarly to the case above we have that $f^{\oplus}$ is monotonically decreasing and nonexpansive.

The proof is now complete. □

of Proposition 4.7. Let us consider the improvement function:

$$\mathbf{\Psi}(f)(s) \stackrel{\text{def}}{=} \begin{cases} 0 & \text{if } s \in \widehat{F} \\ \min\limits_{\alpha\in\widehat{A}_{\mathrm{Min}}(s)} \Big\{ \max\limits_{\beta\in\widehat{A}_{\mathrm{Max}}(s)} \big\{\widehat{\pi}(s,\alpha,\beta) + \sum\limits_{s'\in\widehat{S}} \widehat{p}(s'|s,\alpha,\beta)\cdot f(s')\big\}\Big\} & \text{if } s \notin \widehat{F} \end{cases}$$

We wish to show that if f is regionally quasi-simple, then $\mathbf{\Psi}(f)$ is regionally quasi-simple. If $s \in \widehat{F}$ then the proposition is trivial, therefore we assume that $s \notin \widehat{F}$. From Lemma E.1 we have that $s \mapsto \widehat{\pi}(s,\alpha,\beta) + \sum_{s'\in\widehat{S}} \widehat{p}(s'|s,\alpha,\beta)\cdot f(s')$ is regionally quasi-simple for every α and β. Since pointwise maximum of a set of quasi-simple function is quasi-simple, if follows that the function:

$$s \mapsto \max_{\beta\in\widehat{A}_{\mathrm{Max}}(s)} \big\{\widehat{\pi}(s,\alpha,\beta) + \sum_{s'\in\widehat{S}} \widehat{p}(s'|s,\alpha,\beta)\cdot f(s')\big\}$$

is quasi-simple. Noting that pointwise minimum of quasi-simple function is also quasi-simple, the function:

$$s \mapsto \min_{\alpha \in \widehat{A}_{\mathrm{Min}}(s)} \Big\{ \max_{\beta \in \widehat{A}_{\mathrm{Max}}(s)} \big\{ \widehat{\pi}(s,\alpha,\beta) + \sum_{s' \in \widehat{S}} \widehat{p}(s'|s,\alpha,\beta) \cdot f(s') \big\} \Big\}$$

is regionally quasi-simple, and hence $\mathbf{\Psi}(f)$ is regionally quasi-simple as required. □

F Proof of Lemma 4.10

of Lemma 4.10. Let $s = (\ell,\nu) \in S$ and $\zeta \in \mathcal{R}$ such that $[\nu] \xrightarrow{*} \zeta$. Let $P : \widehat{S} \to \mathbb{R}$ is regionally quasi-simple. Recall that $\widetilde{P} : S \to \mathbb{R}$ denote the function where $\widetilde{P}(\ell,\nu) = P(\ell,\nu,[\nu])$ for all $(\ell,\nu) \in S$.

We need to show that the functions

$$t \;\mapsto\; t + \textstyle\sum_{s' \in S} p(s'|s,(t,a),\bot) \cdot \widetilde{P}(s') \tag{7}$$

$$t \;\mapsto\; t + \textstyle\sum_{s' \in S} p(s'|s,\bot,(t,b)) \cdot \widetilde{P}(s') \tag{8}$$

are continuous and nondecreasing on the domain $\{t \in \mathbb{R}_{\geq 0} \,|\, \nu{+}t \in \zeta\}$. We only show the proof for the function in (7) as the proof for the function in (8) is similar.

To show that the function in (7) is continuous and nondecreasing, we need to show that the function $P^{\boxplus} : \{t \in \mathbb{R}_{\geq 0} \,|\, \nu{+}t \in \zeta\} \to \mathbb{R}$ defined by

$$P^{\boxplus}(t) \stackrel{\text{def}}{=} t + \sum_{(C,\ell') \in 2^{\mathcal{C}} \times L} \delta[\ell,a](C,\ell') \cdot P(\ell', \nu + t[C{:=}0], \zeta[C{:=}0])$$

is continuous and nondecreasing. To ease notation, let $I = \{t \in \mathbb{R}_{\geq 0} \,|\, \nu{+}t \in \zeta\}$, $\nu^t_C = \nu + t[C{:=}0]$ and $\zeta^C = \zeta[C{:=}0]$ and consider any $t_1, t_2 \in I$ such that $t_1 \leq t_2$. To prove lemma holds it is sufficient to show that $P^{\boxplus}(t_2) - P^{\boxplus}(t_1)$ is nonnegative. Now by definition we have:

$$\begin{aligned}
P^{\boxplus}(t_2) - P^{\boxplus}(t_1) &= t_2 - t_1 + \sum_{(C,\ell') \in 2^{\mathcal{C}} \times L} \delta[\ell,a](C,\ell') \cdot \big(P(\ell',\nu^{t_2}_C,\zeta^C) - P(\ell',\nu^{t_1}_C,\zeta^C)\big) \\
&= t_2 - t_1 - \sum_{(C,\ell') \in 2^{\mathcal{C}} \times L} \delta[\ell,a](C,\ell') \cdot \big(P(\ell',\nu^{t_1}_C,\zeta^C) - P(\ell',\nu^{t_2}_C,\zeta^C)\big) \\
&\geq t_2 - t_1 - \sum_{(C,\ell') \in 2^{\mathcal{C}} \times L} \delta[\ell,a](C,\ell') \cdot (t_2 - t_1) \\
&\geq 0
\end{aligned}$$

where the first inequality is due to the fact the P is monotonically decreasing and nonexpansive w.r.t $\trianglelefteq$, and $\nu^{t_1}_C \trianglelefteq \nu^{t_2}_C$, while the last inequality follows from δ being a probability distribution. □

G Proof of Theorem 4.9

Consider ETRG $\mathcal{T}$, let the semantics ERG of $\mathcal{T}$ be $[\![\mathcal{T}]\!] = (S, F, A_{\mathrm{Min}}, A_{\mathrm{Max}}, p_{\mathrm{Min}}, p_{\mathrm{Max}}, \pi_{\mathrm{Min}}, \pi_{\mathrm{Max}})$, while the BRA be $\widehat{\mathcal{T}} = (\widehat{S}, \widehat{F}, \widehat{A}_{\mathrm{Min}}, \widehat{A}_{\mathrm{Max}}, \widehat{p}_{\mathrm{Min}}, \widehat{p}_{\mathrm{Max}}, \widehat{\pi}_{\mathrm{Min}}, \widehat{\pi}_{\mathrm{Max}})$. Suppose that $P \models \mathrm{Opt}(\widehat{\mathcal{T}})$, to prove the theorem holds it is sufficient to show that for any $s{=}(\ell,\nu) \notin F$ we have:

$$\widetilde{P}(s) \quad = \quad \inf_{\alpha \in A_{\mathrm{Min}}(s)} \Big\{ \sup_{\beta \in A_{\mathrm{Max}}(s)} \big\{ \pi(s,\alpha,\beta) + \sum_{s' \in S} p(s'|s,\alpha,\beta) \cdot \widetilde{P}(s') \big\} \Big\} \tag{9}$$

where $\widetilde{P} : S \to \mathbb{R}$ is such that $\widetilde{P}(\ell,\nu) = P(\ell,\nu,[\nu])$ for all $(\ell,\nu) \in S$. We analyse the function on the RHS of (9) in detail by breaking into more manageable pieces, and by showing that in every possible scenario, player Min and Max prefer to choose their actions on region boundaries. For this purpose we introduce functions $P^{\perp}(s)$ (where player Min chooses to play $\perp$) and $P^{\zeta_a,a}(s)$ (where Min intends to delay for some time until region ζ_a and then take an action a). Formally,

$$\text{RHS } of(9) \quad \stackrel{\text{def}}{=} \quad \min \Big\{ P^{\perp}(s), \min_{\zeta_a \in \mathcal{R}, a \in Act_{\mathrm{Min}}} \big\{ P^{(\zeta_a,a)}(s) \big\} \Big\}, \tag{10}$$

To keep the notation simple, we assume that such ζ_a is in the future of $[\nu]$ and invariant of ℓ is satisfied during such delay, and moreover action a is enabled in location ℓ and region ζ_a. Now we further break $P^{\perp}(s)$ as maximum over all choices of actions b and clock regions ζ_b of function $P^{\perp}_{\zeta_b,b}(s)$ (where player Max chooses to delay until region ζ_b and then choose an enabled action b), i.e.,

$$P^{\perp}(s) \stackrel{\text{def}}{=} \max_{\zeta_b \in \mathcal{R}, b \in Act_{\mathrm{Max}}} \Big\{ P^{\perp}_{(\zeta_b,b)}(s) \Big\}. \tag{11}$$

The function $P^{\perp}_{\zeta_b,b}(s)$ represents the supremum over all time delays t_b such that $\nu + t_b \in \zeta_b$, i.e.

$$P^{\perp}_{(\zeta_b,b)}(s) \stackrel{\text{def}}{=} \sup_{t_b \,|\, \nu + t_b \in \zeta_b} \big\{ P^{\perp}_{(\zeta_b,b)}(s,t_b) \big\}, \tag{12}$$

where the function $P^{\perp}_{(\zeta_b,b)}(s,t_b)$ defined in the following manner.

