IST AUSTRIA

Institute of Science and Technology

Robustness of Structurally Equivalent Concurrent Parity Games

Krishnendu Chatterjee

IST Austria (Institute of Science and Technology Austria)

Am Campus 1

A-3400 Klosterneuburg

Technical Report No. IST-2011-0006

http://pub.ist.ac.at/Pubs/TechRpts/2011/IST-2011-0006.pdf

June 27, 2011

Robustness of Structurally Equivalent Concurrent Parity Games

Krishnendu Chatterjee

IST Austria (Institute of Science and Technology Austria)

Abstract. We consider two-player stochastic games played on a finite state space for an infinite number of rounds. The games are *concurrent*: in each round, the two players (player 1 and player 2) choose their moves independently and simultaneously; the current state and the two moves determine a probability distribution over the successor states. We also consider the important special case of turn-based stochastic games where players make moves in turns, rather than concurrently. We study concurrent games with ω-regular winning conditions specified as *parity* objectives. The value for player 1 for a parity objective is the maximal probability with which the player can guarantee the satisfaction of the objective against all strategies of the opponent. We study the problem of continuity and robustness of the value function in concurrent and turn-based stochastic parity games with respect to imprecision in the transition probabilities. We present quantitative bounds on the difference of the value function (in terms of the imprecision of the transition probabilities) and show the value continuity for structurally equivalent concurrent games (two games are structurally equivalent if the support of the transition function is same and the probabilities differ). We also show robustness of optimal strategies for structurally equivalent turn-based stochastic parity games. Finally we show that the value continuity property breaks without the structurally equivalent assumption (even for Markov chains) and show that our quantitative bound is asymptotically optimal. Hence our results are tight (the assumption is both necessary and sufficient) and optimal (our quantitative bound is asymptotically optimal).

1 Introduction

Concurrent stochastic games are played by two players on a finite state space for an infinite number of rounds. In every round, the two players simultaneously and independently choose moves (or actions), and the current state and the two chosen moves determine a probability distribution over the successor states. The outcome of the game (or a *play*) is an infinite sequence of states. These games were introduced by Shapley [24], and has been one of the most fundamental and well studied game models in stochastic graph games. We consider ω-regular objectives specified as parity objectives; that is, given an ω-regular set Φ of infinite state sequences, player 1 wins if the outcome of the game lies in Φ. Otherwise, player 2 wins, i.e., the game is zero-sum. The class of concurrent stochastic games subsumes many other important classes of games as sub-classes: (1) *turn-based stochastic* games where at every round only one-player chooses moves (i.e., the players make moves in turns); and (2) *Markov decision processes* (one-player stochastic games). Concurrent games and its sub-classes provide a rich framework to model various classes of dynamic reactive systems, and ω-regular objectives provide a robust specification language to express all commonly used properties in verification. Thus concurrent games with parity objectives provide the mathematical framework to study many important problems in the synthesis and verification of reactive systems [6, 23, 21] (see also [1, 13, 2]).

The player-1 *value* $v_1(s)$ of the game at a state s is the limit probability with which player 1 can ensure that the outcome of the game lies in Φ; that is, the value $v_1(s)$ is the maximal probability with which player 1 can guarantee Φ against all strategies of player 2. Symmetrically, the player-2 *value* $v_2(s)$ is the limit probability with which player 2 can ensure that the outcome of the game lies outside Φ. The problem of studying the computational complexity of MDPs, turn-based stochastic games, and concurrent games with parity objectives has received a lot of attention in literature. The problem of Markov decision processes with ω-regular objectives has been studied in [8, 9, 4] and the results show existence of pure (deterministic) memoryless (stationary) optimal strategies for parity objectives and the problem of value computation is achievable in polynomial time. Turn-based stochastic games with the special case of reachability objectives has been studied in [7] and existence of pure memoryless optimal strategies has been established and the decision problem of whether the value at a state is at least a given rational value lie in NP $\cap$ coNP. The existence of pure memoryless optimal strategies for turn-based stochastic games with parity objectives was established in [5, 28], and again the decision problem lie in NP $\cap$ coNP. Concurrent parity games has been studied in [10, 12, 3, 14] and for concurrent parity games optimal strategies need not exist, and ε-optimal strategies (for $\varepsilon > 0$) require both infinite memory and randomization, and the decision problem can be solved in PSPACE.

Almost all results in the literature consider the problem of computing values and optimal strategies when the game model is given precisely along with the objective. However it is often unrealistic to know the precise probabilities of transition which are only estimated through observation. Since the transition probabilities are not known precisely, an extremely important question is how robust is the analysis of concurrent games and its sub-classes with parity objectives with respect to small changes in the transition probabilities. This question has been largely ignored in the study of concurrent and turn-based stochastic parity games. In this paper we study the following problems related to continuity and robustness of values: (1) *(continuity of values).* under what conditions can continuity of the value function be proved for concurrent parity games; (2) *(robustness of values).* can quantitative bounds be obtained on the difference of the value function in terms of the difference of the transition probabilities; and (3) *(robustness of optimal strategies).* does optimal strategies of a game remain ε-optimal, for $\varepsilon > 0$, if the transition probabilities are slightly changed.

Our contributions. Our contributions are as follows:

1. We consider *structurally equivalent* game structures, where the support of the transition probabilities are the same, but the precise transition probabilities may differ. We show the following results for structurally equivalent concurrent parity games:
   (a) *Quantitative bound.* We present a quantitative bound on the difference of the value function of two structurally equivalent game structures in terms of the difference of the transition probabilities. We show when the difference in the transition probabilities are small, our bound is asymptotically optimal. Our example to show the matching lower bound is on a Markov chain, and thus our result shows that the bound for a Markov chain can be generalized to concurrent games.
   (b) *Value continuity.* We show *value continuity* for structurally equivalent concurrent parity games, i.e., as the difference in transition probabilities goes to 0, the difference in value functions also goes to 0. We then show that the structurally equivalent assumption is necessary: we show a family of Markov chains (that are not structurally equivalent) where the difference of the transition probabilities goes to 0, but the difference in the value function is 1. It follows that the structural equivalence assumption is both necessary (even for Markov chains) and sufficient (even for concurrent games).

   It follows from above that our results are both optimal (quantitative bounds) as well as tight (assumption both necessary and sufficient). Our result for concurrent parity games is also a significant quantitative generalization of a result for concurrent parity games of [10] which shows that the set of states with value 1 remains same if the games are structurally equivalent. We also argue that the structurally equivalent assumption is not unrealistic in many cases: a reactive system consists of many state variables, and given a state (valuation of variables) it is typically known which variables are possibly updated, and what is unknown is the precise transition probabilities (which are estimated by observation). Thus the system that is obtained for analysis is structurally equivalent to the underlying original system and it only differs in precise transition probabilities.
2. For turn-based stochastic parity games the value continuity and the quantitative bounds are same as for concurrent games. We also prove a stronger result for structurally equivalent turn-based stochastic games that shows that along with continuity of value function, there is also robustness property for pure memoryless optimal strategies. More precisely, for all $\varepsilon > 0$, we present a bound $\beta > 0$, such that any pure memoryless optimal strategy in a turn-based stochastic parity game is an ε-optimal strategy in a structurally equivalent turn-based stochastic game such that the transition probabilities differ by at most β. Our result has deep significance as it allows the rich literature of work on turn-based stochastic games to carry over robustly for structurally equivalent turn-based stochastic games. As argued before the model of turn-based stochastic game obtained to analyze may differ slightly in precise transition

probabilities, and our results shows that the analysis on the slightly imprecise model using the classical results carry over to the underlying original system with small error bounds.

Our results are obtained as follows. The result of [11] shows that the value function for concurrent parity games can be characterized as the limit of the value function of concurrent multi-discounted games. There exists bound on difference on value function of discounted games [15], however, the bound depends on the discount factor, and in the limit gives trivial bounds (and in general this approach does not work as value continuity cannot be proven in general and the structural equivalence assumption is necessary). We use a classical result on Markov chains by Friedlin and Wentzell [16] and generalize a result of Solan [25] from Markov chains with single discount to Markov chains with multi-discounted objective to obtain a bound that is independent of the discount factor for structurally equivalent games. Then the bound also applies when we take the limit of the discount factors, and gives us the desired bound.

Our paper is organized as follows: in Section 2 we present the basic definitions, in Section 3 we consider Markov chains with multi-discounted and parity objectives; in Section 4 (Subsection 4.1) we prove the results related to turn-based stochastic games (item (2) of our contributions) and finally in Subsection 4.2 we present the quantitative bound and value continuity for concurrent games along with the two examples to illustrate the asymptotic optimality of the bound and the structural equivalence assumption is necessary.

2 Definitions

In this section we define game structures, strategies, objectives, values and present other preliminary definitions.

