Computing Laboratory

CONCAVELY-PRICED PROBABILISTIC TIMED AUTOMATA

Marcin Jurdziński
Marta Z. Kwiatkowska
Gethin Norman
Ashutosh Trivedi

CL-RR-09-06

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

Abstract

Concavely-priced probabilistic timed automata, an extension of probabilistic timed automata, are introduced. In this paper we consider expected reachability, discounted, and average price problems for concavely-priced probabilistic timed automata for arbitrary initial states. We prove that these problems are EXPTIME-complete for probabilistic timed automata with two or more clocks and PTIME-complete for automata with one clock. Previous work on expected price problems for probabilistic timed automata was restricted to expected reachability for linearly-priced automata and integer valued initial states. This work uses the boundary region graph introduced by Jurdziński and Trivedi to analyse properties of concavely-priced (non-probabilistic) timed automata.

1 Introduction

Markov decision processes [27] (MDPs) extend finite automata by providing a probability distribution over successor states for each transition. Timed automata [1] extend finite automata by providing a mechanism to constrain the transitions with real-time. Probabilistic timed automata (PTAs) [20, 15, 3] generalise both timed automata and MDPs by allowing both probabilistic and real-time behaviour.

Priced timed automata are timed automata with (time-dependent) prices attached to locations. Optimisation problems on priced timed automata are fundamental to the *verification* of (quantitative timing) properties of systems modelled as timed automata. In linearly-priced timed automata [2, 24] the price information is given by a real-valued function over locations which returns the price to be paid for each time-unit spent in the location. Jurdziński and Trivedi [18] have recently proposed a generalisation to *concave prices*, where the price of remaining in a location is a concave function over time, and demonstrated that, for such prices, a number of optimisation problems including reachability-price, discounted-price, and average-price are PSPACE-complete.

In this paper we present *concavely-priced probabilistic timed automata*, that is, probabilistic timed automata with concave prices. Concave functions appear frequently in many modelling scenarios such as in economics when representing resource utilization, sales or productivity. Examples include:

1. when renting equipment the rental rate decreases as the rental duration increases [11];
2. the expenditure by *vacation travellers* in an airport is typically a concave function of the waiting time [28];
3. the price of perishable products with fixed stock over a finite time period is usually a concave function of time [12], e.g., the price of a low-fare air-ticket as a function of the time remaining before the departure date.

Contribution. We show that finite-horizon expected total price, and infinite-horizon expected reachability, discounted price and average price objectives on concavely-priced PTAs are decidable. We also show that the complexity of solving expected reachability price, expected discounted price and expected average price problems are EXPTIME-complete for concavely-priced

PTAs with two or more clocks and PTIME-complete for concavely-priced PTAs with one clock. An important contribution of this paper are the proof techniques which complement the techniques of [18]. We extend the boundary region graph construction for timed automata [18] to PTAs and demonstrate that all the optimisation problems considered can be reduced to similar problems on the boundary region graph. We characterise the values of the optimisation problems by *optimality equations* and prove the correctness of the reduction by analysing the solutions of the optimality equations on the boundary region graph. This allows us to obtain an efficient algorithm matching the EXPTIME lower bound for the problems. This is a the technical report version of the paper [16].

Related Work. For a review of work on optimisation problems for (non-probabilistic) timed automata we refer the reader to [18]. For priced probabilistic timed automata work has been limited to considering linearly-priced PTAs. Based on the digital clocks approach [13], Kwiatkowska et. al. [19] present a method for solving infinite-horizon expected reachability problems for a subclass of probabilistic timed automata. In [6] an algorithm for calculating the maximal probability of reaching some goal location within a given cost (and time) bound is presented. The algorithm is shown to be partially correct in that if it terminates it terminates with the correct value. Following this, [5] demonstrates the undecidability of this problem. We also mention the approaches for analysing unpriced probabilistic timed automata against temporal logic specifications based on the region graph [20, 15] and either forwards [20] or backwards [21] reachability. The complexity of performing such verification is studied in [23, 17].

2 Preliminaries

We assume, wherever appropriate, sets $\mathbb{N}$ of non-negative integers, $\mathbb{R}$ of reals and $\mathbb{R}_{\oplus}$ of non-negative reals. For $n \in \mathbb{N}$, let $[\![n]\!]_{\mathbb{N}}$ and $[\![n]\!]_{\mathbb{R}}$ denote the sets $\{0, 1, \dots, n\}$, and $\{r \in \mathbb{R} \,|\, 0{\le}r{\le}n\}$ respectively. A *discrete probability distribution* over a countable set Q is a function $\mu : Q \to [0,1]$ such that $\sum_{q \in Q} \mu(q){=}1$. For a possible uncountable set Q', we define $\mathcal{D}(Q')$ to be the set of functions $\mu : Q' \to [0,1]$ such that the set *supp*$(\mu) \stackrel{\text{def}}{=} \{q \in Q \,|\, \mu(q){>}0\}$ is countable and, over the domain *supp*(μ), μ is a distribution. We say that $\mu \in \mathcal{D}(Q)$ is a *point distribution* if $\mu(q){=}1$ for some $q \in Q$.

A set $\mathcal{D} \subseteq \mathbb{R}^n$ is *convex* if $\theta{\cdot}x + (1{-}\theta){\cdot}y \in \mathcal{D}$ for all $x, y \in \mathcal{D}$ and $\theta \in [0,1]$. A function $f : \mathbb{R}^n \to \mathbb{R}$ is *concave* (on its domain $\mathrm{dom}(f) \subseteq \mathbb{R}^n$), if $\mathrm{dom}(f) \subseteq \mathbb{R}^n$ is a convex set and $f(\theta{\cdot}x{+}(1{-}\theta){\cdot}y) \ge \theta{\cdot}f(x){+}(1{-}\theta){\cdot}f(y)$ for all $x, y \in \mathrm{dom}(f)$ and $\theta \in [0,1]$. We require the following well known [8] properties of concave functions.

Lemma 2.1.
1. *If $f_1, \dots, f_k : \mathbb{R}^n \to \mathbb{R}$ are concave and $w_1, \dots, w_k \in \mathbb{R}_{\oplus}$, then $\sum_{i=1}^{k} w_i{\cdot}f_i : \mathbb{R}^n \to \mathbb{R}$ is concave on the domain $\bigcap_{i=1}^{k} \mathrm{dom}(f_i)$.*
2. *If $f : \mathbb{R}^n \to \mathbb{R}$ is concave and $g : \mathbb{R}^m \to \mathbb{R}^n$ linear, then $x \mapsto f(g(x))$ is concave.*
3. *If $f_1, \dots f_k : \mathbb{R}^n \to \mathbb{R}$ are concave, then $x \mapsto \min_{i=1}^{k} f_i(x)$ is concave on the domain $\bigcap_{i=1}^{k} \mathrm{dom}(f_i)$.*

4. *If $f_i : \mathbb{R}^n \to \mathbb{R}$ is concave for $i \in \mathbb{N}$, then $x \mapsto \lim_{i\to\infty} f_i(x)$ is concave.*

Lemma 2.2. *If $f : (a,b) \to \mathbb{R}$ is a concave function and $\overline{f}$ is the (unique) continuous extension of f to the closure of the interval (a,b), then $\inf_{x\in(a,b)} f(x) = \min\{\overline{f}(a), \overline{f}(b)\}$.*

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *Lipschitz continuous* on its domain, if there exists a constant $K{\geq}0$, called a Lipschitz constant of f, such that $\|f(x){-}f(y)\|_\infty \leq K\|x{-}y\|_\infty$ for all $x, y \in \mathrm{dom}(f)$; we then also say that f is K-Lipschitz continuous.

3 Markov Decision Processes

In this section we introduce Markov decision processes (MDPs), a form of transition systems which exhibit both probabilistic and nondeterministic behaviour.

Definition 3.1. *A* priced Markov decision process *is a tuple $\mathcal{M} = (S, A, p, \pi)$ where:*

- *S is the set of* states;
- *A is the set of* actions;
- *$p : S \times A \to \mathcal{D}(S)$ is a partial function called the* probabilistic transition function;
- *$\pi : S \times A \to \mathbb{R}$ is a bounded and measurable* price function *assigning real-values to state-action pairs.*

We write $A(s)$ for the set of actions available at s, i.e., the set of actions a for which $p(s,a)$ is defined. For technical convenience we assume that $A(s)$ is nonempty for all $s \in S$. We say $\mathcal{M}$ is finite, if the sets S and A are finite.

In the priced MDP $\mathcal{M}$, if the current state is s, then a strategy chooses an action $a \in A(s)$ after which a probabilistic transition is made according to the distribution $p(s,a)$, i.e., state $s' \in S$ is reached with probability $p(s'|s,a) \stackrel{\text{def}}{=} p(s,a)(s')$. We say that (s,a,s') is a transition of $\mathcal{M}$ if $p(s'|s,a){>}0$ and a *run* of $\mathcal{M}$ is a sequence $\langle s_0, a_1, s_1, \ldots\rangle \in S{\times}(A{\times}S)^*$ such that (s_i, a_{i+1}, s_{i+1}) is a transition for all $i{\geq}0$. We write $\mathit{Runs}^{\mathcal{M}}$ ($\mathit{Runs}^{\mathcal{M}}_{\mathit{fin}}$) for the sets of infinite (finite) runs and $\mathit{Runs}^{\mathcal{M}}(s)$ ($\mathit{Runs}^{\mathcal{M}}_{\mathit{fin}}(s)$) for the sets of infinite (finite) runs starting from state s. For a finite run $r{=}\langle s_0, a_1, \ldots, s_n\rangle$ we write $\mathit{last}(r){=}s_n$ for the last state of the run. Furthermore, let X_i and Y_i denote the random variables corresponding to i^{th} state and action of a run.

A *strategy* in $\mathcal{M}$ is a function $\sigma : \mathit{Runs}^{\mathcal{M}}_{\mathit{fin}} \to \mathcal{D}(A)$ such that $\mathit{supp}(\sigma(r)) \subseteq A(\mathit{last}(r))$ for all $r \in \mathit{Runs}^{\mathcal{M}}_{\mathit{fin}}$. Let $\mathit{Runs}^{\mathcal{M}}_{\sigma}(s)$ denote the subset of $\mathit{Runs}^{\mathcal{M}}(s)$ which correspond to the strategy σ when starting in state s. Let $\Sigma_{\mathcal{M}}$ be the set of all strategies in $\mathcal{M}$. We say that a strategy σ is *pure* if $\sigma(r)$ is a point distribution for all $r \in \mathit{Runs}^{\mathcal{M}}_{\mathit{fin}}$. We say that a strategy σ is *stationary* if $\mathit{last}(r){=}\mathit{last}(r')$ implies $\sigma(r){=}\sigma(r')$ for all $r, r' \in \mathit{Runs}^{\mathcal{M}}_{\mathit{fin}}$. A strategy σ is *positional* if it is both pure and stationary. To analyse the behaviour of an MDP $\mathcal{M}$ under a strategy σ, for each state s of $\mathcal{M}$, we define a probability space $(\mathit{Runs}^{\mathcal{M}}_{\sigma}(s), \mathcal{F}_{\mathit{Runs}^{\mathcal{M}}_{\sigma}(s)}, \mathrm{Prob}^{\sigma}_{s})$ over the set of infinite runs of σ with s as the initial state. For details on this construction see, for example, [27]. Given a *real-valued random variable* over the set of infinite runs $f : \mathit{Runs}^{\mathcal{M}} \to \mathbb{R}$, using standard techniques from probability theory, we can define the expectation of this variable $\mathbb{E}^{\sigma}_{s}\{f\}$ with respect to σ when starting in s.

3.0.1 Performance objectives.

For a priced MDP $\mathcal{M} = (S, A, p, \pi)$, under any strategy σ and starting from any state s, there is a sequence of random prices $\{\pi(X_{i-1}, Y_i)\}_{i\geq 1}$. Depending on the problem under study there are a number of different performance objectives that can be studied. Below are the objectives most often used.

