

# Games Where You Can Play Optimally with Arena-Independent Finite Memory<sup>★</sup>

Patricia Bouyer<sup>1</sup>, Stéphane Le Roux<sup>1</sup>, Youssouf Oualhadj<sup>2</sup>, Mickael Randour<sup>3</sup>, and Pierre Vandenhover<sup>1,3</sup>

<sup>1</sup> LSV – CNRS & ENS Paris-Saclay, Université Paris-Saclay, France

<sup>2</sup> LACL – UPEC, France

<sup>3</sup> F.R.S.-FNRS & UMONS – Université de Mons, Belgium

**Abstract.** For decades, two-player (antagonistic) games on graphs have been a framework of choice for many important problems in theoretical computer science. A notorious one is controller synthesis, which can be rephrased through the game-theoretic metaphor as the quest for a winning strategy of the system in a game against its antagonistic environment. Depending on the specification, *optimal* strategies might be simple or quite complex, for example having to use (possibly infinite) memory. Hence, research strives to understand which settings allow for simple strategies.

In 2005, Gimbert and Zielonka [26] provided a complete characterization of preference relations (a formal framework to model specifications and game objectives) that admit *memoryless* optimal strategies for both players. In the last fifteen years however, practical applications have driven the community toward games with complex or multiple objectives, where memory — finite or infinite — is almost always required. Despite much effort, the exact frontiers of the class of preference relations that admit *finite-memory* optimal strategies still elude us.

In this work, we establish a complete characterization of preference relations that admit optimal strategies using *arena-independent* finite memory, generalizing the work of Gimbert and Zielonka to the finite-memory case. We also prove an equivalent to their celebrated corollary of great practical interest: if both players have optimal (arena-independent-)finite-memory strategies in all one-player games, then it is also the case in all two-player games. Finally, we pinpoint the boundaries of our results with regard to the literature: our work completely covers the case of arena-independent memory (e.g., multiple parity objectives, lower- and upper-bounded energy objectives), and paves the way to the arena-dependent case (e.g., multiple lower-bounded energy objectives).

## 1 Introduction

**Controller synthesis through the game-theoretic metaphor.** Two-player games on (finite) graphs are studied extensively, in particular for their application to controller synthesis for reactive systems (see, e.g., [28,36,6,2]). The seminal model is *antagonistic* (i.e., *zero-sum* if one chooses a quantitative view): player 1 ( $\mathcal{P}_1$ ) is seen as the system to control, player 2 ( $\mathcal{P}_2$ ) as its antagonistic environment, and the game models their interaction. Each vertex of the game graph (called *arena*) models a *state* of the system and belongs to one of the players. Players take turns moving a pebble from state to state according to the edges, each player choosing the destination whenever the pebble is on one of his states. These choices are made according to the *strategy* of the player, which, in general, might use memory (bounded or not) of the past moves to prescribe the next action.

The resulting infinite sequence of states, called *play*, represents the execution of the system. The objective of  $\mathcal{P}_1$  is to enforce a given *specification*, often encoded as a *winning condition* (i.e., a set of winning plays) or as a *payoff function* to maximize (i.e., a quantitative performance to optimize). This paradigm focuses on the *worst-case* performance of the system, hence  $\mathcal{P}_2$ 's goal is to prevent  $\mathcal{P}_1$  from achieving his objective.

<sup>★</sup> Research supported by F.R.S.-FNRS under Grant n° F.4520.18 (ManySynth), F.R.S.-FNRS mobility funding for scientific missions (Y. Oualhadj in UMONS, 2018), and ENS Paris-Saclay visiting professorship (M. Randour, 2019). Mickael Randour is an F.R.S.-FNRS Research Associate and Pierre Vandenhover is an F.R.S.-FNRS Research Fellow.

The goal of *synthesis* is thus to decide if  $\mathcal{P}_1$  has a *winning strategy*, i.e., one ensuring a given winning condition or guaranteeing a given payoff threshold, against all possible strategies of  $\mathcal{P}_2$ , and to build such a strategy efficiently if it exists.

Winning strategies are essentially *formal blueprints* for controllers to implement in practical applications. Therefore, the complexity of these strategies is of tremendous importance: the simpler the strategy, the easier and cheaper it will be to build the corresponding controller and maintain it. This explains why a lot of research effort is constantly put in identifying the exact complexity (in terms of memory and/or randomness) of strategies needed to play *optimally* (i.e., to the best of the player’s ability) for each specific class of games and objectives (e.g., [26,16,13,41,21,9,1,4,40,5,10]). Alongside the *practical interest* of this question lies the *theoretical puzzle*: understanding the underlying mechanisms and implicit properties of games that lead to “simple” strategies being sufficient. Given the numerous connections between two-player games and various branches of mathematics and computer science, this fundamental question has interest in its own right.

**Preference relations.** As hinted above, there are two prominent ways to formalize a game objective in the literature. The first one, dubbed *quantitative* and inspired by games in economics, is to use *payoff functions* mapping plays to numerical values, and to see  $\mathcal{P}_1$  as a maximizer player. This is for example the case of mean-payoff games [18]. The second one, called *qualitative*, is to define a *set of winning plays* — called *winning condition* — induced by some property, as in, e.g., parity games [19,42]. The two formalisms are strongly linked: the classical decision problem for quantitative games is to fix a payoff threshold and ask if  $\mathcal{P}_1$  has a strategy to guarantee it, essentially transforming the problem into a qualitative one (where the winning plays are all those achieving a payoff at least equal to the threshold). To define payoff functions or winning conditions, one often uses weights, priorities, colors, etc, on states or edges of the arena.

In this work, we walk in the footsteps of Gimbert and Zielonka [26]: we associate a *color* to each edge of our arenas, and we adopt the abstract formalism of *preference relations* over infinite sequences of colors (induced by plays). This general formalism permits to encode virtually all classical game objectives, both qualitative and quantitative, and lets us reason in a well-founded framework under minimal assumptions. See Example 3 for illustrations of classical objectives encoded as preference relations.

**Memoryless optimal strategies.** Remarkably, several canonical classes of games that have been around for decades and proved their usefulness over and over — e.g., mean-payoff [18], parity [19,42], or energy games [11] — share a desirable property: they all admit *memoryless optimal strategies for both players*. That is, for every strategy  $\sigma_i$  of  $\mathcal{P}_i$ , there is a strategy  $\sigma_i^{\text{ML}}$  which is *at least as good* (i.e., wins whenever  $\sigma_i$  wins or ensures at least the same payoff) and that uses no memory at all. Such a memoryless strategy always picks the same edge when in the same state, regardless of what happened before in the game.

Memoryless strategies are the simplest kind of strategies one can use in a turn-based game on a graph. Therefore, it is quite interesting that they suffice for objectives as rich as the ones we just discussed. Following this observation, a lot of effort has been put in understanding which games admit memoryless optimal strategies, and in identifying the exact frontiers of *memoryless determinacy*. Let us mention, non-exhaustively, works by Gimbert and Zielonka [25,26] (culminating in a complete characterization), Aminof and Rubin [1] (through the prism of first-cycle games), and Kopczyński [32] (half-positional determinacy). All these advances were built by identifying the common underlying mechanisms in ad hoc proofs for specific classes of games, and generalizing them to wide classes (e.g., the first-cycle games of Aminof and Rubin are inspired by the seminal paper of Ehrenfeucht and Mycielski on mean-payoff games [18]).

**Gimbert and Zielonka’s approach.** Arguably, the most important result in this direction is the *complete characterization* of preference relations admitting memoryless optimal strategies, established in [26], fifteen years ago. By complete characterization, we mean sufficient *and* necessary conditions on the preference relations.

This result can be stated as follows: a preference relation admits memoryless optimal strategies for both players on all arenas if and only if the relation (used by  $\mathcal{P}_1$ ) and its inverse (used by  $\mathcal{P}_2$ ) are *monotone* and *selective*. These concepts will be defined formally in Section 3.1, but let us give an intuition here. Roughly, a preference relation is monotone if it is stable under *prefix addition*: that is, given two sequences of colors such that one is *strictly* preferred to the other, it is impossible to reverse this order of preference by adding the same prefix to both sequences. Selectivity is similarly defined with regard to *cycle mixing*: if a preference

relation is selective, then, starting from two sequences of colors, it is impossible to create a third one by mixing the first two in such a way that the third one is strictly preferred to the first two. Observe that these elegant notions coincide with the natural intuition that memoryless strategies suffice if there is no interest in behaving differently in a state depending on what happened before.

FINITE  
GAMES

In addition to this complete characterization, Gimbert and Zielonka proved another great result, of high interest in practice [26, Corollary 7]: as a by-product of their approach, they obtain that if memoryless strategies suffice in all one-player games of  $\mathcal{P}_1$  and all one-player games of  $\mathcal{P}_2$ , they also suffice in all two-player games. Such a *lifting corollary* provides a neat and easy way to prove that a preference relation admits memoryless optimal strategies *without proving monotony and selectivity at all*: proving it in the two one-player subcases, which is generally much easier as it boils down to graph reasoning, and then lifting the result to the general two-player case through the corollary.

**The rise of memory.** Over the last decade, the increasing need to model complex specifications has shifted research toward games where multiple (quantitative and qualitative) objectives co-exist and interact, requiring the analysis of *interplay* and *trade-offs* between several objectives. Hence, a lot of effort is put in studying games where objectives are actually conjunctions of objectives, or even richer Boolean combinations. See for example [15] for combinations of parity, [12,16,31] for combinations of energy and parity, [41] for combinations of mean-payoff, [5,4] for combinations of energy and average-energy, [10] for combinations of energy and mean-payoff, [13] for combinations of total-payoff, or [13,9,7] for combinations of window objectives.

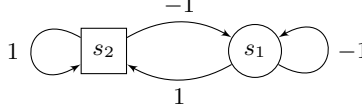
When considering such rich objectives, *memoryless strategies usually do not suffice*, and one has to use an amount of memory that can quickly become an obstacle to implementation (e.g., exponential memory) or that can prevent it completely (infinite memory). Establishing precise memory bounds for such general combinations of objectives is tricky and sometimes counterintuitive. For example, while energy games and mean-payoff games are inter-reducible in the single-objective setting, exponential-memory strategies are both sufficient and necessary for conjunctions of energy objectives [16,31] while *infinite-memory* strategies are required for conjunctions of mean-payoff ones [41].

A natural question arises: *which preference relations do admit finite-memory optimal strategies?* Surprisingly, whether an equivalent to Gimbert and Zielonka’s characterization could be obtained in the finite-memory case or not has remained an open question up to now. It is worth noticing that such an equivalent could be of tremendous help in practice, especially if a *lifting corollary* also holds: see for example [5,4,10], where proving that finite-memory strategies suffice in one-player games was fairly easy, in contrast to the high complexity of the two-player case — a lifting corollary could grant the two-player case for free!

Having said that, one has to hope that the following corollary can be established: “if *finite-memory* strategies suffice in all one-player games of  $\mathcal{P}_1$  and all one-player games of  $\mathcal{P}_2$ , they also suffice in all two-player games.” Unfortunately, this hope is but a delusion.

**Lifting corollary: a counterexample.** Consider games where the colors are integers, and the objective of  $\mathcal{P}_1$  is to create a play such that (a) the running sum of weights grows up to infinity (e.g., consider its  $\liminf$  to define it properly), *or* (b) this running sum of weights takes value zero infinitely often. As this defines a qualitative objective, the corresponding preference relation induces only two equivalence classes: *winning* and *losing* plays. The inverse relation, used by  $\mathcal{P}_2$ , is trivial to obtain. It is fairly easy to prove that  $\mathcal{P}_1$  always has finite-memory optimal strategies in his one-player games (i.e., games where he owns all states), and so does  $\mathcal{P}_2$  in his one-player games. See Appendix A for the details.

Now, consider the very simple two-player game depicted in Figure 1. First, observe that  $\mathcal{P}_1$  (circle) has an *infinite-memory* strategy to win:  $\mathcal{P}_1$  should keep track of the running sum of weights (which is unbounded, hence the need for infinite memory) and loop in  $s_1$  up to the point where this sum hits zero, when  $\mathcal{P}_1$  should then go to  $s_2$ . This strategy ensures victory because either  $\mathcal{P}_2$  always goes back to  $s_1$ , in which case (b) is satisfied; or  $\mathcal{P}_2$  eventually loops forever on  $s_2$ , in which case (a) is satisfied. It remains to argue that  $\mathcal{P}_1$  has no *finite-memory* winning strategy in this game. This can be done using a standard argument: whatever the amount of memory used by  $\mathcal{P}_1$ ,  $\mathcal{P}_2$  may loop in  $s_2$  long enough as to exceed the bound up to which  $\mathcal{P}_1$  can track the sum accurately; thus dooming  $\mathcal{P}_1$  to fail to reset the sum to zero in  $s_1$  infinitely often.



**Fig. 1.**  $\mathcal{P}_1$  (circle) needs infinite memory to win in this game (by always resetting the sum of weights to zero by looping long enough in  $s_1$  before going back to  $s_2$ ), whereas both players have finite-memory optimal strategies in all one-player games using the same preference relations.

This modest example proves that *Gimbert and Zielonka’s approach cannot work in full generality in the finite-memory case*, and for good reasons. Informally, in this case, the corollary breaks down because of (the absence of some sort of) monotony. In the case of *memoryless* strategies, as in [26],  $\mathcal{P}_1$  is already doomed in one-player games in the absence of monotony: two prefixes to distinguish — in order to play optimally — can be hardcoded as different paths leading to the same state in a game arena, as if they were chosen by  $\mathcal{P}_2$  in a two-player game. In the case of *finite-memory* strategies, however, the situation is different. In one-player games, the number of such paths that can be hardcoded in an arena is always bounded, hence finite memory might suffice to react, i.e., to keep track of which prefix is the current one and how to behave accordingly. However, in two-player games,  $\mathcal{P}_2$  might create an *infinite* number of prefixes to distinguish (using a cycle), thus requiring  $\mathcal{P}_1$  to use infinite memory to be able to do so. This is exactly what happens in the example above: in any one-player game, the largest sum that  $\mathcal{P}_1$  has to track is bounded, whereas  $\mathcal{P}_2$  can make this sum as large as he wants in two-player games.

**Our approach.** In a nutshell, we generalize Gimbert and Zielonka’s results — characterization *and* lifting corollary — to the case of *arena-independent* finite memory. That is, we encompass *all* situations where the amount of memory needed by the two players is *solely dependent on the preference relation* (e.g., colors, dimensions of weight vectors), and *not* on the game arena (i.e., number of edges/states). Let us take some classical examples to illustrate this notion.

- All memoryless-determined relations — studied in [26] — use arena-independent memory: the amount of memory required, *none*, is the same for all arenas.
- Combinations of parity objectives use arena-independent memory [15]: the amount of memory only depends on the number of objectives and the number of priorities — both parameters of the preference relation, not on the size of the arena.
- Lower- and upper-bounded energy objectives also use arena-independent memory (see, e.g., [3,5,4]): the memory only depends on the bounds and the weights — parameters of the preference relation, not on the size of the arena.
- On the contrary, combinations of lower-bounded energy objectives (with no upper bound) require arena-dependent memory [16,31]: the memory depends on the size of the arena in addition to the weights used in it. Such a setting falls outside the scope of our results.

This informal concept of arena-independent memory is transparent in our work: in all our results, we use *memory skeletons* — essentially Mealy machines without a next-action function (Section 2) — that suffice for *all* arenas, and that are at the basis of the strategies we build. A quick look at our main concepts (Section 3.1) and results (Section 3.2) suffices to grasp the formalism behind this intuition.

This restriction to arena-independent memory is natural given the counterexample to a general approach presented above. It is also important to note that it is not as restrictive as it may seem, as hinted by the examples above: we are *not restricted to constant memory* but to memory only depending on the *parameters of the preference relation* (or equivalently, objective), and not of the arena. This framework thus already encompasses many objectives from the literature, as well as possible extensions. We discuss this topic in more details in Section 6, where we provide a precise description of the frontiers of our results within the current research landscape.

Let us also highlight that the *arena-independent* case, which we solve here, is an exact equivalent to Gimbert and Zielonka’s results in the finite-memory case: the memoryless case is de facto arena-independent.

Therefore, this paper strictly generalizes [26] by allowing to study any arena-independent memory skeleton instead of the unique trivial one corresponding to memoryless strategies.

**Outline of our contributions.** Informally, our characterization can be stated as follows: given a preference relation and a memory skeleton  $\mathcal{M}$ , both players have optimal finite-memory strategies based on skeleton  $\mathcal{M}$  in all games if and only if the relation and its inverse are  $\mathcal{M}$ -monotone and  $\mathcal{M}$ -selective.

These last two concepts are keys to our approach. Intuitively, they correspond to Gimbert and Zielonka’s monotony and selectivity, *modulo a memory skeleton*. Recall that monotony and selectivity are related to stability of the preference relation with regard to prefix addition and cycle mixing, respectively. Our more general concepts of  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity serve the same purpose, but they only compare sequences of colors that are deemed equivalent by the memory skeleton. For the sake of illustration, take selectivity: it implies that one has no interest in mixing different cycles of the game arena. For its generalization, the memory skeleton is taken into account:  $\mathcal{M}$ -selectivity implies that one has no interest in mixing cycles of the game arena that are read as cycles on the same memory state in the skeleton  $\mathcal{M}$ .

Let us give a quick breakdown of our approach. In Section 2, we introduce all basic notions, including the memory skeletons, and we establish several technical results. We also discuss *optimal strategies* and *Nash equilibria*, their relationship, and their roles in our approach.

Section 3 is dedicated to our characterization, and consists of three parts. In Section 3.1, we introduce the concepts of  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity, cornerstones of our work. We also present two essential tools to establish the characterization: *prefix-covers* and *cyclic-covers* of arenas. Section 3.2 states formally our characterization (Theorem 15), as well as the corresponding lifting corollary (Corollary 18), from one-player to two-player games. We close this overview with an example of application, in Section 3.3.

The proof of the characterization (Theorem 15) is split in two. In Section 4, we establish the implication from (the sufficiency of) finite memory based on  $\mathcal{M}$  to  $\mathcal{M}$ -monotony (Theorem 19) and  $\mathcal{M}$ -selectivity (Theorem 20) of the preference relation. The main idea here is to build game arenas based on automata recognizing the languages involved in the two concepts, and to use the existence of finite-memory optimal strategies in these arenas to prove that  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity hold.

In Section 5, we prove the converse implication. We proceed in two steps, first establishing the existence of *memoryless* optimal strategies in “covered” arenas (Lemma 21 and Theorem 23), and then building on it to obtain the existence of *finite-memory* optimal strategies in general arenas (Corollary 25). The main technical tools we use are Nash equilibria and the aforementioned notions of prefix-covers and cyclic-covers.

We close the paper with a discussion of our characterization, presented in Section 6: we highlight some limitations and interesting features, compare its scope with the current research landscape, and sketch directions for future work.

**Technical overview.** Naturally, our technical approach is inspired by the one of Gimbert and Zielonka for the memoryless case [26], which can actually be rediscovered through our results using a trivial memory skeleton. Two of the most important challenges we had to overcome were:

1. establishing natural concepts of *monotony and selectivity modulo memory* that are exactly as powerful as required to maintain a complete characterization (i.e., sufficient and necessary conditions) in the finite-memory case;
2. circumventing the seemingly unavoidable *coupling between the memory skeleton and the arena* in the inductive argument needed to prove the implication from  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity to finite-memory optimal strategies — which we were able to do using our notions of prefix-covers and cyclic-covers.

All along our paper, we highlight the similarities and discrepancies between our work and Gimbert and Zielonka’s [26]. Whenever possible, we also go further, using weaker hypotheses and proving stronger results, along with addressing core problems left untouched in [26] — while they do have an important impact on the approach (e.g., the role of the zero-sum hypothesis). In that respect, we hope to shed a new light on the seminal results of [26] while generalizing them.

**Critical analysis.** Before jumping to the technical part of this work, let us take a step back and assess the place of our work in its larger line of research. The natural endgame is characterizing all preference

Proved  
finiteness



relations admitting finite-memory optimal strategies, including those using *arena-dependent* memory, and pinpointing the frontiers of application of the lifting corollary — that is, under which conditions is finite-memory determinacy preserved when going from one-player to two-player games?

The road is long from Gimbert and Zielonka’s characterization in the memoryless case [26] to such a general result, and this work is but a first step. We have already established that Gimbert and Zielonka’s approach cannot be fully transposed for finite memory. Our focus on *arena-independent* memory is a way to study the frontiers of this approach while providing an extension of practical interest. While it may seem limited at first, note that our framework already encompasses arguably rich classes of games such as, e.g., generalized parity games and fully-bounded energy games.

Let us stress that our result — relating a memory skeleton  $\mathcal{M}$  and preference relations for which this skeleton suffices — cannot be obtained by simply considering product arenas and invoking Gimbert and Zielonka’s result on memoryless determinacy [26]. While, of course, memoryless strategies on product arenas correspond to memoryfull strategies on original arenas (as we will formally establish in Lemma 4), invoking [26] requires to be able to quantify on *all* arenas, not only *product* arenas. Filling this gap is exactly the goal of this paper, and it is made possible through the new concepts we sketched above.

From a practical point of view, the equivalence result has limitations as it inherently uses the memory skeleton  $\mathcal{M}$ . At this point, our approach neither helps in finding an appropriate skeleton, nor in determining the minimal one; two highly interesting questions from a practical standpoint. Nonetheless, to advance toward answering these questions and to be able to find good skeletons automatically, one first has to understand their theoretical characteristics, which we do here as a necessary stepping stone. Focusing on applications, let us note that the equivalence result is often not the most suited tool: this is instead where the lifting corollary shines. As noted before, reasoning on one-player games (i.e., graphs) is generally much easier than in two-player games (e.g., [5,4,10]). Hence, a reasonably easy way to tackle practical cases is to find skeletons sufficient for  $\mathcal{P}_1$  and  $\mathcal{P}_2$  in their respective one-player games and to use our constructive result to build a skeleton that suffices for both in two-player games.

Finally, we believe it should be possible to generalize our approach to some extent to the *arena-dependent* case, through some *function* associating memory skeletons to arenas. Again, the previous example proves that this would not hold in full generality, but our hope is to establish conditions on this function (which is induced by the preference relation) under which the approach would hold. We leave this question open for now: this paper paves the way to this more general setting.

**Related work.** We already discussed the most important related papers, notably [26]. Let us highlight here some works where similar approaches have been considered to establish “meta-theorems” applying to general classes of games. First and foremost is the determinacy theorem by Martin that guarantees determinacy (without considering the complexity of strategies) for Borel winning conditions [34].

Following the same motivation as our work — the need to characterize (combinations of) objectives admitting finite-memory optimal strategies, Le Roux et al. [40] take another road: whereas our work permits to lift results from one-player games to two-player games, they provide a lifting from the single-objective case to the multi-objective one.

Our work focuses on *deterministic turn-based* two-player games. Sufficient conditions have been published for stochastic models but to the best of our knowledge, no complete characterization, even for the simplest case of Markov decision processes (e.g., [23]). Two unpublished articles contain interesting results on stochastic games [27,24], including an extension of Gimbert and Zielonka’s original work, by the same authors [27]. Whether part of our approach can be useful to tackle the finite-memory case in this context, or in richer contexts mixing games and stochastic models (e.g., [8]) is a question for future research. Some sufficient criteria, orthogonal to our approach, were studied for *concurrent* games in [38].

