

Arena-Independent Finite-Memory Determinacy in Stochastic Games

Patricia Bouyer 

Université Paris-Saclay, CNRS, ENS Paris-Saclay,
Laboratoire Méthodes Formelles, 91190, Gif-sur-Yvette, France

Youssef Oualhadj

Univ Paris Est Creteil, LACL, F-94010 Creteil, France

Mickael Randour

F.R.S.-FNRS & UMONS – Université de Mons, Belgium

Pierre Vandenhover 

F.R.S.-FNRS & UMONS – Université de Mons, Belgium
Université Paris-Saclay, CNRS, ENS Paris-Saclay,
Laboratoire Méthodes Formelles, 91190, Gif-sur-Yvette, France

Abstract

We study stochastic zero-sum games on graphs, which are prevalent tools to model decision-making in presence of an antagonistic opponent in a random environment. In this setting, an important question is the one of *strategy complexity*: what kinds of strategies are sufficient or required to play optimally (e.g., randomization or memory requirements)? Our contributions further the understanding of *arena-independent finite-memory (AIFM) determinacy*, i.e., the study of objectives for which memory is needed, but in a way that only depends on limited parameters of the game graphs. First, we show that objectives for which pure AIFM strategies suffice to play optimally also admit pure AIFM *subgame perfect* strategies. Second, we show that we can reduce the study of objectives for which pure AIFM strategies suffice in two-player stochastic games to the easier study of one-player stochastic games (i.e., Markov decision processes). Third, we *characterize* the sufficiency of AIFM strategies through two intuitive properties of objectives. This work extends a line of research started on deterministic games in [7] to stochastic ones.

2012 ACM Subject Classification Theory of computation → Formal languages and automata theory

Keywords and phrases two-player games on graphs, stochastic games, Markov decision processes, finite-memory determinacy, optimal strategies

Digital Object Identifier 10.4230/LIPIcs.CONCUR.2021.26

Related Version *Full Version*: <https://arxiv.org/abs/2102.10104> [9]

Funding Research supported by F.R.S.-FNRS under Grant n° F.4520.18 (ManySynth), and ENS Paris-Saclay visiting professorship (M. Randour, 2019). This work has been partially supported by the ANR Project MAVeriQ (ANR-ANR-20-CE25-0012). Mickael Randour is an F.R.S.-FNRS Research Associate. Pierre Vandenhover is an F.R.S.-FNRS Research Fellow.

1 Introduction

Controller synthesis consists, given a system, an environment, and a specification, in automatically generating a controller of the system that guarantees the specification in the environment. This task is often studied through a game-theoretic lens: the system is a game, the controller is a player, the uncontrollable environment is its adversary, and the specification is a game objective [45]. A *game on graph* consists of a directed graph, called an *arena*, partitioned into two kinds of vertices: some of them are controlled by the system (called *player 1*) and the others by the environment (called *player 2*). Player 1 is given a



© Patricia Bouyer, Youssef Oualhadj, Mickael Randour, and Pierre Vandenhover;
licensed under Creative Commons License CC-BY 4.0

32nd International Conference on Concurrency Theory (CONCUR 2021).

Editors: Serge Haddad and Daniele Varacca; Article No. 26; pp. 26:1–26:18

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

game objective (corresponding to the specification) and must devise a *strategy* (corresponding to the controller) to accomplish the objective or optimize an outcome. The strategy can be seen as a function that dictates the decisions to make in order to react to every possible chain of events. In case of uncertainty in the system or the environment, probability distributions are often used to model transitions in the game graph, giving rise to the *stochastic game* model. We study here *stochastic turn-based zero-sum games on graphs* [29], also called perfect-information stochastic games. We also discuss the case of *deterministic games*, which can be seen as a subcase of stochastic games in which only Dirac distributions are used in transitions.

Strategy complexity. A common question underlying all game objectives is the one of *strategy complexity*: how *complex* must optimal strategies be, and how *simple* can optimal strategies be? For each distinct game objective, multiple directions can be investigated, such as the need for randomization [19] (must optimal strategies make stochastic choices?), the need for memory [33, 34, 7] (how much information about the past must optimal strategies remember?), or what trade-offs exist between randomization and memory [16, 26, 41]. With respect to memory requirements, three cases are typically distinguished: *memoryless-determined objectives*, for which *memoryless* strategies suffice to play optimally; *finite-memory-determined objectives*, for which finite-memory strategies suffice (memory is then usually encoded as a deterministic finite automaton); and objectives for which infinite memory is required. High memory requirements (such as exponential memory and obviously infinite memory) are a major drawback when it comes to implementing controllers; hence specific approaches are often developed to look for *simple* strategies (e.g., [30]).

Many classical game objectives (reachability [29], Büchi and parity [22], discounted sum [48], energy [10]...) are memoryless-determined, both in deterministic and stochastic arenas. Nowadays, multiple general results allow for a more manageable proof for most of these objectives: we mention [35, 4, 1] for sufficient conditions in deterministic games, and [31, 32] for similar conditions in one-player and two-player stochastic games. One milestone for memoryless determinacy in deterministic games was achieved by Gimbert and Zielonka [33], who provide *two characterizations* of it: the first one states two necessary and sufficient conditions (called *monotony* and *selectivity*) for memoryless determinacy, and the second one states that memoryless determinacy in both players' *one-player games* suffices for memoryless determinacy in two-player games (we call this result the *one-to-two-player lift*). Together, these characterizations provide a theoretical and practical advance. On the one hand, monotony and selectivity improve the high-level understanding of what conditions well-behaved objectives verify. On the other hand, only having to consider the one-player case thanks to the one-to-two-player lift is of tremendous help in practice. A generalization of the one-to-two-player lift to *stochastic games* was shown also by Gimbert and Zielonka in an unpublished paper [34] and is about memoryless strategies that are *pure* (i.e., not using randomization).

The need for memory. Recent research tends to study increasingly complex settings – such as combinations of qualitative/quantitative objectives or of behavioral models – for which finite or infinite memory is often required; see examples in deterministic games [17, 50, 12, 8, 14], Markov decision processes – i.e., one-player stochastic games [46, 47, 24, 3, 11], or stochastic games [28, 18, 25, 23, 40]. Motivated by the growing memory requirements of these endeavors, research about strategy complexity often turns toward *finite-memory determinacy*. Proving finite-memory determinacy is sometimes difficult (already in deterministic games, e.g., [6]),

and as opposed to memoryless strategies, there are few widely applicable results. We mention [38], which provides sufficient conditions for finite-memory determinacy in Boolean combinations of finite-memory-determined objectives in deterministic games. Results for multi-player non-zero-sum games are also available [37].

Arena-independent finite-memory. A middle ground between the well-understood memoryless determinacy and the more puzzling finite-memory determinacy was proposed for deterministic games in [7]: an objective is said to admit *arena-independent finite-memory (AIFM) strategies* if a *single finite memory structure* suffices to play optimally in any arena. In practice, this memory structure may depend on parameters of the objective (for instance, largest weight, number of priorities), but not on parameters intrinsically linked to the arena (e.g., number of states or transitions). AIFM strategies include as a special case memoryless strategies, since they can be implemented with a trivial memory structure with a single state.

AIFM strategies have a remarkable feature: in deterministic arenas, AIFM generalizations of *both* characterizations from [33] hold, including the one-to-two-player lift [7]. From a practical point of view, it brings techniques usually linked to memoryless determinacy to many finite-memory-determined objectives. The aim of this article is to show that this also holds true in stochastic arenas.

Contributions. We provide an overview of desirable properties of objectives in which pure AIFM strategies suffice to play optimally in *stochastic* games, and tools to study them. This entails: (a) a proof of a specific feature of objectives for which pure AIFM strategies suffice to play optimally: for such objectives, there also exist pure AIFM *subgame perfect (SP)* strategies (Theorem 7), which is a stronger requirement than optimality; (b) a more general one-to-two-player lift: we show the equivalence between the existence of pure AIFM optimal strategies in two-player games for both players and the existence of pure AIFM optimal strategies in one-player games, thereby simplifying the proof of memory requirements for many objectives (Theorem 8); (c) two intuitive conditions generalizing monotony and selectivity in the stochastic/AIFM case, which are equivalent to the sufficiency of pure AIFM strategies to play optimally in *one-player stochastic arenas* (Theorem 13) for all objectives that can be encoded as *real payoff functions*. In practice, (c) can be used to prove memory requirements in one-player arenas, and then (b) can be used to lift these to the two-player case.

These results reinforce both sides on the frontier between AIFM strategies and general finite-memory strategies: on the one hand, objectives for which pure AIFM strategies suffice indeed share interesting properties with objectives for which pure memoryless strategies suffice, rendering their analysis easier, even in the stochastic case; on the other hand, our novel result about SP strategies does not hold for (arena-*dependent*) finite-memory strategies, and therefore further distinguishes the AIFM case from the finite-memory case.

The one-to-two-player lift for pure AIFM strategies in stochastic games is not surprising, as it holds for pure memoryless strategies in stochastic games [34], and for AIFM strategies in deterministic games [7]. However, although the monotony/selectivity characterization is inspired from the deterministic case [33, 7], it had not been formulated for stochastic games, even in the pure memoryless case – its proof involves new technical difficulties to which our improved understanding of subgame perfect strategies comes in handy.

All our results are about the optimality of *pure* AIFM strategies in various settings: they can be applied in an independent way for deterministic games and for stochastic games, and they can also consider optimality under restriction to different classes of strategies (allowing or not the use of randomization and infinite memory).

The proof technique for the one-to-two-player lift shares a similar outline in [33, 34, 7] and in this paper: it relies on an induction on the number of edges in arenas to show the existence of memoryless optimal strategies. This edge-induction technique is frequently used in comparable ways in other works about memoryless determinacy [35, 31, 32, 18]. In the AIFM case, the extra challenge consists of applying such an induction to the right set of arenas in order for a result about memoryless strategies to imply something about AIFM strategies. Work in [7] paved the way to neatly overcome this technical hindrance and we were able to factorize the main argument in Lemma 6.

Applicability. Let us discuss objectives that admit, or not, pure AIFM optimal strategies in stochastic arenas.

