MDPs with Energy-Parity Objectives

Richard Mayr[*], Sven Schewe[†], Patrick Totzke[*], Dominik Wojtczak[†]

[*]University of Edinburgh, UK
[†]University of Liverpool, UK

Abstract—Energy-parity objectives combine ω-regular with quantitative objectives of reward MDPs. The controller needs to avoid to run out of energy while satisfying a parity objective.

We refute the common belief that, if an energy-parity objective holds almost-surely, then this can be realised by some finite memory strategy. We provide a surprisingly simple counterexample that only uses coBüchi conditions.

We introduce the new class of bounded (energy) storage objectives that, when combined with parity objectives, preserve the finite memory property. Based on these, we show that almost-sure and limit-sure energy-parity objectives, as well as almost-sure and limit-sure storage parity objectives, are in NP ∩ coNP and can be solved in pseudo-polynomial time for energy-parity MDPs.

I. Introduction

Context. Markov decision processes (MDPs) are a standard model for dynamic systems that exhibit both stochastic and controlled behaviour [1]. Such a process starts in an initial state and makes a sequence of transitions between states. Depending on the type of the current state, either the controller gets to choose an enabled transition (or a distribution over transitions), or the next transition is chosen randomly according to a predefined distribution. By fixing a strategy for the controller, one obtains a Markov chain. The goal of the controller is to optimise the (expected) value of some objective function on runs of such an induced Markov chain.

Our Focus and Motivation. In this paper we study MDPs with a finite number of states, where numeric rewards (which can be negative) are assigned to transitions. We consider quantitative objectives, e.g. the total expected reward or the limit-average expected reward [1], [2]. Note that the total reward is not bounded *a priori*. We also consider ω-regular objectives that can be expressed by parity conditions on the sequence of visited states (subsuming many simpler objectives like Büchi and coBüchi).

When reasoning about controllers for mechanical and electrical systems, one may need to consider quantitative objectives such as the remaining stored energy of the system (which must not fall below zero, or else the system fails), and, at the same time, parity objectives that describe the correct behaviour based on the temporal specification. Thus one needs to study the combined *energy-parity* objective.

Status Quo. Previous work in [2] (Sec. 3) considered the decidability and complexity of the question whether the energy-parity objective can be satisfied *almost-surely*, i.e. whether there exists a strategy (or: a controller) that satisfies the objective with probability 1. They first show that in the restricted case of energy-Büchi objectives, finite memory optimal strategies exist, and that almost-sure satisfiability is in NP ∩ coNP and can be solved in pseudo-polynomial time.

They then describe a direct reduction from almost-sure energy-parity to almost-sure energy-Büchi. This reduction claimed that the winning strategy could be chosen among a certain subclass of strategies, that we call *colour-committing*. Such a strategy eventually commits to a particular winning even colour, where this colour must be seen infinitely often almost-surely and no smaller colour must ever been seen after committing.

However, this reduction from almost-sure energy-parity to almost-sure energy-Büchi in [2] (Sec. 3) contains a subtle error (which also appears in the survey in [3] (Theorem 4)). In fact, we show that strategies for almost-sure energy-parity may require infinite memory.

Our contributions can be summarised as follows.

1) We provide a simple counterexample that shows that, even for almost-sure energy-coBüchi objectives, the winning strategy requires infinite memory and cannot be chosen among the colour-committing strategies.

2) We introduce an energy *storage* objective, which requires that the energy objective is met using a finite energy store. The size of the store can be fixed by the controller, but it cannot be changed. We argue that the almost-sure winning sets for energy-Büchi and storage-Büchi objectives coincide. Moreover, we show that the reduction in [2] actually works for storage-parity instead of for energy-parity conditions. I.e. [2] shows that almost-sure storage parity objectives require just finite memory, are in NP ∩ coNP, and can be solved in pseudo-polynomial time.

3) We develop a solution for the original almost-sure energy-parity objective. It requires a more involved argument and infinite-memory strategies that are obtained by composing three other strategies. We show that almost-sure energy-parity objectives are in NP ∩ coNP and can be solved in pseudo-polynomial time.

4) We then study the *limit-sure problem*. Here one asks whether, for every $\epsilon > 0$, there exists a strategy that satisfies the objective with probability $\geq 1-\epsilon$. This is a question about the existence of a family of ϵ-optimal strategies, not about a single strategy as in the almost-sure problem. The limit-sure problem is equivalent to the question whether the *value* of a given state and initial energy level (w.r.t. the objective) is 1. For the *storage-parity* objective, the limit-sure condition coincides with the almost-sure condition, and thus the

complexity results from [2] apply. In contrast, for the *energy-parity* objective the limit-sure condition does *not* coincide with the almost-sure condition. While almost-sure energy-parity implies limit-sure energy-parity, we give examples that show that the reverse implication does *not* hold. We develop an algorithm to decide the limit-sure energy-parity objective and show that the problem is in NP ∩ coNP and can be solved in pseudo-polynomial time. Moreover, each member in the family of ϵ-optimal strategies that witnesses the limit-sure energy-parity condition can be chosen as a finite-memory strategy (unlike winning strategies for almost-sure energy-parity that may require infinite memory).

Related work. Energy games were introduced in [4] to reason about systems with multiple components and bounded resources. Energy objectives were later also considered in the context of timed systems [5], synthesis of robust systems [6], and gambling [7] (where they translate to "not going bankrupt"). The first analysis of a combined qualitative–quantitative objective was done in [8] for mean-payoff parity games. Almost-sure winning in energy-parity MDPs was considered in [2] (cf. [3] as a survey). However, it was shown in [2] that almost-sure winning in energy-parity MDPs is at least as hard as two player energy-parity games [9]. A recent paper [10] considers a different combined objective: maximising the expected mean-payoff while satisfying the energy objective. The proof of Lemma 3 in [10] uses a reduction to the (generally incorrect) result on energy-parity MDPs of [2], but their result still holds because it only uses the correct part about energy-Büchi MDPs.

Closely related to energy MDPs and games are one-counter MDPs and games, where the counter value can be seen as the current energy level. One-counter MDPs and games with a termination objective (i.e. reaching counter value 0) were studied in [11] and [12], respectively.

Outline of the Paper. The following section introduces the necessary notations. Section III discusses combined energy-parity objectives and formally states our results. In Section IV we explain the error in the construction of [2] (Sec. 3), define the bounded energy storage condition and derive the results on combined storage-parity objectives. Section V discusses the almost-sure problem for energy-parity conditions and provides our new proof of their decidability. The limit-sure problem for energy-parity is discussed in Sections VI and VII. Section VIII discusses lower bounds and the relation between almost/limit-sure parity MDPs and mean-payoff games. Due to the space constraints some details had to be omitted and can be found in the full version [13].

II. NOTATIONS

A probability distribution over a set X is a function $f : X \to [0, 1]$ such that $\sum_{x \in X} f(x) = 1$. We write $\mathcal{D}(X)$ for the set of distributions over X.

Markov Chains. A *Markov chain* is an edge-labeled, directed graph $\mathcal{C} \stackrel{\text{def}}{=} (V, E, \lambda)$, where the elements of V are called *states*, such that the labelling $\lambda : E \to [0, 1]$ provides a

probability distribution over the set of outgoing transitions of any state $s \in V$. A *path* is a finite or infinite sequence $\rho \stackrel{\text{def}}{=} s_1 s_2 \ldots$ of states such that $(s_i, s_{i+1}) \in E$ holds for all indices i; an infinite path is called a *run*. We use $w \in V^*$ to denote a finite path. We write $s \xrightarrow{x} t$ instead of $(s, t) \in E \wedge \lambda(s, t) = x$ and omit superscripts whenever clear from the context. We write $Runs_w^{\mathcal{C}}$ for the cone set wV^{ω}, i.e., the set of runs with finite prefix $w \in V^*$, and assign to it the probability space $(Runs_w^{\mathcal{C}}, \mathcal{F}_w^{\mathcal{C}}, \mathcal{P}_w^{\mathcal{C}})$, where $\mathcal{F}_w^{\mathcal{C}}$ is the σ-algebra generated by all cone sets $Runs_{wx}^{\mathcal{C}} \subseteq Runs_w^{\mathcal{C}}$, for $x = x_1 x_2 \ldots x_l \in V^*$. The probability measure $\mathcal{P}_w^{\mathcal{C}} : \mathcal{F}_w^{\mathcal{C}} \to [0, 1]$ is defined as $\mathcal{P}_w^{\mathcal{C}}(Runs_{wx}) \stackrel{\text{def}}{=} \Pi_{i=1}^{l-1} \lambda(x_i, x_{i+1})$ for cone sets. By Carathéodory's extension theorem [14], this defines a unique probability measure on all measurable subsets of runs.

Markov Decision Processes. A *Markov Decision Process (MDP)* is a sinkless directed graph $\mathcal{M} \stackrel{\text{def}}{=} (V_C, V_P, E, \lambda)$, where V is a set of states, partitioned as $V \stackrel{\text{def}}{=} V_C \uplus V_P$ into *controlled* (V_C) and *probabilistic* states (V_P). The set of *edges* is $E \subseteq V \times V$ and $\lambda : V_P \to \mathcal{D}(E)$ assigns each probabilistic state a probability distribution over its outgoing edges.

A *strategy* is a function $\sigma : V^* V_C \to \mathcal{D}(E)$ that assigns each word $ws \in V^* V_C$ a probability distribution over the outgoing edges of s, that is $\sigma(ws)(e) > 0$ implies $e = (s, t) \in E$ for some $t \in V$. A strategy is called *memoryless* if $\sigma(xs) = \sigma(ys)$ for all $x, y \in V^*$ and $s \in V_C$, and *deterministic* if if $\sigma(w)$ is Dirac for all $w \in V^* V_C$. Each strategy induces a Markov chain $\mathcal{M}(\sigma)$ with states V^* and where $ws \xrightarrow{x} wst$ if $(s, t) \in E$ and either $s \in V_P \wedge x = \lambda(s, t)$ or $s \in V_C \wedge x = \sigma(ws, t)$. We write $Runs_w^{\mathcal{M}}$ for the set of runs in $\mathcal{M}$ (with prefix w), consisting of all runs in $Runs_w^{\mathcal{M}(\sigma)}$ for some strategy σ, and $Runs^{\mathcal{M}}$ for the set of all such paths.