1. *Expected Reachability Price* (with *target set* F):

$$\text{EReach}_{\mathcal{M}}(F)(s,\sigma) \stackrel{\text{def}}{=} \mathbb{E}^{\sigma}_{s}\left\{\textstyle\sum_{i=1}^{\min\{i\,|\,X_i \in F\}} \pi(X_{i-1}, Y_i)\right\}.$$

2. *Expected Total Price* (with *horizon* N):

$$\text{ETotal}_{\mathcal{M}}(N)(s,\sigma) \stackrel{\text{def}}{=} \mathbb{E}^{\sigma}_{s}\left\{\textstyle\sum_{i=1}^{N} \pi(X_{i-1}, Y_i)\right\}.$$

3. *Expected Discounted Price* (with *discount factor* $\lambda \in (0,1)$):

$$\text{EDisct}_{\mathcal{M}}(\lambda)(s,\sigma) \stackrel{\text{def}}{=} \mathbb{E}^{\sigma}_{s}\left\{(1-\lambda)\textstyle\sum_{i=1}^{\infty} \lambda^{i-1} \pi(X_{i-1}, Y_i)\right\} .$$

4. *Expected Average Price:*

$$\text{EAvg}_{\mathcal{M}}(s,\sigma) \stackrel{\text{def}}{=} \limsup_{n\to\infty} \frac{1}{n} \mathbb{E}^{\sigma}_{s}\left\{\textstyle\sum_{i=1}^{n} \pi(X_{i-1}, Y_i)\right\}.$$

For an objective $\text{ECost}_{\mathcal{M}}$ and state s we let $\text{ECost}^*_{\mathcal{M}}(s) = \inf_{\sigma \in \Sigma_M} \text{ECost}_{\mathcal{M}}(s,\sigma)$. A strategy σ of $\mathcal{M}$ is *optimal* for $\text{ECost}_{\mathcal{M}}$ if $\text{ECost}_{\mathcal{M}}(s,\sigma) = \text{ECost}^*_{\mathcal{M}}(s)$ for all $s \in S$. Note that an optimal strategy need not exist, and in such cases one can consider, for each $\varepsilon{>}0$, a ε-optimal strategy, that is, a strategy σ such that $\text{ECost}^*_{\mathcal{M}}(s) \geq \text{ECost}_{\mathcal{M}}(s,\sigma){-}\varepsilon$ for all $s \in S$. For technical convenience we make the follows assumptions for reachability objectives [9].

Assumption 3.2. *If $s \in F$, then s is* absorbing *and* price free*, i.e. $p(s|s,a){=}1$ and $\pi(s,a){=}0$ for all $a \in A(s)$.*

Assumption 3.3. *For all $\sigma \in \Sigma_M$ and $s \in S$ we have* $\lim_{i\to\infty} \text{Prob}^{\sigma}_{s}(X_i \in F) = 1$.

Optimality Equations. We now review optimality equations for determining the objectives given above.

1. Let $P : S \to \mathbb{R}$ and $F \subseteq S$; we write $P \models \text{Opt}^F_R(\mathcal{M})$, and we say that P is a solution of optimality equations $\text{Opt}^F_R(\mathcal{M})$ if, for all $s \in S$, we have:

$$P(s) = \begin{cases} 0 & \text{if } s \in F \\ \inf_{a\in A(s)} \left\{\pi(s,a) + \sum_{s' \in S} p(s'|s,a) \cdot P(s')\right\} & \text{otherwise.} \end{cases}$$

2. Let $T_0, \ldots, T_N : S \to \mathbb{R}$; we say that $\langle T_i \rangle_{i=0}^N$ is a solution of optimality equations $\mathrm{Opt}_T^N(\mathcal{M})$ if, for all $s \in S$, we have:

$$T_i(s) = \begin{cases} 0 & \text{if } i = 0 \\ \inf_{a \in A(s)} \left\{ \pi(s,a) + \sum_{s' \in S} p(s'|s,a) \cdot T_{i-1}(s') \right\} & \text{otherwise.} \end{cases}$$

3. Let $D : S \to \mathbb{R}$; we write $D \models \mathrm{Opt}_D^\lambda(\mathcal{M})$, and we say that D is a solution of optimality equations $\mathrm{Opt}_D^\lambda(\mathcal{M})$ if, for all $s \in S$, we have:

$$D(s) = \inf_{a \in A(s)} \left\{ (1-\lambda) \cdot \pi(s,a) + \lambda \cdot \sum_{s' \in S} p(s'|s,a) \cdot D(s') \right\}.$$

4. Let $G : S \to \mathbb{R}$ and $B : S \to \mathbb{R}$; we write $(G, B) \models \mathrm{Opt}_A(\mathcal{M})$, and we say that (G, B) is a solution of optimality equations[1] $\mathrm{Opt}_A(\mathcal{M})$, if for all $s \in S$, we have:

$$\begin{aligned} G(s) &= \inf_{a \in A(s)} \left\{ \sum_{s' \in S} p(s'|s,a) \cdot G(s') \right\} \\ B(s) &= \inf_{a \in A(s)} \left\{ \pi(s,a) - G(s) + \sum_{s' \in S} p(s'|s,a) \cdot B(s') \right\} \end{aligned}$$

The proof of the following proposition is routine and for details see, for example, [10].

Proposition 3.4. *Let $\mathcal{M}$ be a priced MDP.*

1. *If $P \models \mathrm{Opt}_R^F(\mathcal{M})$, then $P(s) = \mathrm{EReach}_{\mathcal{M}}^*(F)(s)$ for all $s \in S$.*

2. *If $\langle T_i \rangle_{i=0}^N \models \mathrm{Opt}_R^N(\mathcal{M})$, then $T_i(s) = \mathrm{ETotal}_{\mathcal{M}}^*(i)(s)$ for all $i \leq N$ and $s \in S$.*

3. *If $D \models \mathrm{Opt}_D^\lambda(\mathcal{M})$ and D is bounded, then $D(s) = \mathrm{EDisct}_{\mathcal{M}}^*(\lambda)(s)$ for all $s \in S$.*

4. *If $(G, B) \models \mathrm{Opt}_A(\mathcal{M})$ and G, B are bounded, then $G(s) = \mathrm{EAvg}_{\mathcal{M}}^*(s)$ for all $s \in S$.*

Notice that, for each objective, if, for every state $s \in S$, the infimum is attained in the optimality equations, then there exists an optimal positional strategy. An important class of MDPs with this property are finite MDPs, which gives us the following proposition.

Proposition 3.5. *For every finite MDP, the existence of a solution of the optimality equations for the expected reachability, discounted and average price implies the existence of a positional optimal strategy for the corresponding objective.*

For a finite MDP $\mathcal{M}$ a solution of optimality equations for expected reachability, total, discounted, and average price objectives can be obtained by value iteration or strategy improvement algorithms [27].

Proposition 3.6. *For every finite MDP, there exist solutions of the optimality equations for expected reachability, total, discounted, and average price objectives.*

[1] These optimality equations are slightly different from Howard's optimality equations for expected average price, and correspond to Puterman's [27] modified optimality equations.

Proposition 3.5 together with Proposition 3.6 provide a proof of the following well-known result for priced MDPs [27].

Theorem 3.7. *For a finite priced MDP the reachability, discounted, and average price objectives each have an optimal positional strategy.*

Notice that the total price objective need not have an optimal positional strategy since, unlike the other objectives, it is a finite horizon problem. However, for this reason, its analysis concerns only finitely many strategies.

4 Concavely-Priced Probabilistic Timed Automata

In this section we introduce concavely-priced probabilistic timed automata and begin by defining clocks, clock valuations, clock regions and zones.

4.1 Clocks, clock valuations, regions and zones.

We fix a constant $k \in \mathbb{N}$ and finite set of *clocks* $\mathcal{C}$. A (k-bounded) *clock valuation* is a function $\nu : \mathcal{C} \to [\![k]\!]_{\mathbb{R}}$ and we write V for the set of clock valuations.

Assumption 4.1. *Although clocks in (probabilistic) timed automata are usually allowed to take arbitrary non-negative values, we have restricted the values of clocks to be bounded by some constant k. More precisely, we have assumed the models we consider are* bounded *probabilistic timed automata. This standard restriction [7] is for technical convenience and comes without significant loss of generality.*

If $\nu \in V$ and $t \in \mathbb{R}_{\oplus}$ then we write $\nu{+}t$ for the clock valuation defined by $(\nu{+}t)(c) = \nu(c){+}t$, for all $c \in \mathcal{C}$. For $C \subseteq \mathcal{C}$ and $\nu \in V$, we write $\nu[C{:=}0]$ for the clock valuation where $\nu[C{:=}0](c) = 0$ if $c \in C$, and $\nu[C{:=}0](c) = \nu(c)$ otherwise.

For a clock valuation ν we define its *fractional signature* $\lbag \nu \rbag$ 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 $\lbag \nu(c) \rbag = f_i$, and for every clock $c \in \mathcal{C}$, there is $i \leq m$, such that $\lbag \nu(c) \rbag = f_i$. 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 nonnegative integers $i \leq m$ we have $f'_i = \lbag f_i + 1 - f_k \rbag$. 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 $f'_i \leq f'_{i+1}$ for all $i < n$;
- for any $i \leq n$ there exists $j \leq m$ such that $f'_i = f_j$.

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

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

4.2 Probabilistic timed automata

We are now in a position to introduce probabilistic timed automata.

Definition 4.2. *A* probabilistic timed automaton $\mathsf{T} = (L, \mathcal{C}, inv, Act, E, \delta)$ *consists of:*

- *a finite set of locations L;*
- *a finite set of clocks $\mathcal{C}$;*
- *an* invariant condition $inv : L \to \mathcal{Z}$*;*
- *a finite set of* actions Act*;*
- *an* action enabledness function $E : L{\times}Act \to \mathcal{Z}$*;*
- *a* transition probability function $\delta : (L{\times}Act) \to \mathcal{D}(2^{\mathcal{C}}{\times}L)$.

When we consider a probabilistic timed automaton as an input of an algorithm, its size should be understood as the sum of the sizes of encodings of L, $\mathcal{C}$, inv, Act, E, and δ. A *configuration* of a probabilistic timed automaton T is a pair (ℓ, ν), where $\ell \in L$ is a location and $\nu \in V$ is a clock valuation over $\mathcal{C}$ such that $\nu \triangleleft inv(\ell)$. For any $t \in \mathbb{R}$, we let $(\ell, \nu){+}t$ equal the configuration $(\ell, \nu{+}t)$. Informally, the behaviour of a probabilistic timed automaton is as follows. In configuration (ℓ, ν) time passes before an available action is triggered, after which a discrete probabilistic transition occurs. Time passage is available only if the invariant condition $inv(\ell)$ is satisfied while time elapses, and the action a can be chosen after time t if the action is enabled in the location ℓ, i.e., if $\nu{+}t \triangleleft E(\ell, a)$. Both the amount of time and the action chosen are nondeterministic. If the action a is chosen, then the probability of moving to the location ℓ' and resetting all of the clocks in C to 0 is given by $\delta[\ell, a](C, \ell')$.

Formally, the semantics of a probabilistic timed automaton is given by an MDP which has both an infinite number of states and an infinite number of transitions.

Definition 4.3. *Let $\mathsf{T} = (L, \mathcal{C}, inv, Act, E, \delta)$ be a probabilistic timed automaton. The semantics of T is the MDP $[\![\mathsf{T}]\!] = (S_{\mathsf{T}}, A_{\mathsf{T}}, p_{\mathsf{T}})$ where*

- $S_{\mathsf{T}} \subseteq L{\times}V$ *such that $(\ell, \nu) \in S_{\mathsf{T}}$ if and only if $\nu \triangleleft inv(\ell)$;*
- $A_{\mathsf{T}} = \mathbb{R}_{\oplus}{\times}Act$*;*

- *for $(\ell,\nu) \in S_{\mathsf{T}}$ and $(t,a) \in A_{\mathsf{T}}$, we have $p_{\mathsf{T}}((\ell,\nu),(t,a)) = \mu$ if and only if*
  - $\nu{+}t' \triangleleft inv(\ell)$ *for all* $t' \in [0,t]$*;*
  - $\nu{+}t \triangleleft E(\ell,a)$*;*
  - $\mu(\ell',\nu') = \sum_{C\subseteq\mathcal{C}\wedge(\nu+t)[C:=0]=\nu'} \delta[\ell,a](C,\ell')$ *for all* $(\ell',\nu') \in S$*.*

We assume the following—standard and easy to syntactically verify—restriction on PTAs which ensures *time divergent* behaviour.

