

# Bounds for Synchronizing Markov Decision Processes

Laurent Doyen<sup>1</sup> and Marie van den Bogaard<sup>2</sup>

<sup>1</sup> CNRS & LMF, ENS Paris-Saclay, France

<sup>2</sup> LIGM, Université Gustave Eiffel, France

**Abstract.** We consider Markov decision processes with synchronizing objectives, which require that a probability mass of  $1 - \varepsilon$  accumulates in a designated set of target states, either once, always, infinitely often, or always from some point on, where  $\varepsilon = 0$  for sure synchronizing, and  $\varepsilon \rightarrow 0$  for almost-sure and limit-sure synchronizing.

We introduce two new qualitative modes of synchronizing, where the probability mass should be either *positive*, or *bounded* away from 0. They can be viewed as dual synchronizing objectives. We present algorithms and tight complexity results for the problem of deciding if a Markov decision process is positive, or bounded synchronizing, and we provide explicit bounds on  $\varepsilon$  in all synchronizing modes. In particular, we show that deciding positive and bounded synchronizing always from some point on, is coNP-complete.

## 1 Introduction

Markov decision processes (MDP) are finite-state probabilistic systems with controllable (non-deterministic) choices. They play a central role in several application domains for practical purpose [14, 19], and in theoretical computer science as a basic model for the analysis of stochastic transition systems [2, 9].

In the traditional state-based semantics, we consider the *sequences of states* that form a path in the underlying graph of the MDP. When a control policy (or strategy) for the non-deterministic choices is fixed, we obtain a purely stochastic process that induces a probability measure over sets of paths [2, 7].

In the more recent distribution-based semantics, the outcome of a stochastic process is a sequence of distributions over states [3, 18]. This alternative semantics has received some attention recently for theoretical analysis of probabilistic bisimulation [16] and is adequate to describe large populations of agents [12, 8] with applications in system biology [18, 1]. The behaviour of an agent is modeled as an MDP with some state space  $Q$ , and a large population of identical agents is described by a (continuous) distribution  $d : Q \rightarrow [0, 1]$  that gives the fraction  $d(q)$  of agents in the population that are in each state  $q \in Q$ . The control problem is to construct a strategy for the agents that guarantees a specified global outcome of the agents, defined in terms of *sequences of distributions*. Specifications of interest include safety objectives [1] and synchronization objectives [12]. A distribution is  $p$ -synchronized in a set  $T$  of states if it assigns to the states

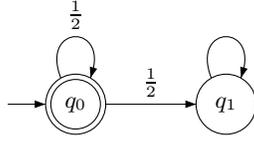


Fig. 1: A Markov chain that is positively but not boundedly synchronizing (for all modes except eventually).

in  $T$  a mass of probability at least  $p$ . Synchronization objectives require that  $p$ -synchronized distributions occur in the outcome sequence, either, at some position (eventually), at all positions (always), infinitely often (weakly), or always from some point on (strongly), where synchronization is sure winning for  $p = 1$ , and almost-sure or limit-sure winning for  $p$  arbitrarily close to 1 [12].

Consider eventually synchronizing as an illustration. Formally, denoting by  $d_i^\sigma(T)$  the probability mass in a set  $T$  under strategy  $\sigma$  at position  $i$  in a given MDP, the three winning modes for eventually synchronizing correspond to the following three possible orders of the quantifiers:

- $\forall \varepsilon > 0 \cdot \exists \sigma \cdot \exists i : d_i^\sigma(T) \geq 1 - \varepsilon$ , for limit-sure winning,
- $\exists \sigma \cdot \forall \varepsilon > 0 \cdot \exists i : d_i^\sigma(T) \geq 1 - \varepsilon$ , for almost-sure winning,
- $\exists \sigma \cdot \exists i \cdot \forall \varepsilon > 0 : d_i^\sigma(T) \geq 1 - \varepsilon$ , for sure winning.

Note that the formula  $\forall \varepsilon > 0 : d_i^\sigma(T) \geq 1 - \varepsilon$  is equivalent to  $d_i^\sigma(T) = 1$  in the case of sure winning. Defining the value of a strategy  $\sigma$  as  $\text{val}(\sigma) = \sup_i d_i^\sigma(T)$ , the question for limit-sure winning is analogous to the cutpoint isolation problem for value 1, i.e. whether the value 1 can be approached arbitrarily closely [20, 4]. Previous work [12] shows that the above three questions are PSPACE-complete, and presents a construction of the (existentially quantified) strategy  $\sigma$  when one exists.

In this paper, we consider dual synchronization objectives obtained either by taking the negation of the synchronization objectives, or by replacing the existential quantifier on strategies by a universal quantifier.

1. Taking the negation corresponds to the control player having no strategy to satisfy the synchronization objective. In that case, we show that a more precise information can be derived, namely bounds on the value of  $\varepsilon$ , which is existentially quantified, and we construct explicit values for the four synchronizing modes. These values give bounds on the isolation distance of the value 1. For instance, the negation of limit-sure eventually synchronizing in  $T$  is given by the formula:

$$\exists \varepsilon > 0 \cdot \forall \sigma \cdot \forall i : d_i^\sigma(T) \leq 1 - \varepsilon.$$

We show that the statement holds for a value  $\varepsilon = \varepsilon_e(n, \alpha, \alpha_0)$  that depends on the number  $n$  of states of the MDP, the smallest positive probability  $\alpha$  in the transitions of the MDP, and the smallest positive probability  $\alpha_0$  in the initial

Table 1: Positive and Bounded winning modes for always, strongly, weakly, and eventually synchronizing objectives.

	Always			Strongly		
Positively	$\exists\sigma$	$\forall i$	$d_i^\sigma(T) > 0$	$\exists\sigma$	$\exists N$	$\forall i \geq N$ $d_i^\sigma(T) > 0$
Boundedly	$\exists\sigma$	$\inf_i$	$d_i^\sigma(T) > 0$	$\exists\sigma$	$\liminf_{i \rightarrow \infty}$	$d_i^\sigma(T) > 0$
	Weakly			Eventually		
Positively	$\exists\sigma$	$\forall N$	$\exists i \geq N$ $d_i^\sigma(T) > 0$	$\exists\sigma$	$\exists i$	$d_i^\sigma(T) > 0$
Boundedly	$\exists\sigma$	$\limsup_{i \rightarrow \infty}$	$d_i^\sigma(T) > 0$	$\exists\sigma$	$\sup_i$	$d_i^\sigma(T) > 0$

distribution  $d_0$  (see Theorem 1). The most interesting case is when limit-sure weakly synchronizing does not hold, that is:

$$\exists \varepsilon > 0 \cdot \forall \sigma \cdot \exists N \cdot \forall i \geq N : d_i^\sigma(T) \leq 1 - \varepsilon.$$

Given the value  $\varepsilon = \varepsilon_w$  that satisfies this condition (see Theorem 2), the value of  $N$  can be arbitrarily large (depending on the strategy  $\sigma$ ). Nevertheless, we can effectively construct a constant  $N_w$  such that, for all strategies  $\sigma$ , in the sequence  $(d_i^\sigma)_{i \in \mathbb{N}}$  there are at most  $N_w$  distributions that are  $(1 - \varepsilon_w)$ -synchronized in  $T$ .

2. Replacing the existential strategy quantifier by a universal quantifier corresponds to an adversarial MDP where all strategies need to satisfy the requirement, or after taking the negation, to the existence of a strategy that violates a dual of the synchronizing requirement. Note that there is no more alternation of quantifiers on  $\varepsilon$  and on  $\sigma$  ( $\forall \varepsilon \cdot \forall \sigma$  is the same as  $\forall \sigma \cdot \forall \varepsilon$ ), which gives rise to only two new winning modes in existential form:

- $\exists \sigma \cdot \exists \varepsilon > 0 \cdot \forall i : d_i^\sigma(T) \geq \varepsilon$ , that we call *bounded* winning,
- $\exists \sigma \cdot \forall i \cdot \exists \varepsilon > 0 : d_i^\sigma(T) \geq \varepsilon$ , that we call *positive* winning (since this is equivalent to  $\exists \sigma \cdot \forall i : d_i^\sigma(T) > 0$ ).

Table 1 presents the analogous definitions of bounded and positive winning for the four synchronizing modes. It is easy to see that for eventually synchronizing, the positive and bounded mode coincide, while for the other synchronizing modes the positive and bounded modes are distinct, already in Markov chains (see Fig. 1).

We establish the complexity of deciding bounded and positive winning in the four synchronizing modes, given an MDP and initial distribution (which we call the membership problem), and we also construct explicit values for  $\varepsilon$ . Adversarial MDPs are a special case of two-player stochastic games [7] in which only the second player (the adversary of the first player) is non-trivial. The results of this paper will be useful for the analysis of adversarial MDPs obtained from a game by fixing a strategy of the first player. The complexity results are summarized in Table 2. For positive winning, memoryless winning strategies exist (playing all actions uniformly at random is sufficient), and the problem

Table 2: Computational complexity of the membership problem (for eventually synchronizing, the positive and bounded modes coincide).

	Always	Eventually	Weakly	Strongly
Positively	coNP-C	NL-C	NL-C	coNP-C
Boundedly	coNP-C		NL-C	coNP-C

can be solved by graph-theoretic techniques on Markov chains. For bounded winning, the most challenging case is strongly synchronizing, where we show that a simple form of strategy with memory is winning, and that the decision problem is coNP-complete. We give a structural characterization of bounded strongly synchronizing MDPs, and show that it can be decided in coNP. Note that the coNP upper bound is not obtained by guessing a strategy, since the coNP lower bound holds in the case of Markov chains where strategies play no role. Omitted proofs and additional material can be found in an extended version of this paper [13].

