A reduction from parity games to simple stochastic games [∗]

Krishnendu Chatterjee

IST Austria (Institute of Science and Technology, Austria)

`krishnendu.chatterjee@ist.ac.at`

Nathanaël Fijalkow

IST Austria (Institute of Science and Technology, Austria)

ÉNS Cachan (École Normale Supérieure de Cachan, France)

`nathanael.fijalkow@gmail.com`

Games on graphs provide a natural model for reactive non-terminating systems. In such games, the interaction of two players on an arena results in an infinite path that describes a run of the system. Different settings are used to model various open systems in computer science, as for instance turn-based or concurrent moves, and deterministic or stochastic transitions. In this paper, we are interested in turn-based games, and specifically in deterministic parity games and stochastic reachability games (also known as simple stochastic games). We present a simple, direct and efficient reduction from deterministic parity games to simple stochastic games: it yields an arena whose size is linear up to a logarithmic factor in size of the original arena.

Keywords. *Stochastic games, parity objectives, reachability objectives.*

1 Introduction

Graph games. Graph games are used to model reactive systems. A finite directed graph, whose vertices represent states and edges represent transitions, models the system. Its evolution consists in interactions between a controller and the environment, which is naturally turned into a game on the graph between two players, Eve and Adam. In the turn-based setting, in each state of the system, either the controller chooses the evolution of the system (the corresponding vertex is then controlled by Eve), or the system evolves in an uncertain way, then aiming at the worst-case scenario Adam controls the corresponding vertex. This defines a 2-player arena as a finite directed graph and a partition of the vertex set into Eve and Adam vertices. However, in many applications, systems are randomized, leading to the definition of stochastic arenas: in addition to Eve and Adam vertices, the graph also has random vertices where the evolution is chosen according to a given probability distribution.

A pebble is initially placed on the vertex representing the initial state of the system, then Eve, Adam and random move this pebble along the edges, constructing an infinite sequence of vertices. The sequence built describes a run of the system: Eve tries to ensure that it satisfies some specification of the system, while Adam tries to spoil it.

Parity objectives. The theory of graph games with ω-regular winning conditions is the foundation for modelling and synthesizing reactive processes with fairness constraints. The *parity* objectives provide an adequate model, as the fairness constraints of reactive processes are ω-regular, and every ω-regular

---

winning condition can be specified as a parity objective [8]. We consider 2-player games with parity objectives: deciding the winner in polynomial time is a longstanding open question, despite many efforts from a large community. The best known upper-bound is UP∩coUP [6].

Simple stochastic games. Considering probabilistic games instead of deterministic allows the description much more reactive systems by modelling uncertainty, but leads to higher complexity for corresponding decision problems. We consider stochastic games with reachability objectives: a given vertex is distinguished, and Eve tries to reach it. Those games were introduced by Condon, and named simple stochastic games [4]. We consider the following decision problem: can Eve ensure to reach the target vertex with probability more than half? As for the above decision problem, the best known upper-bound is NP∩coNP [4].

Reduction: from parity games to simple stochastic games. The notion of reduction between games is an important aspect in the study of games as it allows to understand which classes of games are subsumed by others. A classical reduction of 2-player parity games to simple stochastic games is through a sequence of three reductions: (a) from 2-player parity games to 2-player mean-payoff (or limit-average) games [6]; (b) from 2-player mean-payoff games to 2-player discounted-payoff games [9]; and (c) from 2-player discounted-payoff games to stochastic reachability games [9]. The sequence of reductions yields the following result: given a 2-player parity game with n vertices, m edges, and a parity objective with d priorities, the simple stochastic game obtained through the sequence of reductions has $n + m$ vertices, including m probabilistic ones, $4 \cdot m$ edges and the size of the arena is $O(m \cdot d \cdot \log(n))$.

Our results: we present a direct reduction of 2-player parity games to simple stochastic games, and thus show that one can discount the step of going through mean-payoff and discounted games. Moreover, our reduction is more efficient: given a 2-player parity game with n vertices, m edges, and a parity objective with d priorities, the simple stochastic game obtained by our direct reduction has $n + m$ vertices among which m are probabilistic, $3 \cdot m$ edges and the size of the arena is $O(m \cdot \log(n))$. Finally, we conclude following proof ideas from [5] that the decision problem for simple stochastic games is in UP∩coUP, and from [2, 3] we obtain that the decision problems in stochastic parity, mean-payoff and discounted games all are in UP∩coUP.

2   Definitions

Given a finite set A, a probability distribution μ on A is a function $\mu : A \to [0, 1]$ such that $\sum_{a \in A} \mu(a) = 1$. We denote by $\mathscr{D}(A)$ the set of all probability distributions on A.

Stochastic arena. A stochastic (or $2\frac{1}{2}$-player) arena $G = ((V, E), (V_E, V_A, V_R), \delta)$ consists of a finite directed graph (V, E) with vertex set V and edge set E, a partition (V_E, V_A, V_R) of the vertex set V and a probabilistic transition function $\delta : V_R \to \mathscr{D}(V)$ that given a vertex in V_R gives the probability of transition to the next vertex. Eve chooses the successor of vertices in V_E, while Adam chooses the successor of vertices in V_A; vertices in V_R are random vertices and their successor is chosen according to δ. We assume that for all $u \in V_R$ and $v \in V$ we have $(u, v) \in E$ if and only if $\delta(u)(v) > 0$. We assume that the underlying graph has no deadlock: every vertex has a successor. The special case where $V_R = \emptyset$ corresponds to 2-player arenas (for those we omit δ from the description of the arena).

Size of an arena. The size of a stochastic arena $G = ((V, E), (V_E, V_A, V_R), \delta)$ is the number of bits required to store it:

$$\text{size}(G) = \underbrace{\log(n)}_{\text{vertices}} + \underbrace{2 \cdot m \cdot \log(n)}_{\text{edges}} + \underbrace{n + n_R \cdot \log(n)}_{\text{vertex partition}} + \underbrace{\text{size}(\delta)}_{\text{probabilistic transitions}}$$

where $n = |V|$, $m = |E|$, $n_R = |V_R|$ and $\text{size}(\delta) = \sum_{u \in V_R} \sum_{v \in V} |\delta(u)(v)|$, where $|\delta(u)(v)|$ is the length of the binary representation of $\delta(u)(v)$.

