Polytopal Stochastic Games ⋆

Pablo F. Castro[1,3] and Pedro R. D'Argenio[2,3]

[1] Universidad Nacional de Río Cuarto, FCEFQyN, Departamento de Computación, Río Cuarto, Córdoba, Argentina, castro@dc.exa.unrc.edu.ar
[2] Universidad Nacional de Córdoba, FAMAF, Córdoba, Argentina, pedro.dargenio@unc.edu.ar
[3] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina

Abstract. In this paper we introduce *polytopal stochastic games*, an extension of two-player, zero-sum, turn-based stochastic games, in which we may have uncertainty over the transition probabilities. In these games the uncertainty over the probability distributions is captured via linear (in)equalities whose space of solutions forms a polytope. We give a formal definition of these games and prove their basic properties: determinacy and existence of optimal memoryless and deterministic strategies. We do this for reachability and different types of reward objectives and show that the solution exists in a finite representation of the game. We also state that the corresponding decision problems are in $\mathsf{NP} \cap \mathsf{coNP}$. We motivate the use of polytopal stochastic games via a simple example.

1 Introduction

In the last decades, stochastic systems have become ubiquitous in computer science: communication and security protocols, fault analysis in critical systems, autonomous devices, to name a few examples, typically use techniques coming from probability theory. Furthermore, well-known techniques in artificial intelligence, such as reinforcement learning [28], are based on stochastic models. In view of this, the verification and formal analysis of stochastic systems is one of the most active areas of research in software verification. Christel Baier and Joost-Pieter Katoen's book [4] is considered a standard reference in the area, it introduces common concepts and techniques for model checking probabilistic systems, this includes algorithms for verifying temporal assertions over Markov chains (MCs) and Markov Decision processes (MDPs). The latter can be considered as one player stochastic games, in which the system has to select strategies to solve non-determinism in stochastic settings. In general, game theory offers a powerful mathematical framework for specifying and verifying computing systems. The idea is appealing, a computing system can be thought of as a player playing against an environment, or another system, while trying to achieve certain goals. For instance, a security system can be seen as a player that selects different

⋆ This work was supported by Agencia I+D+i PICT 2019-03134, SeCyT-UNC 33620230100384CB (MECANO), and EU Grant agreement ID: 101008233 (MISSION).

countermeasures to possibly different types of maneuvers executed by an attacker (a second player) each of which may succeed with certain probabilities. The objective of the defense system is to minimize the probability that the attack succeeds while the attacker wants to maximize it. This scenario can be modeled as a stochastic game, and then analysed using techniques coming from game theory. Examples of applications of game theory to the analysis of systems can be found almost everywhere in the last years: self-driving cars [31], robotics [18], UAVs [14], security [3], etc. Furthermore, in recent years, some model checkers have been extended to provide support for stochastic games, e.g., this is the case of PRISM-Games [11], which offers support for several versions of stochastic games.

In this paper we focus on two-player, zero-sum, turn-based perfect-information stochastic games. Intuitively, they are non-deterministic probabilistic transition systems in which the vertices are partitioned into two sets: vertices belonging to player $\Box$ and vertices belonging to player $\Diamond$. When the current state belongs to a given player, say $\Box$, she performs an action by selecting one of the non-deterministic outgoing transitions which would lead to different states with some given probabilities. Typically, the players want to fulfill or maximize/minimize some objectives. Standard quantitative objectives are discounted sum (the players collect an amount of rewards during the play which are multiplied by a discount factor in each step), total sum (the players want to maximize/minimize the cumulative sum of the rewards collected during a play), mean-payoff (the objective is to maximize or minimize the long-run average reward), or simply a reachability objective, that is, they aim to maximize/minimize the probability of reaching certain subset of states. These kinds of objectives can be used and combined to model different kinds of systems, e.g., the case of a self-driving car intending to maximize the probability of reaching some zone in a city can be seen as a multiobjective game [12].

Most of the time, when modeling stochastic systems, one assumes that the probability distributions are exactly known, which may not always be the case due to measurement inaccuracies, lack of data, or other issues. In this paper we propose an extension of stochastic games that adds the possibility of having uncertainty over the probabilities. Games with some kinds of uncertainty have been considered for $1\frac{1}{2}$-player games, i.e., Markov Decision Processes (MDPs). For instance, Interval-valued Discrete-Time Markov Chains (IDTMCs) [17,20,29], Interval Markov Decision Process (IMDP) [29], and Convex MDPs [26]. To the best of our knowledge, these approaches have not been extended to stochastic games (i.e., $2\frac{1}{2}$-player games). A key challenge for doing so is that in multiplayer games one needs to prove determinacy results, this ensures that the games possess a well-defined value, which does not depend on the players' knowledge. In the aforementioned approaches the notion of uncertainty is usually adversarially resolved, that is, each time a state is visited, the adversary picks a transition distribution that respects the constraints, and takes a probabilistic step according to the chosen distribution. However, it is interesting to note that, in two-player games we may adopt two ways of resolving uncertainty: a controllable one, in which the actual player resolves the uncertainty following her goals; and an adversarial one in which the adversary resolves the uncertainty in her favor. The former approach is useful in those scenarios in which the uncertainty affects the adversary as she does not precisely know the possible movements of our player; while the latter is helpful to reason in worst-case scenarios.

We therefore introduce *polytopal stochastic games (PSG)*. PSGs, as defined in Section 4, allow one to model uncertainties over probability distributions using linear (closed) inequalities. Geometrically, these linear inequalities correspond to polytopes, i.e., bounded polyhedra. As PSGs are two-player games, both ways of resolving uncertainty are possible: the adversarial approach and the controllable one. Furthermore, we show that in all the cases these kinds of games preserve some good properties of standard stochastic games for several objectives: reachability, total rewards, average sum, and mean payoff. In particular we show that these games are determined and admit optimal memoryless and deterministic strategies. We also show that these inherently infinite games can be reduced to equivalent finite stochastic games that traverse exclusively through the vertices of the original polytopes. As such, they are amenable to standard algorithmic solutions. Finally, we prove that the complexity of these games for the aforementioned objectives remain in $\mathsf{NP} \cap \mathsf{coNP}$, that is, they stay in the standard complexity class of simple stochastic games, even when polytopal games support for an uncountable number of actions for the players and the discretization may grow exponentially.

Related work. Definitions of infinite stochastic games do exist (see, for instance, [21]) though they are of discrete nature, contrary to the type of games presented here. In fact, PSGs are related to IDTMCs [17,20,29], IMDPs [29], and Convex MDPs [26], but they are variants of MDPs and hence they are $1\frac{1}{2}$-games. In particular, PSGs adopt a semantics similar to IMDPs and Convex MDPs [26] in which the uncertainty introduced by the polytope is interpreted as an uncountable non-deterministic branching. While in [26] interior-point algorithms are used to solve Convex MDPs, we use a discretization through the vertices of polytopes to solve PSGs. Though this has an exponential impact, this is very mild in practice as we will show later. A much simpler variant of PSG was used in [8] to provide an algorithmic solution for a fault tolerant measure. This incipient idea served as the starting point for the generalization presented here.

Somewhat related are the stochastic timed games (STGs) [7,1]. However, the continuous non-determinism introduced by the time in STGs is resolved by uniform and exponential distributions and the remnant non-determinism (resolved by the strategies) is still discrete. This does not make these models simpler since undecidability has been shown for games with at least 3 clocks [7].

Outline of the paper. Section 2 presents a motivating example. Section 3 introduces the background needed for tackling the rest of the paper. The definition of PSGs, their semantics and basic properties are given in Section 4. The main results are presented in Section 5. Full proofs are gathered in the Appendix.

2 Roberta vs. Rigoborto in the land of uncertainties

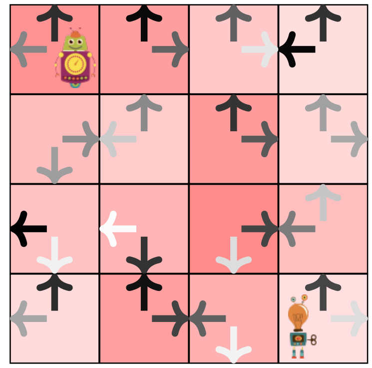

Fig. 1: An example of a grid for the Roborta vs Rigoborto game.

We illustrate our approach by means of a simple example. Consider a field represented as a bidimensional grid and two robots –which we call Roborta and Rigoborto– that navigate it. Roborta can move sideways and forward, Rigoborto can move sideways and backward. The robots start at a certain initial position. Roborta intends to reach the end of the grid, i.e., she wants to reach position $(i, n+1)$ for any i, whereas Rigoborto wants to stop Roborta. He can achieve this by reaching Roborta's location. The robots play in turns. The objective of Roborta is to maximize the probability of reaching the exit, while the objective of Rigoborto is to minimize this value. We spice up this example by considering the terrain quality (which depends on factors like, e.g., stones, mud or grass) and slope, which may cause imprecisions and uncertainties in the robots mobility, probably making them slide towards some undesired direction. The terrain quality and slope may vary in each grid position. In Fig. 1, we show an example of such a scenario. Therein, the robots start at the corners, the arrows indicate the slopes in the terrain, and the colors in the cells indicate the terrain quality. Darker arrows correspond to sharper slopes. Similarly, cells with lower quality are colored with stronger red colors.

More precisely, for each (x, y)-cell, the terrain quality $q_{xy} \in [0, 0.5]$ gives the uncertainty factor, where $q_{xy} = 0$ means that probabilities are completely determined, and, as q_{xy} grows, the probability values become increasingly fuzzier. In addition, we consider two factors associated with the terrain slopes: $l_{xy}, f_{xy} \in [-1, 1]$, representing the inclination of the lateral and frontal slopes respectively. Thus, as l_{xy} get closer to 1 (-1), the likelihood of shifting to the right (left) increases, with $l_{xy} = 0$ not favouring any particular side. Similarly $f_{xy} > 1$ ($f_{xy} < -1$) biases the robot towards the front (back). Let p_c be the probability that the robot command is successful (that is, that it moves in the intended direction), and let p_l, p_r, p_f, and p_b be the probabilities that the command is unsuccessful and the robot uncontrollably slides respectively to the left, right, front, and back. Then, the space of all probability values can be defined by the following set of inequalities:

$$
\begin{aligned}
1 &= p_c + p_l + p_r + p_f + p_b \\
p_c &\geq 0, \quad p_l \geq 0, \quad p_r \geq 0, \quad p_f \geq 0, \quad p_b \geq 0 \\
p_c &\leq 1 - (q_{xy} + \tfrac{1}{2} \cdot (1 - (1 - |l_{xy}|) \cdot (1 - |f_{xy}|))) \\
0 &\leq (1 - \max(0, -l_{xy})) \cdot p_l - (1 - q_{xy}) \cdot (1 - \max(0, l_{xy})) \cdot p_r \\
0 &\leq (1 - \max(0, l_{xy})) \cdot p_r - (1 - q_{xy}) \cdot (1 - \max(0, -l_{xy})) \cdot p_l \\
0 &\leq (1 - \max(0, f_{xy})) \cdot p_f - (1 - q_{xy}) \cdot (1 - \max(0, -f_{xy})) \cdot p_b
\end{aligned}
$$

```
// action specification for Roberta moving to the left
[robl] (turn = 0) & (roby<L) & !Collision -> (rob_mov'=1) & (turn'=1)

[robl-cont] (turn = 1) & (rob_mov = 1) ->
  //The first four probabilistic options  correspond to environments setbacks
  pl : (robx'=max(0,robx-1)) & (rob_mov'=0) + pr : (robx'=min(W-1,robx+1)) & (rob_mov'=0)
  + pf : (roby'=roby+1) & (rob_mov'=0) + pb : (roby'=max(0,roby+1)) & (rob_mov'=0)
  + pc : (robx'=max(0,robx-1)) & (rob_mov'=0)
  {// inequations for uncertainty
    1-(Q[robx,roby]+(1-(1-abs(L[robx,roby]))*(1-abs(F[robx,roby])))/2) >= pc,
    (1-max(0,-L[robx,roby]))*pl - (1-Q[robx,roby])*(1-max(0,L[robx,roby]))*pr >= 0,
    (1-max(0,L[robx,roby]))*pr - (1-Q[robx,roby])*(1-max(0,-L[robx,roby]))*pl >= 0,
    (1-max(0,F[robx,roby]))*pf - (1-Q[robx,roby])*(1-max(0,-F[robx,roby]))*pb >= 0,
    (1-max(0,-F[robx,roby]))*pb - (1-Q[robx,roby])*(1-max(0,F[robx,roby]))*pf >= 0
  };
```

(a) Roberta moves left

```
// action specification for Rigoborto moving to the left
[rigl] (turn = 1) & (rob_mov = 0) & (rigy<L) & !Collision ->
  //The first four probabilistic options  correspond to environments setbacks
  pl : (rigx'=max(0,rigx-1)) & (turn'=0) & (Collision'=(robx=rigx && roby=rigy))
  + pr : (rigx'=min(W-1,rigx+1)) & (turn'=0) & (Collision'=(robx=rigx && roby=rigy))
  + pf : (rigy'=rigy+1) & (turn'=0) & (Collision'=(robx=rigx && roby=rigy))
  + pb : (rigy'=max(0,rigy+1)) & (turn'=0) & (Collision'=(robx=rigx && roby=rigy))
  + pc : (rigx'=max(0,rigx-1)) & (turn'=0) & (Collision'=(robx=rigx && roby=rigy))
  {// inequations for uncertainty
    1-(Q[rigx,rigy]+(1-(1-abs(L[rigx,rigy]))*(1-abs(F[rigx,rigy])))/2) >= pc,
    (1-max(0,-L[rigx,rigy]))*pl - (1-Q[rigx,rigy])*(1-max(0,L[rigx,rigy]))*pr >= 0,
    (1-max(0,L[rigx,rigy]))*pr - (1-Q[rigx,rigy])*(1-max(0,-L[rigx,rigy]))*pl >= 0,
    (1-max(0,F[rigx,rigy]))*pf - (1-Q[rigx,rigy])*(1-max(0,-F[rigx,rigy]))*pb >= 0,
    (1-max(0,-F[rigx,rigy]))*pb - (1-Q[rigx,rigy])*(1-max(0,F[rigx,rigy]))*pf >= 0
  };
```

(b) Rigoborto moves left

Fig. 2: Fragment of code for Roberta vs Rigoborto

$$0 \;\leq\; (1 - \max(0, -f_{xy})) \cdot p_b - (1 - q_{xy}) \cdot (1 - \max(0, f_{xy})) \cdot p_f$$

Note that if $q_{xy} = 0$, the system has a unique solution. If, in addition, $l_{xy} > 0$, $1/(1 - l_{xy}) = p_r/p_l$ giving the likelihood ratio of sliding towards the right.

Our aim is to find the best strategy for Roberta to win against all odds. This implies that the terrain uncertainty behaves adversarially to Roberta but favourably to Rigoborto. Thus, in our model, Rigoborto controls the non-determinism introduced by the terrain uncertainty. Assuming an extension of the PRISM-Games language, the code could look like in Fig. 2, where subfigures 2a and 2b show the decisions to move left by Roberta and Rigoborto respectively.

Variable `turn` indicates who is the next player to move (with `0` for Roberta and `1` for Rigoborto). If it is Roberta's turn (see first line in Fig. 2a) and she decides to move left, she indicates it by setting `rob_mov'=1` (`1` indicates a left move while `2`, `3`, and `4` are used for the other directions, and `0` to indicate that Roberta is not moving). At the same time, she yields her turn by setting `turn'=1`. Notice that the action is not yet complete: the reaction of the terrain to the move is encoded in the next line (action `robl-cont` in Fig. 2a). Notice that this action happens in a state in which `turn=1`, making the terrain uncertainty –defined by the polytope– adversarial to Roberta. Here, variables `robx` and `roby` correspond to Roborta's coordinates x and y and constant matrices `Q`, `L`, and `F` contain the respective values for q_{xy}, l_{xy}, and f_{xy}. The rest of the variables are as expected. Once this step is taken, variable `rob_mov` is set to `0`, thus enabling Rigoborto's move. Rigoborto's decision to move left is given in Fig 2b. Notice that this is performed in a single action since we assume that the terrain uncertainty plays in favour of him. Something particular to this transition is the setting of variable `Collision` to indicate whether Rigoborto has caught Roborta.

3 Preliminaries

In this section we introduce notation and basic concepts of polytopes and games. Interested readers are referred to [32,21].

In the following $\mathscr{P}(S)$ denotes the powerset of set S, and $\mathscr{P}_f(S)$ denotes the set of finite subsets of set S. A *convex polytope* in $\mathbb{R}^n$ is a bounded set $K = \{\boldsymbol{x} \in \mathbb{R}^n \mid A\boldsymbol{x} \leq b\}$, with $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$, for some $m \in \mathbb{N}$. By *bounded* we mean (in our case) that there exists $M \in \mathbb{R}_{\geq 0}$ such that $\sum_{i=1}^{n} |x_i| \leq M$ for all $\boldsymbol{x} \in K$ (x_i denotes the ith element of $\boldsymbol{x}$). Let S be a finite set. As functions in $\mathbb{R}^S$ can be equivalently seen as vectors in $\mathbb{R}^{|S|}$, we will in general refer to polytopes in $\mathbb{R}^S$. Let $\mathsf{Poly}(S)$ be the set of all convex polytopes in $\mathbb{R}^S$. Notice that the set of all probability functions on S form the convex polytope $\mathsf{Dist}(S) = \{\mu \in \mathbb{R}^S \mid \sum_{s\in S} \mu(s) = 1 \text{ and } \forall s \in S\colon \mu(s) \geq 0\}$. Let $\mathsf{DPoly}(S) = \{K \cap \mathsf{Dist}(S) \mid K \in \mathsf{Poly}(S)\}$. Thus, $K \in \mathsf{DPoly}(S)$ is a convex polytope whose elements are also probability functions on S and therefore its defining set of inequality $A\boldsymbol{x} \leq b$ already encodes the inequalities $\sum_{s\in S} x_s = 1$ and $x_s \geq 0$ for $s \in S$.

Any convex polytope $K \in \mathsf{Poly}(S)$ can alternatively be characterized as the convex hull of its finite set of vertices. Let $\mathbb{V}(K)$ denote the set of all vertices of polytope K. If $\mathbb{V}(K) = \{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k\}$, then every $\boldsymbol{x} \in K$ is a convex combination of $\{\boldsymbol{v}^1, \ldots, \boldsymbol{v}^k\}$, that is, $\boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{v}^k$ with $\lambda_i \geq 0$, for $i \in [1..k]$, and $\sum_{i=1}^{k} \lambda_i = 1$. A *simplex* is any convex polytope $K \in \mathsf{Poly}(S)$ whose set of vertices $\mathbb{V}(K)$ is affinely independent, that is, for any family $\{\lambda_{\boldsymbol{v}} \in \mathbb{R}\}_{\boldsymbol{v}\in\mathbb{V}(K)}$ such that $\sum_{\boldsymbol{v}\in\mathbb{V}(K)} \lambda_{\boldsymbol{v}} = 0$, $\sum_{\boldsymbol{v}\in\mathbb{V}(K)} \lambda_{\boldsymbol{v}} \boldsymbol{v} = 0$ implies that $\lambda_{\boldsymbol{v}} = 0$ for all $\boldsymbol{v} \in \mathbb{V}(K)$. This implies that for every $\boldsymbol{x} \in K$, with K being a simplex, the convex combination $\boldsymbol{x} = \sum_{i=1}^{k} \lambda_i \boldsymbol{v}^k$ is unique. We also remark that any convex polytope K can be expressed as the union of a (finite) set of simplices $\{K_i\}_{i\in I}$ so that $\mathbb{V}(K) = \bigcup_{i\in I} \mathbb{V}(K_i)$ (this is a consequence of Charathéodory's Theorem [32,24]). We will call such decomposition a *vertex-preserving triangulation*. Let $\mathsf{Simp}(S)$ denote the set of all simplices in $\mathbb{R}^S$ and $\mathsf{DSimp}(S) = \mathsf{Simp}(S) \cap \mathsf{DPoly}(S)$.