2.1 Game structures

Probability distributions. For a finite set A, a *probability distribution* on A is a function $\delta : A \mapsto [0,1]$ such that $\sum_{a \in A} \delta(a) = 1$. We denote the set of probability distributions on A by $\mathcal{D}(A)$. Given a distribution $\delta \in \mathcal{D}(A)$, we denote by $\mathrm{Supp}(\delta) = \{x \in A \mid \delta(x) > 0\}$ the *support* of the distribution δ.

Concurrent game structures. A (two-player) *concurrent stochastic game structure* $G = \langle S, A, \Gamma_1, \Gamma_2, \delta \rangle$ consists of the following components.

- A finite state space S.
- A finite set A of moves (or actions).
- Two move assignments $\Gamma_1, \Gamma_2 : S \mapsto 2^A \setminus \emptyset$. For $i \in \{1,2\}$, assignment Γ_i associates with each state $s \in S$ the nonempty set $\Gamma_i(s) \subseteq A$ of moves available to player i at state s.
- A probabilistic transition function $\delta : S \times A \times A \mapsto \mathcal{D}(S)$, which associates with every state $s \in S$ and moves $a_1 \in \Gamma_1(s)$ and $a_2 \in \Gamma_2(s)$ a probability distribution $\delta(s, a_1, a_2) \in \mathcal{D}(S)$ for the successor state.

Plays. At every state $s \in S$, player 1 chooses a move $a_1 \in \Gamma_1(s)$, and simultaneously and independently player 2 chooses a move $a_2 \in \Gamma_2(s)$. The game then proceeds to the successor state t with probability $\delta(s, a_1, a_2)(t)$, for all $t \in S$. For all states $s \in S$ and moves $a_1 \in \Gamma_1(s)$ and $a_2 \in \Gamma_2(s)$, we indicate by $\mathrm{Dest}(s, a_1, a_2) = \mathrm{Supp}(\delta(s, a_1, a_2))$ the set of possible successors of s when moves a_1, a_2 are selected. A *path* or a *play* of G is an infinite sequence $\omega = \langle s_0, s_1, s_2, \ldots \rangle$ of states in S such that for all $k \geq 0$, there are moves $a_1^k \in \Gamma_1(s_k)$ and $a_2^k \in \Gamma_2(s_k)$ such that $s_{k+1} \in \mathrm{Dest}(s_k, a_1^k, a_2^k)$. We denote by Ω the set of all paths. We denote by θ_i the random variable that denotes the i-th state of a path. For a play $\omega = \langle s_0, s_1, s_2, \ldots \rangle \in \Omega$, we define $\mathrm{Inf}(\omega) = \{s \in S \mid s_k = s \text{ for infinitely many } k \geq 0\}$ to be the set of states that occur infinitely often in ω.

Special classes of concurrent games. We will consider the following special classes of concurrent games.

1. *Turn-based stochastic games.* A game structure G is *turn-based stochastic* if at every state at most one player can choose among multiple moves; that is, for every state $s \in S$ there exists at most one $i \in \{1,2\}$ with $|\Gamma_i(s)| > 1$.

2. *Markov decision processes.* A game structure is a *player-1 Markov decision process (MDP)* if for all $s \in S$ we have $|\Gamma_2(s)| = 1$, i.e., only player-1 has choice of actions in the game. Similarly, a game structure is a *player-2 MDP* if for all $s \in S$ we have $|\Gamma_1(s) = 1$.
3. *Markov chains.* A game structure is a Markov chain if for all $s \in S$ we have $|\Gamma_1(s)| = 1$ and $|\Gamma_2(s)| = 1$. Hence in a Markov chain the players do not matter, and for the rest of the paper a Markov chain consists of a tuple (S, δ) where $\delta : S \to \mathcal{D}(S)$ is the probabilistic transition function.

2.2 Strategies

A *strategy* for a player is a recipe that describes how to extend a play. Formally, a strategy for player $i \in \{1, 2\}$ is a mapping $\pi_i : S^+ \mapsto \mathcal{D}(A)$ that associates with every nonempty finite sequence $x \in S^+$ of states, representing the past history of the game, a probability distribution $\pi_i(x)$ used to select the next move. The strategy π_i can prescribe only moves that are available to player i; that is, for all sequences $x \in S^*$ and states $s \in S$, we require that $\mathrm{Supp}(\pi_i(x \cdot s)) \subseteq \Gamma_i(s)$. We denote by Π_i the set of all strategies for player $i \in \{1, 2\}$.

Given a state $s \in S$ and two strategies $\pi_1 \in \Pi_1$ and $\pi_2 \in \Pi_2$, we define $\mathrm{Outcome}(s, \pi_1, \pi_2) \subseteq \Omega$ to be the set of paths that can be followed by the game, when the game starts from s and the players use the strategies π_1 and π_2. Formally, $\langle s_0, s_1, s_2, \ldots \rangle \in \mathrm{Outcome}(s, \pi_1, \pi_2)$ if $s_0 = s$ and if for all $k \geq 0$ there exist moves $a_1^k \in \Gamma_1(s_k)$ and $a_2^k \in \Gamma_2(s_k)$ such that

$$\pi_1(s_0, \ldots, s_k)(a_1^k) > 0, \quad \pi_2(s_0, \ldots, s_k)(a_2^k) > 0, \quad s_{k+1} \in \mathrm{Dest}(s_k, a_1^k, a_2^k).$$

Once the starting state s and the strategies π_1 and π_2 for the two players have been chosen, the probabilities of events are uniquely defined [27], where an *event* $\mathcal{A} \subseteq \Omega$ is a measurable set of paths[1]. For an event $\mathcal{A} \subseteq \Omega$, we denote by $\Pr_s^{\pi_1,\pi_2}(\mathcal{A})$ the probability that a path belongs to $\mathcal{A}$ when the game starts from s and the players use the strategies π_1 and π_2.

Classification of strategies. We consider the following special classes of strategies.

1. *(Pure).* A strategy π is *pure (deterministic)* if for all $x \in S^+$ there exists $a \in A$ such that $\pi(x)(a) = 1$. Thus, deterministic strategies are equivalent to functions $S^+ \mapsto A$.
2. *(Finite-memory).* Strategies in general are *history-dependent* and can be represented as follows: let $\mathtt{M}$ be a set called *memory* to remember the history of plays (the set $\mathtt{M}$ can be infinite in general). A strategy with memory can be described as a pair of functions: (a) a *memory update* function $\pi_u : S \times \mathtt{M} \mapsto \mathtt{M}$, that given the memory $\mathtt{M}$ with the information about the history and the current state updates the memory; and (b) a *next move* function $\pi_n : S \times \mathtt{M} \mapsto \mathcal{D}(A)$ that given the memory and the current state specifies the next move of the player. A strategy is *finite-memory* if the memory $\mathtt{M}$ is finite.
3. *(Memoryless).* A *memoryless* strategy is independent of the history of play and only depends on the current state. Formally, for a memoryless strategy π we have $\pi(x \cdot s) = \pi(s)$ for all $s \in S$ and all $x \in S^*$. Thus memoryless strategies are equivalent to functions $S \mapsto \mathcal{D}(A)$.
4. *(Pure memoryless).* A strategy is *pure memoryless* if it is both pure and memoryless. The pure memoryless strategy neither use memory, nor use randomization and are equivalent to functions $S \mapsto A$.

2.3 Objectives

Qualitative objectives. We specify *qualitative* objectives for the players by providing the set of *winning plays* $\Phi \subseteq \Omega$ for each player. In this paper we study only zero-sum games [22, 15], where the objectives of the two players are complementary. A general class of objectives are the Borel objectives [18]. A *Borel objective* $\Phi \subseteq S^\omega$ is a Borel set in the Cantor topology on S^ω. In this paper we consider *ω-regular objectives* specified as parity objectives, which lie in the first $2\,{}^1\!/_2$ levels of the Borel hierarchy (i.e., in the intersection of Σ_3 and Π_3) [26].

[1] To be precise, we should define events as measurable sets of paths *sharing the same initial state,* and we should replace our events with families of events, indexed by their initial state. However, our (slightly) improper definition leads to more concise notation.

– *Parity objectives.* For $c, d \in \mathbb{N}$, we let $[c..d] = \{c, c+1, \ldots, d\}$. Let $p : S \mapsto [0..d]$ be a function that assigns a *priority* $p(s)$ to every state $s \in S$, where $d \in \mathbb{N}$. The *Even parity objective* requires that the minimum priority visited infinitely often is even. Formally, the set of winning plays is defined as Parity$(p) = \{\omega \in \Omega \mid \min\big(p(\mathrm{Inf}(\omega))\big)$ is even $\}$.

