

Strategy Synthesis for Multi-Dimensional Quantitative Objectives

Krishnendu Chatterjee^{1,*}, Mickael Randour^{2,†}, and Jean-François Raskin^{3,‡}

¹ IST Austria (Institute of Science and Technology Austria)

² Computer Science Department, Université de Mons (UMONS), Belgium

³ Département d'Informatique, Université Libre de Bruxelles (U.L.B.), Belgium

Abstract. Multi-dimensional mean-payoff and energy games provide the mathematical foundation for the quantitative study of reactive systems, and play a central role in the emerging quantitative theory of verification and synthesis. In this work, we study the strategy synthesis problem for games with such multi-dimensional objectives along with a parity condition, a canonical way to express ω -regular conditions. While in general, the winning strategies in such games may require infinite memory, for synthesis the most relevant problem is the construction of a finite-memory winning strategy (if one exists). Our main contributions are as follows. First, we show a tight exponential bound (matching upper and lower bounds) on the memory required for finite-memory winning strategies in both multi-dimensional mean-payoff and energy games along with parity objectives. This significantly improves the triple exponential upper bound for multi energy games (without parity) that could be derived from results in literature for games on VASS (vector addition systems with states). Second, we present an optimal symbolic and incremental algorithm to compute a finite-memory winning strategy (if one exists) in such games. Finally, we give a complete characterization of when finite memory of strategies can be traded off for randomness. In particular, we show that for one-dimension mean-payoff parity games, randomized memoryless strategies are as powerful as their pure finite-memory counterparts.

1 Introduction

Two-player games on graphs provide the mathematical foundation to study many important problems in computer science. Game-theoretic formulations have especially proved useful for synthesis [25,42,40], verification [3], refinement [36], and compatibility checking [26] of reactive systems, as well as in analysis of emptiness of automata [45].

Games played on graphs are repeated games that proceed for an infinite number of rounds. The *state* space of the graph is partitioned into player 1 states and player 2 states (player 2 is adversary to player 1). The game starts at an initial state, and if the current state is a player 1 (resp. player 2) state, then player 1 (resp. player 2) chooses an outgoing *edge*. This choice is made according to a *strategy* of the player: given the sequence of visited states, a *pure* (resp. *randomized*) strategy chooses an outgoing edge (resp. probability distribution over outgoing edges). This process of choosing edges is repeated forever, and gives rise to an outcome of the game, called a *play*, that consists of the infinite sequence of states that are visited. When randomized strategies are used, there is in general not a unique outcome, but a set of possible outcomes, as the choice of edges is stochastic rather than deterministic.

Traditionally, games on graphs have been studied with Boolean objectives such as reachability, liveness, ω -regular conditions formalized as the canonical parity objectives, strong fairness objectives, etc [35,31,32,48,45,34]. While games with *quantitative* objectives have been studied in the game theory literature [30,49,38], their application in synthesis and other problems in verification is quite recent. The two classical quantitative objectives that are most relevant in verification and synthesis are the *mean-payoff* and *energy* objectives. In games on graphs with quantitative objectives, the game graph is equipped with a weight function that assigns integer-valued weights to every edge. For mean-payoff objectives, the goal of player 1 is to ensure that the long-run average of the weights is above a threshold. For energy objectives, the goal of player 1 is to ensure that the sum of the weights stays above 0 at all times. In applications of verification and synthesis, the quantitative objectives that typically arise are (i) multi-dimensional quantitative objectives (i.e., conjunction of several quantitative objectives), e.g., to express properties like the average response

* Author supported by Austrian Science Fund (FWF) Grant No P 23499-N23, FWF NFN Grant No S11407 (RiSE), ERC Start Grant (279307: Graph Games), Microsoft faculty fellowship.

† Author supported by F.R.S.-FNRS. fellowship.

‡ Author supported by ERC Starting Grant (279499: inVEST).

time between a grant and a request is below a given threshold v_1 , and the average number of unnecessary grants is below threshold v_2 ; and (ii) conjunction of quantitative objectives with a Boolean objective, such as a mean-payoff parity objective that can express properties like the average response time is below a threshold along with satisfying a liveness property. In summary, the quantitative objectives can express properties related to resource requirements, performance, and robustness; multiple objectives can express the different, potentially dependent or conflicting objectives; and the Boolean objective specifies functional properties such as liveness or fairness. The game theoretic framework of multi-dimensional quantitative games and games with conjunction of quantitative and Boolean objectives has recently been shown to have many applications in verification and synthesis, such as synthesizing systems with quality guarantee [6], synthesizing robust systems [7], performance aware synthesis of concurrent data structure [14], analyzing permissivity in games and synthesis [11], simulation between quantitative automata [19], generalizing Boolean simulation to quantitative simulation distance [15], etc. Moreover, multi-dimensional energy games are equivalent to a decidable class of games on VASS (vector addition systems with states). This model is equivalent to games over multi-counter systems and Petri nets [12].

In literature, there are many recent works on the theoretical analysis of multi-dimensional quantitative games, such as, mean-payoff parity games [22,11], energy-parity games [17], multi-dimensional energy games [20], and multi-dimensional mean-payoff games [20,47]. Most of these works focus on establishing the computational complexity of the problem of deciding if player 1 *has a winning strategy*. From the perspective of synthesis and other related problems in verification, the most important problem is to obtain a witness *finite-memory* winning strategy (if one exists). The winning strategy in the game corresponds to the desired controller for (or implementation of) the system in synthesis, and for implementability a finite-memory strategy is essential. In this work we consider the problem of finite-memory strategy synthesis in multi-dimensional quantitative games in conjunction with parity objectives, and the problem of existence of memory-efficient randomized strategies for such games. These are some of the core and foundational problems in the emerging theory of quantitative verification and synthesis.

Our contributions. In this work, we give an extended presentation of the results of [23], the first study of multi-dimensional energy and mean-payoff objectives in conjunction with parity objectives. Conjunction of parity objectives with multi-dimensional quantitative objectives had never been considered before [23]. Our presentation is based on the journal publication [24]. Since we consider the synthesis of finite-memory strategies, it follows from the results of [20] that both the problems (multi-dimensional energy with parity and multi-dimensional mean-payoff with parity) are equivalent. Our main results for finite-memory strategy synthesis for multi-dimensional energy parity games are as follows. (i) **Optimal memory bounds.** We first show that memory of exponential size is sufficient in multi-dimensional energy parity games. Our result is a significant improvement over the result that can be obtained naively from the results known in literature that yields a triple exponential bound, even in the case of multi-dimensional energy games without parity. Second, we show a matching lower bound by presenting a family of game graphs where exponential memory is necessary in multi-dimensional energy games (without parity), even when all the transition weights belong to $\{-1, 0, +1\}$. Thus we establish *optimal memory bounds* for the finite-memory strategy synthesis problem. (ii) **Symbolic and incremental algorithm.** We present a *symbolic* algorithm (in the sense of [28], i.e., using a compact antichain representation of sets by their minimal elements) to compute a finite-memory winning strategy, if one exists, for multi-dimensional energy parity games. Our algorithm is parameterized by the range of energy levels to consider during its execution. So, we can use it in an *incremental approach*: first, we search for finite-memory winning strategies with a small range, and increment the range only when necessary. We also establish a bound on the maximal range to consider which ensures completeness of the incremental approach. In the worst case the algorithm requires exponential time. Since exponential size memory is required (and also the decision problem is coNP-complete [20]), the worst case exponential bound can be considered as *optimal*. Moreover, as our algorithm is symbolic and incremental, in most relevant problems in practice, it is expected to be efficient. (iii) **Randomized strategies.** We also consider when the (pure) finite-memory strategies can be traded off for conceptually much simpler randomized strategies. We show that for energy objectives randomization is not helpful (as energy objectives are similar in spirit with safety objectives), even with only one player, neither it is for two-player multi-dimensional mean-payoff objectives. However, randomized memoryless strategies suffice for one-player multi-dimensional mean-payoff parity games. For the important special case of mean-payoff parity objectives (conjunction of a single mean-payoff and parity objectives), we show that in games, finite-memory strategies can be traded off for randomized memoryless strategies.

Related works. This paper extends the results presented in its preceding conference version [23] and gives a full presentation of the technical details published in [24]. Games with a single mean-payoff objective have been studied in [30,49], and games with a single energy objective in [16]; their equivalence was established in [10]. One-dimensional mean-payoff parity games problem has been studied in [22]: an exponential algorithm was given to decide if there exists a winning strategy (which in general was shown to require infinite memory); and an improved algorithm was presented in [11]. One-dimensional energy parity games problem has been studied in [17]: it was shown that deciding the existence of a winning strategy is in $\text{NP} \cap \text{coNP}$, and an exponential algorithm was given. It was also shown in [17] that, for one-dimensional energy parity objectives, finite-memory strategies with exponential memory are sufficient, and the decision problem for mean-payoff parity objective can be reduced to energy parity objective. Alternative objectives based on the mean-payoff but with improved tractability in the one-dimensional setting were considered in [21]. Extension of the worst-case threshold problem - the classical decision problem on mean-payoff games - with guarantees on the expected performance faced to a stochastic adversary was studied in [13].

Games on VASS with several different winning objectives have been studied in [12], and from the results of [12] it follows that in multi-dimensional energy games, winning strategies with finite memory are sufficient (and a triple exponential bound on memory can be derived from the results). The complexity of multi-dimensional energy and mean-payoff games was studied in [20,47]. It was shown in [20] that in general, winning strategies in multi-dimensional mean-payoff games require infinite memory, whereas for multi-dimensional energy games, finite-memory strategies are sufficient. Moreover, for finite-memory strategies, the multi-dimensional mean-payoff and energy games coincide, and optimal computational complexity for deciding the existence of a winning strategy was established as coNP -complete [20,47]. Multi-dimensional mean-payoff games with infinite-memory strategies were studied in [47], and optimal computational complexity results were established. Various decision problems over multi-dimensional energy games were studied in [33].

2 Preliminaries

We consider two-player game structures and denote the two *players* by \mathcal{P}_1 and \mathcal{P}_2 .

Multi-weighted two-player game structures. A *multi-weighted two-player game structure* is a tuple $G = (S_1, S_2, s_{\text{init}}, E, k, w)$ where (i) S_1 and S_2 resp. denote the finite sets of *states* belonging to \mathcal{P}_1 and \mathcal{P}_2 , with $S_1 \cap S_2 = \emptyset$; (ii) $s_{\text{init}} \in S = S_1 \cup S_2$ is the initial state; (iii) $E \subseteq S \times S$ is the set of *edges* such that for all $s \in S$, there exists $s' \in S$ such that $(s, s') \in E$; (iv) $k \in \mathbb{N}$ is the *dimension* of the weight vectors; and (v) $w: E \rightarrow \mathbb{Z}^k$ is the multi-weight labeling function. The game structure G is *one-player* if $S_2 = \emptyset$. A *play* in G is an infinite sequence of states $\pi = s_0 s_1 s_2 \dots$ such that $s_0 = s_{\text{init}}$ and for all $i \geq 0$, we have $(s_i, s_{i+1}) \in E$. The *prefix* up to the n -th state of play $\pi = s_0 s_1 \dots s_n \dots$ is the finite sequence $\pi(n) = s_0 s_1 \dots s_n$. Let $\text{First}(\pi(n))$ and $\text{Last}(\pi(n))$ resp. denote s_0 and s_n , the first and last states of $\pi(n)$. A prefix $\pi(n)$ belongs to \mathcal{P}_i , $i \in \{1, 2\}$, if $\text{Last}(\pi(n)) \in S_i$. The set of plays of G is denoted by $\text{Plays}(G)$ and the corresponding set of prefixes is denoted by $\text{Pref}(G)$. The set of prefixes that belong to \mathcal{P}_i is denoted by $\text{Pref}_i(G)$. The *energy level vector* of a sequence of states $\rho = s_0 s_1 \dots s_n$ such that for all $i \geq 0$, we have $(s_i, s_{i+1}) \in E$, is $\text{EL}(\rho) = \sum_{i=0}^{n-1} w(s_i, s_{i+1})$ and the *mean-payoff vector* of a play $\pi = s_0 s_1 \dots$ is $\text{MP}(\pi) = \liminf_{n \rightarrow \infty} \frac{1}{n} \text{EL}(\pi(n))$.

Parity. A game structure G is extended with a priority function $p: S \rightarrow \mathbb{N}$ to the structure $G_p = (S_1, S_2, s_{\text{init}}, E, k, w, p)$. Given a play $\pi = s_0 s_1 s_2 \dots$, we define $\text{Inf}(\pi) = \{s \in S \mid \forall m \geq 0, \exists n > m \text{ such that } s_n = s\}$, the set of states that appear infinitely often along π . The *parity* of a play π is defined as $\text{Par}(\pi) = \min \{p(s) \mid s \in \text{Inf}(\pi)\}$. In the following definitions, we denote any game by G_p with no loss of generality.

Strategies. Given a finite set A , a *probability distribution* on A is a function $p: A \rightarrow [0, 1]$ such that $\sum_{a \in A} p(a) = 1$. We denote the set of probability distributions on A by $\mathcal{D}(A)$. A *pure strategy* for \mathcal{P}_i , $i \in \{1, 2\}$, in G_p is a function $\lambda_i: \text{Pref}_i(G_p) \rightarrow S$ such that for all $\rho \in \text{Pref}_i(G_p)$, we have $(\text{Last}(\rho), \lambda_i(\rho)) \in E$. A *(behavioral) randomized strategy* is a function $\lambda_i: \text{Pref}_i(G_p) \rightarrow \mathcal{D}(S)$ such that for all $\rho \in \text{Pref}_i(G_p)$, we have $\{(\text{Last}(\rho), s) \mid s \in S, \lambda_i(\rho)(s) > 0\} \subseteq E$. A pure strategy λ_i for \mathcal{P}_i has *finite memory* if it can be encoded by a deterministic Moore machine $(M, m_0, \alpha_u, \alpha_n)$ where M is a finite set of states (the memory of the strategy), $m_0 \in M$ is the initial memory state, $\alpha_u: M \times S \rightarrow M$ is an update function, and $\alpha_n: M \times S_i \rightarrow S$ is the next-action function. If the game is in $s \in S_i$ and $m \in M$ is the current memory value, then the strategy chooses $s' = \alpha_n(m, s)$ as the next state of the game. When the game leaves a state $s \in S$, the memory is updated to $\alpha_u(m, s)$. Formally, $(M, m_0, \alpha_u, \alpha_n)$ defines the strategy λ_i such that $\lambda_i(\rho \cdot s) =$

$\alpha_n(\hat{\alpha}_u(m_0, \rho), s)$ for all $\rho \in \mathcal{S}^*$ and $s \in S_i$, where $\hat{\alpha}_u$ extends α_u to sequences of states as expected. A pure strategy is *memoryless* if $|M| = 1$, i.e., it does not depend on history but only on the current state of the game. Similar definitions hold for finite-memory randomized strategies, such that the next-action function α_n is randomized, while the update function α_u remains deterministic. We resp. denote by $\Lambda_i, \Lambda_i^{PF}, \Lambda_i^{PM}, \Lambda_i^{RM}$ the sets of general (i.e., possibly randomized and infinite-memory), pure finite-memory, pure memoryless and randomized memoryless strategies for player \mathcal{P}_i .

Given a prefix $\rho \in \text{Prefix}_i(G_p)$ belonging to player \mathcal{P}_i , and a strategy $\lambda_i \in \Lambda_i$ of this player, we define the *support* of the probability distribution defined by λ_i as $\text{Supp}_{\lambda_i}(\rho) = \{s \in S \mid \lambda_i(\rho)(s) > 0\}$, with $\lambda_i(\rho)(s) = 1$ if λ_i is pure and $\lambda_i(\rho) = s$. A play π is said to be *consistent* with a strategy λ_i of \mathcal{P}_i if for all $n \geq 0$ such that $\text{Last}(\pi(n)) \in S_i$, we have $\text{Last}(\pi(n+1)) \in \text{Supp}_{\lambda_i}(\pi(n))$. Given two strategies, λ_1 for \mathcal{P}_1 and λ_2 for \mathcal{P}_2 , we define $\text{Outcome}_{G_p}(\lambda_1, \lambda_2) = \{\pi \in \text{Plays}(G_p) \mid \pi \text{ is consistent with } \lambda_1 \text{ and } \lambda_2\}$, the set of possible *outcomes* of the game. Note that if both strategies λ_1 and λ_2 are pure, we obtain a unique play $\pi = s_0 s_1 s_2 \dots$ such that for all $j \geq 0, i \in \{1, 2\}$, if $s_j \in S_i$, then we have $s_{j+1} = \lambda_i(s_j)$.

Given the initial state s_{init} and strategies for both players $\lambda_1 \in \Lambda_1, \lambda_2 \in \Lambda_2$, we obtain a Markov chain. Thus, every event $\mathcal{A} \subseteq \text{Plays}(G_p)$, a measurable set of plays, has a uniquely defined probability [46] (Carathéodory's extension theorem induces a unique probability measure on the Borel σ -algebra over $\text{Plays}(G_p)$). We denote by $\mathbb{P}_{s_{init}}^{\lambda_1, \lambda_2}(\mathcal{A})$ the probability that a play belongs to \mathcal{A} when the game starts in s_{init} and is played consistently with λ_1 and λ_2 . Let $f : \text{Plays}(G_p) \rightarrow \mathbb{R}$ be a measurable function, we denote $\mathbb{E}_{s_{init}}^{\lambda_1, \lambda_2}(f)$ the expected value of function f over a play when the game starts in s_{init} and is played consistently with λ_1 and λ_2 . We use the same notions for prefixes by naturally extending them to their infinite counterparts.