Objective Functions. An *objective* is a subset $\mathsf{Obj} \subseteq Runs^{\mathcal{M}}$. We write $\overline{\mathsf{Obj}} \stackrel{\text{def}}{=} Runs^{\mathcal{M}} \setminus \mathsf{Obj}$ for its complement. It is satisfied *surely* if there is a strategy σ such that $Runs^{\mathcal{M}(\sigma)} \subseteq \mathsf{Obj}$, *almost-surely* if there exists a σ such that $\mathcal{P}^{\mathcal{M}(\sigma)}(\mathsf{Obj}) = 1$ and *limit-surely* if $\sup_{\sigma} \mathcal{P}^{\mathcal{M}(\sigma)}(\mathsf{Obj}) = 1$. In other words, the limit-sure condition asks that there exists some infinite sequence $\sigma_1, \sigma_2, \ldots$ of strategies such that $\lim_{n \to \infty} \mathcal{P}^{\mathcal{M}(\sigma_n)}(\mathsf{Obj}) = 1$. We call a strategy ε-*optimal* if $\mathcal{P}^{\mathcal{M}(\sigma)}(\mathsf{Obj}) \geq 1 - \varepsilon$. Relative to a given MDP $\mathcal{M}$ and some finite path w, we define the *value* of Obj as $\mathsf{Val}_w^{\mathcal{M}}(\mathsf{Obj}) \stackrel{\text{def}}{=} \sup_{\sigma} \mathcal{P}_w^{\mathcal{M}(\sigma)}(\mathsf{Obj})$. We use the following objectives, defined by conditions on individual runs.

A *reachability condition* is defined by a set of target states $T \subseteq V$. A run $s_0 s_1 \ldots$ satisfies the reachability condition iff there exists an $i \in \mathbb{N}$ s.t. $s_i \in T$. We write $\Diamond T \subseteq Runs$ for the set of runs that satisfy the reachability condition.

A *parity condition* is given by a function $parity : V \to \mathbb{N}$, that assigns a priority (non-negative integer) to each state. A run $\rho \in Runs$ satisfies the parity condition if the minimal priority that appears infinitely often on the run is even. The *parity objective* is the subset $\mathsf{PAR} \subseteq Runs$ of runs that satisfy the parity condition.

Energy conditions are given by a function $cost : E \to \mathbb{Z}$, that assigns a *cost* value to each edge. For a given initial energy

value $k \in \mathbb{N}$, a run $s_0 s_1 \ldots$ satisfies the k-energy condition if, for every finite prefix, the energy level $k + \sum_{i=0}^{l} cost(s_i, s_{i+1})$ stays greater or equal to 0. Let $\mathsf{EN}(k) \subseteq Runs$ denote the k-energy objective, consisting of those runs that satisfy the k-energy condition.

Mean-payoff conditions are defined w.r.t. the same cost function $cost : E \to \mathbb{Z}$ as the energy conditions. A run $s_0 s_1 \ldots$ satisfies the *positive mean-payoff condition* iff $\liminf_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} cost(s_i, s_{i+1}) > 0$. We write $\mathsf{PosMP} \subseteq Runs$ for the positive mean-payoff objective, consisting of those runs that satisfy the positive mean-payoff condition.

III. PARITY CONDITIONS UNDER ENERGY CONSTRAINTS

We study the combination of energy and parity objectives for finite MDPs. That is, given a MDP and both cost and parity functions, we consider objectives of the form $\mathsf{EN}(k) \cap \mathsf{PAR}$ for integers $k \in \mathbb{N}$. We are interested in identifying those control states and values $k \in \mathbb{N}$ for which the combined k-energy-parity objective is satisfied almost-surely and limit-surely, respectively.

Example 1. Consider a controlled state s that can go left or right with cost 0, or stay with cost 1. The probabilistic state on the left increases or decreases energy with equal chance, whereas the probabilistic state on the right has a positive energy updrift. State s has priority 1, all other states have priority 0.
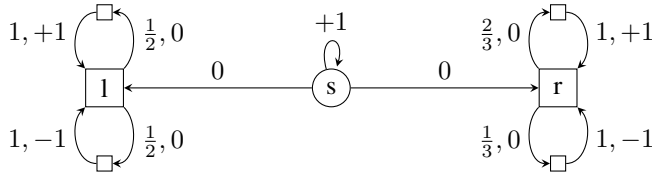
From states other than s there is only one strategy. It holds that $\mathsf{Val}_l(\mathsf{PAR}) = 1$ but $\mathsf{Val}_l(\mathsf{EN}(k)) = 0$ for any $k \in \mathbb{N}$ and so $\mathsf{Val}_l(\mathsf{EN}(k) \cap \mathsf{PAR}) = 0$. For state r we have that $\mathsf{Val}_r(\mathsf{EN}(k) \cap \mathsf{PAR}) = \mathsf{Val}_r(\mathsf{EN}(k)) = 1 - (1/2)^k$, due to the positive drift. For all $k \in \mathbb{N}$ the state s does not satisfy the k-energy-parity objective almost-surely but limit-surely: $\mathsf{Val}_s(\mathsf{EN}(k) \cap \mathsf{PAR}) = 1$ (by going ever higher and then right).

Notice that these energy-parity objectives are trivially monotone in the parameter k because $\mathsf{EN}(k) \subseteq \mathsf{EN}(k+1)$ holds for all $k \in \mathbb{N}$. Consequently, for every fixed state p, if there exists some $k \in N$ such that the k-energy-parity objective holds almost-surely (resp. limit-surely), then there is a minimal such value k. By *solving* the almost-sure/limit-sure problems for these monotone objectives we mean to compute these minimal sufficient values for all initial states.

We now state our two main technical results. We fix a finite MDP $\mathcal{M} \overset{\text{def}}{=} (V_C, V_P, E, \lambda)$, a parity function $parity : V \to \mathbb{N}$ with maximal colour $d \in \mathbb{N}$ and a cost-function $cost : E \to \mathbb{Z}$ with maximal absolute value $W \overset{\text{def}}{=} \max_{e \in E} |cost(e)|$. Let $|\lambda|$ and $|cost|$ be the size of the transition table λ and the cost function $cost$, written as tables with valuations in binary. We

use $\widetilde{\mathcal{O}}(f(n))$ as a shorthand for $\mathcal{O}(f(n) \log^k f(n))$ for some constant k.

Theorem 2. *(1) Almost-sure optimal strategies for k-energy-parity objectives may require infinite memory. (2) The almost-sure problem for k-energy-parity objectives is in* $\mathsf{NP} \cap \mathsf{coNP}$ *and can be solved in pseudo-polynomial time* $\widetilde{\mathcal{O}}(d \cdot |V|^{4.5} \cdot (|\lambda| + |cost|)^2 + |E| \cdot d \cdot |V|^5 \cdot W)$.

Theorem 3. *(1) The limit-sure problem for k-energy-parity objectives is in* $\mathsf{NP} \cap \mathsf{coNP}$ *and can be solved in pseudo-polynomial time* $\widetilde{\mathcal{O}}(d \cdot |V|^{4.5} \cdot (|\lambda| + |cost|)^2 + |E| \cdot d \cdot |V|^5 \cdot W)$. *(2) If the k-energy-parity objective holds limit-surely then, for each $\varepsilon > 0$, there exists a* finite memory ε-optimal strategy.

Remark 4. The claimed algorithms are *pseudo polynomial* in the sense that they depend (linearly) on the value W. If the cost-deltas are $-1, 0$, or 1 only, and not arbitrary binary encoded numbers, this provides a polynomial time algorithm.

Part (2) of Theorem 2 was already claimed in [2], Theorem 1. However, the proof there relies on a particular finiteness assumption that is not true in general. In the next section we discuss this subtle error and describe the class of (bounded) *storage* objectives, for which this assumption holds and the original proof goes through. Our new proof of Theorem 2 is presented in Section V.

The proof of Theorem 3 is deferred to Sections VI and VII. It is based on a reduction to checking almost-sure satisfiability of storage-parity objectives, which can be done in pseudo polynomial time (cf. Theorem 8). We first establish in Section VI that certain *limit values* are computable for each state. In Section VII we then provide the actual reduction, which is based on precomputing these limit values and produces an MDP which is only linearly larger and has no new priorities.

IV. ENERGY STORAGE CONSTRAINTS

The argument of [2] to show computability of almost-sure energy-parity objectives relies on the claim that the controller, if it has a winning strategy, can eventually commit to visiting an even colour infinitely often and *never* visiting smaller colours. We show that this claim already fails for coBüchi conditions (i.e. for MDPs that only use colours 1 and 2). We then identify a stronger kind of energy condition—the storage energy condition we introduce below—that satisfies the above claim and for which the original proof of [2] goes through.

Let us call a strategy *colour-committing* if, for some colour $2i$, almost all runs eventually visit a position such that *almost all* possible continuations visit colour $2i$ infinitely often and *no continuation* (as this is a safety constraint) visits a colour smaller than $2i$.

Claim 5. If there exists some strategy that almost-surely satisfies $\mathsf{EN}(k) \cap \mathsf{PAR}$ then there is also a colour-committing strategy that does.

Proof (that Claim 5 is false). Consider the following example, where the controller owns states A, B, C and tries to avoid visiting state B infinitely often while maintaining the

energy condition. This can be expressed as an energy-parity condition where $parity(A) = parity(C) = parity(D) = 2$ and $parity(B) = 1$.
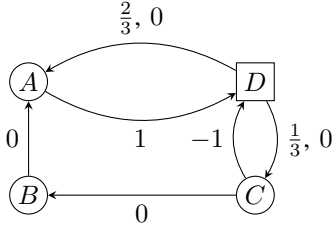
First notice that all states almost-surely satisfy the 0-energy-coBüchi condition $\mathsf{EN}(0) \cap \mathsf{PAR}$. One winning strategy is to chooses the edge $C \xrightarrow{-1} D$ over $C \xrightarrow{0} B$, unless the energy level is 0, in which case $C \xrightarrow{0} B$ is favoured over $C \xrightarrow{-1} D$. This strategy is not colour-committing but clearly energy safe: the only decreasing step is avoided if the energy level is 0.

To see why this strategy also almost-surely satisfies the parity (in this case coBüchi) objective, first observe that it guarantees a positive updrift: from state D with positive energy level, the play returns to D in two steps with expected energy gain $+1/3$, from state D with energy level 0, the play returns to D in either two or three steps, in both cases with energy gain $+1$. The chance to visit state C with energy level 0 when starting at state D with energy level $k \in \mathbb{N}$ is $(1/2)^{k+1}$. This is the same likelihood with which state B is eventually visited. However, every time state B is visited, the system restarts from state D with energy level 1. Therefore, the chance of re-visiting B from B is merely $1/4$. More generally, the chance of seeing state B at least n further times is $(1/4)^n$. The chance of visiting B infinitely often is therefore $\lim_{n\to\infty}(1/4)^n = 0$. This strategy thus satisfies the parity—in this case coBüchi—objective almost-surely. Consequently, the combined 0-energy-parity objective is almost-surely met from all states.

To contradict Claim 5, we contradict the existence of an initial state and a colour-committing strategy that almost-surely satisfies the 0-energy-parity objective. By definition, such a strategy will, on almost all runs, eventually avoid state B completely.