Plays and strategies. A *play* π in a stochastic arena G is an infinite sequence $\langle v_0, v_1, v_2, \ldots \rangle$ of vertices such that for all $i \geq 0$ we have $(v_i, v_{i+1}) \in E$. We denote by Π the set of all plays. A *strategy* for a player is a recipe that prescribes how to play, *i.e*, given a finite history of play, a strategy defines the next move. Formally, a strategy for Eve is a function $\sigma : V^* \cdot V_E \to V$ such that for all $w \in V^*$ and $v \in V_E$ we have $(v, \sigma(w \cdot v)) \in E$. We define strategies for Adam analogously, and denote by Σ and Γ the set of all strategies for Eve and Adam, respectively. A strategy is *memoryless*, or *positional* if it is independent of the history of play and only depends on the current vertex, *i.e*, for all $w, w' \in V^*$ and $v \in V_E$ we have $\sigma(w \cdot v) = \sigma(w' \cdot v)$. Hence a memoryless strategy can be described as a function $\sigma : V_E \to V$.

Once a starting vertex $v \in V$ and strategies σ for Eve and τ for Adam are fixed, the outcome of the game is a random walk $\pi(v, \sigma, \tau)$ for which the probabilities of events are uniquely defined, where an *event* $\mathscr{A} \subseteq \Pi$ is a measurable set of plays. For an event $\mathscr{A} \subseteq \Pi$, we write $\mathbb{P}_v^{\sigma, \tau}(\mathscr{A})$ for the probability that a play belongs to $\mathscr{A}$ if the game starts from the vertex v and the players follow the strategies σ and τ. In case of 2-player arenas, if we fix positional strategies σ, τ, and a starting vertex v, then the play $\pi(v, \sigma, \tau)$ obtained is unique and consists in a simple path $\langle v_0, v_1, \ldots v_{l-1} \rangle$ and a cycle $\langle v_l, v_{l+1}, \ldots, v_k \rangle$ executed infinitely often, *i.e*, the play is a "lasso-play": $\langle v_0, v_1, \ldots, v_{l-1} \rangle \cdot \langle v_l, v_{l+1}, \ldots, v_k \rangle^{\omega}$.

Qualitative objectives. We specify *qualitative* objectives for the players by providing a set of *winning* plays $\Phi \subseteq \Pi$ for each player. We say that a play π *satisfies* the objective Φ if $\pi \in \Phi$. We study only zero-sum games, where the objectives of the two players are complementary, *i.e*, if Eve has the objective Φ, then Adam has the objective $\Pi \setminus \Phi$.

- *Reachability objectives.* Given a set $T \subseteq V$ of "target" vertices, the reachability objective requires that some vertex of T be visited. The set of winning plays is $\text{Reach}(T) = \{ \langle v_0, v_1, v_2, \ldots \rangle \in \Pi \mid v_k \in T \text{ for some } k \geq 0 \}$.

- *Parity objectives.* Let $p : V \to \mathbb{N}$ be a function that assigns a *priority* $p(v)$ to every vertex $v \in V$. For a play $\pi = \langle v_0, v_1, \ldots \rangle \in \Pi$, we define $\text{Inf}(\pi) = \{ v \in V \mid v_k = v \text{ for infinitely many } k \}$ to be the set of vertices that occur infinitely often in π. The *parity objective* is defined as $\text{Parity}(p) = \{ \pi \in \Pi \mid \min(p(\text{Inf}(\pi))) \text{ is even} \}$. In other words, the parity objective requires that the minimum priority visited infinitely often is even.

Quantitative objectives. A *quantitative* objective is specified as a measurable function $f : \Pi \to \mathbb{R}$. In zero-sum games the objectives of the players are functions f and $-f$, respectively. We consider two classes of quantitative objectives, namely, mean-payoff and discounted-payoff objectives.

- *Mean-payoff objectives.* Let $r : V \to \mathbb{R}$ be a real-valued reward function that assigns to every vertex v the reward $r(v)$. The *mean-payoff* objective $\text{MeanPayoff}(r)$ assigns to every play the "long-run" average of the rewards appearing in the play. Formally, for a play $\pi = \langle v_0, v_1, v_2, \ldots \rangle$ we have

$$\text{MeanPayoff}(r)(\pi) = \liminf_{n \to \infty} \frac{1}{n+1} \cdot \sum_{i=0}^{n} r(v_i)$$

- *Discounted-payoff objectives.* Let $r : V \to \mathbb{R}$ be a reward function and $0 < \lambda < 1$ be a discount factor, the discounted-payoff objective $\text{DiscPayoff}(\lambda, r)$ assigns to every play the discounted sum of the rewards in the play. Formally, for a play $\pi = \langle v_0, v_1, v_2, \ldots \rangle$ we have

$$\text{DiscPayoff}(\lambda, r)(\pi) = (1 - \lambda) \cdot \lim_{n \to \infty} \sum_{i=0}^{n} \lambda^i \cdot r(v_i)$$

Values and optimal strategies. Given objectives $\Phi \subseteq \Pi$ for Eve and $\Pi \setminus \Phi$ for Adam, and measurable functions f and $-f$ for Eve and Adam, respectively, we define the *value* functions $\langle E \rangle$ and $\langle A \rangle$ for Eve and Adam, respectively, as the following functions from the vertex space V to the set $\mathbb{R}$ of reals: for all vertices $v \in V$, let

$$\langle E \rangle(\Phi)(v) = \sup_{\sigma \in \Sigma} \inf_{\tau \in \Gamma} \mathbb{P}_v^{\sigma, \tau}(\Phi); \qquad\qquad \langle E \rangle(f)(v) = \sup_{\sigma \in \Sigma} \inf_{\tau \in \Gamma} \mathbb{E}_v^{\sigma, \tau}[f];$$

$$\langle A \rangle(\Pi \setminus \Phi)(v) = \sup_{\tau \in \Gamma} \inf_{\sigma \in \Sigma} \mathbb{P}_v^{\sigma, \tau}(\Pi \setminus \Phi); \qquad \langle A \rangle(-f)(v) = \sup_{\tau \in \Gamma} \inf_{\sigma \in \Sigma} \mathbb{E}_v^{\sigma, \tau}[-f].$$

In other words, the values $\langle E \rangle(\Phi)(v)$ and $\langle E \rangle(f)(v)$ give the maximal probability and expectation with which Eve can achieve her objectives Φ and f from vertex v, and analogously for Adam. The strategies that achieve those values are called optimal: a strategy σ for Eve is *optimal* from the vertex v for the objective Φ if $\langle E \rangle(\Phi)(v) = \inf_{\tau \in \Gamma} \mathbb{P}_v^{\sigma, \tau}(\Phi)$; and σ is *optimal* from the vertex v for f if $\langle E \rangle(f)(v) = \inf_{\tau \in \Gamma} \mathbb{E}_v^{\sigma, \tau}[f]$. The optimal strategies for Adam are defined analogously.