Quantitative objectives. Quantitative objectives are measurable functions $f : \Omega \to \mathbb{R}$. We will consider *multi-discounted* objective function, as there is a close connection established between concurrent games with multi-discounted objectives and concurrent games with parity objectives. Given a concurrent game structure with state space S, let $\boldsymbol{\lambda}$ be a *discount vector* that assigns for all $s \in S$ a discount factor $0 < \lambda(s) < 1$ (unless otherwise mentioned we will always consider discount vectors $\boldsymbol{\lambda}$ such that for all $s \in S$ we have $0 < \lambda(s) < 1$). Let $r : S \to \mathbb{R}$ be a reward function that assigns a real-valued reward $r(s)$ to every state $s \in S$. The multi-discounted objective function $\mathsf{MDT}(\boldsymbol{\lambda}, r) : \Omega \to \mathbb{R}$ maps every path to the mean-discounted reward of the path. Formally, the function is defined as follows: for a path $\omega = s_0 s_1 s_2 \ldots$ we have

$$\mathsf{MDT}(\omega, \boldsymbol{\lambda}, r) = \frac{\sum_{j=0}^{\infty}(\prod_{i=0}^{j} \lambda(s_i)) \cdot r(s_j)}{\sum_{j=0}^{\infty}(\prod_{i=0}^{j} \lambda(s_i))}.$$

Values, optimality, ε-optimality. Given an objective Φ which is a measurable function $\Phi : \Omega \to \mathbb{R}$, we define the *value* for player 1 of game G with objective Φ from the state $s \in S$ as

$$\mathsf{Val}(G, \Phi)(s) = \sup_{\pi_1 \in \Pi_1} \inf_{\pi_2 \in \Pi_2} \mathbb{E}_s^{\pi_1, \pi_2}(\Phi);$$

i.e., the value is the maximal expectation with which player 1 can guarantee the satisfaction of Φ against all player 2 strategies. Given a player-1 strategy π_1, we use the notation

$$\mathsf{Val}^{\pi_1}(G, \Phi)(s) = \inf_{\pi_2 \in \Pi_2} \mathbb{E}_s^{\pi_1, \pi_2}(\Phi).$$

A strategy π_1 for player 1 is *optimal* for an objective Φ if for all states $s \in S$, we have

$$\mathsf{Val}^{\pi_1}(G, \Phi)(s) = \mathsf{Val}(G, \Phi)(s).$$

For $\varepsilon > 0$, a strategy π_1 for player 1 is ε-*optimal* if for all states $s \in S$, we have

$$\mathsf{Val}^{\pi_1}(G, \Phi)(s) \geq \mathsf{Val}(G, \Phi)(s) - \varepsilon.$$

The notion of values, optimal and ε-optimal strategies for player 2 are defined analogously. The following theorem summarizes the results in literature related to determinacy and memory complexity of concurrent games and its sub-classes for parity and multi-discounted objectives.

Theorem 1. *The following assertions hold:*

1. (Determinacy [19]). *For all concurrent game structures and for all parity and multi-discounted objectives Φ we have*
$$\sup_{\pi_1 \in \Pi_1} \inf_{\pi_2 \in \Pi_2} \mathbb{E}_s^{\pi_1, \pi_2}(\Phi) = \inf_{\pi_2 \in \Pi_2} \sup_{\pi_1 \in \Pi_1} \mathbb{E}_s^{\pi_1, \pi_2}(\Phi).$$
2. (Memory complexity). *For all concurrent game structures and for all multi-discounted objectives Φ, randomized memoryless optimal strategies exist [24]. For all turn-based stochastic game structures and for all multi-discounted objectives Φ, pure memoryless optimal strategies exist [15]. For all turn-based stochastic game strucutures and for all parity objectives Φ, pure memoryless optimal strategies exist [5, 28]. In general optimal strategies need not exist in concurrent games with parity objectives, and ε-optimal strategies, for $\varepsilon > 0$, need both randomization and infinite memory in general [10].*

The results of [11] established that the value of concurrent games with certain special multi-discounted objectives can be characterized as valuations of quantitaive discounted μ-calculus formula. In the limit, the value function of the discounted μ-calculus formula characterizes the value function of concurrent games with parity objectives. An elegant interpretation of the result was given in [17], and from the interpretation we obtain the following theorem.

Theorem 2 ([11]). *Let G be a concurrent game structure with a parity objective Φ defined by a priority function p. Let r be a reward function that assigns reward 1 to even priority states and reward 0 to odd priority states. Then there exists an order $s_1 s_2 \ldots s_n$ on the states (where $S = \{s_1, s_2, \ldots, s_n\}$) dependent only on the priority function p such that*

$$\mathsf{Val}(G, \Phi) = \lim_{\lambda(s_1)\to 1} \lim_{\lambda(s_2)\to 1} \ldots \lim_{\lambda(s_n)\to 1} \mathsf{Val}(G, \mathsf{MDT}(\boldsymbol{\lambda}, r));$$

in other words, if we consider the value function $\mathsf{Val}(G, \mathsf{MT}(\boldsymbol{\lambda}, r))$ with the multi-discounted objective and take the limit of the discount factors to 1 in the order of the states we obtain the value function for the parity objective.

2.4 Structure equivalent game structures and distance of game structures

In this sub-section we present notions related to *structure equivalent* game structures.

Structure equivalent game structures. Given two game structures $G_1 = \langle S, A, \Gamma_1, \Gamma_2, \delta_1 \rangle$ and $G_2 = \langle S, A, \Gamma_1, \Gamma_2, \delta_2 \rangle$ on the same state and action space, with different transition function, we say that G_1 and G_2 are *structure equivalent* (denoted $G_1 \equiv G_2$) if for all $s \in S$ and all $a_1 \in \Gamma_1(s)$ and $a_2 \in \Gamma_2(s)$ we have $\mathrm{Supp}(\delta_1(s, a_1, a_2)) = \mathrm{Supp}(\delta_2(s, a_1, a_2))$. Similarly, two Markov chains $G_1 = (S, \delta_1)$ and $G_2 = (S, \delta_2)$ are structurally equivalent (denoted $G_1 \equiv G_2$) if for all $s \in S$ we have $\mathrm{Supp}(\delta_1(s)) = \mathrm{Supp}(\delta_2(s))$. For a game structure G (resp. Markov chain G) we denote by $[\![G]\!]_{\equiv}$ the set of all game structures (resp. Markov chains) that are structurally equivalent to G.

Ratio and absolute distances. Given two game structures $G_1 = \langle S, A, \Gamma_1, \Gamma_2, \delta_1 \rangle$ and $G_2 = \langle S, A, \Gamma_1, \Gamma_2, \delta_2 \rangle$, the *absolute distance* of the game structures is maximum absolute difference in the transition probabilities. Formally,

$$dist_A(G_1, G_2) = \max_{s,t \in S, a \in \Gamma_1(s), b \in \Gamma_2(s)} |\delta_1(s, a, b)(t) - \delta_2(s, a, b)(t)|.$$

The absolute distance for two Markov chains $G_1 = (S, \delta_1)$ and $G_2 = (S, \delta_2)$ is $dist_A(G_1, G_2) = \max_{s,t \in S} |\delta_1(s)(t) - \delta_2(s)(t)|$. We now define the ratio distance between two structurally equivalent game structures and Markov chains. Let G_1 and G_2 be two structurally equivalent game structures. The *ratio* distance is defined on the ratio of the transition probabilities. Formally,

$$dist_R(G_1, G_2) = \max \left\{ \frac{\delta_1(s, a, b)(t)}{\delta_2(s, a, b)(t)}, \frac{\delta_2(s, a, b)(t)}{\delta_1(s, a, b)(t)} \mid s \in S, a \in \Gamma_1(s), b \in \Gamma_2(s), \right.$$
$$\left. t \in \mathrm{Supp}(\delta_1(s, a, b)) = \mathrm{Supp}(\delta_2(s, a, b)) \right\} - 1$$

The ratio distance between two structurally equivalent Markov chains G_1 and G_2 is $\max \left\{ \frac{\delta_1(s)(t)}{\delta_2(s)(t)}, \frac{\delta_2(s)(t)}{\delta_1(s)(t)} \mid s \in S, t \in \mathrm{Supp}(\delta_1(s)) = \mathrm{Supp}(\delta_2(s)) \right\} - 1$.