As every run will surely visit state D infinitely often, we can w.l.o.g. pick a finite possible prefix $s_1 s_2 \ldots s_j$ (i.e. a prefix that can occur with a likelihood $\delta > 0$) of a run that ends in state $s_j = D$ and assume that none (or only a 0 set, but these two conditions coincide for safety objectives) of its continuations visits state B again. Let $l \stackrel{\text{def}}{=} \sum_{i=1}^{j} cost(s_i, s_{i+1})$ denote the sum of rewards collected on this prefix. Note that there is a $(1/3)^{l+1} > 0$ chance that some continuation alternates between states D and C for $l + 1$ times and thus violates the l-energy condition. Consequently, the chance of violating the 0-energy parity condition from the initial state is at least $\delta \cdot (1/2)^{l+1} > 0$. □

Notice that every finite memory winning strategy for the PAR objective must also be colour-committing. The system

above therefore also proves part (1) of Theorem 2, that infinite memory is required for k-energy-parity objectives.

In the rest of this section we consider a stronger kind of energy condition, for which Claim 5 hold and the original proof of [2] goes through. The requirement is that the strategy achieves the energy condition without being able to store an infinite amount of energy. Instead, it has a finite energy store, say s, and cannot store more energy than the size of this storage. Thus, when a transition would lead to an energy level $s' > s$, then it would result in an available energy of s. These are typical behaviours of realistic energy stores, e.g. a rechargeable battery or a storage lake. An alternative view (and a consequence) is that the representation of the system becomes finite-state once the bound s is fixed, and only finite memory is needed to remember the current energy level.

For the definition of a storage objective, we keep the infinite storage capacity, but instead require that no subsequence loses more than s energy units. The definitions are interchangeable, and we chose this one in order not to change the transitions of the system.

Definition 6. For a finite MDP with associated cost function, a run $s_0 s_1 \ldots$ satisfies the *s-storage condition* if, for every infix $s_l s_{l+1} \ldots s_u$, it holds that $s + \sum_{i=l}^{u+1} cost(s_i, s_{i+1}) \geq 0$. Let $\mathsf{ST}(k, s) \subseteq Runs$ denote the k-energy s-storage objective, consisting of those runs that satisfy both the k-energy and the s-storage condition.

Example 7. The two parameters can sometimes be traded against each other, as shown in the following example.
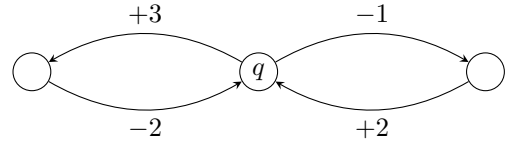
From state q in the middle, one can win with an initial energy level 0 by always going left, provided that one has an energy store of size at least 2. With an energy store of size 1, however, going left is not an option, as one would not be able to return from the state on the left. But with an initial energy level of 1, one can follow the strategy to always turn to the right. So the $\mathsf{ST}(0, 2)$ and $\mathsf{ST}(1, 1)$ objectives hold almost-surely but the $\mathsf{ST}(0, 1)$ objective does not.

We sometimes want to leave the size of the energy store open. For this purpose, we define $\mathsf{ST}(k)$ as the objective that says "there is an s, such that $\mathsf{ST}(k, s)$ holds" and ST for "there is an s such that $\mathsf{ST}(s, s)$ holds". Note that this is not a path property; we rather require that the s is fixed globally. In order to meet an $\mathsf{ST}(k)$ property almost-surely, there must be a strategy σ and an $s \in \mathbb{N}$ such that almost all runs satisfy $\mathsf{ST}(k, s)$: $\exists \sigma, s$ s.t. $\mathcal{P}^{\mathcal{M}(\sigma)}(\mathsf{ST}(k, s)) = 1$. Likewise, for limit-sure satisfaction of ST, we require $\exists s \, \forall \varepsilon > 0 \, \exists \sigma$ s.t. $\mathcal{P}^{\mathcal{M}(\sigma)}(\mathsf{ST}(s, s)) \geq 1 - \varepsilon$.

We now look at combined storage-parity and storage-mean-payoff objectives.

Theorem 8 (Storage-Parity)**.** *For finite MDPs with states V, edges E and associated cost and parity functions, with maximal absolute cost W and maximal colour $d \in \mathbb{N}$:*

1) *The almost-sure problem for storage-parity objectives is in* NP ∩ coNP*, and there is an algorithm to solve it in $\mathcal{O}(|E| \cdot d \cdot |V|^4 \cdot W)$ deterministic time.*
2) *Memory of size $\mathcal{O}(|V| \cdot W)$ is sufficient for almost-sure winning strategies. This also bounds the minimal values $k, s \in \mathbb{N}$ such that $\mathsf{ST}(k,s) \cap \mathsf{PAR}$ holds almost-surely.*

The proof is provided by Chatterjee and Doyen [2]: they first show the claim for energy-Büchi objectives $\mathsf{EN}(k) \cap \mathsf{PAR}$ (where $d = 1$) by reduction to two-player energy-Büchi *games* ([2], Lemma 2). Therefore, almost-sure winning strategies come from first-cycle games and operate in a bounded energy radius. As a result, almost-sure satisfiability for energy-Büchi and storage-Büchi coincide. They then ([2], Lemma 3) provide a reduction for general parity conditions to the Büchi case, assuming Claim 5. Although this fails for energy-parity objectives, as we have shown above, almost-sure winning strategies for storage-parity can be assumed to be finite memory and therefore colour committing. The construction of [2] then goes through without alteration. The complexity bound follows from improvements for energy parity games [9].

Theorem 9 (Storage-Mean-payoff)**.** *For finite MDPs with combined storage and positive mean-payoff objectives:*

1) *The almost-sure problem is in* NP ∩ coNP *and can be solved in $\mathcal{O}(|E| \cdot d \cdot |V|^4 \cdot W)$ deterministic time.*
2) *Memory of size $\mathcal{O}(|V| \cdot W)$ is sufficient for almost-sure winning strategies. This also bounds the minimal value $k, s \in \mathbb{N}$ such that $\mathsf{ST}(k,s) \cap \mathsf{PosMP}$ holds almost-surely.*

Proof. We show that, for every MDP $\mathcal{M}$ with associated *cost* function, there is a linearly larger system $\mathcal{M}'$ with associated $cost'$ and $parity$ function —where the parity function is Büchi, i.e. has image $\{0, 1\}$—that, for every $k \in \mathbb{N}$, $\mathsf{PosMP} \cap \mathsf{ST}(k)$ holds almost-surely in $\mathcal{M}$ iff $\mathsf{PAR} \cap \mathsf{ST}(k)$ holds almost-surely in $\mathcal{M}'$.

For every state q of $\mathcal{M}$, the new system $\mathcal{M}'$ contains two new states, q' and q'', edges (q, q') and (q, q'') with costs 0 and -1, respectively. Each original edge (q, r) is replaced by two edges, (q', r) and (q'', r). All original states become controlled, and the primed and double primed copies of a state q are controlled if, and only if, q was controlled in $\mathcal{M}$. The double primed states have colour 0, while all original and primed states have colour 1. See Figure 2 (on the left) for an illustration.

To give the idea of this construction in a nutshell, the Büchi condition in $\mathcal{M}'$ intuitively sells one energy unit for visiting an accepting state (or: for visiting a state with colour 0, the double primed copy). $\mathsf{ST}(k)$ implies that, as soon as $s + 1$ energy is available, one can sell off one energy unit for a visit of an intermediate accepting state. PosMP implies that this can almost-surely be done infinitely often. Vice-versa, $\mathsf{ST}(k)$ implies non-negative mean payoff. $\mathsf{ST}(k)$ plus Büchi can always be realised with finite memory by Theorem 8 (2),
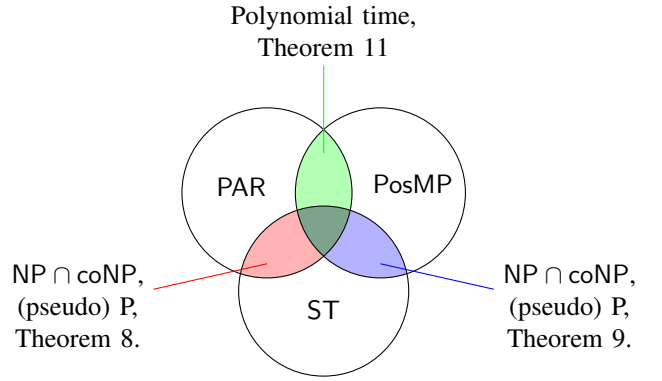

Fig. 1. Results on the almost-sure problems for combined objectives. Positive mean-payoff-parity PosMP ∩ PAR (in green) requires infinite memory while storage-parity (in red) and storage-mean-payoff (in blue) have finite-memory optimal strategies.

and such a strategy then implies that $\mathsf{PosMP} \cap \mathsf{ST}(k)$ holds almost-surely in $\mathcal{M}$. Now the claim holds by Theorem 8. $\qquad\square$

Remark 10. Note that the order of quantification in the limit-sure problems for storage objectives ($\exists s. \forall \varepsilon \ldots$) means that limit-sure and almost-sure winning coincides for storage-parity objectives: if there is an s such that $\mathsf{ST}(s, s) \cap \mathsf{PAR}$ holds limit-surely then one can get rid of the storage condition by hardcoding energy-values up to s into the states. The same is true for mean-payoff-storage objectives. The claims in Theorems 8 and 9 thus also hold for the limit-sure problems.

Finally, we restate the result from [2], Theorem 2 (1) on positive mean-payoff-parity objectives and add to it an explicit computational complexity bound that we will need later.

Theorem 11 (Mean-payoff-Parity)**.** *For finite MDPs with combined positive mean-payoff and parity objectives,*

1) *The almost-sure problem is in* P *and can be solved in $\widetilde{\mathcal{O}}(d \cdot |V|^{3.5} \cdot (|\lambda| + |cost|)^2)$ time.*
2) *Finite memory suffices.*

Proof. The computation complexity bound follows from the analysis of Algorithm 1 in [2]. It executes $d/2$ iterations of a loop, in which Step 3.3 of computing the mean-payoff of maximal end components dominates the cost. This can be formulated as a linear program (LP) that uses two variables, called *gain* and *bias*, for each state [1]. This LP can be solved using Karmarkar's algorithm [15] in time $\widetilde{\mathcal{O}}(|V|^{3.5} \cdot (|\lambda| + |cost|)^2)$. Note that the complexity refers to *all* (not each) maximal end-components.