Finally, we recently discovered unpublished content in Kopczyński’s PhD thesis [33]. Kopczyński distinguishes *chromatic* memory (which corresponds to our definition of memory skeleton), and the more powerful *chaotic* memory, where transitions of the memory can depend on the actual edges of the arenas, rather than simply on the colors of the edges. Chaotic memory is thus intrinsically *arena-dependent*. Our notion of an arena being both prefix- and cyclic-covered by a memory skeleton  $\mathcal{M}$  is equivalent to a notion in [33, Definition 8.12], which defines that an arena *adheres to chromatic memory*  $\mathcal{M}$  if it is possible to assign a

state of  $\mathcal{M}$  to every state of the arena such that moving along the edges of the arena updates these memory states in a consistent way. Our definitions of *prefix-* and *cyclic-cover* can be seen as two distinct sides of this idea of *adherence*, which when added up, are actually equivalent to it.

**Acknowledgments.** We extend our warmest thanks to Mathieu Sassolas, for inspiring discussions that were essential in starting this work.

## 2 Preliminaries

**Automata and languages of colors.** Let  $C$  be an arbitrary set of *colors*.

We recall classical notions on automata on finite words. A *non-deterministic finite-state automaton (NFA)* is a tuple  $\mathcal{N} = (Q, B, \delta, Q_{\text{init}}, Q_{\text{fin}})$ , where  $Q$  is a finite set of *states*,  $B \subseteq C$  is a finite *alphabet* of colors,  $\delta \subseteq Q \times B \times Q$  is a set of *transitions*,  $Q_{\text{init}} \subseteq Q$  is a set of *initial* states, and  $Q_{\text{fin}} \subseteq Q$  is a set of *final* states. Given a state  $q \in Q$  and a word  $w \in B^*$ , we denote by  $\widehat{\delta}(q, w)$  the set of states that can be reached from  $q$  after reading  $w$ . Without loss of generality, we assume all NFA to be *coaccessible*, i.e., for all  $q \in Q$ , there exists  $w \in B^*$ , such that  $\widehat{\delta}(q, w) \cap Q_{\text{fin}} \neq \emptyset$ . Recall that NFA precisely recognize *regular languages*.

For any finite subset  $B \subseteq C$ , we denote by  $\text{Reg}(B)$  the set of all regular languages over  $B$ . Let  $\mathcal{R}(C) = \bigcup_{B \subseteq C, |B| < \infty} \text{Reg}(B)$ , that is, all the regular languages built over  $C$ .

Let  $K \subseteq C^*$  be a language of *finite* words. We denote by  $\text{Pref}(K)$  the set of all *prefixes* of the words in  $K$ . We define the set of *infinite* words  $[K] = \{w = c_1 c_2 \dots \in C^\omega \mid \forall n \geq 1, c_1 \dots c_n \in \text{Pref}(K)\}$ , which contains all infinite words for which every finite prefix is a prefix of a word in  $K$ . Intuitively, if  $K$  is regular, language  $[K]$  corresponds to all the infinite words that can be read on an automaton for  $K$ : we will formalize this in Lemma 2.

The following observation, already noted in [26], will come in handy too.

**Lemma 1.** *Let  $K_1, K_2 \subseteq C^*$ . Then  $[K_1 \cup K_2] = [K_1] \cup [K_2]$ .*

*Proof.* Let  $w \in [K_1 \cup K_2]$ . Every finite prefix of  $w$  is in  $\text{Pref}(K_1 \cup K_2)$ . Assume w.l.o.g. that infinitely many prefixes of  $w$  are in  $\text{Pref}(K_1)$ . This implies that *all* prefixes of  $w$  are in  $\text{Pref}(K_1)$  (intuitively, because there is a continuity in the prefix relation). Hence,  $w \in [K_1] \cup [K_2]$ .

Now, let  $w \in [K_1] \cup [K_2]$ . If  $w \in [K_1]$  (resp.  $[K_2]$ ), every finite prefix of  $w$  is in  $\text{Pref}(K_1)$  (resp.  $\text{Pref}(K_2)$ ), so in particular it is in  $\text{Pref}(K_1 \cup K_2)$ . Hence,  $w \in [K_1 \cup K_2]$ .  $\square$

**Arenas.** We consider two players: player 1 ( $\mathcal{P}_1$ ) and player 2 ( $\mathcal{P}_2$ ). An *arena* is a tuple  $\mathcal{A} = (S_1, S_2, E)$  such that  $S = S_1 \uplus S_2$  (disjoint union) is a finite set of *states* partitioned into states of  $\mathcal{P}_1$  ( $S_1$ ) and  $\mathcal{P}_2$  ( $S_2$ ), and  $E \subseteq S \times C \times S$  is a finite set of *edges*. Let  $\text{col}: E \rightarrow C$  be the projection of edges to *colors* and  $\widehat{\text{col}}$  its natural extension to *sequences* of edges. For an edge  $e \in E$ , we use  $\text{in}(e)$  and  $\text{out}(e)$  to denote its starting state and arrival state respectively, i.e.,  $e = (\text{in}(e), \text{col}(e), \text{out}(e))$ . We assume all arenas to be non-blocking, i.e., for all  $s \in S$ , there exists  $e \in E$  such that  $\text{in}(e) = s$ . For  $i \in \{1, 2\}$ , we call an arena  $\mathcal{A} = (S_1, S_2, E)$  a  $\mathcal{P}_i$ 's *one-player arena* if for all  $s \in S_{3-i}$ ,  $|\{e \in E \mid \text{in}(e) = s\}| = 1$  — that is,  $\mathcal{P}_{3-i}$  has no choice.

Let  $\text{Hists}(\mathcal{A}, s)$  denote the set of *histories* in  $\mathcal{A}$  from initial state  $s \in S$ , i.e., finite sequences of edges  $\rho = e_1 \dots e_n \in E^+$  such that  $\text{in}(e_1) = s$  and for all  $i$ ,  $1 \leq i < n$ ,  $\text{out}(e_i) = \text{in}(e_{i+1})$ . Let  $\text{Plays}(\mathcal{A}, s)$  denote the set of *plays* in  $\mathcal{A}$  from initial state  $s \in S$ , i.e., infinite sequences of edges  $\pi = e_1 e_2 \dots \in E^\omega$  such that  $\text{in}(e_1) = s$  and for all  $i \geq 1$ ,  $\text{out}(e_i) = \text{in}(e_{i+1})$ . We write  $\text{Hists}(\mathcal{A}, S')$  and  $\text{Plays}(\mathcal{A}, S')$  for the unions over subsets of initial states  $S' \subseteq S$ , and write  $\text{Hists}(\mathcal{A})$  and  $\text{Plays}(\mathcal{A})$  for the unions over all states of  $\mathcal{A}$ .

Let  $\rho = e_1 \dots e_n \in \text{Hists}(\mathcal{A})$  (resp.  $\pi = e_1 e_2 \dots \in \text{Plays}(\mathcal{A})$ ): we extend the operator  $\text{in}$  to histories (resp. plays) by identifying  $\text{in}(\rho)$  (resp.  $\text{in}(\pi)$ ) to  $\text{in}(e_1)$ . We proceed similarly for  $\text{out}$  and histories:  $\text{out}(\rho) = \text{out}(e_n)$ . For the sake of convenience, we consider that any set  $\text{Hists}(\mathcal{A}, s)$  contains the *empty history*  $\lambda_s$  such that  $\text{in}(\lambda_s) = \text{out}(\lambda_s) = s$ . We write  $\text{Hists}_i(\mathcal{A}, s)$  and  $\text{Hists}_i(\mathcal{A})$  for the subsets of histories  $\rho$  such that  $\text{out}(\rho) \in S_i$ ,  $i \in \{1, 2\}$ , i.e., histories whose last state belongs to  $\mathcal{P}_i$ .

For any set of histories  $H \subseteq \text{Hists}(\mathcal{A})$ , we write  $\widehat{\text{col}}(H)$  for its projection to colors, i.e.,  $\widehat{\text{col}}(H) = \{\widehat{\text{col}}(\rho) \mid \rho \in H\}$ . We do the same for sets of plays.

**Memory skeletons.** A *memory skeleton* is a tuple  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  where  $M$  is a *finite* set of *states*,  $m_{\text{init}} \in M$  is a fixed *initial* state and  $\alpha_{\text{upd}}: M \times C \rightarrow M$  is an *update function*. We write  $\widehat{\alpha_{\text{upd}}}$  for the natural extension of  $\alpha_{\text{upd}}$  to sequences of colors in  $C^*$ . Note that memory skeletons are deterministic and might have an *infinite* number of transitions, in contrast to NFA. We define the *trivial memory skeleton* with only one state as  $\mathcal{M}_{\text{triv}} = (M = \{m_{\text{init}}\}, m_{\text{init}}, \alpha_{\text{upd}}: \{m_{\text{init}}\} \times C \rightarrow \{m_{\text{init}}\})$ : it permits to formalize memoryless strategies [26] in our framework.

Let  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  be a memory skeleton. For  $m, m' \in M$ , we define the language  $L_{m, m'} = \{w \in C^* \mid \widehat{\alpha_{\text{upd}}}(m, w) = m'\}$  that contains all words that can be read from  $m$  to  $m'$  in  $\mathcal{M}$ .

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 two memory skeletons. We define their *product*  $\mathcal{M}^1 \otimes \mathcal{M}^2$  as the memory skeleton  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  obtained as follows:  $M = M^1 \times M^2$ ,  $m_{\text{init}} = (m_{\text{init}}^1, m_{\text{init}}^2)$ , and, for all  $m^1 \in M^1$ ,  $m^2 \in M^2$ ,  $c \in C$ ,  $\alpha_{\text{upd}}((m^1, m^2), c) = (\alpha_{\text{upd}}^1(m^1, c), \alpha_{\text{upd}}^2(m^2, c))$ . That is, the memories are updated in parallel when a color is read.

**Product arenas.** Let  $\mathcal{A} = (S_1, S_2, E)$  be an arena and  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  be a memory skeleton. We define their *product*  $\mathcal{A} \ltimes \mathcal{M}$  as the arena  $(S'_1, S'_2, E')$  where  $S'_1 = S_1 \times M$ ,  $S'_2 = S_2 \times M$ , and  $E' \subseteq S' \times C \times S'$ , with  $S' = S'_1 \uplus S'_2$ , is such that  $((s_1, m_1), c, (s_2, m_2)) \in E'$  if and only if  $(s_1, c, s_2) \in E$  and  $\alpha_{\text{upd}}(m_1, c) = m_2$ . That is, the memory is updated according to the colors of the edges in  $E$ . Note that even though  $\mathcal{M}$  might contain an infinite number of transitions since  $C$  might be infinite,  $\mathcal{A} \ltimes \mathcal{M}$  is always finite, as  $E$  is finite in  $\mathcal{A}$ . Since we assume arena  $\mathcal{A}$  is non-blocking, it is also the case of arena  $\mathcal{A} \ltimes \mathcal{M}$ .

**Arena induced by an NFA.** Let  $\mathcal{N} = (Q, B, \delta, Q_{\text{init}}, Q_{\text{fin}})$  be an NFA. We say that a state  $q \in Q$  is *essential* if there exists an infinite path in  $\mathcal{N}$  starting in  $q$ . Let  $Q_{\text{ess}} = \{q \in Q \mid q \text{ is essential}\}$ . We define the corresponding *one-player* arena  $\text{Arena}(\mathcal{N}) = (S_1 = Q_{\text{ess}}, S_2 = \emptyset, E \subseteq Q_{\text{ess}} \times B \times Q_{\text{ess}})$ , where  $e = (q, c, q') \in E$  if  $(q, c, q') \in \delta$ . Intuitively,  $\text{Arena}(\mathcal{N})$  transforms  $\mathcal{N}$  into a *non-blocking* arena thanks to the restriction to essential states.

We may now state formally the link between  $[K]$  and the underlying automaton for  $K$ . Our proof is similar to [26, Lemma 4].

**Lemma 2.** *Let  $\mathcal{N} = (Q, B, \delta, Q_{\text{init}}, Q_{\text{fin}})$  be a (coaccessible) NFA recognizing the regular language  $K \subseteq C^*$ . Let  $Q'_{\text{init}} = Q_{\text{init}} \cap Q_{\text{ess}}$ . The following equality holds:*

$$[K] = \widehat{\text{col}}(\text{Plays}(\text{Arena}(\mathcal{N}), Q'_{\text{init}})).$$

*In particular,  $[K]$  is non-empty if and only if there exists an essential initial state in  $\mathcal{N}$ .*

*Proof.* If  $Q'_{\text{init}}$  is empty, the equality trivially holds:  $\widehat{\text{col}}(\text{Plays}(\text{Arena}(\mathcal{N}), Q'_{\text{init}}))$  and  $[K]$  are both empty. Hence, from now on, we assume  $Q'_{\text{init}} \neq \emptyset$ .

We start with the left-to-right inclusion. Let  $w = c_1 c_2 \dots \in [K]$ . We first prove that for all  $n \geq 1$ , it holds that

$$c_1 \dots c_n \in \widehat{\text{col}}(\text{Hists}(\text{Arena}(\mathcal{N}), Q'_{\text{init}})).$$

We assume on the contrary that there exists  $n \geq 1$  such that  $c_1 \dots c_n \notin \widehat{\text{col}}(\text{Hists}(\text{Arena}(\mathcal{N}), Q'_{\text{init}}))$ . As  $\text{Arena}(\mathcal{N})$  is a restriction of the states of  $\mathcal{N}$  to  $Q_{\text{ess}}$ , this means that no matter how  $c_1 \dots c_n$  is read on  $\mathcal{N}$ , it goes through a state in  $Q \setminus Q_{\text{ess}}$ . As there is no infinite path from these states, this contradicts that  $w \in [K]$ ; there cannot be arbitrarily long prefixes starting with  $c_1 \dots c_n$ .

We now use the property that we have just proved along with König's lemma to show that  $w \in \widehat{\text{col}}(\text{Plays}(\text{Arena}(\mathcal{N}), Q'_{\text{init}}))$ . We build a forest of trees  $\mathcal{F}$ . The vertices of  $\mathcal{F}$  are paths  $\rho \in \text{Hists}(\text{Arena}(\mathcal{N}), Q'_{\text{init}})$  such that  $\widehat{\text{col}}(\rho)$  is a prefix of  $w$  and  $\text{in}(\rho) \in Q'_{\text{init}}$ . For every  $q \in Q'_{\text{init}}$ , there is one tree in  $\mathcal{F}$  whose root is the empty path  $\lambda_q$ . There is a transition from a vertex  $\rho$  to a vertex  $\rho'$  if there exists  $e \in E$  such that  $\rho \cdot e = \rho'$ . As there is at least one vertex for each prefix  $c_1 \dots c_n$ , (at least) one of the trees of  $\mathcal{F}$  must be infinite. Moreover,  $\mathcal{F}$  is finitely branching. By König's lemma, we obtain that there must be an infinite path  $\pi$  starting from a root  $\lambda_q$  for some  $q \in Q'_{\text{init}}$ . By construction,  $\widehat{\text{col}}(\pi) = w$ , so  $w \in \widehat{\text{col}}(\text{Plays}(\text{Arena}(\mathcal{N}), Q'_{\text{init}}))$ .

We now prove the right-to-left inclusion. Let  $\pi = e_1 e_2 \dots \in \text{Plays}(\text{Arena}(\mathcal{N}), Q'_{\text{init}})$ . For  $n \geq 1$ , the word  $\widehat{\text{col}}(e_1 \dots e_n)$  is the color of a path in  $\mathcal{N}$ , since every edge of  $\text{Arena}(\mathcal{N})$  corresponds to a transition of  $\mathcal{N}$ . As



$\mathcal{N}$  is coaccessible, there is a path in  $\mathcal{N}$  from the state corresponding to  $\text{out}(e_n)$  to a final state in  $Q_{\text{fin}}$ . Thus, the word  $\widehat{\text{col}}(e_1 \dots e_n)$  is a prefix of an accepted word of  $\mathcal{N}$ , i.e., a prefix of a word in  $K$ ; as this holds for all  $n \geq 1$ , we obtain that  $\widehat{\text{col}}(\pi) \in [K]$ .  $\square$

**Strategies.** A *strategy*  $\sigma_i$  for  $\mathcal{P}_i$ ,  $i \in \{1, 2\}$ , on arena  $\mathcal{A} = (S_1, S_2, E)$ , is a function  $\sigma_i: \text{Hists}_i(\mathcal{A}) \rightarrow E$  such that for all  $\rho \in \text{Hists}_i(\mathcal{A})$ ,  $\text{in}(\sigma_i(\rho)) = \text{out}(\rho)$ . Let  $\Sigma_i(\mathcal{A})$  be the set of all strategies of  $\mathcal{P}_i$  on  $\mathcal{A}$ .

A *finite-memory strategy*  $\sigma_i$  is a strategy that can be encoded as a *Mealy machine*, i.e., a memory skeleton  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  with transitions over a *finite subset of colors*  $B \subseteq C$ , enriched with a *next-action function*  $\alpha_{\text{nxt}}: M \times S_i \rightarrow E$  such that for all  $m \in M$ ,  $s \in S_i$ ,  $\text{in}(\alpha_{\text{nxt}}(m, s)) = s$ . Given a Mealy machine  $\Gamma_{\sigma_i} = (\mathcal{M}, \alpha_{\text{nxt}})$ , strategy  $\sigma_i$  is defined as follows:

- $\forall s \in S_i, \sigma_i(\lambda_s) = \alpha_{\text{nxt}}(m_{\text{init}}, s),$
- $\forall \rho \cdot e \in \text{Hists}_i(\mathcal{A}), e \in E, \sigma_i(\rho \cdot e) = \alpha_{\text{nxt}}\left(\widehat{\alpha_{\text{upd}}}\left(m_{\text{init}}, \widehat{\text{col}}(\rho \cdot e)\right), \text{out}(e)\right).$

We denote by  $\Sigma_i^{\text{FM}}(\mathcal{A})$  the set of all finite-memory strategies of  $\mathcal{P}_i$  on  $\mathcal{A}$ . We say that a strategy  $\sigma_i \in \Sigma_i^{\text{FM}}(\mathcal{A})$  is *based on memory skeleton*  $\mathcal{M}$  if it can be encoded as a Mealy machine  $\Gamma_{\sigma_i} = (\mathcal{M}, \alpha_{\text{nxt}})$ , as above. We always implicitly assume that strategies of  $\Sigma_i^{\text{FM}}(\mathcal{A})$  are built by restricting the transitions of their skeleton  $\mathcal{M}$  to the actual subset of colors appearing in  $\mathcal{A}$ . A strategy  $\sigma_i$  is *memoryless* if it is a function  $\sigma_i: S_i \rightarrow E$ , or equivalently, if it is based on the trivial memory skeleton  $\mathcal{M}_{\text{triv}}$ . We denote by  $\Sigma_i^{\text{ML}}(\mathcal{A})$  the set of all memoryless strategies of  $\mathcal{P}_i$  on  $\mathcal{A}$ .

We denote by  $\text{Plays}(\mathcal{A}, s, \sigma_i)$  the set of plays *consistent* with a strategy  $\sigma_i$  of  $\mathcal{P}_i$  from an initial state  $s$ , i.e., all plays  $\pi = e_1 e_2 \dots \in \text{Plays}(\mathcal{A}, s)$  such that for all prefixes  $\rho = e_1 \dots e_n$ ,  $\text{out}(\rho) \in S_i \implies \sigma_i(\rho) = e_{n+1}$ . We write  $\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2)$  for the singleton set containing the unique play consistent with a couple of strategies for the two players. We use similar notations for histories.

**Preference relations.** Let  $\sqsubseteq$  be a total preorder on  $C^\omega$ , called *preference relation*. We consider *antagonistic* games, where the objective of  $\mathcal{P}_1$  is to create the best possible play with regard to  $\sqsubseteq$  whereas the objective of  $\mathcal{P}_2$  is to obtain the worst possible one. That is,  $\mathcal{P}_2$  uses the inverse relation  $\sqsubseteq^{-1}$ . This corresponds to *zero-sum* games when using a quantitative framework.

Given  $w, w' \in C^\omega$ , we write  $w \sqsubset w'$  if we have  $\neg(w' \sqsubseteq w)$  since the preorder is total. We extend the relation  $\sqsubseteq$  to subsets of  $C^\omega$  as follows: for  $W, W' \subseteq C^\omega$ ,

$$W \sqsubseteq W' \iff \forall w \in W, \exists w' \in W', w \sqsubseteq w'.$$

We also write

$$W \sqsubset W' \iff \exists w' \in W', \forall w \in W, w \sqsubset w'.$$

Note that  $W \sqsubset W'$  if and only if  $\neg(W' \sqsubseteq W)$ .

We sometimes compare words  $w \in C^\omega$  with languages  $K \subseteq C^\omega$ , by simply identifying word  $w$  to its singleton language  $\{w\}$ .

**Games.** A *game* is a tuple  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  where  $\mathcal{A}$  is an arena and  $\sqsubseteq$  is a preference relation. As discussed in Section 1, all the classical objectives from the literature (both qualitative and quantitative) can be expressed in the general framework of preference relations. For  $i \in \{1, 2\}$ , a  $\mathcal{P}_i$ 's *one-player game* is a game  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  such that  $\mathcal{A}$  is a  $\mathcal{P}_i$ 's one-player arena.

*Example 3.* There are two prominent ways to formalize game objectives in the literature: through *payoff functions* and through *winning conditions*. We take an example of each.

First, consider (lim inf) *mean-payoff* games [18]. In this setting, colors are integers, i.e.,  $C = \mathbb{Z}$ , and the goal of  $\mathcal{P}_1$  is to create a play  $\pi$ , with  $w = \widehat{\text{col}}(\pi) = c_1 c_2 \dots$ , maximizing the following *payoff function*:

$$\text{MP}(w) = \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n c_i.$$

Such a payoff function induces a natural preference relation  $\sqsubseteq_{\text{MP}}$  between sequences of colors as follows: for all  $w, w' \in C^\omega$ ,  $w \sqsubseteq_{\text{MP}} w'$  if and only if  $\text{MP}(w) \leq \text{MP}(w')$ . Such quantitative games are *zero-sum*, hence  $\mathcal{P}_2$  uses the natural inverse relation  $\sqsubseteq_{\text{MP}}^{-1}$ : he is a minimizer player in the payoff formulation of these games.

Second, consider *reachability* games (e.g., [20]). In this setting, only two colors are needed: one for edges in the *target set*, and one for the other edges. Let us use  $\top$  and  $\perp$  respectively to color these two sets of edges, i.e.,  $C = \{\top, \perp\}$ . Then, the *winning condition* can be simply written as  $W = \{w \in C^\omega \mid \top \in w\} \subsetneq C^\omega$ . In such games, the goal of  $\mathcal{P}_1$  is to create a play  $\pi$  such that  $\widehat{\text{col}}(\pi) \in W$ , called *winning play*. Defining a corresponding preference relation  $\sqsubseteq_{\text{reach}}$  is straightforward: for all  $w, w' \in C^\omega$ ,  $w \sqsubseteq_{\text{reach}} w'$  if and only if  $w' \in W$  and  $w \notin W$ . That is,  $\sqsubseteq_{\text{reach}}$  defines two equivalence classes: losing and winning plays. This qualitative setting is *antagonistic*, hence  $\mathcal{P}_2$  uses the inverse relation  $\sqsubseteq_{\text{reach}}^{-1}$ : his winning condition is  $\overline{W} = C^\omega \setminus W$  in the classical formulation of these games.