Objectives. An *objective* for \mathcal{P}_1 in G_p is a set of plays $\phi \subseteq \text{Plays}(G_p)$. We consider several kinds of objectives:

- *Multi Energy objectives.* Given an initial natural energy vector $v_0 \in \mathbb{N}^k$, the objective $\text{PosEnergy}_{G_p}(v_0) = \{\pi \in \text{Plays}(G_p) \mid \forall n \geq 0 : v_0 + \text{EL}(\pi(n)) \in \mathbb{N}^k\}$ requires that the energy level in all dimensions stays positive at all times.
- *Multi Mean-payoff objectives.* Given a rational threshold vector $v \in \mathbb{Q}^k$, the objective $\text{MeanPayoff}_{G_p}(v) = \{\pi \in \text{Plays}(G_p) \mid \text{MP}(\pi) \geq v\}$ requires that for all dimension j , the mean-payoff on this dimension is at least $v(j)$.
- *Parity objectives.* Objective $\text{Parity}_{G_p} = \{\pi \in \text{Plays}(G_p) \mid \text{Par}(\pi) \bmod 2 = 0\}$ requires that the minimum priority visited infinitely often be even. When the set of priorities is restricted to $\{0, 1\}$, we have a *Büchi objective*. Note that every multi-weighted game structure G without parity can trivially be extended to G_p with $p : S \rightarrow \{0\}$.
- *Combined objectives.* Parity objectives can naturally be combined with multi mean-payoff and multi energy objectives, resp. yielding $\text{MeanPayoff}_{G_p}(v) \cap \text{Parity}_{G_p}$ and $\text{PosEnergy}_{G_p}(v_0) \cap \text{Parity}_{G_p}$.

Sure, satisfaction and expectation semantics. A strategy λ_1 for \mathcal{P}_1 is *surely winning* for an objective ϕ in G_p if for all plays $\pi \in \text{Plays}(G_p)$ that are consistent with λ_1 , we have $\pi \in \phi$. When at least one of the players plays a randomized strategy, the notion of sure winning in general is too restrictive and inadequate, as the set of consistent plays that do not belong to ϕ may have zero probability measure. Therefore, it is useful to use *satisfaction* or *expectation* criteria. Let $\lambda_1 \in \Lambda_1$ be the strategy of \mathcal{P}_1 .

- Given a threshold $\alpha \in [0, 1]$ and a measurable objective $\phi \subseteq \text{Plays}(G_p)$, α -*satisfaction* asks that for all $\lambda_2 \in \Lambda_2$, we have $\mathbb{P}_{s_{init}}^{\lambda_1, \lambda_2}(\phi) \geq \alpha$. If λ_1 satisfies ϕ with probability $\alpha = 1$, we say that λ_1 is *almost-surely winning* for ϕ in G_p .
- Given a threshold $\beta \in \mathbb{Q}^k$, a function $f : \text{Plays}(G_p) \rightarrow \mathbb{Q}$, β -*expectation* asks that for all $\lambda_2 \in \Lambda_2$, we have $\mathbb{E}_{s_{init}}^{\lambda_1, \lambda_2}(f) \geq \beta$.

Note that energy objectives are naturally more enclined towards satisfaction semantics, as they model safety properties.

Strategy synthesis problem. For multi energy parity games, the problem is to synthesize a finite initial credit $v_0 \in \mathbb{N}^k$ and a pure *finite-memory* strategy $\lambda_1^{pf} \in \Lambda_1^{PF}$ that is surely winning for \mathcal{P}_1 in G_p for the objective $\text{PosEnergy}_{G_p}(v_0) \cap \text{Parity}_{G_p}$, *if one exists*. So, the initial credit is not fixed, but is part of the strategy to synthesize. For multi mean-payoff games, given a threshold $v \in \mathbb{Q}^k$, the problem is to synthesize a pure *finite-memory* strategy $\lambda_1^{pf} \in \Lambda_1^{PF}$ that is surely winning for \mathcal{P}_1 in G_p for the objective $\text{MeanPayoff}_{G_p}(v) \cap \text{Parity}_{G_p}$, *if one exists*. Note that multi energy and multi

mean-payoff games are equivalent for finite-memory strategies, while in general, infinite memory may be necessary for the latter [20].

Trading finite memory for randomness. We study when finite memory can be traded for randomization. The question is: given a strategy $\lambda_1^{Pf} \in \Lambda_1^{Pf}$ which ensures surely winning of some objective ϕ , does there exist a strategy $\lambda_1^{RM} \in \Lambda_1^{RM}$ which ensures almost-surely winning for the same objective ϕ ? For mean-payoff objectives, one can also ask for a weaker equivalence, that is: can randomized memoryless strategies achieve the same expectation as pure finite-memory ones?

3 Optimal memory bounds

In this section, we establish optimal memory bounds for pure finite-memory winning strategies on multi-dimensional energy parity games (MEPGs). Also, as a corollary, we obtain results for pure finite-memory winning strategies on multi-dimensional mean-payoff parity games (MMPPGs). We show that single exponential memory is both sufficient and necessary for winning strategies. Additionally, we show how the parity condition in a MEPG can be removed by adding additional energy dimensions.

Multi energy parity games. A sample game is depicted on Fig. 1. The key point in the upper bound proof on memory is to understand that for \mathcal{P}_1 to win a multi energy parity game, he must be able to force cycles whose energy level is positive in all dimensions and whose minimal parity is even. As stated in the next lemma, finite-memory strategies are sufficient for multi energy parity games for both players.

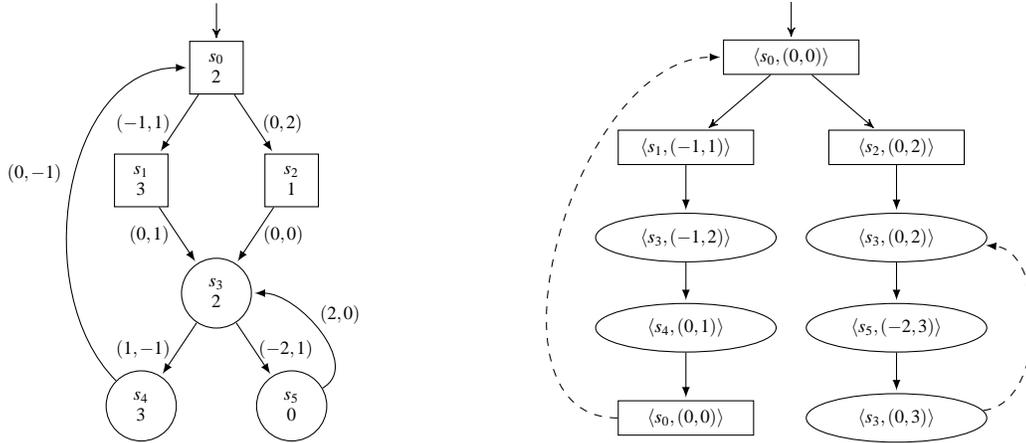


Fig. 1. Two-dimensional energy parity game and even-parity self-covering tree representing an arbitrary finite-memory winning strategy. Circle states belong to \mathcal{P}_1 , square states to \mathcal{P}_2 .

Lemma 1 (Extension of [20, Lemma 2 and 3]). *If \mathcal{P}_1 has a winning strategy in a multi energy parity game, then he has a pure finite-memory one. If \mathcal{P}_2 has a winning strategy in a multi energy parity game, then he has a pure memoryless one.*

Proof. The first part of the result follows using the standard well-quasi ordering argument (straightforward extension of [20, Lemma 2]). The second part follows by the classical edge induction argument: Lemma 3 of [20] and Lemma 3 of [17] show the result using edge induction for multi energy and energy parity games, respectively. Repeating the arguments of Lemma 3 of [17], and replacing the part on single energy objectives by the argument of Lemma 3 of [20] for multi energy objectives, we obtain the desired result. \square

By Lemma 1, we know that w.l.o.g. both players can be restricted to play pure finite-memory strategies. The property on the cycles can then be formalized as follows.

Lemma 2. Let $G_p = (S_1, S_2, s_{\text{init}}, E, k, w, p)$ be a multi energy parity game. Let $\lambda_1^{pf} \in \Lambda_1^{PF}$ be a winning strategy of \mathcal{P}_1 for initial credit $v_0 \in \mathbb{N}^k$. Then, for all $\lambda_2^{pm} \in \Lambda_2^{PM}$, the outcome is a regular play $\pi = \rho \cdot (\eta_\infty)^\omega$, with $\rho \in \text{Prefs}(G), \eta_\infty \in S^+$, such that $\text{EL}(\eta_\infty) \geq 0$ and $\text{Par}(\pi) = \min \{p(s) \mid s \in \eta_\infty\}$ is even.

Proof. Recall that both players play with pure finite memory strategies. Therefore, a finite number of decisions are made and the outcome is a regular play $\pi = \rho \cdot (\eta_\infty)^\omega$. Note that $\text{EL}(\rho)$ does not have to be positive, as \mathcal{P}_1 may have $v_0 > \text{EL}(\rho)$. Similarly, priorities of states visited in ρ have no impact on winning as they are only visited a finite number of times. First, suppose $\text{EL}(\eta_\infty) < 0$ on some dimension $1 \leq j \leq k$. Then, after $m > 0$ cycles, for some $n > 0$, the energy level will be $\text{EL}(\pi(n)) = \text{EL}(\rho \cdot (\eta_\infty)^m) = \text{EL}(\rho) + m \cdot \text{EL}(\eta_\infty)$. Since v_0 is finite and $m \rightarrow \infty$, there exist some $m, n > 0$, such that $v_0 + \text{EL}(\pi(n)) < 0$ on dimension j and λ_1 is not winning. Second, suppose $\min \{p(s) \mid s \in \eta_\infty\}$ is odd. Since the set of states visited infinitely often is exactly the set of states in η_∞ , this implies that $\text{Par}(\pi)$ is odd, and thus λ_1 is not winning. \square

A *self-covering path* in a game, straightforwardly extending the notion introduced by Rackoff [41] for *Vector Addition Systems* (VAS), is a sequence of states $s_0 s_1 s_2 \dots s_m$ such that there exist two positions i and j that verify $0 \leq i < j \leq m, s_i = s_j$ and $\text{EL}(s_0 \dots s_i) \leq \text{EL}(s_0 \dots s_i \dots s_j)$. In other words, such a path describes a finite prefix followed by a cycle which has a non-negative effect on the energy level. Ensuring such cycles is crucial to win the energy objective. With the notion of regular play of Lemma 2, we generalize the notion of self-covering path to include the parity condition. We show here that, if such a path exists, then the lengths of its cycle and the prefix needed to reach it can be bounded. Bounds on the strategy follow. In [41], Rackoff showed how to bound the length of self-covering paths in VAS. This work was extended to Vector Addition Systems with States (VASS) by Rosier and Yen [43]. Recently, Brázdil *et al.* introduced reachability games on VASS and the notion of *self-covering trees* [12]. Their Zero-safety problem with ω initial marking is equivalent to multi energy games with weights in $\{-1, 0, 1\}$, and without the parity condition. They showed that if winning strategies exist for \mathcal{P}_1 , then some of them can be represented as *self-covering trees* of bounded depth. Trees have to be considered instead of paths, as in a game setting all the possible choices of the adversary (\mathcal{P}_2) must be considered. Here, we extend the notion of self-covering trees to *even-parity self-covering trees*, in order to handle parity objectives.

Even-parity self-covering tree. An *even-parity self-covering tree* (epSCT) for $s \in S$ is a finite tree $T = (Q, R)$, where Q is the set of nodes, $\Theta: Q \rightarrow S \times \mathbb{Z}^k$ is a labeling function and $R \subset Q \times Q$ is the set of edges, such that

- The root of T is labeled $\langle s, (0, \dots, 0) \rangle$.
- If $\zeta \in Q$ is not a leaf, then let $\Theta(\zeta) = \langle t, u \rangle, t \in S, u \in \mathbb{Z}^k$, such that
 - if $t \in S_1$, then ζ has a unique child ϑ such that $\Theta(\vartheta) = \langle t', u' \rangle, (t, t') \in E$ and $u' = u + w(t, t')$;
 - if $t \in S_2$, then there is a bijection between children of ζ and edges of the game leaving t , such that for each successor $t' \in S$ of t in the game, there is one child ϑ of ζ such that $\Theta(\vartheta) = \langle t', u' \rangle, u' = u + w(t, t')$.
- If ζ is a leaf, then let $\Theta(\zeta) = \langle t, u \rangle$ such that there is some ancestor ϑ of ζ in T such that $\Theta(\vartheta) = \langle t, u' \rangle$, with $u' \leq u$, and the downward path from ϑ to ζ , denoted by $\vartheta \rightsquigarrow \zeta$, has minimal priority even. We say that ϑ is an *even-descendance energy ancestor* of ζ .

Intuitively, each path from root to leaf is a self-covering path of even parity in the game graph so that plays unfolding according to such a tree correspond to winning plays of Lemma 2. Thus, the epSCT fixes how \mathcal{P}_1 should react to actions of \mathcal{P}_2 in order to win the MEPG (Fig. 1). Note that as the tree is finite, one can take the largest negative number that appears on a node in each dimension to compute an initial credit for which there is a winning strategy (i.e., the one described by the tree). In particular, let W denote the maximal absolute weight appearing on an edge in G_p . Then, for an epSCT T of depth l , it is straightforward to see that the maximal initial credit required is at most $l \cdot W$ as the maximal decrease at each level of the tree is bounded by W . We suppose $W > 0$ as otherwise, any strategy of \mathcal{P}_1 is winning for the energy objective, for any initial credit vector $v_0 \in \mathbb{N}^k$.

Let us explicitly state how \mathcal{P}_1 can deploy a strategy $\lambda_1^T \in \Lambda_1^{PF}$ based on an epSCT $T = (Q, R)$. We refer to such a strategy as an *epSCT strategy*. It consists in following a path in the tree T , moving a pebble from node to node and playing in the game depending on edges taken by this pebble. Each time a node ζ such that $\Theta(\zeta) = \langle t, u \rangle$ is encountered, we do the following.

- If ζ is a leaf, the pebble directly goes up to its oldest even-descendance energy ancestor ϑ . By oldest we mean the first encountered when going down in the tree from the root. Note that this choice is arbitrary, in an effort to ease following proof formulations, as any one would suit.
- Otherwise, if ζ is not a leaf,
 - if $t \in S_2$ and \mathcal{P}_2 plays state $t' \in S$, the pebble is moved along the edge going to the only child ϑ of ζ such that $\Theta(\vartheta) = \langle t', u' \rangle$, $u' = u + w(t, t')$;
 - if $t \in S_1$, the pebble moves to ϑ , $\Theta(\vartheta) = \langle t', u' \rangle$, the only child of ζ , and \mathcal{P}_1 strategy is to choose the state t' in the game.

If such an epSCT T of depth l exists for a game G_p , then \mathcal{P}_1 can play the strategy $\lambda_1^T \in \Lambda_1^{PF}$ to win the game with initial credit bounded by $l \cdot W$.

Bounding the depth of epSCTs. Consider a multi energy game *without* parity. Then, the priority condition on downward paths from ancestor to leaf is not needed and self-covering trees (i.e., epSCTs without the condition on priorities) suffice to describe winning strategies. One can bound the size of SCTs using results on the size of solutions for linear diophantine equations (i.e., with integer variables) [9]. In particular, recent work on reachability games over VASS with weights $\{-1, 0, 1\}$, Lemma 7 of [12], states that if \mathcal{P}_1 has a winning strategy on a VASS, then he can exhibit one that can be described as an SCT whose *depth* is at most $l = 2^{(d-1) \cdot |S|} \cdot (|S| + 1)^{c \cdot k^2}$, where c is a constant independent of the considered VASS and d its branching degree (i.e., the highest number of outgoing edges on any state). Naive use of this bound for multi energy games with arbitrary integer weights would induce a *triple* exponential bound for memory. Indeed, recall that W denotes the maximal absolute weight that appears in a game $G_p = (S_1, S_2, s_{init}, E, k, w, p)$. A straightforward translation of a game with arbitrary weights into an equivalent game that uses only weights in $\{-1, 0, 1\}$ induces a blow-up by W in the size of the state space, and thus an exponential blow-up by W in the depth of the tree, which becomes doubly exponential as we have

$$l = 2^{(d-1) \cdot W \cdot |S|} \cdot (W \cdot |S| + 1)^{c \cdot k^2} = 2^{(d-1) \cdot 2^V \cdot |S|} \cdot (W \cdot |S| + 1)^{c \cdot k^2},$$

where V denotes the number of bits used by the encoding of W . Moreover, the width of the tree increases as d^l , i.e., it increases exponentially with the depth. So straight application of previous results provides an overall tree of triple exponential size. In this paper we improve this bound and prove a single exponential upper bound, even for multi energy *parity* games. We proceed in two steps, first studying the depth of the epSCT, and then showing how to compress the tree into a *directed acyclic graph* (DAG) of *single* exponential size.

Lemma 3. *Let $G_p = (S_1, S_2, s_{init}, E, k, w, p)$ be a multi energy parity game such that W is the maximal absolute weight appearing on an edge and d the branching degree of G_p . Suppose there exists a finite-memory winning strategy for \mathcal{P}_1 . Then there is an even-parity self-covering tree for s_{init} of depth at most $l = 2^{(d-1) \cdot |S|} \cdot (W \cdot |S| + 1)^{c \cdot k^2}$, where c is a constant independent of G_p .*

Lemma 3 eliminates the exponential blow-up in depth induced by a naive coding of arbitrary weights into $\{-1, 0, 1\}$ weights, and implies an overall doubly exponential upper bound. Our proof is a generalization of [12, Lemma 7], using a more refined analysis to handle both *parity* and *arbitrary integer weights*. The idea is the following. First, consider the one-player case. The epSCT is reduced to a path. By Lemma 2, it is composed of a finite prefix, followed by an infinitely repeated sequence of positive energy level and even minimal priority. The point is to bound the length of such a sequence by eliminating cycles that are not needed for energy or parity. Second, to extend the result to two-player games, we use an induction on the number of choices available for \mathcal{P}_2 in a given state. Intuitively, we show that if \mathcal{P}_1 can win with an epSCT T_A when \mathcal{P}_2 plays edges from a set A in a state s , and if he can also win with an epSCT T_B when \mathcal{P}_2 plays edges from a set B , then he can win when \mathcal{P}_2 chooses edges from both A and B , with an epSCT whose depth is bounded by the sum of depths of T_A and T_B .