*Related works.* The distribution-based semantics of MDPs [3, 18] has received an increased amount of attention recently, with works on safety objectives [1] and synchronizing objectives [12] (see also references therein). Logic and automata-based frameworks express distribution-based properties, by allowing different order of the logical quantifiers, such as  $\forall\sigma\exists i$  in standard reachability which becomes  $\exists i\forall\sigma$  in synchronized reachability [3, 17, 6]. The bounded and positive winning modes introduced in this paper have not been considered before. They bear some similarity with the qualitative winning modes in concurrent games [11].

Applications are found in modeling of large populations of identical agents, such as molecules, yeast, bacteria, etc. [18, 1] where the probability distributions represent concentrations of each agent in the system. Analogous models have been considered in a discrete setting where the number of agents is a parameter  $n$ , giving rise to control problems for parameterized systems, asking if there exists a strategy that brings all  $n$  agents synchronously to a target state [5, 8].

## 2 Definitions

A *probability distribution* over a finite set  $Q$  is a function  $d : Q \rightarrow [0, 1]$  such that  $\sum_{q \in Q} d(q) = 1$ . The *support* of  $d$  is the set  $\text{Supp}(d) = \{q \in Q \mid d(q) > 0\}$ . We denote by  $\mathcal{D}(Q)$  the set of all probability distributions over  $Q$ . Given a set  $T \subseteq Q$ , let  $d(T) = \sum_{s \in T} d(s)$ .

A *Markov decision process* (MDP) is a tuple  $\mathcal{M} = \langle Q, \mathbf{A}, \delta \rangle$  where  $Q$  is a finite set of states,  $\mathbf{A}$  is a finite set of labels called actions, and  $\delta : Q \times \mathbf{A} \rightarrow \mathcal{D}(Q)$  is a probabilistic transition function. A *Markov chain* is an MDP with singleton action set  $|\mathbf{A}| = 1$ . Given a state  $q \in Q$  and an action  $a \in \mathbf{A}$ , the successor state of  $q$  under action  $a$  is  $q'$  with probability  $\delta(q, a)(q')$ .

Given  $X, Y \subseteq Q$ , let

$$\text{APre}(Y, X) = \{q \in Q \mid \exists a \in \mathbf{A} : \text{Supp}(\delta(q, a)) \subseteq Y \wedge \text{Supp}(\delta(q, a)) \cap X \neq \emptyset\},$$

be the set of states from which there is an action to ensure that all successor states are in  $Y$  and that with positive probability the successor state is in  $X$ , and for  $X = Q$ , let  $\text{Pre}(Y) = \text{APre}(Y, Q) = \{q \in Q \mid \exists a \in \mathbf{A} : \text{Supp}(\delta(q, a)) \subseteq Y\}$  be the set of states from which there is an action to ensure (with probability 1) that the successor state is in  $Y$ . For  $k > 0$ , let  $\text{Pre}^k(Y) = \text{Pre}(\text{Pre}^{k-1}(Y))$  with  $\text{Pre}^0(Y) = Y$ . Note that the sequence  $\text{Pre}^k(Y)$  of iterated predecessors is ultimately periodic, precisely there exist  $k_1 < k_2 \leq 2^{|Q|}$  such that  $\text{Pre}^{k_1}(Y) = \text{Pre}^{k_2}(Y)$ .

*Strategies.* A (randomized) *strategy* in  $\mathcal{M}$  is a function  $\sigma : (QA)^*Q \rightarrow \mathcal{D}(\mathbf{A})$  that, given a finite sequence  $\rho = q_0a_0q_1a_1 \dots q_k$ , chooses the next action  $a_k$  with probability  $\sigma(\rho)(a_k)$ . We write  $\sigma_1 \subseteq \sigma_2$  if  $\text{Supp}(\sigma_1(\rho)) \subseteq \text{Supp}(\sigma_2(\rho))$  for all  $\rho \in (QA)^*Q$ . A strategy  $\sigma$  is *pure* if for all  $\rho \in (QA)^*Q$ , there exists an action  $a \in \mathbf{A}$  such that  $\sigma(\rho)(a) = 1$ . In all problems considered in this paper, it is known that pure strategies are sufficient [12]. However, the bounds we provide in case there is no winning strategy hold for all strategies, pure or randomized.

Given an initial distribution  $d_0 \in \mathcal{D}(Q)$  and a strategy  $\sigma$  in  $\mathcal{M}$ , the probability of a finite sequence  $\rho = q_0a_0q_1a_1 \dots q_k$  is defined by  $\text{Pr}_{d_0}^\sigma(\rho) = d_0(q_0) \cdot \prod_{j=0}^{k-1} \sigma(q_0a_0 \dots q_j)(a_j) \cdot \delta(q_j, a_j)(q_{j+1})$ . For an initial distribution  $d_0$  such that  $d_0(q_0) = 1$ , we sometimes write  $\text{Pr}_{q_0}^\sigma(\cdot)$  and say that  $q_0$  is the initial state. We say that  $\rho$  is *compatible* with  $\sigma$  and  $d_0$  if  $\text{Pr}_{d_0}^\sigma(\rho) > 0$ . By extension, an infinite sequence  $\pi \in (QA)^\omega$  is compatible with  $\sigma$  and  $d_0$  if all prefixes of  $\pi$  that end in a state are compatible. It is standard to extend (in a unique way)  $\text{Pr}_{d_0}^\sigma$  over Borel sets of infinite paths in  $(QA)^\omega$  (called events), by assigning probability  $\text{Pr}_{d_0}^\sigma(\rho)$  to the basic cylinder set containing all infinite paths with prefix  $\rho$  [22, 2]. Given a set  $T \subseteq Q$  of target states, and  $k \in \mathbb{N}$ , we define the following events (sometimes called objectives):

- $\square T = \{q_0a_0q_1 \dots \in (QA)^\omega \mid \forall i : q_i \in T\}$  the safety event of staying in  $T$ ;
- $\diamond T = \{q_0a_0q_1 \dots \in (QA)^\omega \mid \exists i : q_i \in T\}$  the event of reaching  $T$ ;
- $\diamond^k T = \{q_0a_0q_1 \dots \in (QA)^\omega \mid q_k \in T\}$  the event of reaching  $T$  after exactly  $k$  steps;

A distribution  $d_0$  is *almost-sure winning* for an event  $\Omega$  if there exists a strategy  $\sigma$  such that  $\text{Pr}_{d_0}^\sigma(\Omega) = 1$ , and *limit-sure winning* if  $\sup_\sigma \text{Pr}_{d_0}^\sigma(\Omega) = 1$ , that is the event  $\Omega$  can be realized with probability arbitrarily close to 1. Finally  $d_0$  is *sure winning* for  $\Omega$  if there exists a strategy  $\sigma$  such that all paths compatible with  $\sigma$  and  $d_0$  belong to  $\Omega$ .

Safety and reachability events are dual, in the sense that  $\diamond T$  and  $\square(Q \setminus T)$  form a partition of  $(QA)^\omega$ . It is known for safety objectives  $\square T$  that the three winning regions (sure, almost-sure winning, and limit-sure winning) coincide in MDPs, and for reachability objectives  $\diamond T$ , almost-sure and limit-sure winning coincide [10]. It follows that if the negation of almost-sure reachability holds, that

is  $\Pr_{d_0}^\sigma(\diamond T) < 1$  for all strategies  $\sigma$ , then equivalently  $\inf_\sigma \Pr_{d_0}^\sigma(\square(Q \setminus T)) > 0$  (note the strict inequality), the probability mass that remains always outside  $T$  can be bounded. An explicit bound can be obtained from the classical characterization of the winning region for almost-sure reachability [9].

**Lemma 1.** *If a distribution  $d_0$  is not almost-sure winning for a reachability objective  $\diamond T$  in an MDP  $\mathcal{M}$ , then for all strategies  $\sigma$ , for all  $i \geq 0$ , we have  $\Pr_{d_0}^\sigma(\diamond^i T) \leq 1 - \alpha_0 \cdot \alpha^n$  where  $n = |Q|$  is the number of states and  $\alpha$  the smallest positive probability in  $\mathcal{M}$ , and  $\alpha_0 = \min\{d_0(q) \mid q \in \text{Supp}(d_0)\}$  is the smallest positive probability in the initial distribution  $d_0$ .*

In Lemma 1 it is crucial to notice that the bound  $\alpha_0 \cdot \alpha^n$  is independent of the number  $i$  of steps.

*Synchronizing objectives.* We consider MDPs as generators of sequences of probability distributions over states [18]. Given an initial distribution  $d_0 \in \mathcal{D}(Q)$  and a strategy  $\sigma$  in  $\mathcal{M}$ , the sequence  $\mathcal{M}^\sigma = (\mathcal{M}_i^\sigma)_{i \in \mathbb{N}}$  of probability distributions (from  $d_0$ , which we assume is clear from the context) is defined by  $\mathcal{M}_i^\sigma(q) = \Pr_{d_0}^\sigma(\diamond^i \{q\})$  for all  $i \geq 0$  and  $q \in Q$ . Hence,  $\mathcal{M}_i^\sigma$  is the probability distribution over states after  $i$  steps under strategy  $\sigma$ . Note that  $\mathcal{M}_0^\sigma = d_0$ .