As explained in Section 1, quantitative games are often reduced to qualitative ones by fixing a *threshold* to achieve.  $\triangleleft$

**Optimal strategies.** Let  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  be a game on arena  $\mathcal{A} = (S_1, S_2, E)$ . Given a  $\mathcal{P}_i$ -strategy  $\sigma_i \in \Sigma_i(\mathcal{A})$  and a state  $s \in S$ , we define

$$\begin{aligned} \text{UCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma_i) &= \{w \in C^\omega \mid \exists \sigma_{3-i} \in \Sigma_{3-i}(\mathcal{A}), \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2)) \sqsubseteq w\}, \\ \text{DCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma_i) &= \{w \in C^\omega \mid \exists \sigma_{3-i} \in \Sigma_{3-i}(\mathcal{A}), w \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2))\}. \end{aligned}$$

Note that  $\text{DCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma_i) = \text{UCol}_{\sqsubseteq^{-1}}(\mathcal{A}, s, \sigma_i)$ . Intuitively,  $\text{UCol}_{\sqsubseteq}$  and  $\text{DCol}_{\sqsubseteq}$  represent the *upward* and *downward closures* of sequences of colors (consistent with a strategy) with respect to the preference relation.

Taking the standpoint of  $\mathcal{P}_1$ , we say that a strategy  $\sigma_1 \in \Sigma_1(\mathcal{A})$  is *at least as good as* a strategy  $\sigma'_1 \in \Sigma_1(\mathcal{A})$  from a state  $s \in S$  if

$$\text{UCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma_1) \subseteq \text{UCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma'_1).$$

Intuitively,  $\sigma_1$  is at least as good as  $\sigma'_1$  if the “worst-case” plays consistent with  $\sigma_1$  are at least as good as the ones consistent with  $\sigma'_1$ . The  $\text{UCol}$  operator is useful to define this notion properly even in the case where there is no “worst-case” play for a strategy (i.e., if the infimum used in the classical quantitative setting is not reached). Similar notions have been used before, e.g., in [37].

Symmetrically, for  $\mathcal{P}_2$ , we say that a strategy  $\sigma_2 \in \Sigma_2(\mathcal{A})$  is at least as good as a strategy  $\sigma'_2 \in \Sigma_2(\mathcal{A})$  from a state  $s \in S$  if

$$\text{DCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma_2) \subseteq \text{DCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma'_2).$$

Now, we say that a strategy  $\sigma_i \in \Sigma_i(\mathcal{A})$  of  $\mathcal{P}_i$  is *optimal* from a state  $s \in S$ , aka *s-optimal*, if it is at least as good as every other strategy  $\sigma'_i \in \Sigma_i(\mathcal{A})$  from  $s$ . We extend this notation to subsets of states in the natural way, and we say that a strategy  $\sigma_i$  is *uniformly-optimal* if it is *S-optimal*.

The goal of our paper is to characterize the preference relations that admit *uniformly-optimal finite-memory (UFM) strategies based on a given skeleton*  $\mathcal{M}$  in all arenas. We also discuss the simpler case of *uniformly-optimal memoryless (UML) strategies*, which corresponds to the subset of preference relations studied by Gimbert and Zielonka [26], using the trivial skeleton  $\mathcal{M}_{\text{triv}}$ .

In that respect, the following link is important to observe.

**Lemma 4.** *Let  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  be a game on arena  $\mathcal{A} = (S_1, S_2, E)$ . Let  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  be a memory skeleton and let  $\sigma_i \in \Sigma_i^{\text{FM}}(\mathcal{A})$  be a finite-memory strategy encoded by the Mealy machine  $\Gamma_{\sigma_i} = (\mathcal{M}, \alpha_{\text{next}})$ . Then,  $\sigma_i$  is a UFM strategy in  $\mathcal{G}$  if and only if  $\alpha_{\text{next}}$  corresponds to an  $(S \times \{m_{\text{init}}\})$ -optimal memoryless strategy in  $\mathcal{G}' = (\mathcal{A} \ltimes \mathcal{M}, \sqsubseteq)$ .*

*Proof.* We first aim to define a bijection  $\mathcal{H}: \text{Hists}(\mathcal{A}) \rightarrow \text{Hists}(\mathcal{A} \ltimes \mathcal{M}, S \times \{m_{\text{init}}\})$ . Let  $\rho = e_1 \dots e_n \in \text{Hists}(\mathcal{A})$ , with  $e_j = (s_j, c_j, s_{j+1})$ . We set  $m_1 = m_{\text{init}}$ , and for  $2 \leq j \leq n$ ,  $m_j = \alpha_{\text{upd}}(m_{j-1}, c_j)$ . We define  $e'_j = ((s_j, m_j), c_j, (s_{j+1}, m_{j+1}))$ , and  $\mathcal{H}(\rho) = e'_1 \dots e'_n$ . Notice that  $\widehat{\text{col}}(\mathcal{H}(\rho)) = \widehat{\text{col}}(\rho)$ . Furthermore,  $\mathcal{H}$  is bijective; as the initial state of the memory  $m_{\text{init}}$  is fixed and the memory skeleton is deterministic, the memory states added to  $\rho$  to obtain  $\mathcal{H}(\rho)$  are uniquely determined.

We now show that there is a correspondence between strategies of  $\Sigma_i(\mathcal{A})$  and strategies of  $\Sigma_i(\mathcal{A} \ltimes \mathcal{M})$ : intuitively, augmenting the arena with the skeleton allows some strategies to be played using less memory, but does not fundamentally change each player's possibilities. We define a function  $f: \Sigma_i(\mathcal{A}) \rightarrow \Sigma_i(\mathcal{A} \ltimes \mathcal{M})$ . For  $\tau_i \in \Sigma_i(\mathcal{A})$  and  $\rho' \in \text{Hists}_i(\mathcal{A} \ltimes \mathcal{M})$  with  $\text{in}(\rho') \in S \times \{m_{\text{init}}\}$  and  $\text{out}(\rho') = (s, m) \in S_i \times M$ , if  $\tau_i(\mathcal{H}^{-1}(\rho')) = (s, c, s')$ , we define  $f(\tau_i)(\rho') = ((s, m), c, (s', \alpha_{\text{upd}}(m, c)))$ . The histories induced by strategies  $\tau_i$  and  $f(\tau_i)$  correspond: if  $\rho = \mathcal{H}^{-1}(\rho')$ , then we have

$$\mathcal{H}(\rho \cdot \tau_i(\rho)) = \rho' \cdot (f(\tau_i)(\rho')). \quad (1)$$

We are only interested in the behavior of strategies of  $\Sigma_i(\mathcal{A} \ltimes \mathcal{M})$  on histories  $\rho'$  with  $\text{in}(\rho') \in S \times \{m_{\text{init}}\}$  (in what follows, we will only consider histories and plays starting in such states). If we restrict the image of  $f$  to the set of strategies  $\tau'_i: \text{Hists}_i(\mathcal{A} \ltimes \mathcal{M}, S \times \{m_{\text{init}}\}) \rightarrow E'$ , then  $f$  is a bijection.

Consider the next-action function  $\alpha_{\text{next}}$  of strategy  $\sigma_i$ . Formally, we have to transform  $\alpha_{\text{next}}$  into a proper memoryless strategy in  $\Sigma_i^{\text{ML}}(\mathcal{A} \ltimes \mathcal{M})$ . This can be done through the bijection  $f$ , yielding the memoryless strategy  $\alpha'_{\text{next}} = f(\sigma_i)$ , which corresponds to  $\alpha_{\text{next}}$  interpreted over the product arena, and well-defined for all histories starting in  $S \times \{m_{\text{init}}\}$ .

We now show a second fact related to  $f$ : we have that for all  $s \in S$ , for all  $\tau_1 \in \Sigma_1(\mathcal{A})$ ,  $\tau_2 \in \Sigma_2(\mathcal{A})$ ,

$$\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \tau_1, \tau_2)) = \widehat{\text{col}}(\text{Plays}(\mathcal{A} \ltimes \mathcal{M}, (s, m_{\text{init}}), f(\tau_1), f(\tau_2))). \quad (2)$$

This can easily be proved by induction using Equation (1). Indeed, at each step, both strategy  $\tau_1$  (resp.  $\tau_2$ ) and strategy  $f(\tau_1)$  (resp.  $f(\tau_2)$ ) pick an edge with the same color.

We finish the proof assuming that  $\sigma_i = \sigma_1 \in \Sigma_1^{\text{FM}}(\mathcal{A})$ ; the proof is symmetric for  $\mathcal{P}_2$ . Equation (2) implies that for all  $s \in S$ ,  $\tau_1 \in \Sigma_1(\mathcal{A})$ ,

$$\begin{aligned} \text{UCol}_{\sqsubseteq}(\mathcal{A}, s, \tau_1) &= \{w \in C^\omega \mid \exists \tau_2 \in \Sigma_2(\mathcal{A}), \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \tau_1, \tau_2)) \sqsubseteq w\} \\ &= \{w \in C^\omega \mid \exists \tau_2 \in \Sigma_2(\mathcal{A}), \widehat{\text{col}}(\text{Plays}(\mathcal{A} \ltimes \mathcal{M}, (s, m_{\text{init}}), f(\tau_1), f(\tau_2))) \sqsubseteq w\} \\ &= \{w \in C^\omega \mid \exists \tau'_2 \in \Sigma_2(\mathcal{A} \ltimes \mathcal{M}), \widehat{\text{col}}(\text{Plays}(\mathcal{A} \ltimes \mathcal{M}, (s, m_{\text{init}}), f(\tau_1), \tau'_2)) \sqsubseteq w\} \\ &= \text{UCol}_{\sqsubseteq}(\mathcal{A} \ltimes \mathcal{M}, (s, m_{\text{init}}), f(\tau_1)). \end{aligned} \quad (3)$$

where the penultimate equality uses that the aforementioned restriction of  $f$  is bijective.

Using Equation (3), we can obtain that a strategy  $\tau_1$  is uniformly-optimal in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  if and only if  $f(\tau_1) \in \Sigma_1(\mathcal{A} \ltimes \mathcal{M})$  is  $(S \times \{m_{\text{init}}\})$ -optimal in  $\mathcal{G}' = (\mathcal{A} \ltimes \mathcal{M}, \sqsubseteq)$ . In particular,  $\sigma_1$  is uniformly-optimal in  $\mathcal{G}$  if and only if  $f(\sigma_1) = \alpha'_{\text{next}}$  is  $(S \times \{m_{\text{init}}\})$ -optimal in  $\mathcal{G}'$ .  $\square$

**Nash equilibria.** We use Nash equilibria [35] as tools to establish the existence of optimal strategies in some of our proofs. Let  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  be a game on arena  $\mathcal{A} = (S_1, S_2, E)$ . Formally, a *Nash equilibrium* (NE) from a state  $s \in S$  is a couple of strategies  $(\sigma_1, \sigma_2) \in \Sigma_1(\mathcal{A}) \times \Sigma_2(\mathcal{A})$  such that, for all  $\sigma'_1 \in \Sigma_1(\mathcal{A})$ ,  $\sigma'_2 \in \Sigma_2(\mathcal{A})$ ,

$$\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma'_1, \sigma_2)) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2)) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma'_2)). \quad (4)$$

Similarly to optimal strategies, we call an NE *uniform* if it is an NE from all states  $s \in S$ .

It is worth taking a moment to discuss the link between *optimal strategies* and *Nash equilibria* in our specific context of *antagonistic* games. Both notions seem closely related, and indeed, in [26], Gimbert and Zielonka did choose Equation (4) — i.e., the definition of a Nash equilibrium — as their definition of a *pair of optimal strategies*. This could lead the reader to believe that both notions coincide. However, they do not in full generality, as we discuss in the following.

Additionally, defining optimality for a *pair* of strategies gives rise to difficulties as one naturally wants to reason about optimal strategies of a player without talking about the (possibly optimal) strategy of its adversary. Our definition of *optimal strategy* has the advantage of giving a clear and precise definition that does not involve the standpoint of the adversary.

As stated before, the ultimate goal of our paper is to characterize preference relations that admit *finite-memory optimal strategies*, but Nash equilibria will serve as tools in our endeavor. Let us establish two interesting properties of Nash equilibria *in antagonistic games*.

First, it is possible to mix different Nash equilibria.

**Lemma 5.** *Let  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  be a game on arena  $\mathcal{A} = (S_1, S_2, E)$ , and let  $s \in S$  be a state. Let  $(\sigma_1^a, \sigma_2^a)$  and  $(\sigma_1^b, \sigma_2^b) \in \Sigma_1(\mathcal{A}) \times \Sigma_2(\mathcal{A})$  be two Nash equilibria from  $s$ . Then,  $(\sigma_1^a, \sigma_2^b)$  is also a Nash equilibrium from  $s$ .*

*Proof.* We need to prove that for all  $\sigma'_1 \in \Sigma_1(\mathcal{A})$ ,  $\sigma'_2 \in \Sigma_2(\mathcal{A})$ ,

$$\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma'_1, \sigma_2^b)) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^a, \sigma_2^b)) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^a, \sigma'_2)). \quad (5)$$

Since  $(\sigma_1^a, \sigma_2^a)$  is an NE, we know that

$$\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^b, \sigma_2^a)) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^a, \sigma_2^a)) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^a, \sigma_2^b)),$$

instantiating  $\sigma'_1$  and  $\sigma'_2$  to  $\sigma_1^b$  and  $\sigma_2^b$  respectively in Equation (4). Similarly, since  $(\sigma_1^b, \sigma_2^b)$  is an NE, we know that

$$\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^a, \sigma_2^b)) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^b, \sigma_2^b)) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^b, \sigma_2^a)),$$

instantiating  $\sigma'_1$  and  $\sigma'_2$  to  $\sigma_1^a$  and  $\sigma_2^a$  respectively in Equation (4).

Now it is easy to see from the last two lines that all six sequences of colors are equivalent under  $\sqsubseteq$  as the inequalities form a cycle. Hence, since Equation (4) holds for  $(\sigma_1^a, \sigma_2^a)$  and  $(\sigma_1^b, \sigma_2^b)$ , and since  $\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^a, \sigma_2^b))$  is  $\sqsubseteq$ -equivalent to both  $\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^a, \sigma_2^a))$  and  $\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^b, \sigma_2^b))$ , Equation (5) is trivially verified. That concludes our proof.  $\square$

*Remark 6.* Lemma 5 crucially relies on the assumption (transparent in our definition of Nash equilibrium) that we consider *antagonistic* games, that is,  $\mathcal{P}_2$  uses the inverse preference relation  $\sqsubseteq^{-1}$ . Actually, our approach almost completely carries over to the general case where  $\mathcal{P}_1$  and  $\mathcal{P}_2$  use two different, unrelated, preference relations: the single breaking point being the use of Lemma 5 in Theorem 23 to preserve *memoryless* Nash equilibria in the induction step.  $\triangleleft$

We now establish that Nash equilibria induce optimal strategies (again, in our antagonistic context).

**Lemma 7.** *Let  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  be a game on arena  $\mathcal{A} = (S_1, S_2, E)$ , and let  $s \in S$  be a state. Let  $(\sigma_1, \sigma_2) \in \Sigma_1(\mathcal{A}) \times \Sigma_2(\mathcal{A})$  be a Nash equilibrium from  $s$ . Then, both  $\sigma_1$  and  $\sigma_2$  are  $s$ -optimal strategies.*

*Proof.* We do the proof for  $\mathcal{P}_1$  (it works symmetrically for  $\mathcal{P}_2$ ). Let  $(\sigma_1, \sigma_2) \in \Sigma_1(\mathcal{A}) \times \Sigma_2(\mathcal{A})$  be a Nash equilibrium from  $s \in S$ . Consider the rightmost inequality of Equation (4). From it, we deduce that

$$\text{UCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma_1) = \{w \in C^\omega \mid \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2)) \sqsubseteq w\}. \quad (6)$$

Indeed,  $\sigma_2$  is the *best response* [35] to  $\sigma_1$ .

We claim that  $\sigma_1$  is at least as good as every other strategy, hence that  $\sigma_1$  is optimal. Let  $\sigma'_1 \in \Sigma_1(\mathcal{A})$ . We need to prove that  $\text{UCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma_1) \subseteq \text{UCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma'_1)$ . Let  $w \in \text{UCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma_1)$ . From Equation (6) and the leftmost inequality of Equation (4), we have  $\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma'_1, \sigma_2)) \sqsubseteq w$ . Hence, by definition,  $w \in \text{UCol}_{\sqsubseteq}(\mathcal{A}, s, \sigma'_1)$ , which concludes our proof.  $\square$

As noted above, optimal strategies do not always coincide with Nash equilibria. Intuitively, they do coincide in the classical quantitative formulation (using payoff functions) if the *value* of the game exists, that is, if the best payoff that  $\mathcal{P}_1$  can guarantee is equal to the worst payoff that  $\mathcal{P}_2$  can guarantee [35]. In a sense, our concepts of  $\text{UCol}_{\sqsubseteq}$  and  $\text{DCol}_{\sqsubseteq}$  are meant to mimic the classical sup inf and inf sup formulations in our abstract context where objectives are described as preference relations.

So (quantitative) games do not always have a value, and similarly, in our context, optimal strategies as defined above do not always induce a Nash equilibrium. That being said, they do for arguably all *reasonable*

preference relations, as Martin’s determinacy result grants the existence of winning strategies — in our formalism, Nash equilibria — for all Borel winning conditions [34]. That is, for the equivalence to fail, we would need a preference relation capable of inducing non-Borel sets of winning plays (once a threshold is chosen). Hence, the two notions are virtually equivalent — yet, the above clarification is needed to circumvent slight technical issues in the original proofs of Gimbert and Zielonka [26].

*Remark 8.* In one-player games, the two visions — optimal strategies and Nash equilibria — coincide.  $\triangleleft$

*Remark 9.* Lemma 4 can be restated in terms of Nash equilibria, using a similar reasoning.  $\triangleleft$

### 3 Characterization

We are now able to establish our characterization of preference relations admitting finite-memory optimal strategies based on a given memory skeleton  $\mathcal{M}$ . We proceed in three steps. First, in Section 3.1, we present the core concepts of this characterization, i.e., the properties that preference relations must verify to yield UFM strategies. Second, we state our equivalence result in Section 3.2, alongside a corollary of practical interest that lets one lift results from the one-player case to the two-player one. We defer the formal proofs of both directions of the equivalence to Section 4 and Section 5, and only explain here how to combine them. Finally, we provide an illustrative application of our characterization in Section 3.3.

#### 3.1 Concepts

**Generalizing monotony and selectivity.** As discussed in Section 1, Gimbert and Zielonka’s characterization [26] relies on notions of *monotony* and *selectivity* of the preference relation.

Intuitively, the main difference between Gimbert and Zielonka’s technical approach and ours is the following. In the memoryless setting, all the reasoning can be *abstracted away* from the underlying arena and done at the level of sequences of colors. In the finite-memory one, however, one has to pay attention to how sequences of colors are composed and compared, to maintain consistency with regard to the memory and the underlying game arena. This need to intertwine abstract reasoning on arbitrary sequences of colors with concrete tracking of memory updates is the key obstacle to overcome.

Much of our effort was thus spent on trying to define concepts that would preserve the elegance of monotony and selectivity while allowing us to lift the theory to the finite-memory case. As often the case in these endeavors, the good concepts turned out to be the most natural ones, capturing the intuitive idea that one needs monotony and selectivity *modulo a memory skeleton*.

**Definition 10 ( $\mathcal{M}$ -monotony).** Let  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  be a memory skeleton. A preference relation  $\sqsubseteq$  is  $\mathcal{M}$ -monotone if for all  $m \in M$ , for all  $K_1, K_2 \in \mathcal{R}(C)$ ,

$$\exists w \in L_{m_{\text{init}}, m}, [wK_1] \sqsubset [wK_2] \implies \forall w' \in L_{m_{\text{init}}, m}, [w'K_1] \sqsubseteq [w'K_2]. \quad (7)$$

Recall that a memory skeleton  $\mathcal{M}$  has a fixed initial state  $m_{\text{init}}$ . Intuitively,  $\mathcal{M}$ -monotony extends Gimbert and Zielonka’s monotony by asking one to compare prefixes *belonging to the same language*  $L_{m_{\text{init}}, m}$ , that is, prefixes that are deemed equivalent by the memory skeleton. This property roughly captures that  $\sqsubseteq$  is *stable with regard to prefix addition*, for memory-equivalent prefixes.

The original monotony notion is exactly equivalent to our  $\mathcal{M}$ -monotony with  $\mathcal{M}$  being the trivial skeleton  $\mathcal{M}_{\text{triv}}$ : that is, the memoryless case is naturally a specific subcase of our framework.

**Definition 11 ( $\mathcal{M}$ -selectivity).** Let  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  be a memory skeleton. A preference relation  $\sqsubseteq$  is  $\mathcal{M}$ -selective if for all  $w \in C^*$ ,  $m = \widehat{\alpha_{\text{upd}}}(m_{\text{init}}, w)$ , for all  $K_1, K_2 \in \mathcal{R}(C)$  such that  $K_1, K_2 \subseteq L_{m, m}$ , for all  $K_3 \in \mathcal{R}(C)$ ,

$$[w(K_1 \cup K_2)^*K_3] \sqsubseteq [wK_1^*] \cup [wK_2^*] \cup [wK_3]. \quad (8)$$



Similarly,  $\mathcal{M}$ -selectivity extends Gimbert and Zielonka's selectivity by asking one to compare sequences of colors *belonging to the same language*  $L_{m,m}$ , that is, sequences read as cycles on the memory skeleton. Note also that the memory state  $m$  should be consistent with the prefix  $w$  read from the initial memory state  $m_{\text{init}}$ . This property roughly captures that  $\sqsubseteq$  is *stable with regard to cycle mixing*, for memory-equivalent cycles.

Again, the original selectivity notion is exactly equivalent to  $\mathcal{M}_{\text{triv}}$ -selectivity.

Our notions respect the natural intuition that access to additional memory should always be helpful: if a skeleton  $\mathcal{M}$  is sufficient to classify sequences of colors in a way that guarantees  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity, then it should also be the case for “more powerful” skeletons.

**Lemma 12.** *Let  $\mathcal{M}$  and  $\mathcal{M}'$  be two memory skeletons. If  $\sqsubseteq$  is  $\mathcal{M}$ -monotone (resp.  $\mathcal{M}$ -selective) then, it is also  $(\mathcal{M} \otimes \mathcal{M}')$ -monotone (resp.  $(\mathcal{M} \otimes \mathcal{M}')$ -selective).*

*Proof.* We write  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  and  $\mathcal{M}' = (M', m'_{\text{init}}, \alpha'_{\text{upd}})$ .

Let us assume that  $\sqsubseteq$  is  $\mathcal{M}$ -monotone, that is, for all  $m \in M$ , for all  $K_1, K_2 \in \mathcal{R}(C)$ ,

$$\exists w \in L_{m_{\text{init}}, m}, [wK_1] \sqsubset [wK_2] \implies \forall w' \in L_{m_{\text{init}}, m}, [w'K_1] \sqsubseteq [w'K_2]. \quad (9)$$

We show that  $\sqsubseteq$  is  $(\mathcal{M} \otimes \mathcal{M}')$ -monotone, that is, for all  $(m, m') \in M \times M'$ , for all  $K_1, K_2 \in \mathcal{R}(C)$ ,

$$\exists w \in L_{(m_{\text{init}}, m'_{\text{init}}), (m, m')}, [wK_1] \sqsubset [wK_2] \implies \forall w' \in L_{(m_{\text{init}}, m'_{\text{init}}), (m, m')}, [w'K_1] \sqsubseteq [w'K_2]. \quad (10)$$

To do so, we notice that  $L_{(m_{\text{init}}, m'_{\text{init}}), (m, m')} \subseteq L_{m_{\text{init}}, m}$  (the product of memory skeletons simply updates both memories in parallel). Thus, if the premise of Equation (10) holds, we obtain by Equation (9) that the conclusion of Equation (10) also holds.