Proof. The proof is made in two steps. First, we consider the one-player case, where $S_2 = \emptyset$. Second, we use an induction scheme over the choice degree of \mathcal{P}_2 to extend our results to the two-player case.

We start with $S_2 = \emptyset$, the one-player game. By Lemma 2, a winning play is of the form $\pi = \rho \cdot (\eta_\infty)^\omega$ such that $\text{EL}(\eta_\infty) \geq 0$ and $\text{Par}(\pi) = \min\{p(s) \mid s \in \eta_\infty\}$ is even. Notice that such a play corresponds to the epSCT defined above, as it reduces to an even-parity self-covering path $\langle s_{\text{init}}, (0, \dots, 0) \rangle \rightsquigarrow \langle s, u \rangle \rightsquigarrow \langle s, u' \rangle$ with $u' \geq u$. Therefore its existence is guaranteed and it remains to bound its length. Given such a path, the idea is to eliminate unnecessary cycles, in order to reduce its length while maintaining the needed properties (i.e., positive energy and even minimal priority). First, notice that cycles in the sub-path $\langle s_{\text{init}}, (0, \dots, 0) \rangle \rightsquigarrow \langle s, u \rangle$ can be trivially erased as they are only visited a finite number of times and thus (a) the initial credit can compensate for the loss of their potential positive energy effect, and (b) they do not contribute in the parity. Now consider the sub-path $\langle s, u \rangle \rightsquigarrow \langle s, u' \rangle$. Since it induces a winning play, its minimal priority is even. Let p_m be this priority. We may suppose w.l.o.g. that $p(s) = p_m$, otherwise it suffices to shift this sub-path to $\langle s', v \rangle \rightsquigarrow \langle s', v' \rangle$ for some state s' such that $p(s') = p_m$ and $v' \geq v$, and add the sub-path $\langle s, u \rangle \rightsquigarrow \langle s', v \rangle$ to the finite prefix. Now we may eliminate each cycle of $\langle s, u \rangle \rightsquigarrow \langle s, u' \rangle$ safely in regards to the parity objective as they only contain states with greater or equal priority. Thus, we only need to take care of the energy, and fall under the scope of [12, Lemma 15] for the special case of weights in $\{-1, 0, 1\}$, where an upper bound $h(|S|, k) = (|S| + 1)^{c \cdot k^2}$ on the length of such a path is shown.

We claim that for a one-player game G , with weights in $\{-W, -W + 1, \dots, W - 1, W\}$, an upper bound $h(W, |S|, k) = (W \cdot |S| + 1)^{c \cdot k^2}$ is obtained. Indeed, one can translate $G_p = (S_1, S_2, s_{\text{init}}, E, k, w, p)$ into an equivalent game $G'_{p'} = (S'_1, S_2, s_{\text{init}}, E', k, w', p')$ such that each edge of G_p is split into at most W edges in $G'_{p'}$, with at most $(W - 1)$ dummy states in between, so that each edge of $G'_{p'}$ only uses weights in $\{-1, 0, 1\}$. Let S_d denote the set of these added dummy states. We define this translation $\text{Tr}: G_p \mapsto G'_{p'}$ with $\text{Tr}(S_1) = S_1 \cup S_d$, $\text{Tr}(S_2) = S_2$, $\text{Tr}(s_{\text{init}}) = s_{\text{init}}$, $\text{Tr}(E) = \bigcup_{(s,t) \in E} \text{Tr}((s,t))$, $\text{Tr}(k) = k$, $\text{Tr}(w) = w': E' \rightarrow \{-1, 0, 1\}^k$, $\text{Tr}(p) = p': S' \rightarrow \mathbb{N}$ such that for all $(s,t) \in E$ such that $m = \max\{w(s,t)(j) \mid 1 \leq j \leq k\} - 1$, we have that $\text{Tr}((s,t)) = \{(s, s_d^1), (s_d^1, s_d^2), \dots, (s_d^{m-1}, s_d^m), (s_d^m, t)\}$ such that

$$\left(\forall j > 0, s_d^j \in S_d \wedge p'(s_d^j) = p(s) \right) \wedge \sum_{(q,r) \in \text{Tr}((s,t))} w'(q,r) = w(s,t).$$

To be formally correct, we have to add that for all $s_d \in S_d$, we have $\text{degree}_{\text{in}}(s_d) = \text{degree}_{\text{out}}(s_d) = 1$, and for all $s \notin S_d$, we have $p'(s) = p(s)$. This translation does not hinder the outcome of the game as each edge in G_p has a unique corresponding path in $G'_{p'}$ that preserves the weights and the visited priorities, and that offers no added choice to \mathcal{P}_1 . Since G_p possesses $|E| \leq |S|^2$ edges, and for each edge of G_p , we add at most $(W - 1)$ dummy states in $G'_{p'}$, we have $|S'| \leq |S| + |S|^2 \cdot (W - 1) \leq |S|^2 \cdot W$. Therefore, by applying [12, Lemma 15] on $G'_{p'}$, we obtain the following upper bound:

$$h(W, |S|, k) = h(|S'|, k) \leq (|S|^2 \cdot W + 1)^{c \cdot k^2} \leq (W \cdot |S| + 1)^{c' \cdot k^2}$$

for some constant c' that is independent of G_p .

Now, consider $S_2 \neq \emptyset$. (I) We extend [12, Lemma 16] for parity. This will help us to establish an induction scheme over the choice degree of \mathcal{P}_2 . Suppose $s \in S_2$ has more than one outgoing edge. Let $\tau = (s, t) \in E$ be one of them and $R \subset E$ denote the nonempty set of other outgoing edges. Let G_p^τ (resp. G_p^R) be the game induced when removing R (resp. τ) from G_p . Suppose that (a) s is winning for \mathcal{P}_1 in G_p^τ for initial credit $v_R \in \mathbb{N}^k$, and (b) there exists some state $s' \in S$ such that s' is winning for \mathcal{P}_1 in G_p^τ for initial credit $v_\tau \in \mathbb{N}^k$. We claim that s' is winning in G_p for initial credit $v_0 = v_\tau + v_R$. Indeed, let λ_1^τ and λ_1^R resp. denote winning strategies for \mathcal{P}_1 in G_p^τ and G_p^R . Let \mathcal{P}_1 use the following strategy. Player \mathcal{P}_1 plays λ_1^τ as long as \mathcal{P}_2 does not play any edge of R . If such an edge is played, then \mathcal{P}_1 switches to strategy λ_1^R and plays it until edge τ is played again by \mathcal{P}_2 , in which case \mathcal{P}_1 switches back to λ_1^τ , and so on. In this way, the outcome of the game is guaranteed to be a play $\pi = s' \dots s \dots s \dots$ resulting from a merge between a play consistent with λ_1^τ over G_p^τ (whose energy level is bounded by $-v_\tau$ at all times), and a play consistent with λ_1^R over G_p^R (whose energy level is bounded by $-v_R$ at all times). Therefore, the combined overall energy level of any prefix ρ of this play is bounded by $(-v_\tau - v_R)$ as positive cycles in G_p^τ and G_p^R do remain positive in G_p . Furthermore, the parity condition is preserved in G_p . Indeed, suppose it is not. Thus, there exists a state visited infinitely often in the outcome such that its priority is minimal and odd. However, as the outcome results from merging plays resp. consistent with λ_1^τ and λ_1^R , this implies that one of those strategies yields an odd minimal priority, which contradicts the fact that they are winning. This proves the claim.

(II) We apply the induction scheme of [12, Lemma 18] on $r = |\{(s, t) \in E \mid s \in S_2\}| - |S_2| \leq (d-1) \cdot |S|$, the choice degree of \mathcal{P}_2 . Notice that our translation $\text{Tr}: G_p \mapsto G'_p$ maintains this choice degree unchanged. The claim is that for a winning state s' , there is an epSCT of depth bounded by $2^r \cdot h(W, |S|, k)$. We have proved that for the base case $r = 0$, similar to $S_2 = \emptyset$, this claim is true. So assume it holds for r , it remains to prove that it is preserved for $r + 1$. Let $s \in S_2$ be such that \mathcal{P}_2 has at least two outgoing edges. As before, we define G_p^τ and G_p^R . Clearly, the choice degree of \mathcal{P}_2 is at most r in both games. Let s' be a winning state in G_p . As \mathcal{P}_2 has less choices in both G_p^τ and G_p^R , clearly s' is still winning in those games. If an epSCT in either of them (which are guaranteed to exist and have depth bounded by $2^r \cdot h(W, |S|, k)$ by hypothesis) do not contain the state s , then the claim is verified. Now suppose we have two epSCTs for games G_p^τ and G_p^R such that they both contain state s . Notice that s is winning in those two games and as such, is the root of two respective epSCTs of depth less than $2^r \cdot h(W, |S|, k)$. Applying (I) on states s' and s , we get an epSCT for s' in G_p of depth $2 \cdot 2^r \cdot h(W, |S|, k)$, which concludes the proof. \square

From multi energy parity games to multi energy games. Let G_p be a MEPG and assume that \mathcal{P}_1 has a winning strategy in that game. By Lemma 3, there exists an epSCT whose depth is bounded by l . As a direct consequence of that bounded depth, we have that \mathcal{P}_1 , by playing the strategy prescribed by the epSCT, enforces a stronger objective than the parity objective. Namely, this strategy ensures to “never visit more than l states of odd priorities before seeing a smaller even priority” (which is a safety objective). Then, the parity condition can be transformed into additional energy dimensions.

While our transformation shares ideas with the classical transformation of parity objectives into safety objectives, first proposed in [5] (see also [29, Lemma 6.4]), it is technically different because energy levels cannot be reset (as it would be required by those classical constructions). The reduction is as follows. For each odd priority, we add one dimension. The energy level in this dimension is decreased by 1 each time this odd priority is visited, and it is increased by l each time a smaller even priority is visited. If \mathcal{P}_1 is able to maintain the energy level positive for all dimensions (for a given initial energy level), then he is clearly winning the original parity objective; on the other hand, an epSCT strategy that wins the original objective also wins the new game.

Lemma 4. *Let $G_p = (S_1, S_2, s_{\text{init}}, E, k, w, p)$ be a multi energy parity game with priorities in $\{0, 1, \dots, 2 \cdot m\}$, such that W is the maximal absolute weight appearing on an edge. Then we can construct a multi energy game G with the same set of states, $(k + m)$ dimensions and a maximal absolute weight bounded by l , as defined by Lemma 3, such that \mathcal{P}_1 has a winning strategy in G iff he has one in G_p .*

Proof. Let $G_p = (S_1, S_2, s_{\text{init}}, E, k, w, p)$ be a MEPG with priorities in $\{0, 1, \dots, 2 \cdot m\}$. Let $G = (S_1, S_2, E, (k + m), w')$ be the multi energy game (MEG) obtained by applying the following transformation: $\forall (s, t) \in E, \forall 1 \leq j \leq k, w'((s, t))(j) = w((s, t))(j)$, and (a) if $p(t)$ is even, $\forall k < j \leq k + \frac{p(t)}{2}, w'((s, t))(j) = 0$ and $\forall k + \frac{p(t)}{2} < j \leq k + m, w'((s, t))(j) = l$, or (b) if $p(t)$ is odd, $\forall k < j \leq k + m, j \neq k + \lceil \frac{p(t)}{2} \rceil, w'((s, t))(j) = 0$ and $w'((s, t))(k + \lceil \frac{p(t)}{2} \rceil) = -1$. We have to prove both ways of the equivalence.

First, suppose $\lambda_1 \in \Lambda_1^{PF}$ is a winning strategy for \mathcal{P}_1 in the MEPG G_p . By Lemma 3, there is an epSCT of depth at most l for s_{init} . Thus, we know that in every repeated sequence of l states, the minimal visited priority will be even. Therefore, for all additional dimensions, ranging from $k + 1$ to $k + m$, the effect of a sequence of l states will be bounded from below by $-1 \cdot (l - 1) + l$, which is positive. Thus strategy λ_1 is also winning in G (with initial credit bounded by l on additional dimensions).

Second, suppose $\lambda_1 \in \Lambda_1^{PF}$ is a winning strategy for \mathcal{P}_1 in the MEG G , as defined above. Since λ_1 is winning, it yields an SCT (epSCT without the parity condition) of bounded depth such that \mathcal{P}_1 is able to enforce positive energy cycles. By definition of weights over G , this cannot be the case if the minimal priority infinitely often visited is odd. Thus this strategy is winning for parity on G_p , and stays winning for energy over dimensions 1 to k as weights are unchanged. \square

Bounding the width. Thanks to Lemma 4, we continue with multi energy games without parity. In order to bound the overall size of memory for winning strategies, we consider the width of self-covering trees. The following lemma states that SCTs, whose width is at most doubly exponential by application of Lemma 3, can be compressed into *directed acyclic graphs* (DAGs) of single exponential width. Thus we eliminate the second exponential blow-up and give an overall single exponential bound for memory of winning strategies.

Lemma 5. Let $G = (S_1, S_2, s_{\text{init}}, E, k, w)$ be a multi energy game such that W is the maximal absolute weight appearing on an edge and d the branching degree of G . Suppose there exists a finite-memory winning strategy for \mathcal{P}_1 . Then, there exists $\lambda_1^D \in \Lambda_1^{PF}$ a winning strategy for \mathcal{P}_1 described by a DAG D of depth at most $l = 2^{(d-1) \cdot |S|} \cdot (W \cdot |S| + 1)^{c \cdot k^2}$ and width at most $L = |S| \cdot (2 \cdot l \cdot W + 1)^k$, where c is a constant independent of G . Thus the overall memory needed to win this game is bounded by the single exponential $l \cdot L$.

The sketch of this proof is the following. By Lemma 3, we know that there exists a tree T , and thus a DAG, that satisfies the bound on depth. We construct a finite sequence of DAGs, whose first element is T , so that (1) each DAG describes a winning strategy for the same initial credit, (2) each DAG has the same depth, and (3) the last DAG of the sequence has its width bounded by $|S| \cdot (2 \cdot l \cdot W + 1)^k$. This sequence $D_0 = T, D_1, D_2, \dots, D_n$ is built by merging nodes on the same level of the initial tree depending on their labels, level by level. The key idea of this procedure is that what actually matters for \mathcal{P}_1 is only the current energy level, which is encoded in node labels in the self-covering tree T . Therefore, we merge nodes with identical states and energy levels: since \mathcal{P}_1 can essentially play the same strategy in both nodes, we only keep one of their subtrees.

It is possible to further reduce the practical size of the compressed resulting DAG by merging nodes according to a “greater or equal” relation over energy levels rather than simply equality (Fig. 2). This improvement is part of the algorithm that follows, and it has a significant impact on the practical width of DAGs as it can then be bounded by the number of *incomparable* labeling vectors instead of *unequivalent* ones.

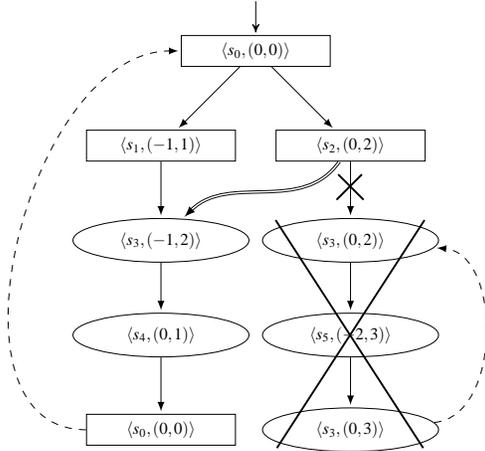


Fig. 2. Merge between comparable nodes.

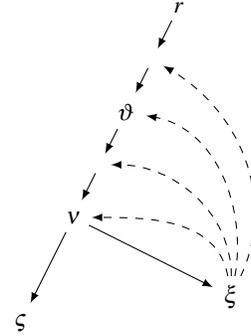


Fig. 3. Cycles have positive energy levels.

The remainder of this subsection is dedicated to the proof of Lemma 5. We need to introduce some notations and two intermediate lemmas. If he so wishes, the reader may directly proceed to the next subsection and Lemma 8 for results on lower memory bounds.

We first introduce some notations. Let $T = (Q, R)$ be a self-covering tree (i.e., epSCT without the parity condition). We define the partial order \preceq on Q such that for all $\zeta_1, \zeta_2 \in Q$ such that $\Theta(\zeta_1) = \langle t_1, u_1 \rangle$ and $\Theta(\zeta_2) = \langle t_2, u_2 \rangle$, we have $\zeta_1 \preceq \zeta_2$ iff $t_1 = t_2$ and $u_1 \leq u_2$. We denote the equivalence by \simeq such that $\zeta_1 \simeq \zeta_2$ iff $\zeta_1 \preceq \zeta_2$ and $\zeta_2 \preceq \zeta_1$. For all $\zeta \in Q$, let Anc and EnAnc resp. denote the set of *ancestors* and *energy ancestors* of ζ in T : $\text{Anc}(\zeta) = \{\vartheta \in Q \setminus \{\zeta\} \mid \vartheta \models \exists \diamond \zeta\}$, where we use the classical CTL notation to denote that there exists a path from ϑ to ζ in T , and $\text{EnAnc}(\zeta) = \{\vartheta \in \text{Anc}(\zeta) \mid \vartheta \preceq \zeta\}$.

We build a sequence of DAGs $(D_i)_{0 \leq i \leq n} \equiv D_0 = T, D_1, D_2, \dots, D_n$ such that for all $0 < i \leq n$, D_i is obtained from D_{i-1} by *merging* two equivalent nodes of the same minimal level (i.e., closest to the root) of D_{i-1} . The sequence stops when we obtain a DAG $D_n = (Q_n, R_n)$ such that for all level j of D_n , there does not exist two distinct equivalent nodes on level j . This construction induces merges by increasing depth, starting with level one. Moreover, if a DAG D_i of the sequence is the result on merges up to level j , then it has the tree property (i.e., every node has a unique father)

for levels greater than j . As the depth and the branching degree of T are finite, the defined sequence of DAGs is finite (and actually bounded).