A *stochastic game* [30,13,15] is a tuple $\mathcal{G} = (\mathcal{S}, (\mathcal{S}_\Box, \mathcal{S}_\Diamond), \mathcal{A}, \theta)$, where $\mathcal{S}$ is a finite set of *states* with $\mathcal{S}_\Box, \mathcal{S}_\Diamond \subseteq \mathcal{S}$ being a partition of $\mathcal{S}$, $\mathcal{A}$ is a (finite) set of *actions*, and $\theta : \mathcal{S} \times \mathcal{A} \times \mathcal{S} \to [0, 1]$ is a *probabilistic transition function* such that for every $s \in \mathcal{S}$ and $a \in \mathcal{A}$, $\theta(s, a, \cdot) \in \mathsf{Dist}(\mathcal{S})$ or $\theta(s, a, \mathcal{S}) = 0$. Let $\mathcal{A}(s) = \{a \in \mathcal{A} \mid \theta(s, a, \mathcal{S}) = 1\}$ be the set of actions enabled at state s. If $\mathcal{S}_\Box = \emptyset$ or $\mathcal{S}_\Diamond = \emptyset$, then $\mathcal{G}$ is a *Markov decision process* (or MDP). If, in addition, $|\mathcal{A}(s)| = 1$ for all $s \in \mathcal{S}$, $\mathcal{G}$ is a *Markov chain* (or MC). A *path* in the game $\mathcal{G}$ is an infinite sequence of states $\rho = s_0 s_1 \ldots$ such that, for every $k \in \mathbb{N}$, there is an $a \in \mathcal{A}$ with $\theta(s_k, a, s_{k+1}) > 0$. For $i \geq 0$, ρ_i indicates the ith state in the path ρ (notice that ρ_0 is the first state in ρ). $\mathsf{Paths}_{\mathcal{G}}$ denotes the set of all paths, and $\mathsf{FPaths}_{\mathcal{G}}$ denotes the set of finite prefixes of paths. Similarly, $\mathsf{Paths}_{\mathcal{G},s}$ and $\mathsf{FPaths}_{\mathcal{G},s}$ denote the set of paths and the set of finite paths starting at state s. A *strategy* for the i-player (for $i \in \{\Box, \Diamond\}$) in a game $\mathcal{G}$ is a function $\pi_i : \mathcal{S}^*\mathcal{S}_i \to \mathsf{Dist}(\mathcal{A})$ that assigns a probabilistic distribution to each finite sequence of states such that $\pi_i(\hat{\rho}s)(a) > 0$ only if $a \in \mathcal{A}(s)$. The set of all strategies for the i-player is named Π_i. Whenever convenient, we indicate that the set of strategies Π_i belongs to the game $\mathcal{G}$ by writing by $\Pi_{\mathcal{G},i}$ A strategy π_i is said to be *pure* or *deterministic* if, for every $\hat{\rho}s \in \mathcal{S}^*\mathcal{S}_i$, $\pi_i(\hat{\rho}s)$ is a Dirac distribution (that is a distribution δ_a s.t., $\delta_a(a) = 1$ and $\delta_a(b) = 0$ for all $b \neq a$), and it is called *memoryless* if $\pi_i(\hat{\rho}s) = \pi_i(s)$, for every $\hat{\rho} \in \mathcal{S}^*$. Let Π_i^M be the set of all memoryless strategies for the i-player and Π_i^{MD} be the set of all its deterministic and memoryless strategies. Note that the definition of strategy given above works for set of actions that are finite, in Section 4 we define strategies for uncountable sets of actions.

Given strategies $\pi_\Box \in \Pi_\Box$ and $\pi_\Diamond \in \Pi_\Diamond$, and an initial state s, the *result* of the game is a Markov chain [10], denoted $\mathcal{G}_s^{\pi_\Box,\pi_\Diamond}$. The Markov chain $\mathcal{G}_s^{\pi_\Box,\pi_\Diamond}$ defines a probability measure $\mathbb{P}_{\mathcal{G},s}^{\pi_\Box,\pi_\Diamond}$ on the Borel σ-algebra generated by the cylinders of $\mathsf{Paths}_{\mathcal{G},s}$. If ξ is a measurable set in such a Borel σ-algebra, $\mathbb{P}_{\mathcal{G},s}^{\pi_\Box,\pi_\Diamond}(\xi)$ is the probability that strategies $\pi_\Box$ and $\pi_\Diamond$ follow a path in ξ starting from state s. We use LTL notation to represent specific set of paths, in particular, $D \,\mathsf{U}^n\, C = \{\rho \in \mathcal{S}^\omega \mid \rho_n \in C \wedge \forall j < n \colon \rho_j \in D\} = D^n \times C \times \mathcal{S}^\omega$ is the set of paths that reach $C \subseteq \mathcal{S}$ in exactly $n \geq 0$ steps traversing before only states in $D \subseteq \mathcal{S}$; $\Diamond^n C = \mathcal{S} \,\mathsf{U}^n\, C$ is the set of all paths reaching states in C in exactly n steps; and $\Diamond C = \bigcup_{n\geq 0}(\mathcal{S} \setminus C) \,\mathsf{U}^n\, C$ is the set of all paths that reach a state in C.

A stochastic game is said to be *almost surely stopping* [13,15] if for all pair of strategies $\pi_\Box$, $\pi_\Diamond$ the probability of reaching a terminal state is 1. A state s is *terminal* if $\theta(s, a, s) = 1$, for all $a \in \mathcal{A}(s)$. In other words, a game is stopping if $\inf_{\pi_\Diamond \in \Pi_\Diamond} \inf_{\pi_\Box \in \Pi_\Box} \mathbb{P}_s^{\pi_\Box,\pi_\Diamond}(\Diamond T) = 1$, where $T \subseteq \mathcal{S}$ is the set of terminal states. A stochastic game is *irreducible* [15] if for all pair of strategies, the probability of reaching a state from any other state is positive, that is, if $\inf_{\pi_\Diamond \in \Pi_\Diamond} \inf_{\pi_\Box \in \Pi_\Box} \mathbb{P}_s^{\pi_\Box,\pi_\Diamond}(\Diamond s') > 0$ for all pair of states $s, s' \in \mathcal{S}$.

A *quantitative objective* or *payoff function* is a measurable function $f : \mathcal{S}^\omega \to \mathbb{R}$. Let $\mathbb{E}_{\mathcal{G},s}^{\pi_\Box,\pi_\Diamond}[f]$ be the expectation of measurable function f under probability $\mathbb{P}_{\mathcal{G},s}^{\pi_\Box,\pi_\Diamond}$. The goal of the $\Box$-player is to maximize this value whereas the goal of the $\Diamond$-player is to minimize it. Sometimes quantitative objective functions can be defined via *rewards*. These are assigned by a *reward function* $r : S \to \mathbb{R}^+$. We usually consider stochastic games augmented with a reward function. Moreover, we assume that for every terminal state s, $r(s) = 0$. The value of the game for the $\Box$-player at state s under strategy $\pi_\Box$ is defined as the infimum over all the values resulting from the $\Diamond$-player strategies in that state, i.e., $\inf_{\pi_\Diamond \in \Pi_\Diamond} \mathbb{E}_{\mathcal{G},s}^{\pi_\Box,\pi_\Diamond}[f]$. The *value of the game* for the $\Box$-player is defined as the supremum of the values of all the $\Box$-player strategies, i.e., $\sup_{\pi_\Box \in \Pi_\Box} \inf_{\pi_\Diamond \in \Pi_\Diamond} \mathbb{E}_{\mathcal{G},s}^{\pi_\Box,\pi_\Diamond}[f]$. Similarly, the value of the game for the $\Diamond$-player under strategy $\pi_\Diamond$ and the value of the game for the $\Diamond$-player are defined as $\sup_{\pi_\Box \in \Pi_\Box} \mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G},s}[f]$ and $\inf_{\pi_\Diamond \in \Pi_\Diamond} \sup_{\pi_\Box \in \Pi_\Box} \mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G},s}[f]$, respectively. We say that a game is *determined* if both values are the same, that is, $\sup_{\pi_\Box \in \Pi_\Box} \inf_{\pi_\Diamond \in \Pi_\Diamond} \mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G},s}[f] = \inf_{\pi_\Diamond \in \Pi_\Diamond} \sup_{\pi_\Box \in \Pi_\Box} \mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G},s}[f]$.

In this paper we focus on *total accumulated reward*, where the payoff function is defined by $\mathsf{rew_t}(\rho) = \lim_{n\to\infty} \sum_{i=0}^{n} r(\rho_i)$, *total discounted reward*, defined by $\mathsf{rew}_\gamma(\rho) = \lim_{n\to\infty} \sum_{i=0}^{n} \gamma^i r(\rho_i)$, where $\gamma \in (0,1)$ is the discount factor, and *average reward*, defined by $\mathsf{rew_a}(\rho) = \lim_{n\to\infty} \frac{1}{n+1} \sum_{i=0}^{n} r(\rho_i)$. By taking, respectively, $\mathsf{f}(i,n) = 1$, $\mathsf{f}(i,n) = \gamma^i$, or $\mathsf{f}(i,n) = \frac{1}{n+1}$, we refer simultaneously to the above payoff functions with the single function $\mathsf{rew_f}(\rho) = \lim_{n\to\infty} \sum_{i=0}^{n} \mathsf{f}(i,n) r(\rho_i)$.

We also focus on *reachability objective*. In this case, the goal of the $\Box$-player is to maximize the probability of reaching a state on a goal set $G \subseteq \mathcal{S}$ whereas the goal of the $\Diamond$-player is to minimize it. Therefore, similar to quantitative objectives, the *value of the reachability game for the* $\Box$*-player* is defined by $\sup_{\pi_\Box \in \Pi_\Box} \inf_{\pi_\Diamond \in \Pi_\Diamond} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G},s}(\Diamond G)$ and the *value of the reachability game for the* $\Diamond$*-player* is defined by $\inf_{\pi_\Diamond \in \Pi_\Diamond} \sup_{\pi_\Box \in \Pi_\Box} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G},s}(\Diamond G)$, and the game is *determined* if both values are the same.

4 Polytopal Stochastic Games

A polytopal stochastic game is characterized through a structure that contains a finite set of states divided into two sets, each owned by a different player. In addition, each state has assigned a finite set of convex polytopes of probability distributions over states. The formal definition is as follows.

Definition 1. *A* polytopal stochastic game *(PSG, for short) is a structure* $\mathcal{K} = (\mathcal{S}, (\mathcal{S}_\Box, \mathcal{S}_\Diamond), \Theta)$ *such that* $\mathcal{S}$ *is a finite set of states partitioned into* $\mathcal{S} = \mathcal{S}_\Box \uplus \mathcal{S}_\Diamond$ *and* $\Theta : \mathcal{S} \to \mathscr{P}_\mathsf{f}(\mathsf{DPoly}(\mathcal{S}))$*. If, in particular,* $\Theta : \mathcal{S} \to \mathscr{P}_\mathsf{f}(\mathsf{DSimp}(\mathcal{S}))$*, we call* $\mathcal{K}$ *a* simplicial stochastic game *(SSG for short).*

The idea of a PSG is as expected: in a state $s \in \mathcal{S}_i$ ($i \in \{\Box, \Diamond\}$), player i chooses to play a polytope $K \in \Theta(s)$ and a distribution $\mu \in K$. The next state s' is sampled according to distribution μ and the game continues from s' repeating the same process.

As a particular example, one can devise a stochastic game variant of Interval Markov Decision Processes (IMDPs) [17,20]. This type of games can be interpreted as a PSG where every polytope $K \in \Theta(s)$, for all $s \in \mathcal{S}$, is defined by $\mu \in K$ iff $\sum_{s' \in \mathcal{S}} \mu(s') = 1$ and, for all $s' \in \mathcal{S}$ and some fixed $0 \leq l_{s'} \leq u_{s'} \leq 1$, $l_{s'} \leq \mu(s') \leq u_{s'}$ (note that the intervals need to be closed).

The behaviour of a polytopal stochastic game is formally interpreted in terms of a stochastic game where the number of transitions outgoing the players' states may be uncountably large. We choose a controllable view on the uncertainty introduced by the polytope since the adversarial alternative can be encoded as was shown in Sec. 2. Formally, the interpretation of a PSG is as follows.

Definition 2. *The* interpretation of the polytopal stochastic game $\mathcal{K}$ *is defined by the stochastic game* $\mathcal{G}_{\mathcal{K}} = (\mathcal{S}, (\mathcal{S}_\Box, \mathcal{S}_\Diamond), \mathcal{A}, \theta)$*, where* $\mathcal{A} = \left(\bigcup_{s\in\mathcal{S}} \Theta(s)\right) \times \mathsf{Dist}(\mathcal{S})$ *and*

$$\theta(s, (K, \mu), s') = \begin{cases} \mu(s') & \text{if } K \in \Theta(s) \text{ and } \mu \in K \\ 0 & \text{otherwise} \end{cases}$$

Notice that the set of actions $\mathcal{A}$ can be uncountably large, as well as each set $\mathcal{A}(s) = \bigcup_{K\in\Theta(s)}\{K\}\times K$. Therefore we need to extend the strategies to this uncountable domain which should be properly endowed with a σ-algebra. For this we make use of a standard construction to provide a σ-algebra to $\mathsf{Dist}(\mathcal{S})$ [16]: $\boldsymbol{\Sigma}_{\mathsf{Dist}(\mathcal{S})}$ is defined as the smallest σ-algebra containing the sets $\{\mu \in \mathsf{Dist}(\mathcal{S}) \mid \mu(S) \geq p\}$ for all $S \subseteq \mathcal{S}$ and $p \in [0, 1]$. Now, we endow $\mathcal{A}$ with the product σ-algebra $\boldsymbol{\Sigma}_{\mathcal{A}} = \mathscr{P}\left(\bigcup_{s\in\mathcal{S}} \Theta(s)\right) \otimes \boldsymbol{\Sigma}_{\mathsf{Dist}(\mathcal{S})}$ (i.e., the smallest σ-algebra containing all rectangles $\boldsymbol{K} \times \boldsymbol{M}$ with $\boldsymbol{K} \subseteq \bigcup_{s\in\mathcal{S}} \Theta(s)$ and $\boldsymbol{M} \in \boldsymbol{\Sigma}_{\mathsf{Dist}(\mathcal{S})}$) and let $\mathsf{PMeas}(\mathcal{A})$ be the set of all probability measures on $\boldsymbol{\Sigma}_{\mathcal{A}}$. It is not difficult to check that each set of enabled actions $\mathcal{A}(s)$ is measurable (i.e., $\mathcal{A}(s) \in \boldsymbol{\Sigma}_{\mathcal{A}}$) and that function $\theta(s, \cdot, s')$ is measurable (i.e., $\{a \in \mathcal{A} \mid \theta(s, a, s') \leq p\} \in \boldsymbol{\Sigma}_{\mathcal{A}}$ for all $p \in [0, 1]$).

We extend the definition of *strategy* for the i-player ($i \in \{\Box, \Diamond\}$) in $\mathcal{G}_{\mathcal{K}}$ to be a function $\pi_i : \mathcal{S}^*\mathcal{S}_i \to \mathsf{PMeas}(\mathcal{A})$ that assigns a probability measure to each finite sequence of states such that $\pi_i(\hat{\rho}s)(\mathcal{A}(s)) = 1$. All other concepts on strategies defined in Sec. 3 apply to this new definition as well.

In the following we present the formal definition of $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_{\mathcal{K}},s}$. First, for each $n \geq 0$ and $s \in \mathcal{S}$, define $\mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_{\mathcal{K}},s} : \mathcal{S}^{n+1} \to [0, 1]$ for all $s' \in \mathcal{S}$ and $\hat{\rho} \in \mathcal{S}^{n+1}$ inductively as follows:

$$\mathbb{P}^{\pi_\Box,\pi_\Diamond,0}_{\mathcal{G}_{\mathcal{K}},s}(s') = \delta_s(s')$$

$$\mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_{\mathcal{K}},s}(\hat{\rho}s') = \begin{cases} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_{\mathcal{K}},s}(\hat{\rho}) \displaystyle\int_{\mathcal{A}} \theta(\mathit{last}(\hat{\rho}), \cdot, s') \, \mathbf{d}(\pi_\Box(\hat{\rho})(\cdot)) & \text{if } \mathit{last}(\hat{\rho}) \in \mathcal{S}_\Box \\ \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_{\mathcal{K}},s}(\hat{\rho}) \displaystyle\int_{\mathcal{A}} \theta(\mathit{last}(\hat{\rho}), \cdot, s') \, \mathbf{d}(\pi_\Diamond(\hat{\rho})(\cdot)) & \text{if } \mathit{last}(\hat{\rho}) \in \mathcal{S}_\Diamond \end{cases}$$

and extend $\mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_{\mathcal{K}},s} : \mathscr{P}(\mathcal{S}^{n+1}) \to [0, 1]$ to sets as the sum of the points.

Let $\boldsymbol{\Sigma}_{\mathcal{S}}$ denote the discrete σ-algebra on $\mathcal{S}$ and $\boldsymbol{\Sigma}_{\mathcal{S}^\omega}$ the usual product σ-algebra on $\mathcal{S}^\omega$. By Carathéodory extension theorem [2], $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_{\mathcal{K}},s} : \boldsymbol{\Sigma}_{\mathcal{S}^\omega} \to [0, 1]$ is defined as the unique probability measure such that for all $n \geq 0$, and $S_i \in \boldsymbol{\Sigma}_{\mathcal{S}}$, $0 \leq i \leq n$,

$$\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_{\mathcal{K}},s}(S_0 \times \cdots \times S_n \times \mathcal{S}^\omega) = \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_{\mathcal{K}},s}(S_0 \times \cdots \times S_n)$$

The notions of *deterministic* and *memoryless* extends directly to this type of strategies. In addition, a strategy π_i, $i \in \{\Box, \Diamond\}$, is *semi-Markov* if for every $\hat{\rho}, \hat{\rho}' \in \mathcal{S}^*$ and $s \in \mathcal{S}_i$, $|\hat{\rho}| = |\hat{\rho}'|$ implies $\pi_i(\hat{\rho}s) = \pi_i(\hat{\rho}'s)$, that is, the decisions of π_i depend only on the length of the run and its last state. Thus, we write $\pi_i(n, s)$ instead of $\pi_i(\hat{\rho}s)$ whenever $|\hat{\rho}| = n$. Let Π^S_i denote the set of all semi-Markov strategies for the i-player. Also, we say that a strategy $\pi_i \in \Pi_i$ is *extreme* if for all $\hat{\rho} \in \mathcal{S}^*$, $\pi_i(\hat{\rho}s)(\{(K,\mu) \in \mathcal{A}(s) \mid \mu \in \mathbb{V}(K)\}) = 1$. Notice that extreme strategies only selects transitions on vertices of polytopes. Let Π_i^{XS} and Π_i^{XMD} be, respectively, the set of all extreme semi-Markov strategies and the set of all extreme deterministic and memoryless srategies for the i-player.