Informally, synchronizing objectives require that the probability of some set  $T$  of states tends to 1 in the sequence  $(\mathcal{M}_i^\sigma)_{i \in \mathbb{N}}$ , either always, once, infinitely often, or always after some point [12]. Given a target set  $T \subseteq Q$ , we say that a probability distribution  $d$  is *p-synchronized* in  $T$  if  $d(T) \geq p$  (and strictly *p-synchronized* in  $T$  if  $d(T) > p$ ), and that a sequence  $d_0 d_1 \dots$  of probability distributions is:

- (a) *always p-synchronizing* if  $d_i$  is *p-synchronized* (in  $T$ ) for all  $i \geq 0$ ;
- (b) *event(ually) p-synchronizing* if  $d_i$  is *p-synchronized* (in  $T$ ) for some  $i \geq 0$ ;
- (c) *weakly p-synchronizing* if  $d_i$  is *p-synchronized* (in  $T$ ) for infinitely many  $i$ 's;
- (d) *strongly p-synchronizing* if  $d_i$  is *p-synchronized* (in  $T$ ) for all but finitely many  $i$ 's.

Given an initial distribution  $d_0$ , we say that for the objective of {always, eventually, weakly, strongly} synchronizing from  $d_0$ , the MDP  $\mathcal{M}$  is:

- *sure winning* if there exists a strategy  $\sigma$  such that the sequence  $\mathcal{M}^\sigma$  from  $d_0$  is {always, eventually, weakly, strongly} 1-synchronizing in  $T$ ;
- *almost-sure winning* if there exists a strategy  $\sigma$  such that for all  $\varepsilon > 0$  the sequence  $\mathcal{M}^\sigma$  from  $d_0$  is {always, eventually, weakly, strongly}  $(1 - \varepsilon)$ -synchronizing in  $T$ ;
- *limit-sure winning* if for all  $\varepsilon > 0$ , there exists a strategy  $\sigma$  such that the sequence  $\mathcal{M}^\sigma$  from  $d_0$  is {always, eventually, weakly, strongly}  $(1 - \varepsilon)$ -synchronizing in  $T$ ;

For  $\lambda \in \{\text{always, event, weakly, strongly}\}$ , we denote by  $\langle\langle 1 \rangle\rangle_{\text{sure}}^\lambda(T)$  the *winning region* defined as the set of initial distributions  $d_0$  such that  $\mathcal{M}$  is sure winning for  $\lambda$ -synchronizing in  $T$  (in this notation, we assume that  $\mathcal{M}$  is clear

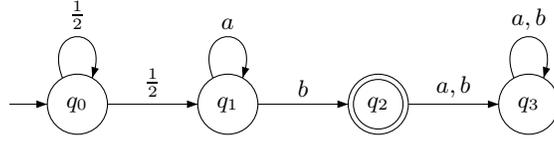


Fig. 2: An MDP where  $\{q_0\} \in \langle\langle 1 \rangle\rangle_{limit}^{event}(\{q_2\})$ .

from the context). We define analogously the winning regions  $\langle\langle 1 \rangle\rangle_{almost}^\lambda(T)$  and  $\langle\langle 1 \rangle\rangle_{limit}^\lambda(T)$  of almost-sure and limit-sure winning distributions.

It is known that for all winning modes, only the support of the initial distributions is relevant, that is for every winning region  $W = \langle\langle 1 \rangle\rangle_\mu^\lambda(T)$  (where  $\mu \in \{sure, almost, limit\}$ ), for all distributions  $d, d'$ , if  $\text{Supp}(d) = \text{Supp}(d')$ , then  $d \in W$  if and only if  $d' \in W$  [12]. Therefore, in the sequel we sometimes write  $S \in \langle\langle 1 \rangle\rangle_\mu^\lambda(T)$ , which can be read as any distribution  $d$  with support  $S$  is in  $\langle\langle 1 \rangle\rangle_\mu^\lambda(T)$ . For each synchronizing mode  $\lambda$  and winning mode  $\mu$ , the membership problem asks to decide, given an MDP  $\mathcal{M}$ , a target set  $T$ , and a set  $S$ , whether  $S \in \langle\langle 1 \rangle\rangle_\mu^\lambda(T)$ .

Consider the MDP in Fig. 2, with initial state  $q_0$  and target set  $T = \{q_2\}$ . The probability mass cannot loop through  $q_2$  and therefore, it is immediate that the MDP is neither always, nor weakly, nor strongly  $(1 - \varepsilon)$ -synchronizing, thus  $\{q_0\} \notin \langle\langle 1 \rangle\rangle_{almost}^\lambda(T)$  for  $\lambda = always, weakly, strongly$ , and thus also  $\{q_0\} \notin \langle\langle 1 \rangle\rangle_{sure}^\lambda(T)$ .

For eventually synchronizing in  $q_2$ , at every step, half of the probability mass in  $q_0$  stays in  $q_0$  while the other half is sent to  $q_1$ . Thus, the probability mass in  $q_0$  tends to 0 but is strictly positive at every step, and the MDP is not sure eventually synchronizing,  $\{q_0\} \notin \langle\langle 1 \rangle\rangle_{sure}^{event}(T)$ . In state  $q_1$ , action  $a$  keeps the probability mass in  $q_1$ , while action  $b$  sends it to the target state  $q_2$ . If action  $b$  is never chosen, then  $q_2$  is never reached, and whenever  $b$  is chosen, a strictly positive probability mass remains in  $q_0$ , thus the MDP is not almost-sure eventually synchronizing,  $\{q_0\} \notin \langle\langle 1 \rangle\rangle_{almost}^{event}(T)$ . On the other hand, for every  $\varepsilon > 0$ , the strategy that plays  $a$  in  $q_1$  for  $k$  steps such that  $\frac{1}{2^k} < \varepsilon$ , and then plays  $b$ , is winning for eventually  $(1 - \varepsilon)$ -synchronizing in  $T$ . Thus the MDP is limit-sure eventually synchronizing,  $\{q_0\} \in \langle\langle 1 \rangle\rangle_{limit}^{event}(T)$ .

The MDP in Fig. 3 is also limit-sure eventually synchronizing in  $\{q_2\}$ . As the probability mass is sent back to  $q_0$  from  $q_2$ , the MDP is even almost-sure weakly (and eventually) synchronizing, using a strategy that plays action  $a$  in  $q_1$  for  $k$  steps to accumulate probability mass  $1 - \frac{1}{2^k}$  in  $q_1$ , then plays action  $b$  and repeats the same pattern for increasing values of  $k$ .

*End-components.* Given a state  $q \in Q$  and a set  $S \subseteq Q$ , let  $A_S(q)$  be the set of all actions  $a \in A$  such that  $\text{Supp}(\delta(q, a)) \subseteq S$ . A *closed* set in an MDP is a set  $S \subseteq Q$  such that  $A_S(q) \neq \emptyset$  for all  $q \in S$ . A set  $S \subseteq Q$  is an *end-component* [10, 2] if (i)  $S$  is closed, and (ii) the graph  $(S, E_S)$  is strongly connected where  $E_S = \{(q, q') \in S \times S \mid \delta(q, a)(q') > 0 \text{ for some } a \in A_S(q)\}$  denote the set

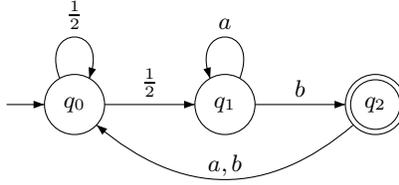


Fig. 3: An MDP where  $\{q_0\} \in \langle\langle 1 \rangle\rangle_{almost}^{weakly}(\{q_2\})$ .

of edges given the actions. In the sequel, end-components should be considered *maximal*, that is such that no strict superset is an end-component. We denote by  $\mathcal{E}$  the union of all end-components, and for  $q \in \mathcal{E}$ , we denote by  $\mathcal{E}(q)$  the maximal end-component containing  $q$ . A fundamental property of end-components is that under arbitrary strategies, with probability 1 the set of states visited infinitely often along a path is an end-component.

**Lemma 2 ([9, 10]).** *Let  $\mathcal{M}$  be an MDP. For all strategies  $\sigma$ , we have  $\liminf_{i \rightarrow \infty} \mathcal{M}_i^\sigma(\mathcal{E}) = 1$ .*

*Tracking counter in MDPs.* It will be useful to track the number of steps (modulo a given number  $r$ ) in MDPs. Given a number  $r \in \mathbb{N}$ , define the MDP  $\mathcal{M} \times [r] = \langle Q_r, \mathbf{A}, \delta_r \rangle$  where  $Q_r = Q \times \{r-1, \dots, 1, 0\}$  and  $\delta_r$  is defined as follows, for all  $\langle q, i \rangle, \langle q', j \rangle \in Q_r$  and  $a \in \mathbf{A}$ :

$$\delta_r(\langle q, i \rangle, a)(\langle q', j \rangle) = \begin{cases} \delta(q, a)(q') & \text{if } j = i - 1 \pmod r, \\ 0 & \text{otherwise.} \end{cases}$$

For a distribution  $d \in \mathcal{D}(Q)$  and  $0 \leq t < r$ , we denote by  $d \times \{t\}$  the distribution defined, for all  $q \in Q$ , by  $d \times \{t\}(\langle q, i \rangle) = d(q)$  if  $t = i$ , and  $d \times \{t\}(\langle q, i \rangle) = 0$  otherwise. Given a finite sequence  $\rho = q_0 a_0 q_1 a_1 \dots q_n$  in  $\mathcal{M}$ , and  $0 \leq t < r$ , there is a corresponding sequence  $\rho' = \langle q_0, k_0 \rangle a_0 \langle q_1, k_1 \rangle a_1 \dots \langle q_n, k_n \rangle$  in  $\mathcal{M} \times [r]$  where  $k_0 = t$  and  $k_{i+1} = k_i - 1 \pmod r$  for all  $0 \leq i < n$ . Since the sequence  $\rho'$  is uniquely defined from  $\rho$  and  $t$ , there is a clear bijection between the paths in  $\mathcal{M}$  starting in  $q_0$  and the paths in  $\mathcal{M} \times [r]$  starting in  $\langle q_0, t \rangle$ . In the sequel, we freely omit to apply and mention this bijection. In particular, we often consider that a strategy  $\sigma$  in  $\mathcal{M}$  can be played directly in  $\mathcal{M} \times [r]$ .