$$P^{\perp}_{(\zeta_b,b)}(s,t_b) \stackrel{\text{def}}{=} t_b + \sum_{s' \in S} p(s'|s,\perp,(t_b,b)) \cdot \widetilde{P}(s'). \tag{13}$$

Since P is regionally quasi-simple, from Lemma 4.10 it follows that $P^{\perp}_{(\zeta_b,b)}(s,t_b)$ is continuous and nondecreasing in t_b. Notice that if ζ_b is a thin region then the set $\{t_b \,|\, \nu + t_b \in \zeta_b\}$ is a singleton and its member t_b is such that $t_b = e - \nu(c)$ for some $e \in [\![K]\!]_{\mathbb{N}}$ and $c \in C$. Similarly if ζ_b is thick then there exists $e, e' \in [\![K]\!]_{\mathbb{N}}$ and $c, c' \in C$ such that $\{t_b \,|\, \nu + t_b \in \zeta_b\} = \{t_b \,|\, e' - \nu(c') < t_b < e - \nu(c)\}$. Observe that for a continuous and nondecreasing function f we have that $\sup_{t_1 < t < t_2} f(t) = f(t_2)$. Hence the optimal strategy for player Max in (12) is as close as possible to farthest region boundary.

Now consider the function $P^{(\zeta_a,a)}(s)$. It can be written as infimum over all time delays t_a such that $\nu + t_a \in \zeta_a$, i.e.

$$P^{(\zeta_a,a)}(s) \stackrel{\text{def}}{=} \inf_{t_a \,|\, \nu + t_a \in \zeta_a} P^{(\zeta_a,a)}(s,t_a) \tag{14}$$

where the function $P^{(\zeta_a,a)}(s,t_a)$ is defined as follows:

$$P^{(\zeta_a,a)}(s,t_a) \quad \stackrel{\text{def}}{=} \quad \max \Big\{ P^{(\zeta_a,a)}_{\perp}(s,t_a), \max_{\zeta_b \in \mathcal{R},\, \zeta_b \xrightarrow{*} \zeta_a,\, b \in Act_{\mathrm{Max}}} \big\{ P^{(\zeta_a,a)}_{(\zeta_b,b)}(s,t_a) \big\} \Big\}$$

The function $P_{\perp}^{(\zeta_a,a)}(s,t_a)$ corresponds to the situation when player Max finds it better to execute player Min's choice, and thus plays the action $\perp$. Formally,

$$P_{\perp}^{(\zeta_a,a)}(s,t_a) \stackrel{\text{def}}{=} t_a + \sum_{s' \in S} p(s'|s,(t_a,a),\perp) \cdot \widetilde{P}(s')$$

On the other hand, the function $P_{(\zeta_b,b)}^{(\zeta_a,a)}(s,t_a)$ corresponds to player Max's intention of choosing a delay till region ζ_b and then executing the action b. Note that in doing so player Max must choose a time delay smaller or equal to t_a, hence ζ_b must be in the past of ζ_a, and moreover if $\zeta_a = \zeta_b$ then the time delay must be smaller than t_a. Formally,

$$P_{(\zeta_b,b)}^{(\zeta_a,a)}(s,t_a) \stackrel{\text{def}}{=} \begin{cases} \sup\limits_{t_b \,|\, \nu + t_b \in \zeta_b} \left\{ P_{(\zeta_b,b)}^{\perp}(s,t_b) \right\} & \zeta_a \neq \zeta_b \\ \sup\limits_{t_b \leq t_a \,|\, \nu + t_b \in \zeta_b} \left\{ P_{(\zeta_b,b)}^{\perp}(s,t_b) \right\} & \zeta_a = \zeta_b \end{cases} \tag{15}$$

where $P_{(\zeta_b,b)}^{\perp}(s,t_b)$ is defined in (13). Since P is regionally quasi-simple, from Lemma 4.10 it follows that $P_{(\zeta_b,b)}^{\perp}(s,t_b)$ is continuous and nondecreasing in t_b. Hence in (15) for the first case, the best choice for t_b equal to the delay that corresponds to furthest boundary of region ζ_b, while in the second case the optimal delay is equal to t_a. It follows that the function $P^{(\zeta_a,a)}(s,t_a)$ is nondecreasing and continuous in t_a, and that in (14) the best choice of t_a for player Min is on nearest region boundary of region ζ_a. Since the moves of both players in every possible situation are on region boundaries, the theorem follows. $\square$