A similar argument can be laid out to show that  $\mathcal{M}$ -selectivity implies  $(\mathcal{M} \otimes \mathcal{M}')$ -selectivity. It is enough to notice that for all  $(m, m') \in M \times M'$ , we have  $L_{(m, m'), (m, m')} \subseteq L_{m, m}$ : the definition of  $\mathcal{M}$ -selectivity is thus clearly stronger than the definition of  $(\mathcal{M} \otimes \mathcal{M}')$ -selectivity.  $\square$

**Prefix-covers and cyclic-covers.** While the aforementioned concepts of  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity are the primordial ones for stating the characterization, we still need two additional notions to prove it.

Let us sketch the issue here already. To prove that monotone and selective preference relations yield UML strategies, Gimbert and Zielonka deploy an *inductive argument* on the number of choices in an arena. Intuitively, we want to use a similar approach for UFM strategies, but because of the unavoidable coupling between the memory skeleton and the arena (e.g., Lemma 4), the induction argument breaks, as adding one choice in the arena results in adding many in the *product arena* (as many as there are memory states), where the reasoning needs to take place. New insight and techniques are thus needed to patch this induction scheme.

To solve this issue, *we decouple the two aspects* (see Section 5). Intuitively, we first establish that, on arenas that inherently share the same good properties as product arenas (that is, they already “classify” prefixes and cycles as the memory would), we can deploy the induction argument and obtain UML strategies. Then, we obtain the result for UFM strategies on *general* arenas as a corollary. The crux is identifying such “good” arenas: this is done through the following notions.

**Definition 13 (Prefix-covers and cyclic-covers).** *Let  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  be a memory skeleton and  $\mathcal{A} = (S_1, S_2, E)$  be an arena. Let  $S_{\text{cov}} \subseteq S$ .*

*We say that  $\mathcal{M}$  is a prefix-cover of  $S_{\text{cov}}$  in  $\mathcal{A}$  if for all  $s \in S$ , there exists  $m_s \in M$  such that, for all  $\rho \in \text{Hists}(\mathcal{A})$  such that  $\text{in}(\rho) \in S_{\text{cov}}$ ,  $\text{out}(\rho) = s$  and such that for all  $\rho'$  proper prefix of  $\rho$ ,  $\text{out}(\rho') \neq s$ , we have  $\widehat{\alpha_{\text{upd}}}(m_{\text{init}}, \text{col}(\rho)) = m_s$ .*

*We say that  $\mathcal{M}$  is a cyclic-cover of  $S_{\text{cov}}$  in  $\mathcal{A}$  if for all  $\rho \in \text{Hists}(\mathcal{A})$  such that  $\text{in}(\rho) \in S_{\text{cov}}$ , if  $s = \text{out}(\rho)$  and  $m = \widehat{\alpha_{\text{upd}}}(m_{\text{init}}, \text{col}(\rho))$ , for all  $\rho' \in \text{Hists}(\mathcal{A})$  such that  $\text{in}(\rho') = \text{out}(\rho) = s$ ,  $\widehat{\alpha_{\text{upd}}}(m, \text{col}(\rho')) = m$ .*

Intuitively,  $\mathcal{M}$  is a prefix-cover for a set of states  $S_{\text{cov}}$  if the histories starting in  $S_{\text{cov}}$  and visiting a given state  $s \in S$  for the first time are read up to the same memory state in the memory skeleton. Similarly,  $\mathcal{M}$  is a cyclic-cover of  $\mathcal{A}$  if the cycles<sup>4</sup> of  $\mathcal{A}$  are read as cycles in the memory skeleton, once the memory has been initialized properly.

As hinted above, the canonical example of a prefix-covered and cyclic-covered arena is a product arena (but many more may be in this case, hence it is beneficial to be general with these concepts).

**Lemma 14.** *Let  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  be a memory skeleton and  $\mathcal{A} = (S_1, S_2, E)$  be an arena. Then  $\mathcal{M}$  is both a prefix-cover and a cyclic-cover for  $S_{\text{cov}} = S \times \{m_{\text{init}}\}$  in the product arena  $\mathcal{A} \ltimes \mathcal{M}$ .*

*Proof.* The main argument that we will be using in this proof is that if there is a history  $\rho$  with  $\text{in}(\rho) = (s, m)$  and  $\text{out}(\rho) = (s', m')$  in the product arena  $\mathcal{A} \ltimes \mathcal{M}$ , then reading  $\widehat{\text{col}}(\rho)$  from  $m$  in the memory skeleton  $\mathcal{M}$  leads to  $m'$  (i.e.,  $\widehat{\alpha}_{\text{upd}}(m, \widehat{\text{col}}(\rho)) = m'$ ). This can be easily proved by induction on the length of  $\rho$ , thanks to how the product arena is built.

We first show that  $\mathcal{M}$  is a prefix-cover for  $S_{\text{cov}} = S \times \{m_{\text{init}}\}$  in the product arena  $\mathcal{A} \ltimes \mathcal{M}$ . What we have to prove, instantiating the definition of prefix-cover in this case, is that for all  $(s, m) \in S \times M$ , there exists  $m_{(s, m)} \in M$  such that, for all  $\rho \in \text{Hists}(\mathcal{A} \ltimes \mathcal{M})$  such that  $\text{in}(\rho) \in S_{\text{cov}}$ ,  $\text{out}(\rho) = (s, m)$  and such that for all  $\rho'$  proper prefix of  $\rho$ ,  $\text{out}(\rho') \neq s$ , we have  $\widehat{\alpha}_{\text{upd}}(m_{\text{init}}, \widehat{\text{col}}(\rho)) = m_{(s, m)}$ . Let  $(s, m) \in M$ ; we take  $m_{(s, m)} = m$ . Then, if  $\rho \in \text{Hists}(\mathcal{A} \ltimes \mathcal{M})$  is such that  $\text{in}(\rho) \in S_{\text{cov}}$  (that is, is equal to  $(s', m_{\text{init}})$  for some  $s' \in S$ ), and  $\text{out}(\rho) = (s, m)$ , we have by construction of the product arena that  $\widehat{\alpha}_{\text{upd}}(m_{\text{init}}, \widehat{\text{col}}(\rho)) = m = m_{(s, m)}$ , as required.

To prove that  $\mathcal{M}$  is a cyclic-cover for  $S_{\text{cov}}$  in  $\mathcal{A} \ltimes \mathcal{M}$ , we have to prove that for all  $\rho \in \text{Hists}(\mathcal{A} \ltimes \mathcal{M})$  such that  $\text{in}(\rho) \in S_{\text{cov}}$ , if  $(s, m) = \text{out}(\rho)$  and  $m' = \widehat{\alpha}_{\text{upd}}(m_{\text{init}}, \widehat{\text{col}}(\rho))$ , for all  $\rho' \in \text{Hists}(\mathcal{A} \ltimes \mathcal{M})$  such that  $\text{in}(\rho') = \text{out}(\rho') = (s, m)$ ,  $\widehat{\alpha}_{\text{upd}}(m', \widehat{\text{col}}(\rho')) = m'$ . Let  $\rho \in \text{Hists}(\mathcal{A} \ltimes \mathcal{M})$  such that  $\text{in}(\rho) \in S_{\text{cov}}$  (that is,  $\text{in}(\rho) = (s', m_{\text{init}})$  for some  $s' \in S$ ). Then, if  $(s, m) = \text{out}(\rho)$ , we have by construction of the product arena that  $m' = \widehat{\alpha}_{\text{upd}}(m_{\text{init}}, \widehat{\text{col}}(\rho)) = m$ . Let  $\rho' \in \text{Hists}(\mathcal{A} \ltimes \mathcal{M})$  such that  $\text{in}(\rho') = \text{out}(\rho') = (s, m)$ . By construction of the product arena, we therefore have that  $\widehat{\alpha}_{\text{upd}}(m, \widehat{\text{col}}(\rho')) = m$ , as required.  $\square$

### 3.2 Main results

**Equivalence.** We now have the necessary ingredients to state our general equivalence result formally.

**Theorem 15 (Equivalence).** *Let  $\sqsubseteq$  be a preference relation and let  $\mathcal{M}$  be a memory skeleton. Then, both players have UFM strategies based on memory skeleton  $\mathcal{M}$  in all games  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  if and only if  $\sqsubseteq$  and  $\sqsubseteq^{-1}$  are  $\mathcal{M}$ -monotone and  $\mathcal{M}$ -selective.*

We state this theorem broadly and with a *focus on UFM strategies*. The actual results we have for each direction of the equivalence — which we develop in Section 4 and Section 5 — are a bit stronger, of wider applicability and/or more interesting, but this statement carries the take-home message of our work. It is also meant to mirror the seminal result of Gimbert and Zielonka [26, Theorem 2]: their result can be retrieved from Theorem 15 by taking the trivial memory skeleton  $\mathcal{M}_{\text{triv}}$ . As such, our work brings a *strict generalization of Gimbert and Zielonka's results* [26] to the finite-memory case.

*Remark 16.* We will refine the statement of our intermediate results in order to use weaker hypotheses and/or grant stronger conclusions, whenever possible. For example, we only need optimal strategies in the left-to-right direction (Theorem 19 and Theorem 20) — and not the stronger notion of Nash equilibrium — while we do prove the existence of finite-memory Nash equilibria in the other direction (Theorem 23 and Corollary 25).

Similarly, we study the two implications of the equivalence in a *compositional* way: we split the reasoning for  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity, using different skeletons for each whenever meaningful, as well as for the

<sup>4</sup> Definition 13 can be equivalently stated by considering simple cycles only.

players, again when beneficial. Additionally, we distinguish between arenas where the players do not need memory and the ones where they do, the first essentially being arenas that already share the good properties of product arenas (as in “product with a memory skeleton”).

While such a level of care is not necessary to obtain Theorem 15, it has two advantages. First, from a practical standpoint, it permits to obtain *more useful results*<sup>5</sup> when focusing on a particular direction of the equivalence (as often the case in applications). Second, from a theoretical standpoint, it permits to isolate each concept and each element of the reasoning and to *highlight their true roles*<sup>6</sup> in the underlying mechanisms that lead to the existence of UFM strategies.  $\triangleleft$

To prove Theorem 15, we invoke the results we will prove in Section 4 and Section 5.

*Proof (Theorem 15).* The left-to-right implication trivially follows from Theorem 19 (for  $\mathcal{M}$ -monotony) and Theorem 20 (for  $\mathcal{M}$ -selectivity), applied to each player with respect to his preference relation. The converse implication is established in Corollary 25, which can be restated in terms of UFM strategies through Lemma 7.  $\square$

As a by-product of our method, we also obtain a similar equivalence by solely considering one of the two players and the corresponding one-player arenas.

**Theorem 17 (One-player equivalence).** *Let  $\sqsubseteq$  be a preference relation and let  $\mathcal{M}$  be a memory skeleton. Then,  $\mathcal{P}_1$  has UFM strategies based on memory skeleton  $\mathcal{M}$  in all his one-player games  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  if and only if  $\sqsubseteq$  is  $\mathcal{M}$ -monotone and  $\mathcal{M}$ -selective.*

Although this looks like a weak version of Theorem 15 at first sight, this is actually a distinct result as both sides of the equivalence are weaker: on the left side, it only handles the memory requirements for  $\mathcal{P}_1$ ’s one-player games; on the right side, it does not assume anything about the inverse preference relation  $\sqsubseteq^{-1}$ .

Albeit close, this is also distinct from the *half-positional determinacy* result from [32, Theorem 4], which gives *sufficient* conditions about a winning condition for a player to admit memoryless optimal strategies on every *two-player* arena — in Theorem 17, we give a *necessary and sufficient* condition for a player to admit UFM strategies on his *one-player* arenas only. The sufficient conditions from [32] (*prefix-independence* and *concavity*) imply  $\mathcal{M}_{\text{triv}}$ -monotony and  $\mathcal{M}_{\text{triv}}$ -selectivity, but not the other way around. Given a preference relation, it is possible for a player to have UFM strategies on his one-player arenas, but not on all two-player arenas: e.g., the example used in [32, Proposition 2]). In such an example, Theorem 17 could be applied, but not the result from [32].

*Proof (Theorem 17).* The left-to-right implication trivially follows from Theorem 19 (for  $\mathcal{M}$ -monotony) and Theorem 20 (for  $\mathcal{M}$ -selectivity). The converse implication is established in Corollary 26.  $\square$

**Lifting corollary.** As discussed in Section 1, the work of Gimbert and Zielonka contains not one, but *two* great results. Alongside the aforementioned equivalence result, Gimbert and Zielonka provide a corollary of high practical interest [26, Corollary 7]: they essentially obtain as a by-product of their approach that if memoryless strategies suffice in all one-player games of  $\mathcal{P}_1$  and all one-player games of  $\mathcal{P}_2$ , they also suffice in all two-player games.

This provides an elegant way to prove that a preference relation (or equivalently an objective) admits memoryless optimal strategies *without proving monotony and selectivity at all*: proving it in the two one-player subcases, which is generally much easier<sup>7</sup> as it boils down to graph reasoning, and then lifting the result to the general two-player case through the corollary.

Again, we are able to lift this corollary to the finite-memory case, as follows.

**Corollary 18.** *Let  $\sqsubseteq$  be a preference relation and  $\mathcal{M}_1, \mathcal{M}_2$  be two memory skeletons. Assume that*

<sup>5</sup> E.g., of wider applicability, or avoiding the use of memory on covered arenas.

<sup>6</sup> E.g.,  $\mathcal{M}$ -monotony deals with prefixes up to the first cycle and  $\mathcal{M}$ -selectivity deals with the cycles thereafter; memory skeletons can be built in a compositional way based on these two orthogonal yet complementary tasks.

<sup>7</sup> See examples of one-player vs. two-player complexity in [5, 4, 10].

1. for all one-player arenas  $\mathcal{A} = (S_1, S_2 = \emptyset, E)$ ,  $\mathcal{P}_1$  has a UFM strategy  $\sigma_1 \in \Sigma_1^{\text{FM}}(\mathcal{A})$  based on memory skeleton  $\mathcal{M}_1$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ ;
2. for all one-player arenas  $\mathcal{A} = (S_1 = \emptyset, S_2, E)$ ,  $\mathcal{P}_2$  has a UFM strategy  $\sigma_2 \in \Sigma_2^{\text{FM}}(\mathcal{A})$  based on memory skeleton  $\mathcal{M}_2$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ .

Then, for all two-player arenas  $\mathcal{A} = (S_1, S_2, E)$ , both  $\mathcal{P}_1$  and  $\mathcal{P}_2$  have UFM strategies  $\sigma_i \in \Sigma_i^{\text{FM}}(\mathcal{A})$  based on memory skeleton  $\mathcal{M} = \mathcal{M}_1 \otimes \mathcal{M}_2$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ .

We highlight the two (possibly different) skeletons of the two players to maintain a compositional approach, but if the same skeleton  $\mathcal{M}$  works in both one-player versions, it also suffices in the two-player version.

*Proof.* By Theorem 19 and Theorem 20 — which essentially state that the left-to-right implication of Theorem 15 holds already in one-player games, the hypothesis yields that  $\sqsubseteq$  is  $\mathcal{M}_1$ -monotone and  $\mathcal{M}_1$ -selective, while  $\sqsubseteq^{-1}$  is  $\mathcal{M}_2$ -monotone and  $\mathcal{M}_2$ -selective. Now it suffices to apply Corollary 25 — essentially the right-to-left implication of Theorem 15 — to get the claim.  $\square$

### 3.3 Example of application

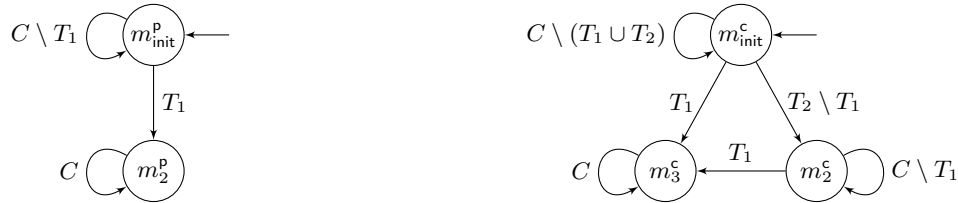
We present an illustrative application of our results, thereby proving the existence of UFM strategies for a specific preference relation: the conjunction of two reachability objectives, a subcase of *generalized reachability games*, studied extensively in [20]. Let  $C$  be an arbitrary set of colors, and  $T_1, T_2 \subseteq C$  be two target sets of colors that have to be reached at least once. Formally, let the winning condition  $W \subseteq C^\omega$  be the set of infinite words  $w = c_1 c_2 \dots$  such that

$$\exists i, j \in \mathbb{N}, c_i \in T_1 \wedge c_j \in T_2.$$

This winning condition induces a preference relation  $\sqsubseteq$  as discussed in Example 3.

In this example, we will use Theorem 15 directly in order to provide one thorough illustration of the definitions of  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity. However, in practice, using Corollary 18 is preferable, as it yields a much shorter proof: by exhibiting the right skeletons for  $\mathcal{P}_1$  and  $\mathcal{P}_2$ , we simply have to show that these skeletons are sufficient to play optimally on both players' *one-player arenas*, which amounts to graph reasoning.

We exhibit two memory skeletons  $\mathcal{M}^p = (M^p, m_{\text{init}}^p, \alpha_{\text{upd}}^p)$  and  $\mathcal{M}^c = (M^c, m_{\text{init}}^c, \alpha_{\text{upd}}^c)$  such that  $\sqsubseteq$  is  $\mathcal{M}^p$ -monotone and  $\mathcal{M}^c$ -selective: they are pictured in Figure 2. Note that such skeletons are obviously not unique.



**Fig. 2.** Memory skeletons  $\mathcal{M}^p$  (left) and  $\mathcal{M}^c$  (right) for two-target reachability games.

Let us prove that  $\sqsubseteq$  is  $\mathcal{M}^p$ -monotone. Let  $m \in M^p$ ,  $K_1, K_2 \in \mathcal{R}(C)$ ; we want to show that Equation (7) is satisfied. We assume that there exists  $w \in L_{m_{\text{init}}^p, m}$  such that  $[wK_1] \sqsubset [wK_2]$ : this means that all words of  $[wK_1]$  are losing, and that there exists a winning word in  $[wK_2]$ . Let  $w' \in L_{m_{\text{init}}^p, m}$ ; we show that we necessarily have that  $[w'K_1] \sqsubseteq [w'K_2]$ . Note that if  $[K_1]$  is empty, this always holds; we now assume that  $[K_1]$  is non-empty. We study the two possible values of  $m$  separately.

- If  $m = m_{\text{init}}^p$ , then  $w$  and  $w'$  do not reach  $T_1$ . If  $w$  does not reach  $T_2$  either, as there is a winning word in  $[wK_2]$ , then there must be a winning word in  $[K_2]$ . This word is still winning after prepending  $w'$  to it, so there is a winning word in  $[w'K_2]$ , and  $[w'K_1] \sqsubseteq [w'K_2]$ . If  $w$  reaches  $T_2$ , then  $[K_1]$  cannot have a word reaching  $T_1$ . As  $w'$  does not reach  $T_1$  either, all words of  $[w'K_1]$  are losing, so  $[w'K_1] \sqsubseteq [w'K_2]$ .
- If  $m = m_2^p$ , then  $w$  and  $w'$  reach  $T_1$ . Clearly,  $w$  cannot reach  $T_2$  (as  $[wK_1]$  would be winning). This implies that  $[K_2]$  must contain a word reaching  $T_2$ ; as  $w'$  reaches  $T_1$ , the concatenation of  $w'$  with the word of  $[K_2]$  reaching  $T_2$  means that there is a winning word in  $[w'K_2]$ , so  $[w'K_1] \sqsubseteq [w'K_2]$ .

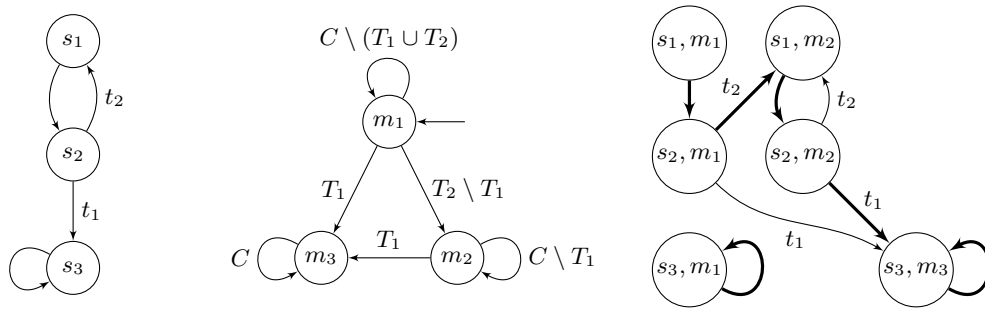
Let us now prove that  $\sqsubseteq$  is  $\mathcal{M}^c$ -selective. Let  $w \in C^*$ ,  $m = \widehat{\alpha}_{\text{upd}}^c(m_{\text{init}}^c, w)$ ,  $K_1, K_2 \in \mathcal{R}(C)$  such that  $K_1, K_2 \subseteq L_{m,m}$ , and  $K_3 \in \mathcal{R}(C)$ . We show that Equation (8) is satisfied, i.e., that

$$[w(K_1 \cup K_2)^* K_3] \sqsubseteq [wK_1^*] \cup [wK_2^*] \cup [wK_3].$$

If all words of  $[w(K_1 \cup K_2)^* K_3]$  are losing, this equation trivially holds; we thus assume that this set contains a winning word. We therefore have to show that there is a winning word in  $[wK_1^*]$ ,  $[wK_2^*]$ , or  $[wK_3]$ . We study the three possible values of  $m$  separately.

- If  $m = m_{\text{init}}^c$ , then  $w$  does not reach  $T_1$  nor  $T_2$ , and the same holds for all words of  $K_1$  and  $K_2$ , as  $K_1, K_2 \subseteq L_{m_{\text{init}}^c, m_{\text{init}}^c}$ . Therefore, if a word of  $[w(K_1 \cup K_2)^* K_3]$  is winning, this must be because a word of  $[wK_3]$  is winning.
- If  $m = m_2^c$ , then  $w$  reaches  $T_2$  but not  $T_1$ , and  $K_1, K_2$  do not reach  $T_1$ . Thus, a word of  $[K_3]$  must reach  $T_1$ ; in particular, a word of  $[wK_3]$  must reach both  $T_1$  and  $T_2$ .
- If  $m = m_3^c$ , we distinguish two cases. If  $w$  reaches  $T_2$  and  $T_1$ , then  $[wK_1^*] \cup [wK_2^*] \cup [wK_3]$  trivially contains only winning words. If  $w$  reaches  $T_1$  but not  $T_2$ , then there must be a word reaching  $T_2$  in  $[(K_1 \cup K_2)^* K_3]$ . Hence, at least one set among  $[K_1^*]$ ,  $[K_2^*]$ , and  $[K_3]$  must contain a word reaching  $T_2$ , so  $[wK_1^*]$ ,  $[wK_2^*]$ , or  $[wK_3]$  contains a winning word.