Theorem 1 (Memoryless determinacy [9]). *For all stochastic arenas,*

1. *For all objectives Φ such that Φ is either a reachability or a parity objective, for all vertices v we have*

$$\langle E \rangle(\Phi)(v) + \langle A \rangle(\Pi \setminus \Phi)(v) = 1.$$

   Memoryless optimal strategies exist for both players from all vertices. Furthermore, for the case of 2-player arena, then for all $v \in V$, $\langle E \rangle(\Phi)(v) \in \{0, 1\}$.

2. *For all objectives $f : \Pi \to \mathbb{R}$ such that f is either a mean-payoff or discounted-payoff objective, for all vertices v we have*

$$\langle E \rangle(f)(v) + \langle A \rangle(-f)(v) = 0.$$

   Memoryless optimal strategies exist for both players from all vertices.

Games. A stochastic game is given by an arena and an objective. As a special case, a 2-player game is given by a 2-player arena and an objective. For instance, a 2-player parity game is a couple $(G, \mathrm{Parity}(p))$, where G is a 2-player arena, and a stochastic reachability game is a couple $(G, \mathrm{Reach}(T))$, where G is a stochastic arena.

We define simple stochastic games to be special case of stochastic reachability game where V contains two distinguished absorbing vertices v_{win} and v_{lose} and the reachability set is $T = \{v_{\mathrm{win}}\}$. Once a play reached one of the two vertices v_{win} or v_{lose}, the game is stopped as its outcome is fixed. A simple stochastic game has the stopping property if for all strategies σ and τ and all vertices $v \in V$, $\mathbb{P}_v^{\sigma, \tau}(\mathrm{Reach}\{v_{\mathrm{win}}, v_{\mathrm{lose}}\}) = 1$.

Decision problems for games. Given an arena G, an objective Φ, a starting vertex $v \in V$ and a rational threshold $q \in \mathbb{Q}$, the decision problem we consider is whether $\langle E \rangle(\Phi)(v) \geq q$. It follows from Theorem 1 that in 2-player arenas, with a parity objective $\mathrm{Parity}(p)$, for a vertex v we have $\langle E \rangle(\mathrm{Parity}(p))(v) \in \{0, 1\}$. If the value is 1, then we say that Eve is winning, otherwise Adam is winning.

3   A direct reduction

In this section, we present a direct reduction to show that determining the winner in 2-player parity games can be reduced to the decision problem of simple stochastic games with the threshold $\frac{1}{2}$. Specifically,
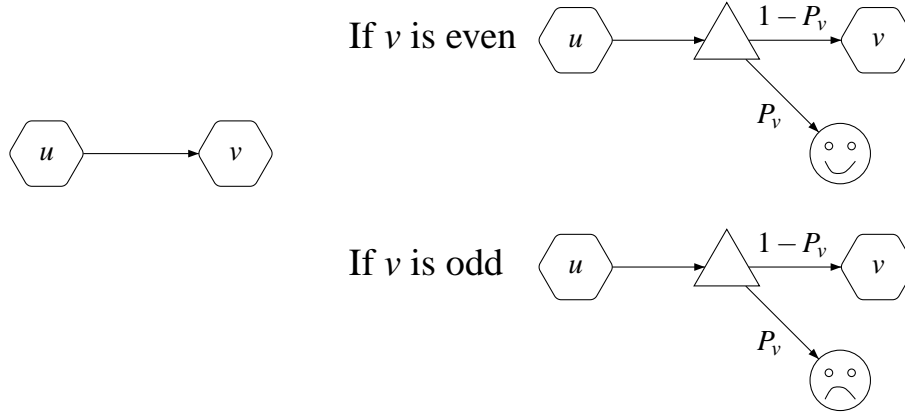
Figure 1: From parity games to simple stochastic games

from a 2-player parity game $(G, \text{Parity}(p))$ and a starting vertex v we show how to construct in polynomial time a stochastic arena $\mathfrak{R}(G)$ with a reachability objective $\text{Reach}(v_{\text{win}})$ such that Eve is winning in G from v for the parity condition if and only if $\langle E \rangle (\text{Reach}(v_{\text{win}}))(v) \geq \frac{1}{2}$ in $\mathfrak{R}(G)$.

Construction of the stochastic arena. We now present the construction of a stochastic arena $\mathfrak{R}(G)$:

$$\mathfrak{R}(G) = ((V \uplus E \uplus \{v_{\text{win}}, v_{\text{lose}}\}, E'), (V_E \uplus \{v_{\text{win}}\}, V_A \uplus \{v_{\text{lose}}\}, E), \delta) \quad ,$$

where $\uplus$ denotes the disjoint union. The set of edges and the transition function is as follows:

$$
\begin{aligned}
E' \;=\; & \{(u, (u,v)); ((u,v), v); ((u,v), v_{\text{win}}) \mid (u,v) \in E, p(v) \text{ even}\} \\
\cup\; & \{(u, (u,v)); ((u,v), v); ((u,v), v_{\text{lose}}) \mid (u,v) \in E, p(v) \text{ odd}\} \\
\cup\; & \{(v_{\text{win}}, v_{\text{win}}); (v_{\text{lose}}, v_{\text{lose}})\}
\end{aligned}
$$

The transition function δ is defined as follows: if $p(v)$ is even, $\delta((u,v))(v_{\text{win}}) = P_v$ and $\delta((u,v))(v) = 1 - P_v$, if $p(v)$ is odd, $\delta((u,v))(v_{\text{lose}}) = P_v$ and $\delta((u,v))(v) = 1 - P_v$. We will describe P_v as reals in the interval $(0,1)$ satisfying certain conditions, and we will prove correctness of the reduction as long as the conditions are satisfied.