Assumption 4.4. *We restrict attention to* structurally non-Zeno *probabilistic timed automata [29, 17]. A PTA is structurally non-Zeno if, for any run $\langle s_0,(t_1,a_1),\dots,s_n\rangle$, such that $s_0{=}(\ell_0,\nu_0)$, $s_n{=}(\ell_n,\nu_n)$ and $\ell_0{=}\ell_n$ (i.e., the run forms a cycle in the finite graph of the locations and transitions of the automaton) we have $\sum_{i=1}^{n} t_i \geq 1$.*

4.3 Priced probabilistic timed automata

We now introduce priced probabilistic timed automata which extend probabilistic timed automata with price functions over state and time-action pairs.

Definition 4.5. *A* priced probabilistic timed automaton $\mathcal{T} = (\mathsf{T},\pi)$ *consists of a probabilistic timed automaton* T *and a price function* $\pi : (L{\times}V) \times (\mathbb{R}_{\oplus}{\times}Act) \to \mathbb{R}$.

The semantics of a priced PTA $\mathcal{T}$ is the priced MDP $[\![\mathcal{T}]\!] = ([\![\mathsf{T}]\!],\pi)$ where $\pi(s,(t,a))$ is the price of taking the action (t,a) from state s in $[\![\mathsf{T}]\!]$. In a *linearly-priced* PTA [19], the price function is represented as a function $r : L \cup Act \to \mathbb{R}$, which gives a *price rate* to every location ℓ, and a price to every action a; the price of taking the timed move (a,t) from state (ℓ,ν) is then defined by $\pi((\ell,\nu),(t,a)) = r(\ell){\cdot}t{+}r(a)$.

In this paper we restrict attention to *concave* price functions requiring that for any location $\ell \in L$ and action $a \in Act$ the function $\pi((\ell,\cdot),(\cdot,a)) : V{\times}\mathbb{R}_{\oplus}{\to}\mathbb{R}$ is concave. From Assumption 4.1 it follows that any such price function π is also K-Lipschitz continuous for some $K \in \mathbb{N}$. However, the results in this paper for PTAs with more than 1 clock, also hold for the more general *region-wise concave* price functions of [18]. Notice that every linearly-priced PTA is also concavely-priced.

Considering the optimisation problems introduced for priced MDPs in Section 3, the following is the main result of the paper.

Theorem 4.6. *The minimisation problems for reachability, total, discounted, and average cost functions for concavely-priced PTAs are decidable.*

In the next section we introduce the boundary region graph, an abstraction whose size is exponential in the size of PTA. In Section 6 we show that to solve the above mentioned optimisation problems on concavely-priced PTAs, it is sufficient to solve them on the corresponding boundary region graph.

5 Boundary Region Graph Construction

Before introducing the boundary region graph we review the standard region graph construction for timed automata [1] extended in [20] to probabilistic timed automata.

5.1 The region graph

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

For $R, R' \in \mathcal{R}$, we say that R' is in the future of R, or that R is in the past of R', if there is $s \in R$, $s' \in R'$ and $t \in \mathbb{R}_\oplus$ such that $s' = s{+}t$; we then write $R \rightarrow_* R'$. We say that R' is the *time successor* of R if $R \rightarrow_* R'$, $R{\neq}R'$, and $R \rightarrow_* R'' \rightarrow_* R'$ implies $R''{=}R$ or $R''{=}R'$ and write $R \rightarrow_{+1} R'$ and $R' \leftarrow_{+1} R$. Similarly we say that R' is the n^{th} successor of R and write $R \rightarrow_{+n} R'$, if there is a sequence of regions $\langle R_0, R_1, \ldots, R_n\rangle$ such that $R_0{=}R$, $R_n{=}R'$ and $R_i \rightarrow_{+1} R_{i+1}$ for every $0{\leq}i{<}n$.

The region graph is an MDP with both a finite number of states and transitions.

Definition 5.1. *Let* $\mathsf{T} = (L, \mathcal{C}, \mathit{inv}, \mathit{Act}, E, \delta)$ *be a probabilistic timed automaton. The* region graph *of* T *is the MDP* $\mathsf{T}_{\text{RG}} = (S_{\text{RG}}, A_{\text{RG}}, p_{\text{RG}})$ *where:*

- $S_{\text{RG}} = \mathcal{R}$*;*
- $A_{\text{RG}} \subseteq \mathbb{N}{\times}\mathit{Act}$ *such that if* $(n,a) \in A_{\text{RG}}$ *then* $n \leq (2{\cdot}|\mathcal{C}|)^k$*;*
- *for* $(\ell,\zeta) \in S_{\text{RG}}$ *and* $(n,a) \in A_{\text{RG}}$ *we have* $p_{\text{RG}}((\ell,\zeta),(n,a)) = \mu$ *if and only if*
  - $(\ell,\zeta) \rightarrow_{+n} (\ell,\zeta_n)$*;*
  - $\zeta_n \lhd E(\ell,a)$*;*
  - $\mu(\ell',\zeta') = \sum_{C\subseteq\mathcal{C}\wedge\zeta_n[C:=0]=\zeta'} \delta[\ell,a](C,l')$ *for all* $(\ell',\zeta') \in S_{\text{RG}}$.

Since the region graph abstracts away the precise timing information, it can not be used to solve expected reachability, total, discounted, and average price problems for the original probabilistic timed automaton. In the next section, we define a new abstraction of probabilistic timed automata, called the boundary region graph, which retains sufficient timing information to solve these performance objectives.

5.2 The boundary region graph

We say that a region $R \in \mathcal{R}$ is *thin* if $[s] \neq [s{+}\varepsilon]$ for every $s \in R$ and $\varepsilon{>}0$; other regions are called *thick*. We write $\mathcal{R}_{\text{Thin}}$ and $\mathcal{R}_{\text{Thick}}$ for the sets of thin and thick regions, respectively. Note that if $R \in \mathcal{R}_{\text{Thick}}$ then, for every $s \in R$, there is an $\varepsilon > 0$, such that $[s] = [s{+}\varepsilon]$. Observe that the time successor of a thin region is thick, and vice versa.

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

The motivation for the boundary region graph is the following. Let $a \in A$, $s = (\ell, \nu)$ and $R = (\ell, \zeta) \to_* R' = (\ell, \zeta')$ such that $s \in R$ and $R' \lhd E(\ell, a)$.

- If $R' \in \mathcal{R}_{\text{Thick}}$, then there are infinitely many $t \in \mathbb{R}_{\oplus}$ such that $s+t \in R'$. One of the main results that we establish is that in the state s, 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 each performance objective. However, since R' is a thick region, the set $\{t \in \mathbb{R}_{\oplus} \mid s+t \in R'\}$ is an open interval, and hence does not contain its boundary values. Observe that the infimum equals $b_- - \nu(c_-)$ where $R \to_{b_-,c_-} R_- \to_{+1} R'$ and the supremum equals $b_+ - \nu(c_+)$ where $R \to_{b_+,c_+} R_+ \leftarrow_{+1} R'$. In the boundary region graph we include these 'best' timed action through the actions $((b_-, c_-, a), R')$ and $((b_+, c_+, a), R')$.
- If $R' \in \mathcal{R}_{\text{Thin}}$, then there exists a unique $t \in \mathbb{R}_{\oplus}$ such that $(\ell, \nu+t) \in R'$. Moreover since R' is a thin region there exists a clock $c \in C$ and a number $b \in \mathbb{N}$ such that $R \to_{b,c} R'$ and $t = b - \nu(c)$. In the boundary region graph we summarize this 'best' timed action from region R via region R' through the action $((b, c, a), R')$.

With this intuition in mind, let us present the definition of a boundary region graph.

Definition 5.2. *Let* $\mathsf{T} = (L, \mathcal{C}, inv, A, E, \delta)$ *be a probabilistic timed automaton. The* boundary region graph *of* T *is defined as the MDP* $\mathsf{T}_{\text{BRG}} = (S_{\text{BRG}}, A_{\text{BRG}}, p_{\text{BRG}})$ *such that:*

- $S_{\text{BRG}} = \{((\ell, \nu), (\ell, \zeta)) \mid (\ell, \zeta) \in \mathcal{R} \wedge \nu \in \overline{\zeta}\}$*;*
- $A_{\text{BRG}} \subseteq ([\![k]\!]_{\mathbb{N}} \times \mathcal{C} \times Act) \times \mathcal{R}$*;*
- *for any state* $((\ell, \nu), (\ell, \zeta)) \in S_{\text{BRG}}$ *and action* $((b, c, a), (\ell, \zeta_a)) \in A_{\text{BRG}}$ *we have that* $p_{\text{BRG}}((\ell, \nu), (\ell, \zeta), ((b, c, a), (\ell, \zeta_a))) = \mu$ *if and only if*

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

  for all $((\ell', \nu'), (\ell', \zeta')) \in S_{\text{BRG}}$ *where* $\nu_a = \nu + time(\nu, (b, c))$ *and one of the following conditions holds:*

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

Although the boundary region graph is infinite, for a fixed initial state we can restrict attention to a finite state subgraph, thanks to the following observation [18].

Proposition 5.3. *For every state* $s \in S_{\text{BRG}}$ *of a boundary region graph* T_{BRG}*, the reachable subgraph* $\mathsf{T}^s_{\text{BRG}}$ *is a finite MDP.*

Proposition 5.4. *If* $s = ((\ell, \nu), (\ell, \zeta)) \in S_{\text{BRG}}$ *is such that* ν *is an integer valuation, then the MDP* $\mathsf{T}^s_{\text{BRG}}$ *is equivalent to the digital clock semantics [19] of* T *and extends the corner point abstraction of [7] to the probabilistic setting.*

Definition 5.5. *Let $\mathcal{T} = (\mathsf{T}, \pi)$ be a priced probabilistic timed automaton. The* priced boundary region graph *of $\mathcal{T}$ equals the priced MDP $\mathcal{T}_{\text{BRG}} = (\mathsf{T}_{\text{BRG}}, \pi_{\text{BRG}})$ where for any state $((\ell, v), (\ell, \zeta)) \in S_{\text{BRG}}$ and action $((b, c, a), (\ell, \zeta')) \in A_{\text{BRG}}$ available in the state:*

$$\pi_{\text{BRG}}\big(((\ell, v), (\ell, \zeta)), ((b, c, a), (\ell, \zeta'))\big) = \pi\big((\ell, \nu), (time(\nu, (b, c)), a)\big) .$$

6 Correctness of the Reduction to Boundary Region Automata

For the remainder of this section we fix a concavely-priced PTA $\mathcal{T}$. Proposition 5.3 together with Proposition 3.6 yield the following important result.

Proposition 6.1. *For the priced MDP $\mathcal{T}_{\text{BRG}}$ there exist solutions of the optimality equations for expected reachability, total, discounted, and average price objectives.*

We say that a function $f : S_{\text{BRG}} \to \mathbb{R}$ is *regionally concave* if for all $(\ell, \zeta) \in \mathcal{R}$ the function $f(\cdot, (\ell, \zeta)) : \{(\ell, \nu) \,|\, \nu \in \overline{\zeta}\} \to \mathbb{R}$ is concave.

Lemma 6.2. *Assume that $P \models \text{Opt}^F_R(\mathcal{T}_{\text{BRG}})$, $\langle T_i \rangle_{i=1}^N \models \text{Opt}^N_T(\mathcal{T}_{\text{BRG}})$, and $D \models \text{Opt}^\lambda_D(\mathcal{T}_{\text{BRG}})$ We have that P, T_N, and D are regionally concave.*

Proof. (Sketch.) Using an elementary, but notationally involved, inductive proof we can show that T_N is regionally concave. The proof uses closure properties of concave functions (see Lemma 2.1), along with the fact that price functions π are concave. The concavity of P and D follows from the observation that they can be characterised as the limit (concave due to Lemma 2.1) of certain optimal expected total price objectives. □

For any function $f : S_{\text{BRG}} \to \mathbb{R}$, we define $\widetilde{f} : S_{\mathsf{T}} \to \mathbb{R}$ by $\widetilde{f}(\ell, \nu) = f((\ell, \nu), (\ell, [\nu]))$.