Let us give a formal definition of the *merge* operation. Consider such a DAG $D_i = (Q_i, R_i)$. Let j the minimal level of D_i that contains two equivalent nodes. Let $\zeta_1, \zeta_2 \in Q_i(j)$ (i.e., nodes of level j) be two nodes such that $\zeta_1 \neq \zeta_2$ and $\zeta_1 \simeq \zeta_2$. We suppose w.l.o.g. an arbitrary order on nodes of the same level so that ζ_1, ζ_2 are the two leftmost nodes that satisfy this condition. We define $D_{i+1} = (Q_{i+1}, R_{i+1}) = \text{merge}(D_i)$ as the result of the following transformation:

- $Q_{i+1} = Q_i \setminus \{\zeta_2\} \cup \{\zeta_d \in Q_i \mid \zeta_2 \in \text{Anc}(\zeta_d)\}$,
- $R_{i+1} = (R_i \cap (Q_{i+1} \times Q_{i+1})) \cup \{(\vartheta, \zeta_1) \mid (\vartheta, \zeta_2) \in R_i\}$.

Thus, we eliminate the subtree starting in ζ_2 and replace all edges that point to ζ_2 by edges pointing to ζ_1 . This follows the idea that the same strategy can be played in ζ_2 as in ζ_1 since the present state and the energy level are the same.

Let $D_i = (Q_i, R_i)$ be a DAG of the sequence $(D_i)_{0 \leq i \leq n}$. Given $\zeta \in Q_i$, $\vartheta \in \text{Anc}(\zeta)$, we denote by $\vartheta \rightsquigarrow \zeta$ an arbitrary downward path from ϑ to ζ in D_i . Given a leaf $\zeta \in Q_i$, we denote its oldest energy ancestor by $\text{oea}(\zeta)$. Recall that a strategy is described by such a DAG according to moves of a pebble. Given a leaf $\zeta \in Q_i$ and one of its energy ancestors $\vartheta \in \text{EnAnc}(\zeta)$, we represent the pebble going up from ζ to ϑ by $\zeta \circ \vartheta$. Given $\alpha, \beta \in (Q_i)^*$, $\alpha \circ \beta$ naturally extends this notation such that we have $\text{Last}(\alpha) \circ \text{First}(\beta)$. We consider energy levels of paths in the tree by referring to their counterparts in the game. Note that given $\vartheta, \zeta \in Q_i$, $\Theta(\vartheta) = \langle t, u \rangle$, $\Theta(\zeta) = \langle t', u' \rangle$, we have $\text{EL}(\vartheta \rightsquigarrow \zeta) = u' - u$. We start with two useful lemmas.

Lemma 6. *Let $D_i = (Q_i, R_i)$ be a DAG of $(D_i)_{0 \leq i \leq n}$. For all nodes $\zeta_1, \zeta_2 \in Q_i$ such that $\zeta_1 \simeq \zeta_2$, we have that $\forall \vartheta \in \text{Anc}(\zeta_1) \cap \text{Anc}(\zeta_2)$, $\text{EL}(\vartheta \rightsquigarrow \zeta_1) = \text{EL}(\vartheta \rightsquigarrow \zeta_2)$.*

Proof. The proof is straightforward. □

Lemma 7. *Let $D_i = (Q_i, R_i)$ be a DAG of $(D_i)_{0 \leq i \leq n}$. Let $\zeta, \vartheta, v, \xi \in Q_i$ be four nodes such that ζ and ξ are leaves, v is the deepest common ancestor of ζ and ξ , and ϑ is an ancestor of v . Let the oldest energy ancestor of ξ be an ancestor of ζ , i.e., $\text{oea}(\xi) \in \text{Anc}(\zeta)$. We have that $\text{EL}(\vartheta \rightsquigarrow \zeta) \leq \text{EL}(\vartheta \rightsquigarrow v \rightsquigarrow \xi \circ \text{oea}(\xi) \rightsquigarrow \zeta)$.*

This lemma states that we can extract pebble cycles, which have positive energy levels, from a given path, in order to obtain some canonical path whose energy level is lower or equal (Fig. 3).

Proof. Let $\chi = \text{oea}(\xi)$ and $\rho = \vartheta \rightsquigarrow v \rightsquigarrow \xi \circ \chi \rightsquigarrow \zeta$. Since $\chi \in \text{Anc}(\zeta) \cap \text{Anc}(\xi)$, we have $\chi \in \text{Anc}(v) \cup \{v\}$. Therefore, and applying Lemma 6, four cases are possible: $\chi \in \text{Anc}(\vartheta)$, $\chi = \vartheta$, $\chi \in \text{Anc}(v) \setminus (\text{Anc}(\vartheta) \cup \{\vartheta\})$, and $\chi = v$. Consider the first case, $\chi \in \text{Anc}(\vartheta)$. Then $\rho = \vartheta \rightsquigarrow v \rightsquigarrow \xi \circ \chi \rightsquigarrow \vartheta \rightsquigarrow v \rightsquigarrow \zeta$. We have $\text{EL}(\rho) = \text{EL}(\vartheta \rightsquigarrow v) + \text{EL}(v \rightsquigarrow \xi) + \text{EL}(\chi \rightsquigarrow \vartheta) + \text{EL}(\vartheta \rightsquigarrow v) + \text{EL}(v \rightsquigarrow \zeta) = \text{EL}(\chi \rightsquigarrow \vartheta \rightsquigarrow v \rightsquigarrow \xi) + \text{EL}(\vartheta \rightsquigarrow \zeta)$. By definition of $\chi = \text{oea}(\xi)$, the first term is positive. Thus, $\text{EL}(\rho) \geq \text{EL}(\vartheta \rightsquigarrow \zeta)$. Arguments are similar for the other cases. □

We proceed with the proof of Lemma 5.

Proof (Lemma 5). Let $(D_i)_{0 \leq i \leq n}$ be the sequence of DAGs defined above. We claim that **(i)** each DAG describes a winning strategy for the same initial credit, **(ii)** each DAG has the same depth l , and **(iii)** the last DAG of the sequence has its width bounded by $|S| \cdot (2 \cdot l \cdot W + 1)^k$.

(i) First, recall that \mathcal{P}_1 can play a strategy $\lambda_1^T \in \Lambda_1^{PF}$ based on edges taken by a pebble on T . Notice that moving the pebble as we previously defined is possible because nodes belonging to \mathcal{P}_1 have only one child, and nodes of \mathcal{P}_2 have childs covering all his choices once, and only once. Fortunately, the merge operation maintains this property. Therefore, it is straightforward to see that \mathcal{P}_1 can also play a strategy $\lambda_1^{D_i} \in \Lambda_1^{PF}$ for a DAG D_i resulting of some merges on T . However, while this would be a valid strategy for \mathcal{P}_1 , we have to prove that it is still a winning one, for the same initial credit v_0 as λ_1^T . Precisely, we claim that $\forall i \geq 0$, we have that $\lambda_1^{D_i}$ is winning for v_0 .

We show it by induction on D_i . The base case is trivial as $D_0 = T$: the strategy λ_1^T is winning for v_0 by definition. Our induction hypothesis is that our claim is valid for D_{i-1} , and we now prove it for D_i , by contradiction. Let $\zeta_1, \zeta_2 \in Q_{i-1}(j)$ be the merged nodes, for some level j of D_{i-1} . Suppose $\lambda_1^{D_i}$ is not winning for v_0 . Thus there exists a finite path ζ of the pebble in D_i , which corresponds to a strategy $\lambda_2^{D_i} \in \Lambda_2^{PF}$ of \mathcal{P}_2 , such that it achieves a negative value on at least one dimension m , $1 \leq m \leq k$. We have that $(v_0 + \text{EL}(\zeta))(m) < 0$. We aim to find a similar path η in D_{i-1} such that $\text{EL}(\eta) \leq \text{EL}(\zeta)$, thus yielding contradiction, as it would witness that $\lambda_1^{D_{i-1}}$ is not winning for v_0 .

We denote by ζ_m the father of ζ_2 in D_{i-1} . The only edge added by the merge operation is (ζ_m, ζ_1) . Obviously, if ζ does not involve this edge, then we can take $\eta = \zeta$ and immediately obtain contradiction. Thus, we can decompose the witness path

$$\zeta = \alpha(1) \zeta_m \zeta_1 \beta(1) \circ \alpha(2) \zeta_m \zeta_1 \beta(2) \circ \dots \circ \alpha(q) \zeta_m \zeta_1 \xi,$$

for some $q \geq 1$ such that for all $1 \leq p \leq q$, we have that $\alpha(p), \beta(p), \xi \in (\mathcal{Q}_i \cup \{\circ\})^*$ are valid paths of the pebble in D_i (and D_{i-1}); they do not involve edge (ζ_m, ζ_1) , i.e., $\{\zeta_m \zeta_1\} \not\subseteq \alpha(p), \beta(p), \xi$; and $\beta(p) \cap (\text{Anc}_{D_i}(\zeta_m) \setminus \text{Anc}_{D_{i-1}}(\zeta_1)) = \emptyset$, $\text{Last}(\beta(p))$ is a leaf and $\text{oea}(\text{Last}(\beta(p))) \in \text{Anc}_{D_i}(\zeta_m)$.

Intuitively, ζ is split into several parts in regard to q , the number of times it takes the added edge (ζ_m, ζ_1) . Each time, this transition is preceded by some path α . It is then followed by some path β where all visited ancestors of ζ_m were already ancestors of ζ_1 in D_{i-1} (thus, β paths can be kept in η). Finally, after the q -th transition $\zeta_m \zeta_1$ is taken, the path ζ ends with a finite sub-path ξ .

We define the witness path η in D_{i-1} as $\eta = \kappa(1)\beta(1) \circ \kappa(2)\beta(2) \circ \dots \circ \kappa(q)\xi$, with the following transformation of sub-paths $\alpha(p) \zeta_m \zeta_1$:

- $\kappa(1) = r \rightsquigarrow_{D_{i-1}} \zeta_1$,
- $\forall 2 \leq p \leq q, \kappa(p) = \text{oea}(\text{Last}(\beta(p-1))) \rightsquigarrow_{D_{i-1}} \zeta_1$,

where $\rightsquigarrow_{D_{i-1}}$ denotes a valid path in D_{i-1} . Note that given preceding definitions, this indeed constitutes a valid path in D_{i-1} . We have to prove that $\text{EL}(\eta) \leq \text{EL}(\zeta)$. We have

$$\text{EL}(\eta) = \sum_{1 \leq p \leq q} \text{EL}(\kappa(p)) + \sum_{1 \leq p \leq q-1} \text{EL}(\beta(p)) + \text{EL}(\xi),$$

and

$$\text{EL}(\zeta) = \sum_{1 \leq p \leq q} \text{EL}(\alpha(p) \zeta_m \zeta_1) + \sum_{1 \leq p \leq q-1} \text{EL}(\beta(p)) + \text{EL}(\xi).$$

Thus, it remains to show that

$$\sum_{1 \leq p \leq q} \text{EL}(\kappa(p)) \leq \sum_{1 \leq p \leq q} \text{EL}(\alpha(p) \zeta_m \zeta_1).$$

In particular, we claim that for all $1 \leq p \leq q$, we have $\text{EL}(\kappa(p)) \leq \text{EL}(\alpha(p) \zeta_m \zeta_1)$. Indeed, notice that $\kappa(p)$ and $\alpha(p)$ share their starting and ending nodes and that $\alpha(p)$ contains a finite number of pebble cycles. Let ϑ denote the common starting node of both $\kappa(p)$ and $\alpha(p)$. Applying Lemma 7 on $\alpha(p)$, we can eliminate cycles one at a time, without ever increasing the energy level, and obtain a path $\vartheta \rightsquigarrow_{D_i} \zeta_m \zeta_1$ such that $\text{EL}(\vartheta \rightsquigarrow_{D_i} \zeta_m \zeta_1) \leq \text{EL}(\alpha(p) \zeta_m \zeta_1)$. Since $\zeta_1 \simeq \zeta_2$, we have by Lemma 6 that $\text{EL}(\vartheta \rightsquigarrow_{D_i} \zeta_m \zeta_1) = \text{EL}(\vartheta \rightsquigarrow_{D_{i-1}} \zeta_m \zeta_2) = \text{EL}(\vartheta \rightsquigarrow_{D_{i-1}} \zeta_1)$, implying the claim.

Consequently, we obtain $\text{EL}(\eta) \leq \text{EL}(\zeta)$, which witnesses that D_{i-1} was not winning. This contradicts our induction hypothesis and concludes our proof that for all $0 \leq i \leq n$, $\lambda_1^{D_i}$ is winning for v_0 .

(ii) Second, the merge operation only prunes some parts of the tree T , without ever adding any new state, and added edges are on existing successive levels. Therefore, each D_i has noticeably the same depth l .

(iii) Third, the last DAG of the sequence, D_n , is such that for all level j , for all $\zeta_1, \zeta_2 \in \mathcal{Q}_n(j)$, we have $(\zeta_1 \neq \zeta_2) \Rightarrow (\zeta_1 \not\sim \zeta_2)$. Therefore the width of this DAG is bounded by the number of possible non-equivalent nodes. Recall that two nodes are equivalent if they have the same labels, i.e., they represent the same state of the game and are marked with exactly the same energy level vector. Since the maximal change in energy level on an edge is W , and the depth of the DAG is bounded by $l = 2^{(d-1) \cdot |S|} \cdot (W \cdot |S| + 1)^{c \cdot k^2}$ thanks to Lemma 3, we have possible vectors in $\{-l \cdot W, -l \cdot W + 1, \dots, l \cdot W - 1, l \cdot W\}^k$ for each state. Consequently, the width of D_n is bounded by

$$|S| \cdot (2 \cdot l \cdot W + 1)^k = |S| \cdot \left(2^{d \cdot |S|} \cdot (W \cdot |S| + 1)^{c \cdot k^2} \cdot W + 1 \right)^k,$$

which is still single exponential. □

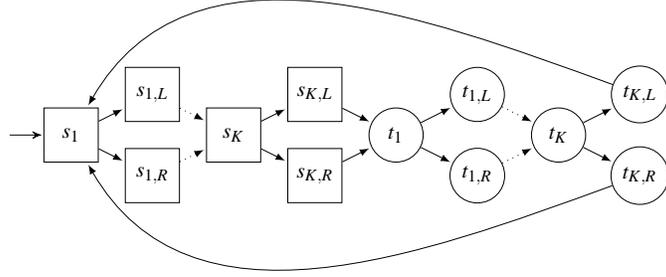


Fig. 4. Family of games requiring exponential memory.

Lower bound. In the next lemma, we show that the upper bound is tight in the sense that there exist families of games which require exponential memory (in the number of dimensions), even for the simpler case of multi energy objectives without parity and weights in $\{-1, 0, 1\}$ (Fig. 4). Note that for one-dimension energy parity, it was shown in [17] that exponential memory (in the encoding of weights) may be necessary.

Lemma 8. *There exists a family of multi energy games $(G(K))_{K \geq 1} = (S_1, S_2, s_{init}, E, k = 2 \cdot K, w : E \rightarrow \{-1, 0, 1\}^k)$ such that for any initial credit, \mathcal{P}_1 needs exponential memory to win.*

The idea is the following: in the example of Fig. 4, if \mathcal{P}_1 does not remember the exact choices of \mathcal{P}_2 (which requires an exponential size Moore machine), there will exist some sequence of choices of \mathcal{P}_2 such that \mathcal{P}_1 cannot counteract a decrease in energy. Thus, by playing this sequence long enough, \mathcal{P}_2 can force \mathcal{P}_1 to lose, whatever his initial credit is.

Proof. We define a family of games $(G(K))_{K \geq 1}$ which is an assembly of $k = 2 \cdot K$ gadgets, the first K belonging to \mathcal{P}_2 , and the remaining K belonging to \mathcal{P}_1 (Fig. 4). Precisely, we have $|S_1| = |S_2| = 3 \cdot K$, $|S| = |E| = 6 \cdot K = 3 \cdot k$ (linear in k), $k = 2 \cdot K$, and w defined as:

$$\begin{aligned} \forall 1 \leq i \leq K, w((\circ, s_i)) &= w((\circ, t_i)) = (0, \dots, 0), \\ w((s_i, s_{i,L})) &= -w((s_i, s_{i,R})) = w((t_i, t_{i,L})) = -w((t_i, t_{i,R})), \\ \forall 1 \leq j \leq k, w((s_i, s_{i,L}))(j) &= \begin{cases} 1 & \text{if } j = 2 \cdot i - 1 \\ -1 & \text{if } j = 2 \cdot i \\ 0 & \text{otherwise} \end{cases}, \end{aligned}$$

where \circ denotes any valid predecessor state.

There exists a winning strategy λ_1^{exp} for \mathcal{P}_1 , for initial credit $v_0^{exp} = (1, \dots, 1)$. Indeed, for any strategy of \mathcal{P}_2 , for any state t_i belonging to \mathcal{P}_1 , it suffices to play the *opposite* choice as \mathcal{P}_2 made on its last visit of s_i to maintain at all times an energy vector which is positive on all dimensions. This strategy thus requires to remember the last choice of \mathcal{P}_2 in all gadgets, which means \mathcal{P}_1 needs K bits to encode these decisions. Thus, this winning strategy is described by a Moore machine containing $2^K = 2^{\frac{k}{2}}$ states, which is exponential in the number of dimensions k .