We present a pictorial description of the reduction in Figure 1: for each edge $(u,v) \in E$, we consider the simple gadget, if $p(v)$ is even (resp. odd), that has an edge to the sink v_{win} (resp. v_{lose}) with probability P_v, and follows the original edge otherwise with probability $1 - P_v$. Hexagonal vertices can be either Eve's or Adam's and triangle vertices are random vertices. v_{win} will be depicted by a smiling face and v_{lose} by a sad face.

The new arena simulates the initial arena, and additionally features two absorbing vertices v_{win} and v_{lose}. To simulate a transition from u to v, the new arena includes a random vertex that follows the transition with high probability $1 - P_v$ or stops the game by going to v_{win} or v_{lose} with small probability P_v. The intuition is that if v has even priority, then Eve is rewarded to visit it by having a small yet positive chance of winning, and symmetrically if v has odd priority for Adam.

Playing forever in $\mathfrak{R}(G)$, the outcome of the play will be favorable for Eve (*i.e* reach v_{win}) if she manages to see even priorities many times. Furthermore, the reward a player receives for visiting a vertex with good priority must depend on this priority: seeing a very small even priority gives more chance to win than a higher one. Indeed, if different priorities are seen infinitely often, the outcome of
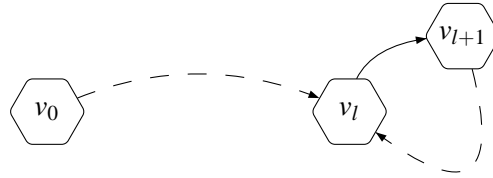
Figure 2: General shape of a play in 2-player game where Eve and Adam play positionally

the play must be in favor of the parity of the lowest priority. This leads to the following assumptions on the probabilities P_v's.

Assumptions on the transition function. We consider the following assumptions on P_v's:

$$\sum_{v \in V} P_v \leq \frac{1}{6} \quad (A_0)$$

and for all $v \in V$, let $J_{\text{odd}}^{>v} = \{u \mid p(u) \text{ odd}, p(u) > p(v)\}$ and $J_{\text{even}}^{>v} = \{u \mid p(u) \text{ even}, p(u) > p(v)\}$:

$$\sum_{u \in J_{\text{odd}}^{>v}} P_u \leq \frac{2}{3} \cdot P_v \qquad (A_1) \qquad\qquad \sum_{u \in J_{\text{even}}^{>v}} P_u \leq \frac{2}{3} \cdot P_v \qquad (A_2)$$

We provide the reader with intuitions on the three assumptions $(A_0) - (A_2)$. The assumption (A_0) ensures that probabilities are small enough such that plays in $\mathfrak{R}(G)$ will last enough to take into account the priorities seen infinitely often, and not only the first ones. The assumptions (A_1) and (A_2) ensure that if v has the lowest priority and is seen infinitely often, no matters how often higher priorities are seen, the outcome will only depend on the priority of v and not on the others.

We present a sequence of properties to prove correctness of the reduction given the three assumptions $(A_0) - (A_2)$ hold.

First note that the set $\{v_{\text{win}}, v_{\text{lose}}\}$ is reached with probability 1, since at each step there is a positive probability to reach it. Another remark is that there is a one-to-one correspondence between strategies in G and $\mathfrak{R}(G)$, so we identify strategies in G or $\mathfrak{R}(G)$.

We will prove that for all $v \in V$, Eve wins from v if and only if $\langle E \rangle(\text{Reach}(v_{\text{win}}))(v) \geq \frac{1}{2}$.

Thanks to Theorem 1, there are memoryless optimal strategies in both games: from now on, we consider σ and τ two memoryless strategies. The *key* property is that the resulting play $\pi = \langle v_0, v_1, \ldots \rangle$ has a simple shape (shown in Figure 2): the play consists in a simple path $\mathcal{P}$ from v_0 to v_l, and then a simple cycle $\mathcal{C}$ is executed forever. Let c be the lowest priority infinitely visited,

$$\pi = \langle v_0, v_1, \ldots, v_{l-1} \rangle \cdot \langle v_l, v_{l+1}, \ldots, v_{l+q-1} \rangle^{\omega}$$

where $p(v_{l+1}) = c$, $\mathcal{P} = \{v_0, v_1, \ldots v_{l-1}\}$ and $\mathcal{C} = \{v_l, v_{l+1}, \ldots, v_{l+q-1}\}$ are pairwise disjoint.

We now consider the corresponding situation in $\mathfrak{R}(G)$: the random walk $\widehat{\pi} = \pi(v_0, \sigma, \tau)$ mimics π until it takes an edge to v_{win} or v_{lose}, stopping the game. We denote by $\widehat{\pi}_i$ the random variable for the ith vertex of $\widehat{\pi}$. Since the starting vertex and the strategies are fixed, we will abbreviate $\mathbb{P}_{v_0}^{\sigma, \tau}$ by $\mathbb{P}$.

We consider the possible different scenarios. There are two possibilities to reach v_{win} or v_{lose}: the first is to reach it during the first l steps, *i.e* during the simple path, the second is to reach it after that, *i.e* during the simple cycle, after the simple path has been crossed.

Notations for events. We define, for $v \in \{v_{\text{win}}, v_{\text{lose}}\}, k, j \geq 0$ the following measurable events.

- The event $\mathrm{Reach}(v, j)$ denotes that v has been reached within j steps, *i.e*, $\mathrm{Reach}(v, j) = \{\pi \mid \widehat{\pi}_j = v\}$ (note that this is equivalent to $\{\pi \mid \exists i \le j, \widehat{\pi}_i = v\}$).

- The event $\mathrm{Cross}(j)$ denotes that neither $\mathrm{v_{win}}$ nor $\mathrm{v_{lose}}$ has been reached within j steps, *i.e*, $\mathrm{Cross}(j) = \{\pi \mid \widehat{\pi}_j \notin \{\mathrm{v_{win}}, \mathrm{v_{lose}}\}\}$.