Lemma 6.3. *If $P \models \text{Opt}^F_R(\mathcal{T}_{\text{BRG}})$, then $\widetilde{P} \models \text{Opt}^F_R([\![\mathcal{T}]\!])$.*

Proof. Assuming $P \models \text{Opt}^F_R(\mathcal{T}_{\text{BRG}})$, to prove this proposition it is sufficient to show that for any $s{=}(\ell, \nu) \in S_{\mathsf{T}}$ we have:

$$\widetilde{P}(s) = \inf_{(t,a)\in A(s)} \{\pi(s,(t,a)) + \sum_{(C,\ell')\in 2^{\mathcal{C}}\times L} \delta[\ell,a](C,\ell')\cdot \widetilde{P}(\ell',(\nu{+}t)[C{:=}0])\}. \tag{1}$$

We therefore fix a state $s{=}(\ell, \nu) \in S_{\mathsf{T}}$ for the remainder of the proof. For any $a \in Act$, let $\mathcal{R}^a_{\text{Thin}}$ and $\mathcal{R}^a_{\text{Thick}}$ denote the set of thin and think regions respectively that are successors of $[\nu]$ and are subsets of $E(\ell, a)$. Considering the RHS of (1) we have:

$$\text{RHS of (1)} \quad = \quad \min_{a\in Act} \{T_{\text{Thin}}(s, a), T_{\text{Thick}}(s, a)\}, \tag{2}$$

where $T_{\text{Thin}}(s, a)$ ($T_{\text{Thick}}(s, a)$) is the infimum of the RHS of (1) over all actions (t, a) such that

$[\nu{+}t] \in \mathcal{R}^a_{\text{Thin}}$ ($[\nu{+}t] \in \mathcal{R}^a_{\text{Thick}}$). For the first term we have:

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

where ν^t_C denote the clock valuation $(\nu{+}t)[C{:=}0]$, $t^{(\ell,\zeta)}$ the time to reach the region R from s and ζ^C the region $\zeta[C{:=}0]$. Considering the second term of (2) we have

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

From Lemma 6.2 we have that $P((\ell',\cdot),(\ell',\zeta^C))$ is concave and, from Lemma 2.1, since ν^t_C is an affine mapping and $\delta[\ell,a](C,\ell'){\geq}0$ for all $2^{\mathcal{C}}\times L$, the weighted sum over (C,ℓ') of the functions $P((\ell',\nu^t_C),(\ell',\zeta^C))$ is concave on the domain $\{t\,|\,\nu{+}t\in\zeta\}$. From the concavity assumption of price functions, $\pi(s,(\cdot,a))$ is concave over the domain $\{t\,|\,\nu{+}t\in\zeta\}$, and therefore, again using Lemma 2.1, we have that the function:

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

is concave over $\{t\,|\,\nu{+}t\in\zeta\}$. Therefore using Lemma 2.2 we have $T_{\text{Thick}}(s,a)$ equals

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

Substituting the values of $T_{\text{Thin}}(s,a)$ and $T_{\text{Thick}}(s,a)$ into (2) and observing that for any thin region $(\ell,\zeta)\in\mathcal{R}^a_{\text{Thin}}$ there exist $b\in\mathbb{Z}$ and $c\in C$ such that $\nu{+}(b{-}\nu(c))\in\zeta$, it follows from Definition 5.2 that RHS of (1) equals:

$$\min_{(\alpha,R)\in A_{\text{BRG}}(s,[s])} \left\{ \pi_{\text{BRG}}((s,[s]),(\alpha,R)) + \sum_{(s',R')\in S_{\text{BRG}}} p_{\text{BRG}}((s',R')|(s,[s]),(\alpha,R))\cdot P(s',R') \right\}$$

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

Lemma 6.4. *If* $\langle T_i \rangle_{i=1}^N \models \mathrm{Opt}_T^N(\mathcal{T}_{\text{BRG}})$*, then* $\langle \widetilde{T}_i \rangle_{i=0}^N \models \mathrm{Opt}_T^N(\llbracket \mathcal{T} \rrbracket)$.

Lemma 6.5. *If* $D \models \mathrm{Opt}_D^\lambda(\mathcal{T}_{\text{BRG}})$*, then* $\widetilde{D} \models \mathrm{Opt}_D^\lambda(\llbracket \mathcal{T} \rrbracket)$.

Since it is not known to us whether there exists a solution $(G, B) \models \mathrm{Opt}_A(\mathcal{T}_{\text{BRG}})$ such that both G and B are regionally concave, we can not show that $(G, B) \models \mathrm{Opt}_A(\mathcal{T}_{\text{BRG}})$, implies $(G, B) \models \mathrm{Opt}_A(\llbracket \mathcal{T} \rrbracket)$. Instead, we use the following result to reduce the average price problem on PTA to that over the corresponding boundary region graph.

Lemma 6.6. *If* $(G, B) \models \mathrm{Opt}_A(\mathcal{T}_{\text{BRG}})$*, then* $\widetilde{G} = \mathrm{EAvg}^*_{\llbracket \mathcal{T} \rrbracket}$.

The proof of this result follows from Lemma 6.8 and Corollary 6.10 below.

Lemma 6.7. *For an arbitrary priced MDP* $\mathcal{M} = (S, A, p, \pi)$ *and state* $s \in S$*, the following inequality holds:*

$$\inf_{\sigma \in \Sigma_M} \limsup_{n \to \infty} \frac{1}{n} \cdot \mathbb{E}_s^\sigma \left\{ \textstyle\sum_{i=1}^n \pi(X_{i-1}, Y_i) \right\} \geq \limsup_{n \to \infty} \inf_{\sigma \in \Sigma_M} \frac{1}{n} \cdot \mathbb{E}_s^\sigma \left\{ \textstyle\sum_{i=1}^n \pi(X_{i-1}, Y_i) \right\}$$

Lemma 6.8. *For every state* $s \in S_{\mathsf{T}}$ *we have* $\mathrm{EAvg}^*_{\llbracket \mathcal{T} \rrbracket}(s) \geq \mathrm{EAvg}^*_{\mathcal{T}_{\text{BRG}}}(s, [s])$.

Proof. Consider any $s \in S_{\mathsf{T}}$, using Lemma 6.7 we have:

$$\begin{aligned}
\mathrm{EAvg}^*_{\llbracket \mathcal{T} \rrbracket}(s) \;&\geq\; \limsup_{n \to \infty} \inf_{\sigma \in \Sigma_{\llbracket \mathcal{T} \rrbracket}} \frac{1}{n} \cdot \mathbb{E}_s^\sigma \{ \textstyle\sum_{i=1}^n \pi(X_{i-1}, Y_i) \} \\
&= \limsup_{n \to \infty} \frac{1}{n} \cdot \mathrm{ETotal}^*_{\llbracket \mathcal{T} \rrbracket}(n)(s) && \text{by definition of } \mathrm{ETotal}^*_{\llbracket \mathcal{T} \rrbracket}(n) \\
&\geq \limsup_{n \to \infty} \frac{1}{n} \cdot \mathrm{ETotal}^*_{\mathcal{T}_{\text{BRG}}}(n)((s, [s])) && \text{by Lemma 6.4 and Proposition 3.4} \\
&= \limsup_{n \to \infty} \frac{1}{n} \cdot \inf_{\sigma \in \Sigma_{\mathcal{T}_{\text{BRG}}}} \mathbb{E}_{(s,[s])}^\sigma \{ \textstyle\sum_{i=1}^n \pi(X_{i-1}, Y_i) \} && \text{by definition} \\
&= \inf_{\sigma \in \Sigma_{\mathcal{T}_{\text{BRG}}}} \limsup_{n \to \infty} \frac{1}{n} \cdot \mathbb{E}_{(s,[s])}^\sigma \{ \textstyle\sum_{i=1}^n \pi(X_{i-1}, Y_i) \} && \text{since } [\mathcal{T}_{\text{BRG}}, (s, [s])] \text{ is finite} \\
&= \mathrm{EAvg}^*_{\mathcal{T}_{\text{BRG}}}(s, [s]) && \text{as required}
\end{aligned}$$

□

Using the Lipschitz continuity of price functions and a slight variant of Lemma 3 of [7] we show that the following proposition and corollary hold.

Proposition 6.9. *There exists* $\delta > 0$ *such that for any* $\varepsilon < \delta$, $\sigma \in \Sigma_{\mathcal{T}_{\text{BRG}}}$ *and* $s \in S_{\mathsf{T}}$*, there exists* $\sigma_\varepsilon \in \Sigma_{\llbracket \mathcal{T} \rrbracket}$ *such that* $|\, \mathrm{ETotal}_{\mathcal{T}_{\text{BRG}}}(N)((s, [s]), \sigma) - \mathrm{ETotal}_{\llbracket \mathcal{T} \rrbracket}(N)(s, \sigma_\varepsilon) \,| \leq N \cdot \varepsilon$ *for all* $N \in \mathbb{N}$.

Corollary 6.10. *For every* $\varepsilon > 0$ *and* $s \in S_{\mathsf{T}}$ *we have* $\mathrm{EAvg}^*_{\llbracket \mathcal{T} \rrbracket}(s) \leq \mathrm{EAvg}^*_{\mathcal{T}_{\text{BRG}}}(s, [s]) + \varepsilon$.

7 Complexity

To show EXPTIME-hardness we present a reduction to the EXPTIME-complete problem of solving countdown games [17]. The lemma concerns only expected reachability price problem as similar reductions follow for the other problems.

Lemma 7.1. *The expected reachability problem is EXPTIME-hard for concavely-priced PTA with two or more clocks.*

Proof. Let $\mathcal{G} = (N, M, \pi_\mathcal{G}, n_0, B_0)$ be a countdown game. N is a finite set of nodes; $M \subseteq N \times N$ is a set of moves; $\pi_\mathcal{G} : M \to \mathbb{N}_+$ assigns a positive integer to every move; $(n_0, B_0) \in N \times \mathbb{N}_+$ is the initial configuration. From $(n, B) \in N \times \mathbb{N}_+$, a move consists of player 1 choosing $k \in \mathbb{N}_+$, such that $k \leq B$ and $\pi_\mathcal{G}(n, n') = k$ for some $(n, n') \in M$, then player 2 choosing $(n, n'') \in M$ such that $\pi_\mathcal{G}(n, n'') = k$; the new configuration is $(n'', B - k)$. Player 1 wins if a configuration of the form $(n, 0)$ is reached, and loses when a configuration (n, B) is reached such that $\pi_\mathcal{G}(n, n') > B$ for all $(n, n') \in M$.