We claim that, for any initial credit v_0 , there exists no winning strategy λ_1 that can be described with less than 2^K states and prove it by contradiction. Suppose \mathcal{P}_1 plays according to such a strategy λ_1 . Then there exists some $1 \leq x \leq K$ such that $\lambda_1(s_1 \dots s_x s_{x,L} \dots t_x) = \lambda_1(s_1 \dots s_x s_{x,R} \dots t_x)$, i.e., \mathcal{P}_1 chooses the same action in t_x against both choices of the adversary. Suppose that \mathcal{P}_1 chooses to play $t_{x,L}$ in both cases, that is $\lambda_1(s_1 \dots s_x s_{x,L} \dots t_x) = \lambda_1(s_1 \dots s_x s_{x,R} \dots t_x) = t_{x,L}$. By playing $s_{x,L}$, \mathcal{P}_2 can force a decrease of the energy vector by 2 on dimension $2 \cdot x$ every visit in gadget x . Similarly, if the strategy of \mathcal{P}_1 is to play $t_{x,R}$, \mathcal{P}_2 wins by choosing to play $s_{x,R}$ as dimension $2 \cdot x - 1$ decreases by 2 every visit. Therefore, whatever the finite initial vector of \mathcal{P}_1 , \mathcal{P}_2 can enforce a negative dimension by playing long enough. This contradicts the fact that λ_1 is winning and concludes our proof that exponential memory is necessary for this simple family of games $(G(K))_{K \geq 1}$. \square

We summarize our results in Theorem 1.

Theorem 1 (Optimal memory bounds). *The following assertions hold: (1) In multi energy parity games, if there exists a winning strategy, then there exists a finite-memory winning strategy. (2) In multi energy parity and multi mean-payoff games, if there exists a finite-memory winning strategy, then there exists a winning strategy with at most exponential memory. (3) There exists a family of multi energy games (without parity) with weights in $\{-1, 0, 1\}$ where all winning strategies require at least exponential memory.*

Proof. Thanks to [20, Theorem 3], we have equivalence between finite-memory winning for multi energy and multi mean-payoff games. The rest follows from straightforward application of Lemma 1, Lemma 4, Lemma 5, and Lemma 8. \square

4 Symbolic synthesis algorithm

We now present a *symbolic, incremental* and *optimal* algorithm to synthesize a finite-memory winning strategy in a MEG.⁴ This algorithm outputs a (set of) winning initial credit(s) and a derived finite-memory winning strategy (if one exists) which is exponential in the worst-case. Its running time is at most exponential. So our symbolic algorithm can be considered (worst-case) optimal in the light of the results of previous section.

This algorithm computes the greatest fixed point of a monotone operator that defines the sets of winning initial (vectors of) credits for each state of the game. As those sets are upward-closed, they are symbolically represented by their minimal elements. To ensure convergence, the algorithm considers only credits that are below some *threshold*, noted \mathbb{C} . This is without giving up completeness because, as we show below, for a game $G = (S_1, S_2, s_{init}, E, k, w)$, it is sufficient to take the value $2 \cdot l \cdot W$ for \mathbb{C} , where l is the bound on the depth of epSCTs obtained in Lemma 3 and W is the largest absolute value of weights used in the game. We also show how to extract a finite state Moore machine representing a corresponding winning strategy (states of the Moore machine encode the memory of the strategy) from this set of minimal winning initial credits and how to obtain an *incremental* algorithm by increasing values for the threshold \mathbb{C} starting from small values.

A controllable predecessor operator. Let $G = (S_1, S_2, s_{init}, E, k, w)$ be a MEG, $\mathbb{C} \in \mathbb{N}$ be a constant, and $U(\mathbb{C})$ be the set $(S_1 \cup S_2) \times \{0, 1, \dots, \mathbb{C}\}^k$. Let $\mathcal{U}(\mathbb{C}) = 2^{U(\mathbb{C})}$, i.e., the powerset of $U(\mathbb{C})$, and the operator $\text{Cpre}_{\mathbb{C}}: \mathcal{U}(\mathbb{C}) \rightarrow \mathcal{U}(\mathbb{C})$ be defined as follows:

$$\begin{aligned} \mathcal{E}(V) &= \{(s_1, e_1) \in U(\mathbb{C}) \mid s_1 \in S_1 \wedge \exists (s_1, s) \in E, \exists (s, e_2) \in V : e_2 \leq e_1 + w(s_1, s)\}, \\ \mathcal{A}(V) &= \{(s_2, e_2) \in U(\mathbb{C}) \mid s_2 \in S_2 \wedge \forall (s_2, s) \in E, \exists (s, e_1) \in V : e_1 \leq e_2 + w(s_2, s)\}, \\ \text{Cpre}_{\mathbb{C}}(V) &= \mathcal{E}(V) \cup \mathcal{A}(V). \end{aligned} \tag{1}$$

Intuitively, $\text{Cpre}_{\mathbb{C}}(V)$ returns the set of energy levels from which \mathcal{P}_1 can force an energy level in V in one step. The operator $\text{Cpre}_{\mathbb{C}}$ is \subseteq -monotone over the complete lattice $\mathcal{U}(\mathbb{C})$, and so there exists a *greatest fixed point* for $\text{Cpre}_{\mathbb{C}}$ in the lattice $\mathcal{U}(\mathbb{C})$, denoted by $\text{Cpre}_{\mathbb{C}}^*$. As usual, the greatest fixed point of the operator $\text{Cpre}_{\mathbb{C}}$ can be computed by successive approximations as the last element of the following finite \subseteq -descending chain. We define the algorithm CpreFP that computes this greatest fixed point:

$$U_0 = U(\mathbb{C}), U_1 = \text{Cpre}_{\mathbb{C}}(U_0), \dots, U_n = \text{Cpre}_{\mathbb{C}}(U_{n-1}) = U_{n-1}. \tag{2}$$

The set U_i contains all the energy levels that are sufficient to maintain the energy positive in all dimensions for i steps. Note that the length of this chain can be bounded by $|U(\mathbb{C})|$ and the time needed to compute each element of the chain can be bounded by a polynomial in $|U(\mathbb{C})|$. As a consequence, we obtain the following lemma.

Lemma 9. *Let $G = (S_1, S_2, s_{init}, E, k, w)$ be a multi energy game and $\mathbb{C} \in \mathbb{N}$ be a constant. Then $\text{Cpre}_{\mathbb{C}}^*$ can be computed in time bounded by a polynomial in $|U(\mathbb{C})|$, i.e., an exponential in the size of G .*

⁴ Note that the symbolic algorithm can be applied to MPEGs and MMPPGs after removal of the parity condition by applying the construction of Lemma 4.

Symbolic representation. To define a symbolic representation of the sets manipulated by the $\text{Cpre}_{\mathbb{C}}$ operator, we exploit the following partial order: let $(s, e), (s', e') \in U(\mathbb{C})$, we define

$$(s, e) \preceq (s', e') \text{ iff } s = s' \text{ and } e \leq e'. \quad (3)$$

A set $V \in U(\mathbb{C})$ is *closed* if for all $(s, e), (s', e') \in U(\mathbb{C})$, if $(s, e) \in V$ and $(s, e) \preceq (s', e')$, then $(s', e') \in V$. By definition of $\text{Cpre}_{\mathbb{C}}$, we get the following property.

Lemma 10. *All sets U_i in Eq. (2) are closed for \preceq .*

Therefore, all sets U_i in the descending chain of Eq. (2) can be symbolically represented by their minimal elements $\text{Min}_{\preceq}(U_i)$ which is an antichain of elements for \preceq . Even if the largest antichain can be exponential in G , this representation is, in practice, often much more efficient, even for small values of the parameters. For example, with $\mathbb{C} = 4$ and $k = 4$, we have that the cardinality of a set can be as large as $|U_i| \leq 625$ whereas the size of the largest antichain is bounded by $|\text{Min}_{\preceq}(U_i)| \leq 35$. Antichains have proved to be very efficient: see for example [2,27,28]. Therefore, our algorithm is expected to have good performance in practice.

Correctness and completeness. The following two lemmas relate the greatest fixed point $\text{Cpre}_{\mathbb{C}}^*$ and the existence of winning strategies for \mathcal{P}_1 in G . We start with the correctness of the symbolic algorithm.

Lemma 11 (Correctness). *Let $G = (S_1, S_2, s_{\text{init}}, E, k, w)$ be a multi energy game, let $\mathbb{C} \in \mathbb{N}$ be a constant. If there exists $(c_1, \dots, c_k) \in \mathbb{N}^k$ such that $(s_{\text{init}}, (c_1, \dots, c_k)) \in \text{Cpre}_{\mathbb{C}}^*$, then \mathcal{P}_1 has a winning strategy in G for initial credit (c_1, \dots, c_k) and the memory needed by \mathcal{P}_1 can be bounded by $|\text{Min}_{\preceq}(\text{Cpre}_{\mathbb{C}}^*)|$ (the size of the antichain of minimal elements in the fixed point).*

Given the set of winning initial credits output by CpreFP , it is straightforward to derive a corresponding winning strategy of at most exponential size. Indeed, for winning initial credit $\bar{c} \in \mathbb{N}^k$, we build a Moore machine which (i) states are the minimal elements of the fixed point (antichain at most exponential in G), (ii) initial state is any element (t, u) among them such that $t = s_{\text{init}}$ and $u \leq \bar{c}$, (iii) next-action function prescribes an action that ensures remaining in the fixed point, and (iv) update function maintains an accurate energy level in the memory.

Proof. We denote by \bar{c} the k -dimension credit vector (c_1, \dots, c_k) . W.l.o.g. we assume that states of G alternate between positions of \mathcal{P}_1 and positions of \mathcal{P}_2 (otherwise, we split needed edges by introducing dummy states). From $\text{Cpre}_{\mathbb{C}}^*$, we construct a Moore machine $\mathcal{M} = (M, m_0, \alpha_u, \alpha_n)$ which respects the following definitions:

- $M = \text{Min}_{\preceq}\{(t, u) \in S_1 \times \{0 \dots \mathbb{C}\}^k \mid (t, u) \in (\text{Cpre}_{\mathbb{C}}^*)\}$. The set of states of the machine is the antichain of \preceq -minimal elements that belong to \mathcal{P}_1 in the fixed point. Note that the length of this antichain is bounded by an exponential in the size of the game.
- m_0 is any element (t, u) in M such that $t = s_{\text{init}}$ and $u \leq \bar{c}$. Note that such an element is guaranteed to exist as $(s_{\text{init}}, \bar{c}) \in \text{Cpre}_{\mathbb{C}}^*$.
- For all $(t, u) \in M$, we define $\alpha_n((t, u))$ by choosing any element $(t, t') \in E$ such that there exists $(t', u') \in \text{Cpre}_{\mathbb{C}}^*$ with $u' = u + w(t, t')$. Such an element is guaranteed to exist by definition of $\text{Cpre}_{\mathbb{C}}$ and the fact that $(t, u) \in \text{Cpre}_{\mathbb{C}}^*$.
- $\alpha_u: M \times ((S_2 \times S) \cap E) \rightarrow M$ is any partial function that respects the following constraint: if $\alpha_n((t, u)) = (t, t')$ then $\alpha_u((t, u), (t', t''))$ is defined for any $(t', t'') \in E$ and can be chosen to be equal to any (t'', u'') such that $u'' \leq u + w(t, t') + w(t', t'')$, and such an u'' is guaranteed to exist by definition of $\text{Cpre}_{\mathbb{C}}$ and because $\text{Cpre}_{\mathbb{C}}^*$ is a fixed point.

Now, let us prove that for any initial prefix $s_0 s_1 \dots s_{2n}$ of even length in G , which is compatible with \mathcal{M} , we have that $\bar{c} + \text{EL}(s_0 s_1 \dots s_{2n-1}) \geq 0$ and that $\bar{c} + \text{EL}(s_0 s_1 \dots s_{2n}) \geq 0$. To establish this property, we first prove the following property by induction on n : $\bar{c} + \text{EL}(s_0 s_1 \dots s_{2n}) \geq u$ where u is the energy level of the label of the state reached after reading the prefix $s_0 s_1 \dots s_{2n}$ with the Moore machine \mathcal{M} . Base case $n = 0$ is trivial. Induction: assume that the property is true for $n - 1$, and let us establish it for n . By induction hypothesis, we have that $\bar{c} + \text{EL}(s_0 s_1 \dots s_{2(n-1)}) \geq u$ where u is the energy level of the label of state m that is reached after reading $s_0 s_1 \dots s_{2(n-1)}$ with the Moore machine. Now, assume that $\alpha_n(m) = (t, t')$. So, $s_{2(n-1)} = t$ and the choice of \mathcal{P}_1 is to play (t, t') . So, $s_{2(n-1)+1} = t'$. Now for all possible choices (t', t'') of \mathcal{P}_2 , we know by definition of \mathcal{M} that the energy level u'' that labels the state $\alpha_u(m, (t', t''))$ is $u'' \leq u + w(t, t') + w(t', t'')$, which establishes our property. Therefore, the strategy of \mathcal{P}_1 based on \mathcal{M} is such that the energy always stays positive for initial credit \bar{c} , which concludes the proof. \square

Completeness of the symbolic algorithm is guaranteed when a sufficiently large threshold \mathbb{C} is used as established in the following lemma.

Lemma 12 (Completeness). *Let $G = (S_1, S_2, s_{init}, E, k, w)$ be a multi energy game in which all absolute values of weights are bounded by W . If \mathcal{P}_1 has a winning strategy in G and $T = (Q, R)$ is a self-covering tree for G of depth l , then $(s_{init}, (\mathbb{C}, \dots, \mathbb{C})) \in \text{Cpre}_{\mathbb{C}}^*$ for $\mathbb{C} = 2 \cdot l \cdot W$.*

Remark 1. This algorithm is complete in the sense that if a winning strategy exists for \mathcal{P}_1 , it outputs at least a winning initial credit (and the derived strategy) for $\mathbb{C} = 2 \cdot l \cdot W$. However, this is different from the *fixed initial credit problem*, which consists in deciding if a particular given credit vector is winning and is known to be EXPSPACE-hard by equivalence with deciding the existence of an infinite run in a Petri net given an initial marking [12,33]. In general, there may exist winning credits incomparable to those captured by algorithm CpreFP. More precisely, given a constant $\mathbb{C} \in \mathbb{N}$, the algorithm fully captures all the winning initial credits smaller than $(\mathbb{C}, \dots, \mathbb{C})$. Indeed, the fixed point computation considers the whole range of initial credits up to the given constant exhaustively, and only removes credits if they do not suffice to win. By Lemma 12, it is moreover guaranteed that if an arbitrary winning initial credit exists, then there exists one in the range defined by the constant $\mathbb{C} = 2 \cdot l \cdot W$. Nevertheless, since our algorithm works in exponential time while the problem of finding *all* the winning initial credits is EXPSPACE-hard, there may be some incomparable credits outside that range that are not captured by the algorithm (comparable credits are captured since we work with upper closed sets). Indeed, if our algorithm was able to compute exhaustively all winning credits in exponential time, this would induce that EXPTIME is equal to EXPSPACE. Notice that defining a class of games for which the algorithm CpreFP proves to be incomplete (in the sense that incomparable winning credits exist outside the region captured by constant $\mathbb{C} = 2 \cdot l \cdot W$) is an interesting open problem.

Proof. To establish this property, we first prove that from the set of labels of T , we can construct a set f which is increasing for the operator $\text{Cpre}_{\mathbb{C}}$, i.e., $\text{Cpre}_{\mathbb{C}}(f) \supseteq f$, and such that $(s_{init}, (\mathbb{C}, \dots, \mathbb{C})) \in f$. We define f from $T = (Q, R)$ as follows. Let $C \in \mathbb{N}$ be the smallest non-negative integer such that for all $q \in Q$, with $\Theta(q) = (t, u)$, for all dimensions i , $1 \leq i \leq k$, we have that $u(i) + C \geq 0$. Integer C is bounded from above by $l \cdot W$ because on every path from the root to a leaf in T , every dimension is at most decreased l times by an amount bounded by W , and at the root all the dimensions are equal to 0. For any $q \in Q$, we denote by $\Theta(q) + C$ the label of q where the energy level has been increased by C in all the dimensions, i.e., if $\Theta(q) = (t, u)$ then $\Theta(q) + C = (t, u + (C, \dots, C))$. Note that for all nodes in Q , the label is at most $l \cdot W$ and thus the shifted label remains under $\mathbb{C} = 2 \cdot l \cdot W$. Now, we define the set f as follows:

$$f = \{(t, u) \in U(\mathbb{C}) \mid \exists q \in Q, \Theta(q) + C \preceq (t, u)\}. \quad (4)$$

So, f is defined as the \preceq -closure of the set of labels in T shifted by C in all the dimensions.

First, note that $(s_{init}, (\mathbb{C}, \dots, \mathbb{C})) \in f$ as the label of the root in T is $(s_{init}, (0, \dots, 0))$. Second, let us show that $\text{Cpre}_{\mathbb{C}}(f) \supseteq f$. Take any $(t, u) \in f$ and let us show that $(t, u) \in \text{Cpre}_{\mathbb{C}}(f)$. We decompose the proof in two cases. **(A)** $t \in S_1$. By definition of f , there exists $q \in Q$ such that $\Theta(q) + C \preceq (t, u)$. W.l.o.g. we can assume that q is not a leaf as otherwise there exists an ancestor q' of q such that $\Theta(q') \preceq \Theta(q)$ (recall the set is described by its minimal elements). By definition of T , there exists $(t, t') \in E$ and $q' \in Q$ such that $(q, q') \in R$ and $\Theta(q') = \Theta(q) + w(t, t')$. Let $(t', v) = \Theta(q') + C$. By definition of f , we have $(t', v) \in f$. By Eq. (1), it follows that $(t, u) \in \text{Cpre}_{\mathbb{C}}(f)$. **(B)** $t \in S_2$. By definition of f , there exists $q \in Q$ such that $\Theta(q) + C \preceq (t, u)$. Again, w.l.o.g. we can assume that q is not a leaf as otherwise there exists an ancestor q' of q such that $\Theta(q') \preceq \Theta(q)$. By definition of T , for all $(t, t') \in E$, there is $q' \in Q$ such that $(q, q') \in R$ and $\Theta(q') = \Theta(q) + w(t, t')$. Let $(t', v) = \Theta(q') + C$. By definition of f , we have $(t', v) \in f$. By Eq. (1), it follows that $(t, u) \in \text{Cpre}_{\mathbb{C}}(f)$.