- The event $\mathrm{ReachPath}(v)$ denotes that v has been reached within l steps, *i.e*, $\mathrm{ReachPath}(v) = \mathrm{Reach}(v, l)$.

- The event $\mathrm{CrossPath}$ denotes that neither $\mathrm{v_{win}}$ nor $\mathrm{v_{lose}}$ has been reached within l steps, *i.e*, $\mathrm{CrossPath} = \mathrm{Cross}(l)$.

- The event $\mathrm{ReachLoop}(v, k)$ denotes that v has been reached within $l + k \cdot q$ steps, *i.e*, $\mathrm{ReachLoop}(v, k) = \mathrm{Reach}(v, l + k \cdot q)$: intuitively, v has been reached either during the path or one of the k first crossings of the loop.

- The event $\mathrm{CrossLoop}(k)$ denotes that neither $\mathrm{v_{win}}$ nor $\mathrm{v_{lose}}$ has been reached within $l + k \cdot q$ steps, *i.e*, $\mathrm{CrossLoop}(k) = \mathrm{Cross}(l + k \cdot q)$: intuitively, during the path and the k first crossings of the loop neither $\mathrm{v_{win}}$ nor $\mathrm{v_{lose}}$ has been reached.

We define the following probabilities:

- $\alpha = \mathbb{P}(\mathrm{CrossPath})$ is the probability to cross the path $\mathscr{P}$;

- $\beta = \mathbb{P}(\mathrm{ReachLoop}(\mathrm{v_{win}}, 1) \mid \mathrm{CrossPath})$ is the probability to reach $\mathrm{v_{win}}$ while following the simple cycle $\mathscr{C}$ for the first time, assuming the path $\mathscr{P}$ was crossed, and similarly

- $\gamma = \mathbb{P}(\mathrm{ReachLoop}(\mathrm{v_{lose}}, 1) \mid \mathrm{CrossPath})$.

We take two steps: the first step is to approximate $\mathbb{P}(\mathrm{Reach}(\mathrm{v_{win}}))$ and $\mathbb{P}(\mathrm{Reach}(\mathrm{v_{lose}}))$ using α, β and γ, and the second is to make use of assumptions $(A_0) - (A_2)$ to evaluate α, β and γ.

Approximations for $\mathbb{P}(\mathrm{Reach}(\mathrm{v_{win}}))$ **and** $\mathbb{P}(\mathrm{Reach}(\mathrm{v_{lose}}))$**.** We rely on the following four properties.

Property 1. *For $k \ge 1$, we have* $\mathbb{P}(\mathrm{ReachLoop}(\mathrm{v_{win}}, k) \mid \mathrm{CrossLoop}(k-1)) = \beta$ *and similarly* $\mathbb{P}(\mathrm{ReachLoop}(\mathrm{v_{lose}}, k) \mid \mathrm{CrossLoop}(k-1)) = \gamma$.

Proof. Since σ and τ are memoryless, the random walk $\widehat{\pi}$ is "memoryless": from v_l, crossing the loop for the first time or for the k-th time will give the same probability to escape to $\mathrm{v_{win}}$ or $\mathrm{v_{lose}}$. ∎

Property 2. *We have, for all $k \ge 1$,* $\mathbb{P}(\mathrm{CrossLoop}(k-1) \mid \mathrm{CrossPath}) = (1 - (\beta + \gamma))^{k-1}$.

Proof. By induction on $k \ge 1$. The case $k = 1$ follows from $\mathrm{CrossLoop}(0) = \mathrm{CrossPath}$. Let $k > 1$:

$$\begin{aligned}
&\mathbb{P}(\mathrm{CrossLoop}(k) \mid \mathrm{CrossPath}) \\
&= \mathbb{P}(\mathrm{CrossLoop}(k) \mid \mathrm{CrossLoop}(k-1)) \cdot \mathbb{P}(\mathrm{CrossLoop}(k-1) \mid \mathrm{CrossPath}) \\
&= (1 - \beta - \gamma) \cdot \mathbb{P}(\mathrm{CrossLoop}(k-1) \mid \mathrm{CrossPath})
\end{aligned}$$

The first equality is a restatement and the second is a result of Property 1. We conclude thanks to the induction hypothesis. ∎

Property 3. *We have* $\mathbb{P}(\mathrm{Reach}(\mathrm{v_{win}}) \mid \mathrm{CrossPath}) = \frac{\beta}{\beta + \gamma}$ *and similarly* $\mathbb{P}(\mathrm{Reach}(\mathrm{v_{lose}}) \mid \mathrm{CrossPath}) = \frac{\gamma}{\beta + \gamma}$.

A simple intuition on this calculation is by referring to a "looping" game. Eve and Adam play a game divided in possibly infinitely many rounds. Each round corresponds to cross the loop once: while doing so, Eve wins with probability β, Adam wins with probability γ and the round is a draw otherwise, with probability $1 - (\beta + \gamma)$. In case of a draw, the game goes on another round. Once a player won, the game is stopped, which corresponds to reach v_{win} or v_{lose}. In this game, Eve wins with probability $\frac{\beta}{\beta+\gamma}$ and Adam with probability $\frac{\gamma}{\beta+\gamma}$.

Proof. We have the following equalities:

$$
\begin{aligned}
\mathbb{P}(\text{Reach}(v_{\text{win}}) \mid \text{CrossPath}) \\
&= \; \sum_{k=1}^{\infty} \mathbb{P}(\text{ReachLoop}(v_{\text{win}}, k) \cap \text{CrossLoop}(k-1) \mid \text{CrossPath}) \\
&= \; \sum_{k=1}^{\infty} \mathbb{P}(\text{ReachLoop}(v_{\text{win}}, k) \mid \text{CrossLoop}(k-1)) \cdot \mathbb{P}(\text{CrossLoop}(k-1) \mid \text{CrossPath}) \\
&= \; \sum_{k=1}^{\infty} \beta \cdot (1 - (\beta + \gamma))^{k-1} \\
&= \; \frac{\beta}{\beta+\gamma}
\end{aligned}
$$

The disjoint union $\text{Reach}(v_{\text{win}}) \cap \text{CrossPath} = \biguplus_{k \geq 1}(\text{ReachLoop}(v_{\text{win}}, k) \cap \text{CrossLoop}(k-1))$ gives the first equality. The second is a restatement, the third equality follows from Property 2 and Property 1. The other equality is achieved by the same proof, replacing v_{win} by v_{lose} and using Property 1 accordingly. ∎

Property 4. *We have* $\mathbb{P}(\text{Reach}(v_{\text{win}})) \geq \alpha \cdot \frac{\beta}{\beta+\gamma}$ *and similarly* $\mathbb{P}(\text{Reach}(v_{\text{lose}})) \geq \alpha \cdot \frac{\gamma}{\beta+\gamma}$.