Polytopal stochastic games can be translated into simplicial stochastic games preserving all the stochastic behaviour. More precisely, for every PSG $\mathcal{K}$ there is a SSG $\mathcal{K}'$ such that for every pair of strategies for $\mathcal{K}$ in a particular class (i.e., memoryless, semi-Markov, etc.), there is a pair of strategies for $\mathcal{K}'$ in the same class that yields the same probability measure and vice versa. Let $\mathsf{Triang}\colon \mathsf{DPoly} \to \mathscr{P}(\mathsf{DSimp})$ be a function that assigns a vertex-preserving triangulation $\mathsf{Triang}(K)$ to each polytope K. Then:

Proposition 1. *Let $\mathcal{K} = (\mathcal{S}, (\mathcal{S}_\Box, \mathcal{S}_\Diamond), \Theta)$ be a PSG and define the SSG $\mathcal{K}' = (\mathcal{S}, (\mathcal{S}_\Box, \mathcal{S}_\Diamond), \Theta')$ such that $\Theta'(s) = \bigcup_{K \in \Theta(s)} \mathsf{Triang}(K)$. Let $\mathcal{G}_\mathcal{K}$ and $\mathcal{G}_{\mathcal{K}'}$ be their respective interpretations. Then,*

1. *for all pair of strategies $\pi_\Box$ and $\pi_\Diamond$ for $\mathcal{G}_\mathcal{K}$ there is a pair of strategies $\pi'_\Box$ and $\pi'_\Diamond$ for $\mathcal{G}_{\mathcal{K}'}$ such that (a) $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s} = \mathbb{P}^{\pi'_\Box,\pi'_\Diamond}_{\mathcal{G}_{\mathcal{K}'},s}$ for all $s \in \mathcal{S}$, and (b) if π_i, $i \in \{\Box, \Diamond\}$, is memoryless (resp. deterministic, semi-Markov or extreme) then so is π'_i; and*
2. *the same holds with the roles of $\mathcal{G}_\mathcal{K}$ and $\mathcal{G}_{\mathcal{K}'}$ exchanged.*

Proof (Sketch). Let $\mathcal{G}_\mathcal{K} = (\mathcal{S}, (\mathcal{S}_\Box, \mathcal{S}_\Diamond), \mathcal{A}, \theta)$ and $\mathcal{G}_{\mathcal{K}'} = (\mathcal{S}, (\mathcal{S}_\Box, \mathcal{S}_\Diamond), \mathcal{A}', \theta')$. To prove item 1, the new strategies are defined so that they preserve the same measure on the probability part of the labels in $\mathcal{A}'$ as the one the old strategies measure on the probability part of $\mathcal{A}$ while properly distributing the probabilities on the simplices of the triangulation of the original polytopes. For this, first fix a function $f_K : \mathsf{Triang}(K) \to \mathscr{P}(K)$ for each polytope $K \in \mathsf{DPoly}(\mathcal{S})$ satisfying (i) $\forall K' \in \mathsf{Triang}(K)\colon f_K(K') \subseteq K'$, (ii) $\bigcup_{K' \in \mathsf{Triang}(K)} f_K(K') = K$, and (iii) $\forall K'_1, K'_2 \in \mathsf{Triang}(K)\colon f_K(K'_1) \cap f_K(K'_2) \neq \emptyset \Rightarrow K'_1 = K'_2$. Thus, $f_K(K')$ is almost the simplex K' but ensuring that distributions on the faces of K' are exactly in one of the $f_K(K'')$, $K'' \in \mathsf{Triang}(K)$.

Given strategies π_i, $i \in \{\Box, \Diamond\}$, for $\mathcal{G}_\mathcal{K}$ define π'_i for $\mathcal{G}_{\mathcal{K}'}$, for all $\hat{\rho} \in \mathcal{S}^*$, $s \in \mathcal{S}_i$, and $A' \in \boldsymbol{\Sigma}_{\mathcal{A}'}$ by

$$\pi'_i(\hat{\rho}s)(A') = \textstyle\sum_{K \in \Theta(s)} \sum_{K' \in \mathsf{Triang}(K)} \pi_i(\hat{\rho}s)(\{K\} \times (A'|_{K'} \cap f_K(K'))) \quad (1)$$

where $A'|_{K'} = \{\mu \mid (K', \mu) \in A'\}$ is the K' section of the measurable set A'. Notice that f_K ensures that the faces of each $K' \in \mathsf{Triang}(K)$ are considered in exactly one summand of the inner summation of (1).

For item 2, the new strategies preserve the same measure on the probability part of $\mathcal{A}$ as the old strategies while gathering the probability of the simplices in the original polytope. So, for each state $s \in \mathcal{S}$, fix $f_s : \Theta(s) \to \mathscr{P}(\mathsf{DSimp}(\mathcal{S}))$ such that (i) $\forall K \in \Theta(s)\colon f_s(K) \subseteq \mathsf{Triang}(K)$, (ii) $\bigcup_{K \in \Theta(s)} f_s(K) = \bigcup_{K \in \Theta(s)} \mathsf{Triang}(K)$, and (iii) $\forall K_1, K_2 \in \Theta(s)\colon f_s(K_1) \cap f_s(K_2) \neq \emptyset \Rightarrow K_1 = K_2$. Given strategies π'_i, $i \in \{\Box, \Diamond\}$, for $\mathcal{G}_{\mathcal{K}'}$ define π_i for $\mathcal{G}_\mathcal{K}$, for all $\hat{\rho} \in \mathcal{S}^*$, $s \in \mathcal{S}_i$, and $A \in \boldsymbol{\Sigma}_\mathcal{A}$ by

$$\pi_i(\hat{\rho}s)(A) = \textstyle\sum_{K \in \Theta(s)} \sum_{K' \in f_s(K)} \pi'_i(\hat{\rho}s)(\{K'\} \times A|_K) \quad (2)$$

Notice that, by definition, $K' \in \Theta'(s)$. Moreover, notice that f_s ensures that a simplex in a triangulation of a polytope outgoing s is considered in exactly one summand of (2).

In both cases, it requires some straightforward calculations to check that the properties of memoryless, semi-Markov, deterministic, and extreme are preserved by the new strategies. Also in both cases, to prove that $\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s} = \mathbb{P}^{\pi'_\square,\pi'_\diamond}_{\mathcal{G}_{\mathcal{K}'},s}$ it sufficies to state that $\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s} = \mathbb{P}^{\pi'_\square,\pi'_\diamond,n}_{\mathcal{G}_{\mathcal{K}'},s}$ for all $n \geq 0$ which is done by induction using results from measure theory. $\square$

5 Discretizing Polytopal Stochastic Games

In this section we show that a PSG can be solved by translating it into a finite stochastic game that is just like the original PSG but it only has the transitions corresponding to the vertices of the polytopes. We focus on reachability games, and the reward games introduced above: total accumulated reward, total discounted reward, and average reward.

The first lemma we introduce states that the calculation of the expected values of the different reward games only depend on the probability of reaching each state and the reward collected in each state regardless the path that lead to such states. In particular, Lemma 1.1 refers to the reward collected in a finite number of steps while Lemma 1.2 refers to the general case stated before.

For $k \geq 0$ define $\diamond^k s = \mathcal{S}^k \times \{s\} \times \mathcal{S}^\omega$ to be the set of all runs in which $s \in \mathcal{S}$ is reached in exactly k steps. Let $\widehat{\mathsf{rew}}^n_\mathsf{f}(\hat\rho) = \sum_{i=0}^n \mathsf{f}(i,n) r(\hat\rho_i)$ for all $\hat\rho \in \mathcal{S}^{n+1}$. Then $\mathsf{rew}_\mathsf{f}(\rho) = \lim_{n\to\infty} \widehat{\mathsf{rew}}^n_\mathsf{f}(\rho[..n+1])$ where $\rho[..n+1]$ is the $(n+1)$th prefix of ρ, i.e., $\rho[..n+1] = \rho_0\rho_1\rho_2...\rho_n$.

Lemma 1. *Let $\mathcal{G}_\mathcal{K}$ be a stochastic game resulting from interpreting a PSG $\mathcal{K}$. For all strategies $\pi_\square \in \Pi_\square$ and $\pi_\diamond \in \Pi_\diamond$,*

1. $\sum_{\hat\rho\in\mathcal{S}^{n+1}} \mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho)\ \widehat{\mathsf{rew}}^n_\mathsf{f}(\hat\rho) = \sum_{i=0}^n \sum_{s'\in\mathcal{S}} \mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond^i s')\ \mathsf{f}(i,n)\ r(s')$, *for all $n \geq 0$, and*
2. $\mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew}_\mathsf{f}] = \lim_{n\to\infty} \sum_{i=0}^n \sum_{s'\in\mathcal{S}} \mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond^i s')\ \mathsf{f}(i,n)\ r(s')$.

The proof of Lemma 1.1 follows by induction on n while Lemma 1.2 can be calculated using the first item.

The next lemma states that if the $\diamond$-player plays a semi-Markov strategy, the $\square$-player can achieve equal results whether she plays an arbitrary strategy or limits to playing only semi-Markov strategies.

Lemma 2. *Let $\mathcal{G}_\mathcal{K}$ be a stochastic game resulting from interpreting a PSG $\mathcal{K}$. If $\pi_\diamond \in \Pi^S_\diamond$ is a semi-Markov strategy, then, for any $\pi_\square \in \Pi_\square$, there is a semi-Markov strategy $\pi^*_\square \in \Pi^S_\square$ such that:*

1. $\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(D \mathsf{U}^n s') = \mathbb{P}^{\pi^*_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(D \mathsf{U}^n s')$, *for all $n \geq 0$, $D \subseteq \mathcal{S}$ and $s' \in \mathcal{S}$;*
2. $\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C) = \mathbb{P}^{\pi^*_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C)$, *for all $C \subseteq \mathcal{S}$; and*

3. $\mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew_f}] = \mathbb{E}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew_f}]$.

Similarly, if $\pi_\Box \in \Pi^S_\Box$ then, for any $\pi_\Diamond \in \Pi_\Diamond$, there exists $\pi^*_\Diamond \in \Pi^S_\Diamond$ satisfying, mutatis mutandis, the same equalities.

Proof (Sketch). To prove item 1, we define the new strategy $\pi^*_\Box$ so that the probability of choosing from $A \in \boldsymbol{\Sigma}_\mathcal{A}$ after a path of length n ending on a state s with the original strategy is uniformly distributed among the paths of this type in the new strategy. Thus, $\pi^*_\Box$ is formally defined as follows. For $\hat{\rho} \in \mathcal{S}^*$, $s' \in \mathcal{S}$, and $A \in \boldsymbol{\Sigma}_\mathcal{A}$, such that $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(D \mathsf{U}^n\, s') > 0$ and $|\hat{\rho}| = n \geq 0$, let

$$\pi^*_\Box(\hat{\rho}s')(A) = \frac{\sum_{\hat{\rho}' \in D^n} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}'s')\ \pi_\Box(\hat{\rho}'s')(A)}{\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(D \mathsf{U}^n\, s')}$$

For $s' \in \mathcal{S}$ with $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(D \mathsf{U}^n\, s') = 0$ and $|\hat{\rho}s'| = n$, define $\pi^*_\Box(\hat{\rho}s')$ to be $\delta_{\mathsf{f}(s')}$ for a globally fixed function f such that $\mathsf{f}(s') \in \mathcal{A}(s')$. Notice that $\pi^*_\Box \in \Pi^S_\Box$.

Then, the proof of item 1 follows by induction with particular care in the case of $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(D \mathsf{U}^n\, s') = 0$. Item 2 follows straightforwardly from item 1 and item 3 follows directly from item 2 using Lemma 1.2. The proof can be replicated mutatis mutandi with $\Box$ and $\Diamond$ exchanged yielding the last part of the lemma. $\Box$

Since $\Theta(s)$ is finite, there can be finitely many polytopes K such that $(K,\mu) \in \mathcal{A}(s)$. Besides, the set of vertices $\mathbb{V}(K)$ of K is finite. Therefore the set $\{(K,\mu) \in \mathcal{A}(s) \mid \mu \in \mathbb{V}(K)\}$ is also finite and, as a consequence, extreme strategies only resolve with discrete (finite) probability distributions. That is, if π_i is extreme, $\pi_i(\hat{\rho}s)$ has finite support for all $\hat{\rho} \in \mathcal{S}^*$ and $s \in \mathcal{S}$.

It turns out that Lemma 2 can be strengthened to obtain *extreme* semi-Markov strategies. We first prove this new lemma for simplicial stochastic games since simplices have the particular property that any vector in a simplex can be uniquely defined as a convex combination of the simplex vertices which is crucial for the proof of the lemma.

Lemma 3. *Let $\mathcal{G}_\mathcal{K}$ be a stochastic game resulting from interpreting a SSG $\mathcal{K}$. If $\pi_\Diamond \in \Pi^S_\Diamond$ is a semi-Markov strategy, then, for any $\pi_\Box \in \Pi^S_\Box$, there is an extreme semi-Markov strategy $\pi^*_\Box \in \Pi^{XS}_\Box$ such that:*

1. $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(D \mathsf{U}^n\, s') = \mathbb{P}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(D \mathsf{U}^n\, s')$*, for all $n \geq 0$, $D \subseteq \mathcal{S}$ and $s' \in \mathcal{S}$;*
2. $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond C) = \mathbb{P}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond C)$*, for all $C \subseteq \mathcal{S}$; and*
3. $\mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew_f}] = \mathbb{E}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew_f}]$.

Similarly, if $\pi_\Box \in \Pi^S_\Box$ then, for any $\pi_\Diamond \in \Pi^S_\Diamond$, there exists $\pi^*_\Diamond \in \Pi^{XS}_\Diamond$ satisfying, mutatis mutandis, the same equalities.

Proof (Sketch). For any $K \in \mathsf{DSimp}(\mathcal{S})$, $\mu \in K$ and $\hat{\mu} \in \mathbb{V}(K)$ define $\mathsf{p}^K(\mu,\hat{\mu}) \in [0,1]$ such that $\sum_{\hat{\mu} \in \mathbb{V}(K)} \mathsf{p}^K(\mu,\hat{\mu})\ \hat{\mu} = \mu$. That is, all $\mathsf{p}^K(\mu,\hat{\mu})$, $\hat{\mu} \in \mathbb{V}(K)$, are the unique factors that define the convex combination for μ in the simplex K. In any other case, let $\mathsf{p}^K(\mu,\hat{\mu}) = 0$.

Let $\mathsf{p}((K,\mu),(K,\hat{\mu})) = \mathsf{p}^K(\mu,\hat{\mu})$ for all $K \in \mathsf{DSimp}(\mathcal{S})$, $\mu \in K$ and $\hat{\mu} \in \mathbb{V}(K)$, and let $\mathsf{p}(a,b) = 0$ for any other $a,b \in \mathcal{A}$. For every $(K,\mu) \in \mathcal{A}$ such that $\mu \in K$, let $\mathbb{V}(K,\mu) = \{(K,\hat{\mu}) \mid \hat{\mu} \in \mathbb{V}(K)\}$ and let $\mathbb{V}(K,\mu) = \emptyset$ otherwise. Thus, for every $s \in \mathcal{S}$ and $a \in \mathcal{A}$, $\theta(s,a,\cdot) = \sum_{b\in\mathbb{V}(a)} \mathsf{p}(a,b)\ \theta(s,b,\cdot)$.

We also extend p to measurable sets $B \in \boldsymbol{\Sigma}_{\mathcal{A}}$ and $a \in \mathcal{A}$ by $\mathsf{p}(a,B) = \sum_{b\in B\cap\mathbb{V}(a)} \mathsf{p}(a,b)$.

For every $\hat{\rho} \in \mathcal{S}^*$, $s' \in \mathcal{S}$ and $B \in \boldsymbol{\Sigma}_{\mathcal{A}}$, define $\pi^*_\Box$ by

$$\pi^*_\Box(\hat{\rho}s')(B) = \int_{\mathcal{A}} \mathsf{p}(\cdot,B)\ \mathbf{d}(\pi_\Box(\hat{\rho}s')(\cdot)).$$

$\pi^*_\Box(\hat{\rho}s')$ is defined so that it assigns to each vertex of a simplex the weighted contribution (according to $\pi_\Box(\hat{\rho}s')$) of each distribution (in the said simplex) to such vertex.

Because $\pi_\Box$ is semi-Markov, so is $\pi^*_\Box$. Moreover, notice that if b is not a vertex label, then $\mathsf{p}(a,b) = 0$ (and hence $\mathsf{p}(a,B) > 0$ only if B contains vertices). This should hint that $\pi^*_\Box$ is also extreme.

Item 1 proceeds by induction on n. Item 2 follows straightforwardly using 1, and item 3 follows from item 2 using Lemma 1.2. The proof can be replicated mutatis mutandi with $\Box$ and $\Diamond$ exchanged which yields the last part of the lemma. ◻

Because of Proposition 1, Lemma 3 extends immediately to PSG. Moreover, by applying Lemma 3 twice and Proposition 1, we have the next corollary.

Corollary 1. *Let $\mathcal{G}_\mathcal{K}$ be a stochastic game resulting from interpreting a PSG $\mathcal{K}$. For all semi-Markov strategies $\pi_\Diamond \in \Pi^S_\Diamond$ and $\pi_\Box \in \Pi^S_\Box$, there are extreme semi-Markov strategies $\pi^*_\Diamond \in \Pi^{XS}_\Diamond$ and $\pi^*_\Box \in \Pi^{XS}_\Box$ such that*

1. $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond C) = \mathbb{P}^{\pi^*_\Box,\pi^*_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond C)$*, for all $C \subseteq \mathcal{S}$; and*
2. $\mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew_f}] = \mathbb{E}^{\pi^*_\Box,\pi^*_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew_f}]$.

Given $\mathcal{G}_\mathcal{K}$, define the *extreme interpretation of* $\mathcal{K}$ as the stochastic game $\mathcal{H}_\mathcal{K} = (\mathcal{S},(\mathcal{S}_\Box,\mathcal{S}_\Diamond),\mathbb{V}(\mathcal{A}),\theta_{\mathcal{H}_\mathcal{K}})$ where $\theta_{\mathcal{H}_\mathcal{K}}$ is the restriction of θ to actions in $\mathbb{V}(\mathcal{A}) = \{(K,\mu) \in \mathcal{A} \mid \mu \in \mathbb{V}(K)\}$, that is, $\theta_{\mathcal{H}_\mathcal{K}}(s,a,s) = \theta(s,a,s')$ for all $s,s' \in \mathcal{S}$ and $a \in \mathbb{V}(\mathcal{A})$. Since $\mathbb{V}(\mathcal{A})$ is finite, $\mathcal{H}_\mathcal{K}$ is a finite stochastic game.