As we do not need to obtain a maximal payoff $\tau > 0$ but can use any smaller value, like $\tau/2$, finite memory suffices. $\quad\square$

V. ALMOST-SURE ENERGY-PARITY

In this section we prove Theorem 2. Our proof can be explained in terms of the three basic objectives: storage (ST), positive mean-payoff (PosMP), and parity (PAR). It is based on the intuition provided by the counterexample in the

previous section. Namely, in order to almost-surely satisfy the energy-parity condition one needs to combine two strategies:

1) One that guarantees the parity condition and, at the same time, a positive expected mean-payoff. Using this strategy one can achieve the energy-parity objective with some non-zero chance.

2) A *bailout* strategy that guarantees positive expected mean-payoff together with a storage condition. This allows to (almost-surely) set the accumulated energy level to some arbitrarily high value.

We show that, unless there exist some safe strategy that satisfies storage-parity, it is sufficient (and necessary) that such two strategies exist and that the controller can freely switch between them. I.e. they do not leave the *combined* almost-sure winning set unless a state that satisfies storage-parity is reached.

Recall that the combined positive mean-payoff-parity objective (for case 1 above) is independent of an initial energy level and its almost-sure problem is decidable in polynomial time due to Theorem 11. The mean-payoff-storage objective $\mathsf{ST}(k) \cap \mathsf{PosMP}$ (for case 2 above), as well as the storage-parity objective are computable by Theorems 8 and 9, respectively. See Figure 1.

To establish Theorem 2, we encode the almost-sure winning sets of the storage-parity objective directly into the system (Definition 12 and Lemma 13), in order to focus on the two interesting conditions from above. We then show (Definition 14 and Lemma 15) that the existence of the two strategies for bailout and $\mathsf{ST} \cap \mathsf{PosMP}$, and the minimal safe energy levels can be computed in the claimed bounds. In Lemma 16 we show that these values coincide with the minimal energy levels of the energy-parity objective for the original system, which concludes the proof.

Definition 12. For a given MDP $\mathcal{M}$ and associated *cost* and *parity* functions, we define an MDP $\mathcal{M}' \overset{\text{def}}{=} (V_C', V_P', E', \lambda')$ with states $V' \overset{\text{def}}{=} V_C' \uplus V_P'$ as follows. For every state q of $\mathcal{M}$ there are two states, q and q' in V' such that both have the same colour as q in $\mathcal{M}$, every original incoming edge now only goes to q', and every original outgoing edge now only goes from q. Moreover, q' is controlled and has an edge to q with $cost(q', q) = 0$.

Finally, $\mathcal{M}'$ contains a single winning sink state w with colour 0 and a positive self-loop, and every state q' gets an edge to w where the cost of $-k_q$, where $k_q \in \mathbb{N}$ is the minimal value such that, for some $s \in \mathbb{N}$, the storage-parity objective $\mathsf{ST}(k_q, s) \cap \mathsf{PAR}$ holds almost-surely See Figure 2 (on the right) for an illustration.

The next lemma summarises the relevant properties of $\mathcal{M}'$. It follows directly from Definition 12 and the observation that $\mathsf{ST}(k, s) \subseteq \mathsf{EN}(k)$ holds for all $k, s \in \mathbb{N}$.

Lemma 13. *For every MDP $\mathcal{M}$, state q and $k \le s \in \mathbb{N}$,*
1) $\mathsf{EN}(k) \cap \mathsf{PAR}$ *holds almost-surely in $\mathcal{M}$ if, and only if, it holds almost-surely in $\mathcal{M}'$.*
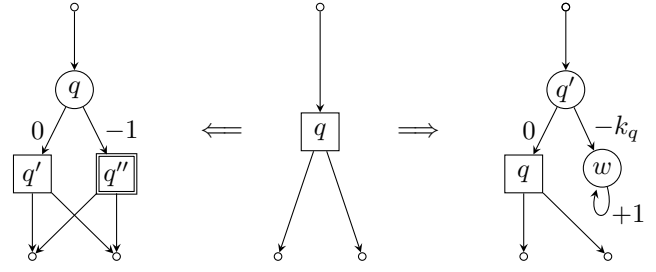

Fig. 2. The reduction from Theorem 9 (middle to left) makes sure that controller can always trade energy values for visiting an accepting state. The reduction from Definition 12 (middle to right) makes sure that in each state q the controller can interrupt and go to a winning sink if the accumulated energy level exceeds the minimal value k_q necessary to almost-surely win the storage-parity objective from q.

2) *If* $\mathsf{ST}(k, s) \cap \mathsf{PAR}$ *holds almost-surely in $\mathcal{M}$ then* $\mathsf{ST}(k, s) \cap \mathsf{PAR} \cap \mathsf{PosMP}$ *holds almost-surely in $\mathcal{M}'$.*

For a set $S \subseteq V'$ of states, we write $\mathcal{M}'|S$ for the restriction of $\mathcal{M}'$ to states in S, i.e. the result of removing all states not in S and their respective edges.

Definition 14. We define $R \subseteq V'$ as the largest set of states such that, in $\mathcal{M}'|R$, every state

1) almost-surely satisfies the $\mathsf{PosMP} \cap \mathsf{PAR}$ objective, and
2) almost-surely satisfies the $\mathsf{PosMP} \cap \mathsf{ST}$ objective.

For every state $q \in V'$ let $safe(q) \in \mathbb{N}$ be the minimal number k such that $\mathsf{PosMP} \cap \mathsf{ST}(k)$ holds almost-surely in $\mathcal{M}'|R$ and $safe(q) \overset{\text{def}}{=} \infty$ if no such number exists (in case $q \notin R$).

The relevance of these numbers for us is, intuitively, that if $safe(q)$ is finite, then there exists a pair of strategies, one for the $\mathsf{PosMP} \cap \mathsf{PAR}$ and one for the $\mathsf{PosMP} \cap \mathsf{ST}(k)$ objective, between which the controller can switch as often as she wants.

Lemma 15. *For a given $\mathcal{M}'$, the values $safe(q)$ are either ∞ or bounded by $\mathcal{O}(|V| \cdot W)$, computable in pseudo-polynomial time $\widetilde{\mathcal{O}}(d \cdot |V|^{4.5} \cdot (|\lambda| + |cost|)^2 + |E| \cdot d \cdot |V|^5 \cdot W)$, and verifiable in $\mathsf{NP} \cap \mathsf{coNP}$.*

Proof. Finite values $safe(q) \in \mathbb{N}$ are clearly bounded by minimal sufficient values for almost-sure satisfiability of $\mathsf{ST}(k) \cap \mathsf{PosMP}$ in $\mathcal{M}'$. Therefore, the claimed bound holds by definition of $\mathcal{M}'$ and Theorems 8 and 9.

The set R is in fact the result of a refinement procedure that starts with all states of $\mathcal{M}'$. In each round, it removes states that fail either of the two conditions. For every projection $\mathcal{M}'|S$, checking Condition 1 takes $\widetilde{\mathcal{O}}(d \cdot |V|^{3.5} \cdot (|\lambda| + |cost|)^2)$ time by Theorem 11 and Condition 2 can be checked in $\mathcal{O}(|E| \cdot d \cdot |V|^4 \cdot W)$ time by Theorem 9. All in all, this provides a pseudo-polynomial time algorithm to compute R. By another application of Theorem 9, we can compute the (pseudo-polynomially bounded) values $safe(q)$. In order to verify candidates for values $safe(q)$ in NP, and also coNP, one can guess a witness, the sequence of sets $R_0 \supset R_1 \supset \ldots \supset R_j = R$, together with certificates for all $i \le j$ that R_{i+1} is the correct set following R_i in the refinement procedure. This

can be checked all at once by considering the disjoint union of all $\mathcal{M}'|R_i$. □

Lemma 16. *For every $k \in \mathbb{N}$ and state q in $\mathcal{M}'$, the energy-parity objective $\mathsf{EN}(k) \cap \mathsf{PAR}$ holds almost-surely from q if, and only if, $\mathrm{safe}(q) \le k$.*

Proof. ($\Longrightarrow$). First observe that the winning sink w in $\mathcal{M}'$ is contained in R, and has $\mathrm{safe}(w) = 0$ since the only strategy from that state satisfies $\mathsf{ST}(0,0) \cap \mathsf{PAR} \cap \mathsf{PosMP}$.

For all other states there are two cases: either there is an $s \in \mathbb{N}$ such that $\mathsf{ST}(k,s) \cap \mathsf{PAR}$ holds almost-surely, or there is no such s. If there is, then the strategy that goes to the sink guarantees the objective $\mathsf{ST}(k,s) \cap \mathsf{PAR} \cap \mathsf{PosMP}$, which implies the claim.

For the second case (there is no s such that $\mathsf{ST}(k,s) \cap \mathsf{PAR}$ holds almost-surely) we see that every almost-surely winning strategy for $\mathsf{EN}(k) \cap \mathsf{PAR}$ must also almost-surely satisfy PosMP. To see this, note that the energy condition implies a non-negative expected mean-payoff, and that an expected mean-payoff of 0 would imply that the storage condition $\mathsf{ST}(k,s)$ is satisfied for some s, which contradicts our assumption. Consequently the $\mathsf{PosMP} \cap \mathsf{PAR}$ objective holds almost-surely.

We now show that the $\mathsf{ST}(k,s) \cap \mathsf{PosMP}$ objective holds almost-surely in state q, where $s > \mathrm{safe}(r)$ for all states r with $\mathrm{safe}(r) < \infty$. We now define a strategy that achieves $\mathsf{ST}(k,s) \cap \mathsf{PosMP}$. For this, we first fix a strategy σ_q that achieves $\mathsf{EN}(h_q) \cap \mathsf{PAR}$ with $h_q = \mathrm{safe}(q)$ for each state q with $\mathrm{safe}(q) < \infty$.

When starting in q, we follow σ_q until one of the following three events happen. We have (1) sufficient energy to move to the winning sink w. In this case we do so. Otherwise, if we (2) have reached a state r and since starting to follow σ_q, the energy balance is strictly greater than[1] $h_r - h_q$. Then we abandon σ_q and follow σ_r as if we were starting the game.

Before we turn to the third event, we observe that, for each strategy σ_q, there is a minimal distance[2] $d_q \in \mathbb{N}$ to (1) or (2) and a positive probability $p_q > 0$ that either event is reached in d_q steps. The third event is now simply that (3) d_q steps have lapsed. When in state r we then also continue with σ_r as if we were starting the game.

It is obvious that no path has negative mean payoff. Moreover, as long as the game does not proceed to the winning sink, a partial run starting at a state q and ending at a state r has energy balance $\ge h_r - h_q$, such that the resulting strategy surely satisfies ST. The expected mean payoff is $\ge p_q/d_q$, and PosMP is obviously satisfied almost-surely. Consequently, $\mathsf{ST}(h_q, s) \cap \mathsf{PosMP}$ holds almost-surely from q.