The intuition behind these two equalities is that we try to ignore what happens while crossing the path, as reaching either v_{win} or v_{lose} is not correlated to the priorities seen infinitely often. In this context, the multiplicative constant α stands for the loss due to crossing the path. As soon as the path is crossed, what happens next will be correlated to the priorities seen infinitely often along the play. We will see that the value of the looping game described above captures the outcome of the parity game.

Proof. We have the following equalities:

$$
\begin{aligned}
\mathbb{P}(\text{Reach}(v_{\text{win}})) &= \; \mathbb{P}(\text{ReachPath}(v_{\text{win}})) + \mathbb{P}(\text{Reach}(v_{\text{win}}) \cap \text{CrossPath}) \\
&= \; \mathbb{P}(\text{ReachPath}(v_{\text{win}})) + \mathbb{P}(\text{CrossPath}) \cdot \mathbb{P}(\text{Reach}(v_{\text{win}}) \mid \text{CrossPath}) \\
&\geq \; \mathbb{P}(\text{CrossPath}) \cdot \mathbb{P}(\text{Reach}(v_{\text{win}}) \mid \text{CrossPath}) \\
&= \; \alpha \cdot \frac{\beta}{\beta+\gamma}
\end{aligned}
$$

From the disjoint union $\text{Reach}(v_{\text{win}}) = \text{ReachPath}(v_{\text{win}}) \uplus (\text{Reach}(v_{\text{win}}) \cap \text{CrossPath})$ follows the first equality. The second is restatement, the inequality is straightforward, the following equality is a restatement and the last equality follows from Property 3 and definition of α. The other claim is achieved by the same proof, replacing v_{win} by v_{lose} and using 3 accordingly. ∎

Approximations for α, β and γ. Note that for $i \geq 1$, we have $\mathbb{P}(\widehat{\pi}_i \in \{v_{\text{win}}, v_{\text{lose}}\} \mid \widehat{\pi}_{i-1} \notin \{v_{\text{win}}, v_{\text{lose}}\}) = P_{v_i}$, which follows from the construction of $\mathfrak{R}(G)$: taking an escape edge comes with probability P_{v_i}.

Property 5. *Given the assumption* (A_0) *is satisfied, we have* $\alpha \geq \frac{5}{6}$.

Intuitively, this property means that the loss due to crossing the path is bounded by a constant.

Proof. Since the path $\mathscr{P}$ is simple, each vertex is visited at most once. Let $I = \{i \mid v_i \in \mathscr{P}\}$ be the set of vertices visited by this path. Then

$$1 - \alpha = \sum_{i \in I} \mathbb{P}(\widehat{\pi}_{i-1} \notin \{v_{\text{win}}, v_{\text{lose}}\} \cap \widehat{\pi}_i \in \{v_{\text{win}}, v_{\text{lose}}\}) \leq \sum_{i \in I} P_{v_i} \leq \frac{1}{6}$$

The first equality follows from the disjoint union:

$$
\begin{aligned}
\Pi \setminus \text{CrossPath} \quad & \\
= \quad & \text{ReachPath}(v_{\text{lose}}) \cup \text{ReachPath}(v_{\text{win}}) \\
= \quad & \biguplus_{i \in I} (\widehat{\pi}_{i-1} \notin \{v_{\text{win}}, v_{\text{lose}}\} \cap \widehat{\pi}_i \in \{v_{\text{win}}, v_{\text{lose}}\})
\end{aligned}
$$

The last inequality follows from assumption (A_0). ∎

Property 6. *Given the assumptions (A_1) and (A_2) are satisfied, if Eve wins the play $\pi(v_0, \sigma, \tau)$ in G, then we have the following inequalities:*

$$
\begin{aligned}
(1) \quad \beta \quad &\geq \quad P_{v_{l+1}} \\
(2) \quad \gamma \quad &\leq \quad \tfrac{2}{3} \cdot P_{v_{l+1}}
\end{aligned}
$$

and similarly if Adam wins the play $\pi(v_0, \sigma, \tau)$ in G, then we have the following inequalities:

$$
\begin{aligned}
(1) \quad \gamma \quad &\geq \quad P_{v_{l+1}} \\
(2) \quad \beta \quad &\leq \quad \tfrac{2}{3} \cdot P_{v_{l+1}}.
\end{aligned}
$$

Intuitively, this property means that if Eve wins in the parity game, then the looping game is winning for her with probability more than $\frac{2}{3}$, and similarly if Adam wins in the parity game, then the looping game is winning for him with probability more than $\frac{2}{3}$.

Proof. We prove inequalities in both cases simultaneously.

1. It relies on the fact that the loop starts by getting to v_{l+1}, *i.e* that either, if Eve wins: $\widehat{\pi}_{l+1} = v_{\text{win}} \cap \text{CrossPath} \subseteq \text{ReachLoop}(v_{\text{win}}, 1) \cap \text{CrossPath}$, or if Adam wins: $\widehat{\pi}_{l+1} = v_{\text{lose}} \cap \text{CrossPath} \subseteq \text{ReachLoop}(v_{\text{lose}}, 1) \cap \text{CrossPath}$.

2. Assume Eve wins, let $J = \{i \mid v_i \in \mathscr{C} \wedge p(v_i) \text{ odd}\}$ be the set of vertices with odd priority visited by the loop. Then

$$\gamma = \sum_{i \in J} \mathbb{P}(\widehat{\pi}_{i-1} \neq v_{\text{lose}} \cap \widehat{\pi}_i = v_{\text{lose}}) \leq \sum_{i \in J} P_{v_i}$$

(A_1) allows to conclude, since v_{l+1} has the lowest priority of the loop, thus $J \subseteq J_{\text{odd}}^{>v_{l+1}}$. Similarly, if Adam wins, the same proof using assumption (A_2) concludes.