Proposition 1. *Let G_1 be a game structure (resp. Markov chain) such that the minimum positive transition probability is $\eta > 0$. For all game structures (resp. Markov chains) $G_2 \in [\![G_1]\!]_{\equiv}$ we have*

$$dist_R(G_1, G_2) \le \frac{dist_A(G_1, G_2)}{\eta - dist_A(G_1, G_2)}$$

Proof. Consider $s \in S$, $a \in \Gamma_1(s), b \in \Gamma_2(s)$, and $t \in \mathrm{Supp}(\delta_1(s,a,b)) = \mathrm{Supp}(\delta_2(s,a,b))$. Then we have the following two inequalities: the first inequality is

$$\frac{\delta_2(s,a,b)(t)}{\delta_1(s,a,b)(t)} \leq \frac{\delta_1(s,a,b)(t) + dist_A(G_1,G_2)}{\delta_1(s,a,b)(t)} \leq 1 + \frac{dist_A(G_1,G_2)}{\delta_1(s,a,b)(t)} \leq 1 + \frac{dist_A(G_1,G_2)}{\eta}$$

and the second inequality is

$$\begin{aligned}\frac{\delta_1(s,a,b)(t)}{\delta_2(s,a,b)(t)} \leq \frac{\delta_1(s,a,b)(t)}{\delta_1(s,a,b)(t) - dist_A(G_1,G_2)} &\leq 1 + \frac{dist_A(G_1,G_2)}{\delta_1(s,a,b)(t) - dist_A(G_1,G_2)} \\ &\leq 1 + \frac{dist_A(G_1,G_2)}{\eta - dist_A(G_1,G_2)}\end{aligned}$$

The desired result follows from the above inequalities. ∎

Notation for fixing strategies. Given a concurrent game structure $G = \langle S, A, \Gamma_1, \Gamma_2, \delta\rangle$, let π_1 be a randomized memoryless strategy. Fixing the strategy π_1 in G we obtain a player-2 MDP, denoted as $G \upharpoonright \pi_1$, defined as follows: (1) the state space is S; (2) for all $s \in S$ we have $\Gamma_1(s) = \{\bot\}$ (hence it is a player-2 MDP); (3) the new transition function δ_{π_1} is defined as follows: for all $s \in S$ and all $b \in \Gamma_2(s)$ we have $\delta_{\pi_1}(s, \bot, b)(t) = \sum_{a \in \Gamma_1(s)} \pi_1(s)(a) \cdot \delta(s,a,b)(t)$. Similarly if we fix a randomized memoryless strategy π_1 in an MDP G we obtain a Markov chain, denoted as $G \upharpoonright \pi_1$. The following proposition is straight forward to verify from the definitions.

Proposition 2. *Let G_1 and G_2 be two concurrent game structures (resp. MDPs) that are structurally equivalent. Let π_1 be a randomized memoryless strategy. Then $dist_A(G_1 \upharpoonright \pi_1, G_2 \upharpoonright \pi_1) = dist_A(G_1, G_2)$ and $dist_R(G_1 \upharpoonright \pi_1, G_2 \upharpoonright \pi_1) = dist_R(G_1, G_2)$.*

3 Markov Chains with Multi-discounted and Parity Objectives

In this section we consider Markov chains with multi-discounted and parity objectives. We present a bound on the difference of value functions of two structurally equivalent Markov chains that is dependent on the distance between the Markov chains and is *independent* of the discount factors. The result for parity objectives is then a consequence of our result for multi-discounted objectives and Theorem 2. Our result crucially depends on a result of Friedlin and Wentzell for Markov chains and we present the result below, and then use the result to present the main result of the section.

3.1 Result of Friedlin and Wentzell

Let (S, δ) be a Markov chain and let s_0 be the initial state. Let $C \subset S$ be a proper subset of S and let us denote by $\mathsf{ex}_C = \inf\{n \in \mathbb{N} \mid \theta_n \not\in C\}$ the first hitting time to the set $S \setminus C$ of states (or the first exit time from set C) (recall that θ_n is the random variable to denote the n-th state of a path). Let $\mathcal{F}(C,S) = \{f : C \to S\}$ denote the set of all functions from C to S. For every $f \in \mathcal{F}(C,S)$ we define a directed graph $G_f = (S, E_f)$ where $(s,t) \in E_f$ iff $f(s) = t$. Let $\alpha_f = 1$ if the directed graph G_f has no directed cycles (i.e., G_f is a directed acyclic graph); and $\alpha_f = 0$ otherwise. Observe that since f is a function, for every $s \in C$ there is exactly one path that leaves C. For every $s \in C$ and every $t \in S$, let $\beta_f(s,t) = 1$ if the directed path that leaves s in G_f reaches t, otherwise $\beta_f(s,t) = 0$. We now state a result that can be obtained as a special case of the result from Friedlin and Wentzell [16].

Theorem 3 (see Lemma 6.3.3 of [20]). *Let (S, δ) be a Markov chain, and let $C \subset S$ be a proper subset of S such that $\Pr_s(\mathsf{ex}_C < \infty) > 0$ for every $s \in C$ (i.e., from all $s \in C$ with positive probability the first hitting time to the complement set is finite). Then for every initial state $s_1 \in C$ and for every $t \not\in C$ we have*

$$\Pr_{s_1}(\theta_{\mathsf{ex}_C} = t) = \frac{\sum_{f \in \mathcal{F}(C,S)} (\beta_f(s_1,t) \cdot \prod_{s \in C} \delta(s)(f(s)))}{\sum_{f \in \mathcal{F}(C,S)} (\alpha_f \cdot \prod_{s \in C} \delta(s)(f(s)))}, \tag{1}$$

in other words, the probability that the exit state is t when the starting state is s_1 is given by the expression on the right hand side.

We present an argument that the assumption that for all $s \in C$ we have $\Pr_s(\mathsf{ex}_C < \infty) > 0$ implies that the denominator of Equation (1) is positive (also see [20, 16, 25]). Since all terms in the summation of the denominator is non-negative, we show a witness function $f \in \mathcal{F}(C, S)$ such that $\alpha_f = 1$ and $\prod_{s \in C} \delta(s)(f(s)) > 0$. Let $s \in C$, and since $\Pr_s(\mathsf{ex}_C < \infty) > 0$, it follows that there exists $\ell > 1$ and a sequence of states $s_1 s_2 \ldots s_\ell$ with $s_1 = s$ such that $s_2, \ldots, s_{\ell-1} \in C$, $s_\ell \in (S \setminus C)$ and for all $i = 1, 2, \ldots, \ell - 1$ we have $\delta(s_i)(s_{i+1}) > 0$. Let us denote by ℓ_s the length of the shortest such sequence. We have the following two cases: (1) $\ell_s = 2$, i.e., there exists $t \in (S \setminus C)$ and $\delta(s)(t) > 0$; or (2) $\ell_s > 2$, and then there exists $t \in C$ with $\delta(s)(t) > 0$ and $\ell_s = \ell_t + 1$. We define the witness f as follows: (1) if $\ell_s = 2$, then $f(s) = t$, where t is any state in $S \setminus C$ with $\delta(s)(t) > 0$; (2) if $\ell_s > 2$, then $f(s) = t$, where $t \in C$ is a state in C such that $\delta(s)(t) > 0$ and $\ell_s = \ell_t + 1$. Since $s \in S$ is chosen arbitrarily, f is a function from C to S, and by construction we have $\prod_{s \in C} \delta(s)(f(s)) > 0$. Since for every $s \in C$, if $f(s) \in C$, then $\ell_{f(s)} + 1 = \ell_s$, it follows that the directed graph induced by f has no cycles and hence $\alpha_f = 1$.

3.2 Value function difference for Markov chains

In this sub-section we will use the result of previous sub-section to obtain bounds on the value functions of Markov chains. We start with the notion of mean-discounted time.

Mean-discounted time. Given a Markov chain (S, δ) and a discount vector $\boldsymbol{\lambda}$, we define for every state $s \in S$, the *mean-discounted time* the process is in the state s. We first define the mean-discounted time function $\mathsf{MDT}(\boldsymbol{\lambda}, s) : \Omega \to \mathbb{R}$ that maps every path to the mean-discounted time that the state s is visited, and the function is formally defined as follows: for a path $\omega = s_0 s_1 s_2 \ldots$ we have

$$\mathsf{MDT}(\boldsymbol{\lambda}, s)(\omega) = \frac{\sum_{j=0}^{\infty} (\prod_{i=0}^{j} \lambda(s_i)) \cdot \mathbf{1}_{s_j = s}}{\sum_{j=0}^{\infty} (\prod_{i=0}^{j} \lambda(s_i))};$$

where $\mathbf{1}_{s_j = s}$ is the indicator function. The expected mean-discounted time function for a Markov chain G with transition function δ is defined as follows: $\mathsf{MT}(s_1, s, G, \boldsymbol{\lambda}) = \mathbb{E}_{s_1}[\mathsf{MDT}(\boldsymbol{\lambda}, s)]$, i.e., it is the expected mean-discounted time for s when the starting state is s_1, where the expectation measure is defined by the Markov chain with transition function δ. We now present a lemma that shows the value function for multi-discounted Markov chains can be expressed as ratio of two polynomials.