- Objectives for which AIFM optimal strategies exist include the aforementioned memoryless-determined objectives [29, 22, 48, 10], as explained earlier. Such objectives could already be studied through the lens of a one-to-two-player lift [34], but our two other main results also apply to these.
- Pure AIFM optimal strategies also exist in lexicographic reachability-safety games [23, Theorem 4]: the memory depends only on the number of targets to visit or avoid, but not on parameters of the arena (number of states or transitions).
- Muller objectives whose probability must be maximized [15] also admit pure AIFM optimal strategies: the number of memory states depends only on the colors and on the Muller condition.
- In general, every ω -regular objective admits pure AIFM optimal strategies, as it can be seen as a parity objective (for which pure memoryless strategies suffice) after taking the product of the game graph with a deterministic parity automaton accepting the objective [42, 20]. This parity automaton can be taken as an arena-independent memory structure. It is therefore possible to use our results to investigate precise memory bounds in stochastic games for multiple ω -regular objectives which have been studied in deterministic games or in one-player stochastic games: generalized parity games [21], lower- and upper-bounded energy games [5], some window objectives [13, 11], weak parity games [49].
- There are objectives for which finite-memory strategies suffice for some player, but with an underlying memory structure depending on parameters of the arena (an example is provided by the *Gain* objective in [40, Theorem 6]). Many objectives also require infinite memory, such as generalized mean-payoff games [18] (both in deterministic and stochastic games) and energy-parity games (only in stochastic games [17, 39]). Our characterizations provide a more complete understanding of why AIFM strategies do not suffice.

Deterministic and stochastic games. There are natural ways to extend classical objectives for deterministic games to a stochastic context: typically, for qualitative objectives, a natural stochastic extension is to maximize the probability to win. Still, in general, memory requirements may increase when switching to the stochastic context. To show that understanding the deterministic case is insufficient to understand the stochastic case, we outline three situations displaying different behaviors.

- As mentioned above, for many classical objectives, memoryless strategies suffice both in deterministic and in stochastic games.
- AIFM strategies may suffice both for deterministic and stochastic games, but with a difference in the size of the required memory structure. One such example is provided by the weak parity objective [49], for which memoryless strategies suffice in deterministic

games, but which requires memory in stochastic games (this was already noticed in [34, Section 4.4]). Yet, it is possible to show that pure AIFM strategies suffice in stochastic games using the results from our paper. This shows that to go from the deterministic to the stochastic case, a “constant” increase in memory may be necessary and sufficient.

- There are also objectives for which memoryless strategies suffice in deterministic games, but even AIFM strategies do not suffice in stochastic games. One such example consists in maximizing the probability to obtain a non-negative discounted sum (which is different from maximizing the expected value of the discounted sum, for which memoryless strategies suffice, as is shown in [48]).

Formal proofs for these last two examples are provided in the full version [9]. These three situations further highlight the significance of establishing results about memory requirements in stochastic games, even for objectives whose deterministic version is well-understood.

Outline. We introduce our framework and notations in Section 2. We discuss AIFM strategies and tools to relate them to memoryless strategies in Section 3, which allows us to prove our result about subgame perfect strategies. The one-to-two-player lift is presented in Section 4, followed by the one-player characterization in Section 5. Due to a lack of space, we choose to focus on Section 5 and only sketch Section 4; the complete proofs and technical details are found in the full version of the article [9].

2 Preliminaries

Let C be an arbitrary set of *colors*.

Arenas. For a measurable space (Ω, \mathcal{F}) (resp. a finite set Ω), we write $\text{Dist}(\Omega, \mathcal{F})$ (resp. $\text{Dist}(\Omega)$) for the set of *probability distributions on* (Ω, \mathcal{F}) (resp. *on* Ω). For Ω a finite set and $\mu \in \text{Dist}(\Omega)$, we write $\text{Supp}(\mu) = \{\omega \in \Omega \mid \mu(\omega) > 0\}$ for the *support of* μ .

We consider stochastic games played by two players, called \mathcal{P}_1 (for player 1) and \mathcal{P}_2 (for player 2), who play in a turn-based fashion on *arenas*. A (two-player stochastic turn-based) *arena* is a tuple $\mathcal{A} = (S_1, S_2, A, \delta, \text{col})$, where: S_1 and S_2 are two disjoint finite sets of *states*, respectively controlled by \mathcal{P}_1 and \mathcal{P}_2 – we denote $S = S_1 \uplus S_2$; A is a finite set of *actions*; $\delta: S \times A \rightarrow \text{Dist}(S)$ is a partial function called *probabilistic transition function*; $\text{col}: S \times A \rightarrow C$ is a partial function called *coloring function*. For a state $s \in S$, we write $A(s)$ for the set of actions that are *available in* s , that is, the set of actions for which $\delta(s, a)$ is defined. For $s \in S$, function col must be defined for all pairs (s, a) such that a is available in s . We require that for all $s \in S$, $A(s) \neq \emptyset$ (i.e., arenas are *non-blocking*).

For $s, s' \in S$ and $a \in A(s)$, we denote $\delta(s, a, s')$ instead of $\delta(s, a)(s')$ for the probability to reach s' in one step by playing a in s , and we write $(s, a, s') \in \delta$ if and only if $\delta(s, a, s') > 0$. An interesting subclass of (stochastic) arenas is the class of *deterministic* arenas: an arena $\mathcal{A} = (S_1, S_2, A, \delta, \text{col})$ is *deterministic* if for all $s \in S$, $a \in A(s)$, $|\text{Supp}(\delta(s, a))| = 1$.

A *play of* \mathcal{A} is an infinite sequence of states and actions $s_0 a_1 s_1 a_2 s_2 \dots \in (SA)^\omega$ such that for all $i \geq 0$, $(s_i, a_{i+1}, s_{i+1}) \in \delta$. A prefix of a play is an element in $S(AS)^*$ and is called a *history*; the set of all histories starting in a state $s \in S$ is denoted $\text{Hists}(\mathcal{A}, s)$. For $S' \subseteq S$, we write $\text{Hists}(\mathcal{A}, S')$ for the unions of $\text{Hists}(\mathcal{A}, s)$ over all states $s \in S'$. For $\rho = s_0 a_1 s_1 \dots a_n s_n$ a history, we write $\text{out}(\rho)$ for s_n . For $i \in \{1, 2\}$, we write $\text{Hists}_i(\mathcal{A}, s)$ and $\text{Hists}_i(\mathcal{A}, S')$ for the corresponding histories ρ such that $\text{out}(\rho) \in S_i$. For $s, s' \in S$, we write $\text{Hists}(\mathcal{A}, s, s')$ for the histories in $\text{Hists}(\mathcal{A}, s)$ such that $\text{out}(\rho) = s'$.

We write $\widehat{\text{col}}$ for the extension of col to histories and plays: more precisely, for a history $\rho = s_0 a_1 s_1 \dots a_n s_n$, $\widehat{\text{col}}(\rho)$ is the finite sequence $\text{col}(s_0, a_1) \dots \text{col}(s_{n-1}, a_n) \in C^*$; for $\pi = s_0 a_1 s_1 a_2 s_2 \dots$ a play, $\widehat{\text{col}}(\pi)$ is the infinite sequence $\text{col}(s_0, a_1) \text{col}(s_1, a_2) \dots \in C^\omega$.

A *one-player arena* of \mathcal{P}_i is an arena $\mathcal{A} = (S_1, S_2, A, \delta, \text{col})$ such that for all $s \in S_{3-i}$, $|A(s)| = 1$. A one-player arena corresponds to a *Markov decision process (MDP)* [44, 2].

An *initialized arena* is a pair $(\mathcal{A}, S_{\text{init}})$ such that \mathcal{A} is an arena and S_{init} is a non-empty subset of the states of \mathcal{A} , called the set of *initial states*. We assume w.l.o.g. that all states of \mathcal{A} are reachable from S_{init} following transitions with positive probabilities in the probabilistic transition function of \mathcal{A} . In case of a single initial state $s \in S$, we write (\mathcal{A}, s) for $(\mathcal{A}, \{s\})$.

We will consider *classes* (sets) of initialized arenas, which are usually denoted by the letter \mathfrak{A} . Typical classes that we will consider consist of all one-player or two-player, deterministic or stochastic initialized arenas. We use *initialized* arenas throughout the paper for technical reasons, but our results can be converted to results using the classical notion of arena.

Memory. We define a notion of memory based on complete deterministic automata on *colors*. The goal of using colors instead of states/actions for transitions of the memory is to define memory structures *independently of arenas*. A *memory skeleton* is a tuple $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$ where M is a set of *memory states*, $m_{\text{init}} \in M$ is an *initial state* and $\alpha_{\text{upd}}: M \times C \rightarrow M$ is an *update function*. We add the following constraint: for all finite sets of colors $B \subseteq C$, the number of states reachable from m_{init} with transitions provided by $\alpha_{\text{upd}}|_{M \times B}$ is finite (where $\alpha_{\text{upd}}|_{M \times B}$ is the restriction of the domain of α_{upd} to $M \times B$).

Memory skeletons with a finite state space are all encompassed by this definition, but this also allows some skeletons with infinitely many states. For example, if $C = \mathbb{N}$, the tuple $(\mathbb{N}, 0, (m, n) \mapsto \max\{m, n\})$, which remembers the largest color seen, is a valid memory skeleton: for any finite $B \subseteq C$, we only need to use memory states up to $\max B$. However, the tuple $(\mathbb{N}, 0, (m, n) \mapsto m + n)$ remembering the current sum of colors seen is *not* a memory skeleton, as infinitely many states are reachable from 0, even if only $B = \{1\}$ can be used. We denote $\widehat{\alpha}_{\text{upd}}: M \times C^* \rightarrow M$ for the natural extension of α_{upd} to finite sequences of colors.

Let $\mathcal{M}^1 = (M^1, m_{\text{init}}^1, \alpha_{\text{upd}}^1)$ and $\mathcal{M}^2 = (M^2, m_{\text{init}}^2, \alpha_{\text{upd}}^2)$ be memory skeletons. Their *product* $\mathcal{M}^1 \otimes \mathcal{M}^2$ is the memory skeleton $(M, m_{\text{init}}, \alpha_{\text{upd}})$ with $M = M^1 \times M^2$, $m_{\text{init}} = (m_{\text{init}}^1, m_{\text{init}}^2)$, and, for all $m^1 \in M^1$, $m^2 \in M^2$, $c \in C$, $\alpha_{\text{upd}}((m^1, m^2), c) = (\alpha_{\text{upd}}^1(m^1, c), \alpha_{\text{upd}}^2(m^2, c))$. The update function of the product simply updates both skeletons in parallel.