Given a countdown game $\mathcal{G}$ we define the PTA $\mathsf{T}_\mathcal{G} = (L, \mathcal{C}, inv, Act, E, \delta)$ where $L = \{\star\} \cup N \cup N_\mathrm{u}$ where $N_\mathrm{u} = \{n_\mathrm{u} \mid n \in N\}$; $\mathcal{C} = \{b, c\}$; $inv(n) = \{\nu \mid 0 \leq \nu(b) \leq B_0 \wedge 0 \leq \nu(c) \leq B_0\}$ and $inv(n_\mathrm{u}) = \{\nu \mid \nu(c) = 0\}$ for any $n \in N$; $Act = \{\star, \mathrm{u}\} \cup \{k \mid \exists m \in M.\, \pi_\mathcal{G}(m) = k\}$; for any $\ell \in L$ and $a \in Act$:

$$E(\ell, a) = \begin{cases} \{\nu \mid \exists n' \in N.\, (\pi_\mathcal{G}(\ell, n') = k \wedge \nu(c) = k)\} & \text{if } \ell \in N \text{ and } a = k \in \mathbb{N}_+ \\ \{\nu \mid \nu(b) = B_0\} & \text{if } \ell \in N_\mathrm{u} \text{ and } a = \star \\ \{\nu \mid \nu(c) = 0\} & \text{if } \ell \in N_\mathrm{u} \text{ and } a = \mathrm{u} \\ \emptyset & \text{otherwise} \end{cases}$$

and for any $n \in N$, $a \in Act(n)$, $C \subseteq \mathcal{C}$ and $\ell' \in L$:

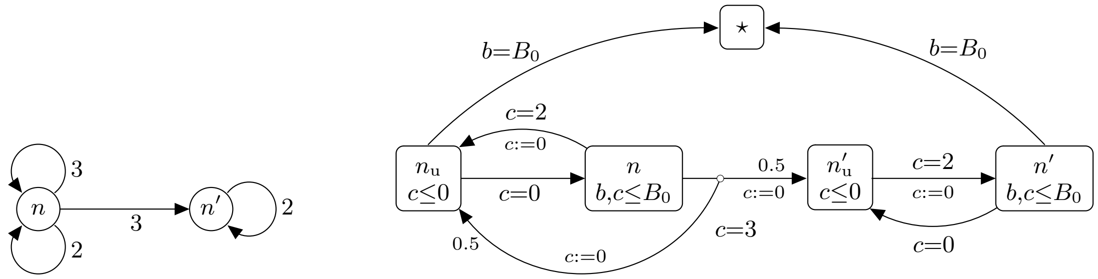
$$\delta(n, a)(C, \ell') = \begin{cases} \frac{1}{|\{n'' \mid \pi_\mathcal{G}(n, n'') = a\}|} & \text{if } a \in \mathbb{N}_+, C = \{c\}, \ell' = n'_\mathrm{u} \text{ and } \pi_\mathcal{G}(n, n') = a \\ 0 & \text{otherwise} \end{cases}$$

$$\delta(n_\mathrm{u}, a)(C, \ell') = \begin{cases} 1 & \text{if } C = \emptyset, \ell' = n \text{ and } a = \mathrm{u} \\ 1 & \text{if } C = \emptyset, \ell' = \star \text{ and } a = \star \\ 0 & \text{otherwise.} \end{cases}$$

An example of a reduction is shown in Figure 1. For the price function $\pi_\mathcal{G}(s, (t, a)) = t$, it routine to verify that the optimal expected reachability price for target $F = \{\star\} \times V$ equals B_0 when starting from $(n_0, (0, 0))$ in the concavely-priced PTA $(\mathsf{T}_\mathcal{G}, \pi_\mathcal{G})$ if and only if player 1 has a winning strategy in the countdown game $\mathcal{G}$. □

On the other hand, we can solve each problem in EXPTIME because:

- we can reduce each problem on probabilistic timed automata to a similar problem on the boundary region graph (see Lemmas 6.3–6.6);
- the boundary region graph has exponential-size and can be constructed in exponential time in the size of the PTA;



(a) Game $\mathcal{G}$ (b) PTA $\mathsf{T}_{\mathcal{G}}$

Figure 1: Countdown Game and the corresponding probabilistic timed automata

- on the boundary region graph (a finite state MDP) we can solve each minimisation problem using a polynomial-time algorithm (see, e.g., [27]) in the size of the graph.

For *one clock* concavely-priced PTAs expected reachability, discounted, and average-price problems are PTIME-hard as these problems are PTIME-complete [26] even for finite MDPs (i.e. PTAs with no clocks). To show PTIME-membership one can adapt the construction of [22]—which shows the NLOGSPACE-membership of the reachability problem for one clock timed automata—to obtain an abstraction similar to the boundary region graph whose size is polynomial in the size of probabilistic timed automata, and then run polynomial-time algorithms to solve this finite MDP.

Theorem 7.2. *The exact complexity of solving expected reachability, discounted and average price problems is EXPTIME-complete for concavely-priced PTA with two or more clocks and PTIME-complete for concavely-priced PTA with one clock.*

8 Conclusion

We presented a probabilistic extension of the boundary region graph originally defined in the case of timed automata for PTAs. We characterize expected total (finite horizon), reachability, discounted and average price using optimality equations. By analysing properties of the solutions of these optimality equations on boundary region graphs, we demonstrated that solutions on the boundary region graph are also solutions to the corresponding optimality equations on the original priced PTA. Using this reduction, we then showed that the exact complexity of solving expected reachability, discounted, and average optimisation problems on concavely-priced PTAs is EXPTIME-complete.

Although the computational complexity is very high, we feel motivated by the success of quantitative analysis tools like UPPAAL [4] and PRISM [14]. We wish to develop more efficient symbolic zone-based algorithms for the problems considered in this paper. Another direction for future work is to consider more involved objectives like price-per-reward average [7, 18] and multi-objective optimisation [25].

Acknowledgments. The authors are supported in part by EPSRC grants EP/D076625 and EP/F001096.

References

[1] R. Alur and D. Dill. A theory of timed automata. In *Theoretical Computer Science*, volume 126, pages 183–235, 1994.

[2] R. Alur, S. La Torre, and G. Pappas. Optimal paths in weighted timed automata. In M. Di Benedetto and A. Sangiovanni-Vincentelli, editors, *Proc. 4th Int Workshop Hybrid Systems: Computation and Control (HSCC'01)*, volume 2034 of *LNCS*, pages 49–62. Springer, 2001.

[3] D. Beauquier. Probabilistic timed automata. *Theoretical Computer Science*, 292(1):65–84, 2003.

[4] G. Behrmann, A. David, K. Larsen, O. Möller, P. Pettersson, and W. Yi. UPPAAL - present and future. In *Proc. 40th IEEE Conf. Decision and Control (CDC'01)*, pages 2881–2886. IEEE CS Press, 2001.

[5] J. Berendsen, T. Chen, and D. Jansen. Undecidability of cost-bounded reachability in priced probabilistic timed automata. In J. Chen and S. Cooper, editors, *Proc. 6th Conf. Theory and Applications of Models of Computation, (TAMC'09)*, volume 5532 of *LNCS*, pages 128–137, 2009.

[6] J. Berendsen, D. Jansen, and J-P. Katoen. Probably on time and within budget - on reachability in priced probabilistic timed automata. In *Proc. 3rd Int. Conf. Quantitative Evaluation of Systems (QEST'06)*, pages 311–322. IEEE CS Press, 2006.

[7] P. Bouyer, E. Brinksma, and K. G. Larsen. Optimal infinite scheduling for multi-priced timed automata. *Formal Methods in System Design*, 32(1):3–23, 2008.

[8] S. Boyd and L. Vandenberghe. *Convex Optimization*. CUP, 2004.

[9] L. de Alfaro. Computing minimum and maximum reachability times in probabilistic systems. In J. Baeten and S. Mauw, editors, *Proc. 10th Int. Conf. Concurrency Theory (CONCUR'99)*, volume 1664 of *LNCS*, pages 66–81. Springer, 1999.

[10] E. Dynkin and A. Yushkevich. *Controlled Markov Processes*. Springer, 1979.

[11] J. E. Falk and J. L. Horowitz. Critical path problems with concave cost-time curves. *Management Science*, 19(4):446–455, 1972.

[12] Y. Feng and B. Xiao. A Continuous-Time Yield Management Model with Multiple Prices and Reversible Price Changes. *Management Science*, 46(5):644–657, 2000.

[13] T. Henzinger, Z. Manna, and A. Pnueli. What good are digital clocks? In W. Kuich, editor, *Proc. 19th Int. Colloq. Automata, Languages and Programming (ICALP'92)*, volume 623 of *LNCS*, pages 545–558. Springer, 1992.

[14] A. Hinton, M. Kwiatkowska, G. Norman, and D. Parker. PRISM: A tool for automatic verification of probabilistic systems. In H. Hermanns and J. Palsberg, editors, *Proc. 12th Int. Conf. Tools and Algorithms for the Construction and Analysis of Systems (TACAS'06)*, volume 3920 of *LNCS*, pages 441–444. Springer, 2006.

[15] H. Jensen. Model checking probabilistic real time systems. In B. Bjerner, M. Larsson, and B. Nordström, editors, *Proc. 7th Nordic Workshop on Programming Theory*, Report 86:247–261. Chalmers University of Technology, 1996.

[16] M. Jurdzinski, M. Kwiatkowska, G. Norman, and A. Trivedi. Concavely-priced probabilistic timed automata. In M. Bravetti and G. Zavattaro, editors, *Proc. 20th Int. Conf. Concurrency Theory (CONCUR'09)*, volume 5710 of *LNCS*, pages 415–430. Springer, 2009.

[17] M. Jurdziński, J. Sproston, and F. Laroussinie. Model checking probabilistic timed automata with one or two clocks. *Logical Methods in Computer Science*, 4(3):12, 2008.

[18] M. Jurdziński and A. Trivedi. Concavely-priced timed automata. In F. Cassez and C. Jard, editors, *Proc. 6th Int. Conf. Formal Modelling and Analysis of Timed Systems (FORMATS'08)*, volume 5215 of *LNCS*, pages 48–62. Springer, 2008.

[19] M. Kwiatkowska, G. Norman, D. Parker, and J. Sproston. Performance analysis of probabilistic timed automata using digital clocks. *Formal Methods in System Design*, 29:33–78, 2006.

[20] M. Kwiatkowska, G. Norman, R. Segala, and J. Sproston. Automatic verification of real-time systems with discrete probability distributions. *Theoretical Computer Science*, 282:101–150, 2002.

[21] M. Kwiatkowska, G. Norman, J. Sproston, and F. Wang. Symbolic model checking for probabilistic timed automata. *Information and Computation*, 205(7):1027–1077, 2007.

[22] F. Laroussinie, N. Markey, and P. Schnoebelen. Model checking timed automata with one or two clocks. In Philippa Gardner and Nobuko Yoshida, editors, *Proc. 15th Int. Conf. Concurrency Theory (CONCUR'04)*, volume 3170 of *LNCS*, pages 387–401. Springer, 2004.

[23] F. Laroussinie and J. Sproston. State explosion in almost-sure probabilistic reachability. *Information Processing Letters*, 102(6):236–241, 2007.

[24] K. G. Larsen, G. Behrmann, E. Brinksma, A. Fehnker, T. Hune, P. Pettersson, and J. Romijn. As cheap as possible: Efficient cost-optimal reachability for priced timed automata. In G. Berry, H. Comon, and A. Finkel, editors, *Proc. 16th Int. Conf. Computer Aided Verification (CAV'01)*, volume 2102 of *LNCS*, pages 493–505. Springer, 2001.

[25] K. G. Larsen and J. I. Rasmussen. Optimal reachability for multi-priced timed automata. *Theoretical Computer Science*, 390(2-3):197–213, 2008.

[26] C. H. Papadimitriou and J. N. Tsitsiklis. The complexity of Markov decision processes. *Mathematics of Operations Research*, 12:441–450, 1987.

[27] M. L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley, 1994.

[28] E. Torres, J.S. Dominnguez, L. Valdes, and R. Aza. Passenger waiting time in an airport and expenditure carried out in the commercial area. *Journal of Air Transport Management*, 11(6):363–367, 2005.

[29] S. Tripakis. Verifying progress in timed systems. In J.-P. Katoen, editor, *Proc. 5th Int. AMAST Workshop on Real-Time and Probabilistic Systems (ARTS'99)*, volume 1601 of *LNCS*, pages 299–314. Springer, 1999.

A Proof of Proposition 3.4

The proof of Proposition 3.4 follows from Propositions A.1-A.4 given below in which we assume a fixed priced MDP $\mathcal{M}$.

Proposition A.1. *If* $P{\models}\mathrm{Opt}_R^F(\mathcal{M})$*, then* $P(s) = \mathrm{EReach}^*_{\mathcal{M}}(F)(s)$ *for all* $s \in S$.

Proof. The proof is in two parts. In the first part we show that $P(s) \leq \mathrm{EReach}_{\mathcal{M}}(F)(s,\sigma)$ for all $\sigma \in \Sigma_{\mathcal{M}}$ and $s \in S \setminus F$, while the second part we show that for any $\varepsilon{>}0$ there exists a pure strategy $\sigma_\varepsilon \in \Sigma_M$ such that $P(s) \geq \mathrm{EReach}_{\mathcal{M}}(F)(s,\sigma){+}\varepsilon$ for all $s \in S \setminus F$.