Lemma 1. *For Markov chains defined on state space S, for all initial states s_0, for all states s, for all discount vectors $\boldsymbol{\lambda}$, there exists two polynomials $g_1(\cdot)$ and $g_2(\cdot)$ in $|S|^2$ variables $x_{t,t'}$, where $t, t' \in S$ such that the following conditions hold:*

1. *the polynomials have degree at most $|S|$ with non-negative coefficients; and*
2. *for all transition functions δ over S we have $\mathsf{MT}(s_0, s, G, \boldsymbol{\lambda}) = \frac{g_1(\delta)}{g_2(\delta)}$, where $G = (S, \delta)$, $g_1(\delta)$ and $g_2(\delta)$ denote the values of the function g_1 and g_2 such that the variables $x_{t,t'}$ is instantiated with values $\delta(t)(t')$ as given by the transition function δ.*

Proof. Fix a discount vector $\boldsymbol{\lambda}$. We construct a Markov chain $\overline{G} = (\overline{S}, \overline{\delta})$ as follows: $\overline{S} = S \cup S_1$, where S_1 is a copy of states of S (and for a state $s \in S$ we denote its corresponding copy as s_1); and the transition function $\overline{\delta}$ is defined below

1. $\overline{\delta}(s_1)(s_1)$ for all $s_1 \in S_1$ (i.e., all copy states are absorbing);
2. for $s \in S$ we have

$$\overline{\delta}(s)(t) = \begin{cases} (1 - \lambda(s)) & t = s_1; \\ \lambda(s) \cdot \delta(s)(t) & t \in S; \\ 0 & t \in S_1 \setminus s_1; \end{cases}$$

   i.e., it goes to the copy with probability $(1 - \lambda(s))$, it follows the transition δ in the original copy with probabilities multiplied by $\lambda(s)$.

We first show that for all s_0 and s we have

$$\mathsf{MT}(s_0, s, G, \boldsymbol{\lambda}) = \Pr_{s_0}^{\overline{\delta}}(\theta_{\mathsf{ex}_S} = s_1);$$

i.e., the expected mean-discounted time in s when the original Markov chain starts in s_0 is the probability in the Markov chain $(\overline{S}, \overline{\delta})$ that the first hitting state out of S is the copy s_1 of the state s. The claim is easy to verify as both $(\mathsf{MT}(s_0, s, G, \boldsymbol{\lambda}))_{s_0 \in S}$ and $(\Pr_{s_0}^{\overline{\delta}}(\theta_{\mathsf{ex}_S} = s_1))_{s_0 \in S}$ are the solutions of the following system of linear equations

$$y_t = (1 - \lambda(t)) \cdot \mathbf{1}_{t=s} + \sum_{z \in S} \lambda(z) \cdot \delta(t)(z) \cdot y_z \quad \forall t \in S.$$

Also the above system of linear equations has a unique solution (this is due to contraction mapping) and we prove this below: let $(y_z^1)_{z \in S}$ and $(y_z^2)_{z \in S}$ be two solutions of the system. We chose $z^* \in S$ such that $z^* = \arg\max_{z \in S} |y_z^1 - y_z^2|$, i.e., z^* is a state that maximizes the difference of the two solutions. Let $\eta = |y_{z^*}^1 - y_{z^*}^2|$. As y^1 and y^2 are solutions of the above system we have by the triangle inequality

$$\begin{aligned} 0 \leq \eta = |y_{z^*}^1 - y_{z^*}^2| &\leq \sum_{t \in S} \lambda(t) \cdot |y_t^1 - y_t^2| \\ &\leq \eta \cdot \sum_{t \in S} \lambda(t) \cdot \delta(s_0)(t) \leq \eta \cdot \max_{t \in S} \lambda(t) \cdot \sum_{t \in S} \delta(s_0)(t). \end{aligned}$$

Since $\sum_{t \in S} \delta(s_0)(t) = 1$, it follows that $\eta \leq \eta \cdot \max_{t \in S} \lambda(t)$. Since $\max_{t \in S} \lambda(t) < 1$ it follows that we must have $\eta = 0$ and hence the two solutions must coincide.

We now claim that $\Pr_{s_0}^{\overline{\delta}}(\mathsf{ex}_S < \infty) > 0$ for all $s_0 \in S$. This follows since for all $s \in S$ we have $\overline{\delta}(s)(s_1) = (1 - \lambda(s)) > 0$ and since $s_1 \notin S$ we have $\Pr_{s_0}^{\overline{\delta}}(\mathsf{ex}_S = 2) = (1 - \lambda(s_0)) > 0$. Now we observe that we can apply Theorem 3 on the Markov chain $\overline{G} = (\overline{S}, \overline{\delta})$ with S as the set C of states of Theorem 3, and obtain the result. Indeed the terms α_f and $\beta_f(s,t)$ are independent of δ, and the two prodtucs of Equation (1) each contains at most $|S|$ terms of the form $\overline{\delta}(s)(t)$ for $s, t \in \overline{S}$. Thus the desired result follows. ∎

Lemma 2. *Let $h(x_1, x_2, \ldots, x_k)$ be a polynomial function with non-negative coefficients of degree at most n. Let $\varepsilon > 0$ and $y, y' \in \mathbb{R}^k$ be two non-negative vectors such that for all $i = 1, 2, \ldots, k$ we have $\frac{1}{1+\varepsilon} \leq \frac{y_i}{y'_i} \leq 1 + \varepsilon$. Then we have*

$$(1+\varepsilon)^{-n} \leq \frac{h(y)}{h(y')} \leq (1+\varepsilon)^n$$

Proof. We first write $h(x)$ as follows:

$$h(x) = \sum_{i=1}^{\ell} a_i \cdot \prod_{j=1}^{n_i} x_{k_{ij}},$$

where $\ell \in \mathbb{N}$, for all $i = 1, 2, \ldots, \ell$ we have $a_i \geq 0$, $n_i \leq n$, and $1 \leq k_{ij} \leq k$ for each $j = 1, 2, \ldots, n_i$. By the hypothesis of the lemma, for all $i = 1, 2, \ldots, \ell$ we have

$$\frac{1}{(1+\varepsilon)^n} \cdot \prod_{j=1}^{n_i} y'_{k_{ij}} \leq \prod_{j=1}^{n_i} y_{k_{ij}} \leq (1+\varepsilon)^n \cdot \prod_{j=1}^{n_i} y'_{k_{ij}}.$$

Since every $a_i \geq 0$, multiplying the above inequalities by a_i and summing over $i = 1, 2, \ldots, \ell$ yields the desired result. ∎

Lemma 3. *Let $G_1 = (S, \delta)$ and $G_2 = (S, \delta')$ be two structurally equivalent Markov chains. For all non-negative reward functions $r : S \to \mathbb{R}$ such that the reward function is bounded by 1, for all discount vectors $\boldsymbol{\lambda}$, for all $s \in S$ we have*

$$|\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) - \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)| \leq (1 + dist_R(G_1, G_2))^{2 \cdot |S|} - 1;$$

i.e., the absolute difference of the value functions for the multi-discounted objective is bounded by $(1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1$.

Proof. We first observe that for a Markov chain G we have $\mathsf{Val}(G, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) = \sum_{t\in S} r(t) \cdot \mathsf{MT}(s, t, G, \boldsymbol{\lambda})$, i.e., the value function for a state s is obtained as the sum of the product of mean-discounted time of states and the rewards with s as the starting state. Hence by Lemma 2 it follows that $\mathsf{Val}(G, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)$ can be expressed as a ratio $\frac{g_1(\cdot)}{g_2(\cdot)}$ of two polynomials of degree at most $|S|$ over $|S|^2$ variables. Hence we have

$$\frac{\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)}{\mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)} = \frac{g_1(\delta)}{g_1(\delta')} \cdot \frac{g_2(\delta')}{g_2(\delta)}$$

Let $\varepsilon = dist_R(G_1, G_2)$. By definition for all $s_1, s_2 \in S$, if $s_2 \in \mathrm{Supp}(\delta(s_1))$, then we have both $\frac{\delta(s_1)(s_2)}{\delta'(s_1)(s_2)}$ and $\frac{\delta'(s_1)(s_2)}{\delta(s_1)(s_2)}$ are between $\frac{1}{1+\varepsilon}$ and $1+\varepsilon$. It follows from Lemma 2, with $k = |S|^2$ that

$$(1+\varepsilon)^{-|S|} \le \frac{g_i(\delta)}{g_i(\delta')} \le (1+\varepsilon)^{|S|}, \qquad \text{for } i \in \{1, 2\}.$$

Thus we have

$$(1+\varepsilon)^{-2\cdot|S|} \le \frac{g_1(\delta)}{g_1(\delta')} \cdot \frac{g_2(\delta')}{g_2(\delta)} \le (1+\varepsilon)^{2\cdot|S|}.$$