Given an extreme semi-Markov strategy $\pi_i \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},i}$ for the i-player in the stochastic game $\mathcal{G}_\mathcal{K}$, $i \in \{\Box,\Diamond\}$, define $\pi^{\mathsf{v}}_i(\hat{\rho}s)(A) = \pi_i(\hat{\rho}s)(A)$ for all $\hat{\rho} \in \mathcal{S}^*$, $s \in \mathcal{S}$, and $A \subseteq \mathbb{V}(\mathcal{A})$ ($A \in \boldsymbol{\Sigma}_{\mathcal{A}}$ since it is finite). Notice that $\pi^{\mathsf{v}}_i(\hat{\rho}s)(\mathcal{A}_{\mathcal{H}_\mathcal{K}}(s)) = \pi_i(\hat{\rho}s)(\mathbb{V}(\mathcal{A}(s))) = 1$. Therefore $\pi^{\mathsf{v}}_i \in \Pi^S_{\mathcal{H}_\mathcal{K},i}$ is a semi-Markov strategy in $\mathcal{H}_\mathcal{K}$. Conversely, for a semi-Markov strategy $\pi_i \in \Pi^S_{\mathcal{H}_\mathcal{K},i}$ for the i-player in the stochastic game $\mathcal{H}_\mathcal{K}$, define $\pi^{\mathsf{x}}_i(\hat{\rho}s)(A) = \pi_i(\hat{\rho}s)(A \cap \mathbb{V}(\mathcal{A}))$ for all $\hat{\rho} \in \mathcal{S}^*$, $s \in \mathcal{S}$, and $A \in \boldsymbol{\Sigma}_{\mathcal{A}}$. $\pi^{\mathsf{x}}_i \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},i}$ is a well defined extreme semi-Markov strategy in $\mathcal{G}_\mathcal{K}$ since $\pi^{\mathsf{x}}_i(\hat{\rho}s)(\mathbb{V}(\mathcal{A}(s))) = \pi_i(\hat{\rho}s)(\mathcal{A}_{\mathcal{H}_\mathcal{K}}(s)) = 1$ and $\pi^{\mathsf{x}}_i(\hat{\rho}s)(\mathcal{A} \setminus \mathbb{V}(\mathcal{A})) = \pi_i(\hat{\rho}s)(\emptyset) = 0$. Then, it can be calculated by induction on n that $\mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s} = \mathbb{P}^{\pi^{\mathsf{v}}_\Box,\pi^{\mathsf{v}}_\Diamond,n}_{\mathcal{H}_\mathcal{K},s}$ and $\mathbb{P}^{\pi^{\mathsf{x}}_\Box,\pi^{\mathsf{x}}_\Diamond,n}_{\mathcal{G}_\mathcal{K},s} = \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{H}_\mathcal{K},s}$ which yield

$$\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s} = \mathbb{P}^{\pi^{\mathsf{v}}_\Box,\pi^{\mathsf{v}}_\Diamond}_{\mathcal{H}_\mathcal{K},s} \text{ and } \mathbb{P}^{\pi^{\mathsf{x}}_\Box,\pi^{\mathsf{x}}_\Diamond}_{\mathcal{G}_\mathcal{K},s} = \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}. \tag{3}$$

This suggests that the solution of a PSG under extreme semi-Markov strategies is equivalent to the solution the game on its extreme interpretation limited to semi-Markov strategies, which is stated in the following:

Proposition 2. *Let $\mathcal{G}_\mathcal{K}$ and $\mathcal{H}_\mathcal{K}$ be respectively the interpretation and the extreme interpretation of $\mathcal{K}$. Then, the following equalities hold*

1. $\inf_{\pi_\Diamond \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Diamond}} \sup_{\pi_\Box \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond C) = \inf_{\pi_\Diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Diamond}} \sup_{\pi_\Box \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond C)$
2. $\sup_{\pi_\Box \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\Diamond \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Diamond}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond C) = \sup_{\pi_\Box \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}} \inf_{\pi_\Diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Diamond}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond C)$
3. $\inf_{\pi_\Diamond \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Diamond}} \sup_{\pi_\Box \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew_f}] = \inf_{\pi_\Diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Diamond}} \sup_{\pi_\Box \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{E}^{\check{\pi}_\Box,\check{\pi}_\Diamond}_{\mathcal{H}_\mathcal{K},s}[\mathsf{rew_f}]$
4. $\sup_{\pi_\Box \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\Diamond \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Diamond}} \mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew_f}] = \sup_{\pi_\Box \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}} \inf_{\pi_\Diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Diamond}} \mathbb{E}^{\check{\pi}_\Box,\check{\pi}_\Diamond}_{\mathcal{H}_\mathcal{K},s}[\mathsf{rew_f}]$

The next proposition, whose proof also uses (3), provides necessary conditions for the polytopal stochastic game to be almost surely stopping or irreducible in terms of the extreme interpretation.

Proposition 3. *Let $\mathcal{G}_\mathcal{K}$ and $\mathcal{H}_\mathcal{K}$ be respectively the interpretation and the extreme interpretation of $\mathcal{K}$. Then, (1) if $\mathcal{G}_\mathcal{K}$ is almost surely stopping, so is $\mathcal{H}_\mathcal{K}$, and (2) if $\mathcal{G}_\mathcal{K}$ is irreducible, so is $\mathcal{H}_\mathcal{K}$.*

Notice that by fixing one strategy in $\mathcal{H}_\mathcal{K}$ to be the memoryless, the remaining structure is a Markov decision process. Then the statements in the following proposition are consequences of standard results in MDP [27].

Proposition 4. *For all $\pi^*_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}$ and $\pi^*_\Diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Diamond}$,*

1. $\sup_{\pi_\Box \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi^*_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond C) = \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi^*_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond C)$*;*
2. $\inf_{\pi_\Diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Diamond}} \mathbb{P}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond C) = \inf_{\pi_\Diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Diamond}} \mathbb{P}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond C)$*;*
3. $\sup_{\pi_\Box \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi^*_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f}) = \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi^*_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f})$*, provided* $\mathbb{E}^{\pi_\Box,\pi^*_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f})$ *is defined for all $\pi_\Box \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}$; and*
4. $\inf_{\pi_\Diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Diamond}} \mathbb{E}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f}) = \inf_{\pi_\Diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Diamond}} \mathbb{E}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f})$*, provided* $\mathbb{E}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f})$ *is defined for all $\pi_\Diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Diamond}$.*

We are now in conditions to present our main result. The following theorem is two folded. On the one hand, it states that the polytopal stochastic games of all quantitative objectives of interest in this paper –namely, quantitative reachability, expected total accumulated reward, expected discounted accumulated rewards, and expected average rewards– are determined. On the other hand, it states that these objectives for PSG can be equivalently solved in its extreme interpretation.

Theorem 1. *Let $\mathcal{G}_\mathcal{K}$ and $\mathcal{H}_\mathcal{K}$ be respectively the interpretation and the extreme interpretation of $\mathcal{K}$. Then,*

1. $$\inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C) = \inf_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\diamond C) = \\ = \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \inf_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\diamond C) = \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C)$$
for all $C \subseteq \mathcal{S}$; and
2. $$\inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) = \inf_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f}) = \\ = \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \inf_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f}) = \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}),$$
provided: (a) $\mathcal{G}_\mathcal{K}$ is almost surely stopping whenever $\mathsf{rew_f} = \mathsf{rew_t}$, and (b) $\mathcal{G}_\mathcal{K}$ is irreducble whenever $\mathsf{rew_f} = \mathsf{rew_a}$.

Proof. For item 2 we calculate as follows:

$$\begin{aligned}
&\inf\nolimits_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \sup\nolimits_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) \\
&\quad \le \inf\nolimits_{\pi_\diamond \in \Pi^{S}_{\mathcal{G}_\mathcal{K},\diamond}} \sup\nolimits_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) && (\Pi^{S}_{\mathcal{G}_\mathcal{K},\diamond} \subseteq \Pi_{\mathcal{G}_\mathcal{K},\diamond}) \\
&\quad = \inf\nolimits_{\pi_\diamond \in \Pi^{S}_{\mathcal{G}_\mathcal{K},\diamond}} \sup\nolimits_{\pi_\Box \in \Pi^{S}_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) && \text{(by Lemma 2.3)} \\
&\quad = \inf\nolimits_{\pi_\diamond \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\diamond}} \sup\nolimits_{\pi_\Box \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) && \text{(by Corollary 1.2)} \\
&\quad = \inf\nolimits_{\pi_\diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\diamond}} \sup\nolimits_{\pi_\Box \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f}) && \text{(by Prop. 2.3)} \\
&\quad \le \inf\nolimits_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \sup\nolimits_{\pi_\Box \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f}) && (\Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond} \subseteq \Pi^{S}_{\mathcal{H}_\mathcal{K},\diamond}) \\
&\quad = \inf\nolimits_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \sup\nolimits_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f}) && \text{(by Prop. 4.3)} \\
&\quad = \sup\nolimits_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \inf\nolimits_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f}) && (*) \\
&\quad = \sup\nolimits_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \inf\nolimits_{\pi_\diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f}) && \text{(by Prop. 4.4)} \\
&\quad \le \sup\nolimits_{\pi_\Box \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}} \inf\nolimits_{\pi_\diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew_f}) && (\Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box} \subseteq \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}) \\
&\quad = \sup\nolimits_{\pi_\Box \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Box}} \inf\nolimits_{\pi_\diamond \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) && \text{(by Prop. 2.4)} \\
&\quad = \sup\nolimits_{\pi_\Box \in \Pi^{S}_{\mathcal{G}_\mathcal{K},\Box}} \inf\nolimits_{\pi_\diamond \in \Pi^{S}_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) && \text{(by Corollary 1.2)} \\
&\quad = \sup\nolimits_{\pi_\Box \in \Pi^{S}_{\mathcal{G}_\mathcal{K},\Box}} \inf\nolimits_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) && \text{(by Lemma 2.3)} \\
&\quad \le \sup\nolimits_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \inf\nolimits_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) && (\Pi^{S}_{\mathcal{G}_\mathcal{K},\Box} \subseteq \Pi_{\mathcal{G}_\mathcal{K},\Box}) \\
&\quad \le \inf\nolimits_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \sup\nolimits_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) && \text{(by prop. of sup and inf)}
\end{aligned}$$

Since the last term is equal to the first term in the calculation, item 2 is concluded. In particular, step (*) is justified as follows, depending on $\mathsf{rew_f}$: For $\mathsf{rew_f} = \mathsf{rew_t}$, (*) follows by [15, Theorem 4.2.6] since, by Proposition 3.(1), the game $\mathcal{H}_\mathcal{K}$ is also almost surely stopping. For $\mathsf{rew_f} = \mathsf{rew}_\gamma$ (*) follows by [15, Theorem 4.3.2]. For $\mathsf{rew_f} = \mathsf{rew_a}$ (*) follows by [15, Theorem 5.1.5] since, by Proposition 3.(2), the game $\mathcal{H}_\mathcal{K}$ is also irreducible.

Item 1 of the theorem follows similarly. In each step, propositions, lemmas and corollaries are the same only differing on the item, while step (*) follows from [13, Lemma 6]. ◻

Since extreme interpretations are finite, the values of the different games can be calculated following known algorithms [13,15]. Thus, Theorem 1 immediately provides an algorithmic solution for PSGs.

The number of vertices of a polytope grows exponentially in the dimension of the polytope [19]. More precisely if d is the dimension of a polytope K and m is the number of inequalities that defines it, $\mathbb{V}(K) \sim \Omega(m^{\lfloor d/2 \rfloor})$. This implies that the extreme interpretation $\mathcal{H}_{\mathcal{K}}$ grows exponentially on the largest size of the support sets of the distributions involved in the original PSG $\mathcal{K}$ which we expect not to be too large. (In our example of Sec. 2, $\lfloor d/2 \rfloor = 2$)

Condon [13] showed that deciding reachability in stochastic games is in $\mathsf{NP} \cap \mathsf{coNP}$. Despite the exponential grow, this is still our case as we show in the following. Let $Val_s(\mathcal{K})$ denote the value of the game at state s, that is, it is equal to $\sup_{\pi_\Box \in \Pi_\Box} \inf_{\pi_\Diamond \in \Pi_\Diamond} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_{\mathcal{K}},s}(\Diamond G)$, or $\sup_{\pi_\Box \in \Pi_\Box} \inf_{\pi_\Diamond \in \Pi_\Diamond} \mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_{\mathcal{K}},s}[\mathsf{rew_f}]$. The problem is then to decide whether $Val_s(\mathcal{K}) \geq q$, for a given $q \in \mathbb{Q}$ and $s \in \mathcal{S}$ Since for all the cases (total reward, discounted reward, average reward and reachability objectives under the respective conditions) the value $Val_s(\mathcal{K})$ of the game can be achieved with an extreme memoryless and deterministic strategies, we can reason as follows: (i) guess a memoryless and deterministic strategy for each player, (ii) on the resulting Markov chain compute the corresponding measure (i.e. total reward, discounted reward, average reward or reachability) on the respective set of linear equations, which can be done in polynomial time (for $\mathsf{rew_a}$ an extra linear summation is needed) [22], (iii) verify if it is a fixpoint of Bellman equations (for reachability, discounted, or total reward), or a fixpoint of the Alg. 5.1.1 of [15], in the case of average reward, and (iv) check whether $Val_s(\mathcal{K}) \geq q$. This puts our problem in NP. With the same process we can check whether $Val_s(\mathcal{K}) < q$ which puts the problem also in coNP. Hence we have the next theorem.

Theorem 2. *For any PSG $\mathcal{K}$, $q \in \mathbb{Q}$, and $s \in \mathcal{S}$, the problem of deciding whether $Val_s(\mathcal{K}) \geq q$ is in $\mathsf{NP} \cap \mathsf{coNP}$. For $\mathsf{rew_f} \in \{\mathsf{rew_t}, \mathsf{rew_a}\}$ the decision problem is restricted to $\mathcal{G}_{\mathcal{K}}$ being almost surely stopping and irreducible, respectively.*

6 Concluding remarks

We believe that polytopal games may have several applications in practice, particularly, in scenarios where the probabilities are not exact but can be characterized with linear equations. We observe that one may expect that the number of vertices of the polytopes keep small in practical examples, hence the game discretization may have no impact on the runtime of a tool implementing the approach described in the paper. However, we leave as further work the implementation of such a tool and an in-depth evaluation of it.

In addition, it would be also be of interest to explore other types of objectives, including ω-regular objectives as already study for standard stochastic games in [9] or even solving stochastic games for conditional probabilities of temporal properties or conditional expectations of rewards models as widely studied by Christel Baier and her team in the context of Markov decision processes [5,6,23,25].

References

1. Akshay, S., Bouyer, P., Krishna, S.N., Manasa, L., Trivedi, A.: Stochastic timed games revisited. In: Faliszewski, P., Muscholl, A., Niedermeier, R. (eds.) 41st International Symposium on Mathematical Foundations of Computer Science, MFCS 2016. LIPIcs, vol. 58, pp. 8:1–8:14. Schloss Dagstuhl - Leibniz-Zentrum für Informatik (2016). https://doi.org/10.4230/LIPICS.MFCS.2016.8
2. Ash, R.B., Doléans-Dade, C.A.: Probability and Measure Theory. Harcourt/Academic Press, 2nd edn. (1999)
3. Aslanyan, Z., Nielson, F., Parker, D.: Quantitative verification and synthesis of attack-defence scenarios. In: IEEE 29th Computer Security Foundations Symposium, CSF 2016. pp. 105–119. IEEE Computer Society (2016). https://doi.org/10.1109/CSF.2016.15
4. Baier, C., Katoen, J.: Principles of model checking. MIT Press (2008)
5. Baier, C., Klein, J., Klüppelholz, S., Märcker, S.: Computing conditional probabilities in Markovian models efficiently. In: Ábrahám, E., Havelund, K. (eds.) Tools and Algorithms for the Construction and Analysis of Systems - 20th International Conference, TACAS 2014, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2014, Grenoble, France, April 5-13, 2014. Proceedings. Lecture Notes in Computer Science, vol. 8413, pp. 515–530. Springer (2014). https://doi.org/10.1007/978-3-642-54862-8_43
6. Baier, C., Klein, J., Klüppelholz, S., Wunderlich, S.: Maximizing the conditional expected reward for reaching the goal. In: Legay, A., Margaria, T. (eds.) Tools and Algorithms for the Construction and Analysis of Systems - 23rd International Conference, TACAS 2017, Proceedings, Part II. Lecture Notes in Computer Science, vol. 10206, pp. 269–285 (2017). https://doi.org/10.1007/978-3-662-54580-5_16
7. Bouyer, P., Forejt, V.: Reachability in stochastic timed games. In: Albers, S., Marchetti-Spaccamela, A., Matias, Y., Nikoletseas, S.E., Thomas, W. (eds.) Automata, Languages and Programming, 36th Internatilonal Colloquium, ICALP 2009, Proceedings, Part II. Lecture Notes in Computer Science, vol. 5556, pp. 103–114. Springer (2009). https://doi.org/10.1007/978-3-642-02930-1_9
8. Castro, P.F., D'Argenio, P.R., Demasi, R., Putruele, L.: Quantifying masking fault-tolerance via fair stochastic games. In: Mezzina, C.A., Caltais, G. (eds.) Proceedings Combined 30th International Workshop on Expressiveness in Concurrency and 20th Workshop on Structural Operational Semantics, EXPRESS/SOS 2023, and 20th Workshop on Structural Operational Semantics. EPTCS, vol. 387, pp. 132–148 (2023). https://doi.org/10.4204/EPTCS.387.10
9. Chatterjee, K., de Alfaro, L., Henzinger, T.A.: The complexity of stochastic Rabin and Streett games'. In: Caires, L., Italiano, G.F., Monteiro, L., Palamidessi, C., Yung, M. (eds.) Automata, Languages and Programming, 32nd International Colloquium, ICALP 2005, Proceedings. Lecture Notes in Computer Science, vol. 3580, pp. 878–890. Springer (2005). https://doi.org/10.1007/11523468_71
10. Chatterjee, K., Henzinger, T.A.: A survey of stochastic ω-regular games. J. Comput. Syst. Sci. **78**(2), 394–413 (2012). https://doi.org/10.1016/j.jcss.2011.05.002
11. Chen, T., Forejt, V., Kwiatkowska, M.Z., Parker, D., Simaitis, A.: Prism-games: A model checker for stochastic multi-player games. In: Piterman, N., Smolka, S.A. (eds.) Tools and Algorithms for the Construction and Analysis of Systems - 19th International Conference, TACAS 2013. Proceedings. Lecture Notes in Computer Science, vol. 7795, pp. 185–191. Springer (2013). https://doi.org/10.1007/978-3-642-36742-7_13

12. Chen, T., Kwiatkowska, M.Z., Simaitis, A., Wiltsche, C.: Synthesis for multi-objective stochastic games: An application to autonomous urban driving. In: Joshi, K.R., Siegle, M., Stoelinga, M., D'Argenio, P.R. (eds.) Quantitative Evaluation of Systems - 10th International Conference, QEST 2013. Proceedings. Lecture Notes in Computer Science, vol. 8054, pp. 322–337. Springer (2013). https://doi.org/10.1007/978-3-642-40196-1_28
13. Condon, A.: The complexity of stochastic games. Inf. Comput. **96**(2), 203–224 (1992). https://doi.org/10.1016/0890-5401(92)90048-K
14. Feng, L., Wiltsche, C., Humphrey, L.R., Topcu, U.: Synthesis of human-in-the-loop control protocols for autonomous systems. IEEE Trans Autom. Sci. Eng. **13**(2), 450–462 (2016). https://doi.org/10.1109/TASE.2016.2530623
15. Filar, J., Vrieze, K.: Competitive Markov Decision Processes. Springer-Verlag, Berlin, Heidelberg (1996)
16. Giry, M.: A categorical approach to probability theory. In: Banaschewski, B. (ed.) Categorical Aspects of Topology and Analysis. Lecture Notes in Mathematics, vol. 915, pp. 68–85. Springer (1982). https://doi.org/10.1007/BFb0092872
17. Jonsson, B., Larsen, K.G.: Specification and refinement of probabilistic processes. In: Proceedings of the 6th Annual Symposium on Logic in Computer Science (LICS'91). pp. 266–277. IEEE Computer Society (1991). https://doi.org/10.1109/LICS.1991.151651
18. Junges, S., Jansen, N., Katoen, J., Topcu, U., Zhang, R., Hayhoe, M.M.: Model checking for safe navigation among humans. In: McIver, A., Horváth, A. (eds.) Quantitative Evaluation of Systems - 15th International Conference, QEST 2018, Proceedings. Lecture Notes in Computer Science, vol. 11024, pp. 207–222. Springer (2018). https://doi.org/10.1007/978-3-319-99154-2_13
19. Kaibel, V., Pfetsch, M.E.: Some algorithmic problems in polytope theory. In: Joswig, M., Takayama, N. (eds.) Algebra, Geometry, and Software Systems [outcome of a Dagstuhl seminar]. pp. 23–47. Springer (2003). https://doi.org/10.1007/978-3-662-05148-1_2
20. Kozine, I., Utkin, L.V.: Interval-valued finite markov chains. Reliab. Comput. **8**(2), 97–113 (2002). https://doi.org/10.1023/A:1014745904458
21. Kucera, A.: Turn-based stochastic games. In: Apt, K.R., Grädel, E. (eds.) Lectures in Game Theory for Computer Scientists, pp. 146–184. Cambridge University Press (2011)
22. Kulkarni, V.G.: Modeling and Analysis of Stochastic Systems. Texts in Statistical Science, CRC Press, 3rd edn. (2017)
23. Märcker, S., Baier, C., Klein, J., Klüppelholz, S.: Computing conditional probabilities: Implementation and evaluation. In: Cimatti, A., Sirjani, M. (eds.) Software Engineering and Formal Methods - 15th International Conference, SEFM 2017, Proceedings. Lecture Notes in Computer Science, vol. 10469, pp. 349–366. Springer (2017). https://doi.org/10.1007/978-3-319-66197-1_22
24. McMullen, P.: Geometric Regular Polytopes. Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge (2020)
25. Piribauer, J., Baier, C.: Partial and conditional expectations in Markov decision processes with integer weights. In: Bojanczyk, M., Simpson, A. (eds.) Foundations of Software Science and Computation Structures - 22nd International Conference, FOSSACS 2019, Proceedings. Lecture Notes in Computer Science, vol. 11425, pp. 436–452. Springer (2019). https://doi.org/10.1007/978-3-030-17127-8_25
26. Puggelli, A., Li, W., Sangiovanni-Vincentelli, A.L., Seshia, S.A.: Polynomial-time verification of PCTL properties of mdps with convex uncertainties. In: Sharygina,