Consider the MDP in Fig. 4 (which is in fact a Markov chain), with initial state  $q_0$  and target set  $T = \{q_2, q_3\}$ . There are two end-components,  $S_1 = \{q_1, q_2\}$  and  $S_2 = \{q_3, q_4\}$ . Although both  $S_1$  and  $S_2$  are sure eventually synchronizing in  $T$  (from  $q_1$  and  $q_3$  respectively), the uniform distribution over  $\{q_1, q_3\}$  is not even limit-sure eventually synchronizing in  $T$ .

### 3 Eventually synchronizing

In the rest of this paper, fix an MDP  $\mathcal{M} = \langle Q, \mathbf{A}, \delta \rangle$  and let  $n = |Q|$  be the size of  $\mathcal{M}$ , and let  $\alpha$  be the smallest positive probability in the transitions of  $\mathcal{M}$ .

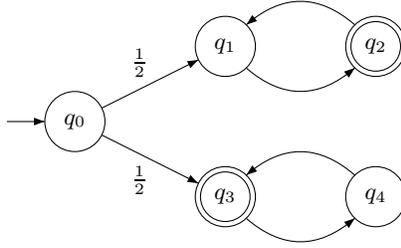


Fig. 4: A Markov chain with two periodic end-components.

We first consider eventually synchronizing, that is synchronization must happen once.

### 3.1 Sure winning

We recall the following characterization of the sure-winning region for eventually synchronizing [12, Lemma 7]:  $d \in \langle\langle 1 \rangle\rangle_{sure}^{event}(T)$  if and only if there exists  $k \geq 0$  such that  $\text{Supp}(d) \subseteq \text{Pre}^k(T)$ . Intuitively, from all states in  $\text{Pre}^i(T)$ , there exists an action to ensure that the successor is in  $\text{Pre}^{i-1}(T)$  (for all  $i > 0$ ), and therefore there exists a strategy to ensure that the probability mass in  $\text{Pre}^k(T)$  reaches  $T$  in exactly  $k$  steps. Now, if  $q_0 \notin \text{Pre}^k(T)$ , then for all sequences of actions  $a_0, \dots, a_{k-1}$  there is a path  $q_0 a_0 q_1 a_1 \dots q_k$  of length  $k$  that ends in  $q_k \in Q \setminus T$ . It is easy to derive the following result from this characterization.

**Lemma 3.** *If  $d_0 \notin \langle\langle 1 \rangle\rangle_{sure}^{event}(T)$  is not sure eventually synchronizing in  $T$ , then for all strategies  $\sigma$ , for all  $i \geq 0$ , we have:*

$$\mathcal{M}_i^\sigma(T) \leq 1 - \alpha_0 \cdot \alpha^i$$

where  $\alpha_0$  is the smallest positive probability in  $d_0$ .

Note that the bound  $1 - \alpha_0 \cdot \alpha^i$  tends to 1 as  $i \rightarrow \infty$ , which is unavoidable since MDPs that are not sure eventually synchronizing may be almost-sure eventually synchronizing [12]. The following variant of Lemma 3 will be useful in the sequel.

*Remark 1.* If  $\{q_0\} \notin \langle\langle 1 \rangle\rangle_{sure}^{event}(T)$  and we take  $\alpha'_0 = d_0(q_0)$ , then from the initial distribution  $d_0$  we have, for all strategies  $\sigma$ , for all  $i \geq 0$ :

$$\mathcal{M}_i^\sigma(T) \leq 1 - \alpha'_0 \cdot \alpha^i.$$

### 3.2 Limit-sure winning

If the MDP  $\mathcal{M}$  is not limit-sure winning for eventually synchronizing in  $T$ , then the probability in  $Q \setminus T$  is bounded away from 0 in all distributions in  $\mathcal{M}^\sigma$  (for all strategies  $\sigma$ ). We give an explicit bound  $\varepsilon_e$  as follows.

**Theorem 1.** *Given an initial distribution  $d_0$ , let  $\alpha_0$  be the smallest positive probability in  $d_0$ , and let  $\varepsilon_e = \alpha_0 \cdot \alpha^{(n+1) \cdot 2^n}$ . If  $d_0 \notin \langle\langle 1 \rangle\rangle_{limit}^{event}(T)$  is not limit-sure eventually synchronizing in  $T$ , then for all strategies  $\sigma$ , for all  $i \geq 0$ , we have:*

$$\mathcal{M}_i^\sigma(T) \leq 1 - \varepsilon_e,$$

that is, no distribution in  $\mathcal{M}^\sigma$  is strictly  $(1 - \varepsilon_e)$ -synchronizing in  $T$ .

*Proof.* We recall the characterization<sup>1</sup> of [12, Lemma 11] for limit-sure synchronizing in an arbitrary set  $T$ . For all  $k \geq 0$ , we have

$$\langle\langle 1 \rangle\rangle_{limit}^{event}(T) = \langle\langle 1 \rangle\rangle_{sure}^{event}(T) \cup \langle\langle 1 \rangle\rangle_{limit}^{event}(R),$$

where  $R = \text{Pre}^k(T)$ .

We use this characterization with a specific  $k$  (and  $R$ ) as follows. Consider the sequence of predecessors  $\text{Pre}^i(T)$  (for  $i = 1, 2, \dots$ ), which is ultimately periodic. Let  $0 \leq k < 2^n$  and  $1 \leq r < 2^n$  be such that  $\text{Pre}^k(T) = \text{Pre}^{k+r}(T)$ , and let  $R = \text{Pre}^k(T)$ . Thus  $R = \text{Pre}^{k+r}(T) = \text{Pre}^r(R)$ .

Since  $d_0 \notin \langle\langle 1 \rangle\rangle_{limit}^{event}(T)$ , we have:

$$(a) \ d_0 \notin \langle\langle 1 \rangle\rangle_{sure}^{event}(T), \quad \text{and} \quad (b) \ d_0 \notin \langle\langle 1 \rangle\rangle_{limit}^{event}(R).$$

By (a), it follows from Lemma 3 that  $\mathcal{M}_i^\sigma(T) \leq 1 - \alpha_0 \cdot \alpha^i$  for all strategies  $\sigma$  and all  $i \geq 0$ , which establishes the bound in the lemma for the first  $2^n$  steps, since  $\alpha_0 \cdot \alpha^i \geq \varepsilon_e$  for all  $i \leq 2^n$ .

We now recall the characterization<sup>1</sup> of [12, Lemma 12] for limit-sure synchronizing in the set  $R$ , which has the property that  $R = \text{Pre}^r(R)$ :  $d_0 \in \langle\langle 1 \rangle\rangle_{limit}^{event}(R)$  if and only if there exists  $0 \leq t < r$  such that  $d_0 \times \{t\}$  is almost-sure winning for the reachability objective  $\diamond(R \times \{0\})$  in the MDP  $\mathcal{M} \times [r]$ . By (b), it follows that for all  $0 \leq t < r$ , the distribution  $d_0 \times \{t\}$  is *not* almost-sure winning for the reachability objective  $\diamond(R \times \{0\})$  in the MDP  $\mathcal{M} \times [r]$ .

Let  $\mathcal{N} = \mathcal{M} \times [r]$ . By Lemma 1, from all distributions  $d_0 \times \{t\}$  (for all  $0 \leq t < r$ ), for all strategies  $\sigma$  and all  $i \geq 0$ , we have:

$$\mathcal{N}_i^\sigma(R \times \{0\}) \leq 1 - \alpha_0 \cdot \alpha^{|Q_r|} = 1 - \alpha_0 \cdot \alpha^{n \cdot 2^n}.$$

Since this holds for all  $t = 0, \dots, r-1$ , we conclude that  $\mathcal{M}_i^\sigma(Q \setminus R) \geq \alpha_0 \cdot \alpha^{n \cdot 2^n}$  in the original MDP  $\mathcal{M}$  from  $d_0$ , for all strategies  $\sigma$  and all  $i \geq 0$ .

Since  $R = \text{Pre}^k(T)$ , it follows from Lemma 3 and Remark 1 that, if at step  $i$  a mass of probability  $p$  is outside  $R$ , then at step  $i+k$  a mass of probability at least  $p \cdot \alpha^k$  is outside  $T$ . Hence we have  $\mathcal{M}_{i+k}^\sigma(Q \setminus T) \geq \alpha_0 \cdot \alpha^{n \cdot 2^n} \cdot \alpha^k \geq \alpha_0 \cdot \alpha^{(n+1) \cdot 2^n}$  for all strategies  $\sigma$  and for all  $i \geq 0$ , which implies  $\mathcal{M}_i^\sigma(Q \setminus T) \geq \alpha_0 \cdot \alpha^{(n+1) \cdot 2^n}$  for all  $i \geq 2^n$  (since  $k < 2^n$ ).