We conclude that every state for which the $\mathsf{EN}(k) \cap \mathsf{PAR}$ objective holds almost-surely must satisfy both criteria of Definition 14 and thus be a member of R. Since almost-sure

[1]Note that the energy balance can never be strictly smaller than $h_r - h_q$ in such a case, as there would not be a safe continuation from r otherwise.

[2]If there is no point where (2) is met, the energy balance on state r is always exactly $h_r - h_q$, such that σ_q satisfies $\mathsf{ST}(h_q, s)$, and (1) is satisfied immediately.

winning strategies cannot leave the respective winning sets, this means that every winning strategy for the above objective also applies in $\mathcal{M}'|R$ and thus justifies that $\mathrm{safe}(q) \le k$.

($\Longleftarrow$). By definition of R, there are two finite memory strategies σ and β which almost-surely satisfy the $\mathsf{PosMP} \cap \mathsf{PAR}$, and the bailout objective $\mathsf{PosMP} \cap \mathsf{ST}(k)$, respectively, from every state q with $\mathrm{safe}(q) \le k$. Moreover, those strategies will never visit any state outside of R.

We start with the bailout strategy β and run it until the energy level is high enough (see below). We then turn to σ and follow it until (if ever) it *could* happen in the next step that a state q is reached while the energy level falls below $\mathrm{safe}(q)$. We then switch back to β.

The "high enough" can be achieved by collecting enough energy that there is a positive probability that one does not change back from σ to β. For this, we can start with a sufficient energy level e such that σ never hits an energy ≤ 0 with a positive probability[3]. The sum $e + s + W$ consisting of this energy, the sufficient storage level for $\mathsf{PosMP} \cap \mathsf{ST}(k)$, and the maximal change W of the energy level obtained in a single step suffices.

The constructed strategy then almost-surely satisfies the $\mathsf{EN}(k_q) \cap \mathsf{PosMP} \cap \mathsf{PAR}$ objective from every state q and $k_q \stackrel{\text{def}}{=} \mathrm{safe}(q)$. In particular, this ensures that the k-energy-parity objective holds almost-surely from q in $\mathcal{M}'|R$ and therefore also in $\mathcal{M}'$. □

Proof of Theorem 2. (1) The fact that infinite memory is necessary follows from our counterexample to Claim 5, and the observation that every finite memory winning strategy for the PAR objective must also be colour-committing.

For parts (2) and (3), it suffices, by Lemma 13(1) and Lemma 16, to construct $\mathcal{M}'$ and compute the values $\mathrm{safe}(q)$ for every state q of $\mathcal{M}'$. The claims then follow from Lemma 15. □

VI. LIMIT VALUES

Since $\mathsf{EN}(k) \subseteq \mathsf{EN}(k+1)$ holds for all $k \in \mathbb{N}$, the chance of satisfying the k-energy-parity objective depends (in a monotone fashion) on the initial energy level: for every state p we have that $\mathsf{Val}_p^{\mathcal{M}}(\mathsf{EN}(k) \cap \mathsf{PAR}) \le \mathsf{Val}_p^{\mathcal{M}}(\mathsf{EN}(k + 1) \cap \mathsf{PAR})$. We can therefore consider the respective *limit values* as the limits of these values for growing k:

$$\mathsf{LVal}_p^{\mathcal{M}} \quad \stackrel{\text{def}}{=} \quad \sup_k \mathsf{Val}_p^{\mathcal{M}}(\mathsf{EN}(k) \cap \mathsf{PAR}). \tag{1}$$

Note that this is *not* the same as the value of PAR alone. For instance, the state l from Example 1 has limit value $\mathsf{LVal}_l = 0 \ne \mathsf{Val}_l(\mathsf{PAR}) = 1$.

The states r and s from Example 1 have $\mathsf{LVal}_r = 1$ and $\mathsf{LVal}_s = 1$. In fact, for any $\mathcal{M}, w, k$ and parity objective PAR it holds that $\mathsf{Val}_p^{\mathcal{M}}(\mathsf{EN}(k) \cap \mathsf{PAR}) \le \mathsf{LVal}_p^{\mathcal{M}} \le \mathsf{Val}_p^{\mathcal{M}}(\mathsf{PAR})$.

Limit values are an important ingredient in our proof of Theorem 3. This is due to the following property, which directly follows from the definition.

[3]We argue in Lemma 21 that, with sufficient energy, this probability can be moved arbitrarily close to 1.

Lemma 17. *Let $\mathcal{M}$ be an MDP and p be a state with $\mathsf{LVal}_p^{\mathcal{M}} = 1$. Then, for all $\varepsilon > 0$, there exist a $k \in \mathbb{N}$ and a strategy σ such that $\mathcal{P}_p^{\mathcal{M}(\sigma)}(\mathsf{EN}(k) \cap \mathsf{PAR}) \geq 1 - \varepsilon$.*

We now show how to compute limit values, based on the following two sets.

$$A \stackrel{\text{def}}{=} \{p \in Q \mid \exists k \exists \sigma \ \mathcal{P}_p^{\mathcal{M}(\sigma)}(\mathsf{EN}(k) \cap \mathsf{PAR}) = 1\}$$

$$B \stackrel{\text{def}}{=} \{p \in Q \mid \exists \sigma \ \mathcal{P}_p^{\mathcal{M}(\sigma)}(\mathsf{PosMP} \cap \mathsf{PAR}) = 1\}$$

The first set, A, contains those states that satisfy the k-energy-parity condition almost-surely for some energy level $k \in \mathbb{N}$. The second set, B, contains those states that almost-surely satisfy the combined positive mean-payoff-parity objective. Our argument for computability of the limit values is based on the following theorem, which claims that limit values correspond to the values of a reachability objective with target $A \cup B$. Formally,

Theorem 18. *For every MDP $\mathcal{M}$ and state p, $\mathsf{LVal}_p^{\mathcal{M}} = \sup_\sigma \mathcal{P}_p^{\mathcal{M}(\sigma)}(\Diamond(A \cup B))$.*

Before we prove this claim by Lemmas 22 and 25 in the remainder of this section, we remark that we can compute $A \cup B$ without constructing A. Let us consider the set

$$A' \stackrel{\text{def}}{=} \{p \in Q \mid \exists k \exists \sigma \ \mathcal{P}_p^{\mathcal{M}(\sigma)}(\mathsf{ST}(k) \cap \mathsf{PAR}) = 1\} \ ,$$

and observe that $A' \subseteq A$ holds by definition and that the construction of A from Theorem 2 establishes $A \subseteq A' \cup B$. Thus, $A \cup B = A' \cup B$ holds, and it suffices to construct A' and B, which is cheaper than constructing A and B.

We now start with some notation.

Definition 19. For an MDP $\mathcal{M} \stackrel{\text{def}}{=} (V_C, V_P, E, \lambda)$, the *attractor* of a set $X \subseteq V$ of states is the set $Att(X) \stackrel{\text{def}}{=} \{q \mid \exists \sigma \ \mathcal{P}_q^{\mathcal{M}(\sigma)}(\Diamond X) = 1\}$ of states that almost-surely satisfy the reachability objective with target X.

Definition 20. For an MDP $\mathcal{M} \stackrel{\text{def}}{=} (V_C, V_P, E, \lambda)$ an *end-component* is a strongly connected set of states $C \subseteq V$ with the following closure properties:

- for all controlled states $v \in C \cap V_C$, some successor v' of v is in C, and
- for all probabilistic states $v \in C \cap V_P$, all successors v' of v are in C.

Given *cost* and *parity* functions and $i \in \mathbb{N}$, we call an end-component

- i *dominated*, if they contain a state p with $parity(p) = i$, but no state q with $parity(q) < i$,
- i *maximal*, if it is a maximal (w.r.t. set inclusion) i dominated end-component, and
- *positive*, if its expected mean-payoff is strictly greater than 0 (recall that the mean-payoff of all states in a strongly connected set of states of an MDP is equal).

Lemma 21. *The states of each positive $2i$-maximal end-component C are contained in B.*

Proof. (sketch) We consider a strategy σ that follows the optimal (w.r.t. the mean-payoff) strategy most of the time and "moves to" a fixed state p with the minimal even parity 2i only sparsely. Such a strategy keeps the mean-payoff value positive while satisfying the parity condition. We show that σ can be defined to use finite memory or no memory, but randomisation. Either way, σ induces a probabilistic one-counter automata [16], whose probability of ever decreasing the counter by some finite k can be analysed, based on the mean-payoff value, using the results in [17]. □

Lemma 22. *For every MDP $\mathcal{M}$ and state p, $\mathsf{LVal}_p^{\mathcal{M}} \geq \sup_\sigma \mathcal{P}_p^{\mathcal{M}(\sigma)}(\Diamond(A \cup B))$.*

Proof. Assume w.l.o.g. that $\tau \stackrel{\text{def}}{=} \sup_\sigma \mathcal{P}_p^{\mathcal{M}(\sigma)}(\Diamond(A \cup B)) > 0$. We show that $\mathsf{LVal}_p^{\mathcal{M}}$ is at least $\tau - 2\varepsilon$ for all $\varepsilon > 0$ as follows.

We start by choosing $k \in \mathbb{N}$ big enough so that for every state $q \in A \cup B$, some strategy satisfies the k-energy-parity objective with probability $> 1 - \varepsilon$. We then consider a memoryless strategy (e.g. from solving the associated linear program), which guarantees that the set $A \cup B$ is reached with likelihood τ, and then determine a natural number l such that it is reached within l steps with probability $> \tau - \varepsilon$. This reachability strategy σ can now be combined with an ε-optimal strategy for states in $A \cup B$: until a state in $A \cup B$ is reached, the controller plays according to σ and then swaps to a strategy that guarantees the k-energy-parity objective with likelihood $> (1 - \varepsilon)$. Such a strategy exists by our assumption on k. This combined strategy will satisfy the $\mathsf{EN}(k+l)$-energy-parity objective with probability $> (\tau - \varepsilon)(1 - \varepsilon) \geq \tau - 2\varepsilon$. □

Definition 23. [Non-losing end-component] We call an end-component non-losing, iff the smallest priority of states in the end-component is even and there is a strategy that allows to

1) almost-surely stay within this end-component,
2) almost-surely visit all states in the end-component, and
3) satisfy the energy condition from some energy level with non-zero probability.

Lemma 24. *Every non-losing end-component I is contained in $Att(A \cup B)$.*

Proof. We start with a case distinction of the mean-payoff value of I. (Recall that, as an end-component in an MDP, all states of I have the same mean-payoff values.) If this value is positive and $2i$ is the lowest priority in I, then I is contained in some $2i$ maximal end-component and by Lemma 21, also in $B \subseteq Att(A \cup B)$. If this value is negative, then the third condition of Definition 23 cannot be satisfied together with the first two. This leaves the case where the mean-payoff value is 0.