Similar arguments can be laid out to show that the preference relation  $\sqsubseteq^{-1}$  of  $\mathcal{P}_2$  is  $\mathcal{M}^p$ -monotone and  $\mathcal{M}_{\text{triv}}$ -selective (where  $\mathcal{M}_{\text{triv}}$  is the trivial memory skeleton defined earlier). Let  $\mathcal{M} = \mathcal{M}^p \otimes \mathcal{M}^c \otimes \mathcal{M}_{\text{triv}}$  be the product of all the considered skeletons. Although  $\mathcal{M}$  formally has six states, only three of them are reachable, and it is similar to skeleton  $\mathcal{M}^c$ . By Lemma 12, we have that both  $\sqsubseteq$  and  $\sqsubseteq^{-1}$  are  $\mathcal{M}$ -monotone and  $\mathcal{M}$ -selective. Using Theorem 15, we obtain that *both players have UFM strategies based on skeleton  $\mathcal{M}$  in all games  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$* . Note that memory skeleton  $\mathcal{M}$  is minimal (no memory skeleton with two states or less suffices for  $\mathcal{P}_1$  to play optimally in all arenas [20]).



**Fig. 3.** Arena  $\mathcal{A}$  (left), memory skeleton  $\mathcal{M}$  (center; with  $m_{\text{init}} = m_1$ ), and their product arena  $\mathcal{A} \times \mathcal{M}$  (right; only states reachable from  $S \times \{m_{\text{init}}\}$  are depicted). We assume that  $T_1 = \{t_1\}$ ,  $T_2 = \{t_2\}$ . The  $(S \times \{m_{\text{init}}\})$ -optimal memoryless strategy is highlighted with bold arrows.

We provide an example of a one-player arena  $\mathcal{A} = (S_1, S_2 = \emptyset, E)$  in Figure 3, and show that there is a UFM strategy for the preference relation  $\sqsubseteq$  based on skeleton  $\mathcal{M}$ . To do so, we invoke Lemma 4: we show



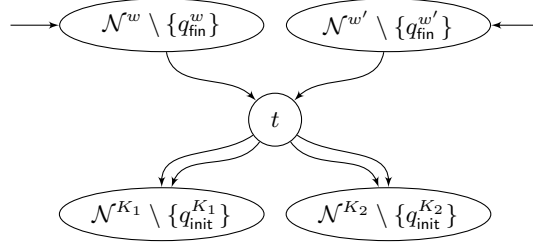
equivalently that the product  $\mathcal{A} \times \mathcal{M}$  admits an  $(S \times \{m_{\text{init}}\})$ -optimal memoryless strategy for  $\sqsubseteq$ . Notice that no memoryless strategy suffices to play optimally in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ , as when starting in  $s_2$ ,  $\mathcal{P}_1$  should first visit  $s_1$  before going to  $s_3$ . Also, the  $(S \times \{m_{\text{init}}\})$ -optimal memoryless strategy for the product arena is only optimal if the initial state is in  $S \times \{m_{\text{init}}\}$ ; it is for instance not optimal from state  $(s_2, m_2)$ .

#### 4 From finite memory based on $\mathcal{M}$ to $\mathcal{M}$ -monotony and $\mathcal{M}$ -selectivity

**Monotony.** We want to keep our approach as compositional as possible, hence we consider the two notions separately. Let us start with  $\mathcal{M}$ -monotony.

**Theorem 19.** *Let  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  be a memory skeleton and  $\sqsubseteq$  be a preference relation. Assume that for all one-player arenas  $\mathcal{A} = (S_1, S_2 = \emptyset, E)$ , for all  $s, s' \in S$ ,  $\mathcal{P}_1$  has an  $s$ -optimal and  $s'$ -optimal strategy  $\sigma \in \Sigma_1^{\text{FM}}(\mathcal{A})$ , encoded as a Mealy machine  $\Gamma_\sigma = (\mathcal{M}, \alpha_{\text{next}})$ , in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ . Then  $\sqsubseteq$  is  $\mathcal{M}$ -monotone.*

Note that the same holds for  $\mathcal{P}_2$  and  $\sqsubseteq^{-1}$  symmetrically. Also, we do not require full *uniformity* of the strategy, but only uniformity with regard to the fixed pair of states (i.e., strategy  $\sigma$  does not need to be optimal from other states).



**Fig. 4.** Automaton  $\mathcal{N}$  built to establish  $\mathcal{M}$ -monotony.

Our proof can be sketched as follows. We need to establish that Equation (7) holds. We first instantiate the four languages involved in it:  $\{w\}$ ,  $\{w'\}$ ,  $K_1$  and  $K_2$ . We take NFA recognizing them and build an NFA  $\mathcal{N}$  that joins them in such a way that, when  $\mathcal{N}$  is considered as a game arena (see Lemma 2), its plays correspond exactly to the languages of infinite words considered in Equation (7). This arena is essentially composed of two chains emulating the two prefixes  $w$  and  $w'$  and leading to a state  $t$  where  $\mathcal{P}_1$  has to pick a side corresponding to the two languages  $[K_1]$  and  $[K_2]$  (Figure 4). Now, establishing the  $\mathcal{M}$ -monotony of  $\sqsubseteq$  boils down to invoking an optimal strategy  $\sigma$  in the corresponding game, the crux being that this strategy always picks the same edge in  $t$  (i.e., the same side between subarenas corresponding to  $[K_1]$  and  $[K_2]$ ) as both prefixes  $w$  and  $w'$  are deemed equivalent by the memory skeleton  $\mathcal{M}$ .

*Proof.* Let  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  be a memory skeleton and  $\sqsubseteq$  be a preference relation satisfying the hypothesis. Let us prove that  $\sqsubseteq$  is  $\mathcal{M}$ -monotone, i.e., that for all  $m \in M$ , for all  $K_1, K_2 \in \mathcal{R}(C)$ ,

$$\exists w \in L_{m_{\text{init}}, m}, [wK_1] \sqsubseteq [wK_2] \implies \forall w' \in L_{m_{\text{init}}, m}, [w'K_1] \sqsubseteq [w'K_2]. \quad (11)$$

Let  $m \in M$ ,  $K_1, K_2 \in \mathcal{R}(C)$ . We assume that  $K_1, K_2 \neq \emptyset$ , otherwise Equation (11) holds trivially: if  $K_1$  is empty, the conclusion of the implication is true regardless of  $K_2$ ; and if  $K_2$  is empty, the premise is false. Now, assume there exists  $w \in L_{m_{\text{init}}, m}$  such that  $[wK_1] \sqsubseteq [wK_2]$ , and let  $w'$  be another prefix in  $L_{m_{\text{init}}, m}$ . We will prove that  $[w'K_1] \sqsubseteq [w'K_2]$ .

Let  $\mathcal{N}^w = (Q^w, B^w, \delta^w, q_{\text{init}}^w, Q_{\text{fin}}^w)$ ,  $\mathcal{N}^{w'} = (Q^{w'}, B^{w'}, \delta^{w'}, q_{\text{init}}^{w'}, Q_{\text{fin}}^{w'})$ ,  $\mathcal{N}^{K_1} = (Q^{K_1}, B^{K_1}, \delta^{K_1}, q_{\text{init}}^{K_1}, Q_{\text{fin}}^{K_1})$  and  $\mathcal{N}^{K_2} = (Q^{K_2}, B^{K_2}, \delta^{K_2}, q_{\text{init}}^{K_2}, Q_{\text{fin}}^{K_2})$  respectively denote NFA recognizing languages  $\{w\}$ ,  $\{w'\}$ ,  $K_1$  and  $K_2$ . They exist since all these languages are regular. We assume w.l.o.g. that automaton  $\mathcal{N}^w$  (resp.  $\mathcal{N}^{w'}$ ,

$\mathcal{N}^{K_1}, \mathcal{N}^{K_2}$ ) is coaccessible and has only one initial state  $q_{\text{init}}^w$  (resp.  $q_{\text{init}}^{w'}, q_{\text{init}}^{K_1}, q_{\text{init}}^{K_2}$ ) with no ingoing transition. We can do this since  $K_1$  and  $K_2$  are non-empty. We also assume w.l.o.g. that  $\mathcal{N}^w$  (resp.  $\mathcal{N}^{w'}$ ) has only one final state  $q_{\text{fin}}^w$  (resp.  $q_{\text{fin}}^{w'}$ ) with no outgoing transition. Actually,  $\mathcal{N}^w$  and  $\mathcal{N}^{w'}$  can be taken as “chains” recognizing a unique word, and being coaccessible and deterministic.

We build an automaton  $\mathcal{N} = (Q, B, \delta, Q_{\text{init}}, Q_{\text{fin}})$  by “merging” states  $q_{\text{init}}^{K_1}, q_{\text{init}}^{K_2}, q_{\text{fin}}^w$ , and  $q_{\text{fin}}^{w'}$ . We call this new merged state  $t$ . Formally, we built it as follows.

- $Q = (Q^w \cup Q^{w'} \cup Q^{K_1} \cup Q^{K_2} \cup \{t\}) \setminus \{q_{\text{init}}^{K_1}, q_{\text{init}}^{K_2}, q_{\text{fin}}^w, q_{\text{fin}}^{w'}\}$ ;
- $B = B^w \cup B^{w'} \cup B^{K_1} \cup B^{K_2}$ ;
- $Q_{\text{init}} = \{q_{\text{init}}^w, q_{\text{init}}^{w'}\}$  and  $Q_{\text{fin}} = Q_{\text{fin}}^{K_1} \cup Q_{\text{fin}}^{K_2}$ ;
- and finally, the transition relation simply takes into account the merging on  $t$ :

$$\begin{aligned} \delta = & \left\{ (q, c, q') \mid (q, c, q') \in (\delta^w \cup \delta^{w'} \cup \delta^{K_1} \cup \delta^{K_2}) \wedge q, q' \notin \{q_{\text{init}}^{K_1}, q_{\text{init}}^{K_2}, q_{\text{fin}}^w, q_{\text{fin}}^{w'}\} \right\} \\ & \cup \left\{ (q, c, t) \mid (q, c, q') \in (\delta^w \cup \delta^{w'}) \wedge q' \in \{q_{\text{fin}}^w, q_{\text{fin}}^{w'}\} \right\} \\ & \cup \left\{ (t, c, q') \mid (q, c, q') \in (\delta^{K_1} \cup \delta^{K_2}) \wedge q \in \{q_{\text{init}}^{K_1}, q_{\text{init}}^{K_2}\} \right\}. \end{aligned}$$

This construction is illustrated in Figure 4. The language recognized by  $\mathcal{N}$  from  $q_{\text{init}}^w$  is  $w(K_1 \cup K_2)$ , whereas from  $q_{\text{init}}^{w'}$ , it is  $w'(K_1 \cup K_2)$ . Observe that  $\mathcal{N}$  is coaccessible since both  $\mathcal{N}^{K_1}$  and  $\mathcal{N}^{K_2}$  are coaccessible.

Recall that we assume  $[wK_1] \sqsubset [wK_2]$ . By definition, this implies that  $[wK_2] \neq \emptyset$ , hence we also have that  $[K_2] \neq \emptyset$ . From this, we get that  $t$  is essential in  $\mathcal{N}$  (Lemma 2). Thus, it is also the case of  $q_{\text{init}}^w$  and  $q_{\text{init}}^{w'}$ .

We will now interpret this NFA as an arena and use the hypothesis. Let  $\mathcal{A} = \text{Arena}(\mathcal{N})$ . By Lemma 2, we have that  $\widehat{\text{col}}(\text{Plays}(\mathcal{A}, q_{\text{init}}^w)) = [w(K_1 \cup K_2)]$  and  $\widehat{\text{col}}(\text{Plays}(\mathcal{A}, q_{\text{init}}^{w'})) = [w'(K_1 \cup K_2)]$ . By hypothesis,  $\mathcal{P}_1$  has a  $q_{\text{init}}^w$ -optimal and  $q_{\text{init}}^{w'}$ -optimal strategy  $\sigma \in \Sigma_1^{\text{FM}}(\mathcal{A})$ , encoded as a Mealy machine  $\Gamma_\sigma = (\mathcal{M}, \alpha_{\text{next}})$ , in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ .

Let  $\pi \in \text{Plays}(\mathcal{A}, q_{\text{init}}^w, \sigma)$  be the only play consistent with strategy  $\sigma$  from  $q_{\text{init}}^w$ . By definition of  $\mathcal{A}$ , this play  $\pi$  necessarily contains a history  $\rho = e_1 \dots e_n$  such that  $\text{out}(e_n) = t$  and for all  $i$ ,  $1 \leq i < n$ ,  $\text{out}(e_i) \neq t$ . Observe that  $\widehat{\text{col}}(\rho) = w$ . Recall that  $m = \widehat{\alpha}_{\text{upd}}(m_{\text{init}}, w)$  is the memory state reached after reading  $w$  since  $w \in L_{m_{\text{init}}, m}$ . Let  $e = \alpha_{\text{next}}(m, t)$  be the edge chosen by  $\sigma$  in  $t$  when  $t$  is visited (note that  $t$  will be visited only once by construction of  $\mathcal{A}$ ).

We will show that  $e$  belongs to the part generated by  $\mathcal{N}^{K_2}$ . By contradiction, assume it belongs to  $\mathcal{N}^{K_1}$ . Then,  $\pi = \rho \cdot \pi'$ , with  $\widehat{\text{col}}(\pi') \in [K_1]$ , hence  $\widehat{\text{col}}(\pi) \in [wK_1]$ . First, observe that

$$[wK_2] \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, q_{\text{init}}^w))$$

since  $[wK_2] \subseteq [wK_1] \cup [wK_2]$ ,  $[wK_1] \cup [wK_2] = [w(K_1 \cup K_2)]$  by Lemma 1, and, as noted above,  $[w(K_1 \cup K_2)] = \widehat{\text{col}}(\text{Plays}(\mathcal{A}, q_{\text{init}}^w))$ . Now since  $\sigma$  is optimal<sup>8</sup> from  $q_{\text{init}}^w$ , we have

$$[wK_2] \sqsubseteq \widehat{\text{col}}(\pi).$$

Finally, we assumed that  $\widehat{\text{col}}(\pi) \in [wK_1]$ , hence we can conclude that

$$[wK_2] \sqsubseteq [wK_1],$$

which contradicts the hypothesis that  $[wK_1] \sqsubset [wK_2]$ . Hence, we have established that  $e$  belongs to  $\mathcal{N}^{K_2}$ .

Now let us consider  $\pi'' \in \text{Plays}(\mathcal{A}, q_{\text{init}}^{w'}, \sigma)$ , the only play consistent with strategy  $\sigma$  from  $q_{\text{init}}^{w'}$ . Again, by definition of  $\mathcal{A}$ , this play  $\pi''$  necessarily contains a history  $\rho'' = e_1 \dots e_n$  such that  $\text{out}(e_n) = t$  and for all  $i$ ,  $1 \leq i < n$ ,  $\text{out}(e_i) \neq t$ . Observe that  $\widehat{\text{col}}(\rho'') = w'$ . Since  $w' \in L_{m_{\text{init}}, m}$ , we also have that  $\widehat{\alpha}_{\text{upd}}(m_{\text{init}}, w') = m$ ,

<sup>8</sup> One can easily get from the definition using the UCol-operator that, in this one-player game,  $\sigma$  is  $q_{\text{init}}^w$ -optimal if and only if for all  $\sigma' \in \Sigma(\mathcal{A})$ ,  $\widehat{\text{col}}(\text{Plays}(\mathcal{A}, q_{\text{init}}^w, \sigma')) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, q_{\text{init}}^w, \sigma))$ .

i.e., the memory state reached after reading  $w'$  is the same as the one reached after reading  $w$ . Recall that  $\alpha_{\text{next}}$  is deterministic by definition: i.e., for a given memory state and state of the arena, it always prescribes the same edge. Hence, we have that  $e = \alpha_{\text{next}}(m, t)$  is exactly the same as before, and therefore belongs to  $\mathcal{N}^{K_2}$ . Thus,  $\widehat{\text{col}}(\pi'') \in [w'K_2]$ .

Finally, since  $\sigma$  is also  $q_{\text{init}}^{w'}$ -optimal and applying the same reasoning as above, we have that

$$\begin{aligned} [w'K_1] \sqsubseteq [w'K_1] \cup [w'K_2] &= [w'(K_1 \cup K_2)] = \widehat{\text{col}}(\text{Plays}(\mathcal{A}, q_{\text{init}}^{w'})) \\ &\sqsubseteq \widehat{\text{col}}(\pi'') \\ &\sqsubseteq [w'K_2], \end{aligned}$$

which proves Equation (11) and concludes our proof.  $\square$

**Selectivity.** We now turn to selectivity, which focuses on stability with regard to cycle mixing.

**Theorem 20.** *Let  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  be a memory skeleton and  $\sqsubseteq$  be a preference relation. Assume that for all one-player arenas  $\mathcal{A} = (S_1, S_2 = \emptyset, E)$ , for all  $s \in S$ ,  $\mathcal{P}_1$  has an  $s$ -optimal strategy  $\sigma \in \Sigma_1^{\text{FM}}(\mathcal{A})$ , encoded as a Mealy machine  $\Gamma_\sigma = (\mathcal{M}, \alpha_{\text{next}})$ , in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ . Then  $\sqsubseteq$  is  $\mathcal{M}$ -selective.*

Note that the same holds for  $\mathcal{P}_2$  and  $\sqsubseteq^{-1}$  symmetrically. Again, observe that our hypothesis is as weak as possible as no uniformity is required.

Our proof bears similarities with the monotone case. We need to establish that Equation (8) holds. We first instantiate the four languages involved in it:  $\{w\}$ ,  $K_1$ ,  $K_2$  and  $K_3$ . We take NFA recognizing them and build an NFA  $\mathcal{N}$  that joins them in such a way that, when  $\mathcal{N}$  is considered as a game arena (see Lemma 2), its plays correspond exactly to the languages of infinite words considered in Equation (8). This arena is essentially composed of a chain emulating the prefix  $w$  and leading to a state  $t$  where  $\mathcal{P}_1$  can visit sides that generate cycles from  $K_1$  and  $K_2$  — forever or for a finite time — or branch to a side corresponding to  $K_3$  (Figure 5). Now, establishing the  $\mathcal{M}$ -selectivity of  $\sqsubseteq$  boils down to invoking an optimal strategy  $\sigma$  in the corresponding game, the crux being that this strategy always picks the same edge in  $t$  (i.e., side between subarenas corresponding to  $[K_1^*]$ ,  $[K_2^*]$  and  $[K_3]$ ) as all cycles on  $t$  are deemed equivalent by the memory skeleton  $\mathcal{M}$ . The main difference with the previous construction is transparent in the last sentence: it is now possible to come back to  $t$ , possibly infinitely often, and our proof takes that into account.

*Proof.* Let  $\mathcal{M} = (M, m_{\text{init}}, \alpha_{\text{upd}})$  be a memory skeleton and  $\sqsubseteq$  a preference relation satisfying the hypothesis. Let us prove that  $\sqsubseteq$  is  $\mathcal{M}$ -selective, i.e., that for all  $w \in C^*$ ,  $m = \widehat{\alpha_{\text{upd}}}(m_{\text{init}}, w)$ , for all  $K_1, K_2 \in \mathcal{R}(C)$  such that  $K_1, K_2 \subseteq L_{m,m}$ , for all  $K_3 \in \mathcal{R}(C)$ ,

$$[w(K_1 \cup K_2)^*K_3] \sqsubseteq [wK_1^*] \cup [wK_2^*] \cup [wK_3]. \quad (12)$$

Let  $w \in C^*$  and  $m = \widehat{\alpha_{\text{upd}}}(m_{\text{init}}, w)$ . Let  $K_1, K_2, K_3 \in \mathcal{R}(C)$ , with  $K_1, K_2 \subseteq L_{m,m}$ . In the following, we assume all three languages  $K_1$ ,  $K_2$  and  $K_3$  to be non-empty. Indeed, if  $K_3$  is empty, so is the left-hand side of Equation (12), hence it trivially holds. If both  $K_1$  and  $K_2$  are empty, Equation (12) compares  $[wK_3]$  to itself, hence it trivially holds again. Finally, if  $K_1$  is the only empty language among the three, then Equation (12) can be restated as follows:

$$[w(K_1 \cup K_2)^*K_3] = [w(K_2 \cup K_2)^*K_3] \sqsubseteq [wK_2^*] \cup [wK_2^*] \cup [wK_3] = [wK_1^*] \cup [wK_2^*] \cup [wK_3],$$

where the middle inequality — the one to prove — involves three non-empty sets. A symmetric argument holds if  $K_2$  is the only empty language. We also assume that  $K_1$  and  $K_2$  do not contain the empty word for technical convenience: this is w.l.o.g. thanks to the Kleene stars used in the regular expressions to consider.

As for monotony, we start by considering NFA for all these languages: let  $\mathcal{N}^w = (Q^w, B^w, \delta^w, q_{\text{init}}^w, Q_{\text{fin}}^w)$ ,  $\mathcal{N}^{K_1} = (Q^{K_1}, B^{K_1}, \delta^{K_1}, q_{\text{init}}^{K_1}, Q_{\text{fin}}^{K_1})$ ,  $\mathcal{N}^{K_2} = (Q^{K_2}, B^{K_2}, \delta^{K_2}, q_{\text{init}}^{K_2}, Q_{\text{fin}}^{K_2})$  and  $\mathcal{N}^{K_3} = (Q^{K_3}, B^{K_3}, \delta^{K_3}, q_{\text{init}}^{K_3}, Q_{\text{fin}}^{K_3})$  respectively denote NFA recognizing languages  $\{w\}$ ,  $K_1$ ,  $K_2$  and  $K_3$ . They exist since all these languages are regular. We assume w.l.o.g. that automaton  $\mathcal{N}^w$  (resp.  $\mathcal{N}^{K_1}$ ,  $\mathcal{N}^{K_2}$ ,  $\mathcal{N}^{K_3}$ ) is coaccessible and has only

one initial state  $q_{\text{init}}^w$  (resp.  $q_{\text{init}}^{K_1}$ ,  $q_{\text{init}}^{K_2}$ ,  $q_{\text{init}}^{K_3}$ ) with no ingoing transition. We can do this since  $K_1$ ,  $K_2$  and  $K_3$  are non-empty. We also assume w.l.o.g. that  $\mathcal{N}^w$  (resp.  $\mathcal{N}^{K_1}$ ,  $\mathcal{N}^{K_2}$ ) has only one final state  $q_{\text{fin}}^w$  (resp.  $q_{\text{fin}}^{K_1}$ ,  $q_{\text{fin}}^{K_2}$ ) with no outgoing transition. Again  $\mathcal{N}^w$  can simply be a “chain” recognizing a unique word, being both coaccessible and deterministic.

Similarly to Theorem 19, we build an automaton  $\mathcal{N} = (Q, B, \delta, Q_{\text{init}}, Q_{\text{fin}})$  by “merging” states  $q_{\text{init}}^{K_1}$ ,  $q_{\text{init}}^{K_2}$ ,  $q_{\text{init}}^{K_3}$ ,  $q_{\text{fin}}^w$ ,  $q_{\text{fin}}^{K_1}$ , and  $q_{\text{fin}}^{K_2}$ . We call this new merged state  $t$ . Formally, we build it as follows.