1. Let $\sigma \in \Sigma_{\mathcal{M}}$. Since $P \models \mathrm{Opt}_R^F(\mathcal{M})$, for any $s \in S \setminus F$ and $a \in A(s)$, we have

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

   and hence, using Assumption 3.2, it follows by induction that for any $n \geq 1$:

   $$P(s) \leq \mathbb{E}_s^\sigma \left\{ \textstyle\sum_{i=1}^{\min\{i \,|\, X_i \in F\}} \pi(X_{i-1},Y_i) \right\} + \textstyle\sum_{s' \in S \setminus F} \mathrm{Prob}_s^\sigma(X_n{=}s'){\cdot}P(s') \qquad (3)$$

   From Assumption 3.3 we have that

   $$\lim_{n\to\infty} \textstyle\sum_{s' \in S \setminus F} \mathrm{Prob}_s^\sigma(X_n{=}s') = 0\,.$$

   Therefore, taking the limit in (3) we get the desired inequality:

   $$\begin{array}{rcl} P(s) & \leq & \mathbb{E}_s^\sigma \left\{ \sum_{i=1}^{\min\{i \,|\, X_i \in F\}} \pi(X_{i-1},Y_i) \right\} \\ & = & \mathrm{EReach}_{\mathcal{M}}(F)(s,\sigma)\,. \end{array}$$

2. Consider any $\varepsilon > 0$ and let σ_ε be the pure strategy where if r is a run of length n, then $\sigma_\varepsilon(r) \in \mathit{last}(r)$ is such that

   $$\pi(\mathit{last}(r),\sigma_\varepsilon(r)) + \textstyle\sum_{s' \in S} p(s'|\mathit{last}(r),\sigma_\varepsilon(r)){\cdot}P(s') \leq P(\mathit{last}(r)) + \varepsilon/2^{n+1}.$$

   Using a similar analysis to the first part, it is straightforward to verify that

   $$P(s) \geq \mathrm{EReach}_{\mathcal{M}}(F)(s,\sigma_\varepsilon) + \varepsilon$$

   which completes the proof.

□

noindent The proof of the following two propositions follow similarly to that of Proposition A.1, and hence their proofs are omitted.

Proposition A.2. *If* $\langle T_i\rangle_{i=0}^{N} \models \mathrm{Opt}_R^N(\mathcal{M})$*, then* $T_i(s) = \mathrm{ETotal}^*_{\mathcal{M}}(i)(s)$ *for all* $i \le N$ *and* $s \in S$.

Proposition A.3. *If* $D \models \mathrm{Opt}_D^{\lambda}(\mathcal{M})$ *and* D *is bounded, then* $D(s) = \mathrm{EDisct}^*_{\mathcal{M}}(\lambda)(s)$ *for all* $s \in S$.

Proposition A.4. *If* $(G, B) \models \mathrm{Opt}_A(\mathcal{M})$ *and* G, B *are bounded, then* $G(s) = \mathrm{EAvg}^*_{\mathcal{M}}(s)$ *for all* $s \in S$.

Proof. The proof is in two parts. In the first part we show that $G(s) \le \mathrm{EAvg}_{\mathcal{M}}(s, \sigma)$ for all $\sigma \in \Sigma_{\mathcal{M}}$ and $s \in S$, while in the second part we show that for $\varepsilon > 0$ there exists a pure strategy $\sigma_\varepsilon \in \Sigma_M$ such that $G(s) \ge \mathrm{EAvg}_{\mathcal{M}}(s, \sigma) + \varepsilon$ for all $s \in S$.

1. Let $\sigma \in \Sigma_{\mathcal{M}}$. Since $(G, B) \models \mathrm{Opt}_A(\mathcal{M})$, for any $s \in S$ and $a \in A(s)$, we have that

$$G(s) \le \textstyle\sum_{s' \in S} p(s'|s, a) \cdot G(s')$$

By induction, it follows that $G(s) \le \mathbb{E}_s^{\sigma}\{G(X_n)\}$ for all $n \ge 0$. Since $(G, B) \models \mathrm{Opt}_A(\mathcal{M})$ we also have that

$$B(s) \le \pi(s, a) - G(s) + \textstyle\sum_{s' \in S} p(s'|s, a) \cdot B(s')$$

for all $s \in S$ and $a \in A(s)$. By induction it follows that for any $n \ge 1$

$$B(s) \le \mathbb{E}_s^{\sigma}\left\{\textstyle\sum_{i=1}^{n} \pi(X_{i-1}, Y_i)\right\} - \mathbb{E}_s^{\sigma}\left\{\textstyle\sum_{i=1}^{n} G(X_{i-1})\right\} + \mathbb{E}_s^{\sigma}(B(X_n)).$$

Substiuting the fact that $G(s) \le \mathbb{E}_s^{\sigma}\{G(X_n)\}$ for any $n \ge 1$ we have:

$$\begin{aligned} B(s) &\le \mathbb{E}_s^{\sigma}\left\{\textstyle\sum_{i=1}^{n} \pi(X_{i-1}, Y_i)\right\} - \mathbb{E}_s^{\sigma}\left\{\textstyle\sum_{i=1}^{n} G(s)\right\} + \mathbb{E}_s^{\sigma}\{B(X_n)\} \\ &= \mathbb{E}_s^{\sigma}\left\{\textstyle\sum_{i=1}^{n} \pi(X_{i-1}, Y_i)\right\} - \textstyle\sum_{i=1}^{n} G(s) + \mathbb{E}_s^{\sigma}\{B(X_n)\} \end{aligned}$$

and therefore, rearranging terms:

$$G(s) \le \tfrac{1}{n} \cdot \mathbb{E}_s^{\sigma}\left\{\textstyle\sum_{i=1}^{n} \pi(X_{i-1}, Y_i)\right\} + \tfrac{1}{n} \cdot \left(\mathbb{E}_s^{\sigma}\{B(X_n)\} - B(s)\right)$$

Since B is bounded, the second term on the right hand side vanishes as $n \to \infty$, and hence taking limit supremum both sides:

$$G(s) \le \limsup_{n \to \infty} \left(\tfrac{1}{n} \cdot \mathbb{E}_s^{\sigma}\left\{\textstyle\sum_{i=1}^{n} \pi(X_{i-1}, Y_i)\right\}\right) = \mathrm{EAvg}_{\mathcal{M}}(s, \sigma).$$

Consider any $\varepsilon > 0$. Let σ_ε by the pure strategy where if r is a run of length n, then

$$\begin{aligned} \textstyle\sum_{s' \in S} p(s'|\mathit{last}(r), \sigma_\varepsilon(r)) \cdot G(s') &\le G(\mathit{last}(r)) + \tfrac{\varepsilon}{2^{n+1}} \\ \pi(\mathit{last}(r), \sigma_\varepsilon(r)) - G(\mathit{last}(r)) + \textstyle\sum_{s' \in S} p(s'|\mathit{last}(r), \sigma_\varepsilon(r)) \cdot B(s') &\le B(\mathit{last}(r)) + \tfrac{\varepsilon}{2^{n+1}} \end{aligned}$$

Using a similar analysis to the first part of the proof, it is straightforward to verify that $G(s) \ge \mathrm{EAvg}_{\mathcal{M}}(s, \sigma_\varepsilon) + \varepsilon$ completing the proof.

□

B Proof of Proposition 5.3

Proof. Let $\mathcal{T}_{\text{BRG}}$ be a boundary region graph $\mathcal{T}_{\text{BRG}}$ and $s \in S_{\text{BRG}}$. We define the fractional signature of a state $s{=}((\ell,\nu),(\ell,\zeta)) \in S_{\text{BRG}}$ of a boundary region graph as $\wr s \int = \wr \nu \int$.

It is easy to verify that starting from a state $s \in S_{\text{BRG}}$ and taking only boundary transitions it is impossible to reach any state whose fractional signature is neither a subsequence, or shift of a subsequence of the fractional signature of s. Therefore the *reachable sub-graph* of $\mathcal{T}_{\text{BRG}}$ from a given initial state $s \in S_{\text{BRG}}$ is the finite MDP $\mathcal{T}^s_{\text{BRG}} = (S^s_{\text{BRG}}, A_{\text{BRG}}, p)$ where S^s_{BRG} is the finite set of states such that $s' \in S'_{\text{BRG}}$ if $\wr s' \int$ is either a subsequence or a k-shift of a subsequence of $\wr s \int$. $\square$

C Proof of Lemma 6.2

Proof. The proof is in two parts. First we show that if $\langle T_i \rangle_{i=0}^N \models \text{Opt}^N_T(\mathcal{T}_{\text{BRG}})$, then T_i is regionally concave for all $i{\le}N$. In the second part, we use this result to show that if $P \models \text{Opt}^F_R(\mathcal{T}_{\text{BRG}})$ and $D \models \text{Opt}^\lambda_D(\mathcal{T}_{\text{BRG}})$, then P and D are regionally concave.

- Consider any $N \in \mathbb{N}$. We now prove by induction that T_i is regionally concave for all $i{\le}N$. The base case is trivial as $T_0(s,R) = 0$ for every state $(s,R) \in S_{\text{BRG}}$. For the inductive step we assume that T_i is regionally concave and show that for any region $R = (\ell,\zeta) \in \mathcal{R}$ the function $T_{i+1}(\cdot,R) : \overline{R} \to \mathbb{R}$ is concave. Therefore consider any $R = (\ell,\zeta) \in \mathcal{R}$, by definition for any $(s,R) = ((\ell,\nu),(\ell,\zeta)) \in S_{\text{BRG}}$:

$$T_{i+1}(s,R) = \min_{(\alpha,R') \in A_{\mathcal{T}_{\text{BRG}}}(s,R)} \left\{ \pi_{\text{BRG}}((s,R),(\alpha,R')) + T^\oplus_i((s,R),(\alpha,R')) \right\}$$

  and $T^\oplus_i((s,R),((b,c,a),R))$ equals

$$\sum_{(C,\ell') \in 2^{\mathcal{C}} \times L} \delta[\ell,a](C,\ell') \cdot T_i((\ell,\nu{+}b{-}\nu(c)[C{:=}0]),(\ell',\zeta[C{:=}0]))\,.$$

  Now the concavity of $T^\oplus_i$ on the domain $\overline{R}$ follows from the following steps:

  - by induction T_i is regionally concave;
  - $T_i((\ell,\nu{+}b{-}\nu(c)[C{:=}0]),(\ell',\zeta[C{:=}0]))$ is concave on the domain $\overline{R}$ for fix (ℓ',C), as function composition of a concave and a linear function is concave (Lemma 2.1);
  - $\sum_{(C,\ell') \in 2^{\mathcal{C}} \times L} \delta[\ell,a](C,\ell') \cdot T_i((\ell,\nu{+}b - \nu(c)[C{:=}0]),(\ell',\zeta[C{:=}0]))$ is concave on the domain $\overline{R}$ as it is a nonnegative (probability) weighted sum of concave functions (again from Lemma 2.1).

  Now by definition of a concavely-priced PTA, the function $\pi((s,R),(\alpha,R'))$ is concave on the domain $\overline{R}$. Therefore, again invoking Lemma 2.1, we have that every $(\alpha,R') \in A_{\mathcal{T}_{\text{BRG}}}(s,R)$ the sum of the concave functions $\pi_{\text{BRG}}((s,R),(\alpha,R'))$ and $T^\oplus_i((s,R),(\alpha,R'))$ is concave on the domain $\overline{R}$. Which implies that $T_{i+1}(\cdot,R) : \overline{R} \to \mathbb{R}$ is concave as it is the minimum of a set of concave functions (Lemma 2.1).

- To show the regional concavity of P and D we demonstrate that they are both limits of regionally concave functions and the fact that the limits of concave functions is concave (Lemma 2.1). Considering P we have for any $(s, R) \in S_{\text{BRG}}$:

$$\begin{aligned}
P(s,R) &= \text{EReach}^*_{\mathcal{T}_{\text{BRG}}}(F)(s,R) && \text{by Proposition 3.4} \\
&= \inf_{\sigma\in\Sigma} \mathbb{E}^\sigma_s \left\{ \textstyle\sum_{i=1}^{\min\{i \mid X_i \in F\}} \pi(X_{i-1}, Y_i) \right\} && \text{by definition} \\
&= \inf_{\sigma\in\Sigma} \lim_{N\to\infty} \mathbb{E}^\sigma_s \left\{ \textstyle\sum_{i=1}^{N} \pi(X_{i-1}, Y_i) \right\} && \text{by Assumption 3.2} \\
&= \lim_{N\to\infty} \inf_{\sigma\in\Sigma} \mathbb{E}^\sigma_s \left\{ \textstyle\sum_{i=1}^{N} \pi(X_{i-1}, Y_i) \right\} && \text{since } \mathcal{T}^s_{\text{BRG}} \text{ is a finite MDP} \\
&= \lim_{N\to\infty} \text{ETotal}^*_{\mathcal{T}_{\text{BRG}}}(N)((s,R)) && \text{by definition} \\
&= \lim_{N\to\infty} T_N(s,R) && \text{by Proposition 3.4.}
\end{aligned}$$

  Note that the limit exists in the fourth equality due to Assumption 3.3. The proof for $D \models \text{Opt}^\lambda_D(\mathcal{T}_{\text{BRG}})$ follows similarly.