N., Veith, H. (eds.) Computer Aided Verification - 25th International Conference. Proceedings. Lecture Notes in Computer Science, vol. 8044, pp. 527–542. Springer (2013). https://doi.org/10.1007/978-3-642-39799-8_35
27. Puterman, M.L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley Series in Probability and Statistics, Wiley (1994). https://doi.org/10.1002/9780470316887
28. R. S. Sutton, A.G.B..: Reinforcement Learning: An Introduction. Bradford Books (2018)
29. Sen, K., Viswanathan, M., Agha, G.: Model-checking markov chains in the presence of uncertainties. In: Hermanns, H., Palsberg, J. (eds.) Tools and Algorithms for the Construction and Analysis of Systems, 12th International Conference, TACAS 2006, Proceedings. Lecture Notes in Computer Science, vol. 3920, pp. 394–410. Springer (2006). https://doi.org/10.1007/11691372_26
30. Shapley, L.S.: Stochastic games. Proc. Natl. Acad. Sci. USA **39**(10), 1095–100 (1953). https://doi.org/10.1073/pnas.39.10.1095
31. Wang, M., Wang, Z., Talbot, J., Gerdes, J.C., Schwager, M.: Game-theoretic planning for self-driving cars in multivehicle competitive scenarios. IEEE Trans. Robotics **37**(4), 1313–1325 (2021). https://doi.org/10.1109/TRO.2020.3047521
32. Ziegler, G.M.: Lectures on polytopes, Graduate texts in mathematics, vol. 152. Springer-Verlag, New York (1995)

A Full proofs

Proof (of Proposition 1). Let $\mathcal{G}_{\mathcal{K}} = (\mathcal{S}, (\mathcal{S}_\Box, \mathcal{S}_\Diamond), \mathcal{A}, \theta)$ with $\mathcal{A} = \left(\bigcup_{s\in\mathcal{S}} \Theta(s)\right) \times \mathsf{Dist}(\mathcal{S})$ and $\mathcal{G}_{\mathcal{K}'} = (\mathcal{S}, (\mathcal{S}_\Box, \mathcal{S}_\Diamond), \mathcal{A}', \theta')$ with $\mathcal{A}' = \left(\bigcup_{s\in\mathcal{S}} \Theta'(s)\right) \times \mathsf{Dist}(\mathcal{S}) = \left(\bigcup_{\substack{s\in\mathcal{S},\\ K\in\Theta(s)}} \mathsf{Triang}(K)\right) \times \mathsf{Dist}(\mathcal{S})$, be the respective interpretations of $\mathcal{K}$ and $\mathcal{K}'$.

To prove item 1, first fix a function $f_K : \mathsf{Triang}(K) \to \mathscr{P}(K)$ for each polytope $K \in \mathsf{DPoly}(\mathcal{S})$ satisfying

(i) $\forall K' \in \mathsf{Triang}(K)\colon f_K(K') \subseteq K'$,
(ii) $\bigcup_{K'\in\mathsf{Triang}(K)} f_K(K') = K$, and
(iii) $\forall K_1', K_2' \in \mathsf{Triang}(K)\colon f_K(K_1') \cap f_K(K_2') \neq \emptyset \Rightarrow K_1' = K_2'$.

Thus, $f_K(K')$ is almost the simplex K' but ensuring that distributions on the faces of K' are exactly in one of the $f_K(K'')$, $K'' \in \mathsf{Triang}(K)$.

Now, let $\pi_\Box$ and $\pi_\Diamond$ be a pair of strategies for $\mathcal{G}_{\mathcal{K}}$. Define π_i', $i \in \{\Box, \Diamond\}$, for all $\hat{\rho} \in \mathcal{S}^*$, $s \in \mathcal{S}_i$, and $A' \in \boldsymbol{\Sigma}_{\mathcal{A}'}$ by

$$\pi_i'(\hat{\rho}s)(A') = \sum_{K\in\Theta(s)} \sum_{K'\in\mathsf{Triang}(K)} \pi_i(\hat{\rho}s)(\{K\} \times (A'|_{K'} \cap f_K(K'))) \tag{4}$$

where $A'|_{K'} = \{\mu \mid (K', \mu) \in A'\}$ is the K' section of the measurable set A'. Notice that f_K ensures that the faces of each $K' \in \mathsf{Triang}(K)$ are considered in exactly one summand of the inner summation of (4). Thus, $\pi_i'(\hat{\rho})$ is a well defined probability measure and hence $\pi_\Box'$ and $\pi_\Diamond'$ is a pair of strategies for $\mathcal{G}_{\mathcal{K}'}$.

It is straightforward to check that if π_i is memoryless or semi-Markov, so is π'_i. Suppose π_i is extreme, then we have that

$$
\begin{aligned}
&\pi'_i(\hat{\rho}s)(\{(K',\mu) \in \mathcal{A}'(s) \mid \mu \in \mathbb{V}(K')\}) \\
&\qquad = \sum_{K \in \Theta(s)} \sum_{K' \in \mathsf{Triang}(K)} \pi_i(\hat{\rho}s)(\{K\} \times (\mathbb{V}(K') \cap f_K(K'))) && (5) \\
&\qquad = \sum_{K \in \Theta(s)} \pi_i(\hat{\rho}s)\Bigg(\{K\} \times \Bigg(\bigcup_{K' \in \mathsf{Triang}(K)} \mathbb{V}(K') \cap f_K(K')\Bigg)\Bigg) && (6) \\
&\qquad = \sum_{K \in \Theta(s)} \pi_i(\hat{\rho}s)(\{K\} \times \mathbb{V}(K)) && (7) \\
&\qquad = \pi_i(\hat{\rho}s)\Bigg(\bigcup_{K \in \Theta(s)} \{K\} \times \mathbb{V}(K)\Bigg) && (8) \\
&\qquad = \pi_i(\hat{\rho}s)(\{(K,\mu) \in \mathcal{A}(s) \mid \mu \in \mathbb{V}(K)\}) && (9) \\
&\qquad = 1 && (10)
\end{aligned}
$$

Equality (5) follows by (4) and (6) is a consequence of π_i being a measure and the fact that f_K guarantees the disjointness of sets in the union. f_K also guarantees that no vertix of K is lost, hence (7). (8) is a consequence of π_i being a measure and (9) by definition of $\mathcal{A}(s)$ and $\theta(s)$. Finally, (10) follows from π_i being extreme.

Let π_i be deterministic and suppose $\pi_i(\hat{\rho}s)(\{(K_\star,\mu)\}) = 1$ for $K_\star \in \Theta(s)$ and $\mu \in K_\star$. Besides, suppose that $K'_\star \in \mathsf{Triang}(K_\star)$ such that $\mu \in f_{K_\star}(K'_\star)$. Then

$$
\begin{aligned}
&\pi'_i(\hat{\rho}s)(\{(K'_\star,\mu)\}) \\
&\qquad = \sum_{K \in \Theta(s)} \sum_{K' \in \mathsf{Triang}(K)} \pi_i(\hat{\rho}s)(\{K\} \times (\{(K'_\star,\mu)\}|_{K'} \cap f_K(K'))) && (11) \\
&\qquad = \pi_i(\hat{\rho}s)(\{(K_\star,\mu)\}) && (12) \\
&\qquad = 1 && (13)
\end{aligned}
$$

Equality (11) follows by (4). (12) follows from the fact that all summands are 0 except for the one in which $K = K_\star$ and $K' = K'_\star$. Finally (13) follows because π_i is deterministic with $\pi_i(\hat{\rho}s)(\{(K_\star,\mu)\}) = 1$ by assumption.

To prove that $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s} = \mathbb{P}^{\pi'_\Box,\pi'_\Diamond}_{\mathcal{G}_{\mathcal{K}'},s}$ it sufficies to state that $\mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s} = \mathbb{P}^{\pi'_\Box,\pi'_\Diamond,n}_{\mathcal{G}_{\mathcal{K}'},s}$ for all $n \geq 0$ which we show by induction in the following.

For $n = 0$, $\mathbb{P}^{\pi_\Box,\pi_\Diamond,0}_{\mathcal{G}_\mathcal{K},s}(s') = \delta_s(s') = \mathbb{P}^{\pi'_\Box,\pi'_\Diamond,0}_{\mathcal{G}_{\mathcal{K}'},s}(s')$. For $n+1 > 0$ we calculate as follows. Suppose $\hat{\rho} \in \mathcal{S}^n$, $s'' \in \mathcal{S}$ and $s' \in \mathcal{S}_i$. Then,

$$
\begin{aligned}
&\mathbb{P}^{\pi'_\Box,\pi'_\Diamond,n+1}_{\mathcal{G}_{\mathcal{K}'},s}(\hat{\rho}s''s') \\
&\quad = \mathbb{P}^{\pi'_\Box,\pi'_\Diamond,n}_{\mathcal{G}_{\mathcal{K}'},s}(\hat{\rho}s'') \int_{\mathcal{A}'} \theta'(s'',\cdot,s') \,\mathbf{d}(\pi'_i(\hat{\rho}s'')(\cdot)) && (14) \\
&\quad = \mathbb{P}^{\pi'_\Box,\pi'_\Diamond,n}_{\mathcal{G}_{\mathcal{K}'},s}(\hat{\rho}s'') \sum_{\substack{K \in \Theta(s'') \\ K' \in \mathsf{Triang}(K)}} \int_{\{K'\} \times f_K(K')} \theta'(s'',\cdot,s') \,\mathbf{d}(\pi'_i(\hat{\rho}s'')(\cdot)) && (15)
\end{aligned}
$$

$$= \mathbb{P}^{\pi'_\square,\pi'_\diamond,n}_{\mathcal{G}_{\mathcal{K}'},s}(\hat{\rho}s'') \sum_{\substack{K\in\Theta(s'')\\ K'\in\mathsf{Triang}(K)}} \int_{f_K(K')} \theta'(s'',(K',\cdot),s')\ \mathbf{d}(\pi'_i(\hat{\rho}s'')(K',\cdot)) \tag{16}$$

$$= \mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_{\mathcal{K}},s}(\hat{\rho}s'') \sum_{\substack{K\in\Theta(s'')\\ K'\in\mathsf{Triang}(K)}} \int_{f_K(K')} \theta(s'',(K,\cdot),s')\ \mathbf{d}(\pi_i(\hat{\rho}s'')(K,\cdot)) \tag{17}$$

$$= \mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_{\mathcal{K}},s}(\hat{\rho}s'') \sum_{K\in\Theta(s'')} \int_{\left(\bigcup_{K'\in\mathsf{Triang}(K)} f_K(K')\right)} \theta(s'',(K,\cdot),s')\ \mathbf{d}(\pi_i(\hat{\rho}s'')(K,\cdot)) \tag{18}$$

$$= \mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_{\mathcal{K}},s}(\hat{\rho}s'') \int_{\left(\bigcup_{K\in\Theta(s'')}\{K\}\times K\right)} \theta(s'',\cdot,s')\ \mathbf{d}(\pi_i(\hat{\rho}s'')(\cdot)) \tag{19}$$

$$= \mathbb{P}^{\pi_\square,\pi_\diamond,n+1}_{\mathcal{G}_{\mathcal{K}},s}(\hat{\rho}s''s') \tag{20}$$

Equality (14) is the definition of $\mathbb{P}^{\pi'_\square,\pi'_\diamond,n+1}_{\mathcal{G}_{\mathcal{K}'},s}$. (15) follows by calculations and noting that $\pi'_i(\hat{\rho}s'')\left(\mathcal{A}'\setminus\left(\bigcup_{\substack{K\in\Theta(s'')\\ K'\in\mathsf{Triang}(K)}}\{K'\}\times f_K(K')\right)\right)=0$. (16) is a consequence of Fubini's theorem. (17) follows by induction hypothesis and the easy-to-check equalities $\theta'(s'',(K',\mu),s')=\theta(s'',(K,\mu),s')$, for all $\mu\in f_K(K')$, and $\pi'_i(\hat{\rho}s'')(\{K'\}\times B\cap f_K(K')))=\pi_i(\hat{\rho}s'')(\{K\}\times B\cap f_K(K')))$, for all $B\in\boldsymbol{\Sigma}_{\mathsf{Dist}(\mathcal{S})}$, $K\in\Theta(s'')$ and $K'\in\mathsf{Triang}(K)$. (18) follows by calculations and (19) follows by noting that $K=\bigcup_{K'\in\mathsf{Triang}(K)} f_K(K')$ and using Fubini's Theorem. Finally, (20) follows by observing that $\pi_i(\hat{\rho}s'')\left(\mathcal{A}\setminus\left(\bigcup_{K\in\Theta(s'')}\{K\}\times K\right)\right)=0$ and by the definition of $\mathbb{P}^{\pi_\square,\pi_\diamond,n+1}_{\mathcal{G}_{\mathcal{K}},s}$.

To prove item 2, first fix a function f_s for each state $s\in\mathcal{S}$ such that (i) $\forall K\in\Theta(s)\colon f_s(K)\subseteq\mathsf{Triang}(K)$, (ii) $\bigcup_{K\in\Theta(s)} f_s(K)=\bigcup_{K\in\Theta(s)}\mathsf{Triang}(K)$, and (iii) $\forall K_1,K_2\in\Theta(s)\colon f_s(K_1)\cap f_s(K_2)\neq\emptyset\Rightarrow K_1=K_2$.

Let $\pi'_\square$ and $\pi'_\diamond$ be a pair of strategies for $\mathcal{G}_{\mathcal{K}'}$. Define π_i, $i\in\{\square,\diamond\}$, for all $\hat{\rho}\in\mathcal{S}^*$, $s\in\mathcal{S}_i$, and $A\in\boldsymbol{\Sigma}_{\mathcal{A}}$ by

$$\pi_i(\hat{\rho}s)(A)=\sum_{K\in\Theta(s)}\sum_{K'\in f_s(K)}\pi'_i(\hat{\rho}s)(\{K'\}\times A|_K) \tag{21}$$

Notice that, by definition, $K'\in\Theta'(s)$. Moreover, notice that f_s ensures that a simplex in a triangulation of a polytope outgoing s is consdered in exactly one summand of (21). Thus $\pi_i(\hat{\rho})$ is a well defined probability measure on $\mathcal{S}$ and hence $\pi_\square$ and $\pi_\diamond$ is a pair of strategies for $\mathcal{G}_{\mathcal{K}}$.

It is straightforward to check that if π'_i is memoryless or semi-Markov, so is π_i. Suppose π'_i is extreme, then we have that

$$\begin{aligned}&\pi_i(\hat{\rho}s)(\{(K'',\mu)\in\mathcal{A}(s)\mid\mu\in\mathbb{V}(K'')\})\\ &\quad=\sum_{K\in\Theta(s)}\sum_{K'\in f_s(K)}\pi'_i(\hat{\rho}s)(\{K'\}\times\{(K,\mu)\in\mathcal{A}(s)\mid\mu\in\mathbb{V}(K)\}|_K)\end{aligned} \tag{22}$$

$$
\begin{aligned}
&= \sum_{K \in \Theta(s)} \sum_{K' \in f_s(K)} \pi'_i(\hat{\rho}s)(\{K'\} \times \mathbb{V}(K')) && (23)\\
&= \pi'_i(\hat{\rho}s)\Big(\bigcup_{K \in \Theta(s)} \bigcup_{K' \in f_s(K)} \{K'\} \times \mathbb{V}(K') \Big) && (24)\\
&= \pi'_i(\hat{\rho}s)\Big(\bigcup_{K' \in \bigcup_{K \in \Theta(s)} f_s(K)} \{K'\} \times \mathbb{V}(K') \Big) && (25)\\
&= \pi'_i(\hat{\rho}s)\Big(\bigcup_{K' \in \Theta'(s)} \{K'\} \times \mathbb{V}(K') \Big) && (26)\\
&= \pi'_i(\hat{\rho}s)(\{(K', \mu) \in \mathcal{A}'(s) \mid \mu \in \mathbb{V}(K')\}) && (27)\\
&= 1 && (28)
\end{aligned}
$$

Equality (22) corresponds to the definition of π_i in (21) and (23) follows from the following easy-to-check equalities: $\{(K, \mu) \in \mathcal{A}(s) \mid \mu \in \mathbb{V}(K)\}|_K = \mathbb{V}(K)$ and $\pi'_i(\hat{\rho}s)(\{K'\} \times (\mathbb{V}(K) \setminus K')) = 0$. (24) follows because $\pi'_i(\hat{\rho}s)$ is a probability measure and f_s guarantees the disjointness of sets in the union while (25) follows by calculations. (26) is a consequence of $\bigcup_{K \in \Theta(s)} f_s(K) = \bigcup_{K \in \Theta(s)} \mathsf{Triang}(K) = \Theta'(s)$ where the first equality is guaranteed by f_s and the second one is the definition of Θ'. Finally, (27) follows by the definition of $\mathcal{A}'(s)$ and (28) because π'_i is extreme.

Suppose now that π'_i is deterministic and assume $\pi'_i(\hat{\rho}s)(\{(K'_\star, \mu)\}) = 1$ for $K'_\star \in \Theta'(s)$ and $\mu \in K'_\star$. Besides, suppose that $K'_\star \in f_s(K_\star)$. Then

$$
\begin{aligned}
\pi_i(\hat{\rho}s)(\{(K_\star, \mu)\}) &= \sum_{K \in \Theta(s)} \sum_{K' \in f_s(K)} \pi'_i(\hat{\rho}s)(\{K'\} \times (\{(K_\star, \mu)\}|_K \cap K')) && (29)\\
&= \sum_{K' \in f_s(K_\star)} \pi'_i(\hat{\rho}s)(\{K'\} \times (\{(K_\star, \mu)\}|_{K_\star} \cap K')) && (30)\\
&= \sum_{K' \in f_s(K_\star)} \pi'_i(\hat{\rho}s)(\{K'\} \times (\{\mu\} \cap K')) && (31)\\
&= \pi'_i(\hat{\rho}s)(\{(K'_\star, \mu)\}) = 1 && (32)
\end{aligned}
$$

(29) follows by (21) while (30) is a consequence that $\{(K_\star, \mu)\}|_K = \emptyset$ whenever $K \neq K_\star$. (31) follows by definition of $|_{K_\star}$ and (32) follows from the fact that, for $K' \neq K'_\star$, either $\mu \notin K'$ or $\pi'_i(\hat{\rho}s)(\{(K', \mu)\}) = 0$. The last equality on (32) follows by the assumptions.