- $Q = (Q^w \cup Q^{K_1} \cup Q^{K_2} \cup Q^{K_3} \cup \{t\}) \setminus \{q_{\text{init}}^{K_1}, q_{\text{init}}^{K_2}, q_{\text{init}}^{K_3}, q_{\text{fin}}^w, q_{\text{fin}}^{K_1}, q_{\text{fin}}^{K_2}\}$ ;
- $B = B^w \cup B^{K_1} \cup B^{K_2} \cup B^{K_3}$ ;
- $Q_{\text{init}} = Q_{\text{init}}^w = \{q_{\text{init}}^w\}$  and  $Q_{\text{fin}} = Q_{\text{fin}}^{K_3}$ ;
- and finally, the transition relation simply takes into account the merging on  $t$ :

$$\begin{aligned} \delta = & \left\{ (q, c, q') \mid (q, c, q') \in (\delta^w \cup \delta^{K_1} \cup \delta^{K_2} \cup \delta^{K_3}) \wedge q, q' \notin \{q_{\text{init}}^{K_1}, q_{\text{init}}^{K_2}, q_{\text{init}}^{K_3}, q_{\text{fin}}^w, q_{\text{fin}}^{K_1}, q_{\text{fin}}^{K_2}\} \right\} \\ & \cup \left\{ (q, c, t) \mid (q, c, q') \in (\delta^w \cup \delta^{K_1} \cup \delta^{K_2}) \wedge q' \in \{q_{\text{fin}}^w, q_{\text{fin}}^{K_1}, q_{\text{fin}}^{K_2}\} \right\} \\ & \cup \left\{ (t, c, q') \mid (q, c, q') \in (\delta^{K_1} \cup \delta^{K_2} \cup \delta^{K_3}) \wedge q \in \{q_{\text{init}}^{K_1}, q_{\text{init}}^{K_2}, q_{\text{init}}^{K_3}\} \right\}. \end{aligned}$$

This construction is illustrated in Figure 5. The language recognized by  $\mathcal{N}$  is  $w(K_1 \cup K_2)^*K_3$ . Observe that  $\mathcal{N}$  is coaccessible since  $\mathcal{N}^{K_1}$ ,  $\mathcal{N}^{K_2}$  and  $\mathcal{N}^{K_3}$  are coaccessible. Also observe that  $t$  is essential by construction: by merging the initial and final states of  $K_1$  (resp.  $K_2$ ), we created cycles on  $t$ . Thus,  $q_{\text{init}}^w$  is also essential.

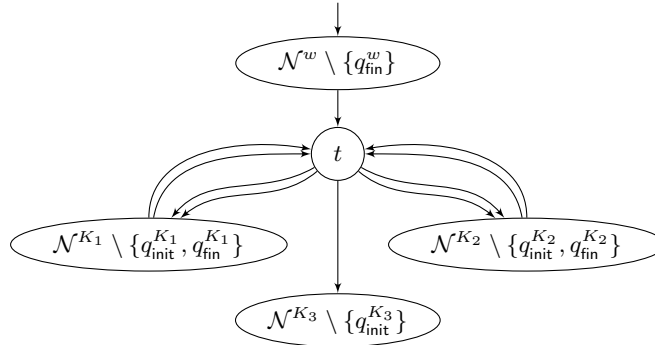


Fig. 5. Automaton  $\mathcal{N}$  built to establish  $\mathcal{M}$ -selectivity.

We will now interpret this NFA as an arena and use the hypothesis. Let  $\mathcal{A} = \text{Arena}(\mathcal{N})$ . By Lemma 2, we have that  $\widehat{\text{col}}(\text{Plays}(\mathcal{A}, q_{\text{init}}^w)) = [w(K_1 \cup K_2)^*K_3]$ . By hypothesis,  $\mathcal{P}_1$  has a  $q_{\text{init}}^w$ -optimal strategy  $\sigma \in \Sigma_1^{\text{FM}}(\mathcal{A})$ , encoded as a Mealy machine  $\Gamma_\sigma = (\mathcal{M}, \alpha_{\text{next}})$ , in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ .

Let  $\pi \in \text{Plays}(\mathcal{A}, q_{\text{init}}^w, \sigma)$  be the only play consistent with  $\sigma$  from  $q_{\text{init}}^w$ . By  $q_{\text{init}}^w$ -optimality, we have that

$$[w(K_1 \cup K_2)^*K_3] \sqsubseteq \widehat{\text{col}}(\pi). \quad (13)$$

By definition of  $\mathcal{A}$ , this play  $\pi$  necessarily contains a history  $\rho = e_1 \dots e_n$  such that  $\text{out}(e_n) = t$  and for all  $i$ ,  $1 \leq i < n$ ,  $\text{out}(e_i) \neq t$ . Observe that  $\widehat{\text{col}}(\rho) = w$ . Recall that  $m = \widehat{\alpha}_{\text{upd}}(m_{\text{init}}, w)$  is the memory state reached after reading  $w$  since  $w \in L_{m_{\text{init}}, m}$ . Let  $e = \alpha_{\text{next}}(m, t)$  be the edge chosen by  $\sigma$  in  $t$  when  $t$  is first visited. Note that in contrast to the construction in Theorem 19,  $t$  could be visited many times here, and even infinitely often (using cycles from  $K_1$  and  $K_2$ ). We consider two cases in the following.

First, assume that  $e$  belongs to the part of the arena generated by  $\mathcal{N}^{K_3}$ . Since  $t$  (originally  $q_{\text{init}}^{K_3}$ ) has no incoming transition in  $\mathcal{N}^{K_3}$ , we conclude that  $\pi$  never visits  $t$  again, and that  $\widehat{\text{col}}(\pi) \in [wK_3]$ . By Equation (13), we verify Equation (12).

Now, assume that  $e$  belongs to the part of the arena generated by  $\mathcal{N}^{K_1}$  (the same reasoning will apply symmetrically for  $\mathcal{N}^{K_2}$ ). We want to show that  $\widehat{\text{col}}(\pi) \in [wK_1^*]$ , i.e., that  $\sigma$  never switches to another part of the arena. Two cases are possible: either (a)  $\pi$  visits  $t$  only once, or (b)  $\pi$  visits  $t$  at least twice.

Case (a). Since  $\pi$  visits  $t$  only once and  $t$  is the only state where the play could switch to a different automaton, we have that  $\pi = \rho \cdot \pi'$  for a suffix  $\pi'$  starting in  $t$  and entirely contained in  $\mathcal{N}^{K_1}$ . Hence, we have  $\widehat{\text{col}}(\pi) = w \cdot \widehat{\text{col}}(\pi')$  with  $\widehat{\text{col}}(\pi') \in [K_1]$ . Thus,  $\widehat{\text{col}}(\pi) \in [wK_1] \subseteq [wK_1^*]$ .

Case (b). Let  $\pi = \rho \cdot \rho' \cdot \pi'$ , such that  $\rho'$  ends with the second visit of  $t$ . Recall that  $w = \widehat{\text{col}}(\rho)$ ,  $m = \widehat{\alpha}_{\text{upd}}(m_{\text{init}}, w)$ , and  $e = \alpha_{\text{next}}(m, t)$ . Now, by definition of  $K_1$ , we have that  $\widehat{\text{col}}(\rho') \in L_{m,m}$ . Hence,  $\widehat{\alpha}_{\text{upd}}(m, \widehat{\text{col}}(\rho')) = m$ . Intuitively, the memory skeleton is back to the same memory state after reading the cycle  $\rho'$ . As argued in Theorem 19,  $\alpha_{\text{next}}$  is deterministic, and both the state of the arena and the memory state are identical after  $\rho$  and after  $\rho \cdot \rho'$ . Therefore  $\sigma(\rho \cdot \rho') = \sigma(\rho) = e$ . Iterating this reasoning (as all cycles on  $t$  in  $\mathcal{N}^{K_1}$  are read as cycles on  $m$  in the memory), we conclude that  $\pi = \rho \cdot (\rho')^\omega$ . This implies that  $\widehat{\text{col}}(\pi) \in [wK_1^*]$ .

Hence, in both cases, we have that  $\widehat{\text{col}}(\pi) \in [wK_1^*]$ . Now, by Equation (13), we verify Equation (12).

Wrapping everything up, we have that whatever the part of the arena to which  $e$  belongs, Equation (12) is verified. Therefore, we have shown that  $\sqsubseteq$  is indeed  $\mathcal{M}$ -selective.  $\square$

**Wrap-up.** We have established that the existence of finite-memory optimal strategies based on a skeleton  $\mathcal{M}$  in one-player games implies both  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity of the preference relation, under very mild uniformity assumption. It is interesting to observe that this holds already for *one-player* games (a fortiori, for *two-player* games too). Next, we consider the converse: we will prove that  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity implies the existence of UFM strategies, not only in one-player games, but even in *two-player* ones, when satisfied by the preference relation *and its inverse*.

## 5 From $\mathcal{M}$ -monotony and $\mathcal{M}$ -selectivity to finite memory based on $\mathcal{M}$

**Induction step.** To prove the sought implication (Theorem 23), we first focus on *memoryless* strategies in “covered” arenas, as discussed in Section 3.1. Intuitively, a “covered” arena resembles a product arena (with a memory skeleton): hence studying memoryless strategies on such arenas is very close to studying finite-memory strategies on general arenas.

We will proceed by induction on the number of choices in an arena, as sketched in Section 3.1. This induction will require us to mix different Nash equilibria (one for each player) in a proper way, to maintain the desired property. For the sake of readability, we thus start by proving the induction step for one player.

For an arena  $\mathcal{A} = (S_1, S_2, E)$ , we write  $n_{\mathcal{A}} = |E| - |S|$  for its number of choices. We also define the notion of *subarena*: we say that an arena  $\mathcal{A}' = (S'_1, S'_2, E')$  is a *subarena* of an arena  $\mathcal{A} = (S_1, S_2, E)$  if  $S_1 = S'_1$ ,  $S_2 = S'_2$ , and  $E' \subseteq E$ . That is, arena  $\mathcal{A}'$  is a subarena of  $\mathcal{A}$  if it can be obtained from  $\mathcal{A}$  by removing some edges of  $\mathcal{A}$  (while keeping it non-blocking). We say that a set of arenas  $\mathfrak{A}$  is closed under subarena operation if for all  $\mathcal{A} \in \mathfrak{A}$ , for all subarenas  $\mathcal{A}'$  of  $\mathcal{A}$ ,  $\mathcal{A}' \in \mathfrak{A}$ .

**Lemma 21.** *Let  $\sqsubseteq$  be a preference relation,  $\mathcal{M}^p$  and  $\mathcal{M}^c$  be two memory skeletons, and  $\mathfrak{A}$  be a set of arenas closed under subarena operation. Assume that  $\sqsubseteq$  is  $\mathcal{M}^p$ -monotone and  $\mathcal{M}^c$ -selective, and that for all  $\mathcal{P}_2$ ’s one-player arenas  $\mathcal{A} = (S_1, S_2, E) \in \mathfrak{A}$ , for all subsets of states  $S_{\text{cov}} \subseteq S$  for which  $\mathcal{M}^p$  is a prefix-cover and  $\mathcal{M}^c$  is a cyclic-cover,  $\mathcal{P}_2$  has an optimal strategy from  $S_{\text{cov}}$ .*

*Let  $n \in \mathbb{N}$ . Assume that for all arenas  $\mathcal{A}' = (S'_1, S'_2, E') \in \mathfrak{A}$  such that  $n_{\mathcal{A}'} < n$ , for all subsets of states  $S'_{\text{cov}} \subseteq S'$  for which  $\mathcal{M}^p$  is a prefix-cover and  $\mathcal{M}^c$  is a cyclic-cover, there exists a memoryless Nash equilibrium  $(\sigma'_1, \sigma'_2) \in \Sigma_1^{\text{ML}}(\mathcal{A}') \times \Sigma_2^{\text{ML}}(\mathcal{A}')$  from  $S'_{\text{cov}}$  in  $\mathcal{G}' = (\mathcal{A}', \sqsubseteq)$ .*

*Then, for all arenas  $\mathcal{A} = (S_1, S_2, E) \in \mathfrak{A}$  such that  $n_{\mathcal{A}} = n$ , for all subsets of states  $S_{\text{cov}} \subseteq S$  for which  $\mathcal{M}^p$  is a prefix-cover and  $\mathcal{M}^c$  is a cyclic-cover, there exists a Nash equilibrium  $(\sigma_1, \sigma_2) \in \Sigma_1^{\text{ML}}(\mathcal{A}) \times \Sigma_2(\mathcal{A})$  from  $S_{\text{cov}}$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  such that  $\sigma_1$  is memoryless.*



Note that the same holds for  $\mathcal{P}_2$  and  $\sqsubseteq^{-1}$  symmetrically.

Intuitively, Lemma 21 states that under the hypotheses of  $\mathcal{M}^P$ -monotony and  $\mathcal{M}^C$ -selectivity, if both players can play optimally with memoryless strategies in “small” and “covered” arenas, the same property holds for at least  $\mathcal{P}_1$  in “covered” arenas where an additional choice exists.

This lemma has to be commented. First, observe that the property is about Nash equilibria. Indeed, as explained in Section 2, the result we prove is actually slightly stronger than the existence of optimal strategies, as it can be stated for Nash equilibria.

Second, this lemma is focused on proving the existence of an NE in which  $\mathcal{P}_1$ ’s strategy is memoryless: proving that this holds for both players will be done in Theorem 23.

Third, as motivated in Section 3.2, we state our result as the existence of memoryless optimal strategies in “covered” arenas: the existence of UFM strategies in general arenas will follow (Corollary 25), but taking this road allows us to keep optimal strategies memoryless for many arenas (which already share the “classifying” properties that a product with a memory skeleton would grant).

Fourth, we use two different skeletons, one for monotony (i.e., dealing with prefixes) and one for selectivity (i.e., dealing with cycles). Obviously, one can use a single combined skeleton using Lemma 12 and Lemma 14, but our approach has the advantage of being *compositional* and highlighting how each skeleton / property impacts the reasoning in the proof: we will see that they have different uses.

Lastly, the notions of prefix-covers and cyclic-covers are defined with regard to a covered set of states  $S_{\text{cov}}$  in order to keep the need for uniformity minimal, in the same spirit as what we did in Section 4.

As mentioned above, our proof is essentially an induction step. Starting from an arena  $\mathcal{A}$  with  $n_{\mathcal{A}} = n$  choices, we identify a state  $t$  in which  $\mathcal{P}_1$  has at least two outgoing edges (the proof is symmetric for  $\mathcal{P}_2$ ). By splitting the edges in  $t$  in two sets, we obtain two corresponding subarenas  $\mathcal{A}_a$  and  $\mathcal{A}_b$  such that  $n_{\mathcal{A}_a}, n_{\mathcal{A}_b} < n$ , along with the corresponding subgames. The induction hypothesis gives us two memoryless Nash equilibria (from  $S_{\text{cov}}$ ) in these subgames:  $(\sigma_1^a, \sigma_2^a)$  and  $(\sigma_1^b, \sigma_2^b)$ . The arguments can then be unfolded intuitively as follows. First, using  $\mathcal{M}^P$ -monotony and  $\mathcal{M}^P$  being a prefix-cover, we identify one subarena (say  $\mathcal{A}_a$ ) which is clearly at least as good as the other for  $\mathcal{P}_1$ . Second, we build a strategy profile  $(\sigma_1^\#, \sigma_2^\#)$ , that we claim to be an NE in  $\mathcal{G}$ , in the following way:  $\mathcal{P}_1$  uses strategy  $\sigma_1^a$  (the one from the best subarena) and  $\mathcal{P}_2$  reacts to  $\mathcal{P}_1$ ’s actions by playing the corresponding best-response strategy. I.e., if  $\mathcal{P}_1$  plays in  $\mathcal{A}_a$ ,  $\mathcal{P}_2$  plays according to  $\sigma_2^a$ , and otherwise he plays according to  $\sigma_2^b$ . Third, it remains to prove the two inequalities of Equation (4). The rightmost one is easy, as well as the leftmost one in the subcase where the unique play  $\pi \in \text{Plays}(\mathcal{A}, s, \sigma_1^\#, \sigma_2^\#)$  does not visit state  $t$ : they can both be proved essentially thanks to the induction hypothesis and easy construction arguments. The crux of the proof is thus in the last step: proving that the leftmost inequality holds when the play visits  $t$ . This can be achieved thanks to  $\mathcal{M}^C$ -selectivity and  $\mathcal{M}^C$  being a cyclic-cover, Lemma 1, inherent properties of the preference relation,  $\mathcal{A}_a$  being the best subarena thanks  $\mathcal{M}^P$ -monotony, and the induction hypothesis, in that order.

Obviously,  $\mathcal{M}$ -monotony (Definition 10),  $\mathcal{M}$ -selectivity (Definition 11), prefix-covers and cyclic-covers (Definition 13) were defined to be sufficient to provide Lemma 21: one of the main challenges was to have them not too powerful as to keep them also necessary, as proved in Section 4.

*Proof.* Let  $\mathcal{M}^P = (M^P, m_{\text{init}}^P, \alpha_{\text{upd}}^P)$  and  $\mathcal{M}^C = (M^C, m_{\text{init}}^C, \alpha_{\text{upd}}^C)$ . Let  $\sqsubseteq$  be a preference relation that is  $\mathcal{M}^P$ -monotone and  $\mathcal{M}^C$ -selective. Let  $n \in \mathbb{N}$  and assume that for all arenas  $\mathcal{A}' = (S'_1, S'_2, E') \in \mathfrak{A}$  such that  $n_{\mathcal{A}'} < n$ , for all subsets of states  $S'_{\text{cov}} \subseteq S'$  for which  $\mathcal{M}^P$  is a prefix-cover and  $\mathcal{M}^C$  is a cyclic-cover, there exists a memoryless Nash equilibrium  $(\sigma'_1, \sigma'_2) \in \Sigma_1^{\text{ML}}(\mathcal{A}') \times \Sigma_2^{\text{ML}}(\mathcal{A}')$  in  $\mathcal{G}' = (\mathcal{A}', \sqsubseteq)$ .

Now, let  $\mathcal{A} = (S_1, S_2, E) \in \mathfrak{A}$  be an arena such that  $n_{\mathcal{A}} = n$ , and let  $S_{\text{cov}} \subseteq S$  be a subset of states for which  $\mathcal{M}^P$  is a prefix-cover and  $\mathcal{M}^C$  is a cyclic-cover. Our goal is to prove that there exists an NE  $(\sigma_1, \sigma_2) \in \Sigma_1^{\text{ML}}(\mathcal{A}) \times \Sigma_2(\mathcal{A})$  from  $S_{\text{cov}}$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  such that  $\sigma_1$  is memoryless. The same proof can be done with respect to  $\mathcal{P}_2$  symmetrically, using its preference relation  $\sqsubseteq^{-1}$  and the appropriate skeletons.

If  $\mathcal{A}$  is such that  $\mathcal{P}_1$  has no choice (i.e., all his states have only one outgoing edge), then  $\mathcal{A}$  is a  $\mathcal{P}_2$ ’s one-player arena. There is only one (memoryless) strategy in  $\Sigma_1(\mathcal{A})$ , and by hypothesis,  $\mathcal{P}_2$  has a strategy to play in  $\mathcal{A}$  that is optimal from states prefix-covered by  $\mathcal{M}^P$  and cyclic-covered by  $\mathcal{M}^C$ . Thus there is indeed a Nash equilibrium in  $\Sigma_1^{\text{ML}}(\mathcal{A}) \times \Sigma_2(\mathcal{A})$  from  $S_{\text{cov}}$ . Let us now assume that  $\mathcal{P}_1$  has at least one choice and let  $t \in S_1$  be a state with at least two outgoing edges, i.e.,  $|\{e \in E \mid \text{in}(e) = t\}| > 1$ . We partition  $\{e \in E \mid \text{in}(e) = t\}$

in two (non-empty) sets  $E_a$  and  $E_b$ , and we define two corresponding subarenas,  $\mathcal{A}_a = (S_1, S_2, E \setminus E_b)$ , and  $\mathcal{A}_b = (S_1, S_2, E \setminus E_a)$ , which are in  $\mathfrak{A}$  as this set is closed under subarena operation. Observe that it remains true that  $\mathcal{M}^p$  is a prefix-cover and  $\mathcal{M}^c$  is a cyclic-cover of  $S_{\text{cov}}$ , both in  $\mathcal{A}_a$  and  $\mathcal{A}_b$ , by definition of prefix- and cyclic-covers (intuitively, we quantify universally over fewer histories than in  $\mathcal{A}$ ).

Thus, by induction hypothesis (since  $n_{\mathcal{A}_a}, n_{\mathcal{A}_b} < n_{\mathcal{A}} = n$ ), we have memoryless NE from  $S_{\text{cov}}$  in the subgames  $\mathcal{G}_a = (\mathcal{A}_a, \sqsubseteq)$  and  $\mathcal{G}_b = (\mathcal{A}_b, \sqsubseteq)$ . Let us denote them by  $(\sigma_1^j, \sigma_2^j)$  for game  $\mathcal{G}_j$ ,  $j \in \{a, b\}$ .

Since  $\mathcal{M}^p$  is a prefix-cover of  $S_{\text{cov}}$  in  $\mathcal{A}$ , there exists  $m_t^p \in M^p$  such that, for all  $\rho \in \text{Hists}(\mathcal{A})$  such that  $\text{in}(\rho) \in S_{\text{cov}}$ ,  $\text{out}(\rho) = t$  and such that for all  $\rho'$  proper prefix of  $\rho$ ,  $\text{out}(\rho') \neq t$ , we have  $\widehat{\alpha}_{\text{upd}}(m_{\text{init}}^p, \widehat{\text{col}}(\rho)) = m_t^p$ . Now, let  $K_j^p = \widehat{\text{col}}(\text{Hists}(\mathcal{A}_j, t, \sigma_2^j))$ , for  $j \in \{a, b\}$ , that is,  $K_j^p$  contains all (projections to colors of) histories consistent with  $\sigma_2^j$  and starting in  $t$  in subarena  $\mathcal{A}_j$ .

By  $\mathcal{M}^p$ -monotony, we can easily deduce that we have

$$\begin{aligned} \forall w \in L_{m_{\text{init}}^p, m_t^p}^p, [wK_a^p] \sqsubseteq [wK_b^p], \\ \text{or } \forall w \in L_{m_{\text{init}}^p, m_t^p}^p, [wK_b^p] \sqsubseteq [wK_a^p] \end{aligned} \quad (14)$$

where  $L_{m_{\text{init}}^p, m_t^p}^p$  stands for the usual language of sequences of colors read from  $m_{\text{init}}^p$  to  $m_t^p$ , the additional superscript being used to highlight that we are considering skeleton  $\mathcal{M}^p$  here. From now on, we assume w.l.o.g. that (14) holds, i.e., that  $\forall w \in L_{m_{\text{init}}^p, m_t^p}^p, [wK_b^p] \sqsubseteq [wK_a^p]$ . Intuitively, this means that, for  $\mathcal{P}_1$ , committing to the subarena  $\mathcal{A}_a$  is always at least as good as committing to the subarena  $\mathcal{A}_b$ . Note that this does not imply anything with regard to alternating between the two subarenas, which could a priori be beneficial: we will deal with that soon thanks to  $\mathcal{M}^c$ -selectivity.