If the mean-payoff value is 0, then there exists a bias function $b : I \to \mathbb{Z}$ that satisfies the following constraints:

- $b(v) = \min \left\{ \mathsf{cost}((v, v')) + b(v') \mid v' \in I \wedge (v, v') \in E \right\}$ holds for all controlled states $v \in I \cap V_C$,
- $b(v) = \sum_{v' \in \{w \in I \mid (v, w) \in E\}} \lambda(v)((v, v')) \cdot \Big(\mathsf{cost}((v, v')) + b(v') \Big)$ for all probabilistic states $v \in I \cap V_P$.

When adjusting b to b' by adding the same constant to all valuations, b' is obviously a bias function, too.

We call a transition (v, v') *invariant* iff $b(v) = \mathrm{cost}((v, v')) + b(v')$ holds. A set $G \subseteq V$ of states invariant if it is strongly connected and contains only controlled states with an invariant transition into G and only probabilistic states with only invariant outgoing transitions, which all go to G. We now make the following case distinction.

Case 1: there is a nonempty, invariant set $G \subseteq I$, such that the state p of G with minimal priority has even priority. First notice that $G \subseteq A$: if the minimal value of the bias function is $b_{\min}$, then the bias of a state in p minus $b_{\min}$ serves as sufficient energy when starting in p: it then holds that $\mathcal{P}_p^{\mathcal{M}(\sigma)}(\mathsf{EN}(k) \cap \mathsf{PAR}) = 1$, where $k \stackrel{\text{def}}{=} b(p) - b_{\min}$, and σ is a memoryless randomised strategy that assigns a positive probability to all transitions into G. Since I is an end-component, it is contained in the attractor of G, which implies the claim, as $Att(G) \subseteq Att(A) \subseteq Att(A \cup B)$.

Case 2: there is no non-empty invariant set $G \subseteq I$ with even minimal priority. We show that this is a contradiction with the assumption that I is a non-losing set, in particular with condition 3 of Definition 23. We assume for contradiction that there is a strategy σ and an energy level k such that we can satisfy the energy parity condition with a positive probability while staying in I and starting at some state $p \in I$. We also assume w.l.o.g. that all bias values are non-negative, and m is the maximal value among them. We set $k' = k + m$.

The 'interesting' events that can happen during a run are selecting a non-invariant transition from a controlled state or reaching a probabilistic state (and making a random decision from this state), where at least one outgoing transition is non-invariant.

We capture both by random variables, where random variables that refer to taking non-invariant transition from controlled states (are deterministic and) have a negative expected value, while random variables that refer to taking a transition from a probabilistic state where at least one outgoing transition is non-invariant refers to a random variable drawn from a finite weight function with expected value 0 and positive variation. Note that random variables corresponding to probabilistic non-invariant transitions are independent and drawn from a *finite* set of distributions.

Let α be any infinite sequence of such random variables. From the results on finitely inhomogeneous controlled random walks [18], we can show that almost-surely the sum of some prefix of α will be lower than $-k'$ (and in fact lower than any finite number). The proof follows the same reasoning as in Proposition 4.1 of [11], where a sufficient and necessary condition was given for not going bankrupt with a positive probability in *solvency games* [7].

We now consider the set of runs induced by σ. As we just showed, almost all runs that have infinitely many interesting events (as described above) will not satisfy the k'-energy condition. Almost all runs that have finitely many interesting events will have an odd dominating priority, and therefore will not satisfy the parity condition. Thus, the probability that the energy parity condition is satisfied by σ is 0. $\square$

Lemma 25. *For every MDP $\mathcal{M}$ and state p, $\mathsf{LVal}_p^{\mathcal{M}} \leq \sup_\sigma \mathcal{P}_p^{\mathcal{M}(\sigma)}(\Diamond(A \cup B))$.*

Proof. Fix p and σ. Every run from p will, with probability 1, eventually reach an end-component and visit all states of the end-component infinitely often [19].

Let C be an end-component such that C forms the infinity set of the runs from p under σ with a positive probability $\tau > 0$. If C does not satisfy the conditions of non-losing end-components, then the probability $\mathcal{P}_q^{\mathcal{M}(\sigma)}(\mathsf{EN}(k) \cap \mathsf{PAR})$ that the k-energy-parity objective is satisfied from some state $q \in C$ is 0, independent of the value k. Thus, the probability of satisfying the k-energy-parity objective from an initial state p is bounded by the chance of reaching a state in some non-losing end-component. These observations hold for every strategy σ and therefore we can bound

$$\mathsf{LVal}_p^{\mathcal{M}} = \sup_k \sup_\sigma \mathcal{P}_p^{\mathcal{M}(\sigma)}(\mathsf{EN}(k) \cap \mathsf{PAR})$$
$$\leq \sup_\sigma \mathcal{P}_p^{\mathcal{M}(\sigma)}(\Diamond(NLE)),$$

where $NLE \subseteq V$ denotes the union of all non-losing end-components. Now Lemma 24 implies that $\sup_\sigma \mathcal{P}_p^{\mathcal{M}(\sigma)}(\Diamond(NLE)) \leq \sup_\sigma \mathcal{P}_p^{\mathcal{M}(\sigma)}(\Diamond(A \cup B))$, which completes the proof. $\square$

Lemma 26. *Determining the limit value of a state p can be done in $\widetilde{\mathcal{O}}(|E| \cdot d \cdot |V|^4 \cdot W + d \cdot |V|^{3.5} \cdot (|\lambda| + |cost|)^2)$ deterministic time. They can also be determined in NP and coNP in the input size when W is given in binary.*

Proof. Recall that $\mathsf{LVal}_p^{\mathcal{M}} = \sup_\sigma \mathcal{P}_p^{\mathcal{M}(\sigma)}(\Diamond(A \cup B))$ by Theorem 18, that $A \cup B = A' \cup B$, and that A' and B are the sets of control states that almost-surely satisfy the storage-parity and mean-payoff-parity objective, respectively. Using the results of Section VI, the algorithm proceeds as follows.

1) Compute A', which can be done in time $\mathcal{O}(|E| \cdot d \cdot |V|^4 \cdot W)$ by Theorem 8.
2) Compute, for each occurring even priority $2i$, the following:
   a) the set of $2i$ maximal end-components, which can be computed in $\mathcal{O}(|E|)$; and
   b) the mean payoff value for the $2i$ maximal end-components can be computed using Karmarkar's algorithm [15] for linear programming in time $\widetilde{\mathcal{O}}(|V|^{3.5} \cdot (|\lambda| + |cost|)^2)$ —note that the complexity refers to *all* (not each) $2i$ maximal end-components.
3) Consider the union of A with all the $2i$ maximal end-components with positive mean payoff computed in Step 2, and compute the maximal achievable probability of reaching this set. (By the results of Section VI, this yields the probability $\sup_\sigma \mathcal{P}_p^{\mathcal{M}(\sigma)}(\Diamond(A \cup B))$.)

The last step costs $\widetilde{\mathcal{O}}(|V|^{3.5} \cdot |\lambda|^2)$ [15] for solving the respective linear program [1], which is dominated by the

estimation of the cost of solving the linear programs from (2b). Likewise, the cost of Step (2a) is dominated by the cost of Step (2b).

This leaves us with once the complexity of (1) and d times the complexity of (2b), resulting in the claimed complexity. Note that it depends on the size of representation (in binary) λ and W (in unary), and the bigger of these values dominates the complexities.

Finally, all steps are in NP and in coNP. $\qquad\square$

VII. LIMIT-SURE ENERGY-PARITY

In this section we provide the reduction from checking if an energy-parity objective holds limit-surely, to checking if such an objective holds almost-surely. The reduction basically extends the MDP so that the controller may "buy" a visit to a good priority (at the expense of energy) if currently in a state p with limit value $\mathsf{LVal}_p = 1$.

Definition 27. The *extension* of a finite MDP $\mathcal{M}$ for given cost and parity functions is the MDP $\mathcal{M}' \supseteq \mathcal{M}$ where, additionally, for every controlled state $s \in V_C$ with $\mathsf{LVal}_s^{\mathcal{M}} = 1$, there is a new state s' with parity 0 and edges $(s, s'), (s', s)$ with $cost(s, s') = -1$ and $cost(s', s) = 0$. We write V' for the set of states of the extension.
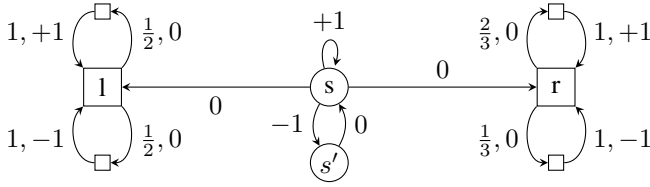

Fig. 3. The extension of the system from Example 1. State s has colour 1, all others 0.

Note that the extension only incurs an $\mathcal{O}(|V_C|)$ blow-up, and $\mathcal{M}'$ satisfies the $\mathsf{EN}(k) \cap \mathsf{PAR}$ objective iff $\mathcal{M}'$ does.

Theorem 28. *Let $\mathcal{M}$ be a MDP with extension $\mathcal{M}'$, p be a state and $k \in \mathbb{N}$. Then, $\mathsf{Val}_p^{\mathcal{M}}(\mathsf{EN}(k) \cap \mathsf{PAR}) = 1$ if, and only if, $\mathcal{P}_p^{\mathcal{M}'(\sigma)}(\mathsf{EN}(k) \cap \mathsf{PAR}) = 1$ for some strategy σ.*

In the remainder of this section we prove this claim. For brevity, let us write $\Delta(w)$ for the cumulative cost $\sum_{i=1}^{k-1} cost(s_i, s_{i+1}) \in \mathbb{Z}$ of all steps in a finite path $w = s_1 s_2 \ldots s_k \in V^*$.

Lemma 29. *Let $\mathcal{M}$ be a MDP with extension $\mathcal{M}'$, p be a state of $\mathcal{M}$, $k \in \mathbb{N}$ and σ' a strategy for $\mathcal{M}'$ such that $\mathcal{P}_p^{\mathcal{M}'(\sigma')}(\mathsf{EN}(k) \cap \mathsf{PAR}) = 1$. Then $\mathsf{Val}_p^{\mathcal{M}}(\mathsf{EN}(k) \cap \mathsf{PAR}) = 1$.*

Proof. Recall Lemma 17, that states with $\mathsf{LVal}_s^{\mathcal{M}} = 1$ have the property that, for every $\varepsilon > 0$, there exists $n_{s,\varepsilon} \in \mathbb{N}$ and a strategy $\sigma_{s,\varepsilon}$ such that

$$\mathcal{P}_s^{\mathcal{M}(\sigma_{s,\varepsilon})}(\mathsf{EN}(n_{s,\varepsilon}) \cap \mathsf{PAR}) \geq 1 - \varepsilon.$$

Consider now a fixed $\varepsilon > 0$ and let $n_\varepsilon \stackrel{\text{def}}{=} \max\{n_{s,\varepsilon} \mid \mathsf{LVal}_s = 1\}$. We show the existence of a strategy σ for $\mathcal{M}$ that satisfies $\mathcal{P}_p^{\mathcal{M}(\sigma)}(\mathsf{EN}(k) \cap \mathsf{PAR}) \geq 1 - \varepsilon$.