Now, let us show that $f \subseteq \text{Cpre}_{\mathbb{C}}^*$. This is a direct consequence of the monotonicity of $\text{Cpre}_{\mathbb{C}}$: it is well known that for any monotone function on a complete lattice, its greatest fixed point is equal to the least upper bound of all post-fixed points (points e such that $e \subseteq \text{Cpre}_{\mathbb{C}}(e)$), i.e., $\text{Cpre}_{\mathbb{C}}^* = \bigcup \{e \mid e \subseteq \text{Cpre}_{\mathbb{C}}(e)\} \supseteq f$. As $(s_{init}, (\mathbb{C}, \dots, \mathbb{C})) \in f$, that concludes the proof. \square

Remark 2. Note that the exponential bound on memory, obtained in Lemma 5, can also be derived from the Moore machine construction of Lemma 11 as this method is complete according to Lemma 12. Still, the DAG construction of Lemma 5 is interesting in its own right, and introduces the concept of node merging, which is underlying to the symbolic algorithm correctness, while transparent in its use.

Incrementality. While the threshold $2 \cdot l \cdot W$ is sufficient, it may be the case that \mathcal{P}_1 can win the game even if its energy level is bounded above by some smaller value. So, in practice, we can use Lemma 11, to justify an incremental algorithm that first starts with small values for the parameter \mathbb{C} and stops as soon as a winning strategy is found or when the value of \mathbb{C} reaches the threshold $2 \cdot l \cdot W$ and no winning strategy has been found.

Application of the symbolic algorithm to MEPGs and MMPGs. Using the reduction of Lemma 4 that allows us to remove the parity condition, and the equivalence between multi energy games and multi mean-payoff games for finite-memory strategies (given by [20, Theorem 3]), along with Lemma 9 (complexity), Lemma 11 (correctness) and Lemma 12 (completeness), we obtain the following result.

Theorem 2 (Symbolic and incremental synthesis algorithm). *Let G_p be a multi energy (resp. multi mean-payoff) parity game. Algorithm CpreFP is a symbolic and incremental algorithm that synthesizes a winning strategy in G_p of at most exponential size memory, if a winning (resp. finite-memory winning) strategy exists. In the worst-case, the algorithm CpreFP takes exponential time.*

Proof. The correctness and completeness for algorithm CpreFP on multi energy games are resp. given by Lemma 11 and Lemma 12. Extension to mean-payoff games (under finite memory) is given by [20, Theorem 3], whereas the parity condition can be encoded as energy thanks to Lemma 4. Exponential worst-case complexity of the algorithm CpreFP is induced by Lemma 9. \square

Integration in synthesis tools. Following the conference version of this paper [23], our results on strategy synthesis have been used in the Acacia+ synthesis tool. This tool originally handled the synthesis of controllers for specifications expressed in LTL (Linear Temporal Logic, a classical formalism for formal specifications [39]) using antichain-based algorithms and has recently been extended to the synthesis from LTL specifications with mean-payoff objectives [8]. The addition of multi mean-payoff objectives to LTL specifications provides a convenient way to enforce that synthesized controllers also satisfy some reasonable behavior from a quantitative standpoint, such as minimizing the number of unsolicited grants in a client-server architecture with prioritized clients. Numerous practical applications may benefit from this multi-dimension framework.

The authors present an approach in which the corresponding synthesis problem ultimately reduces to strategy synthesis on a multi energy game [8, Theorem 26]. Their implementation uses fixed point computations similar to Eq. (2) and has proved efficient (considering the complexity of the problem) in practice. It uses antichains to provide a compact representation of upper-closed sets and implements the incremental approach proposed before (regarding the constant \mathbb{C}). In practical benchmarks, winning strategies can generally be found for rather small values of \mathbb{C} . Hence, the incremental approach overcomes the need to compute up to the exponential theoretical bound $\mathbb{C} = 2 \cdot l \cdot W$ in many cases. Sample benchmarks and experiments can be found in [8], and the tool can be used online [1].

5 Trading finite memory for randomness

In this section, we answer the fundamental question regarding the trade-off of memory for randomness in strategies: we study on which kind of games \mathcal{P}_1 can replace a pure finite-memory winning strategy by an equally powerful, yet conceptually simpler, randomized memoryless one and discuss how memory is encoded into probability distributions. Note that we do not consider wider strategy classes (e.g., randomized finite-memory), nor do we allow randomization for \mathcal{P}_2 (which on most cases is dispensable anyway). Indeed, we aim at a better understanding of the underlying mechanics of memory and randomization, in order to provide alternative strategy representations of practical use; not exploration of more complex games with wider strategy classes (Lemma 21 shows a glimpse of it).

	Multi energy and energy parity	Multi MP (parity)	MP parity
one-player	×	✓	✓
two-player	×	×	✓

Table 1. When pure finite memory for \mathcal{P}_1 can be traded for randomized memorylessness.

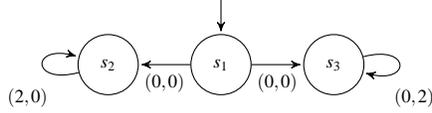


Fig. 5. Randomization can replace memory, but not the opposite.

We present an overview of our results in Tab. 1 and summarize them in Theorem 3. Note that we do not consider the opposite implication, i.e., does there always exist a way of encoding a randomized memoryless strategy into an equivalent finite-memory one. In general, this is not the case even for classes of games where we can trade memory for randomness, and it can easily be witnessed on the one-player multi mean-payoff game depicted on Fig. 5. Indeed, expectation $(1, 1)$ is achievable with a simple uniform distribution while it is not achievable with a pure, arbitrary high memory strategy (even infinite).

We break down these results into three subsections: energy games, multi mean-payoff (parity) games, and single mean-payoff parity games. We start with energy games.

5.1 Randomization and energy games

Randomization is not helpful for energy objectives, even in one-player games. The proof argument is obtained from the intuition that energy objectives are similar in spirit to safety objectives.

Lemma 13. *Randomization is not helpful for almost-sure winning in one-player and two-player energy, multi energy, energy parity and multi energy parity games: if there exists a finite-memory randomized winning strategy, then there exists a pure winning strategy with the same memory requirements.*

Proof. Let G_p be a game fitted with an energy objective. Consider an almost-sure winning strategy λ_1 . If there exists a single path π consistent with λ_1 that violates the energy objective, then there exists a finite prefix witness ρ to violate the energy objective. Moreover, as the finite prefix has positive probability (otherwise the play is not consistent), and the strategy λ_1 is almost-sure winning, it follows that no such path exists. In other words, λ_1 is a sure winning strategy. Since randomization does not help for sure winning strategy, it follows that randomization is not helpful for one-player and two-player energy, multi energy, energy parity and multi energy parity games. \square

5.2 Randomization and multi mean-payoff (parity) games

Randomized memoryless strategies can replace pure finite-memory ones in the one-player multi mean-payoff parity case, but not in the two-player one, even without parity. We first note a useful link between satisfaction and expectation semantics for the mean-payoff objective.

Lemma 14. *Let $G = (S_1, S_2, s_{init}, E, k, w)$ be a game structure with mean-payoff objective $\phi = \text{MeanPayoff}_G(v)$ for some threshold vector $v \in \mathbb{Q}^k$. Let $\lambda_1 \in \Lambda_1$ be a strategy of \mathcal{P}_1 . If λ_1 is almost-sure winning for ϕ (i.e., winning for 1-satisfaction), then λ_1 is also winning for v -expectation for the mean-payoff function MP. The opposite does not hold.*

Proof. We first discuss the claimed implication. Suppose 1-satisfaction is verified. Then, for all strategy $\lambda_2 \in \Lambda_2$ of \mathcal{P}_2 , the set of consistent plays of value $\geq v$ has measure 1, while the one of value $< v$ has measure 0, by definition. Therefore, the expectation $\mathbb{E}_{s_{init}}^{\lambda_1, \lambda_2}(\text{MP})$ is at least v and v -expectation is verified.

To show that the opposite does not hold, consider the simple one-player game depicted on Fig. 5. Let λ_1 be a simple coin flipping on s_1 , i.e., $\lambda_1(s_1)(s_2) = 1/2$, $\lambda_1(s_1)(s_3) = 1/2$, $\lambda_1(s_2)(s_2) = 1$ and $\lambda_1(s_3)(s_3) = 1$. The expectation of this strategy is $v = (1, 1)$. Nevertheless, the probability of achieving mean-payoff of at least v is $0 < 1$, which shows that it does not verify 1-satisfaction for $\text{MeanPayoff}_G(v)$. \square

The fundamental difference between energy and mean-payoff is that energy requires a property to be satisfied *at all times* (in that sense, it is similar to safety), while mean-payoff is a *limit* property. As a consequence, what matters here is the long-run frequencies of weights, not their order of appearance, as opposed to the energy case.

Lemma 15. *Pure finite-memory winning strategies can be traded for equally powerful randomized memoryless ones for one-player multi mean-payoff parity games, for both satisfaction and expectation semantics. For two-player games, randomized memoryless strategies are not as powerful, even limited to expectation semantics, no parity condition, and only 2 dimensions.*

For the one-player case, we extract the frequencies of visit for edges of the graph from the regular outcome that arises from the finite-memory strategy of \mathcal{P}_1 . We build a randomized strategy with probability distributions on edges that yield the exact same frequencies in the long-run. Therefore, if the original pure finite-memory of \mathcal{P}_1 is surely winning, the randomized one is almost-surely winning. For the two-player case, this approach cannot be used as frequencies are not well defined, since the strategy of \mathcal{P}_2 is unknown. Consider a game which needs perfect balance between frequencies of appearance of two sets of edges in a play to be winning (Fig. 6). To almost-surely achieve mean-payoff vector $(0,0)$, \mathcal{P}_1 must ensure that the long-term balance between edges (s_4, s_5) and (s_4, s_6) is the same as the one between edges (s_1, s_3) and (s_1, s_2) . This is achievable with memory as it suffices to react immediately to compensate the choice of \mathcal{P}_2 . However, given a randomized memoryless strategy of \mathcal{P}_1 , \mathcal{P}_2 always has a strategy to enforce that the long-term frequency is unbalanced, and thus the game cannot be won almost-surely by \mathcal{P}_1 with such a strategy. Achieving expected mean-payoff $(0,0)$ is also excluded.

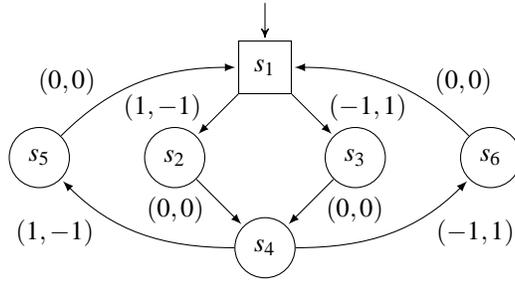


Fig. 6. Memory is needed to enforce perfect long-term balance.

Proof. We begin with the one-player case. Let G_p be a multi mean-payoff parity game. Let $\lambda_1^{pf} \in \Lambda_1^{PF}$ be the pure finite-memory strategy of the player. Since it is pure and finite, its outcome is a regular word $\pi = \rho_1 \cdot (\rho_2)^\omega$, with $\rho_1 \in S^*$, $\rho_2 \in S^+$. Let $\phi = \text{MeanPayoff}_{G_p}(v) \cap \text{Parity}_{G_p}$ be the multi mean-payoff parity objective for some threshold vector $v \in \mathbb{Q}^k$. Suppose this strategy verifies α -satisfaction for ϕ and β -expectation for the MP function, for some α, β . We claim that there exists a randomized memoryless strategy $\lambda_1^{rm} \in \Lambda_1^{RM}$ that is also α -satisfying for ϕ and that satisfies β -expectation for the MP function; and we show how to build it.

We denote concatenation by the \cdot symbol. Given a finite word $\rho \in S^*$, two states $s, s' \in S$, we resp. denote by $\text{occ}(s, \rho)$ and $\text{occ}((s, s'), \rho)$ the number of occurrences of the state s and the transition (s, s') in the word ρ . We add the subscript \circ when we count the first state of the word as the successor of the last one (i.e., the word is a cycle in the game graph). That is, $\text{occ}_\circ(*, \rho) = \text{occ}(*, \rho \cdot \text{First}(\rho))$.

Let us consider the mean-payoff of the outcome of strategy λ_1^{pf} . Recall that for a play $\pi \in \text{Plays}(G)$, $\pi = s^1, s^2, s^3, \dots$, we have $\text{MP}(\pi) = \liminf_{n \rightarrow \infty} \frac{1}{n} \sum_{1 \leq i < n} w(s^i, s^{i+1})$. Since the play induced by λ_1^{pf} is regular, the limit is well defined and we may express the mean-payoff in terms of frequencies, that is

$$\text{MP}(\pi) = \sum_{(s, s') \in E} w(s, s') \cdot \text{freq}_\infty((s, s')),$$

where freq_∞ denotes the long-term frequency of a transition defined as

$$\forall (s, s') \in E, \quad \text{freq}_\infty((s, s')) = \frac{\text{occ}_\circ((s, s'), \rho_2)}{|\rho_2|}.$$

We define the randomized memoryless strategy λ_1^{rm} as follows: $\forall s, s' \in S, (s, s') \in E, X = \{(s, t) \mid t \in S, (s, t) \in (\rho_1 \cdot \text{First}(\rho_2))\}$,

$$\lambda_1^{rm}(s)(s') = \begin{cases} \frac{1}{|X|} & \text{if } s \in \rho_1 \wedge s \notin \rho_2, \\ \frac{\text{occ}_o((s, s'), \rho_2)}{\text{occ}(s, \rho_2)} & \text{if } s \in \rho_2, \\ 0 & \text{otherwise.} \end{cases}$$

Intuitively, we fix a uniform distribution over transitions of the finite prefix ρ_1 as we only need to ensure reaching the bottom strongly connected component (BSCC) defined by ρ_2 with probability 1, and the relative frequencies in ρ_1 do not matter (because these weights and priorities are negligible in the long run). On the contrary, we use the exact frequencies for transitions of ρ_2 as they prevail long-term wise. Note that λ_1^{rm} is a correctly defined randomized memoryless strategy.

Obviously, λ_1^{rm} yields a Markov chain over states of $(\rho_1 \cup \rho_2)$ such that states of $(\rho_1 \setminus \rho_2)$ are transient and states of ρ_2 constitute a BSCC that is reached with probability one. Thus, the mean-payoff induced by λ_1^{rm} is totally dependent on this BSCC mean-payoff value. As a consequence, proving that transition frequencies in the BSCC are exactly the same as frequencies freq_∞ defined by λ_1^{pf} will imply the claim on mean-payoff. Moreover, parity will remain satisfied as the sets of infinitely often visited states will be the same for both the pure and the randomized strategy. Let $T = \{t_1, t_2, \dots, t_m\}$ be the set of states that appear in ρ_2 . This BSCC is an ergodic Markov chain $\mathcal{M}_e = (T, P)$ with the following matrix of transition probabilities:

$$P = \begin{matrix} & t_1 & \dots & t_m \\ \begin{matrix} t_1 \\ \vdots \\ t_m \end{matrix} & \begin{pmatrix} \frac{\text{occ}_o((t_1, t_1), \rho_2)}{\text{occ}(t_1, \rho_2)} & & \\ & \ddots & \\ & & \frac{\text{occ}_o((t_m, t_m), \rho_2)}{\text{occ}(t_m, \rho_2)} \end{pmatrix} \end{matrix}.$$

Classical analysis of ergodic Markov chains grants the existence of a unique probability vector v such that $vP = v$, i.e.,

$$\forall 1 \leq i \leq m, v_i = \sum_{1 \leq j \leq m} \frac{\text{occ}_o((t_j, t_i), \rho_2)}{\text{occ}(t_j, \rho_2)} \cdot v_j.$$

This vector v represents the occurrence frequency of each state in an infinite run over the Markov chain. It is easy to see that the unique probability vector v that satisfies $vP = v$ is

$$v = \left(\frac{\text{occ}(t_1, \rho_2)}{|\rho_2|}, \dots, \frac{\text{occ}(t_m, \rho_2)}{|\rho_2|} \right).$$

Moreover, given a transition of the Markov chain, its frequency is simply the product of the frequency of its starting state by the probability of the transition when the chain is in this state: for all $t, t' \in T$, we have $\text{freq}_\infty^{\mathcal{M}_e}((t, t')) = v(t) \cdot P(t, t')$. By definition of v and P , that is

$$\text{freq}_\infty^{\mathcal{M}_e}((t, t')) = \frac{\text{occ}_o((t, t'), \rho_2)}{|\rho_2|} = \text{freq}_\infty((t, t')),$$

thus proving that the randomized strategy λ_1^{rm} *almost-surely* yields the same mean-payoff and parity as the pure finite-memory one λ_1^{pf} . The expected value threshold is also verified by Lemma 14.