Let us define the strategy  $\sigma_1^\# \in \Sigma_1^{\text{ML}}(\mathcal{A})$  of  $\mathcal{P}_1$  in the game  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$  as  $\sigma_1^\# = \sigma_1^a$ , the strategy used in the NE from  $S_{\text{cov}}$  in the subgame  $\mathcal{G}_a$  (we chose this one because of assumption (14):  $\mathcal{G}_a$  is the better subgame of the two for  $\mathcal{P}_1$ ). Strategy  $\sigma_1^\#$  is thus memoryless by definition. Note that  $\sigma_1^\#$  is well-defined on  $\mathcal{A}$  even though the original strategy was on  $\mathcal{A}_a$ , since it is memoryless (i.e., whether the prefix did visit  $\mathcal{A}_b$  or not does not matter). Now we define a corresponding strategy  $\sigma_2^\# \in \Sigma_2(\mathcal{A})$  for  $\mathcal{P}_2$  in  $\mathcal{G}$  that uses a small amount of memory, as follows:

$$\forall \rho \in \text{Hists}_2(\mathcal{A}), \sigma_2^\#(\rho) = \begin{cases} \sigma_2^a(\rho) & \text{if } \rho \text{ never visited } t, \\ \sigma_2^a(\rho) & \text{if the last visit of } t \text{ was followed by an edge in } \mathcal{A}_a, \\ \sigma_2^b(\rho) & \text{otherwise.} \end{cases}$$

Again this strategy is well-defined on  $\mathcal{A}$ . Our goal is to show that the strategy profile  $(\sigma_1^\#, \sigma_2^\#)$  is an NE from all states in  $S_{\text{cov}}$  in the larger game  $\mathcal{G}$ . In particular, Lemma 7 implies that this profile is a couple of  $S_{\text{cov}}$ -optimal strategies.

Formally, we will establish that for all  $s \in S_{\text{cov}}$ , for all  $\sigma_1 \in \Sigma_1(\mathcal{A})$ , for all  $\sigma_2 \in \Sigma_2(\mathcal{A})$ , we have

$$\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2^\#)) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^\#, \sigma_2^\#)) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^\#, \sigma_2)). \quad (15)$$

We begin with the rightmost inequality of Equation (15). Let  $s \in S_{\text{cov}}$  and let  $\sigma_2 \in \Sigma_2(\mathcal{A})$  be an arbitrary strategy for  $\mathcal{P}_2$  in  $\mathcal{G}$ . We denote by  $\sigma_2[\mathcal{A}_a]$  its restriction to (histories of)  $\mathcal{A}_a$ : note that this strategy is well-defined as only edges belonging to  $\mathcal{P}_1$  have been removed in  $\mathcal{A}_a$ .

We have

$$\begin{aligned} \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^\#, \sigma_2^\#)) &= \widehat{\text{col}}(\text{Plays}(\mathcal{A}_a, s, \sigma_1^a, \sigma_2^a)) && \text{because these strategies stay in } \mathcal{A}_a, \\ &\sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}_a, s, \sigma_1^a, \sigma_2[\mathcal{A}_a])) && \text{because } (\sigma_1^a, \sigma_2^a) \text{ is an NE from } s \text{ in } \mathcal{A}_a, \\ &= \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^\#, \sigma_2)) && \text{because these strategies stay in } \mathcal{A}_a, \end{aligned}$$

hence the rightmost inequality is verified.

Now, consider the leftmost inequality of Equation (15). Let  $s \in S_{\text{cov}}$  and let  $\sigma_1 \in \Sigma_1(\mathcal{A})$  be an arbitrary strategy for  $\mathcal{P}_1$  in  $\mathcal{G}$ . Let  $\pi \in \text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2^\#)$  be the only play consistent with  $\sigma_1$  and  $\sigma_2^\#$  from  $s$ . We first consider the case where  $\pi$  never visits  $t$ . If this is the case, then  $\pi$  is also a play in  $\mathcal{A}_a$ . Let  $\sigma'_1 \in \Sigma_1(\mathcal{A}_a)$  be a strategy of  $\mathcal{P}_1$  that mimics  $\sigma_1$  on all histories that belong to  $\mathcal{A}_a$ , except the ones ending in  $t$ , where it plays an arbitrary edge in  $E_a$ . We have

$$\begin{aligned} \widehat{\text{col}}(\pi) &= \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2^\#)) = \widehat{\text{col}}(\text{Plays}(\mathcal{A}_a, s, \sigma'_1, \sigma_2^a)) && \text{because } \pi \text{ stays in } \mathcal{A}_a \text{ and never visits } t, \\ &\sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}_a, s, \sigma_1^a, \sigma_2^a)) && \text{because } (\sigma_1^a, \sigma_2^a) \text{ is an NE from } s \text{ in } \mathcal{A}_a, \\ &= \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^\#, \sigma_2^\#)) && \text{because these strategies stay in } \mathcal{A}_a, \end{aligned}$$

hence the leftmost inequality is verified in the case where  $\pi$  never visits  $t$ .

It remains to consider the case where  $\pi$  does visit  $t$ . Observe that for the moment, we have not used  $\mathcal{M}^c$ -selectivity and the fact that  $s \in S_{\text{cov}}$ : they will be crucial to solve this (more complex) case.

We define  $K_j^c = \widehat{\text{col}}(\{\rho \in \text{Hists}(\mathcal{A}_j, t, \sigma_2^j) \mid \text{out}(\rho) = t\})$ , for  $j \in \{a, b\}$ , that is,  $K_j^c$  contains all (projections to colors of) cycles on  $t$  consistent with  $\sigma_2^j$  (i.e., the strategy from the subgame NE) in subarena  $\mathcal{A}_j$ . Since the unique play  $\pi \in \text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2^\#)$  visits  $t$  at least once, we write  $\pi = \rho \cdot \pi'$  for  $\rho$ , the prefix ending with the first visit of  $t$ . Let  $w = \widehat{\text{col}}(\rho)$ . Observe that

$$\widehat{\text{col}}(\pi) \in [w(K_a^c \cup K_b^c)^*(K_a^p \cup K_b^p)]$$

since  $\sigma_2^\#$  alternates between  $\sigma_2^a$  and  $\sigma_2^b$  depending on what  $\mathcal{P}_1$  plays in  $t$ . Intuitively, either  $\pi$  cycles infinitely often on  $t$  using cycles of  $(K_a^c \cup K_b^c)$ , or it does it for a while, then switches to  $(K_a^p \cup K_b^p)$ , which induces that  $\pi$  commits to a subarena. Thus, we trivially have  $\widehat{\text{col}}(\pi) \sqsubseteq [w(K_a^c \cup K_b^c)^*(K_a^p \cup K_b^p)]$ .

Since  $\mathcal{M}^c$  is a cyclic-cover of  $S_{\text{cov}}$  in  $\mathcal{A}$ , and  $\rho$  starts in  $S_{\text{cov}}$ , we know that for  $m^c = \widehat{\alpha}_{\text{upd}}^c(m_{\text{init}}^c, \widehat{\text{col}}(\rho))$ , and for all  $\rho' \in \text{Hists}(\mathcal{A})$  such that  $\text{in}(\rho') = \text{out}(\rho') = t$ , we have  $\widehat{\alpha}_{\text{upd}}^c(m^c, \widehat{\text{col}}(\rho')) = m^c$ . That is, all cycles on  $t$  are read as cycles on  $m^c$  in  $\mathcal{M}^c$ . This implies that  $K_a^c, K_b^c \subseteq L_{m^c, m^c}$ . Knowing that, we can invoke the  $\mathcal{M}^c$ -selectivity (Equation (8)) of the preference relation to obtain

$$\widehat{\text{col}}(\pi) \sqsubseteq [w(K_a^c)^*] \cup [w(K_b^c)^*] \cup [w(K_a^p \cup K_b^p)].$$

Using Lemma 1, we have

$$\widehat{\text{col}}(\pi) \sqsubseteq [w(K_a^c)^*] \cup [w(K_b^c)^*] \cup [wK_a^p] \cup [wK_b^p].$$

Observe that  $[w(K_j^c)^*] \subseteq [wK_j^p]$ , for  $j \in \{a, b\}$ . Hence, we have

$$\widehat{\text{col}}(\pi) \sqsubseteq [wK_a^p] \cup [wK_b^p].$$

Now, recall that using the  $\mathcal{M}^p$ -monotony of  $\sqsubseteq$ , we assumed that  $\forall w \in L_{m_{\text{init}}^p, m_t^p}^p$ ,  $[wK_b^p] \sqsubseteq [wK_a^p]$ . Since  $\rho$  starts in  $S_{\text{cov}}$ , ends in its first visit to  $t$  and  $\mathcal{M}^p$  is a prefix-cover of  $S_{\text{cov}}$ , this inequality is in particular true for  $w = \widehat{\text{col}}(\rho)$ . Hence, we have

$$\widehat{\text{col}}(\pi) \sqsubseteq [wK_a^p].$$

Now, recall that  $(\sigma_1^a, \sigma_2^a)$  is an NE from  $S_{\text{cov}}$  in  $\mathcal{G}_a$ . Recall also that  $w$  represents the history up to the first visit of  $t$  consistent with  $(\sigma_1, \sigma_2^\#)$ ; it is also consistent with  $(\sigma_1, \sigma_2^a)$  since  $\sigma_2^\#$  follows  $\sigma_2^a$  up to the first visit of  $t$ . Hence, we also have

$$[wK_a^p] \subseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}_a, s, \sigma_2^a)) \sqsubseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}_a, s, \sigma_1^a, \sigma_2^a)).$$

Therefore,

$$\begin{aligned}\widehat{\text{col}}(\pi) &\subseteq \widehat{\text{col}}(\text{Plays}(\mathcal{A}_a, s, \sigma_1^a, \sigma_2^a)) \\ &= \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1^\#, \sigma_2^\#)) \quad \text{because these strategies stay in } \mathcal{A}_a.\end{aligned}$$

Recalling that  $\pi$  is the only play in  $\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2^\#)$ , we are done with proving the leftmost inequality of Equation (15).

Summing up our arguments, we have established that the couple of strategies  $(\sigma_1^\#, \sigma_2^\#)$  is indeed a Nash equilibrium in  $\mathcal{G}$  from  $S_{\text{cov}}$ . Note that this in particular implies, via Lemma 7, that  $\sigma_1^\#$  is an  $S_{\text{cov}}$ -optimal memoryless strategy in  $\mathcal{G}$ .  $\square$

**Memoryless Nash equilibria.** We are now armed to establish the implication sketched earlier. As motivated before, we first state the result in the context of memoryless NE on “covered” arenas, the finite-memory case on general arenas will follow almost trivially. We first show the result for one-player arenas, and use it to obtain the two-player case.

**Lemma 22.** *Let  $\sqsubseteq$  be a preference relation and  $\mathcal{M}^p, \mathcal{M}^c$  be two memory skeletons. Assume that  $\sqsubseteq$  is  $\mathcal{M}^p$ -monotone and  $\mathcal{M}^c$ -selective. Then, for all  $\mathcal{P}_1$ ’s one-player arenas  $\mathcal{A} = (S_1, S_2, E)$ , for all subsets of states  $S_{\text{cov}} \subseteq S$  for which  $\mathcal{M}^p$  is a prefix-cover and  $\mathcal{M}^c$  is a cyclic-cover, there exists a memoryless optimal strategy  $\sigma_1 \in \Sigma_1^{\text{ML}}(\mathcal{A})$  from  $S_{\text{cov}}$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ .*

*Proof.* Let  $\mathfrak{A}$  be the set of all  $\mathcal{P}_1$ ’s one-player arenas, which is closed under subarena operation. By hypothesis,  $\sqsubseteq$  is  $\mathcal{M}^p$ -monotone and  $\mathcal{M}^c$ -selective, and clearly  $\mathcal{P}_2$  has optimal strategies on all his one-player arenas in  $\mathfrak{A}$ , as  $\mathcal{P}_2$  has no choice in any of these arenas.

We proceed by induction on the number of choices in the arena. The base case,  $n_{\mathcal{A}} = 0$ , is trivial. Now let  $n \in \mathbb{N} \setminus \{0\}$  and assume the result holds for  $n_{\mathcal{A}} < n$ . Let  $\mathcal{A} = (S_1, S_2, E) \in \mathfrak{A}$  be a  $\mathcal{P}_1$ ’s one-player arena such that  $n_{\mathcal{A}} = n$ , and let  $S_{\text{cov}} \subseteq S$  be a subset of states for which  $\mathcal{M}^p$  is a prefix-cover and  $\mathcal{M}^c$  is a cyclic-cover. We can invoke Lemma 21 (note that as we only consider  $\mathcal{P}_1$ ’s one-player arenas, the existence of a Nash equilibrium coincides with the existence of an optimal strategy for  $\mathcal{P}_1$  — see Remark 8), and obtain an optimal strategy for  $\mathcal{P}_1$  from  $S_{\text{cov}}$ .  $\square$

We are now ready for the two-player case, which requires monotony and selectivity assumptions for both relation  $\sqsubseteq$  and relation  $\sqsubseteq^{-1}$ .

**Theorem 23.** *Let  $\sqsubseteq$  be a preference relation and  $\mathcal{M}_1^p, \mathcal{M}_2^p, \mathcal{M}_1^c$  and  $\mathcal{M}_2^c$  be four memory skeletons. Assume that  $\sqsubseteq$  is  $\mathcal{M}_1^p$ -monotone and  $\mathcal{M}_1^c$ -selective, and that  $\sqsubseteq^{-1}$  is  $\mathcal{M}_2^p$ -monotone and  $\mathcal{M}_2^c$ -selective. Then, for all arenas  $\mathcal{A} = (S_1, S_2, E)$ , for all subsets of states  $S_{\text{cov}} \subseteq S$  for which  $\mathcal{M}_1^p$  and  $\mathcal{M}_2^p$  are prefix-covers, and  $\mathcal{M}_1^c$  and  $\mathcal{M}_2^c$  are cyclic-covers, there exists a memoryless Nash equilibrium  $(\sigma_1, \sigma_2) \in \Sigma_1^{\text{ML}}(\mathcal{A}) \times \Sigma_2^{\text{ML}}(\mathcal{A})$  from  $S_{\text{cov}}$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ .*

As always, we want to keep our results as general and compositional as possible, hence we consider different skeletons for the two players. As argued before, one can always take a single skeleton for the two players, as well as for the two notions, by taking their product and using Lemma 12 and Lemma 14.

As discussed previously, this theorem in particular implies the existence of memoryless  $S_{\text{cov}}$ -optimal strategies for both players (via Lemma 7).

It is fairly straightforward to prove Theorem 23 once Lemma 21 is established: the main idea is to invoke Lemma 21 for both players while doing the induction, and obtain two Nash equilibria, both of which being memoryless for only one player. Then, to conclude, we resort to Lemma 5 which gives us the possibility to mix these two NE into one that is now memoryless for *both* players.

*Remark 24.* Recall that a crucial hypothesis for Lemma 5 to hold is that our games are *antagonistic*, i.e., that we consider  $\sqsubseteq$  and its inverse relation  $\sqsubseteq^{-1}$ . It is quite interesting to observe that our use of Lemma 5 is the only circumstance in which this hypothesis matters (and it is indeed essential) in all our reasoning.<sup>9</sup>

<sup>9</sup> To be more precise: we wrote everything in the antagonistic setting, but Equation (4) can be written as two inequalities in the general setting —  $\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1', \sigma_2)) \sqsubseteq_1 \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2))$  and  $\widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2')) \sqsubseteq_2 \widehat{\text{col}}(\text{Plays}(\mathcal{A}, s, \sigma_1, \sigma_2))$  — and all our previous reasoning can be rewritten accordingly.

In other words, most of our arguments would hold for two different preference relations,  $\sqsubseteq_1$  and  $\sqsubseteq_2$ , without the hypothesis that  $\sqsubseteq_2$  equals  $(\sqsubseteq_1)^{-1}$ . The problem would be that we cannot mix the two equilibria in a single equilibrium with both strategies being memoryless — while we do need it in the hypothesis of the induction step Lemma 21.

Whether the same reasoning can be extended to (general) Nash equilibria by adapting Lemma 21 to take into account the unavoidable blow-up of memory is a question we leave open for future work. Note that the memory bounds would be awful in any case: as the induction would unroll, the memory needed in the equilibria would build up (essentially one bit of memory is added at each call of the induction step in our easier setting, which is then discarded thanks to Lemma 21).  $\triangleleft$

*Proof.* Let  $\sqsubseteq$  be a preference relation and  $\mathcal{M}_1^p, \mathcal{M}_2^p, \mathcal{M}_1^c$  and  $\mathcal{M}_2^c$  be four memory skeletons such that  $\sqsubseteq$  is  $\mathcal{M}_1^p$ -monotone and  $\mathcal{M}_1^c$ -selective, and  $\sqsubseteq^{-1}$  is  $\mathcal{M}_2^p$ -monotone and  $\mathcal{M}_2^c$ -selective. We consider  $\mathfrak{A}$  the set of all arenas, which is closed under subarena operation. By Lemma 22, we immediately obtain that for all  $\mathcal{P}_1$ 's (resp.  $\mathcal{P}_2$ 's) one-player arenas  $\mathcal{A} = (S_1, S_2, E)$ , for all subsets of states  $S_{\text{cov}} \subseteq S$  for which  $\mathcal{M}_1^p$  and  $\mathcal{M}_2^p$  are prefix-covers, and  $\mathcal{M}_1^c$  and  $\mathcal{M}_2^c$  are cyclic-covers,  $\mathcal{P}_1$  (resp.  $\mathcal{P}_2$ ) has an optimal strategy from  $S_{\text{cov}}$ .

We will proceed by induction on the number of choices in the arena, as described before. The base case,  $n_{\mathcal{A}} = 0$ , is trivial. Now let  $n \in \mathbb{N} \setminus \{0\}$  and assume the result holds for  $n_{\mathcal{A}} < n$ . Let  $\mathcal{A} = (S_1, S_2, E) \in \mathfrak{A}$  be an arena such that  $n_{\mathcal{A}} = n$ , and let  $S_{\text{cov}} \subseteq S$  be a subset of states for which  $\mathcal{M}_1^p$  and  $\mathcal{M}_2^p$  are prefix-covers, and  $\mathcal{M}_1^c$  and  $\mathcal{M}_2^c$  are cyclic-covers.

Focusing on  $\mathcal{P}_1$  and  $\sqsubseteq$ , we invoke Lemma 21 (using  $\mathcal{M}_1^p \otimes \mathcal{M}_2^p$  and  $\mathcal{M}_1^c \otimes \mathcal{M}_2^c$ , and the induction hypothesis) and obtain an NE  $(\sigma_1^\clubsuit, \sigma_2^\clubsuit) \in \Sigma_1^{\text{ML}}(\mathcal{A}) \times \Sigma_2(\mathcal{A})$  from  $S_{\text{cov}}$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ . Note that this NE is only memoryless for  $\mathcal{P}_1$ ! Symmetrically, focusing on  $\mathcal{P}_2$  and  $\sqsubseteq^{-1}$ , we invoke Lemma 21 (using  $\mathcal{M}_1^p \otimes \mathcal{M}_2^p$  and  $\mathcal{M}_1^c \otimes \mathcal{M}_2^c$ , and the induction hypothesis) and obtain an NE  $(\sigma_1^\clubsuit, \sigma_2^\clubsuit) \in \Sigma_1(\mathcal{A}) \times \Sigma_2^{\text{ML}}(\mathcal{A})$  from  $S_{\text{cov}}$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ . Again, note that this NE is only memoryless for  $\mathcal{P}_2$ .

To conclude, it suffices to use Lemma 5:  $(\sigma_1^\clubsuit, \sigma_2^\clubsuit)$  can be mixed with  $(\sigma_1^\clubsuit, \sigma_2^\clubsuit)$  into an equivalent NE  $(\sigma_1^\clubsuit, \sigma_2^\clubsuit) \in \Sigma_1^{\text{ML}}(\mathcal{A}) \times \Sigma_2^{\text{ML}}(\mathcal{A})$ , which is now memoryless for *both* players. This concludes our induction step and our proof.  $\square$

**Finite-memory Nash equilibria and UFM strategies.** Finally, we conclude this section by establishing our result as a corollary.

**Corollary 25.** *Let  $\sqsubseteq$  be a preference relation and  $\mathcal{M}_1^p, \mathcal{M}_2^p, \mathcal{M}_1^c$  and  $\mathcal{M}_2^c$  be four memory skeletons. Assume that  $\sqsubseteq$  is  $\mathcal{M}_1^p$ -monotone and  $\mathcal{M}_1^c$ -selective, and that  $\sqsubseteq^{-1}$  is  $\mathcal{M}_2^p$ -monotone and  $\mathcal{M}_2^c$ -selective. Then, for all arenas  $\mathcal{A} = (S_1, S_2, E)$ , there exists a uniform finite-memory Nash equilibrium  $(\sigma_1, \sigma_2) \in \Sigma_1^{\text{FM}}(\mathcal{A}) \times \Sigma_2^{\text{FM}}(\mathcal{A})$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ , such that strategies  $\sigma_i$  are encoded as Mealy machines  $\Gamma_{\sigma_i} = (\mathcal{M}, \alpha_{\text{next}}^i)$  based on the joint memory skeleton  $\mathcal{M} = \mathcal{M}_1^p \otimes \mathcal{M}_2^p \otimes \mathcal{M}_1^c \otimes \mathcal{M}_2^c$ .*

As usual in this section, we state our result for the slightly stronger notion of Nash equilibria: it involves in particular the *existence of UFM strategies for both players*. As for Theorem 23, we use four memory skeletons to keep the approach compositional and player-based, and we provide strategies based on their product memory. However, if there exists a skeleton  $\mathcal{M}$  that is already such that both  $\sqsubseteq$  and  $\sqsubseteq^{-1}$  are  $\mathcal{M}$ -monotone and  $\mathcal{M}$ -selective, this skeleton suffices to build both strategies (this is transparent in the following proof).

This corollary is fairly easy to obtain. We build the joint memory skeleton  $\mathcal{M}$  as defined above. By Lemma 12 and Lemma 14, we can invoke Theorem 23 on the product arena  $\mathcal{A} \ltimes \mathcal{M}$  and obtain a memoryless NE on it, or equivalently, a finite-memory one on the original arena, through Lemma 4.

*Proof.* Let  $\sqsubseteq$  be a preference relation and let  $\mathcal{M}_1^p, \mathcal{M}_2^p, \mathcal{M}_1^c$  and  $\mathcal{M}_2^c$  be four memory skeletons such that  $\sqsubseteq$  is  $\mathcal{M}_1^p$ -monotone and  $\mathcal{M}_1^c$ -selective, and that  $\sqsubseteq^{-1}$  is  $\mathcal{M}_2^p$ -monotone and  $\mathcal{M}_2^c$ -selective. Let  $\mathcal{A} = (S_1, S_2, E)$  be an arena.

We define  $\mathcal{M} = (\mathcal{M}, m_{\text{init}}, \alpha_{\text{upd}}) = \mathcal{M}_1^p \otimes \mathcal{M}_2^p \otimes \mathcal{M}_1^c \otimes \mathcal{M}_2^c$ , the joint memory skeleton. By Lemma 12,  $\sqsubseteq$  and  $\sqsubseteq^{-1}$  are both  $\mathcal{M}$ -monotone and  $\mathcal{M}$ -selective.



Consider the product arena  $\mathcal{A}' = \mathcal{A} \ltimes \mathcal{M}$ , as defined in Section 2. Recall that  $S' = S \times M$ . By Lemma 14, the set of states  $S'_{\text{cov}} = S \times \{m_{\text{init}}\} \subseteq S'$  is both prefix-covered and cyclic-covered by  $\mathcal{M}$ .

Putting the last two arguments together, we may invoke Theorem 23 on  $\mathcal{A}'$  and obtain a memoryless Nash equilibrium  $(\sigma'_1, \sigma'_2) \in \Sigma_1^{\text{ML}}(\mathcal{A}') \times \Sigma_2^{\text{ML}}(\mathcal{A}')$  from  $S'_{\text{cov}}$  in  $\mathcal{G}' = (\mathcal{A}', \sqsubseteq)$ .