Strategies. Given an initialized arena $(\mathcal{A} = (S_1, S_2, A, \delta, \text{col}), S_{\text{init}})$ and $i \in \{1, 2\}$, a *strategy* of \mathcal{P}_i on $(\mathcal{A}, S_{\text{init}})$ is a function $\sigma_i: \text{Hists}_i(\mathcal{A}, S_{\text{init}}) \rightarrow \text{Dist}(A)$ such that for all $\rho \in \text{Hists}_i(\mathcal{A}, S_{\text{init}})$, $\text{Supp}(\sigma_i(\rho)) \subseteq A(\text{out}(\rho))$. For $i \in \{1, 2\}$, we denote by $\Sigma_i^G(\mathcal{A}, S_{\text{init}})$ the set of all strategies of \mathcal{P}_i on $(\mathcal{A}, S_{\text{init}})$.

A strategy σ_i of \mathcal{P}_i on $(\mathcal{A}, S_{\text{init}})$ is *pure* if for all $\rho \in \text{Hists}_i(\mathcal{A}, S_{\text{init}})$, $|\text{Supp}(\sigma_i(\rho))| = 1$. If a strategy is not pure, then it is *randomized*. A strategy σ_i of \mathcal{P}_i on $(\mathcal{A}, S_{\text{init}})$ is *memoryless* if for all $\rho, \rho' \in \text{Hists}_i(\mathcal{A}, S_{\text{init}})$, $\text{out}(\rho) = \text{out}(\rho')$ implies $\sigma_i(\rho) = \sigma_i(\rho')$. A *pure memoryless* strategy of \mathcal{P}_i can be simply specified as a function $S_i \rightarrow A$. A strategy σ_i of \mathcal{P}_i on $(\mathcal{A}, S_{\text{init}})$ is *finite-memory* if it can be encoded as a *Mealy machine* $\Gamma = (\mathcal{M}, \alpha_{\text{nxt}})$, with $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$ a memory skeleton and $\alpha_{\text{nxt}}: S_i \times M \rightarrow \text{Dist}(A)$ being the *next-action function*, which is such that for $s \in S_i$, $m \in M$, $\text{Supp}(\alpha_{\text{nxt}}(s, m)) \subseteq A(s)$. Strategy σ_i is encoded by Γ if for all histories $\rho \in \text{Hists}_i(\mathcal{A}, S_{\text{init}})$, $\sigma_i(\rho) = \alpha_{\text{nxt}}(\text{out}(\rho), \widehat{\alpha}_{\text{upd}}(m_{\text{init}}, \widehat{\text{col}}(\rho)))$. If σ_i can be encoded as a Mealy machine $(\mathcal{M}, \alpha_{\text{nxt}})$, we say that σ_i is *based on (memory) \mathcal{M}* . If σ_i is *based on \mathcal{M}* and is *pure*, then the next-action function can be specified as a function $S_i \times M \rightarrow A$. Memoryless strategies correspond to finite-memory strategies based on the *trivial memory skeleton* $\mathcal{M}_{\text{triv}} = (\{m_{\text{init}}\}, m_{\text{init}}, (m_{\text{init}}, c) \mapsto m_{\text{init}})$ that has a single state.

We denote by $\Sigma_i^{\text{PFM}}(\mathcal{A}, S_{\text{init}})$ (resp. $\Sigma_i^{\text{P}}(\mathcal{A}, S_{\text{init}})$, $\Sigma_i^{\text{GFM}}(\mathcal{A}, S_{\text{init}})$, $\Sigma_i^{\text{G}}(\mathcal{A}, S_{\text{init}})$) the set of pure finite-memory (resp. pure, finite-memory, general) strategies of \mathcal{P}_i on $(\mathcal{A}, S_{\text{init}})$. A *type of strategies* is an element $X \in \{\text{PFM}, \text{P}, \text{GFM}, \text{G}\}$ corresponding to these subsets.

Outcomes. Let $(\mathcal{A} = (S_1, S_2, A, \delta, \text{col}), S_{\text{init}})$ be an initialized arena. When both players have decided on a strategy and an initial state has been chosen, the generated object is a (finite or countably infinite) Markov chain, which induces a probability distribution on the plays. For strategies σ_1 of \mathcal{P}_1 and σ_2 of \mathcal{P}_2 on $(\mathcal{A}, S_{\text{init}})$ and $s \in S_{\text{init}}$, we denote $\mathbf{P}_{\mathcal{A},s}^{\sigma_1, \sigma_2}$ for the probability distribution on plays induced by σ_1 and σ_2 , starting from state s .

We define \mathcal{F} to be the smallest σ -algebra on C^ω generated by the set of all *cylinders* on C . In particular, every probability distribution $\mathbf{P}_{\mathcal{A},s}^{\sigma_1, \sigma_2}$ naturally induces a probability distribution over (C^ω, \mathcal{F}) through the $\widehat{\text{col}}$ function, which we denote $\mathbf{Pc}_{\mathcal{A},s}^{\sigma_1, \sigma_2}$.

Preferences. To specify each player's objective, we use the general notion of *preference relation*. A *preference relation* \sqsubseteq (on C) is a total preorder over $\text{Dist}(C^\omega, \mathcal{F})$. The idea is that \mathcal{P}_1 favors the distributions in $\text{Dist}(C^\omega, \mathcal{F})$ that are the largest for \sqsubseteq , and as we are studying *zero-sum* games, \mathcal{P}_2 favors the distributions that are the smallest for \sqsubseteq . For \sqsubseteq a preference relation and $\mu, \mu' \in \text{Dist}(C^\omega, \mathcal{F})$, we write $\mu \sqsubset \mu'$ if $\mu \sqsubseteq \mu'$ and $\mu' \not\sqsubseteq \mu$.

Depending on the context, it might not be necessary to define a preference relation as *total*: it is sufficient to order distributions that can arise as an element $\mathbf{P}_{\mathcal{A},s}^{\sigma_1, \sigma_2}$. For example, in the specific case of deterministic games in which only pure strategies are considered, all distributions that arise are always Dirac distributions on a single infinite word in C^ω . In this context, it is therefore sufficient to define a total preorder over all Dirac distributions (which we can then see as infinite words, giving a definition of preference relation similar to [33, 7]). We give some examples to illustrate our notion of *preference relation*.

► **Example 1.** We give three examples corresponding to three different ways to encode preference relations. First, a preference relation can be induced by an event $W \in \mathcal{F}$ called a *winning condition*, which consists of infinite sequences of colors. The objective of \mathcal{P}_1 is to maximize the probability that the event W happens. An event W naturally induces a preference relation \sqsubseteq_W such that for $\mu, \mu' \in \text{Dist}(C^\omega, \mathcal{F})$, $\mu \sqsubseteq_W \mu'$ if and only if $\mu(W) \leq \mu'(W)$. For $C = \mathbb{N}$, we give the example of the *weak parity* winning condition W_{wp} [49], defined as $W_{\text{wp}} = \{c_1 c_2 \dots \in C^\omega \mid \max_{j \geq 1} c_j \text{ exists and is even}\}$. In finite arenas, the value $\max_{j \geq 1} c_j$ always exists, as there are only finitely many colors that appear. This is different from the classical parity condition, which requires the maximal color *seen infinitely often* to be even, and not just the maximal color seen.

A preference relation can also be induced by a Borel (*real*) *payoff function* $f: C^\omega \rightarrow \mathbb{R}$. For example, if $C = \mathbb{R}$ and $\lambda \in (0, 1)$, a classical payoff function [48] is the *discounted sum* Disc_λ , defined for $c_1 c_2 \dots \in C^\omega$ as $\text{Disc}_\lambda(c_1 c_2 \dots) = \lim_n \sum_{i=0}^n \lambda^i \cdot c_{i+1}$. The goal of \mathcal{P}_1 is to maximize the expected value of f , which is defined for a probability distribution $\mu \in \text{Dist}(C^\omega, \mathcal{F})$ as $\mathbf{E}_\mu[f] = \int f d\mu$. A payoff function f naturally induces a preference relation \sqsubseteq_f : for $\mu_1, \mu_2 \in \text{Dist}(C^\omega, \mathcal{F})$, $\mu_1 \sqsubseteq_f \mu_2$ if and only if $\mathbf{E}_{\mu_1}[f] \leq \mathbf{E}_{\mu_2}[f]$. Payoff functions are more general than winning conditions: for W a winning condition, the preference relation induced by the indicator function of W corresponds to the preference relation induced by W .

It is also possible to specify preference relations that cannot be expressed as a payoff function. An example is given in [27]: we assume that the goal of \mathcal{P}_1 is to see color $c \in C$ with probability *precisely* $\frac{1}{2}$. We denote the event of seeing color c as $\diamond c \in \mathcal{F}$. Then for $\mu, \mu' \in \text{Dist}(C^\omega, \mathcal{F})$, $\mu \sqsubseteq \mu'$ if and only if $\mu(\diamond c) \neq \frac{1}{2}$ or $\mu'(\diamond c) = \frac{1}{2}$. \dashv

A (two-player stochastic turn-based zero-sum) *initialized game* is a tuple $\mathcal{G} = (\mathcal{A}, S_{\text{init}}, \sqsubseteq)$, where $(\mathcal{A}, S_{\text{init}})$ is an initialized arena and \sqsubseteq is a preference relation.

Optimality. Let $\mathcal{G} = (\mathcal{A}, S_{\text{init}}, \sqsubseteq)$ be an initialized game and $X \in \{\text{PFM}, \text{P}, \text{GFM}, \text{G}\}$ be a type of strategies. For $s \in S_{\text{init}}$, $\sigma_1 \in \Sigma_1^X(\mathcal{A}, S_{\text{init}})$, we define

$$\text{UCol}_{\sqsubseteq}^X(\mathcal{A}, s, \sigma_1) = \{\mu \in \text{Dist}(C^\omega, \mathcal{F}) \mid \exists \sigma_2 \in \Sigma_2^X(\mathcal{A}, s), \text{Pc}_{\mathcal{A}, s}^{\sigma_1, \sigma_2} \sqsubseteq \mu\}.$$

The set $\text{UCol}_{\sqsubseteq}^X(\mathcal{A}, s, \sigma_1)$ corresponds to all the distributions that are at least as good for \mathcal{P}_1 (w.r.t. \sqsubseteq) as a distribution that \mathcal{P}_2 can induce by playing a strategy σ_2 of type X against σ_1 ; this set is upward-closed w.r.t. \sqsubseteq . For $\sigma_1, \sigma'_1 \in \Sigma_1^X(\mathcal{A}, S_{\text{init}})$, we say that σ_1 is *at least as good as σ'_1 from $s \in S_{\text{init}}$ under X strategies* if $\text{UCol}_{\sqsubseteq}^X(\mathcal{A}, s, \sigma_1) \subseteq \text{UCol}_{\sqsubseteq}^X(\mathcal{A}, s, \sigma'_1)$. This inclusion means that the best replies of \mathcal{P}_2 against σ'_1 yield an outcome that is at least as bad for \mathcal{P}_1 (w.r.t. \sqsubseteq) as the best replies of \mathcal{P}_2 against σ_1 . We can define symmetrical notions for strategies of \mathcal{P}_2 .