Combining the results for  $i \leq 2^n$  and for  $i \geq 2^n$ , we get  $\mathcal{M}_i^\sigma(T) \leq 1 - \varepsilon_e$  for all  $i \geq 0$ , which concludes the proof.  $\square$

<sup>1</sup> The results of [12, Lemma 11 & 12] consider a more general definition of limit-sure synchronizing, where the support of the  $(1 - \varepsilon)$ -synchronizing distribution is required to have its support contained in a given set  $Z$ . We release this constraint by taking  $Z = Q$ .

Theorem 1 also gives a sufficient condition that can be used as an alternative to [12, Lemma 11] to show that an MDP is limit-sure eventually synchronizing. This will be useful in the proof of our main result (Theorem 2).

A variant of Theorem 1 is obtained by observing that if  $d_0 \notin \langle\langle 1 \rangle\rangle_{\text{limit}}^{\text{event}}(T)$ , there exists a set  $S_0 \subseteq \text{Supp}(d_0)$  such that  $S_0 \notin \langle\langle 1 \rangle\rangle_{\text{limit}}^{\text{event}}(T)$ . It may be that  $S_0$  is a strict subset of  $\text{Supp}(d_0)$ , and then it is sufficient to consider  $\alpha_0$  as the smallest positive probability of  $d_0$  on  $S_0$ .

*Remark 2.* If  $S_0 \notin \langle\langle 1 \rangle\rangle_{\text{limit}}^{\text{event}}(T)$  and  $S_0 \subseteq \text{Supp}(d_0)$ , then we can define  $\alpha_0$  by  $\min\{d_0(q) \mid q \in S_0\}$  in the bound  $\varepsilon_e$  of Theorem 1.

### 3.3 Almost-sure winning

A simple argument shows that the almost-sure winning region for eventually synchronizing consists of the union of the sure winning region for eventually synchronizing and the almost-sure winning region for weakly synchronizing [21, Section 5.1.2], that is  $\langle\langle 1 \rangle\rangle_{\text{almost}}^{\text{event}}(T) = \langle\langle 1 \rangle\rangle_{\text{sure}}^{\text{event}}(T) \cup \langle\langle 1 \rangle\rangle_{\text{almost}}^{\text{weakly}}(T)$ .

It follows that if  $d_0 \notin \langle\langle 1 \rangle\rangle_{\text{almost}}^{\text{event}}(T)$ , then both  $d_0 \notin \langle\langle 1 \rangle\rangle_{\text{sure}}^{\text{event}}(T)$  and  $d_0 \notin \langle\langle 1 \rangle\rangle_{\text{almost}}^{\text{weakly}}(T)$ , and we can use both the results of Lemma 3 and Theorem 2.

## 4 Weakly synchronizing

We now consider weakly synchronizing, which intuitively requires that synchronization happens infinitely often.

### 4.1 Sure winning

We recall the following characterization of the sure-winning region for weakly synchronizing [12, Lemma 18]: for all distributions  $d_0 \in \mathcal{D}(Q)$ , we have  $d_0 \in \langle\langle 1 \rangle\rangle_{\text{sure}}^{\text{weakly}}(T)$  if and only if there exists a set  $S \subseteq T$  such that  $\text{Supp}(d) \subseteq \text{Pre}^k(S)$  for some  $k \geq 0$ , and  $S \subseteq \text{Pre}^r(S)$  for some  $r \geq 1$ .

**Lemma 4.** *If  $d_0 \notin \langle\langle 1 \rangle\rangle_{\text{sure}}^{\text{weakly}}(T)$  is not sure weakly synchronizing in  $T$ , then for all strategies  $\sigma$ , in the sequence  $\mathcal{M}^\sigma$  there are at most  $2^n$  distributions that are 1-synchronized in  $T$ , that is  $\mathcal{M}_i^\sigma(T) = 1$  for at most  $2^n$  values of  $i$ .*

### 4.2 Limit-sure and almost-sure winning

The winning region for limit-sure and almost-sure weakly synchronizing coincide [12, Theorem 7]. Therefore, in the sequel we treat them interchangeably. We recall the following characterization of almost-sure weakly synchronizing.

**Lemma 5 ([12, Lemma 23, Theorem 7]).** *For all distributions  $d_0$ , the following equivalence holds:  $d_0 \in \langle\langle 1 \rangle\rangle_{\text{almost}}^{\text{weakly}}(T)$  if and only if there exists a set  $T' \subseteq T$  such that:*

$$d_0 \in \langle\langle 1 \rangle\rangle_{\text{limit}}^{\text{event}}(T') \quad \text{and} \quad T' \in \langle\langle 1 \rangle\rangle_{\text{limit}}^{\text{event}}(\text{Pre}(T')).$$

The condition in Lemma 5 ensures that from  $d_0$  almost all the probability mass (namely  $1 - \varepsilon$  for arbitrarily small  $\varepsilon > 0$ ) can be injected in a set  $T'$  of target states in 0 or more steps, and that from  $T'$  almost all the probability mass can be injected in  $\text{Pre}(T')$ , thus also in  $T'$  (but, in at least 1 step). Intuitively, by successively halving the value of  $\varepsilon$  one can construct a strategy that ensures almost all the probability mass loops through  $T'$ , thus a limit-sure weakly synchronizing strategy (which is equivalent to the existence of an almost-sure weakly synchronizing strategy).

If  $d_0 \notin \langle\langle 1 \rangle\rangle_{\text{almost}}^{\text{weakly}}(T)$  is not almost-sure weakly synchronizing, we use Lemma 5 to show that for all sets  $T' \subseteq T$ , if  $d_0 \in \langle\langle 1 \rangle\rangle_{\text{limit}}^{\text{event}}(T')$  is limit-sure eventually synchronizing in  $T'$ , then  $T'$  is not limit-sure eventually synchronizing in  $\text{Pre}(T')$  (i.e.,  $T' \notin \langle\langle 1 \rangle\rangle_{\text{limit}}^{\text{event}}(\text{Pre}(T'))$ ). This implies that a bounded number of distributions in the sequence  $\mathcal{M}^\sigma$  can be  $(1 - \varepsilon)$ -synchronized in  $T$  (for sufficiently small  $\varepsilon$ ). We now state the main result of this section.

**Theorem 2.** *Given an initial distribution  $d_0$ , let  $\alpha_0$  be the smallest positive probability in  $d_0$ , and let  $\varepsilon_w = \alpha_0 \cdot \frac{\alpha^{(n+2) \cdot 4^n}}{n^{2^n+1}}$  and  $N_w = 2^n$ .*

*If  $d_0 \notin \langle\langle 1 \rangle\rangle_{\text{almost}}^{\text{weakly}}(T)$  is not almost-sure weakly synchronizing in  $T$ , then for all strategies  $\sigma$ , in the sequence  $\mathcal{M}^\sigma$  at most  $N_w$  distributions are strictly  $(1 - \varepsilon_w)$ -synchronized in  $T$ , that is  $\mathcal{M}_i^\sigma(T) > 1 - \varepsilon_w$  for at most  $N_w$  values of  $i$ .*

*Proof.* Given the assumption of the lemma, we show the following statement by induction on  $k = 0, 1, \dots, 2^n$ : if there are  $k$  distributions in  $\mathcal{M}^\sigma$  that are strictly  $(1 - \varepsilon_w)$ -synchronized in  $T$ , then there exist  $k$  *distinct* nonempty sets  $T_1, \dots, T_k \subseteq T$  such that no distribution after those  $k$  distributions in  $\mathcal{M}^\sigma$  is strictly  $(1 - \varepsilon_w)$ -synchronized in  $T_j$  (for all  $1 \leq j \leq k$ ).

For  $k = 2^n$ , one of the sets  $T_j$  is equal to  $T$  and it follows that at most  $2^n$  distributions in  $\mathcal{M}^\sigma$  can be  $(1 - \varepsilon_w)$ -synchronized in  $T$ , which concludes the base case. For the proof by induction, we use the bound  $\varepsilon_e$  of Theorem 1. Let  $F = (n + 1) \cdot 2^n$  (thus  $\varepsilon_e = \alpha_0 \cdot \alpha^F$ ) and for  $k = 0, 1, \dots$  define  $z_k = \frac{\alpha_0}{n} \cdot \left(\frac{\alpha^{F+1}}{n}\right)^k$ . We prove a slightly stronger statement: for  $k = 0, 1, \dots, 2^n$ , if there are  $k$  positions  $i_1 < i_2 < \dots < i_k$  such that the distributions  $\mathcal{M}_{i_j}^\sigma$  ( $j = 1, \dots, k$ ) are strictly  $(1 - \varepsilon_w)$ -synchronized in  $T$ , then there exist  $k$  distinct nonempty sets  $T_1, \dots, T_k \subseteq T$  such that no distribution after position  $i_j$  in  $\mathcal{M}^\sigma$  is strictly  $(1 - z_j \cdot \alpha^{F+1})$ -synchronized in  $T_j$  (for all  $1 \leq j \leq k$ ).

This statement is indeed stronger since the sequence  $z_k$  is decreasing, and  $\varepsilon_w$  was chosen such that  $\varepsilon_w \leq z_{2^n}$ , from which it follows that  $1 - \varepsilon_w \geq 1 - z_k$  for all  $k \leq 2^n$ .