∎

It follows from Property 4, 5 and 6 that under assumptions $(A_0) - (A_2)$, we have the desired equivalence: Eve wins in G if and only if $\mathbb{P}(\text{Reach}(v_{\text{win}})) \geq \frac{1}{2}$ in $\mathfrak{R}(G)$.

Theorem 2. *Under the three assumptions $(A_0) - (A_2)$, we have: for all 2-player arenas G equipped with parity objective $\text{Parity}(p)$, $\mathfrak{R}(G)$ equipped with reachability condition $\text{Reach}(v_{\text{win}})$ is a simple stochastic game with the stopping property, and for all $v \in V$, Eve wins for the parity condition from v in G if and only if $\langle E \rangle(\text{Reach}(v_{\text{win}}))(v) \geq \frac{1}{2}$ in $\mathfrak{R}(G)$.*

Transition probabilities. We now present transition probabilities satisfying the assumptions $(A_0) - (A_2)$ that can be described with $O(\log(n))$ bits. Let $p : V \to \mathbb{N}$ the priority function in G, we first build an equivalent parity function p'. We sort vertices with respect to p and define the following *monotone* mapping:

- the lowest priority becomes either 4 if it is even or 5 if odd;

- proceeding from the lowest to the greatest, a vertex is assigned the lowest integer greater than the last integer used, matching its parity.

This ensures that all priorities are distinct, and the highest priority is at most $2n + 2$. Then, apply the reduction $\mathfrak{R}$ with $P_j = \frac{1}{2^j}$. We argue that the probability transition function satisfies $(A_0) - (A_2)$. We have

$$\sum_{v \in V} P_v \leq \frac{1}{2^4} + \frac{1}{2^5} + \ldots = \frac{1}{8} \leq \frac{1}{6}$$

Hence (A_0) is satisfied. For all $v \in V$,

$$\sum_{u \in J_{\text{odd}}^{>v}} P_u \leq P_v \cdot \left(\frac{1}{2} + \frac{1}{2 \cdot 2^2} + \frac{1}{2 \cdot 2^4} + \ldots \right) = P_v \cdot \frac{1}{2} \cdot \frac{4}{3} = \frac{2}{3} \cdot P_v$$

Hence (A_1) is satisfied and a similar argument holds for (A_2). Hence we have the following result.

Theorem 3. $\mathfrak{R}$ *is a polynomial-time reduction from* 2*-player parity games to simple stochastic games. Furthermore, for all* 2*-player parity games* $(G, \text{Parity}(p))$, *the size of the stochastic arena* $\mathfrak{R}(G)$ *is* $O(|E| \cdot \log(|V|))$.

4 Reducing parity games to simple stochastic games

In this section, we discuss related works. Deciding the winner in parity games is equivalent to the model-checking problem of modal mu-calculus. A reduction from model-checking games to simple stochastic games was defined in [7]. Another reduction, using a discounted mu-calculus, from concurrent parity games to concurrent discounted games was presented in [1]. Our intend was to propose a direct and simple reduction from 2-player parity games to simple stochastic games. In the first subsection, we discuss its efficiency compared to the previously known three step reduction. In the second subsection, we use remarks from [5] to prove that solving stochastic parity, mean-payoff, discounted-payoff games as well as simple stochastic games is in UP ∩ coUP.

4.1 Discounting the discounted

In this subsection we present the classical sequence of reductions: from 2-player parity games to 2-player mean-payoff games [6], from 2-player mean-payoff games to 2-player discounted-payoff games [9], and from 2-player discounted games to simple stochastic games [9].

Parity games to mean-payoff games. A 2-player parity game with n vertices and d different priorities can be reduced in polynomial time to a 2-player mean payoff game on the same arena using rewards from the set $\{-n^d, \ldots, n^d\}$, such that Eve wins the parity game if and only if the value of Eve in the mean-payoff game is at least 0 [6].

Mean-payoff games to discounted-payoff games. A 2-player mean payoff game with n vertices whose reward function ranges from $-B$ to B can be reduced in polynomial time to a discounted-payoff game on

the same arena with discount factor λ such that $\lambda \geq 1 - \frac{B}{4n^3}$ such that the value of Eve in the mean-payoff game is at least 0 if and only if the value of Eve in the discounted-payoff game is at least 0 [9].

Discounted-payoff games to simple stochastic games. A 2-player discounted-payoff game with n vertices can be reduced in polynomial time to a simple stochastic game using $n + m$ vertices including m random vertices and $4 \cdot m$ edges such that the value of Eve in the discounted-payoff game is at least 0 if and only if the value of Eve in the simple stochastic game is at least $\frac{1}{2}$ [9].

Size of the resulting games. We now analyze the size of the games produced by the three step reduction. Let G a 2-player parity game having n vertices, m edges and d distinct priorities. The first reduction to a 2-player mean payoff game yields a game with n vertices, m edges and rewards can be specified with $O(d \cdot \log(n))$ bits. The second reduction to a discounted-payoff game yields a 2-player game with n vertices, m edges, rewards specified with $O(d \cdot \log(n))$ bits and the discount factor specified with $O(d \cdot \log(n))$ bits. Finally, the last reduction to a simple stochastic game yields a game with $n + m$ vertices, with m random vertices, $4 \cdot m$ edges and each probability of transition specified with $O(d \cdot \log(n))$ bits, thus the size of the transition function is $O(m \cdot (\log(n + m) + d \cdot \log(n))) = O(m \cdot d \cdot \log(n))$. Since d is $O(n)$, in the worst case the size of the game obtained by the three step reduction is $O(m \cdot n \cdot \log(n))$.