Like before, to prove that $\mathbb{P}^{\pi_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s} = \mathbb{P}^{\pi'_\Box, \pi'_\Diamond}_{\mathcal{G}_{\mathcal{K}'}, s}$ it sufficies to state that $\mathbb{P}^{\pi_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s} = \mathbb{P}^{\pi'_\Box, \pi'_\Diamond, n}_{\mathcal{G}_{\mathcal{K}'}, s}$ for all $n \geq 0$ which we show by induction. For $n = 0$, $\mathbb{P}^{\pi_\Box, \pi_\Diamond, 0}_{\mathcal{G}_\mathcal{K}, s}(s') = \delta_s(s') = \mathbb{P}^{\pi'_\Box, \pi'_\Diamond, 0}_{\mathcal{G}_{\mathcal{K}'}, s}(s')$. For $n + 1 > 0$ we calculate as follows. Suppose $\hat{\rho} \in \mathcal{S}^n$, $s'' \in \mathcal{S}$ and $s' \in \mathcal{S}_i$. Then,

$$\mathbb{P}^{\pi_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho}s''s')$$

$$= \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s'') \int_{\mathcal{A}} \theta(s'',\cdot,s')\ \mathbf{d}(\pi_i(\hat{\rho}s'')(\cdot)) \tag{33}$$

$$= \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s'') \sum_{K\in\Theta(s'')} \int_{K} \theta(s'',(K,\cdot),s')\ \mathbf{d}(\pi_i(\hat{\rho}s'')(K,\cdot)) \tag{34}$$

$$= \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s'') \sum_{K\in\Theta(s'')} \int_{K} \theta(s'',(K,\cdot),s')\ \mathbf{d}\Bigg(\sum_{K'\in f_s(K)} \pi'_i(\hat{\rho}s'')(K',\cdot)\Bigg) \tag{35}$$

$$= \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s'') \sum_{K\in\Theta(s'')} \sum_{K'\in f_s(K)} \int_{K'} \theta(s'',(K,\cdot),s')\ \mathbf{d}\Bigg(\pi'_i(\hat{\rho}s'')(K',\cdot)\Bigg) \tag{36}$$

$$= \mathbb{P}^{\pi'_\Box,\pi'_\Diamond,n}_{\mathcal{G}_{\mathcal{K}'},s}(\hat{\rho}s'') \sum_{K\in\Theta(s'')} \sum_{K'\in f_s(K)} \int_{K'} \theta'(s'',(K',\cdot),s')\ \mathbf{d}\Bigg(\pi'_i(\hat{\rho}s'')(K',\cdot)\Bigg) \tag{37}$$

$$= \mathbb{P}^{\pi'_\Box,\pi'_\Diamond,n}_{\mathcal{G}_{\mathcal{K}'},s}(\hat{\rho}s'') \sum_{K'\in\Theta'(s'')} \int_{K'} \theta'(s'',(K',\cdot),s')\ \mathbf{d}\Bigg(\pi'_i(\hat{\rho}s'')(K',\cdot)\Bigg) \tag{38}$$

$$= \mathbb{P}^{\pi'_\Box,\pi'_\Diamond,n}_{\mathcal{G}_{\mathcal{K}'},s}(\hat{\rho}s'') \int_{\left(\bigcup_{K'\in\Theta'(s'')}\{K'\}\times K'\right)} \theta'(s'',\cdot,s')\ \mathbf{d}\Bigg(\pi'_i(\hat{\rho}s'')(\cdot)\Bigg) \tag{39}$$

$$= \mathbb{P}^{\pi'_\Box,\pi'_\Diamond,n}_{\mathcal{G}_{\mathcal{K}'},s}(\hat{\rho}s'') \int_{\mathcal{A}'} \theta'(s'',\cdot,s')\ \mathbf{d}\Bigg(\pi'_i(\hat{\rho}s'')(\cdot)\Bigg) \tag{40}$$

$$= \mathbb{P}^{\pi'_\Box,\pi'_\Diamond,n+1}_{\mathcal{G}_{\mathcal{K}'},s}(\hat{\rho}s''s') \tag{41}$$

Equality (33) applies the definition of $\mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}$. (34) follows from the fact that $\pi_i(\hat{\rho}s'')\left(\mathcal{A}\setminus\left(\bigcup_{K\in\Theta(s'')}\{K\}\times K\right)\right) = 0$ and from Fubini's theorem. (35) follows because, by (21), $\pi_i(\hat{\rho}s'')(\{K\}\times B) = \sum_{K'\in f_s(K)} \pi'_i(\hat{\rho}s'')(\{K'\}\times B)$ since $(\{K''\}\times B)|_K = \emptyset$ for any $K'' \neq K$. (36) follows from calculations taking into account that $\pi'_i(\hat{\rho}s'')(\{K'\}\times(K\setminus K')) = 0$. (37) follows by induction hypothesis and from the fact that $\theta'(s'',(K',\mu),s') = \theta(s'',(K,\mu),s')$ for all $\mu\in K'$ whenever, $K'\in\mathsf{Triang}(K)$. (38) follows from the fact that $\bigcup_{K\in\Theta(s)} f_s(K) = \Theta'(s)$ and every $f_s(K)$ is disjoint from any other $f_s(K'')$. Fubini's theorem yields (39) and the fact that $\pi'_i(\hat{\rho}s'')\left(\mathcal{A}'\setminus\left(\bigcup_{K'\in\Theta'(s'')}\{K'\}\times K'\right)\right) = 0$ yields (40). Finally, by definition of $\mathbb{P}^{\pi'_\Box,\pi'_\Diamond,n+1}_{\mathcal{G}_{\mathcal{K}'},s}$ we conclude in (41). ◻

Proof (of Lemma 1). First of all, notice that $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond^k s') = \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\mathcal{S}^k\times\{s'\}\times\mathcal{S}^\omega) = \mathbb{P}^{\pi_\Box,\pi_\Diamond,k}_{\mathcal{G}_\mathcal{K},s}(\mathcal{S}^k\times\{s'\})$. This fact will be used in the following without making explicit the justification.

For item 1, we start by proving that for all $n\geq 0$ and $\alpha\in\mathbb{R}$,

$$\sum_{\hat{\rho}\in\mathcal{S}^{n+1}} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho})\ \sum_{i=0}^{n}\alpha^i\ r(\hat{\rho}_i) = \sum_{i=0}^{n}\sum_{s'\in\mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond^i s')\ \alpha^i\ r(s'). \tag{42}$$

We proceed by induction on n. For $n = 0$ we calculate:

$$\begin{aligned}\sum_{\hat\rho \in \mathcal{S}^{0+1}} \mathbb{P}^{\pi_\Box,\pi_\Diamond,0}_{\mathcal{G}_\mathcal{K},s}(\hat\rho) \sum_{i=0}^{0} \alpha^i \; r(\hat\rho_i) &= \sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond,0}_{\mathcal{G}_\mathcal{K},s}(s') \; \alpha^0 \; r(s') \\ &= \sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond^0 s') \; \alpha^0 \; r(s') \\ &= \sum_{i=0}^{0} \sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond^i s') \; \alpha^i \; r(s')\end{aligned}$$

All steps follow by there respective definitions.

For $n + 1$ ($n \geq 0$) we proceed as follows:

$$\begin{aligned}&\sum_{\hat\rho \in \mathcal{S}^{n+2}} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho) \sum_{i=0}^{n+1} \alpha^i \; r(\hat\rho_i) \\ &= \sum_{\hat\rho \in \mathcal{S}^{n+1}} \sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s') \; \left(\left(\sum_{i=0}^{n} \alpha^i \; r(\hat\rho_i) \right) + \alpha^{n+1} \; r(s') \right) \qquad (43) \\ &= \sum_{\hat\rho \in \mathcal{S}^{n+1}} \left(\sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s') \right) \sum_{i=0}^{n} \alpha^i \; r(\hat\rho_i) \\ &\qquad + \sum_{s' \in \mathcal{S}} \left(\sum_{\hat\rho \in \mathcal{S}^{n+1}} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s') \right) \; \alpha^{n+1} \; r(s') \\ &= \sum_{\hat\rho \in \mathcal{S}^{n+1}} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho) \sum_{i=0}^{n} \alpha^i \; r(\hat\rho_i) \\ &\qquad + \sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\mathcal{S}^{n+1} \times \{s'\}) \; \alpha^{n+1} \; r(s') \qquad (44) \\ &= \sum_{i=0}^{n} \sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond^i s') \; \alpha^i \; r(s') \\ &\qquad + \sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond^{n+1} s') \; \alpha^{n+1} \; r(s') \qquad (45) \\ &= \sum_{i=0}^{n+1} \sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond^i s') \; \alpha^i \; r(s')\end{aligned}$$

Most of the steps follow by simple calculations. In particular, in (43) we separate the trailing state. In (44), we use the fact that $\sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s') = \mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho \times \mathcal{S}) = \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho)$ in the first summand. Finally, induction hypothesis on the first summand of (45) is applied.

Item 1 of the lemma follows for $\widehat{\mathsf{rew}}^n_{\mathsf{t}}$ and $\widehat{\mathsf{rew}}^n_{\gamma}$ by taking $\alpha = 1$ and $\alpha = \gamma$ in (42), respectively. The case of $\widehat{\mathsf{rew}}^n_{\mathsf{a}}$ follows by observing that $\widehat{\mathsf{rew}}^n_{\mathsf{a}}(\hat\rho) = \frac{\widehat{\mathsf{rew}}^n_{\mathsf{t}}(\hat\rho)}{n+1}$

and calculating as follows $\sum_{\hat\rho\in\mathcal{S}^{n+1}} \mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho)\ \widehat{\mathsf{rew}}^n_{\mathsf{a}}(\hat\rho) = \sum_{\hat\rho\in\mathcal{S}^{n+1}} \mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho)\ \frac{\widehat{\mathsf{rew}}^n_{\mathsf{t}}(\hat\rho)}{n+1} = \sum_{i=0}^{n}\sum_{s'\in\mathcal{S}} \mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond^i s')\ \frac{1}{n+1}\ r(s')$.

Item 2 follows from item 1 as follows:

$$\mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew}_{\mathsf{f}}] = \lim_{n\to\infty} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}[\widehat{\mathsf{rew}}^n_{\mathsf{f}}] \tag{46}$$

$$= \lim_{n\to\infty} \sum_{\hat\rho\in\mathcal{S}^{n+1}} \mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho)\ \widehat{\mathsf{rew}}_{\mathsf{f}}(\hat\rho) \tag{47}$$

$$= \lim_{n\to\infty} \sum_{i=0}^{n}\sum_{s'\in\mathcal{S}} \mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond^i s')\ \mathsf{f}(i,n)\ r(s') \tag{48}$$

Equality (46) folows by convergence properties of random variables, (47) is the definition of $\mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}[\widehat{\mathsf{rew}}^n_{\mathsf{f}}]$, and (47) is item 1. $\square$

Proof (of Lemma 2). Define $\pi^*_\square$ as follows. For $\hat\rho\in\mathcal{S}^*$, $s'\in\mathcal{S}$, and $A\in\boldsymbol{\Sigma}_\mathcal{A}$, such that $\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(D\,\mathsf{U}^n\,s')>0$ and $|\hat\rho| = n \geq 0$, let

$$\pi^*_\square(\hat\rho s')(A) = \frac{\displaystyle\sum_{\hat\rho'\in D^n} \mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho' s')\ \pi_\square(\hat\rho' s')(A)}{\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(D\,\mathsf{U}^n\,s')}$$

For $s'\in\mathcal{S}$ with $\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(D\,\mathsf{U}^n\,s')=0$ and $|\hat\rho s'| = n$, define $\pi^*_\square(\hat\rho s')$ to be $\delta_{\mathsf{f}(s')}$ for a globally fixed function f such that $\mathsf{f}(s')\in\mathcal{A}(s')$. Notice that $\pi^*_\square\in\Pi^S_\square$. Therefore, we write $\pi^*_\square(n,s')$ for $\pi^*_\square(\hat\rho s')$ whenever $|\hat\rho| = n$.

We focus first on item 1 and proceed by induction. For $n = 0$,

$$\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(D\,\mathsf{U}^0\,s') = \mathbb{P}^{\pi_\square,\pi_\diamond,0}_{\mathcal{G}_\mathcal{K},s}(s') = \delta_s(s') = \mathbb{P}^{\pi^*_\square,\pi_\diamond,0}_{\mathcal{G}_\mathcal{K},s}(s') = \mathbb{P}^{\pi^*_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(D\,\mathsf{U}^0\,s').$$

For $n+1\geq 0$, note that $\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(D\,\mathsf{U}^{n+1}\,s') = \sum_{\hat\rho\in D^{n+1}} \mathbb{P}^{\pi_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s') = \sum_{s''\in D}\sum_{\hat\rho\in D^n} \mathbb{P}^{\pi_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'' s')$. Hence, it suffices to show that, for all $s''\in\mathcal{S}$,

$$\sum_{\hat\rho\in D^n} \mathbb{P}^{\pi_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'' s') = \sum_{\hat\rho\in D^n} \mathbb{P}^{\pi^*_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'' s'),$$

for which we differentiate two cases, depending on the player. So, if $s''\in\mathcal{S}_\diamond$ we proceed as follows

$$\sum_{\hat\rho\in D^n} \mathbb{P}^{\pi_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'' s')$$

$$= \sum_{\hat\rho\in D^n} \mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'') \int_\mathcal{A} \theta(s'',\cdot,s')\ \mathbf{d}(\pi_\diamond(\hat\rho s'')(\cdot)) \tag{49}$$

$$= \left(\sum_{\hat\rho\in D^n} \mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')\right) \int_\mathcal{A} \theta(s'',\cdot,s')\ \mathbf{d}(\pi_\diamond(n,s'')(\cdot)) \tag{50}$$

$$= \left(\sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') \right) \int_\mathcal{A} \theta(s'', \cdot, s') \; \mathbf{d}(\pi_\Diamond(n, s'')(\cdot)) \tag{51}$$

$$= \sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'' s') \tag{52}$$

Step (49) follows by definition of $\mathbb{P}^{\pi_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}$. Because $\pi_\Diamond$ is semi-Markov the integral can be factored out of the summation in (50). Induction hypothesis is applied in (51), since $\mathbb{P}^{\pi_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \mathsf{U}^n\, s'') = \left(\sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') \right)$ and similarly for $\mathbb{P}^{\pi^*_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \mathsf{U}^n\, s'')$. Finally, step (52) resolves using the same steps as before in reverse order.

If $s'' \in \mathcal{S}_\Box$ we have two subcases. First suppose $\mathbb{P}^{\pi_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \mathsf{U}^n\, s'') > 0$. Then

$$\sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'' s')$$

$$= \sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') \int_\mathcal{A} \theta(s'', \cdot, s') \; \mathbf{d}(\pi_\Box(\hat{\rho} s'')(\cdot)) \tag{53}$$

$$= \mathbb{P}^{\pi_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \mathsf{U}^n\, s'') \; \int_\mathcal{A} \theta(s'', \cdot, s') \; \frac{\sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') \; \mathbf{d}(\pi_\Box(\hat{\rho} s'')(\cdot))}{\mathbb{P}^{\pi_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \mathsf{U}^n\, s'')} \tag{54}$$

$$= \mathbb{P}^{\pi^*_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \mathsf{U}^n\, s'') \; \int_\mathcal{A} \theta(s'', \cdot, s') \; \mathbf{d}(\pi^*_\Box(n, s'')(\cdot)) \tag{55}$$

$$= \left(\sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') \right) \int_\mathcal{A} \theta(s'', \cdot, s') \; \mathbf{d}(\pi^*_\Box(n, s'')(\cdot)) \tag{56}$$

$$= \sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') \; \int_\mathcal{A} \theta(s'', \cdot, s') \; \mathbf{d}(\pi^*_\Box(\hat{\rho} s'')(\cdot)) \tag{57}$$

$$= \sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'' s') \tag{58}$$

Step (53) follows by definition of $\mathbb{P}^{\pi_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}$ and (54) is obtained by multiplying and dividing by $\mathbb{P}^{\pi_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \mathsf{U}^n\, s'')$. Step (55) follows by induction and using the definition of $\pi^*_\Box$. By noting that $\mathbb{P}^{\pi^*_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \mathsf{U}^n\, s'') = \sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'')$ we obtain (56). Calculations and recalling that $\pi^*_\Box(n, s'') = \pi^*_\Box(\hat{\rho} s'')$, whenever $|\hat{\rho}| = n$, yields (57) which, by definition of $\mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}$, concludes in (58).

For the case $\mathbb{P}^{\pi_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \mathsf{U}^n\, s'') = 0$, we first prove the following claim.

Claim. $\mathbb{P}^{\pi_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') = 0$ implies $\mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') = 0$, for all $n \geq 0$, $\hat{\rho} \in D^n$, and $s'' \in \mathcal{S}$.

Proof of claim. We proceed by induction on n. For $n = 0$ the claim follows after noting that $\mathbb{P}^{\pi_\Box, \pi_\Diamond, 0}_{\mathcal{G}_\mathcal{K}, s}(s') = \delta_s(s') = \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, 0}_{\mathcal{G}_\mathcal{K}, s}(s')$ by the definition of $\mathbb{P}$.

So, let $n+1>0$ and suppose $\mathbb{P}^{\pi_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'' s')=0$. Suppose $s''\in\mathcal{S}_\square$. Since, by definition, $\mathbb{P}^{\pi_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'' s')=\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')\int_\mathcal{A}\theta(s'',\cdot,s')\ \mathbf{d}(\pi_\square(\hat\rho s'')(\cdot))$, then either $\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')=0$ or $\int_\mathcal{A}\theta(s'',\cdot,s')\ \mathbf{d}(\pi_\square(\hat\rho s'')(\cdot))=0$.

If $\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')=0$, then $\mathbb{P}^{\pi^*_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')=0$ by induction hypothesis and hence $\mathbb{P}^{\pi^*_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'' s')=\mathbb{P}^{\pi^*_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')\int_\mathcal{A}\theta(s'',\cdot,s')\ \mathbf{d}(\pi_\square(\hat\rho s'')(\cdot))=0$.

If, instead $\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')>0$, then $\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(D\ \mathsf{U}^n\ s'')>0$ and hence,

$$
\begin{aligned}
&\mathbb{P}^{\pi^*_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'' s')\\
&\quad=\mathbb{P}^{\pi^*_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')\int_\mathcal{A}\theta(s'',\cdot,s')\ \mathbf{d}(\pi^*_\square(\hat\rho s'')(\cdot)) && (59)\\
&\quad=\mathbb{P}^{\pi^*_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')\int_\mathcal{A}\theta(s'',\cdot,s')\ \frac{\sum_{\hat\rho\in D^n}\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')\ \mathbf{d}(\pi_\square(\hat\rho s'')(\cdot))}{\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(D\ \mathsf{U}^n\ s'')} && (60)\\
&\quad=\mathbb{P}^{\pi^*_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')\ \frac{\sum_{\hat\rho\in D^n}\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')}{\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(D\ \mathsf{U}^n\ s'')}\int_\mathcal{A}\theta(s'',\cdot,s')\ \mathbf{d}(\pi_\square(\hat\rho s'')(\cdot)) && (61)\\
&\quad=0 && (62)
\end{aligned}
$$

Step (59) follows by definition of $\mathbb{P}^{\pi^*_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}$, step (60) by definition of $\pi^*_\diamond$ and step (61) by standard calculations. Since necessarily $\int_\mathcal{A}\theta(s'',\cdot,s')\ \mathbf{d}(\pi_\diamond(\hat\rho s'')(\cdot))=0$, the proof concludes with step (62).