□

D Proof of Lemma 6.7

Proof. For every $\sigma' \in \Sigma_M$ and for every $n \in \mathbb{N}$, we have that

$$\frac{1}{n} \cdot \mathbb{E}^{\sigma'}_s \Big\{ \textstyle\sum_{i=1}^n \pi(X_{i-1}, Y_i) \Big\} \geq \inf_{\sigma\in\Sigma_M} \frac{1}{n} \cdot \mathbb{E}^{\sigma}_s \Big\{ \textstyle\sum_{i=1}^n \pi(X_{i-1}, Y_i) \Big\}$$

Taking limit supremum both sides, we have that

$$\limsup_{n\to\infty} \frac{1}{n} \cdot \mathbb{E}^{\sigma'}_s \Big\{ \textstyle\sum_{i=1}^n \pi(X_{i-1}, Y_i) \Big\} \geq \limsup_{n\to\infty} \inf_{\sigma\in\Sigma_M} \frac{1}{n} \cdot \mathbb{E}^{\sigma}_s \Big\{ \textstyle\sum_{i=1}^n \pi(X_{i-1}, Y_i) \Big\}$$

Since the previous inequality holds for arbitrary σ' we have that

$$\inf_{\sigma\in\Sigma_M} \limsup_{n\to\infty} \frac{1}{n} \cdot \mathbb{E}^{\sigma}_s \Big\{ \textstyle\sum_{i=1}^n \pi(X_{i-1}, Y_i) \Big\} \geq \limsup_{n\to\infty} \inf_{\sigma\in\Sigma_M} \frac{1}{n} \cdot \mathbb{E}^{\sigma}_s \Big\{ \textstyle\sum_{i=1}^n \pi(X_{i-1}, Y_i) \Big\}$$

as required. □

E Proof of Proposition 6.9

Before we give the proof of Proposition 6.9 we require the following definitions and results.

Definition E.1. *Let $\nu \in V$ be a clock valuation and $\wr\nu\int = \{f_0, f_1, \ldots f_n\}$ its fractional signature. The* strait *of ν is defined as* $\mathsf{strait}(\nu) = \min\{1-f_n, f_n-f_{n-1}, \ldots, f_1-f_0\}$.

For any clock valuations $\nu, \nu' \in V$, we define a measure of the *scatter* of ν' around ν as follows.

Definition E.2. *Let $\nu, \nu' \in V$ be clock valuations. The scatter of ν' around ν equals*

$$\mathsf{scatter}(\nu, \nu') = \mathsf{left\_scatter}(\nu, \nu') + \mathsf{right\_scatter}(\nu, \nu')$$

where $\mathsf{left\_scatter}(\nu, \nu')$ *and* $\mathsf{right\_scatter}(\nu, \nu')$ *denote the left and right scatter of ν' around ν and are defined as follows:*

$$\begin{array}{rcl} \mathsf{left\_scatter}(\nu, \nu') & = & \max\{\nu(c) - \nu'(c) \mid c \in C \wedge \nu(c) \geq \nu'(c)\} \\ \mathsf{right\_scatter}(\nu, \nu') & = & \max\{\nu'(c) - \nu(c) \mid c \in C \wedge \nu(c) \leq \nu'(c)\} \end{array}$$

Notice that $\|\nu - \nu'\|_\infty \leq \mathsf{scatter}(\nu, \nu') \leq 2 \cdot \|\nu - \nu'\|_\infty$.

Lemma E.3. *If $((\ell, \nu), R)$ and $((b, c, a), R')$ are a state and action of $\mathcal{T}_{\text{BRG}}$ and $\varepsilon{>}0$ such that $\varepsilon{<}\mathsf{strait}(\nu)/2$, then $\varepsilon{<}\mathsf{strait}((\nu{+}t)[C{:=}0])/2$ for all $C \subseteq \mathcal{C}$ where $t = b{-}\nu(c)$.*

Proposition E.4. *Let $((\ell, \nu), R)$ be a state of the boundary region graph $\mathcal{T}_{\text{BRG}}$ and $\varepsilon{>}0$ such that $\varepsilon{<}\mathsf{strait}(\nu)/2$. If (ℓ, ν') is a state of $[\![\mathsf{T}]\!]$ and $\mathsf{scatter}(\nu, \nu'){<}\varepsilon$, then for any action $((b, c, a), R') \in A_{\text{BRG}}((\ell, \nu), R)$ there exists an action $(t', a) \in A_{\mathsf{T}}(\ell, \nu')$, such that $\nu'{+}t' \in R'$, $|t{-}t'| < \varepsilon$ and $\mathsf{scatter}((\nu{+}t), (\nu'{+}t')) < \varepsilon$ where $t = b{-}\nu(c)$.*

Proof. Consider any state $s{=}((\ell, \nu), R)$ of $\mathcal{T}_{\text{BRG}}$, action $((b, c, a), R') \in A_{\text{BRG}}((\ell, \nu), R)$, state of $s'{=}(\ell, \nu')$ of $[\![\mathsf{T}]\!]$ and $\varepsilon{>}0$ such that $\varepsilon{<}\mathsf{strait}(\nu)/2$ and $\mathsf{scatter}(\nu, \nu'){<}\varepsilon$. To proof the proposition, by definition of $\mathcal{T}_{\text{BRG}}$ and $[\![\mathsf{T}]\!]$, it is sufficient to show show that there exists $t' \in \mathbb{R}$ such that $\nu'{+}t' \in R'$ and $\mathsf{scatter}((\nu{+}t), (\nu'{+}t')){<}\varepsilon$ where $t = b - \nu(c)$. We claim that if

$$t' = \begin{cases} b{-}\nu'(c) & \text{if } [s'{+}t(s', \alpha)] = R' \\ b{-}\nu'(c){+}d & \text{if } [s'{+}t(s', \alpha)] \rightarrow_{+1} R' \\ b{-}\nu'(c){-}d & \text{if } [s'{+}t(s', \alpha)] \leftarrow_{+1} R' \end{cases}$$

where $d < \min\{\varepsilon{-}\mathsf{scatter}(\nu, \nu'), \mathsf{strait}(\nu')\}$, then $\mathsf{scatter}((\nu{+}t), (\nu'{+}t')){<}\varepsilon$ and $\nu'{+}t' \in R'$ as required. The idea behind such a value of d is that if $d \geq \mathsf{strait}(\nu')$ then it is possible that $\nu'{+}t' \notin R'$, and if $d \geq \varepsilon{-}\mathsf{scatter}(\nu, \nu')$ then it is possible that $\mathsf{scatter}((\nu{+}t), (\nu'{+}t')) \geq \varepsilon$. Also observe that $|t - t'| < \varepsilon$. We prove our claim for the case when $\nu'(c) < \nu(c)$. The other two cases, i.e., when $\nu(c) > \nu'(c)$ and $\nu(c) = \nu'(c)$ are analogous and hence omitted. Since $((b, c, a), R') \in A_{\text{BRG}}((\ell, \nu), R)$ and $d{<}\mathsf{strait}(\nu')$, it follows that $v'{+}t' \in R'$ and it therefore remains to show that $\mathsf{scatter}((\nu{+}t), (\nu'{+}t')){<}\varepsilon$.

Therefore assume $\nu'(c) < \nu(c)$, we first consider when $\nu(c){-}\nu'(c) = \mathsf{left\_scatter}(\nu, \nu')$ and then the case when $\nu(c){-}\nu'(c) < \mathsf{left\_scatter}(\nu, \nu')$.

- If $\nu(c){-}\nu'(c) = \mathsf{left\_scatter}(\nu, \nu')$, then there are three subcases, either $[s'{+}t(s', \alpha)] = R'$, $[s'{+}t(s', \alpha)] \leftarrow_{+1} R'$ or $[s'{+}t(s', \alpha)] \rightarrow_{+1} R'$. We consider only the subcase when $[s'{+}t(s', \alpha)] \leftarrow_{+1} R'$ as the treatment for the other cases is similar. By construction:

$$\begin{array}{rll} t' = & b - \nu'(c) - d & \\ = & b - (\nu(c) - \mathsf{left\_scatter}(\nu, \nu')) - d & \text{since } \nu(c){-}\nu'(c) = \mathsf{left\_scatter}(\nu, \nu') \\ = & (b - \nu(c)) + (\mathsf{left\_scatter}(\nu, \nu') - d) & \text{rearranging} \\ = & t + (\mathsf{left\_scatter}(\nu, \nu') - d) & \text{by definition of } t. \end{array}$$

Now, for any $c' \in \mathcal{C}$ we have:

$$
\begin{aligned}
(\nu + t)(c') - (\nu' + t')(c') &= \nu(c') + t - \nu'(c') - t' \\
&= (\nu(c') - \nu'(c')) + (t - t') && \text{rearranging} \\
&= (\nu(c') - \nu'(c')) + (t - (t + (\mathsf{left\_scatter}(\nu, \nu') - d))) && \text{from above} \\
&= (\nu(c') - \nu'(c')) - (\mathsf{left\_scatter}(\nu, \nu') - d) && \text{rearranging} \qquad (4)
\end{aligned}
$$

and similarly:

$$
(\nu' + t')(c') - (\nu + t)(c') = (\nu'(c') - \nu(c')) + (\mathsf{left\_scatter}(\nu, \nu') - d)\,. \qquad (5)
$$

Now, from (4) it follows that $\mathsf{left\_scatter}(\nu{+}t, \nu'{+}t'){=}d$ and if $\mathsf{right\_scatter}(\nu, \nu'){\neq}0$, then using (5) we have:

$$
\mathsf{right\_scatter}(\nu{+}t, \nu'{+}t) = \mathsf{right\_scatter}(\nu, \nu') + \mathsf{left\_scatter}(\nu, \nu') - d\,.
$$

On the other hand, if $\mathsf{right\_scatter}(\nu, \nu'){=}0$, then again using (5) we have:

$$
\mathsf{right\_scatter}(\nu{+}t, \nu'{+}t) = (\mathsf{left\_scatter}(\nu, \nu') - d) - \min_{c' \in C}\{\nu(c'){-}\nu'(c')\}\,.
$$

In either case, it is straightforward to verify that $\mathsf{scatter}(\nu{+}t, \nu'{+}t') \leq \mathsf{scatter}(\nu, \nu') < \varepsilon$ as required.

- If $\nu(c){-}\nu'(c) < \mathsf{left\_scatter}(\nu, \nu')$, then again there are three subcases: $[s'{+}t(s', \alpha)] = R'$, $[s'{+}t(s', \alpha)] \leftarrow_{+1} R'$, and $[s'{+}t(s', \alpha)] \rightarrow_{+1} R'$. In this case, we consider the subcase when $[s'{+}t(s', \alpha)] \rightarrow_{+1} R'$. The treatment for the other two cases is similar. By construction we have:

$$
\begin{aligned}
t' &= b - \nu'(c) + d \\
&= (b - \nu(c)) - (\nu(c) - \nu'(c)) + d && \text{rearranging} \\
&= t - (\nu(c) - \nu'(c)) + d && \text{by definition of } t.
\end{aligned}
$$

It then follows that:

$$
\mathsf{left\_scatter}(\nu{+}t, \nu'{+}t') = \mathsf{left\_scatter}(\nu, \nu') - (\nu(c) - \nu'(c)) - d
$$

and if $\mathsf{right\_scatter}(\nu, \nu'){\neq}0$, then

$$
\mathsf{right\_scatter}(\nu{+}t, \nu'{+}t) = \mathsf{right\_scatter}(\nu, \nu') + (\nu(c) - \nu'(c)) + d\,.
$$

On the other hand, if $\mathsf{right\_scatter}(\nu, \nu') = 0$, then

$$
\mathsf{right\_scatter}(\nu{+}t, \nu'{+}t) = \max\left\{0, (\nu(c){-}\nu'(c)) + d - \min_{c \in C}\{\nu(c){-}\nu'(c)\}\right\}.
$$

Again it is straightforward to verify that $\mathsf{scatter}(\nu{+}t, \nu'{+}t') \leq \mathsf{scatter}(\nu, \nu')$.