Hence we have

$$(1+\varepsilon)^{-2\cdot|S|} \le \frac{\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)}{\mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)} \le (1+\varepsilon)^{2\cdot|S|}$$

We consider the case when $\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) \ge \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)$, and the other case argument is symmetric. We also assume without loss of generality that $\mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) > 0$. Otherwise if $\mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) = 0$, since rewards are non-negative, it follows that no state with positive reward is reachable from s both in G_1 and G_2 (because if they are reachable, then they are reachable with positive probability and then the value is positive), and hence $\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r)) = \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r)) = 0$ and the result of the lemma follows trivially. Since we assume that $\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) \ge \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)$ and $\mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) > 0$, we have

$$\begin{aligned}
&|\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) - \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)| \\
&\qquad = \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) \cdot \left(\frac{\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)}{\mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)} - 1 \right) \\
&\qquad \le \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) \cdot \left((1+\varepsilon)^{2\cdot|S|} - 1 \right)
\end{aligned}$$

Since the reward function is bounded by 1, it follows that $\mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) \le 1$, and hence we have

$$|\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) - \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)| \le (1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1.$$

The desired result follows. ∎

Theorem 4. *Let $G_1 = (S, \delta)$ and $G_2 = (S, \delta')$ be two structurally equivalent Markov chains. Let η be the minimum positive transition probability in G_1. The following assertions hold:*

1. *For all non-negative reward functions $r : S \to \mathbb{R}$ such that the reward function is bounded by 1, for all discount vectors $\boldsymbol{\lambda}$, for all $s \in S$ we have*

$$\begin{aligned}
|\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) - \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)| &\le (1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1 \\
&\le \left(1 + \frac{dist_A(G_1, G_2)}{\eta - dist_A(G_1, G_2)}\right)^{2\cdot|S|} - 1
\end{aligned}$$

2. *For all parity objectives Φ and for all $s \in S$ we have*

$$\begin{aligned}|\mathsf{Val}(G_1,\Phi)(s) - \mathsf{Val}(G_2,\Phi)(s)| &\leq (1 + dist_R(G_1,G_2))^{2\cdot|S|} - 1 \\ &\leq \left(1 + \frac{dist_A(G_1,G_2)}{\eta - dist_A(G_1,G_2)}\right)^{2\cdot|S|} - 1\end{aligned}$$

Proof. The first part follows from Lemma 3 and Proposition 1. The second part follows from part 1, the fact the value function for parity objectives is obtained as the limit of multi-discounted objectives (Theorem 2), and the fact the bound for part 1 is independent of the discount factors (hence independent of taking the limit). ∎

4 Value Continuity for Parity Objectives

In this section we show two results: first we show robustness of strategies and present quantitative bounds on value function for turn-based stochastic games and then we show the continuity for concurrent parity games.

4.1 Quantitative bounds for structurally equivalent turn-based stochastic parity games

In this section we present quantitative bounds for robustness of optimal strategies in structurally equivalent turn-based stochastic games. For every $\varepsilon > 0$ we present a bound $\beta > 0$ such that if the distance of the structurally equivalent turn-based stochastic games differ by at most β then any pure memoryless optimal strategy in one game is ε-optimal in the other. We first show the result for MDPs and then extend to turn-based stochastic games.

Theorem 5. *Let G_1 be a player-1 MDP such that the minimum positive transition probability is $\eta > 0$. The following assertions hold:*

1. *For all player-1 MDPs $G_2 \in [\![G_1]\!]_\equiv$, for all parity objectives Φ and for all $s \in S$ we have*

$$|\mathsf{Val}(G_1,\Phi)(s) - \mathsf{Val}(G_2,\Phi)(s)| \leq (1 + dist_R(G_1,G_2))^{2\cdot|S|} - 1 \leq \left(1 + \frac{dist_A(G_1,G_2)}{\eta - dist_A(G_1,G_2)}\right)^{2\cdot|S|} - 1$$

2. *For $\varepsilon > 0$, let $\beta \leq \frac{\eta}{2} \cdot \left((1 + \frac{\varepsilon}{2})^{\frac{1}{2\cdot|S|}} - 1\right)$ such that $\beta \leq \frac{\eta}{2}$. For all $G_2 \in [\![G_1]\!]_\equiv$ such that $dist_A(G_1,G_2) \leq \beta$, for all parity objectives Φ, every pure memoryless optimal strategy π_1 in G_1 is an ε-optimal strategy in G_2. In other words, for the interval $[0,\beta)$, every pure memoryless optimal strategy in G_1 is an ε-optimal strategy in all structurally equivalent MDPs of G_1 such that the distance lie in the interval $[0,\beta)$.*

Proof. We prove the two parts below.

1. Without loss of generality, let $\mathsf{Val}(G_1,\Phi)(s) \geq \mathsf{Val}(G_2,\Phi)(s)$. Let π_1 be a pure memoryless optimal strategy in G_1 and such a strategy exists by Theorem 1. Then we have the following inequality

$$\begin{aligned}\mathsf{Val}(G_2,\Phi)(s) &\geq \mathsf{Val}(G_2 \upharpoonright \pi_1,\Phi)(s) \\ &\geq \mathsf{Val}(G_1 \upharpoonright \pi_1,\Phi)(s) - \left((1 + dist_R(G_1,G_2))^{2\cdot|S|} - 1\right) \\ &= \mathsf{Val}(G_1,\Phi)(s) - \left((1 + dist_R(G_1,G_2))^{2\cdot|S|} - 1\right)\end{aligned}$$

The (in)equalities are obtained: the first inequality follows because the value in G_2 is at least the value in G_2 obtained by fixing a particular strategy (in this case π_1); the second inequality is obtained by appying Theorem 4 on the structurally equivalent Markov chains $G_1 \upharpoonright \pi_1$ and $G_2 \upharpoonright \pi_1$; and the final equality follows since π_1 is an optimal strategy in G_1. The desired result follows.

2. Let $G_2 \in [\![G_1]\!]_{\equiv}$ such that $dist_A(G_1, G_2) \leq \beta$. Let π_1 be any pure memoryless optimal strategy in G_1. Then we have the following inequality

$$\begin{aligned}\mathsf{Val}(G_2 \restriction \pi_1, \Phi)(s) &\geq \mathsf{Val}(G_1 \restriction \pi_1, \Phi)(s) - \left((1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1\right) \\ &= \mathsf{Val}(G_1, \Phi)(s) - \left((1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1\right) \\ &\geq \mathsf{Val}(G_2, \Phi)(s) - 2 \cdot \left((1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1\right).\end{aligned}$$

The first inequality is a consequence of Theorem 4 applied on Markov chains $G_2 \restriction \pi_1$ and $G_1 \restriction \pi_1$; the equality follows from the fact π_1 is an optimal strategy in G_1; and the final equality follows by applying the result of part 1. Hence to prove that π_1 is ε-optimal in G_2 we need to show that

$$2 \cdot \left((1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1\right) \leq \varepsilon \tag{2}$$

We have

$$(1 + dist_R(G_1, G_2)) \leq \left(1 + \frac{dist_A(G_1, G_2)}{\eta - dist_A(G_1, G_2)}\right) \leq \left(1 + \frac{2 \cdot dist_A(G_1, G_2)}{\eta}\right);$$

the first inequality follows from Proposition 1 and the second inequality follows since $dist_A(G_1, G_2) \leq \beta \leq \frac{\eta}{2}$. Hence to prove inequality (2) it suffices to show that

$$\left(1 + \frac{2 \cdot \beta}{\eta}\right)^{2\cdot|S|} \leq 1 + \frac{\varepsilon}{2}.$$

Since $\beta \leq \frac{\eta}{2} \cdot \left((1 + \frac{\varepsilon}{2})^{\frac{1}{2\cdot|S|}} - 1\right)$, we obtain the desired inequality.

The desired result follows. ∎

Theorem 6. *Let G_1 be a turn-based stochastic game such that the minimum positive transition probability is $\eta > 0$. The following assertions hold:*

1. *For all turn-based stochastic games $G_2 \in [\![G_1]\!]_{\equiv}$, for all parity objectives Φ and for all $s \in S$ we have*

$$|\mathsf{Val}(G_1, \Phi)(s) - \mathsf{Val}(G_2, \Phi)(s)| \leq (1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1 \leq \left(1 + \frac{dist_A(G_1, G_2)}{\eta - dist_A(G_1, G_2)}\right)^{2\cdot|S|} - 1$$

2. *For $\varepsilon > 0$, let $\beta \leq \frac{\eta}{2} \cdot \left((1 + \frac{\varepsilon}{2})^{\frac{1}{2\cdot|S|}} - 1\right)$, such that $\beta \leq \frac{\eta}{2}$. For all $G_2 \in [\![G_1]\!]_{\equiv}$ such that $dist_A(G_1, G_2) \leq \beta$, for all parity objectives Φ, every pure memoryless optimal strategy π_1 in G_1 is an ε-optimal strategy in G_2.*

Proof. The proof is essentially to repeat the proof of Theorem 5: as in MDPs pure memoryless optimal strategies exist in turn-based stochastic games with parity objectives (Theorem 1); and once a pure memoryless strategy is fixed in a turn-based stochastic game we obtain an MDP. Since Theorem 5 extend the result of Theorem 4 from Markov chains to MDPs, the proof for the desired result follows by mimicing the proof of Theorem 5 and instead of using the result of Theorem 4 for Markov chains using the result of Theorem 5 for MDPs. ∎

4.2 Value continuity for concurrent parity games

In this section we show value continuity for structurally equivalent concurrent parity games, and show with an example on Markov chains that the continuity property breaks without the structurally equivalent assumption. Finally with an example on Markov chains we show the our quantitative bounds are asymptotically optimal for small distance values. We start with a lemma for MDPs.