The base case for  $k = 0$  holds trivially. For the induction case, assume that the statement holds for a given  $k < 2^n$ , and show that it holds for  $k + 1$  as follows. If there are  $k + 1$  positions  $i_1 < i_2 < \dots < i_{k+1}$  such that all distributions  $d_j = \mathcal{M}_{i_j}^\sigma$  ( $j = 1, \dots, k + 1$ ) are strictly  $(1 - \varepsilon_w)$ -synchronized in  $T$ , then by the induction hypothesis, no distribution after position  $i_j$  in  $\mathcal{M}^\sigma$  is strictly  $(1 - z_j \cdot \alpha^{F+1})$ -synchronized in  $T_j$  (for all  $1 \leq j \leq k$ ).

Now consider the distribution  $d_{k+1}$  at position  $i_{k+1}$ , which is  $(1 - \varepsilon_w)$ -synchronized in  $T$  and appears after position  $i_k$  in  $\mathcal{M}^\sigma$ . We construct the set  $T_{k+1} = \{q \in T \cap \text{Supp}(d_{k+1}) \mid d_{k+1}(q) > z_{k+1}\}$ , which contains the states in  $T$  that carry enough probability mass (namely  $z_{k+1}$ ) according to  $d_{k+1}$ .

Note that not all states in  $T \cap \text{Supp}(d_{k+1})$  carry a probability mass less than  $z_{k+1}$ : otherwise, the total mass of  $T$  in  $d_{k+1}$  would be at most  $n \cdot z_{k+1} \leq 1 - \varepsilon_w$  (this inequality holds thanks to  $n \geq 2$ ), in contradiction with  $d_{k+1}$  being  $(1 - \varepsilon_w)$ -synchronized in  $T$ . Therefore  $T_{k+1}$  is nonempty. Hence the set  $T_{k+1}$  can be obtained from  $T$  by removing at most  $n - 1$  states and we have

$$\begin{cases} d_{k+1}(T_{k+1}) > 1 - \varepsilon_w - (n - 1) \cdot z_{k+1} \geq 1 - n \cdot z_{k+1} = 1 - z_k \cdot \alpha^{F+1} \\ d_{k+1}(q) > z_{k+1} \text{ for all } q \in T_{k+1} \end{cases}$$

So  $d_{k+1}$  is strictly  $(1 - z_k \cdot \alpha^{F+1})$ -synchronized in  $T_{k+1}$ , and therefore also strictly  $(1 - z_j \cdot \alpha^{F+1})$ -synchronized in  $T_{k+1}$  (for all  $1 \leq j \leq k$ ). Then, the induction hypothesis implies that the set  $T_{k+1}$  is distinct from  $T_1, \dots, T_k$ . Since  $1 - z_k \cdot \alpha^{F+1} \geq 1 - z_0 \cdot \alpha^{F+1} > 1 - \varepsilon_e$ , it follows that  $d_{k+1} = \mathcal{M}_{i_{k+1}}^\sigma$  is strictly  $(1 - \varepsilon_e)$ -synchronized in  $T_{k+1}$ , and by Theorem 1, that the initial distribution  $d_0$  is limit-sure eventually synchronizing in  $T_{k+1}$ , that is  $d_0 \in \langle\langle 1 \rangle\rangle_{\text{limit}}^{\text{event}}(T_{k+1})$ .

By Lemma 5, this entails that  $T_{k+1}$  is not limit-sure eventually synchronizing in  $\text{Pre}(T_{k+1})$  (i.e.,  $T_{k+1} \notin \langle\langle 1 \rangle\rangle_{\text{limit}}^{\text{event}}(\text{Pre}(T_{k+1}))$ ), and by Theorem 1, for all distributions  $d$  in  $\mathcal{M}^\sigma$  that occur at or after position  $i_{k+1}$ , we have  $d(\text{Pre}(T_{k+1})) \leq 1 - z_{k+1} \cdot \alpha^F$  where  $z_{k+1} < \min\{d_{k+1}(q) \mid q \in T_{k+1} \cap \text{Supp}(d_{k+1})\}$  is a lower bound on the smallest positive probability of a state of  $T_{k+1}$  in the distribution  $d_{k+1}$ , taken as the initial distribution (see Remark 2). It follows that for all distributions  $d$  in  $\mathcal{M}^\sigma$  that occur (strictly) after position  $i_{k+1}$ , we have  $d(T_{k+1}) \leq 1 - z_{k+1} \cdot \alpha^{F+1}$ . Hence no distribution in  $\mathcal{M}^\sigma$  after  $d_{k+1}$  is strictly  $(1 - z_{k+1} \cdot \alpha^{F+1})$ -synchronized, which concludes the proof of the induction case.  $\square$

## 5 Always and Strongly Synchronizing

The analysis of always and strongly synchronizing modes is relatively straightforward, and we present the bounds in Theorem 3.

**Theorem 3.** *Given an initial distribution  $d_0$ , let  $\alpha_0$  be the smallest positive probability in  $d_0$ , and let  $\varepsilon_a = \alpha_0 \cdot \frac{\alpha^n}{n}$  and  $\varepsilon_s = \alpha_0 \cdot \frac{\alpha^{2n}}{n^2}$ .*

- if  $d_0 \notin \langle\langle 1 \rangle\rangle_{\text{sure}}^{\text{always}}(T)$  is not sure always synchronizing in  $T$ , then for all strategies  $\sigma$ , in the sequence  $\mathcal{M}^\sigma$  there exists a position  $i \leq n$  such that  $\mathcal{M}_i^\sigma$  is not  $(1 - \varepsilon_a)$ -synchronized in  $T$ ,
- if  $d_0 \notin \langle\langle 1 \rangle\rangle_{\text{sure}}^{\text{strongly}}(T)$  is not sure strongly synchronizing in  $T$ , then for all strategies  $\sigma$ , in the sequence  $\mathcal{M}^\sigma$  there exist infinitely many positions  $i_0 < i_1 < i_2 < \dots$  where  $i_0 \leq n$  and  $i_{j+1} - i_j \leq n$  for all  $j \geq 0$  such that  $\mathcal{M}_{i_j}^\sigma$  is not 1-synchronized in  $T$ .

- if  $d_0 \notin \langle\langle 1 \rangle\rangle_{almost}^{strongly}(T)$  is not almost-sure strongly synchronizing in  $T$ , then for all strategies  $\sigma$ , in the sequence  $\mathcal{M}^\sigma$  there exist infinitely many positions  $i_0 < i_1 < i_2 < \dots$  where  $i_0 \leq n$  and  $i_{j+1} - i_j \leq n$  for all  $j \geq 0$  such that  $\mathcal{M}_{i_j}^\sigma$  is not  $(1 - \varepsilon_s)$ -synchronized in  $T$ .

## 6 Adversarial Synchronizing Objectives

In an adversarial MDP the strategies are universally quantified, which corresponds to satisfying an objective regardless of the choice of strategies by an adversary. Replacing  $\exists\sigma$  by  $\forall\sigma$  in the definition of the three winning modes gives, after taking the negation to get existentially quantified strategies, the following new winning modes.

Given a set  $T \subseteq Q$ , we say that a sequence  $d_0 d_1 \dots$  of probability distributions is:

- *positively* {always, eventually, weakly, strongly} winning if  $d_i(T) > 0$  for, respectively, all  $i \geq 0$ , some  $i \geq 0$ , infinitely many  $i$ 's, all but finitely many  $i$ 's.
- *boundedly* {always, eventually, weakly, strongly} winning if there exists  $\varepsilon > 0$  such that  $d_i(T) > \varepsilon$  for, respectively, all  $i \geq 0$ , some  $i \geq 0$ , infinitely many  $i$ 's, all but finitely many  $i$ 's.

For  $\lambda \in \{always, event, weakly, strongly\}$ , we denote by  $\langle\langle 1 \rangle\rangle_{positive}^\lambda(T)$  (resp.,  $\langle\langle 1 \rangle\rangle_{bounded}^\lambda(T)$ ) the set of initial distributions  $d_0$  from which there exists a strategy  $\sigma$  such that the sequence  $\mathcal{M}^\sigma$  is positively (resp., boundedly)  $\lambda$ -synchronizing in  $T$ , and we say that  $\sigma$  is positively (resp., boundedly)  $\lambda$ -synchronizing in  $T$  from  $d_0$ .

Table 1 summarizes the new definitions. Note that replacing the existential quantification on strategies in boundedly winning mode by a supremum gives the same question, since  $\exists\sigma : f(\sigma) > 0$  is equivalent to  $\sup_\sigma f(\sigma) > 0$ . For the same reason, we have the identity  $\langle\langle 1 \rangle\rangle_{positive}^{event}(T) = \langle\langle 1 \rangle\rangle_{bounded}^{event}(T)$ . It is easy to show that the definitions imply the identity  $\langle\langle 1 \rangle\rangle_{bounded}^{always}(T) = \langle\langle 1 \rangle\rangle_{positive}^{always}(T) \cap \langle\langle 1 \rangle\rangle_{bounded}^{strongly}(T)$ , which we also obtain as a corollary of Lemma 6 below.

*Remark 3.* It immediately follows from the definitions that for all synchronizing modes  $\lambda \in \{always, event, weakly, strongly\}$ , and  $\mu \in \{positive, bounded\}$ :

- $\langle\langle 1 \rangle\rangle_\mu^{always}(T) \subseteq \langle\langle 1 \rangle\rangle_\mu^{strongly}(T) \subseteq \langle\langle 1 \rangle\rangle_\mu^{weakly}(T) \subseteq \langle\langle 1 \rangle\rangle_\mu^{event}(T)$ ,
- $\langle\langle 1 \rangle\rangle_{bounded}^\lambda(T) \subseteq \langle\langle 1 \rangle\rangle_{positive}^\lambda(T)$ ,

and moreover,

- $\langle\langle 1 \rangle\rangle_{positive}^{event}(T) = \langle\langle 1 \rangle\rangle_{bounded}^{event}(T)$ , and
- $\langle\langle 1 \rangle\rangle_{bounded}^{always}(T) = \langle\langle 1 \rangle\rangle_{positive}^{always}(T) \cap \langle\langle 1 \rangle\rangle_{bounded}^{strongly}(T)$ .