To conclude, it suffices to use Lemma 4 (stated using NE, as discussed in Remark 9): the memoryless equilibrium  $(\sigma'_1, \sigma'_2)$  in the product game  $\mathcal{G}'$  can be seen as a finite-memory equilibrium  $(\sigma_1, \sigma_2) \in \Sigma_1^{\text{FM}}(\mathcal{A}) \times \Sigma_2^{\text{FM}}(\mathcal{A})$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ , where strategies  $\sigma_i$  are encoded as Mealy machines  $\Gamma_{\sigma_i} = (\mathcal{M}, \alpha_{\text{next}}^i)$ .  $\square$

We can also formulate a version of this last result focusing on one-player arenas.

**Corollary 26.** *Let  $\sqsubseteq$  be a preference relation and  $\mathcal{M}^p, \mathcal{M}^c$  be two memory skeletons. Assume that  $\sqsubseteq$  is  $\mathcal{M}^p$ -monotone and  $\mathcal{M}^c$ -selective. Then, for all  $\mathcal{P}_1$ 's one-player arenas  $\mathcal{A} = (S_1, S_2, E)$ , there exists a UFM strategy  $\sigma_1 \in \Sigma_1^{\text{FM}}(\mathcal{A})$  in  $\mathcal{G} = (\mathcal{A}, \sqsubseteq)$ , such that strategy  $\sigma_1$  is encoded as a Mealy machine  $\Gamma_{\sigma_1} = (\mathcal{M}, \alpha_{\text{next}})$  based on the joint memory skeleton  $\mathcal{M} = \mathcal{M}^p \otimes \mathcal{M}^c$ .*

*Proof (Sketch).* The proof is very similar to (and easier than) the proof of Corollary 25, but uses the one-player implication of Lemma 22 instead of Theorem 23.  $\square$

*Remark 27.* Our whole induction scheme is bottom-up, and presented through the prism of covered arenas. It is possible to obtain a similar proof scheme by going top-down and starting from product arenas — which are particular cases of covered arenas (Lemma 14). Our approach pursues two objectives. First, extracting the main technical elements needed and describing them through the concepts of prefix- and cyclic-covers to give a better grasp of how things work and where. Second, providing memoryless optimal strategies in *all* covered arenas (Theorem 23).  $\triangleleft$

## 6 Discussion

We close our paper with a discussion of the assets and limits of our approach, its applicability with regard to the current research landscape, and the directions we aim to follow in future work.

**Technical features of our approach.** As observed through Remark 16, our results are established using fine-grained assumptions and conclusions, in an effort to push the approach to its limits. They also preserve *compositionality*, splitting the reasoning for  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity, and for the two players.

Alongside  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity, we define two other key concepts to solve the technical issues related to the induction on product arenas: *prefix-covers* and *cyclic-covers*. These notions are crucial tools to prove the results in Section 5. They also bear an interesting connection with the *adherence* notion of Kopczyński [33].

**Some advantages.** The aforementioned concepts of prefix-covers and cyclic-covers also have benefits from a practical point of view: given a preference relation  $\sqsubseteq$  and the corresponding memory skeleton  $\mathcal{M}$ , they let us *identify game arenas where memoryless strategies suffice* whereas finite memory (based on  $\mathcal{M}$ ) might be necessary in general. Such arenas are the ones covered by  $\mathcal{M}$ . Hence in practice, this approach permits to obtain UML strategies for many arenas where a coarser approach would only provide UFM ones.

Our approach yields *two methods* to establish that a preference relation (or equivalently a payoff function or a winning condition) admits UFM strategies. The first one, exhibiting appropriate memory skeletons and proving  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity, is based on Theorem 15 and can be used *compositionally* through Corollary 25. The second one follows the *lifting corollary*, Corollary 18: one only has to study the one-player subcases then invoke this result to lift the existence of UFM strategies to the two-player case, without checking for  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity at all. Hence this second method is often painless in practice.

Two interesting facts can be seen through Corollary 18. First, there is *no blow-up in the memory* required when going from one-player games to two-player games: the overall memory simply combines the memory skeletons of the two players. Second, assuming that one has an algorithm to solve<sup>10</sup> one-player games — say

<sup>10</sup> I.e., decide who has a winning strategy from a given state.

for  $\mathcal{P}_1$  — for a winning condition satisfying our hypotheses, this lifting corollary also induces a *naive algorithm for the two-player case for free*: thanks to the bounds on memory, one may enumerate the strategies of the adversary,  $\mathcal{P}_2$  — or guess one if one aims for a non-deterministic algorithm — and solve the corresponding  $\mathcal{P}_1$ ’s game(s) where the strategy of  $\mathcal{P}_2$  is fixed. Note that while such a simple algorithm might not be optimal, it does correspond to the approach giving the best complexity class known for the renowned family of games in  $\text{NP} \cap \text{coNP}$ , such as, e.g., parity or mean-payoff games (e.g., [30]). These last two cases could already be dealt with thanks to Gimbert and Zielonka’s result since they involve memoryless strategies, but now a similar road can be taken for any objective that admits arena-independent finite-memory optimal strategies, such as, e.g., generalized parity games.

**Applicability.** Let us give a quick tour of some classical (combinations of) objectives — expressed through winning conditions, payoffs or preference relations — and assess whether our approach permits to establish the existence of UFM strategies in the corresponding games.

Note that when considering multiple (quantitative) objectives, optimal strategies usually do not exist, and one has to settle for *Pareto-optimal* ones (e.g., [17]). However, in many cases, the (decision) problem under study is as follows: given a threshold (vector), define the winning condition as all the plays achieving at least this threshold, and check for a winning strategy. Hence multi-objective quantitative games are often de facto reduced to qualitative win-lose games for this so-called *threshold problem*. Observe that, given a multi-objective setting, if UFM strategies exist for all threshold problems, then finite-memory strategies suffice to realize the Pareto front (as each point of this front can be considered as a threshold). Therefore, *our approach also enables reasoning about the existence of finite-memory Pareto-optimal strategies in multi-objective games*.

We start our overview with some game settings that fall under the scope of our approach. Obviously, *all memoryless-determined objectives* are among them, since we generalize Gimbert and Zielonka’s work [26]: this includes, e.g., mean-payoff [18], parity [19,42], energy [11] or average-energy games [5]. As established in Section 1, our results encompass all cases where *arena-independent* memory suffices. Hence they permit to rediscover the existence of UFM strategies for games such as, e.g., generalized reachability [20], generalized parity [15], window parity games [9], some variants of window mean-payoff games [13], or lower- and upper-bounded (multi-dimension) energy games [3,5,4]. Our approach can also be useful to extend these known results to more general combinations, either via appropriate memory skeletons or through the lifting corollary (see an application in Section 3.3).

There are many games that do not fit our approach for *good reasons*, as they do not admit UFM strategies in general: e.g., multi-dimension mean-payoff [41], mean-payoff parity [14], or energy mean-payoff games [10]. More interesting are games for which finite-memory strategies exist, but the memory is *arena-dependent*. These notably include games with multi-dimension lower-bounded energy objectives and no upper bound [16,31], or other variants of window mean-payoff games [13]. In such games, the players usually have to keep track of information such as, e.g., the sum of weights along an acyclic path, which is bounded for any given arena, but by a value that grows when the arena grows. Hence the need for memory that grows with the arena parameters. Our results cannot be applied directly to such cases in order to obtain the existence of finite-memory strategies for all games. An adaptation of our approach could potentially be used for subclasses of arenas where the parameters are bounded (in order to regain a skeleton working on all arenas of the class).

**Comparison with related work.** We already discussed extensively the most important related articles [25,26,32,1,34,40] in Section 1, alongside a technical comparison between our work and Gimbert and Zielonka’s seminal result [26]. Here, we simply want to highlight interesting directions of research inspired by some of these papers.

First, Aminof and Rubin provide a simpler (but incomplete) approach to memoryless determinacy through the prism of first-cycle games in [1]: a similar take on finite-memory determinacy could be appealing — it could provide sufficient conditions easier to test than  $\mathcal{M}$ -monotony and  $\mathcal{M}$ -selectivity.

Second, Kopczyński establishes sufficient (and relaxed) conditions to ensure the existence of UML strategies *for one player, in two-player games*, in [32]: it would be interesting to study the corresponding problem in the finite-memory case. Indeed, in many games where infinite memory is needed, it is only the case for

one of the players (e.g., [41,14,10]) and conditions à la Kopczyński could thus prove useful. Note that this is different from Theorem 17, which gives a sufficient *and necessary* condition but for one-player games only.

Finally, recall that Le Roux et al. give a rather tight characterization of combinations of objectives preserving the sufficiency of finite-memory strategies in [40]. Their techniques, as well as the scope of their results, are somewhat orthogonal to ours. Whether both approaches can be intertwined to obtain results on more general settings remains an open question.

**Limits and future work.** To close this paper, we recall three limits of our approach, and the corresponding open problems.

First, as explained throughout the paper, our results cover all cases where *arena-independent* memory suffices, and are *limited to these cases*. We have argued that the approach cannot be fully lifted to the general case, for good reasons, as the lifting corollary breaks in some situations (Section 1). Still, we have hope to generalize our approach to some extent to the *arena-dependent* case, through some *function* associating memory skeletons to arenas, as discussed in Section 1. Obtaining a lifting corollary — under well-chosen conditions — in the arena-dependent case would be of tremendous help in practice: see for example [5,4,10]. Hence this is clearly the next step in our quest.

Second, our result is a characterization *instantiated by a memory skeleton*  $\mathcal{M}$ . While the lifting corollary is helpful in applications, it would be fantastic to be able to find an appropriate skeleton automatically, and to be able to determine if a given skeleton is minimal (with regard to a preference relation). This paper is a first step toward these long-term objectives.

Lastly, as explained in Remark 6 and Remark 24, most of our arguments carry over to the case of *general Nash equilibria*. That is, when considering not necessarily antagonistic games where the two players use different, not necessarily inverse, preference relations. Whether our approach can be adapted in this case, at the price of an unavoidable blow-up of memory, is an open question worth considering. In particular, we want to study the links between our results (including the lifting from one-player to two-player games) and recent results lifting finite-memory determinacy from two-player to multi-player games [39].

## References

1. Benjamin Aminof and Sasha Rubin. First-cycle games. *Inf. Comput.*, 254:195–216, 2017.
2. Roderick Bloem, Krishnendu Chatterjee, and Barbara Jobstmann. Graph games and reactive synthesis. In Edmund M. Clarke, Thomas A. Henzinger, Helmut Veith, and Roderick Bloem, editors, *Handbook of Model Checking*, pages 921–962. Springer, 2018.
3. 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.
4. 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.
5. Patricia Bouyer, Nicolas Markey, Mickael Randour, Kim G. Larsen, and Simon Laursen. Average-energy games. *Acta Inf.*, 55(2):91–127, 2018.
6. Romain Brenguier, Lorenzo Clemente, Paul Hunter, Guillermo A. Pérez, Mickael Randour, Jean-François Raskin, Ocan Sankur, and Mathieu Sassolas. Non-zero sum games for reactive synthesis. In Adrian-Horia Dediu, Jan Janousek, Carlos Martín-Vide, and Bianca Truthe, editors, *Language and Automata Theory and Applications - 10th International Conference, LATA 2016, Prague, Czech Republic, March 14-18, 2016, Proceedings*, volume 9618 of *Lecture Notes in Computer Science*, pages 3–23. Springer, 2016.
7. Thomas Brihaye, Florent Delgrange, Youssouf Oualhadj, and Mickael Randour. Life is random, time is not: Markov decision processes with window objectives. In Fokkink and van Glabbeek [22], pages 8:1–8:18.
8. 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.

9. 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.
10. Véronique Bruyère, Quentin Hautem, Mickael Randour, and Jean-François Raskin. Energy mean-payoff games. In Fokkink and van Glabbeek [22], pages 21:1–21:17.
11. Arindam Chakrabarti, Luca de Alfaro, Thomas A. Henzinger, and Mariëlle Stoelinga. Resource interfaces. In Rajeev Alur and Insup Lee, editors, *EMSOFT*, volume 2855 of *Lecture Notes in Computer Science*, pages 117–133. Springer, 2003.
12. Krishnendu Chatterjee and Laurent Doyen. Energy parity games. *Theor. Comput. Sci.*, 458:49–60, 2012.
13. Krishnendu Chatterjee, Laurent Doyen, Mickael Randour, and Jean-François Raskin. Looking at mean-payoff and total-payoff through windows. *Inf. Comput.*, 242:25–52, 2015.
14. Krishnendu Chatterjee, Thomas A. Henzinger, and Marcin Jurdzinski. Mean-payoff parity games. In *20th IEEE Symposium on Logic in Computer Science (LICS 2005), 26-29 June 2005, Chicago, IL, USA, Proceedings*, pages 178–187. IEEE Computer Society, 2005.
15. 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.
16. Krishnendu Chatterjee, Mickael Randour, and Jean-François Raskin. Strategy synthesis for multi-dimensional quantitative objectives. *Acta Inf.*, 51(3-4):129–163, 2014.
17. 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.
18. Andrzej Ehrenfeucht and Jan Mycielski. Positional strategies for mean payoff games. *Int. Journal of Game Theory*, 8(2):109–113, 1979.
19. E. Allen Emerson and Charanjit S. Jutla. The complexity of tree automata and logics of programs. In *FOCS*, pages 328–337. IEEE Computer Society, 1988.
20. Nathanaël Fijalkow and Florian Horn. The surprising complexity of reachability games. *CoRR*, abs/1010.2420, 2010.
21. Nathanaël Fijalkow, Florian Horn, Denis Kuperberg, and Michal Skrzypczak. Trading bounds for memory in games with counters. In Halldórsson et al. [29], pages 197–208.
22. 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*. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2019.
23. 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.
24. Hugo Gimbert and Edon Kelmendi. Two-player perfect-information shift-invariant submixing stochastic games are half-positional. Unpublished, 2014.
25. Hugo Gimbert and Wiesław Zielonka. When can you play positionally? In Jirí Fiala, Václav Koubek, and Jan Kratochvíl, editors, *Mathematical Foundations of Computer Science 2004, 29th International Symposium, MFCS 2004, Prague, Czech Republic, August 22-27, 2004, Proceedings*, volume 3153 of *Lecture Notes in Computer Science*, pages 686–697. Springer, 2004.
26. 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.
27. Hugo Gimbert and Wiesław Zielonka. Pure and Stationary Optimal Strategies in Perfect-Information Stochastic Games with Global Preferences. Unpublished, December 2009.
28. Erich Grädel, Wolfgang Thomas, and Thomas Wilke, editors. *Automata, Logics, and Infinite Games: A Guide to Current Research [outcome of a Dagstuhl seminar, February 2001]*, volume 2500 of *Lecture Notes in Computer Science*. Springer, 2002.

29. Magnús M. Halldórsson, Kazuo Iwama, Naoki Kobayashi, and Bettina Speckmann, editors. *Automata, Languages, and Programming - 42nd International Colloquium, ICALP 2015, Kyoto, Japan, July 6-10, 2015, Proceedings, Part II*, volume 9135 of *Lecture Notes in Computer Science*. Springer, 2015.
30. Marcin Jurdzinski. Deciding the winner in parity games is in  $UP \cap co-UP$ . *Inf. Process. Lett.*, 68(3):119–124, 1998.
31. Marcin Jurdzinski, Ranko Lazic, and Sylvain Schmitz. Fixed-dimensional energy games are in pseudo-polynomial time. In Halldórsson et al. [29], pages 260–272.
32. 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.
33. Eryk Kopczyński. *Half-positional Determinacy of Infinite Games*. PhD thesis, Warsaw University, 2008.
34. Donald A. Martin. Borel determinacy. *Annals of Mathematics*, 102(2):363–371, 1975.
35. Martin J. Osborne and Ariel Rubinstein. *A course in game theory*. The MIT Press, Cambridge, USA, 1994.
36. 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.
37. Stéphane Le Roux. Infinite sequential Nash equilibrium. *Logical Methods in Computer Science*, 9(2), 2013.
38. Stéphane Le Roux. Concurrent games and semi-random determinacy. In Igor Potapov, Paul G. Spirakis, and James Worrell, editors, *43rd International Symposium on Mathematical Foundations of Computer Science, MFCS 2018, August 27-31, 2018, Liverpool, UK*, volume 117 of *LIPIcs*, pages 40:1–40:15. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018.
39. Stéphane Le Roux and Arno Pauly. Extending finite-memory determinacy to multi-player games. *Inf. Comput.*, 261(Part):676–694, 2018.
40. 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 fuer Informatik, 2018.
41. 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.
42. Wiesław Zielonka. Infinite games on finitely coloured graphs with applications to automata on infinite trees. *Theor. Comput. Sci.*, 200(1-2):135–183, 1998.

## A Counterexample to a general lifting corollary

We discuss in full details the counterexample presented in Section 1. Let  $C = \mathbb{Z}$ . We consider the following two winning conditions:

$$W_1 = \{c_1 c_2 \dots \in C^\omega \mid \liminf_{n \rightarrow \infty} \sum_{i=1}^n c_i = +\infty\},$$

$$W_2 = \{c_1 c_2 \dots \in C^\omega \mid \sum_{i=1}^n c_i = 0 \text{ for infinitely many } n\text{'s}\}.$$

If the play obtained by playing a game is  $\pi$ ,  $\mathcal{P}_1$  wins if and only if  $\widehat{\text{col}}(\pi)$  lies in  $W = W_1 \cup W_2$ , and  $\mathcal{P}_2$  wins if and only if  $\widehat{\text{col}}(\pi)$  lies in  $\overline{W} = C^\omega \setminus W = \overline{W_1} \cap \overline{W_2}$  (which corresponds to the description given in Section 1). We prove that  $\mathcal{P}_1$  and  $\mathcal{P}_2$  have finite-memory optimal strategies in their respective one-player games.

Let us fix some terminology beforehand: we say that a cycle in an arena is a *zero* cycle if the sum of its weights is zero, a *positive* cycle if this sum is strictly positive, and a *negative* cycle if this sum is strictly negative.

We first consider  $\mathcal{P}_1$ 's one-player games. In a one-player arena,  $\mathcal{P}_1$  can create a play  $\pi$  such that  $\widehat{\text{col}}(\pi) \in W_1$  if and only if there is a reachable positive cycle. In this case,  $\mathcal{P}_1$  can win with a memoryless strategy (simply reaching the cycle and then looping in it). If that is not possible, in order to win,  $\mathcal{P}_1$  has to induce



a play  $\pi$  such that  $\widehat{\text{col}}(\pi) \in W_2$ . We show that if possible, this can be done using finite memory. Let us assume that there exists a play  $\pi = e_1 e_2 \dots$  such that  $\widehat{\text{col}}(\pi) = c_1 c_2 \dots \in W_2$ . Let us consider two indices  $k, l \in \mathbb{N}$  such that  $k < l$ ,  $e_k = e_l$ ,  $\sum_{i=1}^k c_i = 0$ , and  $\sum_{i=1}^l c_i = 0$ . Such two indices necessarily exist, as there are finitely many edges in the arena, but infinitely many indices for which the running sum of weights is 0. Notice in particular that  $\sum_{i=k+1}^l c_i = 0$ . Now, consider the play

$$\pi' = e_1 \dots e_k e_{k+1} \dots e_l e_{k+1} \dots e_l e_{k+1} \dots,$$

with the sequence of edges  $e_{k+1} \dots e_l$  repeating *ad infinitum* ( $\pi'$  is a “lasso”). This is a valid play since  $e_k = e_l$ . Moreover, we have that  $\widehat{\text{col}}(\pi') \in W_2$  as after repeating  $m$  times the sequence  $e_{k+1} \dots e_l$ , the sum of the weights equals  $\sum_{i=1}^k c_i + m \cdot \sum_{i=k+1}^l c_i = 0 + m \cdot 0 = 0$ . The play  $\pi'$  can be implemented with finite memory, as it consists of a finite prefix and a repeated finite sequence, which corresponds to a zero cycle.

We now turn our attention to  $\mathcal{P}_2$ ’s one-player games;  $\mathcal{P}_2$  wins a play  $\pi$  such that  $\widehat{\text{col}}(\pi) = c_1 c_2 \dots \in C^\omega$  if and only if

$$\liminf_{n \rightarrow \infty} \sum_{i=1}^n c_i < +\infty \wedge \sum_{i=1}^n c_i = 0 \text{ for at most finitely many } n\text{'s.}$$

In a one-player arena, if there is a reachable negative cycle,  $\mathcal{P}_2$  can ensure to win by pumping it forever, and can therefore win with a memoryless strategy. We now consider an arena that has no reachable negative cycle. As we did for  $\mathcal{P}_1$ , we show that if  $\mathcal{P}_2$  can win a game in such an arena, then he can do so using finite memory. If  $\mathcal{P}_2$  can win, let  $\pi = e_1 e_2 \dots$  be a winning play for  $\mathcal{P}_2$ , i.e.,  $\widehat{\text{col}}(\pi) = c_1 c_2 \dots \in \overline{W_1} \cap \overline{W_2}$ . Let  $s$  be a state visited infinitely often when  $\pi$  is played, and  $m \in \mathbb{N}$  be the first index such that  $\text{out}(e_m) = s$ . We can decompose  $\pi$  into a finite prefix  $e_1 \dots e_m$  followed by an infinite sequence of cycles, all starting in  $s$ . Since there is no negative cycle, we cannot have that infinitely many of these cycles are positive, as this would imply that  $\widehat{\text{col}}(\pi) \in W_1$ . Thus, infinitely many zero cycles are taken from  $s$ . As  $\widehat{\text{col}}(\pi) \in \overline{W_2}$ , there exists such a cycle  $e_k \dots e_l$  (that is,  $\text{in}(e_k) = \text{out}(e_l) = s$  and  $\sum_{i=k}^l c_i = 0$ ) such that for all  $k \leq n \leq l$ , it holds that  $\sum_{i=1}^n c_i \neq 0$ . This also implies that  $\sum_{i=1}^{k-1} c_i \neq 0$ , i.e., the history up to this cycle has a non-zero sum. Now, let us consider the play

$$\pi' = e_1 \dots e_{k-1} e_k \dots e_l e_k \dots e_l e_k \dots,$$

with the sequence of edges  $e_k \dots e_l$  repeating *ad infinitum* ( $\pi'$  is a “lasso”). This is a valid play as  $\text{in}(e_k) = \text{out}(e_l)$ . As  $\sum_{i=k}^l c_i = 0$ , we have that  $\widehat{\text{col}}(\pi') \in \overline{W_1}$ . Moreover, every time the cycle starts again, the running sum of weights is equal to the same value:  $\sum_{i=1}^{k-1} c_i \neq 0$ . Therefore, as the running sum of weights does not reach zero the first time the cycle is taken, and it also never reaches zero along the cycle, it can never reach zero after index  $k - 1$ . Hence,  $\widehat{\text{col}}(\pi')$  is not in  $W_2$  either, and  $\pi'$  is winning for  $\mathcal{P}_2$ . For the same reason as for  $\mathcal{P}_1$ , play  $\pi'$  only requires finite memory to be implemented.

As argued in Section 1, the two-player game from Figure 1 illustrates that  $\mathcal{P}_1$  might need infinite memory to play optimally in the two-player case. This proves that Gimbert and Zielonka’s approach cannot work in full generality in the finite-memory case, as we cannot obtain the existence of finite-memory optimal strategies in all two-player games from the existence of finite-memory optimal strategies in all one-player games.