Lemma 4. *Let G_1 and G_2 be two structurally equivalent MDPs. Let η be the minimum positive transition probability in G_1. For all non-negative reward functions $r : S \to \mathbb{R}$ such that the reward function is bounded by 1, for all discount vectors $\boldsymbol{\lambda}$, for all $s \in S$ we have*

$$
\begin{aligned}
|\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) - \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)| &\leq (1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1 \\
&\leq \left(1 + \frac{dist_A(G_1, G_2)}{\eta - dist_A(G_1, G_2)}\right)^{2\cdot|S|} - 1
\end{aligned}
$$

Proof. The proof is essentially mimicing the proof of part(1) of Theorem 5. Without loss of generality, let $\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) \geq \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)$. Let π_1 be a pure memoryless optimal strategy in G_1 and such a strategy exists by Theorem 1. Then we have the following inequality

$$
\begin{aligned}
\mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) &\geq \mathsf{Val}(G_2 \upharpoonright \pi_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) \\
&\geq \mathsf{Val}(G_1 \upharpoonright \pi_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) - \left((1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1\right) \\
&= \mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) - \left((1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1\right)
\end{aligned}
$$

The (in)equalities are obtained: the first inequality follows because the value in G_2 is at least the value in G_2 obtained by fixing a particular strategy (in this case π_1); the second inequality is obtained by appying Theorem 4 on the structurally equivalent Markov chains $G_1 \upharpoonright \pi_1$ and $G_2 \upharpoonright \pi_1$; and the final equality follows since π_1 is an optimal strategy in G_1. The desired result follows. ∎

Lemma 5. *Let G_1 and G_2 be two structurally equivalent concurrent game structures. Let η be the minimum positive transition probability in G_1. For all non-negative reward functions $r : S \to \mathbb{R}$ such that the reward function is bounded by 1, for all discount vectors $\boldsymbol{\lambda}$, for all $s \in S$ we have*

$$
\begin{aligned}
|\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) - \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)| &\leq (1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1 \\
&\leq \left(1 + \frac{dist_A(G_1, G_2)}{\eta - dist_A(G_1, G_2)}\right)^{2\cdot|S|} - 1
\end{aligned}
$$

Proof. The proof is essentially mimicing the proof of Lemma 4. Without loss of generality, let $\mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) \geq \mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s)$. Let π_1 be a randomized memoryless optimal strategy in G_1 and such a strategy exists by Theorem 1. Then we have the following inequality

$$
\begin{aligned}
\mathsf{Val}(G_2, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) &\geq \mathsf{Val}(G_2 \upharpoonright \pi_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) \\
&\geq \mathsf{Val}(G_1 \upharpoonright \pi_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) - \left((1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1\right) \\
&= \mathsf{Val}(G_1, \mathsf{MDT}(\boldsymbol{\lambda}, r))(s) - \left((1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1\right)
\end{aligned}
$$

The argument for the inequalities are exactly the same as in Lemma 4. The desired result follows. ∎

Theorem 7. *Let G_1 and G_2 be two structurally equivalent concurrent game structures. Let η be the minimum positive transition probability in G_1. For all parity objectives Φ and for all $s \in S$ we have*

$$
\begin{aligned}
|\mathsf{Val}(G_1, \Phi)(s) - \mathsf{Val}(G_2, \Phi)(s)| &\leq (1 + dist_R(G_1, G_2))^{2\cdot|S|} - 1 \\
&\leq \left(1 + \frac{dist_A(G_1, G_2)}{\eta - dist_A(G_1, G_2)}\right)^{2\cdot|S|} - 1
\end{aligned}
$$

Proof. The result follows from Theorem 2, Lemma 5 and the fact that the bound of Lemma 5 are independent of the discount factors and hence independent of taking the limits. ∎

Theorem 8. *For all concurrent game structures G_1, for all parity objectives Φ*

$$
\lim_{\varepsilon \to 0} \sup_{G_2 \in [\![G_1]\!]_{\equiv}, dist_A(G_1, G_2) \leq \varepsilon} \sup_{s \in S} |\mathsf{Val}(G_1, \Phi)(s) - \mathsf{Val}(G_2, \Phi)(s)| = 0.
$$

Fig. 1. Markov chains G_1 and G_2^{ε}.

Proof. Let $\eta > 0$ be the minimum positive transition probability in G_1. Then by Theorem 7 we have

$$\lim_{\varepsilon\to 0} \sup_{G_2 \in [\![G_1]\!]_{\equiv}, dist_A(G_1,G_2)\le\varepsilon} \sup_{s\in S} |\mathsf{Val}(G_1,\Phi)(s) - \mathsf{Val}(G_2,\Phi)(s)| \le \lim_{\varepsilon\to 0}\left(1+\frac{\varepsilon}{\eta-\varepsilon}\right)^{2\cdot|S|} - 1 = 0.$$

The desired result follows. ∎

Example 1 (Structurally equivalence assumption necessary). In this example we show that in Theorem 8 the structural equivalence assumption is necessary, and there by showing that the result is tight. We show an Markov chain G_1 and a family of Markov chains G_2^{ε}, for $\varepsilon > 0$, such that $dist_A(G_1, G_2^{\varepsilon}) \le \varepsilon$ (but G_1 is not structurally equivalent to G_2^{ε}) with a parity objective Φ and we have

$$\lim_{\varepsilon\to 0}\sup_{s\in S} |\mathsf{Val}(G_1,\Phi)(s) - \mathsf{Val}(G_2^{\varepsilon},\Phi)(s)| = 1.$$

The Markov chains G_1 and G_2^{ε} are defined over the state space $\{s_0, s_1\}$, and in G_1 both states have self-loops with probability 1, and in G_2^{ε} the self-loop at s_0 has probability $1-\varepsilon$ and the transition probability from s_0 to s_1 is ε (see Fig 1). Clearly, $dist_A(G_1, G_2^{\varepsilon}) = \varepsilon$. The parity objective Φ requires to visit the state s_1 infinitely often (i.e., assign priority 2 to s_1 and priority 1 to s_0). Then we have $\mathsf{Val}(G_1,\Phi)(s_0) = 0$ as the state s_0 is never left, whereas in G_2^{ε} the state s_1 is the only closed recurrent set of the Markov chain and hence reached with probability 1 from s_0. Hence $\mathsf{Val}(G_2^{\varepsilon},\Phi)(s_0) = 1$. It follows that $\lim_{\varepsilon\to 0}\sup_{s\in S} |\mathsf{Val}(G_1,\Phi)(s) - \mathsf{Val}(G_2^{\varepsilon},\Phi)(s)| = 1$. ∎

Example 2 (Asymptotically tight bound for small distances). We now show that the our quantitative bound for the value function difference is asymptotically optimal for small distances. Let us denote the absolute distance as ε, and quantitative bound we obtain in Theorem 7 is $(1+\frac{\varepsilon}{\eta-\varepsilon})^{2\cdot|S|}-1$, and if ε is small ($\varepsilon << \eta$ and ε close to zero), we obtain the following approximate bound

$$(1+\frac{\varepsilon}{\eta-\varepsilon})^{2\cdot|S|} - 1 \approx (1+\frac{\varepsilon}{\eta})^{2\cdot|S|} - 1 \approx 1 + 2\cdot|S|\cdot\frac{\varepsilon}{\eta} - 1 = 2\cdot|S|\cdot\frac{\varepsilon}{\eta}.$$