4.2 The complexity of stochastic games

Another motivation to present a clean and direct reduction from 2-player parity games to simple stochastic games was to extend it from stochastic parity games to simple stochastic games. As for the deterministic case, such a reduction is known, but again through stochastic mean-payoff and stochastic discounted-payoff, and is more involved [3]. Although we did not manage to adapt our proofs to extend our direct reduction from stochastic parity games to simple stochastic games, we believe it is possible. Our main difficulty is that the shape of a play, even if both players play positionally, is no more a "lasso-play". Indeed, even if the parity condition is satisfied with probability more than half, we cannot guarantee that an even priority will be visited within a linear number of steps.

In the remaining of this subsection, we gather several results and make two very simple observations of the result of Condon [5] to prove that the decision problem for simple stochastic games is in UP $\cap$ coUP, in a similar fashion to the proof of [6], which was stated for the simpler case of 2-player discounted games.

The reduction of stochastic parity to stochastic mean-payoff games was established in [3] and reduction of stochastic mean-payoff and discounted games to simple stochastic games was established in [2]. The Figure 3 summarizes all the reductions. We now argue that simple stochastic games can be decided in UP $\cap$ coUP.

Simple stochastic games in UP $\cap$ coUP**.** First, it was shown in [4] that simple stochastic games with arbitrary rational transition probabilities can be reduced in polynomial time to stopping simple stochastic games where random vertices have two outgoing edges each with probability half. Second, it follows from the result of [5] that the value vector of a *stopping* simple stochastic game is the unique solution of the following equations set:

$$\begin{cases} \langle E \rangle(v_{\text{win}}) = 1 \\ \langle E \rangle(v_{\text{lose}}) = 0 \\ \langle E \rangle(v) = \max_{(v,v') \in E} \text{val}_E(v') & \text{if } v \in V_E \\ \langle E \rangle(v) = \min_{(v,v') \in E} \text{val}_E(v') & \text{if } v \in V_A \\ \langle E \rangle(v) = \sum_{(v,v') \in E} \delta(v,v') \cdot \text{val}_E(v') & \text{if } v \in V_R \end{cases}$$
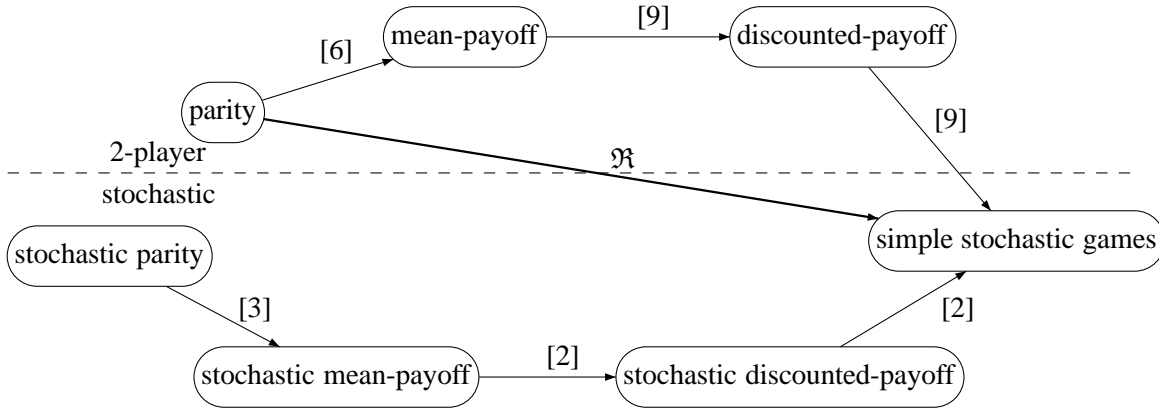
Figure 3: Reductions

Hence an algorithm can guess the value vector and check that it is actually the solution of the equation set. To prove the desired result we need to show that the guess is of polynomial size and the verification can be achieved in polynomial time. It follows from [4] that for simple stochastic games with n vertices and all probabilities one half, the values are of the form p/q, where p,q are integers, $0 \leq p, q \leq 4^{n-1}$. Hence the length of the guess is at most $n \cdot \log(4^{n-1}) = O(n^2)$, which is polynomial. Thus the guess is of polynomial size and the verification can be done in polynomial time. The unique solution implies that simple stochastic games are in UP, and the coUP argument is symmetric. Along with the reductions of [2, 3] we obtain the following result.

Theorem 4 (Complexity of stochastic games)**.** *For all stochastic arenas, for all objectives ϕ such that ϕ is a parity, mean-payoff, discounted-payoff or reachability objective, the decision problem of whether $\langle\mathrm{E}\rangle(\phi)(v) \geq q$, for a rational number q is in* UP$\cap$coUP.

References

[1] Luca de Alfaro, Thomas A. Henzinger & Rupak Majumdar (2003): *Discounting the Future in Systems Theory*. In: *International Colloquium on Automata, Languages and Programming, ICALP*, pp. 1022–1037.

[2] Daniel Andersson & Peter Bro Miltersen (2009): *The Complexity of Solving Stochastic Games on Graphs*. In: *International Symposium on Algorithms and Computation, ISAAC*, pp. 112–121.

[3] Krishnendu Chatterjee & Thomas A. Henzinger (2008): *Reduction of stochastic parity to stochastic mean-payoff games*. *Information Processing Letters, IPL* 106(1), pp. 1–7.

[4] Anne Condon (1992): *The Complexity of Stochastic Games*. *Information and Computation* 96(2), pp. 203–224.

[5] Anne Condon (1993): *On Algorithms for Simple Stochastic Games*. In: *Advances in Computational Complexity Theory*, *DIMACS Series in Discrete Mathematics and Theoretical Computer Science* 13, American Mathematical Society, pp. 51–73.

[6] Marcin Jurdziński (1998): *Deciding the Winner in Parity Games is in UP ∩ co-UP*. *Information Processing Letters, IPL* 68(3), pp. 119–124.

[7] Colin Stirling (1999): *Bisimulation, Modal Logic and Model Checking Games*. *Logic Journal of the IGPL* 7(1), pp. 103–124.

[8] Wolfgang Thomas (1997): *Languages, Automata, and Logic*. In G. Rozenberg & A. Salomaa, editors: *Handbook of Formal Languages*, chapter 7, 3, Beyond Words, Springer, pp. 389–455.

[9] Uri Zwick & Mike Paterson (1996): *The complexity of mean payoff games on graphs*. *Theoretical Computer Science, TCS* 158, pp. 343–359.