For $s''\in\mathcal{S}_\diamond$, it follows as before only differing in the case of $\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')>0$, in which case $\int_\mathcal{A}\theta(s'',\cdot,s')\ \mathbf{d}(\pi_\diamond(\hat\rho s'')(\cdot))=0$. Thus, $\mathbb{P}^{\pi^*_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'' s')=\mathbb{P}^{\pi^*_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')\int_\mathcal{A}\theta(s'',\cdot,s')\ \mathbf{d}(\pi_\diamond(\hat\rho s'')(\cdot))=0$. *(End of claim)* □

Now, notice that $\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(D\ \mathsf{U}^n\ s'')=0$ implies that $\sum_{\hat\rho\in D^n}\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')=0$, and therefore $\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')=0$ for all $\hat\rho\in D^n$. Recall that $s''\in\mathcal{S}_\square$. Then,

$$
\begin{aligned}
\sum_{\hat\rho\in D^n}\mathbb{P}^{\pi_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'' s') &= \sum_{\hat\rho\in D^n}\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')\int_\mathcal{A}\theta(s'',\cdot,s')\ \mathbf{d}(\pi_\square(\hat\rho s'')(\cdot)) && (63)\\
&=0 && (64)\\
&= \sum_{\hat\rho\in D^n}\mathbb{P}^{\pi^*_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')\int_\mathcal{A}\theta(s'',\cdot,s')\ \mathbf{d}(\pi^*_\square(\hat\rho s'')(\cdot)) && (65)\\
&= \sum_{\hat\rho\in D^n}\mathbb{P}^{\pi^*_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'' s') && (66)
\end{aligned}
$$

In the above calculations, (63) and (66) follow by definition of $\mathbb{P}^{\pi_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}$ and $\mathbb{P}^{\pi^*_\square,\pi_\diamond,n+1}_{\mathcal{G}_\mathcal{K},s}$, respectively, step (64) follow from the fact that $\mathbb{P}^{\pi_\square,\pi_\diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s'')=0$ for all $\hat\rho\in D^n$ and because of this, the claim yields (65). This concludes the proof of item 1.

For 2, we have that

$$
\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C)=\sum_{s'\in C}\sum_{n\geq 0}\mathbb{P}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}((\mathcal{S}\setminus C)\ \mathsf{U}^n\ s')
$$

$$= \sum_{s' \in C} \sum_{n \geq 0} \mathbb{P}^{\pi^*_\square, \pi_\diamond}_{\mathcal{G}_\mathcal{K}, s}((\mathcal{S} \setminus C) \,\mathsf{U}^n\, s') = \mathbb{P}^{\pi^*_\square, \pi_\diamond}_{\mathcal{G}_\mathcal{K}, s}(\diamond C)$$

where the middle equality follow from item 1, and the other two because, for all $n \neq m$ and $s' \in C$, $((\mathcal{S} \setminus C) \,\mathsf{U}^n\, s') \cap ((\mathcal{S} \setminus C) \,\mathsf{U}^m\, s') = \emptyset$.

Item 3 can be calculated as follows.

$$\mathbb{E}^{\pi_\square, \pi_\diamond}_{\mathcal{G}_\mathcal{K}, s}[\mathsf{rew_f}] = \lim_{n \to \infty} \sum_{i=0}^{n} \sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi_\square, \pi_\diamond}_{\mathcal{G}_\mathcal{K}, s}(\diamond^i s') \; \mathsf{f}(i, n) \; r(s') \tag{67}$$

$$= \lim_{n \to \infty} \sum_{i=0}^{n} \sum_{s' \in \mathcal{S}} \mathbb{P}^{\pi^*_\square, \pi_\diamond}_{\mathcal{G}_\mathcal{K}, s}(\diamond^i s') \; \mathsf{f}(i, n) \; r(s') \tag{68}$$

$$= \mathbb{E}^{\pi^*_\square, \pi_\diamond}_{\mathcal{G}_\mathcal{K}, s}[\mathsf{rew_f}] \tag{69}$$

Steps (67) and (69) follow from Lemma 1.2, while (68) follows by item 2 of this lemma.

Notice that the proof can be replicated mutatis mutandi with $\square$ and $\diamond$ exchanged. Therefore the last part of the lemma also holds. $\square$

Proof (of Lemma 3). For any $K \in \mathsf{DSimp}(\mathcal{S})$, $\mu \in K$ and $\hat{\mu} \in \mathbb{V}(K)$ define $\mathsf{p}^K(\mu, \hat{\mu}) \in [0, 1]$ such that $\sum_{\hat{\mu} \in \mathbb{V}(K)} \mathsf{p}^K(\mu, \hat{\mu}) \; \hat{\mu} = \mu$. That is, all $\mathsf{p}^K(\mu, \hat{\mu})$, $\hat{\mu} \in \mathbb{V}(K)$, are the unique factors that define the convex combination for μ in the simplex K. Therefore, $\mathsf{p}^K(\mu, \hat{\mu})$ is well defined for all $K \in \mathsf{DSimp}(\mathcal{S})$, $\mu \in K$ and $\hat{\mu} \in \mathbb{V}(K)$. In any other case, let $\mathsf{p}^K(\mu, \hat{\mu}) = 0$.

Let $\mathsf{p}((K, \mu), (K, \hat{\mu})) = \mathsf{p}^K(\mu, \hat{\mu})$ for all $K \in \mathsf{DSimp}(\mathcal{S})$, $\mu \in K$ and $\hat{\mu} \in \mathbb{V}(K)$, and let $\mathsf{p}(a, b) = 0$ for any other $a, b \in \mathcal{A}$. For every $(K, \mu) \in \mathcal{A}$ such that $\mu \in K$, let $\mathbb{V}(K, \mu) = \{(K, \hat{\mu}) \mid \hat{\mu} \in \mathbb{V}(K)\}$ and let $\mathbb{V}(K, \mu) = \emptyset$ otherwise. Thus, for every $s \in \mathcal{S}$ and $a \in \mathcal{A}$,

$$\theta(s, a, \cdot) = \sum_{b \in \mathbb{V}(a)} \mathsf{p}(a, b) \; \theta(s, b, \cdot). \tag{70}$$

We also extend p to measurable sets $B \in \boldsymbol{\Sigma}_\mathcal{A}$ and $a \in \mathcal{A}$ by $\mathsf{p}(a, B) = \sum_{b \in B \cap \mathbb{V}(a)} \mathsf{p}(a, b)$. We observe that $\mathsf{p}(\cdot, B)$ is measurable for any $B \in \boldsymbol{\Sigma}_\mathcal{A}$, which is a consequence of the following calculation, where $p \in [0, 1]$ and $\mathrm{pr}_2(B) = \{\mu \mid (K, \mu) \in B\}$,

$$\{a \mid \mathsf{p}(a, B) \geq p\} = \left\{(K, \mu) \mid \textstyle\sum_{\hat{\mu} \in \mathrm{pr}_2(B) \cap \mathbb{V}(K)} \mathsf{p}^K(\mu, \hat{\mu}) \geq p\right\}$$
$$= \textstyle\bigcup_{K \in \bigcup_{s \in \mathcal{S}} \Theta(s)} \{K\} \times \left\{\mu \mid \sum_{\hat{\mu} \in \mathrm{pr}_2(B) \cap \mathbb{V}(K)} \mathsf{p}^K(\mu, \hat{\mu}) \geq p\right\}$$

Notice that, if $p = 0$ then $\{a \mid \mathsf{p}(a, B) \geq p\} = \mathcal{A}$, which is measurable. If $p > 0$, notice that $\bigcup_{s \in \mathcal{S}} \Theta(s)$ is finite and $\left\{\mu \mid \sum_{\hat{\mu} \in \mathrm{pr}_2(B) \cap \mathbb{V}(K)} \mathsf{p}^K(\mu, \hat{\mu}) \geq p\right\} = \left\{\mu \in K \mid \sum_{\hat{\mu} \in \mathrm{pr}_2(B) \cap \mathbb{V}(K)} \mathsf{p}^K(\mu, \hat{\mu}) \geq p\right\}$ is a convex polytope, hence measurable. Therefore, $\{a \mid \mathsf{p}(a, B) \geq p\}$ is measurable.

For every $\hat{\rho} \in \mathcal{S}^*$, $s' \in \mathcal{S}$ and $B \in \boldsymbol{\Sigma}_{\mathcal{A}}$, define $\pi^*_\square$ by

$$\pi^*_\square(\hat{\rho}s')(B) = \int_{\mathcal{A}} \mathsf{p}(\cdot, B)\ \mathbf{d}(\pi_\square(\hat{\rho}s')(\cdot)) = \int_{\mathcal{A}(s')} \mathsf{p}(\cdot, B)\ \mathbf{d}(\pi_\square(\hat{\rho}s')(\cdot)).$$

The first equality is the definition and the second equality follows from the fact that $\pi_\square(\hat{\rho}s')(\mathcal{A} \setminus \mathcal{A}(s')) = 0$. $\pi^*_\square(\hat{\rho}s')$ is defined so that it assigns to each vertex of a simplex the weighted contribution (according to $\pi_\square(\hat{\rho}s')$) of each distribution (in the said simplex) to such vertex.

Since $\mathsf{p}(\cdot, B)$ is measurable, $\pi^*_\square$ is well defined. Moreover, because $\pi_\square$ is semi-Markov, so is $\pi^*_\square$. The following calculation shows that $\pi^*_\square$ is also extreme.

$$\begin{aligned}
&\pi^*_\square(\hat{\rho}s')(\mathbb{V}(\mathcal{A}(s'))) \\
&\quad = \int_{\mathcal{A}(s')} \mathsf{p}(\cdot, \mathbb{V}(\mathcal{A}(s')))\ \mathbf{d}(\pi_\square(\hat{\rho}s')(\cdot)) && (71)\\
&\quad = \int_{\mathcal{A}(s')} \mathsf{p}(\cdot, \{(K', \hat{\mu}) \mid K' \in \Theta(s'), \hat{\mu} \in \mathbb{V}(K')\})\ \mathbf{d}(\pi_\square(\hat{\rho}s')(\cdot)) && (72)\\
&\quad = \int_{\mathcal{A}(s')} \lambda(K, \mu). \left(\sum_{K' \in \Theta(s'), \hat{\mu} \in \mathbb{V}(K')} \mathsf{p}((K, \mu), (K', \hat{\mu})) \right)\ \mathbf{d}(\pi_\square(\hat{\rho}s')(\cdot)) && (73)\\
&\quad = \int_{\mathcal{A}(s')} \lambda(K, \mu). \left(\mathbf{1}_{\Theta(s')}(K) \sum_{\hat{\mu} \in \mathbb{V}(K)} \mathsf{p}^K(\mu, \hat{\mu}) \right)\ \mathbf{d}(\pi_\square(\hat{\rho}s')(\cdot)) && (74)\\
&\quad = \int_{\mathcal{A}(s')} \mathbf{d}(\pi_\square(\hat{\rho}s')(\cdot)) && (75)\\
&\quad = \pi_\square(\hat{\rho}s')(\mathcal{A}(s')) && (76)\\
&\quad = 1 && (77)
\end{aligned}$$

The definition of $\pi^*_\square$ yields the first step (71). Step (72) follows by observing that $\mathbb{V}(\mathcal{A}(s')) = \{(K', \hat{\mu}) \in \mathcal{A}(s) \mid \hat{\mu} \in \mathbb{V}(K')\} = \{(K', \hat{\mu}) \mid K' \in \Theta(s'), \hat{\mu} \in \mathbb{V}(K')\}$ Using the lambda notation and the definition of p on sets, we obtain (73). By noting that the sum is 0 if $K \notin \Theta(s')$, we introduce the characteristic function $\mathbf{1}_{\Theta(s')}$ in (74), where we also apply the deintinion of p. The fact that $\sum_{\hat{\mu} \in \mathbb{V}(K)} \mathsf{p}^K(\mu, \hat{\mu}) = 1$ and that $\mathbf{1}_{\Theta(s')}(K) = 1$ for all $(K, \mu) \in \mathcal{A}(s')$ (since $K \in \Theta(s')$) yields (75). Finally, step (76) follows by the definition of the integral and step (77) because $\pi_\square$ is a strategy.

We now proceed to prove item 1 by induction on n. For $n = 0$ we calculate:

$$\mathbb{P}^{\pi_\square, \pi_\diamond}_{\mathcal{G}_\mathcal{K}, s}(D\ \mathsf{U}^0\ s') = \mathbb{P}^{\pi_\square, \pi_\diamond, 0}_{\mathcal{G}_\mathcal{K}, s}(s') = \delta_s(s') = \mathbb{P}^{\pi^*_\square, \pi_\diamond, 0}_{\mathcal{G}_\mathcal{K}, s}(s') = \mathbb{P}^{\pi^*_\square, \pi_\diamond}_{\mathcal{G}_\mathcal{K}, s}(D\ \mathsf{U}^0\ s').$$

For $n + 1 > 0$, we have that

$$\mathbb{P}^{\pi_\square, \pi_\diamond}_{\mathcal{G}_\mathcal{K}, s}(D\ \mathsf{U}^{n+1}\ s') = \sum_{\hat{\rho} \in D^n, s'' \in D} \mathbb{P}^{\pi_\square, \pi_\diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho}s''s')$$

$$= \sum_{\hat{\rho} \in D^n, s'' \in (\mathcal{S}_\Box \cap D)} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'' s') + \sum_{\hat{\rho} \in D^n, s'' \in (\mathcal{S}_\Diamond \cap D)} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'' s')$$

and similarly for $\mathbb{P}^{\pi^*_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \,\mathsf{U}^{n+1}\, s')$. Therefore, it suffices to show that

$$\sum_{\hat{\rho} \in D^n, s'' \in \mathcal{S}_\Box} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'' s') = \sum_{\hat{\rho} \in D^n, s'' \in \mathcal{S}_\Box} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'' s') \quad \text{and} \tag{78}$$

$$\sum_{\hat{\rho} \in D^n, s'' \in \mathcal{S}_\Diamond} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'' s') = \sum_{\hat{\rho} \in D^n, s'' \in \mathcal{S}_\Diamond} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'' s'). \tag{79}$$

To prove (78), we calculate as follows:

$$\sum_{\hat{\rho} \in D^n, s'' \in \mathcal{S}_\Box} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'' s')$$

$$= \sum_{\hat{\rho} \in D^n, s'' \in \mathcal{S}_\Box} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') \int_\mathcal{A} \theta(s'', \cdot, s') \; \mathbf{d}(\pi^*_\Box(\hat{\rho} s'')(\cdot)) \tag{80}$$

$$= \sum_{s'' \in \mathcal{S}_\Box} \left(\sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') \right) \int_\mathcal{A} \theta(s'', \cdot, s') \; \mathbf{d}(\pi^*_\Box(n, s'')(\cdot)) \tag{81}$$

$$= \sum_{s'' \in \mathcal{S}_\Box} \mathbb{P}^{\pi^*_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \,\mathsf{U}^n\, s'') \sum_{a \in \mathbb{V}(\mathcal{A}(s''))} \theta(s'', a, s') \; \pi^*_\Box(n, s'')(\{a\}) \tag{82}$$

$$= \sum_{s'' \in \mathcal{S}_\Box} \mathbb{P}^{\pi_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \,\mathsf{U}^n\, s'') \sum_{a \in \mathbb{V}(\mathcal{A}(s''))} \theta(s'', a, s') \int_\mathcal{A} \mathsf{p}(\cdot, \{a\}) \; \mathbf{d}(\pi_\Box(n, s'')(\cdot)) \tag{83}$$

$$= \sum_{s'' \in \mathcal{S}_\Box} \left(\sum_{\hat{\rho} \in D^n} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') \right) \int_\mathcal{A} \lambda x. \left(\sum_{a \in \mathbb{V}(\mathcal{A}(s''))} \mathsf{p}(x, a) \; \theta(s'', a, s') \right) \mathbf{d}(\pi_\Box(n, s'')(\cdot)) \tag{84}$$

$$= \sum_{\hat{\rho} \in D^n, s'' \in \mathcal{S}_\Box} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') \int_\mathcal{A} \lambda x. \left(\sum_{a \in \mathbb{V}(x)} \mathsf{p}(x, a) \; \theta(s'', a, s') \right) \mathbf{d}(\pi_\Box(n, s'')(\cdot)) \tag{85}$$

$$= \sum_{\hat{\rho} \in D^n, s'' \in \mathcal{S}_\Box} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'') \int_\mathcal{A} \theta(s'', \cdot, s') \; \mathbf{d}(\pi_\Box(\hat{\rho} s'')(\cdot)) \tag{86}$$

$$= \sum_{\hat{\rho} \in D^n, s'' \in \mathcal{S}_\Box} \mathbb{P}^{\pi_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}(\hat{\rho} s'' s') \tag{87}$$

(80) follows by the definition of $\mathbb{P}^{\pi^*_\Box, \pi_\Diamond, n+1}_{\mathcal{G}_\mathcal{K}, s}$ while (81) follows from the fact the $\pi^*_\Box$ is semi-Markov. Step (82) follows from the definition of $\mathbb{P}^{\pi^*_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}(D \,\mathsf{U}^n\, s'')$ and the

fact $\mathbb{V}(\mathcal{A}(s))$, the support set of $\pi^*_\Box(n, s'')$, is finite. In (83), induction hypothesis is applied as well as the definition of $\pi^*_\Box(n, s'')(\{a\})$. (84) follows by the definition of $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(D \,\mathsf{U}^n\, s'')$, calculations, the introduction of the λ notation and the fact that $\mathsf{p}(x, \{a\}) = \mathsf{p}(x, a)$. In (85), we observe that $\mathsf{p}(x, a) > 0$ only if $a \in \mathbb{V}(x)$. (86) follows from (70) and the fact $\mathcal{A}(s)$ is the support set of $\pi_\Box(\hat{\rho}s'')$. Finally, the definition of $\mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}$ is applied in (87).

To prove (79), we calculate as follows:

$$
\begin{aligned}
&\sum_{\hat{\rho}\in D^n, s''\in\mathcal{S}_\Diamond} \mathbb{P}^{\pi^*_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s''s') \\
&\qquad = \sum_{\hat{\rho}\in D^n, s''\in\mathcal{S}_\Diamond} \mathbb{P}^{\pi^*_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s'') \int_\mathcal{A} \theta(s'', \cdot, s')\ \mathbf{d}(\pi_\Diamond(\hat{\rho}s'')(\cdot)) && (88)\\
&\qquad = \sum_{s''\in\mathcal{S}_\Diamond} \left(\sum_{\hat{\rho}\in D^n} \mathbb{P}^{\pi^*_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s'') \right) \int_\mathcal{A} \theta(s'', \cdot, s')\ \mathbf{d}(\pi_\Diamond(n, s'')(\cdot)) && (89)\\
&\qquad = \sum_{s''\in\mathcal{S}_\Diamond} \mathbb{P}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(D \,\mathsf{U}^n\, s'') \int_\mathcal{A} \theta(s'', \cdot, s')\ \mathbf{d}(\pi_\Diamond(n, s'')(\cdot)) && (90)\\
&\qquad = \sum_{s''\in\mathcal{S}_\Diamond} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(D \,\mathsf{U}^n\, s'') \int_\mathcal{A} \theta(s'', \cdot, s')\ \mathbf{d}(\pi_\Diamond(n, s'')(\cdot)) && (91)\\
&\qquad = \sum_{\hat{\rho}\in D^n, s''\in\mathcal{S}_\Diamond} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s''s') && (92)
\end{aligned}
$$

Step (88) follows by the definition of $\mathbb{P}^{\pi^*_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}$, (89) follows from the fact the $\pi_\Diamond$ is semi-Markov, and (90) follows from the definition of $\mathbb{P}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(D \,\mathsf{U}^n\, s'')$. Induction hypothesis is applied in (91). Finally (92) follows like the first three steps in the inverse order.