It is easy to see that if there exists a strategy  $\sigma$  that is positively  $\lambda$ -synchronizing in  $T$ , then the strategy  $\sigma_u$  that plays at every round all actions uniformly at random is also positively  $\lambda$ -synchronizing in  $T$ , because the condition  $d_i(T) > 0$  is equivalent to  $\text{Supp}(d_i) \cap T \neq \emptyset$ , and because we have  $\sigma \subseteq \sigma_u$ , which implies that  $\text{Supp}(\mathcal{M}_i^\sigma) \subseteq \text{Supp}(\mathcal{M}_i^{\sigma_u})$  for all  $i \geq 0$ .

Hence, in all four synchronization modes, the question is equivalent to the same question in Markov chains (obtained from the given MDP by fixing the strategy  $\sigma_u$ ) which can be solved as follows. Given a Markov chain, consider the underlying directed graph  $\langle Q, E \rangle$  where  $(q, q') \in E$  if  $\delta(q, a)(q') > 0$  (where  $A = \{a\}$ ). For positive eventually synchronizing, it suffices to find a state in  $T$  that is reachable in that graph, and for positive weakly synchronizing, it suffices to find a state in  $T$  that is both reachable and can reach itself. These questions are NL-complete. For positive always and strongly synchronizing, the question is equivalent to the model-checking problem for the formulas  $G\exists T$  and  $FG\exists T$  in the logic CTL+Sync, which are both coNP-complete [6, Lemma 2 & Section 3].

For boundedly winning, we show that one strategy is good enough in all four synchronization modes, like for positive winning. The strategy plays like  $\sigma_u$  for the first  $2^n$  rounds, and then switches to a strategy  $\sigma_{\mathcal{E}}$  that, in the states  $q \in \mathcal{E}$ , plays uniformly at random all actions that stay in the end-component  $\mathcal{E}(q)$  of  $q$  (thus all actions in  $A_{\mathcal{E}(q)}$ ), and in the transient states  $q \notin \mathcal{E}$ , plays all actions uniformly at random. We call this strategy the *freezing* strategy. Intuitively we use  $\sigma_u$  to scatter the probability mass in all end-components of the MDP, and then  $\sigma_{\mathcal{E}}$  to maintain a bounded probability in each end-component.

**Lemma 6.** *Let  $\mathcal{M}$  be an MDP with  $n$  states and initial distribution  $d_0$ , and let  $T$  be a target set. Consider the following conditions:*

$$(1) \forall i \geq 0 : \mathcal{M}_i^{\sigma_u}(T) > 0 \quad (2) \forall i \geq 2^n : \mathcal{M}_i^{\sigma_u}(\mathcal{E} \cap T) > 0$$

*Then, the following equivalences hold:*

- (a)  $d_0 \in \langle\langle 1 \rangle\rangle_{\text{positive}}^{\text{always}}(T)$  if and only if Condition (1) holds;
- (b)  $d_0 \in \langle\langle 1 \rangle\rangle_{\text{bounded}}^{\text{strongly}}(T)$  if and only if Condition (2) holds;
- (c)  $d_0 \in \langle\langle 1 \rangle\rangle_{\text{bounded}}^{\text{always}}(T)$  if and only if Conditions (1) and (2) hold;

*Proof.* Equivalence (a) follows from the definition of positive always synchronizing, and from the fact that the uniform strategy  $\sigma_u$  is sufficient for positive winning.

We show Equivalence (b) as follows. First, if Condition (2) does not hold, then  $\mathcal{M}_i^{\sigma_u}(\mathcal{E} \cap T) = 0$  for some  $i \geq 2^n$ , and thus also for infinitely many  $i$ 's (since the sequence  $\text{Supp}(\mathcal{M}_i^{\sigma_u})$  is ultimately periodic, after at most  $2^n$  steps). For arbitrary strategy  $\sigma$ , we have  $\text{Supp}(\mathcal{M}_i^\sigma) \subseteq \text{Supp}(\mathcal{M}_i^{\sigma_u})$  for all  $i \geq 0$ , therefore  $\mathcal{M}_i^\sigma(\mathcal{E} \cap T) = 0$  for infinitely many  $i$ 's. By Lemma 2, we have  $\liminf_{i \rightarrow \infty} \mathcal{M}_i^\sigma(\mathcal{E}) = 1$  which entails that  $\limsup \mathcal{M}_i^\sigma(\mathcal{E} \setminus T) = 1$  and  $\limsup \mathcal{M}_i^\sigma(Q \setminus T) = 1$ , that is  $\liminf \mathcal{M}_i^\sigma(T) = 0$ . Since this holds for arbitrary strategy  $\sigma$ , it follows that  $d_0 \notin \langle\langle 1 \rangle\rangle_{\text{bounded}}^{\text{strongly}}(T)$ .

For the converse direction, assuming Condition (2) holds, we show that  $d_0 \in \langle\langle 1 \rangle\rangle_{\text{bounded}}^{\text{strongly}}(T)$ , witnessed by the freezing strategy  $\sigma_f$  (which plays like  $\sigma_u$  for the first  $2^n$  rounds, and then switches to the strategy  $\sigma_{\mathcal{E}}$ ).

We show the key property that

$$\text{Supp}(\mathcal{M}_i^{\sigma_u}) \cap \mathcal{E} = \text{Supp}(\mathcal{M}_i^{\sigma_f}) \cap \mathcal{E} \text{ for all } i \geq 2^n.$$

Fix an arbitrary  $i \geq 2^n$  and let  $p$  be a period of the sequence  $\text{Supp}(\mathcal{M}^{\sigma_u})$  such that  $i - p \leq 2^n$  and  $\text{Supp}(\mathcal{M}_i^{\sigma_u}) = \text{Supp}(\mathcal{M}_{i-p}^{\sigma_u})$ . Consider the Markov chain  $\mathcal{M}_{\mathcal{E}}$  obtained by fixing the strategy  $\sigma_{\mathcal{E}}$  in  $\mathcal{M}$ . From the basic theory of Markov chains, each end-component  $C$  in  $\mathcal{M}$  is a recurrent class in  $\mathcal{M}_{\mathcal{E}}$ . For each  $i \geq 2^n$ , either all or none of the states in a periodic class of  $C$  are in the support of  $\mathcal{M}_i^{\sigma_u}$ .

To show the key property, first consider a state  $q \in \text{Supp}(\mathcal{M}_i^{\sigma_u}) \cap \mathcal{E}$  for  $i \geq 2^n$ , and show that  $q \in \text{Supp}(\mathcal{M}_i^{\sigma_f}) \cap \mathcal{E}$ . Let  $S$  be the periodic class of  $\mathcal{E}(q)$  containing  $q$  (in  $\mathcal{M}_{\mathcal{E}}$ ), and thus

$$S \subseteq \text{Supp}(\mathcal{M}_i^{\sigma_u}) \cap \mathcal{E}, \text{ and thus } S \subseteq \text{Supp}(\mathcal{M}_{i-p}^{\sigma_u}) \cap \mathcal{E}.$$

Since  $\sigma_u$  and  $\sigma_f$  coincide on the first  $2^n$  rounds, we have  $S \subseteq \text{Supp}(\mathcal{M}_{i-p}^{\sigma_f}) \cap \mathcal{E}$ .

Now consider the strategy  $\sigma_{\mathcal{E}}$  and an initial distribution with support  $S$ , and denote by  $S + j$  the support of the probability distribution after playing  $\sigma_{\mathcal{E}}$  for  $j$  steps. Then, since  $\sigma_{\mathcal{E}} \subseteq \sigma_f \subseteq \sigma_u$ ,

$$S + p \subseteq \text{Supp}(\mathcal{M}_i^{\sigma_f}) \cap \mathcal{E}, \text{ and } S + p \subseteq \text{Supp}(\mathcal{M}_i^{\sigma_u}) \cap \mathcal{E}.$$

We can repeat the same argument with  $S + p$  instead of  $S$ , and show by induction that  $S + j \cdot p \subseteq \text{Supp}(\mathcal{M}_i^{\sigma_f}) \cap \mathcal{E}$  for all  $j \geq 1$ . In particular, by taking  $j$  the period of the end-component containing  $q$ , we get  $S + j \cdot p = S$  and thus  $S \subseteq \text{Supp}(\mathcal{M}_i^{\sigma_f}) \cap \mathcal{E}$ , which establishes one direction of the key property (the converse direction follows from  $\sigma_f \subseteq \sigma_u$ ).

From the theory of Markov chains, in every end-component state  $q \in \mathcal{E}$ , the positive probability mass is bounded away from 0 in  $\mathcal{M}^{\mathcal{E}}$ , that is there exists a bound  $\varepsilon > 0$  such that for all  $i \geq 2^n$ , for all  $q \in \mathcal{E}$ , if  $\mathcal{M}_i^{\sigma_f}(q) \neq 0$ , then  $\mathcal{M}_i^{\sigma_f}(q) \geq \varepsilon$ . By the key property and Condition (2), for all  $i \geq 2^n$ , there exists  $q \in \mathcal{E} \cap T$  such that  $\mathcal{M}_i^{\sigma_f}(q) \neq 0$ , which implies that  $\liminf_{i \rightarrow \infty} M_i^{\sigma_f}(T) \geq \varepsilon > 0$  and thus  $d_0 \in \langle\langle 1 \rangle\rangle_{\text{bounded}}^{\text{strongly}}(T)$ .