Let $\mathcal{G} = (\mathcal{A}, S_{\text{init}}, \sqsubseteq)$ be an initialized game and $X \in \{\text{PFM}, \text{P}, \text{GFM}, \text{G}\}$ be a type of strategies. A strategy $\sigma_i \in \Sigma_i^X(\mathcal{A}, S_{\text{init}})$ is *X-optimal in \mathcal{G}* if it is at least as good under X strategies as any other strategy in $\Sigma_i^X(\mathcal{A}, S_{\text{init}})$ from all $s \in S_{\text{init}}$.

When the considered preference relation \sqsubseteq is clear, we often talk about X -optimality in an initialized arena $(\mathcal{A}, S_{\text{init}})$ to refer to X -optimality in the initialized game $(\mathcal{A}, S_{\text{init}}, \sqsubseteq)$. Given a preference relation, a class of arenas, and a type of strategies, our goal is to understand what kinds of strategies are sufficient to play optimally. In the following definition, abbreviations *AIFM* and *FM* stand respectively for *arena-independent finite-memory* and *finite-memory*.

► **Definition 2.** Let \sqsubseteq be a preference relation, \mathfrak{A} be a class of initialized arenas, $X \in \{\text{PFM}, \text{P}, \text{GFM}, \text{G}\}$ be a type of strategies, and \mathcal{M} be a memory skeleton. We say that pure AIFM strategies suffice to play X -optimally in \mathfrak{A} for \mathcal{P}_1 if there exists a memory skeleton \mathcal{M} such that for all $(\mathcal{A}, S_{\text{init}}) \in \mathfrak{A}$, \mathcal{P}_1 has a pure strategy based on \mathcal{M} that is X -optimal in $(\mathcal{A}, S_{\text{init}})$. We say that pure FM strategies suffice to play X -optimally in \mathfrak{A} for \mathcal{P}_1 if for all $(\mathcal{A}, S_{\text{init}}) \in \mathfrak{A}$, there exists a memory skeleton \mathcal{M} such that \mathcal{P}_1 has a pure strategy based on \mathcal{M} that is X -optimal in $(\mathcal{A}, S_{\text{init}})$.

Since memoryless strategies are a specific kind of finite-memory strategies based on the same memory skeleton $\mathcal{M}_{\text{triv}}$, the sufficiency of pure memoryless strategies is equivalent to the sufficiency of pure strategies based on $\mathcal{M}_{\text{triv}}$, and is therefore a specific case of the sufficiency of pure AIFM strategies. Notice the difference between the order of quantifiers for AIFM and FM strategies: the sufficiency of pure AIFM strategies implies the sufficiency of pure FM strategies, but the opposite is false in general (an example is given in [17]).

► **Example 3.** Let us reconsider the weak parity winning condition W_{wp} introduced in Example 1: the goal of \mathcal{P}_1 is to maximize the probability that the greatest color seen is even. To play optimally in any stochastic game, it is sufficient for both players to remember the largest color already seen, which can be implemented by the memory skeleton $\mathcal{M}_{\text{max}} = (\mathbb{N}, 0, (m, n) \mapsto \max\{m, n\})$. As explained above, this memory skeleton has an infinite state space, but as there are only finitely many colors in every (finite) arena, only a finite part of the skeleton is sufficient to play optimally in any given arena. The size of the skeleton used for a fixed arena depends on the appearing colors, but for a fixed number of colors, it does not depend on parameters of the arena (such as its state and action spaces). Therefore pure AIFM strategies suffice to play optimally for both players, and more precisely pure strategies based on \mathcal{M}_{max} suffice for both players. ─

We define a second stronger notion related to optimality of strategies, which is the notion of *subgame perfect strategy*: a strategy is *subgame perfect* in a game if it reacts optimally to all histories consistent with the arena, even histories not consistent with the strategy itself, or histories that only a non-rational adversary would play [43]. This is a desirable property of strategies that is stronger than optimality, since a subgame perfect strategy is not only optimal from the initial position, but from any arbitrary stage (*subgame*) of the game. In particular, if an opponent plays non-optimally, an optimal strategy that is not subgame perfect does not always fully exploit the advantage that the opponent's suboptimal behavior provides, and may yield a result that is not optimal *when starting in a subgame*. We first need extra definitions.

For $w \in C^*$, $\mu \in \text{Dist}(C^\omega, \mathcal{F})$, we define the *shifted distribution* $w\mu$ as the distribution such that for an event $E \in \mathcal{F}$, $w\mu(E) = \mu(\{w' \in C^\omega \mid ww' \in E\})$.

For $(\mathcal{A}, S_{\text{init}})$ an initialized arena, for $\sigma_i \in \Sigma_i^G(\mathcal{A}, S_{\text{init}})$, and for $\rho = s_0 a_1 s_1 \dots a_n s_n \in \text{Hists}(\mathcal{A}, S_{\text{init}})$, we define the *shifted strategy* $\sigma_i[\rho] \in \Sigma_i^G(\mathcal{A}, \text{out}(\rho))$ as the strategy such that, for $\rho' = s_n a_{n+1} s_{n+1} \dots a_m s_m \in \text{Hists}_i(\mathcal{A}, \text{out}(\rho))$, $\sigma_i[\rho](\rho') = \sigma_i(s_0 a_1 s_1 \dots a_m s_m)$.

For \sqsubseteq a preference relation and $w \in C^*$, we define the *shifted preference relation* $\sqsubseteq_{[w]}$ as the preference relation such that for $\mu, \mu' \in \text{Dist}(C^\omega, \mathcal{F})$, $\mu \sqsubseteq_{[w]} \mu'$ if and only if $w\mu \sqsubseteq w\mu'$.

► **Definition 4.** Let $\mathcal{G} = (\mathcal{A}, S_{\text{init}}, \sqsubseteq)$ be an initialized game and $X \in \{\text{PFM}, \text{P}, \text{GFM}, \text{G}\}$ be a type of strategies. A strategy $\sigma_i \in \Sigma_i^X(\mathcal{A}, S_{\text{init}})$ is *X-subgame perfect (X-SP)* in \mathcal{G} if for all $\rho \in \text{Hists}(\mathcal{A}, S_{\text{init}})$, shifted strategy $\sigma_i[\rho]$ is *X-optimal* in the initialized game $(\mathcal{A}, \text{out}(\rho), \sqsubseteq_{[\widehat{\text{col}}(\rho)])}$.

Strategies that are X-SP are in particular X-optimal; the converse is not true in general.

3 Coverability and subgame perfect strategies

In this section, we establish a key tool (Lemma 6) which can be used to reduce questions about the sufficiency of *AIFM strategies* in reasonable classes of initialized arenas to the sufficiency of *memoryless strategies* in a subclass. We then describe the use of this lemma to obtain our first main result (Theorem 7), which shows that the sufficiency of pure AIFM strategies implies the existence of pure AIFM *SP* strategies. Technical details are in [9].

► **Definition 5.** An initialized arena $((S_1, S_2, A, \delta, \text{col}), S_{\text{init}})$ is covered by memory skeleton $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$ if there exists a function $\phi: S \rightarrow M$ such that for all $s \in S_{\text{init}}$, $\phi(s) = m_{\text{init}}$, and for all $(s, a, s') \in \delta$, $\alpha_{\text{upd}}(\phi(s), \text{col}(s, a)) = \phi(s')$.

The *coverability* property means that it is possible to assign a unique memory state to each arena state such that transitions of the arena always update the memory state in a way that is consistent with the memory skeleton. A covered initialized arena carries in some way already sufficient information to play with memory \mathcal{M} without actually using memory – using the memory skeleton \mathcal{M} would not be more powerful than using no memory at all. This property is linked to the classical notion of *product arena with \mathcal{M}* : a strategy based on \mathcal{M} corresponds to a memoryless strategy in a product arena (e.g., [7, Lemma 1] and [9]). For our results, *coverability* is a key technical definition, as the class of initialized arenas covered by a memory skeleton is sufficiently well-behaved to support edge-induction arguments, whereas it is difficult to perform such techniques directly on the class of *product arenas*: removing a single edge from a product arena makes it hard to express as a product arena, whereas it is clear that coverability is preserved. Every initialized arena is covered by $\mathcal{M}_{\text{triv}}$, which is witnessed by the function ϕ associating m_{init} to every state.

Definitions close to our notion of *coverability by \mathcal{M}* were introduced for deterministic arenas in [36, 7]. The definition of *adherence with \mathcal{M}* in [36, Definition 8.12] is very similar, but does not distinguish initial states from the rest (neither in the arena nor in the memory

skeleton). Our property of $(\mathcal{A}, S_{\text{init}})$ being *covered by* \mathcal{M} is also equivalent to \mathcal{A} being both *prefix-covered* and *cyclic-covered* by \mathcal{M} from S_{init} [7]. Distinguishing both notions gives insight in [7] as they are used at different places in proofs (*prefix-covered* along with *monotony*, and *cyclic-covered* along with *selectivity*). Here, we opt for a single concise definition.

The following lemma sums up our practical use of the idea of *coverability*.

► **Lemma 6.** *Let \sqsubseteq be a preference relation, \mathcal{M} be a memory skeleton, and $X \in \{\text{PFM}, \text{P}, \text{GFM}, \text{G}\}$ be a type of strategies. Let \mathfrak{A} be the class of one-player or two-player, stochastic or deterministic initialized arenas. Then, \mathcal{P}_1 has an X -optimal (resp. X -SP) strategy based on \mathcal{M} in all initialized arenas in \mathfrak{A} if and only if \mathcal{P}_1 has a memoryless X -optimal (resp. X -SP) strategy in all initialized arenas covered by \mathcal{M} in \mathfrak{A} .*

We now state one of our main results, which shows that the sufficiency of pure strategies based on the same memory skeleton \mathcal{M} implies that pure SP strategies based on \mathcal{M} exist.