We propose the strategy σ which proceeds in $\mathcal{M}$ just as σ' does in $\mathcal{M}'$ but skips over "buying" loops (s, s') followed by (s', s) in $\mathcal{M}'$. This goes on indefinitely unless the observed path $\rho = s_0 s_1 \ldots s_l$ reaches a *tipping point*: the last state s_l has $\mathsf{LVal}_{s_l}^{\mathcal{M}} = 1$ and the accumulated cost is $\Delta(\rho) \geq n_\varepsilon$. At this point σ continues as $\sigma_{s_l, \varepsilon}$.

We claim that $\mathcal{P}_p^{\mathcal{M}(\sigma)}(\mathsf{EN}(k) \cap \mathsf{PAR}) \geq 1 - \varepsilon$. Indeed, first notice that for any prefix $\tau \in V^*$ of a run $\rho \in Runs_p^{\mathcal{M}(\sigma)}$ until the tipping point, there is a unique corresponding path $\tau' = s'_1 s'_2 \ldots s'_i \in V'^*$ in $\mathcal{M}'$, which is a prefix of some run $\rho' \in Runs_p^{\mathcal{M}'(\sigma')}$. Moreover, the strategy σ maintains the invariant that the accumulated cost of such prefix τ is

$$\Delta(\tau) = \Delta(\tau') + |\{j \mid s'_j \in V' \setminus V\}|,$$

the accumulated cost of the corresponding path τ' plus the number of times τ' visited a new state in $V' \setminus V$. In particular this means that the path τ can only violate the energy condition if also τ' does.

To show the claim, first notice that the error introduced by the runs in $Runs_p^{\mathcal{M}(\sigma)}$ that eventually reach a tipping point cannot exceed ε. This is because from the tipping point onwards, σ proceeds as some $\sigma_{s,\varepsilon}$ and thus achieves the energy-parity condition with chance $\geq 1 - \varepsilon$. So the error introduced by the runs in $Runs_p^{\mathcal{M}(\sigma)}$ is a weighted average of values $\leq \varepsilon$, and thus itself at most ε.

Now suppose a run $\rho \in Runs_p^{\mathcal{M}(\sigma)}$ never reaches a tipping point. Then the corresponding run $\rho' \in Runs_p^{\mathcal{M}'(\sigma')}$ cannot visit new states in $V' \setminus V$ more than n_ε times. Since with chance 1, ρ' and therefore also ρ satisfies the k-energy condition it remains to show that ρ also satisfies the parity condition. To see this, just notice that ρ' satisfies this condition almost-surely and since it visits new states only finitely often, ρ and ρ' share an infinite suffix. $\qquad\square$

The "only if" direction of Theorem 28 is slightly more complicated. We go via an intermediate finite system $\mathcal{B}_k$ defined below. The idea is that if $\mathsf{EN}(k) \cap \mathsf{PAR}$ holds limit-surely in $\mathcal{M}$ then PAR holds limit-surely in $\mathcal{B}_k$ and since $\mathcal{B}_k$ is finite this means that PAR also holds almost-surely in $\mathcal{B}_k$. Based on an optimal strategy in $\mathcal{B}_k$ we then derive a strategy in the extension $\mathcal{M}'$ which satisfies $\mathsf{EN}(k) \cap \mathsf{PAR}$ a.s. The two steps of the argument are shown individually as Lemmas 31 and 32. Together with Lemma 29 these complete the proof of Theorem 28.

Definition 30. Let $\mathcal{B}_k$ be the finite MDP that mimics $\mathcal{M}$ but hardcodes the accumulated costs as long as they remain between $-k$ and $|V|$. That is, the states of $\mathcal{B}_k$ are pairs (s, n) where $s \in V$ and $-k \leq n \leq |V|$. Moreover, a state (s, n)

- is a (losing) sink with maximal odd parity if $n = -k$ or $\mathsf{LVal}_s^{\mathcal{M}} < 1$,
- is a (winning) sink with parity 0 if $n = |V|$.

We reuse strategies for $\mathcal{M}$ in $\mathcal{B}_k$ and write $\mathcal{B}_k(\sigma)$ for the Markov chain that is the result of basing decisions on σ until a sink is reached.

Lemma 31. *If* $\mathsf{Val}_s^{\mathcal{M}}(\mathsf{EN}(k) \cap \mathsf{PAR}) = 1$ *then* $\mathsf{Val}_{(s,0)}^{\mathcal{B}_k}(\mathsf{PAR}) = 1.$

Proof. We show that, for every $\varepsilon > 0$, there is a strategy σ such that $\mathcal{P}_s^{\mathcal{B}_k(\sigma)}(\mathsf{PAR}) \geq 1 - \varepsilon$. This would be trivial (by re-using strategies from $\mathcal{M}$) if not for the extra sinks for states with $\mathsf{LVal}_s^{\mathcal{M}} < 1$. Let's call these states *small* here and let S be the set of all small states. We aim to show that the k-energy-parity condition can be satisfied and at the same time, the chance of visiting a small state with accumulated cost below $|V|$ can be made arbitrary small. More precisely, define $D \subseteq \mathit{Runs}^{\mathcal{M}}$ as the set of runs which never visit a small state with accumulated cost below $|V|$:

$$D \stackrel{\text{def}}{=} \{s_0 s_1 \cdots \mid \forall i \in \mathbb{N}. \ s_i \in S \implies \Delta(s_0 \ldots s_i) \geq |V|\}.$$

We claim that

$$\mathsf{Val}_s^{\mathcal{M}}(\mathsf{EN}(k) \cap \mathsf{PAR} \cap D) = 1 \qquad (2)$$

holds. We show this by contradicting the converse that, for

$$\gamma \stackrel{\text{def}}{=} \mathsf{Val}_s^{\mathcal{M}}(\overline{\mathsf{EN}(k) \cap \mathsf{PAR} \cap D}) = \mathsf{Val}_s^{\mathcal{M}}(\overline{\mathsf{EN}(k) \cap \mathsf{PAR}} \cup \overline{D}),$$

$\gamma > 0$. Equivalently, we contradict that, for every strategy σ,

$$\mathcal{P}_s^{\mathcal{M}(\sigma)}(\overline{\mathsf{EN}(k) \cap \mathsf{PAR}}) < \gamma/2 \ \Rightarrow \ \mathcal{P}_s^{\mathcal{M}(\sigma)}(\overline{D}) > \gamma/2. \quad (3)$$

To do this, we define $\delta < 1$ as the maximum of

$$\{ \ \mathsf{Val}_s^{\mathcal{M}}(\mathsf{EN}(n) \cap \mathsf{PAR}) < 1 \ \mid \ s \in S, \ n \leq k + |V| \ \} \cup \{0\},$$

that is, the maximal value $\mathsf{Val}_s^{\mathcal{M}}(\mathsf{EN}(n) \cap \mathsf{PAR}) < 1$ for any $s \in S$ and $n \leq k + |V|$, and 0 if no such value exists. Notice that this is well defined due to the finiteness of V. This value δ estimates the chance that a run that is not in D fails the k-energy-parity condition. In other words, for any strategy σ and value $0 \leq \beta \leq 1$,

$$\mathcal{P}_s^{\mathcal{M}(\sigma)}(\overline{D}) > \beta \text{ implies } \mathcal{P}_s^{\mathcal{M}(\sigma)}(\overline{\mathsf{EN}(k) \cap \mathsf{PAR}}) \geq \beta \cdot (1 - \delta).$$

This is because $\mathcal{P}_s^{\mathcal{M}(\sigma)}(\overline{D})$ is the chance of a run reaching a state s with accumulated cost $n < |V|$ and because $\mathsf{Val}_s^{\mathcal{M}(\sigma)}(\mathsf{EN}(n) \cap \mathsf{PAR}) \leq \delta$.

We pick an $\varepsilon' > 0$ that is smaller than $(\gamma/2) \cdot (1 - \delta)$. By assumption of the lemma, there is some strategy σ such that $\mathcal{P}_s^{\mathcal{M}(\sigma)}(\overline{\mathsf{EN}(k) \cap \mathsf{PAR}}) < \varepsilon' < \gamma/2$. Then by Equation (3), we get $\mathcal{P}_s^{\mathcal{M}(\sigma)}(\overline{D}) > \gamma/2$ and thus $\mathcal{P}_s^{\mathcal{M}(\sigma)}(\overline{\mathsf{EN}(k) \cap \mathsf{PAR}}) \geq (\gamma/2) \cdot (1 - \delta) > \varepsilon'$, which is a contradiction. We conclude that Equation (2) holds.

To get the conclusion of the lemma just observe that for any strategy σ it holds that $\mathcal{P}_s^{\mathcal{M}(\sigma)}(\mathsf{EN}(k) \cap \mathsf{PAR} \cap D) \leq \mathcal{P}_s^{\mathcal{B}_k(\sigma)}(\mathsf{PAR})$. $\quad\square$

Lemma 32. *If* $\mathsf{Val}_s^{\mathcal{B}_k}(\mathsf{PAR}) = 1$ *then* $\mathcal{P}_s^{\mathcal{M}'(\sigma')}(\mathsf{EN}(k) \cap \mathsf{PAR}) = 1$ *for some* σ'.

Proof. Finite MDPs have pure optimal strategies for the PAR objective [20]. Thus by assumption and because $\mathcal{B}_k$ is finite, we can pick an optimal strategy σ satisfying $\mathcal{P}_s^{\mathcal{B}_k(\sigma)}(\mathsf{PAR}) = 1$. Notice that all runs in $\mathit{Runs}_s^{\mathcal{B}_k(\sigma)}$ according to this optimal

strategy must never see a small state (one with $\mathsf{LVal}_p^{\mathcal{M}} < 1$). Based on σ, we construct the strategy σ' for $\mathcal{M}'$ as follows.

The new strategy just mimics σ until the observed path $s_1 s_2 \ldots s_n$ visits the first controlled state after a cycle with positive cost: it holds that $s_n \in V_C$ and there are $i, j \leq n$ with $s_i = s_j$ and $\Delta(s_i \ldots s_j) > 0$. When this happens, σ' uses the new edges to visit a 0-parity state, forgets about the cycle and continues just as from $s_1 s_2 \ldots s_i s_{j+1} \ldots s_n$.