For 2, we have that

$$
\begin{aligned}
\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond C) &= \sum_{s'\in C} \sum_{n\geq 0} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}((\mathcal{S}\setminus C) \,\mathsf{U}^n\, s') \\
&= \sum_{s'\in C} \sum_{n\geq 0} \mathbb{P}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}((\mathcal{S}\setminus C) \,\mathsf{U}^n\, s') = \mathbb{P}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond C)
\end{aligned}
$$

where the middle equality follow from item 1, and the other two because, for all $n \neq m$ and $s' \in C$, $((\mathcal{S}\setminus C) \,\mathsf{U}^n\, s') \cap ((\mathcal{S}\setminus C) \,\mathsf{U}^m\, s') = \emptyset$.

Item 3 can be calculated as follows:

$$
\begin{aligned}
\mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew}_\mathsf{f}] &= \lim_{n\to\infty} \sum_{i=0}^{n} \sum_{s'\in\mathcal{S}} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond^i s')\ \mathsf{f}(i, n)\ r(s') && \text{(by Lemma 1.2)}\\
&= \lim_{n\to\infty} \sum_{i=0}^{n} \sum_{s'\in\mathcal{S}} \mathbb{P}^{\pi^*_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond^i s')\ \mathsf{f}(i, n)\ r(s') && \text{(by item 2)}
\end{aligned}
$$

$$= \mathbb{E}^{\pi^*_\Box, \pi_\Diamond}_{\mathcal{G}_\mathcal{K}, s}[\mathsf{rew_f}] \qquad \text{(by Lemma 1.2)}$$

Notice that the proof can be replicated mutatis mutandi with $\Box$ and $\Diamond$ exchanged. Therefore the last part of the lemma also holds. $\Box$

Proposition 5. *Let $\mathcal{G}_\mathcal{K}$ and $\mathcal{H}_\mathcal{K}$ be respectively the interpretation and the extreme interpretation of $\mathcal{K}$. Then*

1. *For every $\pi_\Box \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Box}$ and $\pi_\Diamond \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Diamond}$, (a) $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond C) = \mathbb{P}^{\pi^\vee_\Box,\pi^\vee_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond C)$, for all $C \subseteq \mathcal{S}$, and (b) $\mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew_f}] = \mathbb{E}^{\pi^\vee_\Box,\pi^\vee_\Diamond}_{\mathcal{H}_\mathcal{K},s}[\mathsf{rew_f}]$; and*
2. *For every $\pi_\Box \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}$ and $\pi_\Diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\Diamond}$, (a) $\mathbb{P}^{\pi^\times_\Box,\pi^\times_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond C) = \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond C)$, for all $C \subseteq \mathcal{S}$, and (b) $\mathbb{E}^{\pi^\times_\Box,\pi^\times_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew_f}] = \mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}[\mathsf{rew_f}]$.*

Proof (of Proposition 5). To prove item 1, we first prove by induction that for all $n \geq 0$ and $\hat{\rho} \in \mathcal{S}^{n+1}$,

$$\mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}) = \mathbb{P}^{\pi^\vee_\Box,\pi^\vee_\Diamond,n}_{\mathcal{H}_\mathcal{K},s}(\hat{\rho}) \tag{93}$$

The case $n = 0$ is direct. For $n+1 > 0$, $\hat{\rho} \in \mathcal{S}^n$, $s' \in \mathcal{S}$ and $s'' \in \mathcal{S}_\Box$ we calculate as follows

$$\mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s''s') = \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s'') \int_\mathcal{A} \theta(s'',\cdot,s')\ \mathbf{d}(\pi_\Box(\hat{\rho}s'')(\cdot)) \tag{94}$$

$$= \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s'') \sum_{a \in \mathbb{V}(\mathcal{A}(s''))} \theta(s'',a,s')\ \pi_\Box(\hat{\rho}s'')(\{a\}) \tag{95}$$

$$= \mathbb{P}^{\pi^\vee_\Box,\pi^\vee_\Diamond,n}_{\mathcal{H}_\mathcal{K},s}(\hat{\rho}s'') \sum_{a \in \mathcal{A}_{\mathcal{H}_\mathcal{K}}(s'')} \theta_{\mathcal{H}_\mathcal{K}}(s'',a,s')\ \pi^\vee_\Box(\hat{\rho}s'')(\{a\}) \tag{96}$$

$$= \mathbb{P}^{\pi^\vee_\Box,\pi^\vee_\Diamond,n+1}_{\mathcal{H}_\mathcal{K},s}(\hat{\rho}s''s') \tag{97}$$

Step (94) follows from definition of $\mathbb{P}^{\pi_\Box,\pi_\Diamond,n+1}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s''s')$. To obtain (95) we use the fact that the support set $\mathbb{V}(\mathcal{A}(s''))$ of $\pi_\Box(\hat{\rho}s'')$ is finite. In (96) induction hypothesis is applied, together with the definitions of $\theta_{\mathcal{H}_\mathcal{K}}(s'',a,s')$ and $\pi^\vee_\Box(\hat{\rho}s'')(\{a\})$. Finally, the definition of $\mathbb{P}^{\pi^\vee_\Box,\pi^\vee_\Diamond,n+1}_{\mathcal{H}_\mathcal{K},s}(\hat{\rho}s''s')$ yields (97).

The calculations follow similarly for the case of $s'' \in \mathcal{S}_\Diamond$, which proves (93).

Now item 1a follows from the following calculations:

$$\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond C) = \sum_{n \geq 0} \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}((\mathcal{S} \setminus C)\ \mathsf{U}^n\ C) \tag{98}$$

$$= \sum_{n \geq 0} \sum_{s' \in C} \sum_{\hat{\rho} \in (\mathcal{S} \setminus C)^n} \mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{G}_\mathcal{K},s}(\hat{\rho}s') \tag{99}$$

$$= \sum_{n \geq 0} \sum_{s' \in C} \sum_{\hat{\rho} \in (\mathcal{S} \setminus C)^n} \mathbb{P}^{\pi^\vee_\Box,\pi^\vee_\Diamond,n}_{\mathcal{H}_\mathcal{K},s}(\hat{\rho}s') \tag{100}$$

$$= \mathbb{P}^{\pi^\vee_\Box,\pi^\vee_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond C) \tag{101}$$

Equality (98) follows from the fact $((\mathcal{S}\setminus C)\,\mathsf{U}^n\, C)\cap((\mathcal{S}\setminus C)\,\mathsf{U}^m\, C)=\emptyset$ whenever $n\neq m$ and equality (99), from the definition of $\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}((\mathcal{S}\setminus C)\,\mathsf{U}^n\, C)$. The auxiliary result (93) yields (100) and (101) follows as the previous steps in inverse direction.

For item 1b, we calculate as follows

$$
\begin{aligned}
&\mathbb{E}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew}_\mathsf{f}]\\
&\quad=\lim_{n\to\infty}\sum_{i=0}^{n}\sum_{s'\in\mathcal{S}}\mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond^i s')\ \mathsf{f}(i,n)\ r(s') && \text{(by Lemma 1.2)}\\
&\quad=\lim_{n\to\infty}\sum_{i=0}^{n}\sum_{s'\in\mathcal{S}}\sum_{\hat\rho\in\mathcal{S}^i}\mathbb{P}^{\pi_\Box,\pi_\Diamond,i}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s')\ \mathsf{f}(i,n)\ r(s') && \text{(by def. of } \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond^i s'))\\
&\quad=\lim_{n\to\infty}\sum_{i=0}^{n}\sum_{s'\in\mathcal{S}}\sum_{\hat\rho\in\mathcal{S}^i}\mathbb{P}^{\pi_\Box^\vee,\pi_\Diamond^\vee,i}_{\mathcal{G}_\mathcal{K},s}(\hat\rho s')\ \mathsf{f}(i,n)\ r(s') && \text{(because of (93))}\\
&\quad=\lim_{n\to\infty}\sum_{i=0}^{n}\sum_{s'\in\mathcal{S}}\mathbb{P}^{\pi_\Box^\vee,\pi_\Diamond^\vee}_{\mathcal{G}_\mathcal{K},s}(\Diamond^i s')\ \mathsf{f}(i,n)\ r(s') && \text{(by def. of } \mathbb{P}^{\pi_\Box^\vee,\pi_\Diamond^\vee}_{\mathcal{G}_\mathcal{K},s}(\Diamond^i s'))\\
&\quad=\mathbb{E}^{\pi_\Box^\vee,\pi_\Diamond^\vee}_{\mathcal{G}_\mathcal{K},s}[\mathsf{rew}_\mathsf{f}] && \text{(by Lemma 1.2)}
\end{aligned}
$$

To prove item 2 we proceed similarly. Thus the proof boils down to show that for all $n\geq 0$ and $\hat\rho\in\mathcal{S}^{n+1}$, $\mathbb{P}^{\pi_\Box^\times,\pi_\Diamond^\times,n}_{\mathcal{G}_\mathcal{K},s}(\hat\rho)=\mathbb{P}^{\pi_\Box,\pi_\Diamond,n}_{\mathcal{H}_\mathcal{K},s}(\hat\rho)$ which can be done just like for (93). □

Proof (of Proposition 2). Let $\pi_i\in\Pi^{XS}_{\mathcal{G}_\mathcal{K},i}$. Then, for all $\hat\rho\in\mathcal{S}^*$, $s\in\mathcal{S}$, and $A\in\boldsymbol{\Sigma}_\mathcal{A}$, $(\pi_i^\vee)^\times(\hat\rho s)(A)=\pi_i^\vee(\hat\rho s)(A\cap\mathbb{V}(\mathcal{A}))=\pi_i(\hat\rho s)(A\cap\mathbb{V}(\mathcal{A}))=\pi_i(\hat\rho s)(A)$. All equalities follows from definitions except the last one that follows from the fact that π_i is extreme. Similarly, for $\pi_i\in\Pi^S_{\mathcal{H}_\mathcal{K},i}$, and all $\hat\rho\in\mathcal{S}^*$, $s\in\mathcal{S}$, and $A\subseteq\mathbb{V}(\mathcal{A})$, $(\pi_i^\times)^\vee(\hat\rho s)(A)=\pi_i^\times(\hat\rho s)(A)=\pi_i(\hat\rho s)(A\cap\mathbb{V}(\mathcal{A}))=\pi_i(\hat\rho s)(A)$. Again, all equalities follows from definitions except the last one that follows from the fact that $A\subseteq\mathbb{V}(\mathcal{A})$. From these observations, it follows that

$$\Pi^S_{\mathcal{H}_\mathcal{K},i}=\{\pi_i^\vee\mid\pi_i\in\Pi^{XS}_{\mathcal{G}_\mathcal{K},i}\}\text{ and }\Pi^{XS}_{\mathcal{G}_\mathcal{K},i}=\{\pi_i^\times\mid\pi_i\in\Pi^S_{\mathcal{H}_\mathcal{K},i}\}.\tag{102}$$

These equalities and Proposition 5 leads to the result. □

Proof (Proof of Proposition 3). For item (1) we have that

$$
\begin{aligned}
1&=\inf_{\pi_\Diamond\in\Pi_{\mathcal{G}_\mathcal{K},\Diamond}}\ \inf_{\pi_\Box\in\Pi_{\mathcal{G}_\mathcal{K},\Box}}\ \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond T) && \text{(def. of stopping)}\\
&\leq\inf_{\pi_\Diamond\in\Pi^{XS}_{\mathcal{G}_\mathcal{K},\Diamond}}\ \inf_{\pi_\Box\in\Pi^{XS}_{\mathcal{G}_\mathcal{K},\Box}}\ \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{G}_\mathcal{K},s}(\Diamond T) && (\Pi^{XS}_{\mathcal{G}_\mathcal{K},i}\subseteq\Pi_{\mathcal{G}_\mathcal{K},\Diamond})\\
&=\inf_{\pi_\Diamond\in\Pi^{S}_{\mathcal{H}_\mathcal{K},\Diamond}}\ \inf_{\pi_\Box\in\Pi^{S}_{\mathcal{H}_\mathcal{K},\Box}}\ \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond T) && \text{(by observation (102))}\\
&=\inf_{\pi_\Diamond\in\Pi_{\mathcal{H}_\mathcal{K},\Diamond}}\ \inf_{\pi_\Box\in\Pi_{\mathcal{H}_\mathcal{K},\Box}}\ \mathbb{P}^{\pi_\Box,\pi_\Diamond}_{\mathcal{H}_\mathcal{K},s}(\Diamond T)
\end{aligned}
$$

Since $\mathcal{H}_\mathcal{K}$ is finite, the last step follows from standard results on MDP [27].

With a similar reasoning we obtain $0 < \inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \inf_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond s') \leq \inf_{\pi_\diamond \in \Pi_{\mathcal{H}_\mathcal{K},\diamond}} \inf_{\pi_\Box \in \Pi_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\diamond s')$, proving thus also item (2). $\Box$

Proof (of Theorem 1). For item 1, we calculate as follows:

$$
\begin{aligned}
&\inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C) \\
&\qquad \leq \inf_{\pi_\diamond \in \Pi^S_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C) && (\Pi^S_{\mathcal{G}_\mathcal{K},\diamond} \subseteq \Pi_{\mathcal{G}_\mathcal{K},\diamond}) \\
&\qquad = \inf_{\pi_\diamond \in \Pi^S_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi^S_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C) && \text{(by Lemma 2.2)} \\
&\qquad = \inf_{\pi_\diamond \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C) && \text{(by Corollary 1.1)} \\
&\qquad = \inf_{\pi_\diamond \in \Pi^S_{\mathcal{H}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi^S_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\diamond C) && \text{(by Prop. 2.1)} \\
&\qquad \leq \inf_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi^S_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\diamond C) && (\Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond} \subseteq \Pi^S_{\mathcal{H}_\mathcal{K},\diamond}) \\
&\qquad = \inf_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\diamond C) && \text{(by Prop. 4.1)} \\
&\qquad = \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \inf_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\diamond C) && \text{(by [13, Lemma 6])} \\
&\qquad = \sup_{\pi_\Box \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box}} \inf_{\pi_\diamond \in \Pi^S_{\mathcal{H}_\mathcal{K},\diamond}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\diamond C) && \text{(by Prop. 4.2)} \\
&\qquad \leq \sup_{\pi_\Box \in \Pi^S_{\mathcal{H}_\mathcal{K},\Box}} \inf_{\pi_\diamond \in \Pi^S_{\mathcal{H}_\mathcal{K},\diamond}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\diamond C) && (\Pi^{MD}_{\mathcal{H}_\mathcal{K},\Box} \subseteq \Pi^S_{\mathcal{H}_\mathcal{K},\Box}) \\
&\qquad = \sup_{\pi_\Box \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\diamond \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C) && \text{(by Prop. 2.2)} \\
&\qquad = \sup_{\pi_\Box \in \Pi^S_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\diamond \in \Pi^S_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C) && \text{(by Corollary 1.1)} \\
&\qquad = \sup_{\pi_\Box \in \Pi^S_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C) && \text{(by Lemma 2.2)} \\
&\qquad \leq \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C) && (\Pi^S_{\mathcal{G}_\mathcal{K},\Box} \subseteq \Pi_{\mathcal{G}_\mathcal{K},\Box}) \\
&\qquad \leq \inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{P}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\diamond C) && \text{(by prop. of sup and inf)}
\end{aligned}
$$

Since the last term is equal to the first term in the calculation, item 1 is concluded.

For item 2 we calculate as follows:

$$
\begin{aligned}
&\inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) \\
&\qquad \leq \inf_{\pi_\diamond \in \Pi^S_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\Box \in \Pi_{\mathcal{G}_\mathcal{K},\Box}} \mathbb{E}^{\pi_\Box,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew_f}) && (\Pi^S_{\mathcal{G}_\mathcal{K},\diamond} \subseteq \Pi_{\mathcal{G}_\mathcal{K},\diamond})
\end{aligned}
$$

$$
\begin{aligned}
&= \inf_{\pi_\diamond \in \Pi^{S}_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\square \in \Pi^{S}_{\mathcal{G}_\mathcal{K},\square}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && \text{(by Lemma 2.3)}\\
&= \inf_{\pi_\diamond \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\square \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\square}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && \text{(by Corollary 1.2)}\\
&= \inf_{\pi_\diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\diamond}} \sup_{\pi_\square \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\square}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && \text{(by Prop. 2.3)}\\
&\le \inf_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \sup_{\pi_\square \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\square}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && (\Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond} \subseteq \Pi^{S}_{\mathcal{H}_\mathcal{K},\diamond})\\
&= \inf_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \sup_{\pi_\square \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\square}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && \text{(by Prop. 4.3)}\\
&= \sup_{\pi_\square \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\square}} \inf_{\pi_\diamond \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && (*)\\
&= \sup_{\pi_\square \in \Pi^{MD}_{\mathcal{H}_\mathcal{K},\square}} \inf_{\pi_\diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && \text{(by Prop. 4.4)}\\
&\le \sup_{\pi_\square \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\square}} \inf_{\pi_\diamond \in \Pi^{S}_{\mathcal{H}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{H}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && (\Pi^{MD}_{\mathcal{H}_\mathcal{K},\square} \subseteq \Pi^{S}_{\mathcal{H}_\mathcal{K},\square})\\
&= \sup_{\pi_\square \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\square}} \inf_{\pi_\diamond \in \Pi^{XS}_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && \text{(by Prop. 2.4)}\\
&= \sup_{\pi_\square \in \Pi^{S}_{\mathcal{G}_\mathcal{K},\square}} \inf_{\pi_\diamond \in \Pi^{S}_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && \text{(by Corollary 1.2)}\\
&= \sup_{\pi_\square \in \Pi^{S}_{\mathcal{G}_\mathcal{K},\square}} \inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && \text{(by Lemma 2.3)}\\
&\le \sup_{\pi_\square \in \Pi_{\mathcal{G}_\mathcal{K},\square}} \inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && (\Pi^{S}_{\mathcal{G}_\mathcal{K},\square} \subseteq \Pi_{\mathcal{G}_\mathcal{K},\square})\\
&\le \inf_{\pi_\diamond \in \Pi_{\mathcal{G}_\mathcal{K},\diamond}} \sup_{\pi_\square \in \Pi_{\mathcal{G}_\mathcal{K},\square}} \mathbb{E}^{\pi_\square,\pi_\diamond}_{\mathcal{G}_\mathcal{K},s}(\mathsf{rew}_\mathsf{f}) && \text{(by prop. of sup and inf)}
\end{aligned}
$$

Since the last term is equal to the first term in the calculation, item 2 is concluded. In particular, step (*) is justified as follows, depending on rew_f:

- For $\mathsf{rew}_\mathsf{f} = \mathsf{rew}_\mathsf{t}$, (*) follows by [15, Theorem 4.2.6] since, by Proposition 3.(1), the game $\mathcal{H}_\mathcal{K}$ is also almost surely stopping.
- For $\mathsf{rew}_\mathsf{f} = \mathsf{rew}_\gamma$ (*) follows by [15, Theorem 4.3.2].
- For $\mathsf{rew}_\mathsf{f} = \mathsf{rew}_\mathsf{a}$ (*) follows by [15, Theorem 5.1.5] since, by Proposition 3.(2), the game $\mathcal{H}_\mathcal{K}$ is also irreducible. □