We now illustrate with an example (on structurally equivalent Markov chains) where the difference in the value function is $O(|S|\cdot\varepsilon)$, for small ε. Consider the Markov chain defined on state space $S = \{s_0, s_1, \ldots, s_{2n-1}, s_{2n}\}$ as follows: states s_0 and s_{2n} are absorbing (states with self-loops of probability 1) and for a state $1 \le i \le 2n-1$ we have

$$\delta(s_i)(s_{i-1}) = \frac{1}{2}+\varepsilon; \quad \delta(s_i)(s_{i+1}) = \frac{1}{2}-\varepsilon;$$

i.e., we have a Markov chain defined on a line from 0 to $2n$ (with 0 and $2n$ absorbing states) and the chain moves towards 0 with probability $\frac{1}{2}+\varepsilon$ and towards $2n$ with probability $\frac{1}{2}-\varepsilon$ (see Fig 2). Our goal is to estimate the probability to reach the state s_0, and let v_i denote the probability to reach s_0 from the starting state s_i. Then we have the following simple recurrence for $1 \le i \le 2n-1$

$$v_i = (\frac{1}{2}+\varepsilon)\cdot v_{i-1} + (\frac{1}{2}-\varepsilon)\cdot v_{i+1};$$

Fig. 2. Markov chains for Example 2.

and $v_0 = 1$ and $v_{2n} = 0$. We will consider $\varepsilon \geq 0$ such that ε is very small and hence higher order terms (like ε^2) can be ignored. We claim that the values v_i can be expressed as the following recurrence: $v_{i+1} = (\frac{1}{2} + \varepsilon) \cdot c_i \cdot v_i$, where $c_i = \frac{4}{4-c_{i+1}}$. The proof is by induction and is shown below:

$$\begin{aligned} v_i &= (\tfrac{1}{2} + \varepsilon) \cdot v_{i-1} + (\tfrac{1}{2} - \varepsilon) \cdot v_{i+1} \\ &= (\tfrac{1}{2} + \varepsilon) \cdot v_{i-1} + (\tfrac{1}{2} - \varepsilon) \cdot (\tfrac{1}{2} + \varepsilon) \cdot c_i \cdot v_i \qquad \text{(by inductive hypothesis)} \\ &= (\tfrac{1}{2} + \varepsilon) \cdot v_{i-1} + \tfrac{1}{4} \cdot c_i \cdot v_i \qquad \text{(ignoring } \varepsilon^2\text{)} \end{aligned}$$

It follows that $v_i = (\frac{1}{2} + \varepsilon) \cdot \frac{4}{4-c_i} \cdot v_{i-1} = (\frac{1}{2} + \varepsilon) \cdot c_{i-1} \cdot v_{i-1}$. Hence we have

$$v_1 = (\frac{1}{2} + \varepsilon) \cdot 1 + (\frac{1}{2} - \varepsilon) \cdot (\frac{1}{2} - \varepsilon) \cdot c_1 \cdot v_1 \Rightarrow v_1 = \frac{4}{4 - c_1} \cdot (\frac{1}{2} + \varepsilon).$$

Then we have $v_2 = (\frac{1}{2} + \varepsilon) \cdot c_1 \cdot v_1 = \frac{4}{4-c_1} \cdot c_1 \cdot (\frac{1}{2} + \varepsilon)^2$ and then $v_3 = \frac{4}{4-c_1} \cdot c_1 \cdot c_2 \cdot (\frac{1}{2} + \varepsilon)^3$ and so on. Finally we have obtain v_n as follows: $v_n = \frac{4}{4-c_1} \cdot c_1 \cdot c_2 \cdots c_{n-1} \cdot (\frac{1}{2} + \varepsilon)^n$. Observe that for the Markov chain with $\varepsilon = 0$, the states s_0 and s_{2n} are the recurrent states, and since the chain is symmetric from s_n (with $\varepsilon = 0$) the probability to reach s_{2n} and s_0 must be equal and hence is $\frac{1}{2}$. It follows that we must have $\frac{4}{4-c_1} \cdot c_1 \cdot c_2 \cdots c_{n-1} = 2^{n-1}$. Hence we have that for $\varepsilon > 0$, but very small, $v_n \approx \frac{1}{2} + n \cdot \varepsilon$. Thus the difference with the value function when $\varepsilon = 0$ as compared to when $\varepsilon > 0$ but very small is $n \cdot \varepsilon = O(|S| \cdot \varepsilon)$. Also observe that the Markov chain obtained for $\varepsilon = 0$ and $\frac{1}{2} > \varepsilon > 0$ are structurally equivalent. Thus the desired result follows. ∎

5 Conclusion

In this work we studied the robustness and continuity property of concurrent and turn-based stochastic parity games with respect to small imprecision in the transition probabilities. We presented quantitative bounds of difference of value function and proved value continuity for concurrent parity games under the structural equivalence assumption, and showed robustness of all pure memoryless optimal strategies for structurally equivalent turn-based stochastic parity games. We also showed that the structural equivalence assumption is necessary and that our quantitative bounds are asymptotically optimal for small imprecision. We believe our results will find applications in robustness analysis of various other classes of stochastic games.

References

1. M. Abadi, L. Lamport, and P. Wolper. Realizable and unrealizable specifications of reactive systems. In *ICALP'89*, LNCS 372, pages 1–17. Springer, 1989.
2. R. Alur, T.A. Henzinger, and O. Kupferman. Alternating-time temporal logic. *Journal of the ACM*, 49:672–713, 2002.
3. K. Chatterjee, L. de Alfaro, and T.A. Henzinger. The complexity of quantitative concurrent parity games. In *SODA'06*, pages 678–687. ACM-SIAM, 2004.

4. K. Chatterjee, T.A. Henzinger, and M. Jurdziński. Games with secure equilibria. In *LICS'04*, pages 160–169. IEEE, 2004.
5. K. Chatterjee, M. Jurdziński, and T.A. Henzinger. Quantitative stochastic parity games. In *SODA'04*, pages 121–130. SIAM, 2004.
6. A. Church. Logic, arithmetic, and automata. In *Proceedings of the International Congress of Mathematicians*, pages 23–35. Institut Mittag-Leffler, 1962.
7. A. Condon. The complexity of stochastic games. *Information and Computation*, 96(2):203–224, 1992.
8. C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *Journal of the ACM*, 42(4):857–907, 1995.
9. L. de Alfaro. *Formal Verification of Probabilistic Systems*. PhD thesis, Stanford University, 1997.
10. L. de Alfaro and T.A. Henzinger. Concurrent omega-regular games. In *LICS'00*, pages 141–154. IEEE, 2000.
11. L. de Alfaro, T.A. Henzinger, and R. Majumdar. Discounting the future in systems theory. In *ICALP'03*, LNCS 2719, pages 1022–1037. Springer, 2003.
12. L. de Alfaro and R. Majumdar. Quantitative solution of omega-regular games. In *STOC'01*, pages 675–683. ACM Press, 2001.
13. D.L. Dill. *Trace Theory for Automatic Hierarchical Verification of Speed-independent Circuits*. The MIT Press, 1989.
14. K. Etessami and M. Yannakakis. Recursive concurrent stochastic games. In *ICALP'06 (2)*, LNCS 4052, Springer, pages 324–335, 2006.
15. J. Filar and K. Vrieze. *Competitive Markov Decision Processes*. Springer-Verlag, 1997.
16. M. I. Friedlin and A. D. Wentzell. *Random perturbations of dynamical systems*. Springer, 1984.
17. H. Gimbert and W. Zielonka. Discounting infinite games but how and why? *Electr. Notes Theor. Comput. Sci.*, 119(1):3–9, 2005.
18. A. Kechris. *Classical Descriptive Set Theory*. Springer, 1995.
19. D.A. Martin. The determinacy of Blackwell games. *The Journal of Symbolic Logic*, 63(4):1565–1581, 1998.
20. E. Milman. The semi-algebraic theory of stochastic games. *Math. Oper. Research*, pages 401–418, 2002.
21. A. Pnueli and R. Rosner. On the synthesis of a reactive module. In *POPL'89*, pages 179–190. ACM Press, 1989.
22. T.E.S. Raghavan and J.A. Filar. Algorithms for stochastic games — a survey. *ZOR — Methods and Models of Op. Res.*, 35:437–472, 1991.
23. P.J. Ramadge and W.M. Wonham. Supervisory control of a class of discrete-event processes. *SIAM Journal of Control and Optimization*, 25(1):206–230, 1987.
24. L.S. Shapley. Stochastic games. *Proc. Nat. Acad. Sci. USA*, 39:1095–1100, 1953.
25. E. Solan. Continuity of the value of competitive Markov decision processes. *Journal of Theoretical Probability*, 16:831–845, 2003.
26. W. Thomas. Automata on infinite objects. In J. van Leeuwen, editor, *Handbook of Theoretical Computer Science*, volume B, pages 133–191. Elsevier, 1990.
27. M.Y. Vardi. Automatic verification of probabilistic concurrent finite-state systems. In *FOCS'85*, pages 327–338. IEEE Computer Society Press, 1985.
28. W.Zielonka. Perfect-information stochastic parity games. In *FoSSaCS'04*, pages 499–513. LNCS 2987, Springer, 2004.