► **Theorem 7.** *Let \sqsubseteq be a preference relation, \mathcal{M} be a memory skeleton, and $X \in \{\text{PFM}, \text{P}, \text{GFM}, \text{G}\}$ be a type of strategies. Let \mathfrak{A} be the class of one-player or two-player, stochastic or deterministic initialized arenas. If \mathcal{P}_1 has pure X -optimal strategies based on \mathcal{M} in initialized arenas of \mathfrak{A} , then \mathcal{P}_1 has pure X -SP strategies based on \mathcal{M} in initialized arenas of \mathfrak{A} .*

Proof sketch. By Lemma 6, we can prove instead that \mathcal{P}_1 has a pure memoryless X -SP strategy in every initialized arena covered by \mathcal{M} , based on the hypothesis that \mathcal{P}_1 has a pure memoryless X -optimal strategy in every initialized arena covered by \mathcal{M} . For $(\mathcal{A}, S_{\text{init}}) \in \mathfrak{A}$ covered by \mathcal{M} , \mathcal{P}_1 has a pure memoryless X -optimal strategy σ_1^0 . If this strategy is not X -SP, there must be a prefix $\rho \in \text{Hists}(\mathcal{A}, S_{\text{init}})$ such that σ_1^0 is not X -optimal in $(\mathcal{A}, \text{out}(\rho), \sqsubseteq_{[\widehat{\text{col}}(\rho)]})$. Then, we extend $(\mathcal{A}, S_{\text{init}})$ by adding a “chain” of states with colors $\widehat{\text{col}}(\rho)$ up to $\text{out}(\rho)$, and add as an initial state the first state of this chain. *This new arena is still covered by \mathcal{M}* , thus \mathcal{P}_1 has a pure memoryless X -optimal strategy in this arena that is now X -optimal after seeing ρ . If this strategy is not X -SP, we keep iterating our reasoning. This iteration necessarily ends, as we consider *finite* arenas, on which there are *finitely many* pure memoryless strategies. ◀

This result shows a major distinction between the sufficiency of AIFM strategies and the more general sufficiency of FM strategies: if a player can always play optimally with the same memory, then SP strategies may be played with the same memory as optimal strategies – if a player can play optimally but needs arena-*dependent* finite memory, then infinite memory may still be required to obtain SP strategies. One such example is provided in [38, Example 16] for the *average-energy games with lower-bounded energy* in deterministic arenas: \mathcal{P}_1 can always play optimally with pure finite-memory strategies [6, Theorem 13], but infinite memory is needed for SP strategies. As will be further explained later, we will also use Theorem 7 to gain technical insight in the proof of the main result of Section 5.

4 One-to-two-player lift

Our goal in this section is to expose a practical tool to help study the memory requirements of two-player stochastic (or deterministic) games. This tool consists in reducing the study of the sufficiency of pure AIFM strategies for both players in *two*-player games to *one*-player games. We first state our result, and we then explain how it relates to similar results from the literature and sketch its proof. A slightly generalized result with a more fine-grained quantification on the classes of initialized arenas is in [9], with the complete proof.

► **Theorem 8** (Pure AIFM one-to-two-player lift). *Let \sqsubseteq be a preference relation, \mathcal{M}_1 and \mathcal{M}_2 be two memory skeletons, and $X \in \{\text{PFM}, \text{P}, \text{GFM}, \text{G}\}$ be a type of strategies. Let \mathfrak{A} be the class of all initialized stochastic or deterministic arenas.*

Assume that in all initialized one-player arenas of \mathcal{P}_1 in \mathfrak{A} , \mathcal{P}_1 can play X -optimally with a pure strategy based on \mathcal{M}_1 , and in all initialized one-player arenas of \mathcal{P}_2 in \mathfrak{A} , \mathcal{P}_2 can play X -optimally with a pure strategy based on \mathcal{M}_2 . Then in all initialized two-player arenas in \mathfrak{A} , both players have a pure X -SP strategy based on $\mathcal{M}_1 \otimes \mathcal{M}_2$.

The practical usage of this result can be summed up as follows: to determine whether pure AIFM strategies are sufficient for both players in stochastic (resp. deterministic) arenas to play X -optimally, it is sufficient to prove it for stochastic (resp. deterministic) *one-player* arenas. Our theorem deals in a uniform manner with stochastic and deterministic arenas, under different types of strategies. Studying memory requirements of one-player arenas is significantly easier than doing so in two-player arenas, as a one-player arena can be seen as a graph (in the deterministic case) or an MDP (in the stochastic case). Still, we bring more tools to study memory requirements of one-player arenas in Section 5.

Theorem 8 generalizes known one-to-two-player lifts: for pure memoryless strategies in deterministic [33] and stochastic [34] games, and for pure AIFM strategies in deterministic games [7]. Very briefly, our proof technique consists in extending the lift for *pure memoryless strategies* in stochastic games [34] in order to deal with *initialized* arenas. Then, we show that this pure memoryless one-to-two-player lift can be applied to the class of *initialized arenas covered by $\mathcal{M}_1 \otimes \mathcal{M}_2$* (using an edge-induction technique), and Lemma 6 permits to go back from pure memoryless strategies to pure strategies based on $\mathcal{M}_1 \otimes \mathcal{M}_2$. Thanks to Theorem 7, we also go further in our understanding of the optimal strategies: we obtain the existence of X -SP strategies instead of the seemingly weaker existence of X -optimal strategies.

5 AIFM characterization

For this section, we fix \sqsubseteq a preference relation, $X \in \{\text{PFM}, \text{P}, \text{GFM}, \text{G}\}$ a type of strategies, and $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$ a memory skeleton. We distinguish two classes of initialized arenas: the class $\mathfrak{A}_{\mathcal{P}_1}^{\text{D}}$ of all initialized one-player deterministic arenas of \mathcal{P}_1 , and the class $\mathfrak{A}_{\mathcal{P}_1}^{\text{S}}$ of all initialized one-player stochastic arenas of \mathcal{P}_1 . A class of arenas will therefore be specified by a letter $Y \in \{\text{D}, \text{S}\}$, which we fix for the whole section. Our aim is to give a better understanding of the preference relations for which pure strategies based on \mathcal{M} suffice to play X -optimally in $\mathfrak{A}_{\mathcal{P}_1}^Y$, by *characterizing* it through two intuitive conditions. All definitions and proofs are stated from the point of view of \mathcal{P}_1 . As we only work with one-player arenas in this section, we abusively write $P_{\mathcal{A},s}^{\sigma_1}$ and $Pc_{\mathcal{A},s}^{\sigma_1}$ for the distributions on plays and colors induced by a strategy σ_1 of \mathcal{P}_1 on (\mathcal{A}, s) , with the unique, trivial strategy for \mathcal{P}_2 .

For $\mathcal{A} \in \mathfrak{A}_{\mathcal{P}_1}^Y$ and s a state of \mathcal{A} , we write $[\mathcal{A}]_s^X = \{Pc_{\mathcal{A},s}^{\sigma_1} \mid \sigma_1 \in \Sigma_1^X(\mathcal{A}, s)\}$ for the set of all distributions over (C^ω, \mathcal{F}) induced by strategies of type X in \mathcal{A} from s .

For $m_1, m_2 \in M$, we write $L_{m_1, m_2} = \{w \in C^* \mid \widehat{\alpha_{\text{upd}}}(m_1, w) = m_2\}$ for the language of words that are read from m_1 up to m_2 in \mathcal{M} . Such a language can be specified by the deterministic automaton that is simply the memory skeleton \mathcal{M} with m_1 as the initial state and m_2 as the unique final state. We extend the *shifted distribution* notation to *sets* of distributions: for $w \in C^*$, for $\Lambda \subseteq \text{Dist}(C^\omega, \mathcal{F})$, we write $w\Lambda$ for the set $\{w\mu \mid \mu \in \Lambda\}$.

Given \sqsubseteq a preference relation, we also extend \sqsubseteq to sets of distributions: for $\Lambda_1, \Lambda_2 \subseteq \text{Dist}(C^\omega, \mathcal{F})$, we write $\Lambda_1 \sqsubseteq \Lambda_2$ if for all $\mu_1 \in \Lambda_1$, there exists $\mu_2 \in \Lambda_2$ such that $\mu_1 \sqsubseteq \mu_2$; we write $\Lambda_1 \sqsubset \Lambda_2$ if there exists $\mu_2 \in \Lambda_2$ such that for all $\mu_1 \in \Lambda_1$, $\mu_1 \sqsubset \mu_2$. Notice that $\neg(\Lambda_1 \sqsubseteq \Lambda_2)$ is equivalent to $\Lambda_2 \sqsubset \Lambda_1$. If Λ_1 is a singleton $\{\mu_1\}$, we write $\mu_1 \sqsubseteq \Lambda_2$ for

$\{\mu_1\} \sqsubseteq \Lambda_2$ (and similarly for Λ_2 , and similarly using \sqsupseteq). Notice that $\mu_1 \sqsubseteq \mu_2$ is equivalent to $\{\mu_1\} \sqsubseteq \{\mu_2\}$, so this notational shortcut is sound. For two initialized arenas (\mathcal{A}_1, s_1) and (\mathcal{A}_2, s_2) , the inequality $[\mathcal{A}_1]_{s_1}^X \sqsubseteq [\mathcal{A}_2]_{s_2}^X$ means that for every strategy on (\mathcal{A}_1, s_1) , there is a strategy on (\mathcal{A}_2, s_2) that induces a distribution that is at least as good.