Now it remains to show that this does not carry over to two-player games. Indeed, we show that randomized memoryless strategies cannot replace pure finite-memory ones for the expectation semantics, even without parity. By

Lemma 14, this implies that it cannot be verified for 1-satisfaction semantics either. Consider the game depicted on Fig. 6. Player \mathcal{P}_1 has a pure finite-memory strategy λ_1^{pf} that ensures $\text{MP}(\pi) \geq (0,0)$, against all strategy λ_2 of \mathcal{P}_2 . This strategy is simply to take the opposite choice of \mathcal{P}_2 : $\lambda_1^{pf}(*s_2s_4) = s_6$ and $\lambda_1^{pf}(*s_3s_4) = s_5$. Now suppose \mathcal{P}_1 uses a randomized memoryless strategy λ_1^{rm} such that $\lambda_1^{rm}(s_4)(s_5) = p$ and $\lambda_1^{rm}(s_4)(s_6) = 1 - p$, for some $p \in [0, 1]$. We claim that whatever the value of p , there exists a counter-strategy λ_2 for \mathcal{P}_2 such that $\mathbb{E}_{s_1}^{\lambda_1^{rm}, \lambda_2}(\text{MP}) \not\geq (0,0)$. Suppose $p \geq 1/2$ and let $\lambda_2(s_1) = s_2$. Then, we have

$$\mathbb{E}_{s_1}^{\lambda_1^{rm}, \lambda_2}(\text{MP}) = \frac{(1, -1) + [p \cdot (1, -1) + (1 - p) \cdot (-1, 1)]}{4} = \frac{1}{2}(p, -p) \not\geq (0,0).$$

Now suppose $p < 1/2$ and let $\lambda_2(s_1) = s_3$. Then, we have

$$\mathbb{E}_{s_1}^{\lambda_1^{rm}, \lambda_2}(\text{MP}) = \frac{(-1, 1) + [p \cdot (1, -1) + (1 - p) \cdot (-1, 1)]}{4} = \frac{1}{2}(p - 1, 1 - p) \not\geq (0,0).$$

This shows that memory is needed to achieve the $(0,0)$ -expectation objective and concludes our proof. \square

5.3 Randomization and single mean-payoff parity games

Randomized memoryless strategies can replace pure finite-memory ones for single mean-payoff parity games. The proof outline is as follows. We do it in two steps. First, we show that it is the case for the simpler case of *MP Büchi games* (Lemma 18). Suppose \mathcal{P}_1 has a pure finite-memory winning strategy for such a game. We use the existence of particular pure memoryless strategies on winning states: the classical attractor for Büchi states, and a strategy that ensures that cycles of the outcome have positive energy (whose existence follows from [17]). We build an almost-surely randomized memoryless winning strategy for \mathcal{P}_1 by mixing those strategies in the probability distributions, with sufficient probability over the strategy that is good for energy. We illustrate this construction on the simple game G_p depicted on Fig. 7. Let $\lambda_1^{pf} \in \Lambda_1^{PF}$ be a strategy of \mathcal{P}_1 such that \mathcal{P}_1 plays (s_1, s_1) for 8 times, then plays (s_1, s_2) once, and so on. This strategy ensures surely winning for the objective $\phi = \text{MeanPayoff}_{G_p}(3/5) \cap \text{Buchi}_{G_p}(\{s_2\})$. Obviously, \mathcal{P}_1 has a pure memoryless strategy that ensures winning for the Büchi objective: playing (s_1, s_2) . On the other hand, he also has a pure memoryless strategy that ensures cycles of positive energy: playing (s_1, s_1) . Let $\lambda_1^{rm} \in \Lambda_1^{RM}$ be the strategy defined as follows: play (s_1, s_2) with probability γ and (s_1, s_1) with the remaining probability. This strategy is almost-surely winning for ϕ for sufficiently small values of γ (e.g., $\gamma = 1/9$). Second, we extend this result to *MP parity games* using an induction on the number of priorities and the size of games (Lemma 20). We consider *subgames* that reduce to the MP Büchi and MP coBüchi cases. For MP coBüchi games, pure memoryless strategies are known to suffice [22].

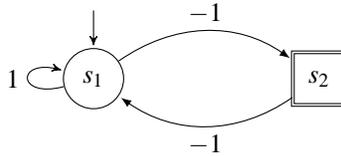


Fig. 7. Mixing strategies that are resp. *good for Büchi* and *good for energy*.

Büchi case. A particular, simpler case of the parity objective is the Büchi objective. It corresponds to parity with priorities $\{0, 1\}$. We denote a Büchi game by $G = (S_1, S_2, s_{init}, E, w, F)$, with F the set of Büchi states such that a play is winning if it visits infinitely often states of the set F . We first state results on these Büchi objectives, as they are conceptually simpler to understand. Proof arguments for parity are more involved and make use of results on Büchi objectives. We sometimes denote the Büchi objective for the set F by $\square \diamond F$ (where \square stands for *globally* and \diamond for *finally*), using the classical LTL formulation [39].

We first introduce the useful notion of ε -optimality. Given a game G_p with a one-dimension⁵ mean-payoff objective, we define its value as

$$\underline{\text{val}} = \sup_{\lambda_1 \in \Lambda_1} \inf_{\lambda_2 \in \Lambda_2} \{v \mid \text{Outcome}_{G_p}(\lambda_1, \lambda_2) \subseteq \text{MeanPayoff}_{G_p}(v)\}.$$

A strategy is said optimal for the mean-payoff objective if it achieves this value. Such a strategy may not need to exist in general, even in one-player games [22,11,18] (Fig. 8, \mathcal{P}_1 has to delay its visits of s_1 for longer and longer intervals in order to tend towards value 1). However, it is known that for all $\varepsilon > 0$, ε -optimal strategies (i.e., that achieve value $(\underline{\text{val}} - \varepsilon)$) always exist in one-dimension mean-payoff games, as a consequence of Martin's theorem on Borel determinacy [37].

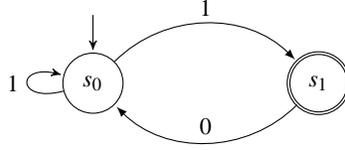


Fig. 8. Mean-payoff Büchi requires infinite memory for optimality.

Here, we show finite-memory strategies can be traded off for randomized memoryless ones for mean-payoff Büchi games. Precisely, we prove that ε -optimality for mean-payoff Büchi games can as well be achieved by randomized memoryless strategies. We first need to state two useful lemmas granting the existence of pure memoryless strategies that are resp. *good-for-energy* or *good-for-Büchi*, in all states that are winning for the mean-payoff Büchi objective. These strategies will help us build the needed ε -optimal strategies.

Lemma 16 (Extension of [17, Lemma 4]). *Let $G = (S_1, S_2, s_{\text{init}}, E, w, F)$, with F the set of Büchi states. Let $\text{Win} \subseteq S$ be the set of winning states for the mean-payoff Büchi objective with threshold 0. For all $s \in \text{Win}$, \mathcal{P}_1 has a uniform (i.e., independent of the starting state) memoryless good-for-energy strategy λ_1^{gfe} whose outcome never leaves the set Win , such that any cycle c of this outcome has energy $\text{EL}(c) \geq 0$.*

Lemma 17 (Classical attractor). *Let $G = (S_1, S_2, s_{\text{init}}, E, w, F)$, with F the set of Büchi states. Let $\text{Win} \subseteq S$ be the set of winning states for the mean-payoff Büchi objective with threshold 0. For all $s \in \text{Win}$, \mathcal{P}_1 has a uniform (i.e., independent of the starting state) memoryless good-for-Büchi strategy $\lambda_1^{\diamond F}$, an attractor strategy for F , whose outcome never leaves the set Win , such that it ensures reaching F in at most $|S|$ steps.*

The randomized memoryless strategy of \mathcal{P}_1 will thus consist in mixing these two strategies, with a very low probability on the good-for-Büchi strategy. Indeed, the Büchi objective will be satisfied whatever this probability is, provided it is strictly positive. On the other hand, by giving more weight to the good-for-energy strategy, \mathcal{P}_1 can obtain a mean-payoff that is arbitrary close to the optimum.

Lemma 18. *In mean-payoff Büchi games, ε -optimality can be achieved surely by pure finite-memory strategies and almost-surely by randomized memoryless strategies.*

Proof. Let $G = (S_1, S_2, s_{\text{init}}, E, w, F)$, with F the set of Büchi states. We consider the mean-payoff objective with threshold 0 (w.l.o.g.). Let $\text{Win} \subseteq S$ be the set of winning states for the mean-payoff Büchi objective. By Lemma 16 and Lemma 17, for all $s \in \text{Win}$, \mathcal{P}_1 has two uniform memoryless strategies λ_1^{gfe} and $\lambda_1^{\diamond F}$, whose outcomes never leave the set Win , such that λ_1^{gfe} ensures that any cycle c of its outcome has energy $\text{EL}(c) \geq 0$, and $\lambda_1^{\diamond F}$, an attractor strategy for F , ensures reaching F in at most $|S|$ steps.

We first build ε -optimal *pure finite-memory* strategies based on these two pure memoryless strategies. Let $\varepsilon > 0$. As usual, W denotes the largest absolute weight on any edge. Let us define λ_1^{pf} such that (a) it plays λ_1^{gfe} for $\left\lceil \frac{2 \cdot W \cdot |S|}{\varepsilon} \right\rceil - |S|$ steps, then (b) it plays $\lambda_1^{\diamond F}$ for $|S|$ steps, then again (a). This ensures that F is visited infinitely often as $\lambda_1^{\diamond F}$ is played

⁵ The multi-dimensional setting gives rise to incomparable outcomes and the need to consider *Pareto-optimality*.

infinitely many times for $|S|$ steps in a row. Furthermore, the total cost of phases (a) + (b) is bounded by $-2 \cdot W \cdot |S|$, and thus the mean-payoff of the outcome is at least $-\varepsilon$, against any strategy of the adversary.

Second, we show that based on the same pure memoryless strategies, it is possible to obtain almost-surely ε -optimal *randomized memoryless* strategies, i.e.,

$$\begin{aligned} & \forall \varepsilon > 0, \exists \lambda_1^{rm} \in \Lambda_1^{RM}, \forall \lambda_2 \in \Lambda_2, \\ & \mathbb{P}_{s_{init}}^{\lambda_1^{rm}, \lambda_2} (\pi \models \square \diamond F) = 1 \wedge \mathbb{P}_{s_{init}}^{\lambda_1^{rm}, \lambda_2} (\text{MP}(\pi) \geq -\varepsilon) = 1. \end{aligned}$$

Note that pure memoryless strategies suffice for \mathcal{P}_2 as he essentially has to win against the Büchi *or* the mean-payoff criterion [11]. Therefore, given $\varepsilon > 0$, we need to build some strategy $\lambda_1^{rm} \in \Lambda_1^{RM}$ such that

$$\forall \lambda_2^{pm} \in \Lambda_2^{PM}, \mathbb{P}_{s_{init}}^{\lambda_1^{rm}, \lambda_2^{pm}} (\pi \models \square \diamond F) = 1 \wedge \mathbb{P}_{s_{init}}^{\lambda_1^{rm}, \lambda_2^{pm}} (\text{MP}(\pi) \geq -\varepsilon) = 1.$$

We build such a strategy as follows:

$$\forall s \in S, \lambda_1^{rm}(s) = \begin{cases} \lambda_1^{gfe}(s) & \text{with probability } 1 - \gamma, \\ \lambda_1^{\diamond F}(s) & \text{with probability } \gamma, \end{cases}$$

for some *well-chosen* $\gamma \in]0, 1[$.

It is straightforward to see that the Büchi objective is almost-surely satisfied for all values of $\gamma > 0$ as at all times, the probability of playing according to $\lambda_1^{\diamond F}$ for $|S|$ steps in a row, and thus ensuring a visit of F , is $\gamma^{|S|}$, which is strictly positive.

It remains to study if it is always possible to choose such a constant γ such that objective $\text{MeanPayoff}_{G_p}(-\varepsilon)$ is almost-surely satisfied. Consider such a strategy $\lambda_1^{rm} \in \Lambda_1^{RM}$ and some fixed strategy $\lambda_2^{pm} \in \Lambda_2^{PM}$ of \mathcal{P}_2 : the game reduces to the finite Markov chain $\mathcal{M}_c = (S, \delta, w)$, where $\delta: E \rightarrow [0, 1]$ is the transition probability function resulting from fixing those strategies. Suppose λ_2^{pm} is winning for \mathcal{P}_2 . Thus, $\mathbb{P}_{s_{init}}^{\mathcal{M}_c} (\text{MP}(\pi) < -\varepsilon) > 0$. The mean-payoff depends on limit behavior: the probability measure of plays that do not enter in a bottom strongly connected component (BSCC) is zero [4], whereas in a BSCC, the expected mean-payoff is the same in all states and it is obtained almost-surely (as follows from definition of BSCCs and prefix-independence of the mean-payoff). This implies that there exists some BSCC \mathcal{C} in \mathcal{M}_c such that $\mathbb{P}_{s_{init}}^{\mathcal{M}_c} (\diamond \mathcal{C}) > 0$ and $\mathbb{E}^{\mathcal{C}} (\text{MP}) < -\varepsilon$.

We claim that it is possible to choose γ such that all BSCCs, in all Markov chains induced by pure memoryless strategies of \mathcal{P}_2 , have expectation greater than or equal to ε , thus proving that strategy λ_1^{rm} is almost-surely ε -optimal with regard to the mean-payoff value function. Intuitively, the smaller this constant γ is chosen, the nearer will the expected mean-payoff induced by λ_1^{rm} be to the one induced by λ_1^{gfe} , that is at least zero. Since the number of pure memoryless strategies of \mathcal{P}_2 is finite, and so is the number of BSCCs induced by λ_1^{rm} (regardless of the exact value of $\gamma \in]0, 1[$, we obtain the same BSCCs in terms of states and edges), one can compute a suitable γ for each of them, and take the minimum to ensure that the property will be satisfied in all possible cases.

Therefore, let us fix some strategy λ_2^{pm} of \mathcal{P}_2 , and some BSCC \mathcal{C} of the induced Markov chain when played against strategy λ_1^{rm} of \mathcal{P}_1 . It remains to show that (*claim*) there exists $\gamma \in]0, 1]$ such that $\mathbb{E}^{\mathcal{C}(\gamma)} (\text{MP}) \geq -\varepsilon$ to conclude this proof. Observe that we write $\mathcal{C}(\gamma)$ as transition probabilities inside \mathcal{C} depend on γ . By contradiction, suppose the claim is false. Precisely, we assume that (*contradiction hypothesis*) for all $\gamma \in]0, 1]$, we have that $\mathbb{E}^{\mathcal{C}(\gamma)} (\text{MP}) < -\varepsilon$.

Besides, observe that for $\gamma = 0$, strategy λ_1^{rm} is exactly equal to λ_1^{gfe} . As we know that λ_1^{gfe} ensures a worst-case mean-payoff at least equal to zero, we trivially deduce that $\mathbb{E}^{\mathcal{C}(0)} \geq 0$. This implies that $\sup_{\gamma \in [0, 1]} \mathbb{E}^{\mathcal{C}(\gamma)} \geq \mathbb{E}^{\mathcal{C}(0)} \geq 0$. Notice that in this case, interval $[0, 1]$ is closed.

By results in the literature, it is known that this supremum is continuous. See for example Solan [44] on the continuity of the optimal expected value function in the general context of competitive Markov decision processes (equivalent to $2\frac{1}{2}$ -player games). Therefore, we have that $\sup_{\gamma \in [0, 1]} \mathbb{E}^{\mathcal{C}(\gamma)} = \sup_{\gamma \in [0, 1]} \mathbb{E}^{\mathcal{C}(\gamma)} \geq \mathbb{E}^{\mathcal{C}(0)} \geq 0$. On the other hand, by (*contradiction hypothesis*), we also have that $\sup_{\gamma \in [0, 1]} \mathbb{E}^{\mathcal{C}(\gamma)} \leq -\varepsilon$. Since ε is strictly positive, there is a clear contradiction, which concludes our proof. \square

Parity Case. Given those results for mean-payoff Büchi games, we now consider the more general case of mean-payoff parity games. We start by introducing the useful notion of *subgames*.

Subgame. Let $G_p = (S_1, S_2, s_{init}, E, k, w, p)$ be a game and $A \subseteq S$ be a subset of states in G_p . If E is such that for all $s \in A$, there exists $s' \in A$ with $(s, s') \in E$, then we define the *subgame* $G_p \downarrow A$ as $(S_1 \cap A, S_2 \cap A, E \cap (A \times A), w', p')$ where w', p' are the functions w, p restricted to the subdomain A . Note that for subgames, we do not consider an initial state.

Let $G_p = (S_1, S_2, s_{init}, E, k, w, p)$ and $U \subseteq S$. We define $\text{Attr}_1(U)$ as the set that is obtained as the limit of the following increasing sequence: $U_0 = U$, and $U_i = U_{i-1} \cup \{s \in S_1 \mid \exists s' \in U_{i-1}, (s, s') \in E\} \cup \{s \in S_2 \mid \forall s', (s, s') \in E, s' \in U_{i-1}\}$, for $i \geq 1$. As this sequence of sets is increasing, there exists $i \leq |S|$ such that $U_j = U_i$ for all $j \geq i$. $\text{Attr}_1(U)$ contains all the states in G from which \mathcal{P}_1 can force a visit to U , and it is well known that \mathcal{P}_1 has a pure memoryless strategy to force such a visit from those states. Also, it is clear that \mathcal{P}_1 does not have a strategy to leave the states in $S \setminus \text{Attr}_1(U)$. Attractors can be defined symmetrically for \mathcal{P}_2 and are noted $\text{Attr}_2(\cdot)$. As direct consequence, we have the following proposition.

Proposition 1. *Let $G_p = (S_1, S_2, s_{init}, E, k, w, p)$ be a game, let $U \subseteq S$ and $\text{Attr}_1(U)$ be such that $B = S \setminus \text{Attr}_1(U)$ is non-empty, then $G_p \downarrow B$ is a subgame.*

The following lemma states that optimal pure memoryless strategies exist for \mathcal{P}_1 in games with mean-payoff coBüchi objectives (i.e., parity with priorities $\{1, 2\}$). For mean-payoff Büchi objectives, we showed in Lemma 18 that, for all $\varepsilon > 0$, ε -optimal randomized memoryless strategies exist.

Lemma 19 ([22, Theorem 5]). *Let $G_p = (S_1, S_2, s_{init}, E, k, w, p)$ be a game with priorities $\{1, 2\}$, and $\text{WIN}_{\geq 0}^p$ be the set of nodes in G_p from which \mathcal{P}_1 wins the mean-payoff coBüchi objective for threshold 0 (w.l.o.g.). Then from all states in $\text{WIN}_{\geq 0}^p$, \mathcal{P}_1 has a pure memoryless winning strategy for the coBüchi mean-payoff objective for threshold 0.*

We now establish that ε -optimal randomized memoryless strategies also exist for mean-payoff parity games, and thus, can replace pure finite-memory ones.