We claim that $\mathcal{P}_s^{\mathcal{M}'(\sigma')}(\mathsf{EN}(k) \cap \mathsf{PAR}) = 1$. To see this, just observe that a run of $\mathcal{M}'(\sigma')$ that infinitely often uses new states in $V' \setminus V$ must satisfy the PAR objective as those states have parity 0. Those runs which visit new states only finitely often have a suffix that directly corresponds to a run of $\mathcal{B}_k(\sigma)$, and therefore also satisfy the parity objective. Finally, almost all runs in $\mathit{Runs}_s^{\mathcal{M}'(\sigma')}$ satisfy the $\mathsf{EN}(k)$ objective because all runs in $\mathit{Runs}_s^{\mathcal{B}_k(\sigma)}$ do, and a negative cost due to visiting a new state in $V' \setminus V$ is always balanced by the strictly positive cost of a cycle. $\quad\square$

This concludes the proof of Theorem 28. The proof of Theorem 3 now follows by Theorems 18 and 28 and the fact that almost-sure reachability, positive mean-payoff and k-energy-parity and k-storage-parity objectives are (pseudo) polynomial time computable (Theorems 2 and 8).

Proof of Theorem 3. Fix a MDP $\mathcal{M} \stackrel{\text{def}}{=} (V_C, V_P, E, \lambda)$ with cost and parity functions. For (1) and (2) we can, by Lemma 26 compute the set of control states p with limit value $\mathsf{LVal}_p^{\mathcal{M}} = 1$. Based on this, we can (in logarithmic space) construct the extension $\mathcal{M}' \stackrel{\text{def}}{=} (V_{C'}, V_{P'}, E', \lambda')$ where $|V_{C'}| = 2 \cdot |V_C|$, $|E'| = |E| + 2 \cdot |V_C|$ and the rest is as in $\mathcal{M}$. By Theorem 28, a state $p \in V_C \cup V_P$ satisfies the k-energy-parity objective limit-surely in $\mathcal{M}$ iff it satisfies it almost-surely in $\mathcal{M}'$. The claim then follows from Theorem 2.

(3) To see that there are finite memory ε-optimal strategies we observe that the strategies we have constructed in Lemma 29 work in phases and in each phase follow some finite memory strategy. In the first state, these strategies follow some almost-surely optimal strategy in the extension $\mathcal{M}'$, but only as long as the energy level remains below some threshold that depends on ε. If this level is exceeded it means that a "tipping point" is reached and the strategy switches to a second phase.

The second phase starts from a state with limit value 1, and our strategy just tries to reach a control state in the set $A' \cup B$ from Section VI. For almost-sure reachability, memoryless deterministic strategies suffice. Finally, when ending up in a state of A', the strategy follows an almost-sure optimal strategy for storage-parity (with finite memory by Theorem 8). Similarly, when ending up in a state of B, the strategy follows almost-sure optimal strategy for the combined positive mean-payoff-parity objective (with finite memory by [2]). $\quad\square$

VIII. LOWER BOUNDS

Polynomial time hardness of all our problems follows, e.g., by reduction from REACHABILITY IN AND-OR GRAPHS

[21], where non-target leaf nodes are energy decreasing sinks. This works even if the energy deltas are encoded in unary.

If we allow binary encoded energy deltas, i.e. $W \gg 1$, then solving two-player energy games is logspace equivalent to solving two-player mean-payoff games ([5], Prop. 12), a well-studied problem in NP ∩ coNP that is not known to be polynomial [22]. Two-player energy games reduce directly to both almost-sure and limit-sure energy objectives for MDPs, where adversarial states are replaced by (uniformly distributed) probabilistic ones: a player max strategy that avoids ruin in the game directly provides a strategy for the controller in the MDP, which means that the energy objective holds almost-surely (thus also limit-surely). Conversely, a winning strategy for the opponent ensures ruin after a fixed number r of game rounds. Therefore the error introduced by any controller strategy in the MDP is at least $(1/d)^r$, where d is the maximal out-degree of the probabilistic states, which means that the energy objective cannot be satisfied even limit-surely (thus not almost-surely). It follows that almost-sure and limit-sure energy objectives for MDPs are at least as hard as mean-payoff games. The same holds for almost-sure and limit-sure storage objectives for MDPs, since in the absence of parity conditions, storage objectives coincide with energy objectives. Finally we obtain that all the more general almost-sure and limit-sure energy-parity and storage-parity objectives for MDPs are at least as hard as mean-payoff games.

IX. Conclusions and Future Work

We have shown that even though strategies for almost-sure energy parity objectives in MDPs require infinite memory, the problem is still in NP ∩ coNP. Moreover, we have shown that the limit-sure problem (i.e. the problem of checking whether a given configuration (state and energy level) in energy-parity MDPs has value 1) is also in NP ∩ coNP. However, the fact that a state has value 1 can always be witnessed by a family of strategies attaining values $1 - \epsilon$ (for every $\epsilon > 0$) where each member of this family uses only finite memory.

We leave open the decidability status of quantitative questions, e.g. whether $\mathsf{Val}_p^{\mathcal{M}}(\mathsf{EN}(k) \cap \mathsf{PAR}) \geq 0.5$ holds.

Energy-parity objectives on finite MDPs correspond to parity objectives on certain types of infinite MDPs where the current energy value is part of the state. More exactly, these infinite MDPs can be described by single-sided vector addition systems with states [23], [24], where the probabilistic transitions cannot change the counter values but only the control-states (thus yielding an upward-closed winning set). I.e. single-sidedness corresponds to energy objectives. For those systems, almost-sure Büchi objectives are decidable (even for multiple energy dimensions) [24], but the decidability of the limit-sure problem was left open. This problem is solved here, even for parity objectives, but only for dimension one. However, decidability for multiple energy dimensions remains open.

If one considers the more general case of MDPs induced by counter *machines*, i.e. with zero-testing transitions, then even for single-sided systems as described above all problems become undecidable from dimension 2 onwards. However,

decidability of almost-sure and limit-sure parity conditions for MDPs induced by one-counter machines (with only one dimension of energy) remains open.

References

[1] M. L. Puterman, *Markov Decision Processes*. Wiley, 1994.

[2] K. Chatterjee and L. Doyen, "Energy and mean-payoff parity Markov decision processes," in *MFCS 2011*, vol. 6907, 2011, pp. 206–218.

[3] ——, "Games and Markov decision processes with mean-payoff parity and energy parity objectives," in *Proc. of MEMICS*, ser. LNCS, vol. 7119. Springer, 2011, pp. 37–46.

[4] A. Chakrabarti, L. De Alfaro, T. A. Henzinger, and M. Stoelinga, "Resource interfaces," in *International Workshop on Embedded Software*. Springer, 2003, pp. 117–133.

[5] P. Bouyer, U. Fahrenberg, K. G. Larsen, N. Markey, and J. Srba, *Infinite Runs in Weighted Timed Automata with Energy Constraints*, ser. LNCS. Springer, 2008, vol. 5215, pp. 33–47.

[6] R. Bloem, K. Chatterjee, T. A. Henzinger, and B. Jobstmann, "Better quality in synthesis through quantitative objectives," in *CAV*, ser. LNCS, vol. 5643, 2009, pp. 140–156.

[7] N. Berger, N. Kapur, L. Schulman, and V. Vazirani, "Solvency Games," in *FSTTCS*, ser. LIPIcs, vol. 2, 2008, pp. 61–72.

[8] K. Chatterjee, T. A. Henzinger, and M. Jurdzinski, "Mean-payoff parity games," in *LICS*, 2005, pp. 178–187.

[9] K. Chatterjee and L. Doyen, "Energy parity games," *Theoretical Computer Science*, vol. 458, pp. 49 – 60, 2012.

[10] T. Brázdil, A. Kučera, and P. Novotný, "Optimizing the expected mean payoff in energy Markov decision processes," in *ATVA*, ser. LNCS, vol. 9938. Springer, 2016, pp. 32–49, full version at https://arxiv.org/abs/1607.00678.

[11] T. Brázdil, V. Brožek, K. Etessami, A. Kučera, and D. Wojtczak, "One-counter Markov decision processes," in *SODA*. Society for Industrial and Applied Mathematics, 2010, pp. 863–874.

[12] T. Brázdil, V. Brožek, and K. Etessami, "One-Counter Stochastic Games," in *FSTTCS*, ser. LIPIcs, vol. 8, 2010, pp. 108–119.

[13] R. Mayr, S. Schewe, P. Totzke, and D. Wojtczak, "MDPs with Energy-Parity Objectives," *CoRR*, vol. abs/1701.02546, 2017. [Online]. Available: http://arxiv.org/abs/1701.02546

[14] P. Billingsley, *Probability and Measure*. New York, NY: Wiley, 1995, third Edition.

[15] N. Karmarkar, "A new polynomial-time algorithm for linear programming," in *STOC*, 1984, pp. 302–311.

[16] K. Etessami, D. Wojtczak, and M. Yannakakis, "Quasi-birth–death processes, tree-like QBDs, probabilistic 1-counter automata, and pushdown systems," *Performance Evaluation*, vol. 67, no. 9, pp. 837–857, 2010.

[17] T. Brázdil, S. Kiefer, and A. Kučera, "Efficient analysis of probabilistic programs with an unbounded counter," *Journal of the ACM*, vol. 61, no. 6, pp. 41:1–41:35, 2014.

[18] R. Durrett, H. Kesten, and G. Lawler, "Making money from fair games," in *Random walks, Brownian motion, and interacting particle systems*. Springer Science+Business Media, LLC, 1991, pp. 255–267.

[19] L. De Alfaro, "Formal verification of probabilistic systems," Ph.D. dissertation, Stanford University, 1997.

[20] K. Chatterjee, M. Jurdziński, and T. A. Henzinger, "Quantitative stochastic parity games," in *Proceedings of the Fifteenth Annual ACM-SIAM Symposium on Discrete Algorithms*, 2004, pp. 121–130.

[21] N. Immerman, "Number of quantifiers is better than number of tape cells," *Journal of Computer and System Sciences*, vol. 22, no. 3, pp. 384–406, 1981.

[22] U. Zwick and M. Paterson, "The complexity of mean payoff games on graphs," *Theoretical Computer Science*, vol. 158, no. 1, pp. 343–359, 1996.

[23] P. Abdulla, R. Mayr, A. Sangnier, and J. Sproston, "Solving parity games on integer vectors," in *Proc. of CONCUR*, ser. LNCS, vol. 8052, 2013.

[24] P. A. Abdulla, R. Ciobanu, R. Mayr, A. Sangnier, and J. Sproston, "Qualitative analysis of VASS-induced MDPs," in *FOSSACS*, ser. LNCS, vol. 9634, 2016.