For two arenas \mathcal{A}_1 and \mathcal{A}_2 with disjoint state spaces, if s_1 and s_2 are two states controlled by \mathcal{P}_1 that are respectively in \mathcal{A}_1 and \mathcal{A}_2 with disjoint sets of available actions, we write $(\mathcal{A}_1, s_1) \sqcup (\mathcal{A}_2, s_2)$ for the *merged arena* in which s_1 and s_2 are merged, and everything else is kept the same. The merged state which comes from the merge of s_1 and s_2 is usually called t . Formally, let $\mathcal{A}_1 = (S_1^1, S_2^1, A^1, \delta^1, \text{col}^1)$, $\mathcal{A}_2 = (S_1^2, S_2^2, A^2, \delta^2, \text{col}^2)$, $s_1 \in S_1^1$, and $s_2 \in S_2^1$. We assume that $S^1 \cap S^2 = \emptyset$ and that $A(s_1) \cap A(s_2) = \emptyset$. We define $(\mathcal{A}_1, s_1) \sqcup (\mathcal{A}_2, s_2)$ as the arena $(S_1, S_2, A, \delta, \text{col})$ with $S_1 = S_1^1 \uplus S_1^2 \uplus \{t\} \setminus \{s_1, s_2\}$, $S_2 = S_2^1 \uplus S_2^2$; we define $A(t) = A^1(s_1) \uplus A^2(s_2)$ and for $i \in \{1, 2\}$, $\delta(t, a) = \delta^i(t, a)$ and $\text{col}(t, a) = \text{col}^i(t, a)$ if $a \in A(s_i)$; all the other available actions, transitions and colors are kept the same as in the original arenas (with transitions going to s_1 or s_2 being directed to t). A symmetrical definition can be written if s_1 and s_2 are both controlled by \mathcal{P}_2 .

We can now present the two properties of preference relations at the core of our characterization. These properties are called *X-Y-M-monotony* and *X-Y-M-selectivity*; they depend on a type of strategies X , a type of arenas Y , and a memory skeleton \mathcal{M} . The first appearance of the monotony (resp. selectivity) notion was in [33], which dealt with deterministic arenas and memoryless strategies; their monotony (resp. selectivity) is equivalent to our P-D- $\mathcal{M}_{\text{triv}}$ -monotony (resp. P-D- $\mathcal{M}_{\text{triv}}$ -selectivity). In [7], these definitions were generalized to deal with the sufficiency of strategies based on \mathcal{M} in deterministic arenas; their notion of \mathcal{M} -monotony (resp. \mathcal{M} -selectivity) is equivalent to our P-D- \mathcal{M} -monotony (resp. P-D- \mathcal{M} -selectivity).

► **Definition 9 (Monotony).** *We say that \sqsubseteq is X-Y-M-monotone if for all $m \in M$, for all $(\mathcal{A}_1, s_1), (\mathcal{A}_2, s_2) \in \mathfrak{A}_{\mathcal{P}_1}^Y$, there exists $i \in \{1, 2\}$ s.t. for all $w \in L_{m_{\text{init}}, m}$, $w[\mathcal{A}_{3-i}]_{s_{3-i}}^X \sqsubseteq w[\mathcal{A}_i]_{s_i}^X$.*

The crucial part of the definition is the order of the last two quantifiers: of course, given a $w \in L_{m_{\text{init}}, m}$, as \sqsubseteq is total, it will always be the case that $w[\mathcal{A}_1]_{s_1}^X \sqsubseteq w[\mathcal{A}_2]_{s_2}^X$ or that $w[\mathcal{A}_2]_{s_2}^X \sqsubseteq w[\mathcal{A}_1]_{s_1}^X$. However, we ask for something stronger: it must be the case that the set of distributions $w[\mathcal{A}_i]_{s_i}^X$ is preferred to $w[\mathcal{A}_{3-i}]_{s_{3-i}}^X$ for *any* word $w \in L_{m_{\text{init}}, m}$.

The original monotony definition [33] states that when presented with a choice *once* among two possible continuations, if a continuation is better than the other one after some prefix, then this continuation is also at least as good after all prefixes. This property is not sufficient for the sufficiency of pure memoryless strategies as it does not guarantee that if the same choice presents itself multiple times in the game, the same continuation should always be chosen, as alternating between both continuations might still be beneficial in the long run – this is dealt with by selectivity. If memory \mathcal{M} is necessary to play optimally, then it makes sense that there are different optimal choices depending on the current memory state and that we should only compare prefixes that reach the same memory state. The point of taking into account a memory skeleton \mathcal{M} in our definition of X-Y-M-monotony is to only compare prefixes that are read up to the same memory state from m_{init} .

► **Definition 10 (Selectivity).** *We say that \sqsubseteq is X-Y-M-selective if for all $m \in M$, for all $(\mathcal{A}_1, s_1), (\mathcal{A}_2, s_2) \in \mathfrak{A}_{\mathcal{P}_1}^Y$ such that for $i \in \{1, 2\}$, $\widehat{\text{col}}(\text{Hists}(\mathcal{A}_i, s_i, s_i)) \subseteq L_{m, m}$, for all $w \in L_{m_{\text{init}}, m}$, $w[(\mathcal{A}_1, s_1) \sqcup (\mathcal{A}_2, s_2)]_t^X \sqsubseteq w[\mathcal{A}_1]_{s_1}^X \cup w[\mathcal{A}_2]_{s_2}^X$ (where t comes from the merge of s_1 and s_2).*

Our formulation of the selectivity concept differs from the original definition [33] and its AIFM counterpart [7] in order to take into account the particularities of the stochastic context, even if it can be proven that they are equivalent in the pure deterministic case. The

idea is still the same: the original selectivity definition states that when presented with a choice among multiple possible continuations after some prefix, if a continuation is better than the others, then as the game goes on, if the same choice presents itself again, it is sufficient to always pick the same continuation to play optimally; there is no need to alternate between continuations. This property is not sufficient for the sufficiency of pure memoryless strategies as it does not guarantee that for all prefixes, the same initial choice is always the one we should commit to – this is dealt with by monotony. The point of memory skeleton \mathcal{M} in our definition is to guarantee that every time the choice is presented, we are currently in the same memory state.

An interesting property is that both notions are stable by product with a memory skeleton: if \sqsubseteq is X-Y- \mathcal{M} -monotone (resp. X-Y- \mathcal{M} -selective), then for all memory skeletons \mathcal{M}' , \sqsubseteq is also X-Y- $(\mathcal{M} \otimes \mathcal{M}')$ -monotone (resp. X-Y- $(\mathcal{M} \otimes \mathcal{M}')$ -selective). The reason is that in each definition, we quantify universally over the class of all prefixes w that reach the same memory state m ; if we consider classes that are subsets of the original classes, then the definition still holds. This property matches the idea that playing with more memory is never detrimental.

Combined together, it is intuitively reasonable that X-Y- \mathcal{M} -monotony and X-Y- \mathcal{M} -selectivity are equivalent to the sufficiency of pure strategies based on \mathcal{M} to play X-optimally in $\mathfrak{A}_{\mathcal{P}_1}^Y$: monotony tells us that when a single choice has to be made given a state of the arena and a memory state, the best choice is always the same no matter what prefix has been seen, and selectivity tells us that once a good choice has been made, we can commit to it in the future of the game. We formalize this idea in Theorem 13. First, we add an extra restriction on preference relations which is useful when stochasticity is involved.

► **Definition 11** (Mixing is useless). *We say that mixing is useless for \sqsubseteq if for all sets I at most countable, for all positive reals $(\lambda_i)_{i \in I}$ such that $\sum_{i \in I} \lambda_i = 1$, for all families $(\mu_i)_{i \in I}$, $(\mu'_i)_{i \in I}$ of distributions in $\text{Dist}(C^\omega, \mathcal{F})$, if for all $i \in I$, $\mu_i \sqsubseteq \mu'_i$, then $\sum_{i \in I} \lambda_i \mu_i \sqsubseteq \sum_{i \in I} \lambda_i \mu'_i$.*

That is, if we can write a distribution as a convex combination of distributions, then it is never detrimental to improve a distribution appearing in the convex combination.

► **Remark 12.** All preference relations encoded as Borel real payoff functions (as defined in Example 1) satisfy this property (it is easy to show the property for indicator functions, and we can then extend this fact to all Borel functions thanks to properties of the Lebesgue integral). The third preference relation from Example 1 (reaching $c \in C$ with probability precisely $\frac{1}{2}$) does not satisfy this property: if $\mu_1(\diamond c) = 0$, $\mu'_1(\diamond c) = \frac{1}{2}$, and $\mu_2(\diamond c) = 1$, we have $\mu_1 \sqsubset \mu'_1$ and $\mu_2 \sqsubseteq \mu_2$, but $\frac{1}{2}\mu'_1 + \frac{1}{2}\mu_2 \sqsubset \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2$. In deterministic games with pure strategies, only Dirac distributions on infinite words occur as distributions induced by an arena and a strategy, so the requirement that mixing is useless is not needed. ─

► **Theorem 13.** *Assume that no stochasticity is involved (that is, $X \in \{P, \text{PFM}\}$ and $Y = D$), or that mixing is useless for \sqsubseteq . Then pure strategies based on \mathcal{M} suffice to play X-optimally in all initialized one-player arenas in $\mathfrak{A}_{\mathcal{P}_1}^Y$ for \mathcal{P}_1 if and only if \sqsubseteq is X-Y- \mathcal{M} -monotone and X-Y- \mathcal{M} -selective.*

Proof sketch. We sketch both directions of the proof (available in [9]). The proof of the necessary condition of Theorem 13 is the easiest direction. The main idea is to build the right arenas (using the arenas occurring in the definitions of monotony and selectivity) so that we can use the hypothesis about the existence of pure X-optimal strategies based on \mathcal{M} to immediately deduce X-Y- \mathcal{M} -monotony and X-Y- \mathcal{M} -selectivity. It is not necessary that mixing is useless for \sqsubseteq for this direction of the equivalence.

For the sufficient condition, we first reduce as usual the statement to the existence of pure memoryless strategies in covered initialized arenas, using Lemma 6. We proceed with an edge-induction in these arenas (as for Theorem 8). The base case is trivial (as in an arena in which all states have a single available action, there is a single strategy which is pure and memoryless). For the induction step, we take an initialized arena $(\mathcal{A}', S_{\text{init}}) \in \mathfrak{A}_{\mathcal{P}_1}^Y$ covered by \mathcal{M} , and we pick a state t with (at least) two available actions. A memory state $\phi(t)$ is associated to t thanks to coverability. We consider arenas $(\mathcal{A}'_a, S_{\text{init}})$ obtained from $(\mathcal{A}', S_{\text{init}})$ by leaving a single action a available in t , to which we can apply the induction hypothesis and obtain a pure memoryless X -optimal strategy σ_1^a . It is left to prove that one of these strategies is also X -optimal in $(\mathcal{A}', S_{\text{init}})$; this is where X - Y - \mathcal{M} -monotony and X - Y - \mathcal{M} -selectivity come into play.