Lemma 20. *Let $G_p = (S_1, S_2, s_{init}, E, k, w, p)$ and $\text{WIN}_{\geq 0}^p$ be the set of nodes in G_p from which \mathcal{P}_1 wins the mean-payoff parity objective for threshold 0. Then for all $\varepsilon > 0$, there exists $\lambda_1^{rm} \in \Lambda_1^{RM}$, such that for all $s \in \text{WIN}_{\geq 0}^p$ and for all $\lambda_2 \in \Lambda_2$, we have that:*

$$\mathbb{P}_s^{\lambda_1^{rm}, \lambda_2} (MP(\pi) \geq -\varepsilon) = 1 \wedge \mathbb{P}_s^{\lambda_1^{rm}, \lambda_2} (Par(\pi) \bmod 2 = 0) = 1.$$

Proof. The proof is by induction on the lexicographic order \preceq on games, defined as follows: $G_p^1 \preceq G_p^2$ if G_p^1 has less priorities than G_p^2 or G_p^1 has the same priorities than in G_p^2 but less states. Clearly, this lexicographic order is well-founded.

The base cases are twofold: one for the number of states, and one for priorities. First, if the game is such that $|S| = 1$, then obviously, if \mathcal{P}_1 can win, he can do so with a pure memoryless strategy, which respects the claim. Second, for two priorities. W.l.o.g., we can assume that all priorities are either in $\{0, 1\}$ or in $\{1, 2\}$. Those cases resp. correspond to mean-payoff Büchi and mean-payoff coBüchi games. The result for mean-payoff Büchi games has been established in Lemma 18, while the result for mean-payoff coBüchi games is a direct consequence of Lemma 19, as pure memoryless strategies are a special case of randomized memoryless strategies.

Let us now consider the inductive case. Suppose we have a mean-payoff parity game G_p with m priorities and $|S|$ states. W.l.o.g., we can make the assumption that the lowest priority in G_p is either 0 or 1, otherwise we subtract an even number to all priorities so that we are in that case. Let $U_0 = \{s \in \text{WIN}_{\geq 0}^p \mid p(s) = 0\}$ and $U_1 = \{s \in \text{WIN}_{\geq 0}^p \mid p(s) = 1\}$.

We consider the two possible following situations corresponding to U_0 empty or not.

1. U_0 empty. In that case U_1 is not empty. Let us consider $A_2 = \text{Attr}_2(U_1)$ the attractor of \mathcal{P}_2 for U_1 . It must be the case that $\text{WIN}_{\geq 0}^p \setminus A_2$ is non-empty, otherwise this would contradict the fact that \mathcal{P}_1 is winning the parity objective from states in $\text{WIN}_{\geq 0}^p$. Indeed, if it was not the case, then \mathcal{P}_2 would be able to force an infinite number of visits to U_1 from all states in $\text{WIN}_{\geq 0}^p$, and the parity would be odd as U_0 is empty, a contradiction with the definition of

$\text{WIN}_{\geq 0}^p$. (i) Let $B = \text{WIN}_{\geq 0}^p \setminus A_2$. First note that, as B is non-empty, by Proposition 1, $G_p \downarrow B$ is a subgame. Also, note that from all states in B , it must be the case that \mathcal{P}_1 has a winning strategy that does not require visits of the states outside B , i.e., states in A_2 , for otherwise this would lead to a contradiction with the fact that \mathcal{P}_1 is winning the parity objective in $\text{WIN}_{\geq 0}^p$. So all states in the subgame $G_p \downarrow B$ are winning for \mathcal{P}_1 . The game $G_p \downarrow B$ does not contain states with priority $\bar{0}$, and so we can apply our induction hypothesis to conclude that \mathcal{P}_1 has a memoryless randomized strategy from all states in B , as $(G_p \downarrow B) \preceq G_p$ since it has one less priority. (ii) Now, let us concentrate on states in A_2 . Let $A_1 = \text{Attr}_1(B)$. From states in A_1 , \mathcal{P}_1 has a pure memoryless strategy to reach states in B , and so from there \mathcal{P}_1 can play as in $G_p \downarrow B$, and we are done. Let $C = A_2 \setminus A_1$. If C is empty, we are done. Otherwise, by Proposition 1, $G_p \downarrow C$ is a subgame (\mathcal{P}_2 can force to stay within C). We conclude that all states in this game must be winning for \mathcal{P}_1 . This game has the same minimal priority than in the original game (i.e., priority 1) but it has at least one state less, and so we can apply our induction hypothesis to conclude that \mathcal{P}_1 has a memoryless randomized strategy from all states in C . Therefore, by (i) and (ii), \mathcal{P}_1 has a memoryless randomized strategy from all states in $\text{WIN}_{\geq 0}^p$, which proves the claim in that case.

2. U_0 is not empty. Let us consider $A_1 = \text{Attr}_1(U_0)$. (iii) First, consider the case where $A_1 = \text{WIN}_{\geq 0}^p$. In this case, it means that \mathcal{P}_1 can force a visit to states in U_0 from any states in $\text{WIN}_{\geq 0}^p$. So, we conclude that \mathcal{P}_1 wins in G_p the mean-payoff Büchi game with threshold 0, and by Lemma 18, we conclude that \mathcal{P}_1 has a memoryless randomized strategy from all states in G_p for almost surely winning the parity game with mean-payoff threshold 0 so we are done. (iv) Second, consider the case where $B = \text{WIN}_{\geq 0}^p \setminus A_1$ is non-empty. Then by Proposition 1, $G_p \downarrow B$ is a subgame. So \mathcal{P}_2 can force to stay within B in the original game and so we conclude that all states in the game $G_p \downarrow B$ are winning for \mathcal{P}_1 . As $G_p \downarrow B$ does not contain states of priority 0, and thus has at least one less priority, we can apply the induction hypothesis to conclude that \mathcal{P}_1 has a memoryless randomized strategy from all states in B . Therefore, by (iii) and (iv), \mathcal{P}_1 has a memoryless randomized strategy from all states in $\text{WIN}_{\geq 0}^p$, which also proves the case.

As we have proved the claim in both possible cases, this concludes the proof. \square

5.4 Summary for randomization

We sum up results for these different classes of games in Theorem 3 (cf. Table 1).

Theorem 3 (Trading finite memory for randomness). *The following assertions hold: (1) Randomized strategies are exactly as powerful as pure strategies for energy objectives. Randomized memoryless strategies are not as powerful as pure finite-memory strategies for almost-sure winning in one-player and two-player energy, multi energy, energy parity and multi energy parity games. (2) Randomized memoryless strategies are not as powerful as pure finite-memory strategies for almost-sure winning in two-player multi mean-payoff games. (3) In one-player multi mean-payoff parity games, and two-player single mean-payoff parity games, if there exists a pure finite-memory sure winning strategy, then there exists a randomized memoryless almost-sure winning strategy.*

Proof. (1) For energy games, results follow from Lemma 13. (2) For two-player multi mean-payoff games, they follow from Lemma 15. (3) For one-player multi mean-payoff games, they follow from Lemma 15. For two-player single mean-payoff parity, they are direct consequence of Lemma 20. \square

We close this section by observing that there are even more powerful classes of strategies. Their study, as well as their practical interest, remains open.

Lemma 21. *Randomized finite-memory strategies are strictly more powerful than both randomized memoryless and pure finite-memory strategies for multi-mean payoff games with expectation semantics, even in the one-player case.*

The intuition is essentially that memory permits to achieve an exact payoff by sticking to a given side, while randomization permits to combine payoffs of pure strategies to achieve any linear combination in between.

Proof. Consider the game G depicted on Fig. 9. Whatever the pure finite-memory strategy of \mathcal{P}_1 , the only achievable mean-payoff values are $(1, -1)$ (if (s_0, s_1) is never taken) and $(-1, 1)$ (if (s_0, s_1) is taken). This is also true for randomized memoryless strategies: either the probability of (s_0, s_1) is null and the mean-payoff has value $(1, -1)$, or this

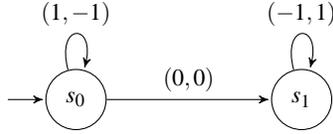


Fig. 9. Randomized finite memory is strictly more powerful than randomized memorylessness and pure finite memory.

probability is strictly positive, and the mean-payoff has value $(-1, 1)$ as the probability mass will eventually reach s_1 . On the contrary, value $(0, 0)$ is achievable by a randomized finite-memory strategy. Indeed, consider the strategy that tosses a coin in its first visit of s_0 to decide if it will always play (s_0, s_0) or if it will play (s_0, s_1) and then always (s_1, s_1) . This strategy only needs one bit of memory and one bit to encode probabilities, and still, it is strictly more powerful than any amount of pure memory or any arbitrary high precision for probabilities without memory. \square

6 Conclusion

In this work, we considered the finite-memory strategy synthesis problem for games with multiple quantitative (energy and mean-payoff) objectives along with a parity objective. We established tight (matching upper and lower) exponential bounds on the memory requirements for such strategies (Theorem 1), significantly improving the previous triple exponential bound for multi energy games (without parity) that could be derived from results in literature for games on VASS. We presented an optimal symbolic and incremental strategy synthesis algorithm (Theorem 2). As discussed in Section 4, the presented algorithm has been used as part of the synthesis tool Acacia+ for specifications combining LTL properties and multi-dimensional quantitative objectives [8] and has proved efficient in practice. Finally, we also presented a precise characterization of the trade-off of memory for randomness in strategies (Theorem 3).

References

1. Acacia+. <http://lit2.ulb.ac.be/acaciaplus/>.
2. P.A. Abdulla, Y.-F. Chen, L. Holík, R. Mayr, and T. Vojnar. When simulation meets antichains. In *Proc. of TACAS*, LNCS 6015. Springer, 2010.
3. R. Alur, T.A. Henzinger, and O. Kupferman. Alternating-time temporal logic. *J. ACM*, 49(5):672–713, 2002.
4. C. Baier and J.-P. Katoen. *Principles of model checking*. MIT Press, 2008.
5. J. Bernet, D. Janin, and I. Walukiewicz. Permissive strategies: from parity games to safety games. *ITA*, 36(3):261–275, 2002.
6. R. Bloem, K. Chatterjee, T.A. Henzinger, and B. Jobstmann. Better quality in synthesis through quantitative objectives. In *Proc. of CAV*, LNCS 5643, pages 140–156. Springer, 2009.
7. R. Bloem, K. Greimel, T.A. Henzinger, and B. Jobstmann. Synthesizing robust systems. In *Proc. of FMCAD*, pages 85–92. IEEE, 2009.
8. A. Bohy, V. Bruyère, E. Filiot, and J.-F. Raskin. Synthesis from LTL specifications with mean-payoff objectives. In *Proc. of TACAS*, LNCS 7795, pages 169–184. Springer, 2013.
9. I. Borosh and B. Treybig. Bounds on positive integral solutions of linear diophantine equations. *Proc. of the American Mathematical Society*, 55(2):299–304, 1976.
10. P. Bouyer, U. Fahrenberg, K.G. Larsen, N. Markey, and J. Srba. Infinite runs in weighted timed automata with energy constraints. In *Proc. of FORMATS*, LNCS 5215, pages 33–47. Springer, 2008.
11. P. Bouyer, N. Markey, J. Olschewski, and M. Ummels. Measuring permissiveness in parity games: Mean-payoff parity games revisited. In *Proc. of ATVA*, LNCS 6996, pages 135–149. Springer, 2011.
12. T. Brázdil, P. Jancar, and A. Kucera. Reachability games on extended vector addition systems with states. In *Proc. of ICALP*, LNCS 6199, pages 478–489. Springer, 2010.
13. V. Bruyère, E. Filiot, M. Randour, and J.-F. Raskin. Meet your expectations with guarantees: Beyond worst-case synthesis in quantitative games. In *Proc. of STACS*, LIPIcs 25, pages 199–213. Schloss Dagstuhl - LZI, 2014.
14. P. Cerný, K. Chatterjee, T.A. Henzinger, A. Radhakrishna, and R. Singh. Quantitative synthesis for concurrent programs. In *Proc. of CAV*, LNCS 6806, pages 243–259. Springer, 2011.
15. P. Cerný, T.A. Henzinger, and A. Radhakrishna. Simulation distances. *Theor. Comput. Sci.*, 413(1):21–35, 2012.

16. A. Chakrabarti, L. de Alfaro, T.A. Henzinger, and M. Stoelinga. Resource interfaces. In *Proc. of EMSOFT*, LNCS 2855, pages 117–133. Springer, 2003.
17. K. Chatterjee and L. Doyen. Energy parity games. In *Proc. of ICALP*, LNCS 6199, pages 599–610. Springer, 2010.
18. K. Chatterjee and L. Doyen. Games and markov decision processes with mean-payoff parity and energy parity objectives. In *Proc. of MEMICS*, LNCS. Springer, 2011.
19. K. Chatterjee, L. Doyen, and T.A. Henzinger. Quantitative languages. *ACM Trans. Comput. Log.*, 11(4), 2010.
20. K. Chatterjee, L. Doyen, T.A. Henzinger, and J.-F. Raskin. Generalized mean-payoff and energy games. In *Proc. of FSTTCS*, LIPIcs 8, pages 505–516. Schloss Dagstuhl - LZI, 2010.
21. K. Chatterjee, L. Doyen, M. Randour, and J.-F. Raskin. Looking at mean-payoff and total-payoff through windows. In *Proc. of ATVA*, LNCS 8172, pages 118–132. Springer, 2013.
22. K. Chatterjee, T.A. Henzinger, and M. Jurdzinski. Mean-payoff parity games. In *Proc. of LICS*, pages 178–187. IEEE Computer Society, 2005.
23. K. Chatterjee, M. Randour, and J.-F. Raskin. Strategy synthesis for multi-dimensional quantitative objectives. In *Proc. of CONCUR*, LNCS 7454, pages 115–131. Springer, 2012.
24. K. Chatterjee, M. Randour, and J.-F. Raskin. Strategy synthesis for multi-dimensional quantitative objectives. *Acta Informatica*, 51(3-4):129–163, 2014.
25. A. Church. Logic, arithmetic, and automata. In *Proceedings of the International Congress of Mathematicians*, pages 23–35. Institut Mittag-Leffler, 1962.
26. L. de Alfaro and T.A. Henzinger. Interface theories for component-based design. In *Proc. of EMSOFT*, LNCS 2211, pages 148–165. Springer, 2001.
27. M. De Wulf, L. Doyen, T.A. Henzinger, and J.-F. Raskin. Antichains: A new algorithm for checking universality of finite automata. In *Proc. of CAV*, LNCS 4144, pages 17–30. Springer, 2006.
28. L. Doyen and J.-F. Raskin. Antichains algorithms for finite automata. In *Proc. of TACAS*, LNCS 6015, pages 2–22. Springer-Verlag, 2010.
29. L. Doyen and J.-F. Raskin. Games with imperfect information: Theory and algorithms. In *Lectures in Game Theory for Computer Scientists*, pages 185–212. 2011.
30. A. Ehrenfeucht and J. Mycielski. Positional strategies for mean payoff games. *Int. Journal of Game Theory*, 8(2):109–113, 1979.
31. E.A. Emerson and C. Jutla. The complexity of tree automata and logics of programs. In *Proc. of FOCS*, pages 328–337. IEEE, 1988.
32. E.A. Emerson and C. Jutla. Tree automata, mu-calculus and determinacy. In *Proc. of FOCS*, pages 368–377. IEEE, 1991.
33. U. Fahrenberg, L. Juhl, K.G. Larsen, and J. Srba. Energy games in multiweighted automata. In *Proc. of ICTAC*, LNCS 6916, pages 95–115. Springer, 2011.
34. E. Grädel, W. Thomas, and T. Wilke, editors. *Automata, Logics, and Infinite Games: A Guide to Current Research*, LNCS 2500. Springer, 2002.
35. Y. Gurevich and L. Harrington. Trees, automata, and games. In *Proc. of STOC*, pages 60–65. ACM, 1982.
36. T.A. Henzinger, O. Kupferman, and S. Rajamani. Fair simulation. *Information and Computation*, 173(1):64–81, 2002.
37. D.A. Martin. Borel determinacy. *Annals of Mathematics*, 102(2):363–371, 1975.
38. D.A. Martin. The determinacy of Blackwell games. *The Journal of Symbolic Logic*, 63(4):1565–1581, 1998.
39. A. Pnueli. The temporal logic of programs. In *Proc. of FOCS*, pages 46–57. IEEE Computer Society, 1977.
40. A. Pnueli and R. Rosner. On the synthesis of a reactive module. In *Proc. of POPL*, pages 179–190, 1989.
41. C. Rackoff. The covering and boundedness problems for vector addition systems. *Theor. Comput. Sci.*, 6:223–231, 1978.
42. P.J. Ramadge and W.M. Wonham. Supervisory control of a class of discrete-event processes. *SIAM Journal of Control and Optimization*, 25(1):206–230, 1987.
43. L.E. Rosier and H.-C. Yen. A multiparameter analysis of the boundedness problem for vector addition systems. *J. Comput. Syst. Sci.*, 32(1):105–135, 1986.
44. E. Solan. Continuity of the value of competitive Markov decision processes. *Journal of Theoretical Probability*, 16(4):831–845, 2003.
45. W. Thomas. Languages, automata, and logic. In *Handbook of Formal Languages*, volume 3, Beyond Words, chapter 7, pages 389–455. Springer, 1997.
46. M.Y. Vardi. Automatic verification of probabilistic concurrent finite-state programs. In *Proc. of FOCS*, pages 327–338. IEEE Computer Society, 1985.
47. Y. Velner and A. Rabinovich. Church synthesis problem for noisy input. In *Proc. of FOSSACS*, LNCS 6604, pages 275–289. Springer, 2011.
48. W. Zielonka. Infinite games on finitely coloured graphs with applications to automata on infinite trees. In *Theoretical Computer Science*, volume 200(1-2), pages 135–183, 1998.
49. U. Zwick and M. Paterson. The complexity of mean payoff games on graphs. *Theoretical Computer Science*, 158:343–359, 1996.