Finally, the proof for Equivalence (c) follows the same steps as above to show that Conditions (1) and (2) imply  $d_0 \in \langle\langle 1 \rangle\rangle_{\text{bounded}}^{\text{always}}(T)$ , where Condition (1) is used to bound  $M_i^{\sigma_f}(T)$  for the first  $2^n$  rounds, and thus to ensure that  $M_i^{\sigma_f}(T) \geq B > 0$  for all  $i \geq 0$ , hence  $d_0 \in \langle\langle 1 \rangle\rangle_{\text{bounded}}^{\text{always}}(T)$ . The converse direction immediately follows from the first part of Remark 3 and Equivalences (a) and (b).  $\square$

We extract explicit bounds from the proof of Lemma 6. All end-components are reached within a most  $n$  steps (under  $\sigma_u$ ), and further all states in (a periodic class of) a recurrent class are reached (synchronously) within a most  $n^2$  steps [15, Theorem 4.2.11], thus all states in the periodic class have probability mass at

least  $\varepsilon_a = \alpha_0 \cdot \left(\frac{\alpha}{|\mathbf{A}|}\right)^{n+n^2}$  where  $\alpha_0$  is the smallest positive probability in the initial distribution  $d_0$  (note that  $\alpha/|\mathbf{A}|$  is the smallest probability in the Markov chain  $\mathcal{M}_{\mathcal{E}}$ ). It follows that the freezing strategy ensures probability at least  $\varepsilon_a$  in  $T$  at every step (if  $\mathcal{M}$  is boundedly always synchronizing), and probability at least  $\varepsilon_a$  in  $T$  at every step after  $N = n + n^2$ .

The conditions (1) and (2) in Lemma 6 can be decided in coNP as follows. For Condition (1) we guess an index  $i \leq 2^n$  (in binary) and compute the  $i$ -th power of the Boolean transition matrix  $M \in \{0, 1\}^{n^2}$  where  $M(q, q') = 1$  if there is a transition from state  $q$  to state  $q'$  in the Markov chain obtained from the given MDP  $\mathcal{M}$  by fixing the strategy  $\sigma_u$ . The matrix  $M^i$  can be computed in polynomial time by successive squaring of  $M$ . Then it suffices to check whether  $M^i(q_0, q) = 0$  for all  $q_0 \in \text{Supp}(d_0)$  and  $q \in T$ . For Condition (2), since the sequence  $\text{Supp}(\mathcal{M}_i^{\sigma_u})$  is ultimately periodic, we guess two indices  $i, p \leq 2^n$  ( $p \geq 1$ ) and check that  $\text{Supp}(\mathcal{M}_i^{\sigma_u}) = \text{Supp}(\mathcal{M}_{i+p}^{\sigma_u})$  and  $\text{Supp}(\mathcal{M}_i^{\sigma_u}) \cap \mathcal{E} \cap T = \emptyset$ , using the same approach by successive squaring. Note that the union  $\mathcal{E}$  of all end-components can be computed in polynomial time [9, 10].

Conditions (1) and (2) are also coNP-hard, using the same reduction that established coNP-hardness of the positive always and positive bounded synchronizing [6, Lemma 2 & Section 3], in which positive and bounded winning mode coincide. It follows that the membership problem for bounded always and bounded strongly synchronizing is coNP-complete.

We now show the solution for bounded weakly synchronizing. It suffices to find a state in  $T \cap \mathcal{E}$  that is reachable in the underlying graph of the Markov chain  $\mathcal{M}_{\sigma_u}$ , which is a NL-complete problem (like for positive weakly synchronizing, except we require a reachable state in  $T \cap \mathcal{E}$ , not just in  $T$ ). Indeed, if all reachable end-components are contained in  $Q \setminus T$ , then by Lemma 2 we have  $\liminf_{i \rightarrow \infty} \mathcal{M}_i^{\sigma_u}(Q \setminus T) = 1$ , that is  $\limsup_{i \rightarrow \infty} \mathcal{M}_i^{\sigma_u}(T) = 0$ . For the converse direction, if a state  $\hat{q} \in T \cap \mathcal{E}$  is reachable, then by a similar argument as above based on the theory of Markov chains, as the probability mass in the states of the periodic classes (that contain some probability mass) is bounded away from 0 in  $\mathcal{M}^{\sigma_{\mathcal{E}}}$ , it follows that within every  $p$  steps, where  $p$  is the period of the recurrent class  $\mathcal{E}(\hat{q})$  the probability mass in  $\hat{q}$  is at least  $\alpha_0 \cdot (\alpha/|\mathbf{A}|)^{n+n^2}$ . Therefore,  $\mathcal{M}$  is boundedly weakly synchronizing in  $T$ . For the sake of completeness, note that for eventually synchronizing MDPs, the probability mass  $\varepsilon_e = \alpha_0 \cdot (\alpha/|\mathbf{A}|)^n$  in  $T$  can be ensured within  $n$  steps (using  $\sigma_u$ ).

**Theorem 4.** *The complexity of the membership problem for positive and bounded synchronizing objectives is summarized in Table 2.*

In Table 2, the merged cells for eventually synchronizing reflect the fact that the winning regions coincide (see Remark 3). The winning regions for the other synchronizing modes do not coincide, already in Markov chains (see Fig. 1).

**Acknowledgment.** The authors are grateful to Jean-François Raskin for logistical support, and to Mahsa Shirmohammadi for interesting discussions about adversarial objectives.

## References

1. S. Akshay, B. Genest, and N. Vyas. Distribution-based objectives for Markov decision processes. In *Proc. of LICS*, pages 36–45. ACM, 2018.
2. C. Baier and J.-P. Katoen. *Principles of Model Checking*. MIT, 2008.
3. D. Beauquier, A. M. Rabinovich, and A. Slissenko. A logic of probability with decidable model-checking. In *Proc. of CSL: Computer Science Logic*, LNCS 2471, pages 306–321. Springer, 2002.
4. P. C. Bell and P. Semukhin. Decidability of cutpoint isolation for probabilistic finite automata on letter-bounded inputs. In *Proc. of CONCUR*, volume 171 of *LIPICs*, pages 22:1–22:16. Schloss Dagstuhl - Leibniz-Zentrum für Informatik, 2020.
5. N. Bertrand, M. Dewaskar, B. Genest, and H. Gimbert. Controlling a population. In *Proc. of CONCUR: Concurrency Theory*, volume 85 of *LIPICs*, pages 12:1–12:16. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2017.
6. K. Chatterjee and L. Doyen. Computation tree logic for synchronization properties. In *Proc. of ICALP: Automata, Languages, and Programming*, volume 55 of *LIPICs*, pages 98:1–98:14. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2016.
7. K. Chatterjee and T. A. Henzinger. A survey of stochastic  $\omega$ -regular games. *J. Comput. Syst. Sci.*, 78(2):394–413, 2012.
8. T. Colcombet, N. Fijalkow, and P. Ohlmann. Controlling a random population. In *Proc. of FoSSaCS*, LNCS 12077, pages 119–135. Springer, 2020.
9. C. Courcoubetis and M. Yannakakis. The complexity of probabilistic verification. *Journal of the ACM*, 42(4):857–907, 1995.
10. L. de Alfaro. *Formal Verification of Probabilistic Systems*. PhD thesis, Stanford University, 1997.
11. L. de Alfaro and T. A. Henzinger. Concurrent omega-regular games. In *Proc. of LICS: Logic in Computer Science*, pages 141–154. IEEE, 2000.
12. L. Doyen, T. Massart, and M. Shirmohammadi. The complexity of synchronizing Markov decision processes. *J. of Computer and System Sciences*, 100:96–129, 2019.
13. L. Doyen and M. van den Bogaard. Bounds for synchronizing Markov decision processes. *CoRR*, abs/2204.12814, 2022.
14. E. Feinberg and A. Shwartz, editors. *Handbook of Markov Decision Processes - Methods and Applications*. Kluwer, 2002.
15. R. G. Gallager. *Stochastic processes: theory for applications*. Cambridge University Press, 2013.
16. H. Hermanns, J. Krcál, and J. Kretínský. Probabilistic bisimulation: Naturally on distributions. In *Proc. of CONCUR: Concurrency Theory*, LNCS 8704, pages 249–265. Springer, 2014.
17. P. Jancar and Z. Sawa. A note on emptiness for alternating finite automata with a one-letter alphabet. *Inf. Process. Lett.*, 104(5):164–167, 2007.
18. V. A. Korthikanti, M. Viswanathan, G. Agha, and Y. Kwon. Reasoning about MDPs as transformers of probability distributions. In *Proc. of QEST: Quantitative Evaluation of Systems*, pages 199–208. IEEE Computer Society, 2010.
19. M. L. Puterman. *Markov Decision Processes*. John Wiley and Sons, 1994.
20. M. O. Rabin. Probabilistic automata. *Information and Control*, 6:230–245, 1963.
21. M. Shirmohammadi. *Qualitative analysis of probabilistic synchronizing systems*. PhD thesis, U. Libre de Bruxelles & École Normale Supérieure de Cachan, 2014.
22. M. Y. Vardi. Automatic verification of probabilistic concurrent finite-state programs. In *Proc. of FOCS*, pages 327–338. IEEE Computer Society, 1985.