This completes the proof. □

The following result follows directly from Proposition E.4 above.

Corollary E.5. *Let* $((\ell,\nu),R)$ *be a state of the boundary region graph* $\mathcal{T}_{\text{BRG}}$ *and* $0{<}\varepsilon{<}\mathsf{strait}(\nu)/2$. *If* (ℓ,ν') *is a state of* $[\![\mathsf{T}]\!]$ *and* $\mathsf{scatter}(\nu,\nu') < \varepsilon$, *then for any* $((b,c,a),R') \in A_{\text{BRG}}((\ell,\nu),R)$, *there exists* $(t',a) \in A_{\mathsf{T}}(\ell,\nu')$, *such that* $\nu'{+}t' \in R'$, $|t{-}t'| < \varepsilon$ *and for every* $C \subseteq \mathcal{C}$, *we have:*

$$\mathsf{scatter}\big(\, (\nu{+}t)[C{:=}0]\, ,\, (\nu'{+}t')[C{:=}0]\,\big) < \varepsilon\,.$$

where $t = b - \nu(c)$.

Before we give the final result required to prove Proposition 6.9 we require the following definition.

Definition E.6. *For any priced MDP* $\mathcal{M} = (S,A,p,\pi)$, *strategy* σ *of* $\mathcal{M}$ *and finite run* r, *let* $\sigma[r]$ *be the strategy such that for any finite run* r' *of* $\mathcal{M}$*:*

$$\sigma[r](r') \stackrel{\text{def}}{=} \begin{cases} \sigma\langle r,a,r''\rangle & \textit{if } r' \textit{ is of the form } \langle \textit{last}(r),a,r''\rangle \\ \sigma(r') & \textit{otherwise}. \end{cases}$$

Intuitively, the strategy $\sigma[r]$ acts essentially as σ assuming that the run r has already occurred.

Proposition E.7. *Suppose that the price function* π *is* K*-Lipschitz continuous. If* σ_{BRG} *and* $((\ell,\nu),R)$ *are a strategy and state of* $\mathcal{T}_{\text{BRG}}$, $0{<}\varepsilon{<}\mathsf{strait}(\nu)/2$ *and* (ℓ,ν') *is a state of* $[\![\mathsf{T}]\!]$ *such that* $\mathsf{scatter}(\nu,\nu'){<}\varepsilon$, *then there exists a strategy* $\sigma_{\mathcal{T}}$ *of* $[\![\mathsf{T}]\!]$ *such that for any* $N \in \mathbb{N}$*:*

$$\big|\, \mathrm{ETotal}_{\mathcal{T}_{\text{BRG}}}(N)(((\ell,\nu),R),\sigma_{\text{BRG}}) - \mathrm{ETotal}_{[\![\mathcal{T}]\!]}(N)((\ell,\nu'),\sigma_{\mathcal{T}})\,\big| < K{\cdot}N{\cdot}\varepsilon\,.$$

Proof. Consider any strategy σ_{BRG} and state $((\ell,\nu),R)$ of the boundary region graph $\mathcal{T}_{\text{BRG}}$ and $\varepsilon{>}0$ such that $\varepsilon{<}\mathsf{strait}(\nu)/2$. To prove that the proposition holds we will prove by induction on M that for any state (ℓ,ν') of $[\![\mathsf{T}]\!]$ such that $\mathsf{scatter}(\nu,\nu'){<}\varepsilon$ there exists a strategy $\sigma_{\mathcal{T}}^{M}$ such that

$$\big|\, \mathrm{ETotal}_{\mathcal{T}_{\text{BRG}}}(N)(((\ell,\nu),R),\sigma_{\text{BRG}}) - \mathrm{ETotal}_{[\![\mathcal{T}]\!]}(N)((\ell,\nu'),\sigma_{\mathcal{T}}^{M})\,\big| < \varepsilon$$

for all $N \leq M$. If $M = 0$, then for any state (ℓ,ν') of $[\![\mathsf{T}]\!]$ such that $\mathsf{scatter}(\nu,\nu') < \varepsilon$ and strategy $\sigma_{\mathcal{T}}^{M}$ by definition of ETotal we have:

$$\big|\, \mathrm{ETotal}_{\mathcal{T}_{\text{BRG}}}(N)(((\ell,\nu),R),\sigma_{\text{BRG}}) - \mathrm{ETotal}_{[\![\mathcal{T}]\!]}(N)((\ell,\nu'),\sigma_{\mathcal{T}}^{M})\,\big| = |0{-}0| = 0 \leq \varepsilon$$

as required.

Now suppose that the result holds for some M and consider any $N \leq M{+}1$ and state (ℓ,ν') of $[\![\mathsf{T}]\!]$ such that $\mathsf{scatter}(\nu,\nu'){<}\varepsilon$. The case when $N{=}0$ follows as above and therefore for the remainder of the proof we restrict attention to the case when $N{>}0$. By definition $\sigma_{\text{BRG}}((\ell,\nu),R) = ((b,c,a),R')$ for some $((b,c,a),R') \in A_{\text{BRG}}((\ell,\nu),R)$. Now, from Corollary E.5 we have that there exists $(t',a) \in A_{\mathsf{T}}(\ell,\nu')$, such that $\nu'{+}t' \in R'$, $|t{-}t'| < \varepsilon$ and for every $C \subseteq \mathcal{C}$, we have:

$$\mathsf{scatter}\big(\, (\nu{+}t)[C{:=}0]\, ,\, (\nu'{+}t')[C{:=}0]\,\big) < \varepsilon$$

where $t = b - \nu(c)$. Letting $\sigma_{\mathcal{T}}^{M+1}(\ell, \nu') = (t', a)$ it follows from Definition E.6 and the definition of ETotal that for any $N \leq M$:

$$\begin{aligned}
&\mathrm{ETotal}_{\mathcal{T}_{\mathrm{BRG}}}(N)((((\ell,\nu),R),\sigma_{\mathrm{BRG}}) \\
&= \pi((\ell,\nu),(t,a)) + \sum_{(C,\ell')\in 2^{\mathcal{C}}\times L} \delta[\ell,a](C,\ell') \cdot \mathrm{ETotal}_{\mathcal{T}_{\mathrm{BRG}}}(N{-}1)(s_{\mathrm{BRG}}^{\ell',C}, \sigma_{\mathrm{BRG}}[r_{\mathrm{BRG}}^{\ell',C}]) \\
&\mathrm{ETotal}_{[\![\mathcal{T}]\!]}(N)(((\ell,\nu'),\sigma_{\mathcal{T}}^{M+1}) \\
&= \pi((\ell,\nu'),(t',a)) + \sum_{(C,\ell')\in 2^{\mathcal{C}}\times L} \delta[\ell,a](C,\ell') \cdot \mathrm{ETotal}_{[\![\mathcal{T}]\!]}(N{-}1)(s_{\mathsf{T}}^{\ell',C}, \sigma_{\mathcal{T}}^{M+1}[r_{\mathsf{T}}^{\ell',C}])
\end{aligned}$$

where in the above $s_{\mathrm{BRG}}^{\ell',C} = ((\ell', (\nu{+}t)[C{:=}0]), R'[C{:=}0])$, $r_{\mathrm{BRG}}^{\ell',C} = \langle((\ell,\nu),R),(a,t),s_{\mathrm{BRG}}^{\ell',C}\rangle$, $s_{\mathsf{T}}^{\ell',C} = (\ell', (\nu'{+}t')[C{:=}0])$ and $r_{\mathsf{T}}^{\ell',C} = \langle(\ell,\nu'),(a,t'),s^{\ell',C,t'}\rangle$.Combining these facts yields:

$$\begin{aligned}
&\Big| \mathrm{ETotal}_{\mathcal{T}_{\mathrm{BRG}}}(N)((((\ell,\nu),R),\sigma_{\mathrm{BRG}}) - \mathrm{ETotal}_{[\![\mathcal{T}]\!]}(N)((\ell,\nu'),\sigma_{\mathcal{T}}^{M+1}) \Big| \\
&\qquad \leq \big|\pi((\ell,\nu),(t,a)) - \pi((\ell,\nu'),(t',a))\big| + \sum_{(C,\ell')\in 2^{\mathcal{C}}\times L} \delta[\ell,a](C,\ell')\cdot \\
&\qquad\quad \Big| \mathrm{ETotal}_{\mathcal{T}_{\mathrm{BRG}}}(N{-}1)(s_{\mathrm{BRG}}^{\ell',C}, \sigma_{\mathrm{BRG}}[r_{\mathrm{BRG}}^{\ell',C}]) - \mathrm{ETotal}_{[\![\mathcal{T}]\!]}(N{-}1)(s_{\mathsf{T}}^{\ell',C}, \sigma_{\mathcal{T}}^{M+1}[r_{\mathsf{T}}^{\ell',C}]) \Big| \, .
\end{aligned} \tag{6}$$

Now by construction, for any $C \subseteq \mathcal{C}$ we have that:

$$\mathsf{scatter}\big(\, (\nu{+}t)[C{:=}0]\, ,\, (\nu'{+}t')[C{:=}0]\, \big) < \varepsilon$$

and hence using Lemma E.3 and induction, for any $\ell' \in L$ and $C \subseteq \mathcal{C}$, there exists a strategy $\sigma_{\mathcal{T}}^{\ell',C}$ of $[\![\mathsf{T}]\!]$ such that

$$\Big| \mathrm{ETotal}_{\mathcal{T}_{\mathrm{BRG}}}(N{-}1)(s_{\mathrm{BRG}}^{\ell',C}, \sigma_{\mathrm{BRG}}[r_{\mathrm{BRG}}^{\ell',C}]) - \mathrm{ETotal}_{[\![\mathcal{T}]\!]}(N{-}1)(s_{\mathsf{T}}^{\ell',C}, \sigma_{\mathcal{T}}^{\ell',C}) \Big| < K\cdot(N{-}1)\cdot\varepsilon \tag{7}$$

Constructing the strategy $\sigma_{\mathcal{T}}^{M+1}$ such that $\sigma_{\mathcal{T}}^{M+1}[r_{\mathsf{T}}^{\ell',C}] = \sigma_{\mathcal{T}}^{\ell',C}$ for all $\ell' \in L$ and $C \subseteq \mathcal{C}$, it follows from (6) and (7) that:

$$\begin{aligned}
&\Big| \mathrm{ETotal}_{\mathcal{T}_{\mathrm{BRG}}}(N)((((\ell,\nu),R),\sigma_{\mathrm{BRG}}) - \mathrm{ETotal}_{[\![\mathcal{T}]\!]}(N)((\ell,\nu'),\sigma_{\mathcal{T}}^{M+1}) \Big| \\
&< \big|\pi((\ell,\nu),(t,a)) - \pi((\ell,\nu'),(t',a))\big| + \sum_{(C,\ell')\in 2^{\mathcal{C}}\times L} \delta[\ell,a](C,\ell') \cdot K\cdot(N{-}1)\cdot\varepsilon \\
&\leq \big|\pi((\ell,\nu),(t,a)) - \pi((\ell,\nu'),(t',a))\big| + K\cdot(N{-}1)\cdot\varepsilon && \text{since } \delta[\ell,a] \in \mathcal{D}(2^{\mathcal{C}}\times L) \\
&\leq K\cdot\varepsilon + K\cdot(N{-}1)\cdot\varepsilon && \text{since } |t-t'| \leq \varepsilon \text{ and } \pi \text{ is } K\text{-Lipschitz continuous} \\
&= K\cdot N\cdot\varepsilon && \text{rearranging}
\end{aligned}$$

which completes the proof by induction. □