The property of X - Y - \mathcal{M} -monotony tells us that one of these subarenas $(\mathcal{A}'_{a^*}, S_{\text{init}})$ is preferred to the others w.r.t. \sqsubseteq after reading any word in $L_{m_{\text{init}}, \phi(t)}$. We now want to use X - Y - \mathcal{M} -selectivity to conclude that there is no reason to use actions different from a^* when coming back to t , and that $\sigma_1^{a^*}$ is therefore also X -optimal in $(\mathcal{A}', S_{\text{init}})$. To do so, we take any strategy $\sigma_1 \in \Sigma_1^X(\mathcal{A}', s)$ for $s \in S_{\text{init}}$ and we condition distribution $P_{\mathcal{A}', s}^{\sigma_1}$ over all the ways it reaches (or not) t , which gives a convex combination of probability distributions. We want to state that once t is reached, no matter how, switching to strategy $\sigma_1^{a^*}$ is always beneficial. For this, we would like to use X -subgame-perfection of $\sigma_1^{a^*}$ rather than simply X -optimality: this is why in the actual proof, our induction hypothesis is about X -SP strategies and not X -optimal strategies. Luckily, Theorem 7 indicates that requiring subgame perfection is not really stronger than what we want to prove. We then use that mixing is useless for \sqsubseteq (Definition 11) to replace all the parts that go through t in the convex combination by a better distribution induced by $\sigma_1^{a^*}$ from t . ◀

The literature provides some *sufficient* conditions for preference relations to admit pure memoryless optimal strategies in one-player stochastic games (for instance, in [31]). Here, we obtain a full characterization when mixing is useless for \sqsubseteq (in particular, this is a full characterization for Borel real payoff functions), which can deal not only with memoryless strategies, but also with the more general AIFM strategies. It therefore provides a more fundamental understanding of preference relations for which AIFM strategies suffice or do not suffice. In particular, there are examples in which the known sufficient conditions are not verified even though pure memoryless strategies suffice (one such example is provided in [10]), and that is for instance where our characterization can help.

It is interesting to relate the concepts of monotony and selectivity to other properties from the literature to simplify the use of our characterization. For instance, if a real payoff function $f: C^\omega \rightarrow \mathbb{R}$ is *prefix-independent*,¹ then it is also X - Y - \mathcal{M} -monotone for any X , Y , and \mathcal{M} ; therefore, the sufficiency of pure AIFM strategies immediately reduces to analyzing selectivity.

6 Conclusion

We have studied stochastic games and gave an overview of desirable properties of preference relations that admit *pure arena-independent finite-memory optimal strategies*. Our analysis provides general tools to help study memory requirements in stochastic games, both with one player (Markov decision processes) and two players, and links both problems. It generalizes both work on deterministic games [33, 7] and work on stochastic games [34].

¹ A function $f: C^\omega \rightarrow \mathbb{R}$ is *prefix-independent* if for all $w \in C^*$, for all $w' \in C^\omega$, $f(ww') = f(w')$.

A natural question that remains unsolved is the link between memory requirements of a preference relation in deterministic and in stochastic games; our results can be called independently to study both problems, but do not describe a bridge to go from one to the other yet. Also, our results can only be used to show the optimality of *pure* strategies with some fixed memory, but in some cases, using *randomized* strategies allows for lesser memory requirements [16, 41]. Investigating whether extensions to our results dealing with randomized strategies hold would therefore be valuable.

References

- 1 Benjamin Aminof and Sasha Rubin. First-cycle games. *Inf. Comput.*, 254:195–216, 2017. doi:10.1016/j.ic.2016.10.008.
- 2 Christel Baier and Joost-Pieter Katoen. *Principles of model checking*. MIT Press, 2008.
- 3 Raphaël Berthon, Mickael Randour, and Jean-François Raskin. Threshold constraints with guarantees for parity objectives in Markov decision processes. In Ioannis Chatzigiannakis, Piotr Indyk, Fabian Kuhn, and Anca Muscholl, editors, *44th International Colloquium on Automata, Languages, and Programming, ICALP 2017, July 10-14, 2017, Warsaw, Poland*, volume 80 of *LIPIcs*, pages 121:1–121:15. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2017. doi:10.4230/LIPIcs.ICALP.2017.121.
- 4 Alessandro Bianco, Marco Faella, Fabio Mogavero, and Aniello Murano. Exploring the boundary of half-positionality. *Ann. Math. Artif. Intell.*, 62(1-2):55–77, 2011. doi:10.1007/s10472-011-9250-1.
- 5 Patricia Bouyer, Ulrich Fahrenberg, Kim Guldstrand Larsen, Nicolas Markey, and Jiri Srba. Infinite runs in weighted timed automata with energy constraints. In Franck Cassez and Claude Jard, editors, *Formal Modeling and Analysis of Timed Systems, 6th International Conference, FORMATS 2008, Saint Malo, France, September 15-17, 2008. Proceedings*, volume 5215 of *Lecture Notes in Computer Science*, pages 33–47. Springer, 2008. doi:10.1007/978-3-540-85778-5_4.
- 6 Patricia Bouyer, Piotr Hofman, Nicolas Markey, Mickael Randour, and Martin Zimmermann. Bounding average-energy games. In Javier Esparza and Andrzej S. Murawski, editors, *Foundations of Software Science and Computation Structures - 20th International Conference, FOSSACS 2017, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2017, Uppsala, Sweden, April 22-29, 2017, Proceedings*, volume 10203 of *Lecture Notes in Computer Science*, pages 179–195, 2017. doi:10.1007/978-3-662-54458-7_11.
- 7 Patricia Bouyer, Stéphane Le Roux, Youssouf Oualhadj, Mickael Randour, and Pierre Vandenhover. Games where you can play optimally with arena-independent finite memory. In Igor Konnov and Laura Kovács, editors, *31st International Conference on Concurrency Theory, CONCUR 2020, September 1-4, 2020, Vienna, Austria (Virtual Conference)*, volume 171 of *LIPIcs*, pages 24:1–24:22. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.CONCUR.2020.24.
- 8 Patricia Bouyer, Nicolas Markey, Mickael Randour, Kim G. Larsen, and Simon Laursen. Average-energy games. *Acta Informatica*, 55(2):91–127, 2018. doi:10.1007/s00236-016-0274-1.
- 9 Patricia Bouyer, Youssouf Oualhadj, Mickael Randour, and Pierre Vandenhover. Arena-independent finite-memory determinacy in stochastic games. *CoRR*, abs/2102.10104, 2021. arXiv:2102.10104.
- 10 Tomáš Brázdil, Václav Brozek, and Kousha Etessami. One-counter stochastic games. In Kamal Lodaya and Meena Mahajan, editors, *IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2010, December 15-18, 2010, Chennai, India*, volume 8 of *LIPIcs*, pages 108–119. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2010. doi:10.4230/LIPIcs.FSTTCS.2010.108.

- 11 Thomas Brihaye, Florent Delgrange, Youssef Oualhadj, and Mickael Randour. Life is random, time is not: Markov decision processes with window objectives. *Log. Methods Comput. Sci.*, 16(4), 2020. URL: <https://lmcs.episciences.org/6975>.
- 12 Véronique Bruyère, Emmanuel Filiot, Mickael Randour, and Jean-François Raskin. Meet your expectations with guarantees: Beyond worst-case synthesis in quantitative games. *Inf. Comput.*, 254:259–295, 2017. doi:10.1016/j.ic.2016.10.011.
- 13 Véronique Bruyère, Quentin Hautem, and Mickael Randour. Window parity games: an alternative approach toward parity games with time bounds. In Domenico Cantone and Giorgio Delzanno, editors, *Proceedings of the Seventh International Symposium on Games, Automata, Logics and Formal Verification, GandALF 2016, Catania, Italy, 14-16 September 2016*, volume 226 of *EPTCS*, pages 135–148, 2016. doi:10.4204/EPTCS.226.10.
- 14 Véronique Bruyère, Quentin Hautem, Mickael Randour, and Jean-François Raskin. Energy mean-payoff games. In Wan Fokkink and Rob van Glabbeek, editors, *30th International Conference on Concurrency Theory, CONCUR 2019, August 27-30, 2019, Amsterdam, the Netherlands*, volume 140 of *LIPICs*, pages 21:1–21:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPICs.CONCUR.2019.21.
- 15 Krishnendu Chatterjee. The complexity of stochastic Müller games. *Inf. Comput.*, 211:29–48, 2012. doi:10.1016/j.ic.2011.11.004.
- 16 Krishnendu Chatterjee, Luca de Alfaro, and Thomas A. Henzinger. Trading memory for randomness. In *1st International Conference on Quantitative Evaluation of Systems (QEST 2004), 27-30 September 2004, Enschede, The Netherlands*, pages 206–217. IEEE Computer Society, 2004. doi:10.1109/QEST.2004.1348035.
- 17 Krishnendu Chatterjee and Laurent Doyen. Energy parity games. *Theor. Comput. Sci.*, 458:49–60, 2012. doi:10.1016/j.tcs.2012.07.038.
- 18 Krishnendu Chatterjee and Laurent Doyen. Perfect-information stochastic games with generalized mean-payoff objectives. In Martin Grohe, Eric Koskinen, and Natarajan Shankar, editors, *Proceedings of the 31st Annual ACM/IEEE Symposium on Logic in Computer Science, LICS '16, New York, NY, USA, July 5-8, 2016*, pages 247–256. ACM, 2016. doi:10.1145/2933575.2934513.
- 19 Krishnendu Chatterjee, Laurent Doyen, Hugo Gimbert, and Thomas A. Henzinger. Randomness for free. In Petr Hliněný and Antonín Kucera, editors, *Mathematical Foundations of Computer Science 2010, 35th International Symposium, MFCS 2010, Brno, Czech Republic, August 23-27, 2010. Proceedings*, volume 6281 of *Lecture Notes in Computer Science*, pages 246–257. Springer, 2010. doi:10.1007/978-3-642-15155-2_23.
- 20 Krishnendu Chatterjee and Thomas A. Henzinger. A survey of stochastic ω -regular games. *J. Comput. Syst. Sci.*, 78(2):394–413, 2012. doi:10.1016/j.jcss.2011.05.002.
- 21 Krishnendu Chatterjee, Thomas A. Henzinger, and Nir Piterman. Generalized parity games. In Helmut Seidl, editor, *Foundations of Software Science and Computational Structures, 10th International Conference, FOSSACS 2007, Held as Part of the Joint European Conferences on Theory and Practice of Software, ETAPS 2007, Braga, Portugal, March 24-April 1, 2007, Proceedings*, volume 4423 of *Lecture Notes in Computer Science*, pages 153–167. Springer, 2007. doi:10.1007/978-3-540-71389-0_12.
- 22 Krishnendu Chatterjee, Marcin Jurdzinski, and Thomas A. Henzinger. Quantitative stochastic parity games. In J. Ian Munro, editor, *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms, SODA 2004, New Orleans, Louisiana, USA, January 11-14, 2004*, pages 121–130. SIAM, 2004. URL: <http://dl.acm.org/citation.cfm?id=982792.982808>.
- 23 Krishnendu Chatterjee, Joost-Pieter Katoen, Maximilian Weininger, and Tobias Winkler. Stochastic games with lexicographic reachability-safety objectives. In Shuvendu K. Lahiri and Chao Wang, editors, *Computer Aided Verification - 32nd International Conference, CAV 2020, Los Angeles, CA, USA, July 21-24, 2020, Proceedings, Part II*, volume 12225 of *Lecture Notes in Computer Science*, pages 398–420. Springer, 2020. doi:10.1007/978-3-030-53291-8_21.

- 24 Krishnendu Chatterjee, Zuzana Kretínská, and Jan Kretínský. Unifying two views on multiple mean-payoff objectives in Markov decision processes. *Log. Methods Comput. Sci.*, 13(2), 2017. doi:10.23638/LMCS-13(2:15)2017.
- 25 Krishnendu Chatterjee and Nir Piterman. Combinations of qualitative winning for stochastic parity games. In Wan J. Fokink and Rob van Glabbeek, editors, *30th International Conference on Concurrency Theory, CONCUR 2019, August 27-30, 2019, Amsterdam, the Netherlands*, volume 140 of *LIPIcs*, pages 6:1–6:17. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019. doi:10.4230/LIPIcs.CONCUR.2019.6.
- 26 Krishnendu Chatterjee, Mickael Randour, and Jean-François Raskin. Strategy synthesis for multi-dimensional quantitative objectives. *Acta Inf.*, 51(3-4):129–163, 2014. doi:10.1007/s00236-013-0182-6.
- 27 Taolue Chen, Vojtech Forejt, Marta Z. Kwiatkowska, Aistis Simaitis, Ashutosh Trivedi, and Michael Ummels. Playing stochastic games precisely. In Maciej Koutny and Irek Ulidowski, editors, *CONCUR 2012 - Concurrency Theory - 23rd International Conference, CONCUR 2012, Newcastle upon Tyne, UK, September 4-7, 2012. Proceedings*, volume 7454 of *Lecture Notes in Computer Science*, pages 348–363. Springer, 2012. doi:10.1007/978-3-642-32940-1_25.
- 28 Taolue Chen, Vojtech Forejt, Marta Z. Kwiatkowska, Aistis Simaitis, and Clemens Wiltsche. On stochastic games with multiple objectives. In Krishnendu Chatterjee and Jiri Sgall, editors, *Mathematical Foundations of Computer Science 2013 - 38th International Symposium, MFCS 2013, Klosterneuburg, Austria, August 26-30, 2013. Proceedings*, volume 8087 of *Lecture Notes in Computer Science*, pages 266–277. Springer, 2013. doi:10.1007/978-3-642-40313-2_25.
- 29 Anne Condon. The complexity of stochastic games. *Inf. Comput.*, 96(2):203–224, 1992. doi:10.1016/0890-5401(92)90048-K.
- 30 Florent Delgrange, Joost-Pieter Katoen, Tim Quatmann, and Mickael Randour. Simple strategies in multi-objective MDPs. In Armin Biere and David Parker, editors, *Tools and Algorithms for the Construction and Analysis of Systems - 26th International Conference, TACAS 2020, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2020, Dublin, Ireland, April 25-30, 2020, Proceedings, Part I*, volume 12078 of *Lecture Notes in Computer Science*, pages 346–364. Springer, 2020. doi:10.1007/978-3-030-45190-5_19.
- 31 Hugo Gimbert. Pure stationary optimal strategies in Markov decision processes. In Wolfgang Thomas and Pascal Weil, editors, *STACS 2007, 24th Annual Symposium on Theoretical Aspects of Computer Science, Aachen, Germany, February 22-24, 2007, Proceedings*, volume 4393 of *Lecture Notes in Computer Science*, pages 200–211. Springer, 2007. doi:10.1007/978-3-540-70918-3_18.
- 32 Hugo Gimbert and Edon Kelmendi. Submixing and shift-invariant stochastic games. *CoRR*, abs/1401.6575, 2014. arXiv:1401.6575.
- 33 Hugo Gimbert and Wiesław Zielonka. Games where you can play optimally without any memory. In Martín Abadi and Luca de Alfaro, editors, *CONCUR 2005 - Concurrency Theory, 16th International Conference, CONCUR 2005, San Francisco, CA, USA, August 23-26, 2005, Proceedings*, volume 3653 of *Lecture Notes in Computer Science*, pages 428–442. Springer, 2005. doi:10.1007/11539452_33.
- 34 Hugo Gimbert and Wiesław Zielonka. Pure and Stationary Optimal Strategies in Perfect-Information Stochastic Games with Global Preferences. Unpublished, 2009. URL: <https://hal.archives-ouvertes.fr/hal-00438359>.
- 35 Eryk Kopczyński. Half-positional determinacy of infinite games. In Michele Bugliesi, Bart Preneel, Vladimiro Sassone, and Ingo Wegener, editors, *Automata, Languages and Programming, 33rd International Colloquium, ICALP 2006, Venice, Italy, July 10-14, 2006, Proceedings, Part II*, volume 4052 of *Lecture Notes in Computer Science*, pages 336–347. Springer, 2006. doi:10.1007/11787006_29.
- 36 Eryk Kopczyński. *Half-positional Determinacy of Infinite Games*. PhD thesis, Warsaw University, 2008.

- 37 Stéphane Le Roux and Arno Pauly. Extending finite-memory determinacy to multi-player games. *Inf. Comput.*, 261(Part):676–694, 2018. doi:10.1016/j.ic.2018.02.024.
- 38 Stéphane Le Roux, Arno Pauly, and Mickael Randour. Extending finite-memory determinacy by boolean combination of winning conditions. In Sumit Ganguly and Paritosh K. Pandya, editors, *38th IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science, FSTTCS 2018, December 11-13, 2018, Ahmedabad, India*, volume 122 of *LIPIcs*, pages 38:1–38:20. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2018. doi:10.4230/LIPIcs.FSTTCS.2018.38.
- 39 Richard Mayr, Sven Schewe, Patrick Totzke, and Dominik Wojtczak. MDPs with energy-parity objectives. In *32nd Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2017, Reykjavik, Iceland, June 20-23, 2017*, pages 1–12. IEEE Computer Society, 2017. doi:10.1109/LICS.2017.8005131.
- 40 Richard Mayr, Sven Schewe, Patrick Totzke, and Dominik Wojtczak. Simple stochastic games with almost-sure energy-parity objectives are in NP and coNP. In Stefan Kiefer and Christine Tasson, editors, *Foundations of Software Science and Computation Structures - 24th International Conference, FOSSACS 2021, Held as Part of the European Joint Conferences on Theory and Practice of Software, ETAPS 2021, Luxembourg City, Luxembourg, March 27 - April 1, 2021, Proceedings*, volume 12650 of *Lecture Notes in Computer Science*, pages 427–447. Springer, 2021. doi:10.1007/978-3-030-71995-1_22.
- 41 Benjamin Monmege, Julie Parreaux, and Pierre-Alain Reynier. Reaching your goal optimally by playing at random with no memory. In Igor Konnov and Laura Kovács, editors, *31st International Conference on Concurrency Theory, CONCUR 2020, September 1-4, 2020, Vienna, Austria (Virtual Conference)*, volume 171 of *LIPIcs*, pages 26:1–26:21. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020. doi:10.4230/LIPIcs.CONCUR.2020.26.
- 42 Andrzej W. Mostowski. Regular expressions for infinite trees and a standard form of automata. In Andrzej Skowron, editor, *Computation Theory - Fifth Symposium, Zaborów, Poland, December 3-8, 1984, Proceedings*, volume 208 of *Lecture Notes in Computer Science*, pages 157–168. Springer, 1984. doi:10.1007/3-540-16066-3_15.
- 43 Martin J. Osborne. *An introduction to game theory*. Oxford University Press, 2004.
- 44 Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. Wiley Series in Probability and Statistics. Wiley, 1994. doi:10.1002/9780470316887.
- 45 Mickael Randour. Automated synthesis of reliable and efficient systems through game theory: A case study. In *Proc. of ECCS 2012*, Springer Proceedings in Complexity XVII, pages 731–738. Springer, 2013. doi:10.1007/978-3-319-00395-5_90.
- 46 Mickael Randour, Jean-François Raskin, and Ocan Sankur. Variations on the stochastic shortest path problem. In Deepak D’Souza, Akash Lal, and Kim Guldstrand Larsen, editors, *Verification, Model Checking, and Abstract Interpretation - 16th International Conference, VMCAI 2015, Mumbai, India, January 12-14, 2015. Proceedings*, volume 8931 of *Lecture Notes in Computer Science*, pages 1–18. Springer, 2015. doi:10.1007/978-3-662-46081-8_1.
- 47 Mickael Randour, Jean-François Raskin, and Ocan Sankur. Percentile queries in multi-dimensional Markov decision processes. *Formal Methods Syst. Des.*, 50(2-3):207–248, 2017. doi:10.1007/s10703-016-0262-7.
- 48 L. S. Shapley. Stochastic games. *Proceedings of the National Academy of Sciences*, 39(10):1095–1100, 1953. doi:10.1073/pnas.39.10.1095.
- 49 Wolfgang Thomas. Church’s problem and a tour through automata theory. In Arnon Avron, Nachum Dershowitz, and Alexander Rabinovich, editors, *Pillars of Computer Science, Essays Dedicated to Boris (Boaz) Trakhtenbrot on the Occasion of His 85th Birthday*, volume 4800 of *Lecture Notes in Computer Science*, pages 635–655. Springer, 2008. doi:10.1007/978-3-540-78127-1_35.
- 50 Yaron Velner, Krishnendu Chatterjee, Laurent Doyen, Thomas A. Henzinger, Alexander Moshe Rabinovich, and Jean-François Raskin. The complexity of multi-mean-payoff and multi-energy games. *Inf. Comput.*, 241:177–196, 2015. doi:10.1016/j.ic.2015.03.001.